

Genetically Encoded Fragment-Based Discovery from Phage-Displayed Macrocyclic Libraries with Genetically Encoded Unnatural Pharmacophores

Arunika I. Ekanayake, Lena Sobze, Payam Kelich, Jihea Youk, Nicholas J. Bennett, Raja Mukherjee, Atul Bhardwaj, Frank Wuest, Lela Vukovic, and Ratmir Derda*

Cite This: *J. Am. Chem. Soc.* 2021, 143, 5497–5507

Read Online

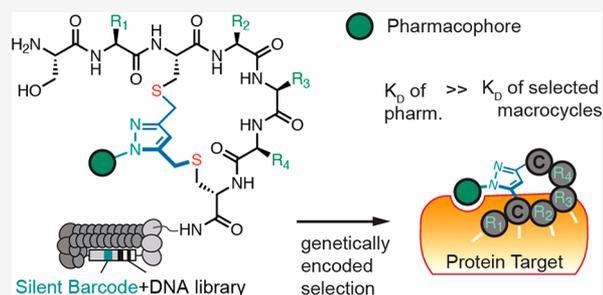
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: Genetically encoded macrocyclic peptide libraries with unnatural pharmacophores are valuable sources for the discovery of ligands for many targets of interest. Traditionally, generation of such libraries employs “early stage” incorporation of unnatural building blocks into the chemically or translationally produced macrocycles. Here, we describe a divergent late-stage approach to such libraries starting from readily available starting material: genetically encoded libraries of peptides. A diketone linchpin 1,5-dichloropentane-2,4-dione converts peptide libraries displayed on phage to 1,3-diketone bearing macrocyclic peptides (DKMP): shelf-stable precursors for Knorr pyrazole synthesis. Ligation of diverse hydrazine derivatives onto DKMP libraries displayed on phage that carries silent DNA-barcodes yields macrocyclic libraries in which the amino acid sequence and the pharmacophore are encoded by DNA. Selection of this library against carbonic anhydrase enriched macrocycles with benzenesulfonamide pharmacophore and nanomolar K_d . The methodology described in this manuscript can graft diverse pharmacophores into many existing genetically encoded phage libraries and significantly increase the value of such libraries in molecular discoveries.



INTRODUCTION

Late-stage functionalization of unprotected peptides composed of natural amino acids in aqueous media provides a powerful approach to modify readily available million-to-billion scale genetically encoded peptide libraries, displayed on phage, mRNA, or DNA.^{1–5} Such functionalization expands existing genetically encoded chemical space to incorporate unnatural pharmacophores not present in the original peptide libraries, allowing the discovery of value-added molecules with properties not offered by peptides alone.^{6–9} Numerous reports demonstrated the power of discovery of potent ligands from phage- and mRNA-displayed libraries in which unnatural pharmacophores were grafted onto the peptides in these genetically encoded libraries.^{10–12} Genetically encoded fragment-based discovery (GE-FBD)¹⁰ from such libraries is conceptually similar to canonical fragment-based design (FBD), which is a powerful method for the development of ligands, drug leads and three FDA-approved drugs to date.^{13–19} Methods for production of GE-macrocyclic libraries with unnatural pharmacophores are bottom-up organic synthesis of DNA-encoded libraries (DEL),^{20–25} *in vitro* translation of mRNA-displayed macrocycles using modified protein translation systems,^{11,12,26} and engineering of biochemical pathways that produce ribosomally made and post translationally modified peptides.²⁷ Among these approaches, the late-stage chemical modification of existing GE-

peptide libraries offers potentially the simplest path to generating chemical diversity. It combines a robust expression of million-to-billion scale peptides libraries made of 20 natural amino acids with a simple diversification of peptides by site-specific chemical conjugation. This report advances both the late stage functionalization of GE-libraries and GE-FBD approaches by introducing three important concepts: (i) ligation of unnatural fragments onto preformed GE-macrocyclic libraries; (ii) production of shelf-stable GE-macrocyclic libraries with a handle for biorthogonal reaction that forms an irreversible covalent bond, (iii) encoding and decoding of ligated unnatural fragments by DNA sequencing.

There exist several examples of late-stage functionalization of GE-peptide libraries to yield macrocycles with unnatural chemotypes. Oximes derived from dichloroacetone (DCA) linchpin²⁸ convert linear peptide libraries displayed on phage to a library of macrocycles and simultaneously introduces a diverse

Received: January 31, 2021

Published: March 30, 2021



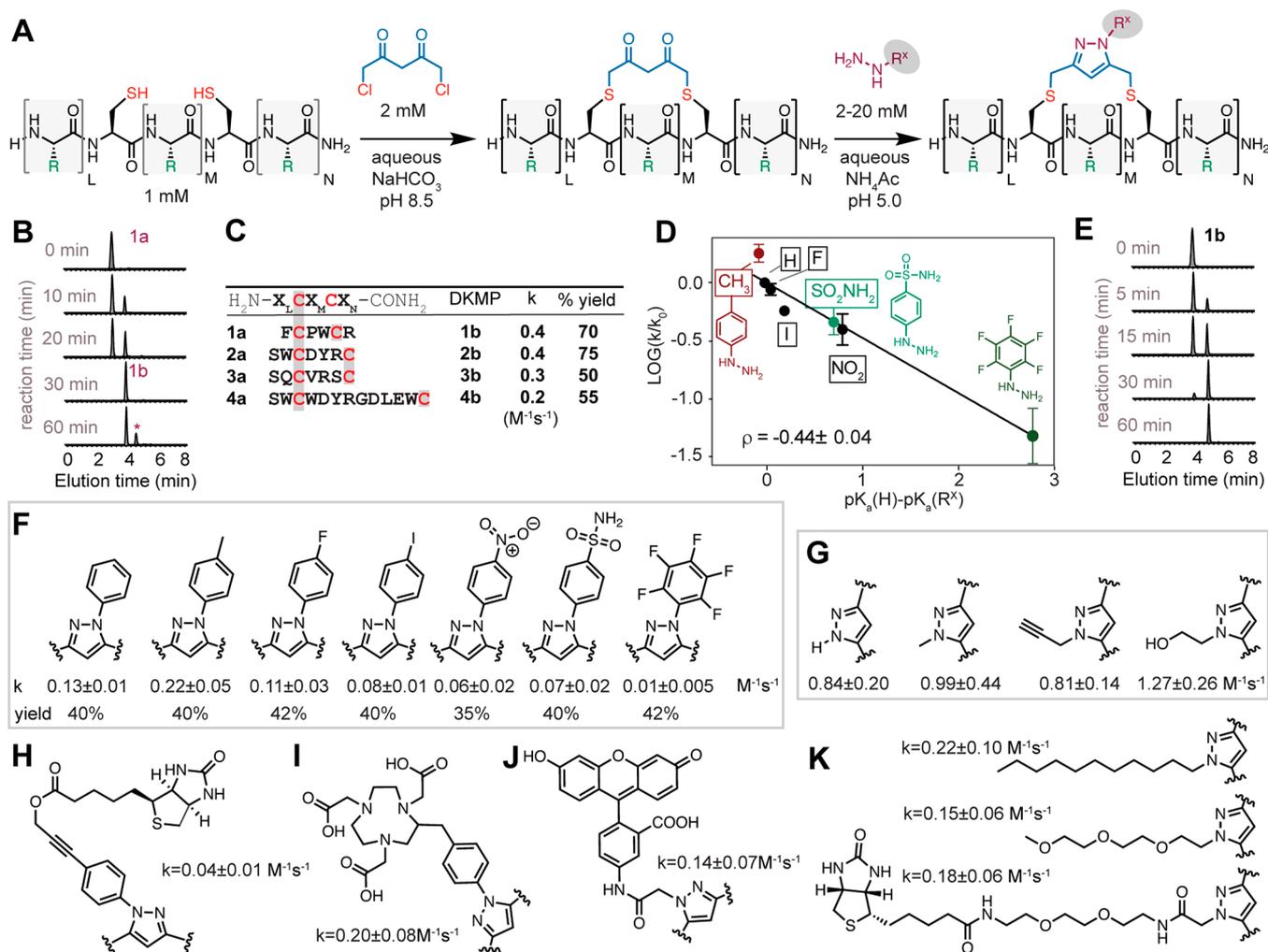


Figure 1. Model reactions on peptides. (A) Conversion of unprotected linear peptides to macrocycles with a diketone linchpin using dichloropentadione (DPD) at pH 8.5 and further functionalization of macrocycles using hydrazines at pH 5.0. (B) Liquid chromatography (LC) traces at 214 nm for the reaction between unprotected peptide **1a** and DPD. The reaction is complete at 30 min; prolonged reaction times lead to an emergence of a minor byproduct (*) that remains less than 5% of the DKMP product at 2 mM DPD concentration for 24 h and does not increase unless the DPD concentration is increased to 20 mM (detailed in Appendix 2 and SI Section 1.3). (C) Reactions between different peptides and DPD show similar reaction rates. (D) Ligation of aromatic hydrazines onto 1,3-diketone peptide displays negative Hammett correlation to the substituents on the phenyl ring indicating positive charge build up during the transition state. (E) LC traces for reaction between 1,3-diketone peptide and phenyl hydrazine (2 mM). The reaction is completed within 60 min. (F) Reaction rates (reaction rates are shown as mean (M) \pm 95% confidence interval, CI) for 1,3-diketone peptides (**3b**) with aryl hydrazines and yields for their HPLC purified products. (G) Reaction rates ($M \pm 95\%$ CI) for 1,3-diketone peptide (**3b**) and alkyl hydrazines. (H–K) Reaction rates ($M \pm 95\%$ CI) for 1,3-diketone peptide and hydrazine groups displaying affinity handles, imaging chelators, long alkyl chains, and oligoethylene glycol units.

range of glycans²⁹ or reactive covalent warheads³⁰ into peptide macrocycles. Alkylation with other bi- and tridentate electrophile linchpins were employed to introduce noncovalent and covalent warheads into T7³¹ and M13-displayed³² phage libraries. These approaches use macrocyclization as a divergent step and require optimization of nontrivial ring closing reactions for every linchpin structure. More robust approaches graft the desired unnatural fragments onto the preformed macrocycles.²¹ For example, Suga and co-workers converted free cysteines in mRNA displayed peptide macrocycles to dehydroalanines (DhA),³³ followed by Michael addition of thioglycosides to DhA.³⁴ Ketone functionality in DCA-modified phage-displayed libraries can be used to introduce glycans and other pharmacophores via oxime ligation.²⁸ Minor limitations of these approaches are the reversibility of formed bonds, slow reactivity of ketone, and reactivity of DhA functionality to

biological nucleophiles. To overcome these limitations, we describe synthesis and late-stage diversification of GE macrocyclic libraries with a 1,3-diketone reactive handle. We demonstrate that this approach offers advantages not present in prior reports such as stability to storage, superior reactivity, and formation of irreversible bond.

Aliphatic 1,3-diketones are *bona fide* bio-orthogonal moieties with long-term stability *in vivo*. The evidence for the stability of 1,3-diketones *in vivo* comes from work of Barbas and Lerner who immunized mice with 1,3-diketone haptens and isolated antibody reactive to 1,3-diketones.^{35,36} They demonstrated that 1,3-diketones injected into blood circulation reacted with the circulating anti-1,3-diketone antibody selectively³⁶ and only rare unique peptide sequences³⁷ had any detectable reactivity with this group. Kate Carroll and co-workers discovered that 1,3-diketones react with biomolecules that contain sulfenic acid: a

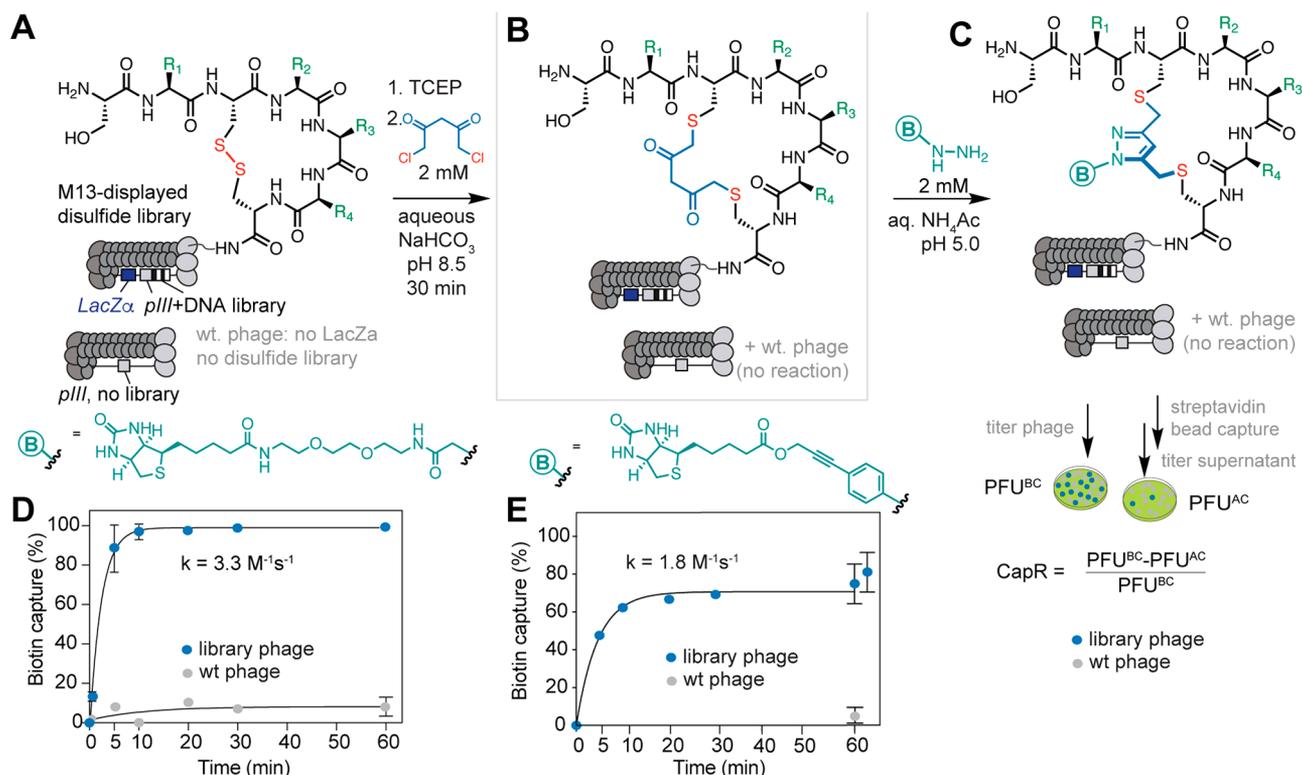


Figure 2. Macrocyclization of genetically encoded phage libraries (A, B) M13-phage displayed disulfide library is reduced by TCEP to yield reactive thiols and then DPD to form 1,3-diketone bearing macrocyclic library. (C) The 1,3-diketone linchpin bearing macrocycles were further functionalized with hydrazines, e.g., biotin hydrazine. A biotin capture experiment was used to measure the conversion of the 1,3-diketone. (D) The rate of conversion of the M13 library bearing 1,3-diketone was calculated using capture of biotinylated phage clones, using alkyl biotin hydrazine (BAH) reagent. (E) The rate of conversion of the M13 library bearing 1,3-diketone was calculated using capture of biotinylated phage clones, using phenyl biotin hydrazine (BPH) reagent.

transient species formed from endogenous cysteines due to oxidative stress.^{38–40} Sulfenic acid rapidly attacked nucleophilic carbon in derivatives of dione, 1,3 diketone in a six-membered ring, but this reaction was significantly slower with open chain aliphatic 1,3-diketones.⁴⁰ Unlike ketones that form reversible adducts with hydrazine-derivatives, 1,3-diketones undergo an irreversible cyclocondensation with hydrazines to generate 1,2-diazoles (pyrazoles). On the basis of these observations, we rationalized that the Knorr pyrazole reaction⁴¹ can be employed as bio-orthogonal late-stage functionalization of GE-libraries in aqueous, mild, biocompatible conditions.

RESULTS AND DISCUSSION

Substrate Scope of Macrocyclization and Late-Stage Modification of Unprotected Peptides in Water. Our design started from aqueous modification of genetically encoded peptide libraries displayed on phage with 1,5-dichloropentane-2,4-dione (DPD) to yield 1,3-diketone modified macrocyclic peptides (DKMP) (Figure 1A). We employed previously reported synthesis of DPD⁴² and confirmed its structure by X-ray crystallography (Figure S1). In model studies, DPD converted synthetic peptides of structure, X_nCX_mC to DKMP's within 30 min in pH 8.5 aqueous bicarbonate buffer (Figure 1A–C). We demonstrated that these DKMP's are ideally poised for late-stage functionalization using diverse alkyl and aryl hydrazine functionalities as precursors. Reaction in aqueous ammonium acetate buffer (pH 5.0) irreversibly grafted these functionalities onto the macrocycle via a hydrolytically stable endocyclic pyrazole (Figure 1A,E). Incubation of pyrazole

peptide product and hydrazine in ammonium acetate buffer for 7 days produced no detectable crossover product confirming the irreversibility of the formed bond (Figure S2). Substituted phenyl hydrazines formed N-aryl pyrazole macrocyclic peptides with rate constants ranging from 0.01 to 0.22 $\text{M}^{-1} \text{s}^{-1}$ (Figure 1D,F). Reaction between DKMP and diverse N-alkyl hydrazines occurred with $k = 0.81\text{--}1.27 \text{ M}^{-1} \text{s}^{-1}$ in the same conditions (Figure 1G). Hammett series plot for substituted phenyl hydrazines gave a $\rho = -0.44$ value, indicating a buildup of positive charge in the transition state of the rate-determining step (Figure 1C). Reaction between DKMP and the most electron poor perfluorophenyl hydrazine was completed after overnight incubation whereas other reactions were completed in <1-h incubation (Figure 1E, Figure S3–S13 and Table S1). The LC-MS monitoring of the pyrazole formation reactions revealed >95% conversion to the desired product irrespective of the hydrazine derivative. The isolated yields for this reaction were between 35 and 42%, and the difference is most likely due to loss in HPLC purification on milligram scale. We employed these reactions to demonstrate grafting of fluorophores, imaging chelators, *n*-alkyls, and poly ethylene glycol moieties onto macrocycles (Figure 1H–K); grafting of long chain *n*-alkyl chains (“lipidation”) or poly ethylene glycol moieties (“pegylation”) are known to enhance pharmacokinetic properties of peptides and diketone ligation can be used for prospective “lipidation” and “pegylation” of macrocyclic peptide libraries.

Optimization of Reactions on Phage-Displayed Libraries. When adapting the reaction conditions to phage-displayed peptide libraries, we employed well-established biotin-capture

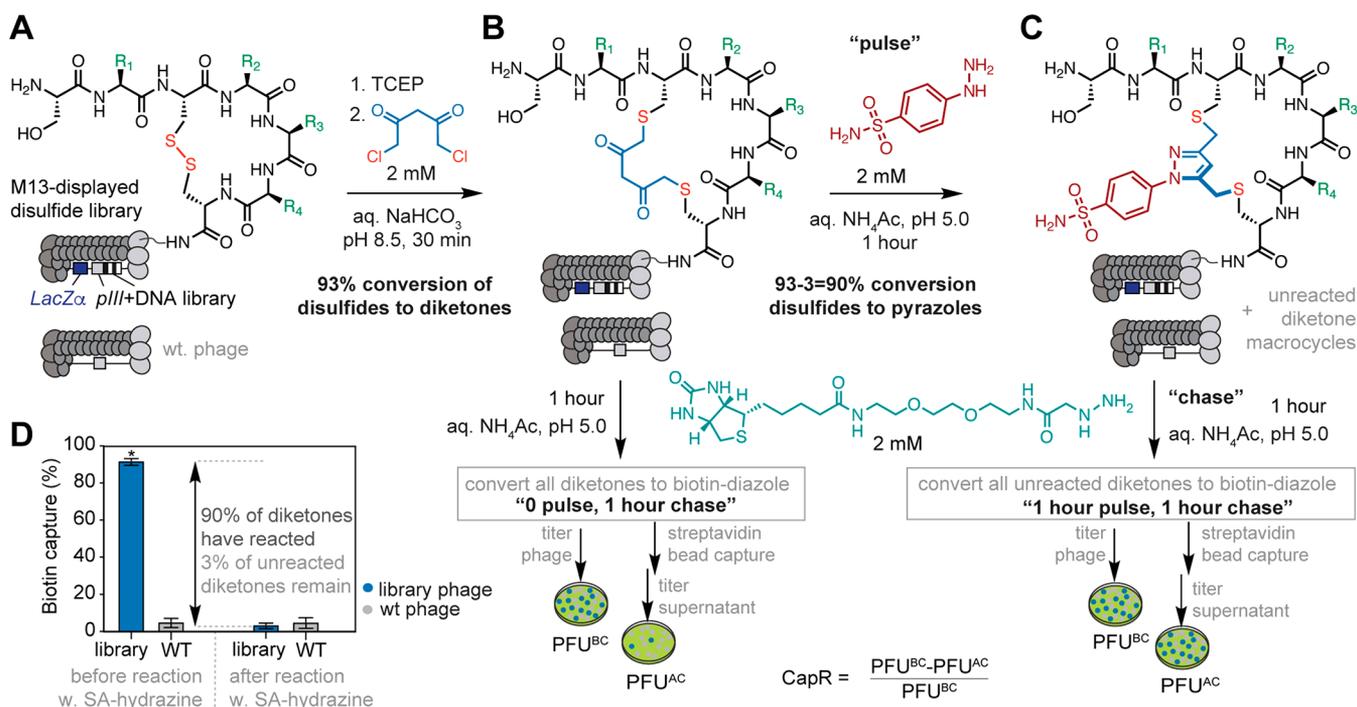


Figure 3. Pulse-chase monitoring of reactions in DKMP libraries. (A) M13-phage displayed disulfide library is reacted with TCEP and then DPD to form DKMP library. (B) Reaction of a portion of DKMP library with biotin hydrazine for 1 h detects that 90% of clones contain reactive 1,3-diketone groups. (C) Reaction of DKMP library with 4-hydrazine benzenesulfonamide (4HBS) for 1 h ("pulse") followed by biotin hydrazine for 1 h ("chase") allows detection of residual 1,3-diketone groups that were not consumed by 4HBS. (D) From 93% library with reactive 1,3-diketones, <3% were not consumed after 1 h reaction with 4HBS and can be biotinylated; these results indicate that 90% of the library were modified by 4HBS ($M \pm SD$, $n = 3$, $*p = 7 \times 10^{-5}$). Note that wild type phage present in the same solution is not biotinylated in any of the experiments indicating that only the disulfide library is modified with diketones and/or hydrazine derivatives.

coupled to plaque forming units (PFU) assay approach^{6,7,29,43–47} to measure the conversion, regioselectivity, and kinetics^{29,44–46} of chemical modification of phage displayed libraries (Figure 2). Hydrazine is known to damage DNA in bacteria, viruses/phages, and even in isolated DNA.⁴⁸ Indeed, PFU assay confirmed that most hydrazine derivatives killed >99.999% of infective phage particles in less than 5 min (Figure S14). Importantly, we observed that addition of metal chelator such as EDTA rescued this toxicity.⁴⁸ For example, in the presence of 5 mM EDTA, incubation of phage with 2 mM phenyl hydrazine for several hours did not show significant decrease in the number of infective particles. Importantly, EDTA did not influence the rate of reaction between diketone and hydrazine (Figure S14) and modification of DKMP-libraries on phage was possible in the presence of EDTA (Figure 2A–C).

Specifically, we modified phage-displayed SXCX₃C library of 130000 heptamer peptides⁴⁹ with 1 mM DPD for 30 min (Figure 2A,B) and then 2 mM biotin hydrazine probe +5 mM EDTA for 1 h and used biotin-capture assay to monitor the reaction efficiency (Figure 2C). We measured PFU before capture of the phage particles with streptavidin beads (PFU^{before}) and after capture (PFU^{after}) and used the capture ratio, $\text{CapR} = (\text{PFU}^{\text{before}} - \text{PFU}^{\text{after}}) / \text{PFU}^{\text{before}}$, to estimate the kinetics of the emergence of biotinylated phage particles at various reaction times (Figure 2D,E). As in previous reports, blue-white PFU assay in agar overlay supplemented with colorimetric substrate X-gal distinguished phage particles that displayed a library of peptides and transduced LacZα reporter (i.e., produced PFU of blue color) in the presence of wild type (wt) phage that displayed no Cys-containing peptides and transduced no reporter (i.e., produced PFU of white color).⁴⁵

We compared the reaction of SXCX₃C library and two different biotin hydrazine reagents: an alkyl biotin hydrazine containing a short peg linker (BAH) and a biotin phenyl hydrazine derivative (BPH) (Figure 2D,E). We observed faster reaction kinetics with BAH reagent ($k = 3.3 \text{ M}^{-1} \text{ s}^{-1}$), leading to the capture of >90% of the DKMP-SXCX₃C library, while the BPH reagent reacted at a lower rate ($k = 1.8 \text{ M}^{-1} \text{ s}^{-1}$) saturating the capture ratio around 75% at 60 min of reaction. Reaction was regioselective because the control wild type phage present in the same solution was not biotinylated and not captured (Figure 2D,E). Interestingly, reaction of a random decamer SXCX₆C library with an optimal BAH reagent reached a plateau at ~80% modification (Figure S15). We exposed the remaining, nonreactive 20% population to biotin peg iodoacetamide reagents (BIA) to cap any unreacted cysteines but could not detect any phage with unreacted cysteine residues (Figure S15D). However, when the nonreactive population was reamplified in bacteria, it could be modified by DPD and BAH again (Figure S16). These observations suggest that phage particles in the unreactive population do not display any peptides (e.g., due to proteolytic cleavage during expression) but the displayed peptides are re-expressed after amplification (Figure S16E,F). The intermediate DKMP library was stable in storage, as evidenced from reproducible yield of biotinylation of DKMP library after 1 month of storage (Figure S17). Reactivity of DKMP handle was unchanged after 10 days of incubation in protein and metabolite rich media (yeast extract, Figure S17).

Synthesis of Phage Displayed Macrocycles Libraries with Diverse Pharmacophores. To allow screening of the mixture of libraries modified by different unnatural fragments, we expressed chemically identical SXCX₃C phage libraries, each

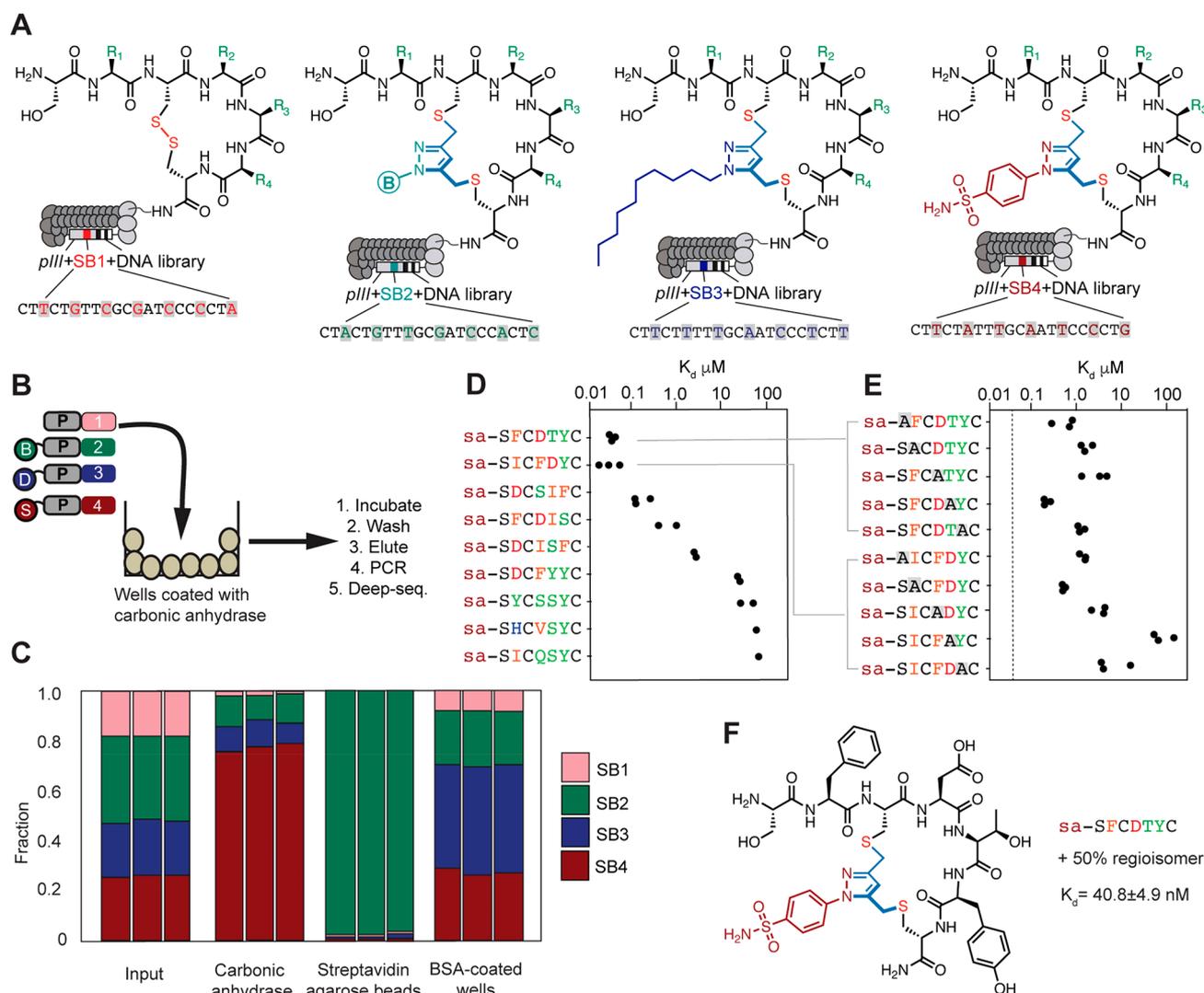


Figure 4. Selection of ligands from functionalized macrocyclic libraries. (A) Phage-displayed libraries of macrocycles with four distinct modifications and sequences of “silent barcodes” that encode these modifications. (B) A mixture of modified libraries was incubated with immobilized protein, followed by washing steps and acid elution. Eluted phage was PCR-amplified. The amplicons were sequenced by Illumina. (C) Abundance of silent barcodes in sequencing of libraries eluted from streptavidin, carbonic anhydrase, or BSA-coated wells showed an enrichment of specific barcodes that encode the fragments that bind to the corresponding protein target. Three bars describe the results of sequencing of three independent panning experiments. (D) K_d of binding to BCA for nine peptide macrocycles selected against BCA, determined with ITC. (E) K_d of binding to BCA for the alanine scans of the best two binders. The dotted line indicates the 40 nM point. (F) Structure of macrocyclic peptide displaying the lowest K_d value (40.8 ± 4.9 nM). Each macrocycle was synthesized as a 1:1 mixture of two regioisomers according to LC trace of LCMS.

containing a unique silent DNA barcode,⁴⁷ and modified each library by a different pharmacophore (Figure 3, Figure 4A). Silent-barcoded ^{SB2}SXCX₃C-DKMP library was modified by a biotin hydrazine forming biotinylated pyrazole-macrocyclic library (Figure 4A) in >95% yield (Figure S18). ^{SB3}SXCX₃C-DKMP library was modified by a *n*-decyl-hydrazine giving rise to “lipidated” pyrazole-macrocyclics; ^{SB4}SXCX₃C-DKMP library was converted to a library of macrocycles with a sulfonamide warhead (Figure 4A). Both reactions occurred in 90–96% modification yield, and to confirm this yield, we employed previously developed “pulse-chase” biotin capture (Figure 3, Figures S18, S19). In short, reaction between DKMP library and biotin hydrazine confirmed that 93% of the clones in ^{SB4}SXCX₃C-DKMP library contained the 1,3-diketone group (Figure 3A,B). Reaction between ^{SB4}SXCX₃C-DKMP and 4-hydrazine benzenesulfonamide for an hour (“pulse”) followed by addition of biotin hydrazine (“chase”) produced <3% of the

biotinylated clones (Figure 3C). This observation confirmed that in 90% of the library, the 1,3 diketone groups has been converted to benzenesulfonamide-pyrazole. Analogous pulse-chase confirmed 96% lipidation of ^{SB3}SXCX₃C-DKMP library (Figure S18).

Ligands Discovery from Phage Displayed Macrocyclic Libraries with Diverse Pharmacophores. Next, we demonstrated the utility of late-stage modified libraries in the discovery of macrocyclic ligands for protein receptors. We mixed the four libraries with DNA-encoded modifications (Figure 4A) in a 1:1:1:1 ratio to produce a library of $4 \times 130000 = 520000$ macrocycles in which the identity of both peptide sequence and the unnatural chemotype can be decoded by simple DNA sequencing. We performed 3–4 parallel instances of one round of panning of this mixed library against bovine carbonic anhydrase (BCA)^{50,51} (Figure 4B), streptavidin (Figure S20) and bovine serum albumin (BSA) and analyzed each output by

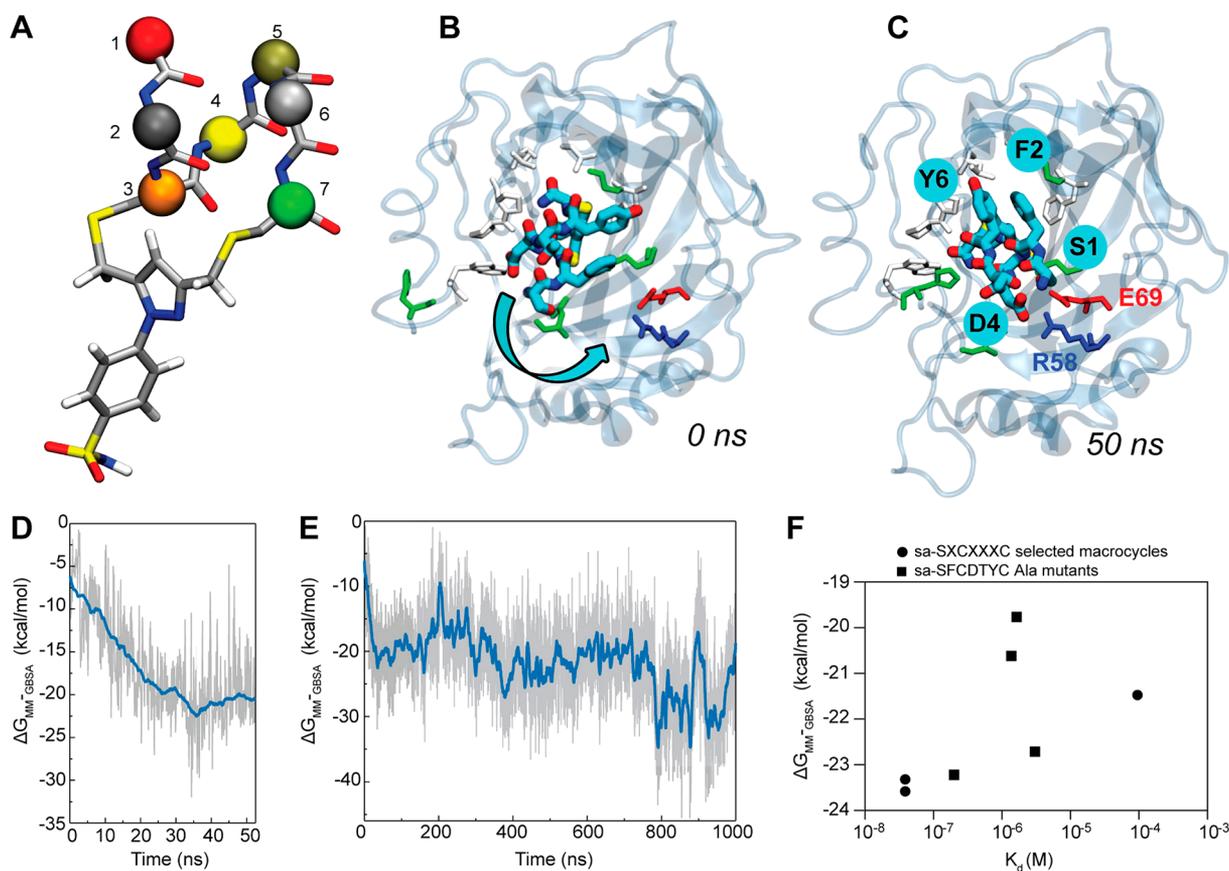


Figure 5. MD simulations of sa-SFCDTYC macrocycle. (A) Isomer of sa-SFCDTYC macrocycle used for MD simulations. (B) Initial (docked) binding pose of the macrocycle. (C) Macrocycle binding pose after 50 ns of simulations; the overall peptide orientation remains similar throughout a 1 μ s trajectory. (D) Free energy of binding between peptide and BCA, $\Delta G_{MM}^{G_{BSA}}$, during the reorientation in the first 50 ns of the trajectory. (E) Free energy of binding between peptide and BCA, $\Delta G_{MM}^{G_{BSA}}$, during the 1 μ s simulation trajectory. (F) The free energies of binding for the same macrocycles as in panel A calculated through MD simulations and their comparison experimental K_d values. In panels B and C, macrocycle is shown in thick licorice (C-teal, N-blue, O-red, S-yellow). BCA is shown in transparent new cartoon, with amino acid residues within 4 Å of the macrocycle at 50 ns shown in licorice (basic-blue, acidic-red, nonpolar-white, polar-green).

deep sequencing (Figure 4C). The sublibrary with the specific modification was enriched in a screen against a cognate target: $^{SB4}SXCX_3C$ -sulfonamide was enriched in panning against BCA and $^{SB2}SXCX_3C$ -biotin was enriched in panning against streptavidin when compared to the input (Figure 4C). Interestingly, we also observed a modest enrichment of $^{SB3}SXCX_3C$ -n-decyl sublibraries in panning on BSA-coated wells in line with known affinity to albumin.^{52–54}

To analyze the deep-sequencing data, we employed the previously published⁵⁵ Bioconductor EdgeR differential enrichment (DE) analysis^{56,57} with a negative binomial model, Trimmed Mean of M-values (TMM) normalization,⁵⁸ and Benjamini–Hochberg (BH)⁵⁹ correction to control the false discovery rate (FDR) at $\alpha = 0.05$. DE analysis of the output from BCA panning identified the 55 families of 688 peptide sequences that were significantly ($p < 0.05$) enriched in panning to BCA when compared to input and panning against BSA control (Figure S21, SI, S1). To confirm that the sequences identified by DE-analysis are the binders for targets of interest, we synthesized nine out of 688 predicted sequences and modified peptides with DPD and sulfonamide hydrazines to form peptide macrocycles with grafted sulfonamides (Figure 4D,E). Isothermal titration calorimetry (ITC) determined the binding constant K_d for the 4-hydrazine benzenesulfonamide to be $\sim 50 \mu M$. K_d of sulfonamide macrocycles produced from a random SICQSYC

sequence was $>100 \mu M$. From eight peptides predicted by DE analysis, six macrocycles exhibited 10–1000 fold enhancement in affinity (Figure 4D); two most potent macrocycles interacted with BCA with $K_d = 40.8 \pm 4.9$ nM (SFCDTYC) (Figure 4F) and 41.1 ± 26.1 nM (SICFDYC). We performed isothermal calorimetry (ITC) measurements for the alanine scans of the two best binders (Figure 4E) to identify the most important residues for binding to BCA. All of the single amino acid modifications to the macrocycles caused over 5-fold increase of K_d value when comparing to the original peptide sequence, indicating all residues are critical for recognition and removal of any results in significant disruptions for binding interactions. These results confirm that late-stage modified macrocycle libraries with multiple diverse unnatural fragments can be used for successful discovery of fragment-containing macrocycles with low-nanomolar potency.

Structural Characterization of Macrocycles:BCA Complex Using Molecular Dynamics. To study structures of complexes between sa-SFCDTYC and sa-SICFDYC macrocycles and BCA, we screened crystallization conditions for BCA both as the apo-form and in the presence of either ligand. Unfortunately, no crystallization conditions, including the previously reported conditions,^{60,61} yielded crystals of sufficient quality. As an alternative, we docked the two isoforms of the nine macrocycles selected from panning (Figure 4D, Table S4) to

previously reported structure of BCA (6SKV), employing Autodock Vina.⁶² In all but one docked binding pose, negatively charged benzenesulfonamide moieties⁶³ were in the substrate cleft, with <4 Å distances between S atoms of benzenesulfonamide and Zn²⁺ (Table S4). Next, we employed MD simulation for 1 μs to examine the docked binding pose of BCA and sa-SFCDTYC macrocycle complex. For one of the regio-isomers, the macrocycle reoriented from the docked pose to a new pose within the first 20 ns (Figure 5A–C, Figure S52A and SI video 1), while the other isomer took longer to reach a stable conformation. Therefore, we moved forward with the isomer that reoriented within a shorter time interval into a low energy conformation. After reorientation, the macrocycle established several favorable contacts. The S1 forms hydrogen bonds with E69 (BCA) with its N-terminus and –OH groups for 68% of 1 μs trajectory (Figure S52A,B). The reorientation enabled a formation of the salt bridge between D4 and R58 (BCA), and formation of nonpolar contacts between F2 and Y6 to BCA (Figure 5B). Overall, the macrocycle reorientation process is characterized by a decrease in its free energy of binding (Figure 5D), with $\Delta G_{\text{MM-GBSA}}$ changing from the initial value of –7 kcal/mol to a more favorable value of –20 kcal/mol within first 50 ns. The overall peptide orientation in the binding pose shown in Figure 5C remained similar throughout a 1 μs long trajectory, with rearrangements of the side chains contributing to further decreases and fluctuations in the value of $\Delta G_{\text{MM-GBSA}}$ (Figure 5E).

We compared the MD simulations of the sa-SFCDTYC:BCA complex to the conformational ensemble of free sa-SFCDTYC macrocycles in aqueous solution. Within first several nanoseconds, the peptide fragment changed from its initial helical conformation to a random coil conformation (Figure S53A) that fluctuated over time. We monitored the structure of the peptide backbone over time with the root-mean-square deviations (RMSD) measured with respect to a reference state with the helical peptide backbone (Figure S53B,D). The RMSD increased within nanoseconds, as the peptide backbone underwent stepwise conformational changes (certain conformations are preserved for tens of nanoseconds). The control SFCDTYC macrocycle with a disulfide bond between C3 and C7 amino acids also transitioned from helical conformation to the random coil within nanoseconds. However, its RMSD plot, evaluated with respect to the helical peptide backbone conformation, fluctuated without stepwise conformational changes observed with pyrazole formation, indicating that the pyrazole modification increases the rigidity of the backbone of sa-SFCDTYC compared to the disulfide peptide (Figure S53C,D).

To juxtapose the results of the MD simulations of BCA:macrocycle complexes and experimental observations, several macrocycles were prepared in the binding pose of Figure 5B and examined in short MD simulations. The examined macrocycle sequences were SFCDTYC, SICFDYC, and SICQSYC, and single alanine mutants of SFCDTYC sequence in positions 2, 4, 5, and 6. The systems with the macrocycles were prepared by extracting a relaxed binding pose from MD simulations of BCA and sa-SFCDTYC and mutating the amino acids to the desired sequence. Extended MD simulation of the complexes was critical because we did not observe any correlation between docking scores produced by Autodock Vina and K_d values for macrocycles (Figure S54A), in line with the acknowledged challenges in docking of static peptide structures.⁶⁴ In contrast to static docking scores, the calculated

free energies of binding of these macrocycles and BCA from the MD simulation, $\Delta G_{\text{MM-GBSA}}$, correlated reasonably with the experimental K_d values obtained with ITC studies (Figure 5F, Table S5). The hydrophobic interactions observed in the pose from MD calculations were consistent with experimentally observed 50-fold increase in K_d value upon Phe → Ala and Tyr → Ala substitutions in both macrocycles (Figure 4E). Changing the Asp to an Ala in both the macrocycles also caused over a 100-fold increase in the K_d values. These experimental observations were corroborated by MD simulations that yielded the hydrogen bonding and ionic interactions between BCA and Asp in the macrocycles (Figure 5F).

CONCLUSION

The broad substrate scope of Knorr pyrazole synthesis makes it an attractive strategy for the diversification of macrocycles with built in 1,3-diketones using a large range of commercially available alkyl and aryl-hydrazine warheads. So far, we identified only two substrates with suboptimal reactivity: reaction between 1,3-diketone-macrocycle and *N*-acyl hydrazines and benzenesulfonohydrazide in water was slow or incomplete. Even when formed, *N*-acyl 1,2-pyrazoles can be readily cleaved by thiols⁶⁵ and other biological nucleophiles; such cleavage makes *N*-acyl 1,2-pyrazoles not suitable for the stable grafting of functionalities onto peptides. Further substrate scope profiling may uncover further limitations; however, we foresee few problems in reactions that employ simple alkyl and aryl hydrazine fragments.

A limitation of the Knorr-pyrazole ligation is formation of two regioisomers of pyrazole: we observed the formation of a 1:1 mixture of two isomers in many LC-traces; interestingly, this ratio was skewed toward one isomer in reaction with perfluorophenyl hydrazine (Figure S10). The reason for this regio-induction is presently not clear. The formation of isomers is a trait of many contemporary modifications of polypeptides: for example, reactions between Cys and maleimide yield mixtures of stereoisomers. However, this reaction is used in the manufacturing of FDA-approved antibody-drug conjugates (ADC) such as Ketruda, Trodelvy, Enhertu, Polivy, Adcetris, and Padcev. A ligation to dehydroalanine on polypeptides^{34,66,67} yields two diastereomers and has been successfully translated to the manufacture of ADCs.³³ Pictet-Spengler⁶⁸ and Hydrazino-iso-Pictet-Spengler (HIPS) reactions that yield diastereomeric linkages are employed in the manufacture of TRPH-222 ADC, which is currently in a Phase 1 clinical trial. Many reagents for the modification of protein, via strain-promoted cycloadditions and inverse demand Diels–Alder reactions, form isomeric linkages. Other bioorthogonal ligations of aldehydes⁶⁹ to hydrazines, oximes,⁴⁵ 2-amino benzamidoximes,⁴⁴ and Wittig ylides⁴⁶ form *E* and *Z* products. In mRNA- and phage-displayed libraries, reactions that yield a mixture of stereo^{7,34} or regioisomers^{7,70,71} have been employed as well. In such mixed-isomer libraries, activity can be attributed to one synthetic isomer postdiscovery.^{7,70,71} Formation of two regioisomers, thus, is not an impediment to a GE-discovery process: In this report, the macrocyclic peptides with phenyl sulfonamide fragments were discovered and synthesized postdiscovery as mixtures of two regioisomers. It is likely that one isomer has higher activity than the other, but we did not attempt to measure the activities of separated isomers.

In this report, important advances were made to late stage modifications of GE-macrocycles described in the pioneering report of GE-FBD by Roberts⁷² and Dwyer et al.⁷³ Simplicity and robustness of chemical modification make it simple to adapt

it to many existing phage and mRNA-displayed cysteine-containing phage libraries.^{34,74} These libraries when modified with DPD should yield shelf-stable, divergent macrocyclic precursors to GE-macrocyclic libraries with unnatural fragments. It is possible to produce GE-macrocyclic libraries with unnatural fragments by direct translation via unnatural amino acid mutagenesis,^{75–77} metabolic suppression,^{78,79} and flexigyme technology.^{11,26} However, the unique advantage of late stage chemical modifications is the ability to introduce large functionalities such as fluorophores, metal chelators (Figure 1I,J), or complex glycans^{34,80,81} (not shown in this report): it might not be possible to introduce such groups via translational machinery. The selection of peptides premodified with fluorescent probes or metal-chelating probes (Figure 1I,J) offers an interesting opportunity to discover peptides for imaging while minimizing the commonly observed decrease in potency or specificity of peptide due to the conjugation of imaging probes. Similarly, the GE-screening of prospective pegylated or lipidated macrocycles (Figure 1K) can provide the pegylated and lipidated ligands and reduce the number of steps for optimization of leads. The “late” nature of 1,3-diketone ligation makes it an interesting candidate for the ligation of reactive warheads for genetically encoded discovery of covalent³⁰ or reversible covalent inhibitors.⁶ A combination of this modification with silent encoding opens new opportunities in the encoding of a diverse number of macrocycles with pharmacophores, warheads, and functionalities that would be difficult to introduce by other methods.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/jacs.1c01186>.

Detailed synthetic methods, biochemical methods describing the synthesis and selection of phage libraries, isothermal titration calorimetry (ITC) assay or protein–ligand binding, data processing methods describing the analysis of the DNA sequencing data, statistical methods, and computational calculations and predictions for bovine carbonic anhydrase (BCA)-macrocyclic binding and solution conformations of the macrocycles (PDF)

LCMS monitoring of byproduct generation, NMR spectra of byproduct with peptide 3b, synthesis and characterization of diketone and extra products from EDT (PDF)

Source data: submitted as “data.zip” contain files describing “Kinetics Matlab” directory with raw data used to monitor the kinetics of reactions and MatLab scripts for curve fit; “Sequence files” directory with *.txt files describing the raw deep-sequencing data, *.xlsx tables describing the silent barcoding, *.xlsx tables describing the differential enrichment (DE) analysis; Supplementary Files S1 and S2 describing the output of differential enrichment analysis and clustering; Titters.xlsx describing the phage titers (PFU) for all experiments described in this manuscript; .gif file showing the binding pose of sa-
SFCDTYC:BCA (ZIP)

Accession Codes

CCDC 2001271 contains the supplementary crystallographic data for this paper. These data can be obtained free of charge via www.ccdc.cam.ac.uk/data_request/cif, or by emailing data_request@ccdc.cam.ac.uk, or by contacting The Cambridge

Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, UK; fax: +44 1223 336033.

■ AUTHOR INFORMATION

Corresponding Author

Ratmir Derda – Department of Chemistry, University of Alberta, Edmonton, AB T6G 2G2, Canada; orcid.org/0000-0003-1365-6570; Email: ratmir@ualberta.ca

Authors

Arunika I. Ekanayake – Department of Chemistry, University of Alberta, Edmonton, AB T6G 2G2, Canada

Lena Sobze – Department of Chemistry, University of Alberta, Edmonton, AB T6G 2G2, Canada

Payam Kelich – Department of Chemistry and Biochemistry, University of Texas at El Paso, El Paso, Texas 79968, United States

Jihe Youk – Department of Chemistry, University of Alberta, Edmonton, AB T6G 2G2, Canada

Nicholas J. Bennett – Department of Chemistry, University of Alberta, Edmonton, AB T6G 2G2, Canada

Raja Mukherjee – Department of Chemistry, University of Alberta, Edmonton, AB T6G 2G2, Canada

Atul Bhardwaj – Cross Cancer Institute, University of Alberta, Edmonton, AB T6G 1Z2, Canada

Frank Wuest – Department of Chemistry and Cross Cancer Institute, University of Alberta, Edmonton, AB T6G 2G2, Canada; orcid.org/0000-0002-6705-6450

Lela Vukovic – Department of Chemistry and Biochemistry, University of Texas at El Paso, El Paso, Texas 79968, United States; orcid.org/0000-0002-9053-5708

Complete contact information is available at:

<https://pubs.acs.org/doi/10.1021/jacs.1c01186>

Notes

The authors declare the following competing financial interest(s): R.D. is the C.E.O. and a shareholder of 48Hour Discovery Inc., the company that licensed the patent application describing silent encoding and chemical modification technologies.

Data Availability: All raw deep-sequencing data is publicly available on <http://48hdcloud.ca/> with data-specific URL listed in Supplementary Table S2. MatLab, Python, and R scripts used for analysis of deep-seq data have been deposited to <https://github.com/derdalab/diketone>.

■ ACKNOWLEDGMENTS

The authors acknowledge funding from NSERC (RGPIN-2016-402511 to R.D.) and NSERC Accelerator Supplement (to R.D.). Infrastructure support was provided by CFI New Leader Opportunity (to R.D.). L.S. acknowledges summer research fellowship from GlycoNet. We thank Dr. Ryan T. McKay at the University of Alberta NMR spectrometry facility, Dr. Randy Whittall and Béla Reiz at the mass spectrometry facility, and Dr. Michael J. Ferguson at the X-ray crystallography laboratory (Chemistry Department). We thank Dr. Jessica Cao at 48Hour Discovery Inc. for assistance with analysis of the sequencing data.

■ REFERENCES

(1) Cohen, D. T.; Zhang, C.; Fadzen, C. M.; Mijalis, A. J.; Hie, L.; Johnson, K. D.; Shriver, Z.; Plante, O.; Miller, S. J.; Buchwald, S. L.; Pentelute, B. L. A chemoselective strategy for late-stage functionaliza-

tion of complex small molecules with polypeptides and proteins. *Nat. Chem.* **2019**, *11* (1), 78–85.

(2) Leroux, M.; Vorherr, T.; Lewis, I.; Schaefer, M.; Koch, G.; Karaghiosoff, K.; Knochel, P. Late-Stage Functionalization of Peptides and Cyclopeptides Using Organozinc Reagents. *Angew. Chem., Int. Ed.* **2019**, *58* (24), 8231–8234.

(3) Noisier, A. F. M.; Johansson, M. J.; Knerr, L.; Hayes, M. A.; Drury, W. J., 3rd; Valeur, E.; Malins, L. R.; Gopalakrishnan, R. Late-Stage Functionalization of Histidine in Unprotected Peptides. *Angew. Chem., Int. Ed.* **2019**, *58* (52), 19096–19102.

(4) Tang, J.; Chen, H.; He, Y.; Sheng, W.; Bai, Q.; Wang, H. Peptide-guided functionalization and macrocyclization of bioactive peptid-sulfonamides by Pd(II)-catalyzed late-stage C-H activation. *Nat. Commun.* **2018**, *9* (1), 3383.

(5) Heinis, C.; Rutherford, T.; Freund, S.; Winter, G. Phage-encoded combinatorial chemical libraries based on bicyclic peptides. *Nat. Chem. Biol.* **2009**, *5* (7), 502–7.

(6) McCarthy, K. A.; Kelly, M. A.; Li, K.; Cambray, S.; Hosseini, A. S.; van Opijnen, T.; Gao, J. Phage Display of Dynamic Covalent Binding Motifs Enables Facile Development of Targeted Antibiotics. *J. Am. Chem. Soc.* **2018**, *140* (19), 6137–6145.

(7) Ng, S.; Lin, E.; Kitov, P. I.; Tjhung, K. F.; Gerlits, O. O.; Deng, L.; Kasper, B.; Sood, A.; Paschal, B. M.; Zhang, P.; Ling, C. C.; Klassen, J. S.; Noren, C. J.; Mahal, L. K.; Woods, R. J.; Coates, L.; Derda, R. Genetically encoded fragment-based discovery of glycopeptide ligands for carbohydrate-binding proteins. *J. Am. Chem. Soc.* **2015**, *137* (16), 5248–51.

(8) Ven Chang, I.; Tsutsumi, H.; Mihara, H. Screening for concanavalin A binders from a mannose-modified alpha-helix peptide phage library. *Mol. BioSyst.* **2017**, *13* (11), 2222–2225.

(9) Vinogradov, A. A.; Yin, Y.; Suga, H. Macrocytic Peptides as Drug Candidates: Recent Progress and Remaining Challenges. *J. Am. Chem. Soc.* **2019**, *141* (10), 4167–4181.

(10) Derda, R.; Ng, S. Genetically encoded fragment-based discovery. *Curr. Opin. Chem. Biol.* **2019**, *50*, 128–137.

(11) Huang, Y.; Wiedmann, M. M.; Suga, H. RNA Display Methods for the Discovery of Bioactive Macrocycles. *Chem. Rev.* **2019**, *119* (17), 10360–10391.

(12) Richardson, S. L.; Dods, K. K.; Abrigo, N. A.; Iqbal, E. S.; Hartman, M. C. In vitro genetic code reprogramming and expansion to study protein function and discover macrocyclic peptide ligands. *Curr. Opin. Chem. Biol.* **2018**, *46*, 172–179.

(13) Murray, C. W.; Rees, D. C. The rise of fragment-based drug discovery. *Nat. Chem.* **2009**, *1* (3), 187–92.

(14) Scott, D. E.; Coyne, A. G.; Hudson, S. A.; Abell, C. Fragment-based approaches in drug discovery and chemical biology. *Biochemistry* **2012**, *51* (25), 4990–5003.

(15) Bollag, G.; Hirth, P.; Tsai, J.; Zhang, J.; Ibrahim, P. N.; Cho, H.; Spevak, W.; Zhang, C.; Zhang, Y.; Habets, G.; Burton, E. A.; Wong, B.; Tsang, G.; West, B. L.; Powell, B.; Shellooe, R.; Marimuthu, A.; Nguyen, H.; Zhang, K. Y.; Artis, D. R.; Schlessinger, J.; Su, F.; Higgins, B.; Iyer, R.; D'Andrea, K.; Koehler, A.; Stumm, M.; Lin, P. S.; Lee, R. J.; Grippo, J.; Puzanov, I.; Kim, K. B.; Ribas, A.; McArthur, G. A.; Sosman, J. A.; Chapman, P. B.; Flaherty, K. T.; Xu, X.; Nathanson, K. L.; Nolop, K. Clinical efficacy of a RAF inhibitor needs broad target blockade in BRAF-mutant melanoma. *Nature* **2010**, *467* (7315), 596–9.

(16) Perera, T. P. S.; Jovcheva, E.; Mevellec, L.; Vialard, J.; De Lange, D.; Verhulst, T.; Paulussen, C.; Van De Ven, K.; King, P.; Freyne, E.; Rees, D. C.; Squires, M.; Saxty, G.; Page, M.; Murray, C. W.; Gilissen, R.; Ward, G.; Thompson, N. T.; Newell, D. R.; Cheng, N.; Xie, L.; Yang, J.; Platero, S. J.; Karkera, J. D.; Moy, C.; Angibaud, P.; Laquerre, S.; Lorenzi, M. V. Discovery and Pharmacological Characterization of JNJ-42756493 (Erdafitinib), a Functionally Selective Small-Molecule FGFR Family Inhibitor. *Mol. Cancer Ther.* **2017**, *16* (6), 1010–1020.

(17) Souers, A. J.; Levenson, J. D.; Boghaert, E. R.; Ackler, S. L.; Catron, N. D.; Chen, J.; Dayton, B. D.; Ding, H.; Enschede, S. H.; Fairbrother, W. J.; Huang, D. C.; Hymowitz, S. G.; Jin, S.; Khaw, S. L.; Kovar, P. J.; Lam, L. T.; Lee, J.; Maecker, H. L.; Marsh, K. C.; Mason, K. D.; Mitten, M. J.; Nimmer, P. M.; Oleksijew, A.; Park, C. H.; Park, C.

M.; Phillips, D. C.; Roberts, A. W.; Sampath, D.; Seymour, J. F.; Smith, M. L.; Sullivan, G. M.; Tahir, S. K.; Tse, C.; Wendt, M. D.; Xiao, Y.; Xue, J. C.; Zhang, H.; Humerickhouse, R. A.; Rosenberg, S. H.; Elmore, S. W. ABT-199, a potent and selective BCL-2 inhibitor, achieves antitumor activity while sparing platelets. *Nat. Med.* **2013**, *19* (2), 202–8.

(18) Tsai, J.; Lee, J. T.; Wang, W.; Zhang, J.; Cho, H.; Mamo, S.; Bremer, R.; Gillette, S.; Kong, J.; Haass, N. K.; Sproesser, K.; Li, L.; Smalley, K. S.; Fong, D.; Zhu, Y. L.; Marimuthu, A.; Nguyen, H.; Lam, B.; Liu, J.; Cheung, I.; Rice, J.; Suzuki, Y.; Luu, C.; Settachatgul, C.; Shellooe, R.; Cantwell, J.; Kim, S. H.; Schlessinger, J.; Zhang, K. Y.; West, B. L.; Powell, B.; Habets, G.; Zhang, C.; Ibrahim, P. N.; Hirth, P.; Artis, D. R.; Herlyn, M.; Bollag, G. Discovery of a selective inhibitor of oncogenic B-Raf kinase with potent antimelanoma activity. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (8), 3041–6.

(19) Erlanson, D. A. Introduction to fragment-based drug discovery. *Top. Curr. Chem.* **2011**, *317*, 1–32.

(20) Franzini, R. M.; Neri, D.; Scheuermann, J. DNA-encoded chemical libraries: advancing beyond conventional small-molecule libraries. *Acc. Chem. Res.* **2014**, *47* (4), 1247–55.

(21) Li, Y.; De Luca, R.; Cazzamalli, S.; Pretto, F.; Bajic, D.; Scheuermann, J.; Neri, D. Versatile protein recognition by the encoded display of multiple chemical elements on a constant macrocyclic scaffold. *Nat. Chem.* **2018**, *10* (4), 441–448.

(22) Denton, K. E.; Wang, S.; Gignac, M. C.; Milosevich, N.; Hof, F.; Dykhuizen, E. C.; Krusemark, C. J. Robustness of In Vitro Selection Assays of DNA-Encoded Peptidomimetic Ligands to CBX7 and CBX8. *SLAS discovery: advancing life sciences R & D* **2018**, *23* (5), 417–428.

(23) Cai, B.; Kim, D.; Akhand, S.; Sun, Y.; Cassell, R. J.; Alpsy, A.; Dykhuizen, E. C.; Van Rijn, R. M.; Wendt, M. K.; Krusemark, C. J. Selection of DNA-Encoded Libraries to Protein Targets within and on Living Cells. *J. Am. Chem. Soc.* **2019**, *141* (43), 17057–17061.

(24) Buller, F.; Steiner, M.; Frey, K.; Mircsof, D.; Scheuermann, J.; Kalisch, M.; Buhlmann, P.; Supuran, C. T.; Neri, D. Selection of Carbonic Anhydrase IX Inhibitors from One Million DNA-Encoded Compounds. *ACS Chem. Biol.* **2011**, *6* (4), 336–44.

(25) Goodnow, R. A., Jr.; Dumelin, C. E.; Keefe, A. D. DNA-encoded chemistry: enabling the deeper sampling of chemical space. *Nat. Rev. Drug Discovery* **2017**, *16* (2), 131–147.

(26) Passioura, T.; Katoh, T.; Goto, Y.; Suga, H. Selection-based discovery of druglike macrocyclic peptides. *Annu. Rev. Biochem.* **2014**, *83*, 727–52.

(27) Hetrick, K. J.; Walker, M. C.; van der Donk, W. A. Development and Application of Yeast and Phage Display of Diverse Lanthipeptides. *ACS Cent. Sci.* **2018**, *4* (4), 458–467.

(28) Assem, N.; Ferreira, D. J.; Wolan, D. W.; Dawson, P. E. Acetone-Linked Peptides: A Convergent Approach for Peptide Macrocyclization and Labeling. *Angew. Chem., Int. Ed.* **2015**, *54* (30), 8665–8.

(29) Ng, S.; Derda, R. Phage-displayed macrocyclic glycopeptide libraries. *Org. Biomol. Chem.* **2016**, *14* (24), 5539–45.

(30) Chen, S.; Lovell, S.; Lee, S.; Fellner, M.; Mace, P. D.; Bogoy, M. Identification of highly selective covalent inhibitors by phage display. *Nat. Biotechnol.* **2020**, DOI: 10.1038/s41587-020-0733-7.

(31) Uematsu, S.; Tabuchi, Y.; Ito, Y.; Taki, M. Combinatorially Screened Peptide as Targeted Covalent Binder: Alteration of Bait-Conjugated Peptide to Reactive Modifier. *Bioconjugate Chem.* **2018**, *29* (6), 1866–1871.

(32) Ernst, C.; Sindlinger, J.; Schwarzer, D.; Koch, P.; Boeckler, F. M. The Symmetric Tetravalent Sulfhydryl-Specific Linker NATBA Facilitates a Combinatorial "Tool Kit" Strategy for Phage Display-Based Selection of Functionalized Bicyclic Peptides. *ACS omega* **2018**, *3* (10), 12361–12368.

(33) Dadova, J.; Galan, S. R.; Davis, B. G. Synthesis of modified proteins via functionalization of dehydroalanine. *Curr. Opin. Chem. Biol.* **2018**, *46*, 71–81.

(34) Jongkees, S. A. K.; Umamoto, S.; Suga, H. Linker-free incorporation of carbohydrates into in vitro displayed macrocyclic peptides. *Chem. Sci.* **2017**, *8* (2), 1474–1481.

(35) Barbas, C. F., 3rd; Heine, A.; Zhong, G.; Hoffmann, T.; Gramatikova, S.; Bjornstedt, R.; List, B.; Anderson, J.; Stura, E. A.;

- Wilson, I. A.; Lerner, R. A. Immune versus natural selection: antibody aldolases with enzymic rates but broader scope. *Science* **1997**, *278* (5346), 2085–92.
- (36) Rader, C.; Turner, J. M.; Heine, A.; Shabat, D.; Sinha, S. C.; Wilson, I. A.; Lerner, R. A.; Barbas, C. F. A humanized aldolase antibody for selective chemotherapy and adaptor immunotherapy. *J. Mol. Biol.* **2003**, *332* (4), 889–99.
- (37) Tanaka, F.; Fuller, R.; Asawapornmongkol, L.; Warsinke, A.; Gobuty, S.; Barbas, C. F., 3rd Development of a small peptide tag for covalent labeling of proteins. *Bioconjugate Chem.* **2007**, *18* (4), 1318–24.
- (38) Gupta, V.; Carroll, K. S. Profiling the Reactivity of Cyclic C-Nucleophiles towards Electrophilic Sulfur in Cysteine Sulfenic Acid. *Chem. Sci.* **2016**, *7* (1), 400–415.
- (39) Gupta, V.; Carroll, K. S. Rational design of reversible and irreversible cysteine sulfenic acid-targeted linear C-nucleophiles. *Chem. Commun. (Cambridge, U. K.)* **2016**, *52* (16), 3414–7.
- (40) Gupta, V.; Paritala, H.; Carroll, K. S. Reactivity, Selectivity, and Stability in Sulfenic Acid Detection: A Comparative Study of Nucleophilic and Electrophilic Probes. *Bioconjugate Chem.* **2016**, *27* (5), 1411–1418.
- (41) Knorr, L. Einwirkung von Acetessigester auf Phenylhydrazin. *Ber. Dtsch. Chem. Ges.* **1883**, *16* (2), 2597–2599.
- (42) Rychnovsky, S. D.; Griesgraber, G.; Zeller, S.; Skalitzky, D. J. Optically pure 1,3-diols from (2R,4R)- and (2S,4S)-1,2:4,5-diepoxy-pentane. *J. Org. Chem.* **1991**, *56* (17), 5161–5169.
- (43) Kalhor-Monfared, S.; Jafari, M. R.; Patterson, J. T.; Kitov, P. I.; Dwyer, J. J.; Nuss, J. M.; Derda, R. Rapid biocompatible macrocyclization of peptides with decafluoro-diphenylsulfone. *Chem. Sci.* **2016**, *7* (6), 3785–3790.
- (44) Kitov, P. I.; Vinals, D. F.; Ng, S.; Tjhung, K. F.; Derda, R. Rapid, hydrolytically stable modification of aldehyde-terminated proteins and phage libraries. *J. Am. Chem. Soc.* **2014**, *136* (23), 8149–52.
- (45) Ng, S.; Jafari, M. R.; Matochko, W. L.; Derda, R. Quantitative synthesis of genetically encoded glycopeptide libraries displayed on M13 phage. *ACS Chem. Biol.* **2012**, *7* (9), 1482–7.
- (46) Triana, V.; Derda, R. Tandem Wittig/Diels-Alder diversification of genetically encoded peptide libraries. *Org. Biomol. Chem.* **2017**, *15* (37), 7869–7877.
- (47) Tjhung, K. F.; Kitov, P. I.; Ng, S.; Kitova, E. N.; Deng, L.; Klassen, J. S.; Derda, R. Silent Encoding of Chemical Post-Translational Modifications in Phage-Displayed Libraries. *J. Am. Chem. Soc.* **2016**, *138* (1), 32–5.
- (48) Kimball, R. F. The mutagenicity of hydrazine and some of its derivatives. *Mutat. Res., Rev. Genet. Toxicol.* **1977**, *39* (2), 111–26.
- (49) He, B.; Tjhung, K. F.; Bennett, N. J.; Chou, Y.; Rau, A.; Huang, J.; Derda, R. Compositional Bias in Naive and Chemically-modified Phage-Displayed Libraries uncovered by Paired-end Deep Sequencing. *Sci. Rep.* **2018**, *8* (1), 1214.
- (50) Krishnamurthy, V. M.; Kaufman, G. K.; Urbach, A. R.; Gitlin, I.; Gudixsen, K. L.; Weibel, D. B.; Whitesides, G. M. Carbonic Anhydrase as a Model for Biophysical and Physical-Organic Studies of Proteins and Protein–Ligand Binding. *Chem. Rev.* **2008**, *108* (3), 946–1051.
- (51) Melkko, S.; Scheuermann, J.; Dumelin, C. E.; Neri, D. Encoded self-assembling chemical libraries. *Nat. Biotechnol.* **2004**, *22* (5), 568–74.
- (52) Bech, E. M.; Pedersen, S. L.; Jensen, K. J. Chemical Strategies for Half-Life Extension of Biopharmaceuticals: Lipidation and Its Alternatives. *ACS Med. Chem. Lett.* **2018**, *9* (7), 577–580.
- (53) Fu, C.; Chen, Q.; Zheng, F.; Yang, L.; Li, H.; Zhao, Q.; Wang, X.; Wang, L.; Wang, Q. Genetically Encoding a Lipidated Amino Acid for Extension of Protein Half-Life in vivo. *Angew. Chem., Int. Ed.* **2019**, *58* (5), 1392–1396.
- (54) Zorzi, A.; Middendorp, S. J.; Wilbs, J.; Deyle, K.; Heinis, C. Acylated heptapeptide binds albumin with high affinity and application as tag furnishes long-acting peptides. *Nat. Commun.* **2017**, *8*, 16092.
- (55) Matochko, W. L.; Derda, R. Next-generation sequencing of phage-displayed peptide libraries. *Methods Mol. Biol.* **2015**, *1248*, 249–66.
- (56) Robinson, M. D.; Smyth, G. K. Moderated statistical tests for assessing differences in tag abundance. *Bioinformatics* **2007**, *23* (21), 2881–7.
- (57) Robinson, M. D.; Smyth, G. K. Small-sample estimation of negative binomial dispersion, with applications to SAGE data. *Biostatistics* **2007**, *9* (2), 321–32.
- (58) Robinson, M. D.; Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **2010**, *11* (3), R25.
- (59) Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **1995**, *57* (1), 289–300.
- (60) Saito, R.; Sato, T.; Ikai, A.; Tanaka, N. Structure of bovine carbonic anhydrase II at 1.95 Å resolution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60* (4), 792–5.
- (61) Warden, A. C.; Williams, M.; Peat, T. S.; Seabrook, S. A.; Newman, J.; Dojchinov, G.; Haritos, V. S. Rational engineering of a mesohalophilic carbonic anhydrase to an extreme halotolerant biocatalyst. *Nat. Commun.* **2015**, *6*, 10278.
- (62) Trott, O.; Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31* (2), 455–61.
- (63) Gaspari, R.; Rechlin, C.; Heine, A.; Bottegoni, G.; Rocchia, W.; Schwarz, D.; Bomke, J.; Gerber, H. D.; Klebe, G.; Cavalli, A. Kinetic and Structural Insights into the Mechanism of Binding of Sulfonamides to Human Carbonic Anhydrase by Computational and Experimental Studies. *J. Med. Chem.* **2016**, *59* (9), 4245–56.
- (64) Rentzsch, R.; Renard, B. Y. Docking small peptides remains a great challenge: an assessment using AutoDock Vina. *Briefings Bioinf.* **2015**, *16* (6), 1045–56.
- (65) Flood, D. T.; Hintzen, J. C. J.; Bird, M. J.; Cistrone, P. A.; Chen, J. S.; Dawson, P. E. Leveraging the Knorr Pyrazole Synthesis for the Facile Generation of Thioester Surrogates for use in Native Chemical Ligation. *Angew. Chem., Int. Ed.* **2018**, *57* (36), 11634–11639.
- (66) Gunnoo, S. B.; Finney, H. M.; Baker, T. S.; Lawson, A. D.; Anthony, D. C.; Davis, B. G. Creation of a gated antibody as a conditionally functional synthetic protein. *Nat. Commun.* **2014**, *5*, 4388.
- (67) Wright, T. H.; Bower, B. J.; Chalker, J. M.; Bernardes, G. J.; Wiewiora, R.; Ng, W. L.; Raj, R.; Faulkner, S.; Vallee, M. R.; Phanumartwath, A.; Coleman, O. D.; Thezenas, M. L.; Khan, M.; Galan, S. R.; Lercher, L.; Schombs, M. W.; Gerstberger, S.; Palm-Espling, M. E.; Baldwin, A. J.; Kessler, B. M.; Claridge, T. D.; Mohammed, S.; Davis, B. G., Posttranslational mutagenesis: A chemical strategy for exploring protein side-chain diversity. *Science* **2016**, *354* (6312).
- (68) Agarwal, P.; van der Weijden, J.; Sletten, E. M.; Rabuka, D.; Bertozzi, C. R. A Pictet-Spengler ligation for protein chemical modification. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (1), 46–51.
- (69) Spears, R. J.; Fascione, M. A. Site-selective incorporation and ligation of protein aldehydes. *Org. Biomol. Chem.* **2016**, *14* (32), 7622–38.
- (70) Kale, S. S.; Villequey, C.; Kong, X. D.; Zorzi, A.; Deyle, K.; Heinis, C. Cyclization of peptides with two chemical bridges affords large scaffold diversities. *Nat. Chem.* **2018**, *10* (7), 715–723.
- (71) Kong, X. D.; Moriya, J.; Carle, V.; Pojer, F.; Abriata, L. A.; Deyle, K.; Heinis, C. De novo development of proteolytically resistant therapeutic peptides for oral administration. *Nat. Biomed. Eng.* **2020**, *4* (5), 560–571.
- (72) Li, S.; Roberts, R. W. A novel strategy for in vitro selection of peptide-drug conjugates. *Chem. Biol.* **2003**, *10* (3), 233–9.
- (73) Dwyer, M. A.; Lu, W.; Dwyer, J. J.; Kossiakoff, A. A. Biosynthetic phage display: a novel protein engineering tool combining chemical and genetic diversity. *Chem. Biol.* **2000**, *7* (4), 263–74.
- (74) Hacker, D. E.; Abrigo, N. A.; Hoinka, J.; Richardson, S. L.; Przytycka, T. M.; Hartman, M. C. T. Direct, Competitive Comparison of Linear, Monocyclic, and Bicyclic Libraries Using mRNA Display. *ACS Comb. Sci.* **2020**, *22* (6), 306–310.

(75) Oller-Salvia, B.; Chin, J. W. Efficient Phage Display with Multiple Distinct Non-Canonical Amino Acids Using Orthogonal Ribosome-Mediated Genetic Code Expansion. *Angew. Chem., Int. Ed.* **2019**, *58* (32), 10844–10848.

(76) 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.* **2020**, *6* (3), 368–381.

(77) Wang, X. S.; Chen, P. C.; Hampton, J. T.; Tharp, J. M.; Reed, C. A.; Das, S. K.; Wang, D. S.; Hayatshahi, H. S.; Shen, Y.; Liu, J.; Liu, W. R. A Genetically Encoded, Phage-Displayed Cyclic-Peptide Library. *Angew. Chem., Int. Ed.* **2019**, *58* (44), 15904–15909.

(78) Van Deventer, J. A.; Yuet, K. P.; Yoo, T. H.; Tirrell, D. A. Cell surface display yields evolvable, clickable antibody fragments. *ChemBioChem* **2014**, *15* (12), 1777–81.

(79) Stieglitz, J. T.; Kehoe, H. P.; Lei, M.; Van Deventer, J. A. A Robust and Quantitative Reporter System To Evaluate Noncanonical Amino Acid Incorporation in Yeast. *ACS Synth. Biol.* **2018**, *7* (9), 2256–2269.

(80) Chou, Y.; Kitova, E. N.; Joe, M.; Brunton, R.; Lowary, T. L.; Klassen, J. S.; Derda, R. Genetically-encoded fragment-based discovery (GE-FBD) of glycopeptide ligands with differential selectivity for antibodies related to mycobacterial infections. *Org. Biomol. Chem.* **2018**, *16* (2), 223–227.

(81) Sojitra, M.; Sarkar, S.; Maghera, J.; Rodrigues, E.; Carpenter, E.; Seth, S.; Vinals, D. F.; Bennett, N.; Reddy, R.; Khalil, A.; Xue, X.; Bell, M.; Zheng, R. B.; Ling, C.-C.; Lowary, T. L.; Paulson, J. C.; Macauley, M. S.; Derda, R. Genetically Encoded, Multivalent Liquid Glycan Array (LiGA) *bioRxiv*2020