ACS chemical biology

Articles

pubs.acs.org/acschemicalbiology

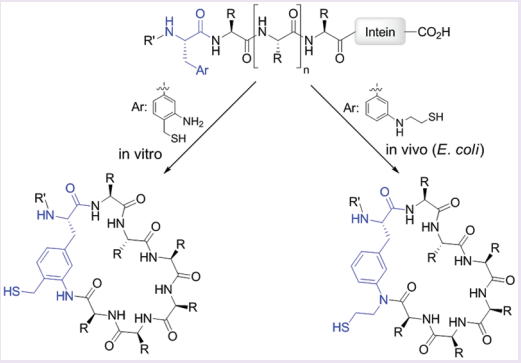
Ribosomal Synthesis of Macrocyclic Peptides *in Vitro* and *in Vivo* Mediated by Genetically Encoded Aminothiol Unnatural Amino Acids

John R. Frost, Nicholas T. Jacob, Louis J. Papa, Andrew E. Owens, and Rudi Fasan*

Department of Chemistry, University of Rochester, Hutchinson Hall, Rochester, New York 14627, United States

Supporting Information

ABSTRACT: A versatile method for orchestrating the formation of side chain-to-tail cyclic peptides from ribosomally derived polypeptide precursors is reported. Upon ribosomal incorporation into intein-containing precursor proteins, designer unnatural amino acids bearing side chain 1,3- or 1,2-aminothiol functionalities are able to promote the cyclization of a downstream target peptide sequence via a C-terminal ligation/ring contraction mechanism. Using this approach, peptide macrocycles of variable size and composition could be generated in a pH-triggered manner *in vitro* or directly in living bacterial cells. This methodology furnishes a new platform for the creation and screening of genetically encoded libraries of conformationally constrained peptides. This strategy was applied to identify and isolate a low-micromolar streptavidin binder (K_D = 1.1 μM) from a library of cyclic peptides produced in *Escherichia coli*, thereby illustrating its potential toward aiding the discovery of functional peptide macrocycles.



Macrocyclic peptides occupy a most relevant region of the chemical space and have attracted increasing attention as molecular scaffolds for addressing challenging pharmacological targets.[1,2] Important benefits arising from the cyclization of peptide sequences include preorganization into a bioactive conformation and reduced entropic costs upon complex formation, which can result in enhanced binding selectivity and affinity, respectively, toward a target biomolecule.[3−10] In addition, macrocyclization has proven to be beneficial toward improving the proteolytic stability[11−13] and cell penetration ability[14−16] of peptide-based molecules. These advantageous features are well-reflected by the widespread occurrence of ring topologies among natural biologically active peptides, such as those produced by nonribosomal peptide synthetases[17,18] (e.g., cyclosporin, polymyxin) or those belonging to the class of ribosomally synthesized and post-translationally modified peptides (RiPPs)[19] (e.g., lanthipeptides,[20] cyclotides,[21] cyanobactins[22]).

Because of the attractive properties of macrocyclic peptides, the development of methods for the preparation of this class of compounds has attracted significant interest.[23−27] Among these, strategies that rely on the cyclization of genetically encoded peptide sequences offer several advantageous features, which include the opportunity to create and evaluate large combinatorial libraries of these molecules by combining genetic randomization with high-throughput display/screening platforms.[24−26] For example, biopanning of phage displayed peptide libraries cyclized via disulfide bridges[28−32] or exogenous cross-linking agents[33−35] has enabled the discovery of potent inhibitors of enzymes and proteins. Coupling *in vitro* translation methods with mRNA display, other groups have successfully isolated potent macrocyclic peptide inhibitors for a variety of other therapeutically relevant protein targets.[36−41]

Despite this progress, there are currently only a few methods that enable the production of macrocyclic peptides of arbitrary sequence within a cell, thereby providing the capability of coupling library production with a selection system or phenotypic screen.[42−48] In this regard, a viable method is the so-called split-intein-mediated circular ligation of peptides and proteins (SICLOPPS) introduced by Benkovic and co-workers,[49] which enables the intracellular formation of head-to-tail cyclic peptides by exploiting the trans splicing reactivity of the natural split intein DnaE. More recently, a strategy useful for the generation *in vitro* and *in vivo* of macrocylic peptides constrained by an inter side chain thioether bond was reported by our group.[50] While these methods provide a route to obtain head-to-tail or side chain-to-side-chain cyclic peptides, complementary strategies to access alternative peptide macrocycle architectures would be highly desirable. For example, neither of these approaches allows the formation of side chain-to-tail cyclic peptides, a molecular topology that is found in many bioactive peptide natural products (e.g., bacitracin A, polymixin D).

As part of our efforts toward the development of methods for the synthesis of macrocyclic organo−peptide hybrids,[51−54] we recently described the ability of bifunctional oxyamine/1,3-aminothiol aryl reagents to induce the efficient cyclization of polypeptide precursors containing a target peptide sequence

Received: February 21, 2015
Accepted: May 1, 2015
Published: May 1, 2015

DOI: 10.1021/acschembio.5b00119
ACS Chem. Biol. 2015, 10, 1805−1816

flanked by a side chain keto group and a C-terminal intein.[52] Although two possible routes for macrocyclization are available in this system (e.g., side chain → C-end versus C-end → side chain ligation), our studies revealed that this process is largely driven by the intermolecular attack of the 1,3-aminothiol aryl moiety onto the intein-catalyzed thioester bond within the polypeptide precursor.[53] These mechanistic insights suggested to us the possibility of directing the biosynthesis of side chain-to-tail macrocyclic peptides through the ribosomal expression of a precursor polypeptide in which an appropriately designed aminothiol-based unnatural amino acid is placed upstream of an intein (Figure 1). Given the expected rate enhancement in

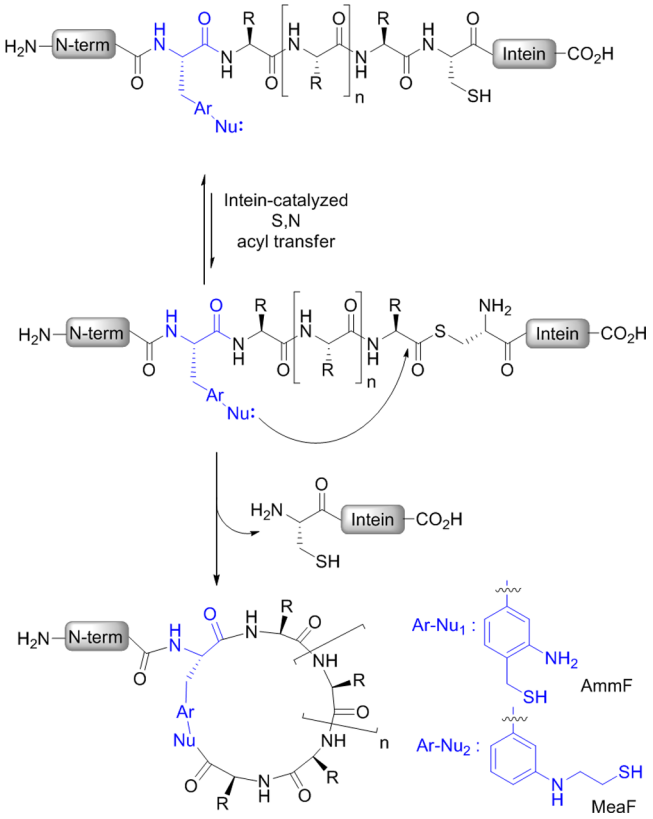


Figure 1. Overview of the strategy for generating side chain-to-tail macrocyclic peptides via cyclization of ribosomally derived, intein-fused precursor proteins by means of aminothiol unnatural amino acids. Nu refers to the nucleophilic 1,3- or 1,2-aminothiol moiety.

moving from an intermolecular to an intramolecular setting, we envisioned this construct could undergo a spontaneous post-translational self-processing reaction through (a) capturing of the transient intein-catalyzed thioester linkage by the side chain thiol functionality followed by (b) irreversible acyl shift to the neighboring amino group to give the desired macrocyclic peptide. Here, we describe the successful implementation of this design strategy and its functionality toward enabling the ribosomal synthesis of macrocyclic peptides both *in vitro* and in living cells. Furthermore, it is shown how this approach could be successfully applied to evolve and isolate a cyclopeptide with improved binding affinity toward a model target protein (streptavidin).

RESULTS AND DISCUSSION

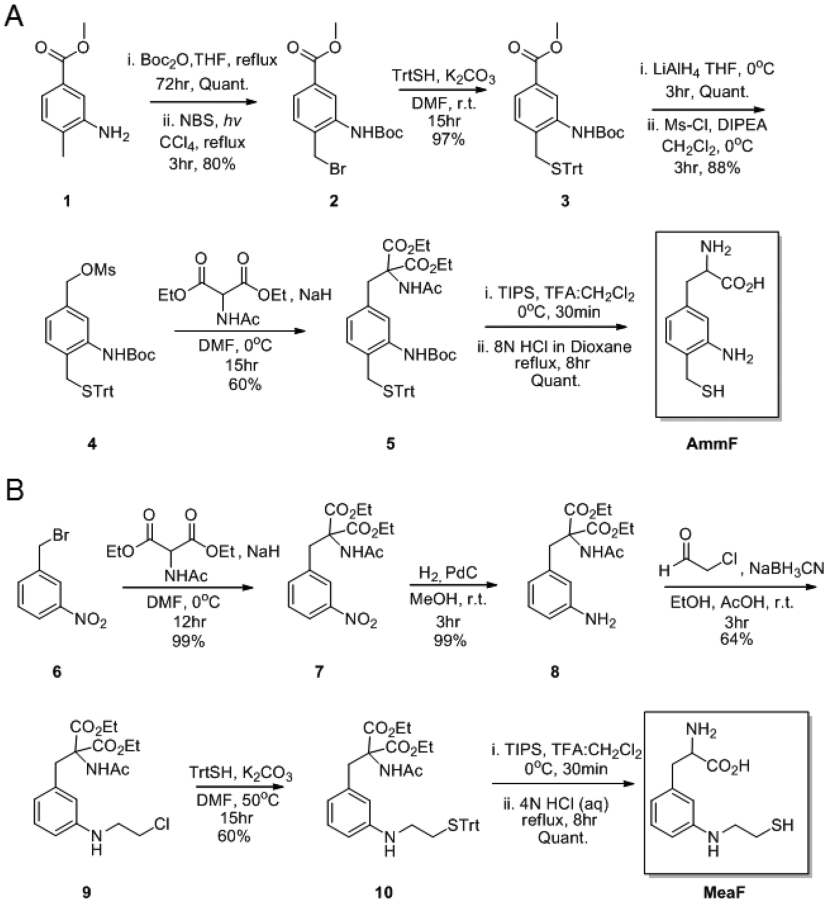
Synthesis of 1,3-Aminothiol Amino Acid, AmmF. In previous work, we recognized the peculiar ability of an *o*-amino-mercaptomethyl-aryl (AMA) group to undergo efficient C-terminal ligation at GyrA intein thioesters under catalyst-free conditions.[53] On the basis of this knowledge, an amino acid incorporating this reactive moiety, 3-amino-4-mercaptomethyl-phenylalanine (AmmF, Scheme 1A), was designed in order to test its ability to mediate peptide cyclization according to the strategy outlined in Figure 1. As outlined in Scheme 1A, gram-scale synthesis of AmmF was carried out starting from commercially available methyl 3-amino-4-methyl-benzoate, followed by N-Boc protection, benzylic bromination, and benzylic substitution with tritylmercaptan to afford intermediate **3**. Lithium aluminum hydride reduction of the methyl ester group in **3** followed by mesylation and substitution of the mesyl group with diethylacetamidomalonate yielded intermediate **5**. Following deprotection, the desired product, AmmF, was obtained in 41% overall yield over eight steps.

Aminoacyl-tRNA Synthetase for Ribosomal Incorporation of AmmF. With AmmF in hand, we sought to identify an orthogonal aminoacyl-tRNA synthetase (AARS)/tRNA pair for the ribosomal incorporation of this unnatural amino acid (UAA) via amberstop codon (TAG) suppression.[55] Interestingly, a number of engineered AARS enzymes have been shown to be polyspecific, that is, capable of charging a cognate tRNA molecule with multiple UAAs while maintaining discriminating selectivity against the natural amino acids.[56,57] We reasoned that this feature could be exploited to find a suitable AARS for recognition of AmmF. To this end, we generated a panel of six previously reported AARSs derived from *Methanocaldococcus jannaschii* tyrosyl-tRNA synthetase (*Mj*Tyr-RS) in combination with the cognate amber stop codon suppressor tRNA ($MjtRNA^{Tyr}_{CUA}$). This panel included *Mj*Tyr-RS variants selected for recognition of 3-aminotyrosine (3AmY),[58] *p*-aminophenylalanine (pAmF),[59] *p*-azidophenylalanine (AzF),[60] *p*-acetylphenylalanine (pAcF),[61] *O*-propargyltyrosine (OpgY),[62] and *O*-(2-bromoethyl)-tyrosine (O2beY)[50](Figure 2A).

Among these, 3AmY-RS was considered to be a particularly promising candidate owing to the structural similarity between 3AmY and AmmF. The relative ability of the AARSs to incorporate AmmF in response to a TAG codon was established via a fluorescence assay with a yellow fluorescent protein (YFP) variant containing an amber stop codon at the N-terminus. In this assay, the suppression efficiency and fidelity of each synthetase is measured based on the relative expression levels of the reporter YFP protein in the presence and absence of AmmF, respectively (Figure 2B). Gratifyingly, the 3AmY-RS synthetase was found to be capable of incorporating AmmF with good efficiency as compared to that of most of the other AARSs. Interestingly, the AzF-RS synthetase exhibited even better performance in that respect, as judged by the higher fluorescence signal (+15%) measured in the assay. This result is consistent with previous observations by Schultz and co-workers regarding the ability of AzF-RS to recognize various *meta*, *para*-disubstituted phenylalanine derivatives, in addition to AzF.[57] Further experiments were then carried out to quantitatively assess the suppression efficiency of AzF-RS with AmmF. For this purpose, YFP(AmmF) and wild-type YFP were expressed under identical conditions and purified by Ni-affinity chromatography. Direct comparison of the expression yield (16 vs 45 mg/L culture, respectively) indicated a suppression efficiency of 36% (Figure 2C), a value comparable to that reported for AARSs specifically evolved for the ribosomal incorporation of other UAAs.[57]

DOI: 10.1021/acschembio.5b00119
ACS Chem. Biol. 2015, 10, 1805−1816

Scheme 1. Synthetic Routes for the Preparation of (A) 3-Amino-4-mercaptomethyl-phenylalanine (AmmF) and (B) 3-(2-Mercapto-ethyl)amino-phenylalanine (MeaF)



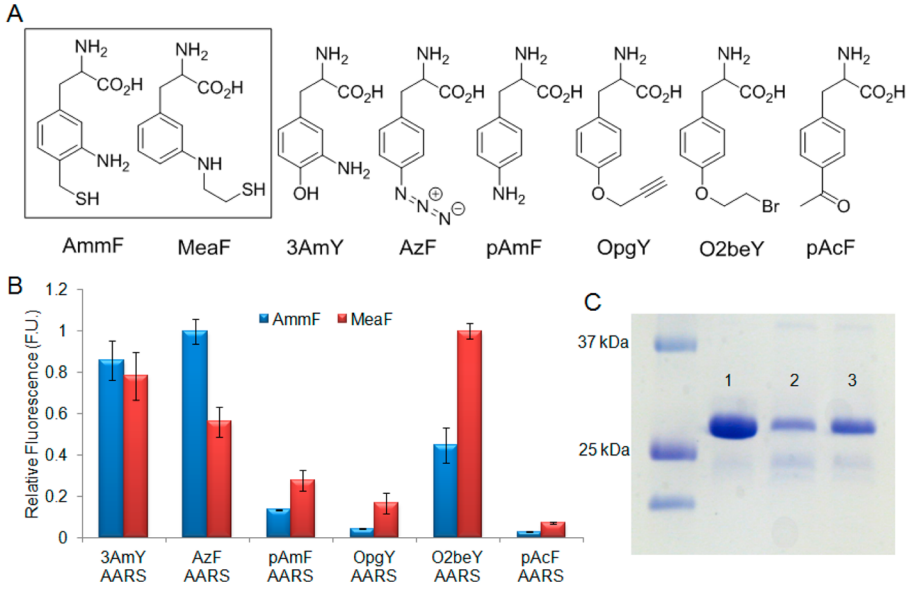


Figure 2. Orthogonal aminoacyl-tRNA synthetases for ribosomal incorporation of the aminothiol amino acids. (A) Structures of the target amino acids, AmmF and MeaF, and other UAAs mentioned in the text. (B) Relative fluorescence measured in the YFP reporter assay for the panel of AARSs in the presence of AmmF or MeaF. Data are normalized to the highest value within each series. (C) SDS-PAGE analysis of purified wild-type YFP (lane 1) and YFP variants incorporating AmmF (lane 2) and MeaF (lane 3).

Characterization of AmmF-Containing Precursor Proteins. Having identified a viable AARS for AmmF incorporation, we tested the ability of this amino acid to promote peptide macrocyclization according to the scheme outlined in Figure 1. To this end, we investigated a panel of six constructs encompassing target peptide sequences of variable

DOI: 10.1021/acschembio.5b00119
ACS Chem. Biol. 2015, 10, 1805−1816

length (4−12 amino acids) framed in between AmmF and the N198A variant of GyrA mini-intein from *Myobacterium xenopi*[63] (entries 1−6, Table 1). These constructs would allow us to probe the formation of peptide macrocycles of varying ring size. In addition, the N-terminal chitin binding domain (CBD) would serve as a large (~8 kDa) affinity tag to facilitate both the isolation of the products formed in cells and measurement of the extent of intein cleavage via SDS-PAGE densitometry.

Table 1. Precursor Protein Constructs Investigated in This Study[a]

entry	construct name	peptide sequence
1	CBD-4T	CBD-(UAA)TGST-GyrA-His_6
2	CBD-5T	CBD-(UAA)TGSGT-GyrA-His_6
3	CBD-6T	CBD-(UAA)TGSYGT-GyrA-His_6
4	CBD-8T	CBD-(UAA)TGSAEYGT-GyrA-His_6
5	CBD-10T	CBD-(UAA)TGSKLAEYGT-GyrA-His_6
6	CBD-12T	CBD-(UAA)TGSWGKLAEYGT-GyrA-His_6
7	MG-8T	MG(UAA)TGSAEYGT-GyrA-His_6
8	CBD-7T	CBD-(UAA)TGSAYGT-GyrA-His_6
9	Strep1	MG(UAA)FTNVHPQFANA-GyrA-His_6
10	Flag-Strep1	FLAG-GSSG(UAA)FTNVHPQFANA-GyrA-His_6
11	Flag-Strep2	FLAG-GSSG(UAA)FTNVHPQSANA-GyrA-His_6
12	Flag-Strep3	FLAG-GSSG(UAA)FTNYHPQDANA-GyrA-His_6

[a]CBD, chitin binding domain; UAA, unnatural amino acid (i.e., AmmF or MeaF); GyrA, *Mxa* GyrA intein (N198A variant).

Accordingly, the precursor proteins were produced in *Escherichia coli* cells cotransformed with a pET-based plasmid encoding for the precursor protein and a pEVOL[64]-based plasmid for the coexpression of the AzF-RS/$tRNA_{CUA}$ pair for the site-selective incorporation of AmmF.

Contrary to our expectations, only low levels of intein cleavage (20−25%) were observed for all of the constructs under the applied expression conditions (12 h, 27 °C), as determined by SDS-PAGE analysis. These values were comparable to those obtained with control constructs in which AmmF was substituted for a non-nucleophilic UAA (= pAcF). Moreover, no traces of the desired macrocyclic compounds were detected after processing the cell lysates with chitin beads.

These results were surprising, given that compounds sharing the *o*-amino-mercaptomethyl-aryl moiety of AmmF can induce nearly quantitative clevage of GyrA-fusion proteins (80−90%) in less than 10−12 h.[53] To investigate the basis of this unexpected lack of reactivity, the AmmF-containing constructs were isolated in full-length form using the C-terminal His tag and then incubated with thiophenol to cleave the intein and release the N-terminal CBD-containing fragment. MALDI-TOF analysis of these reactions showed the accumulation of a single product corresponding to the linear peptide with an additional mass of +26 Da for each of the constructs (Figure S1). In contrast, under identical conditions, control proteins containing pAcF instead of AmmF produced a species corresponding to the expected linear peptide with no

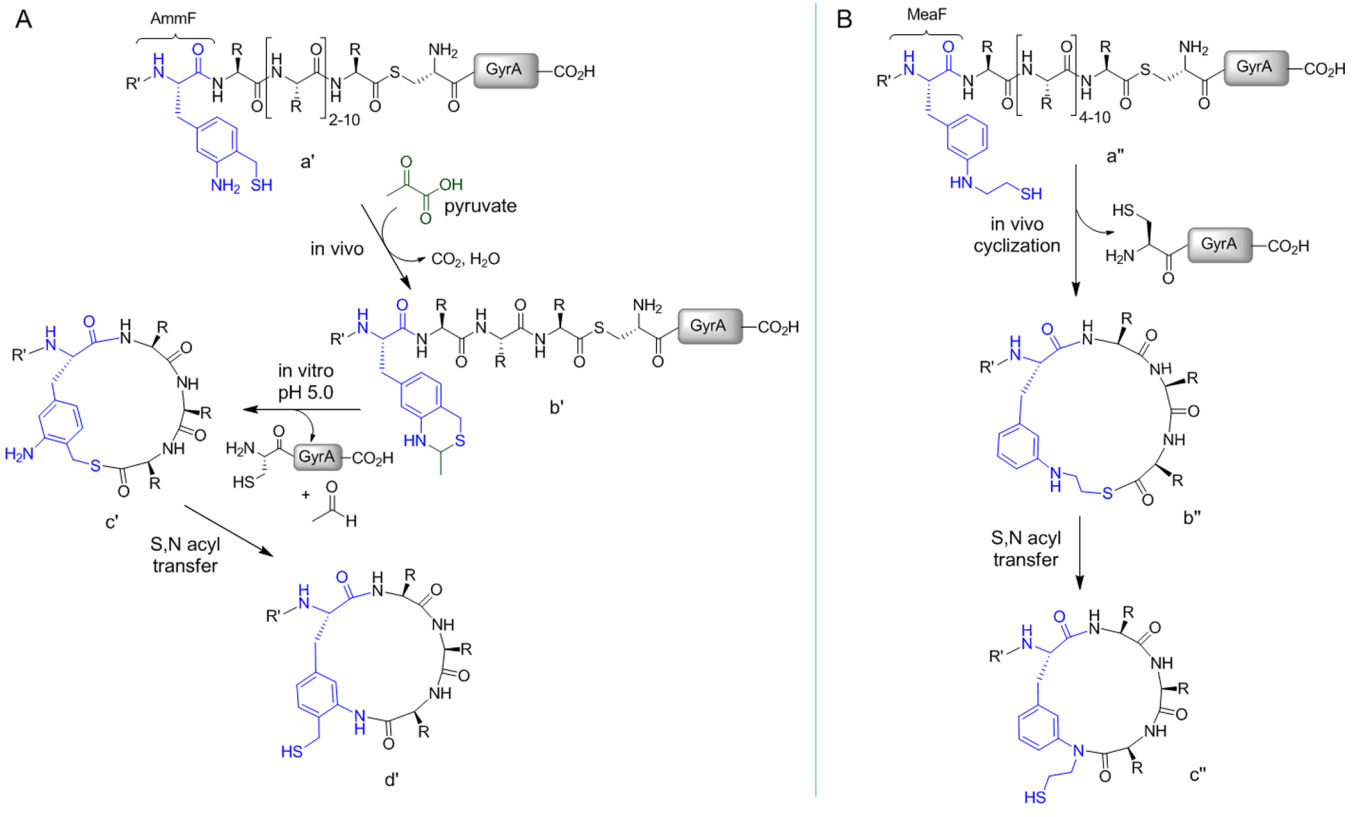


Figure 3. Precursor polypeptides, reaction intermediates, and final macrocyclic peptide products for (A) AmmF-mediated and (B) MeaF-mediated peptide cyclization. (A) Expression of the AmmF-containing precursor protein (a′) in *E. coli* results in the formation of the benzothiazine adduct (b′), which, upon pH-induced deprotection *in vitro*, leads to the target macrocyclic peptide (d′). (B) Expression of the MeaF-containing precursor protein (a″) in *E. coli* results in spontaneous, post-translational peptide cyclization to give the thiolactone intermediate (b″) and then the final macrolactam product (c″) via an S → N acyl shift rearrangement.

DOI: 10.1021/acschembio.5b00119
ACS Chem. Biol. 2015, 10, 1805−1816

modifications. These results pointed at a post-translational modification occurring specifically at the level of the unnatural amino acid in the AmmF-containing constructs. The observed mass difference was found to be consistent with the formation of a 2,4-dihydro-benzo[1,3]thiazine adduct (Figure 3A, species b′) resulting from a condensation/decarboxylation reaction of the AMA moiety in AmmF with pyruvate, an abundant metabolite in *E. coli*.[65] Importantly, a similar reaction, leading to a thiazolidine adduct, was reported for polypeptides carrying an exposed N-terminal cysteine upon expression in this organism.[66] To confirm the assignment, a model reaction was carried out by incubating a small-molecule surrogate of AmmF, methyl 3-amino-4-(mercaptomethyl)benzoate, with pyruvate under physiological conditions (50 mM KPi, 150 mM NaCl, pH 7.5). LC-MS analysis revealed quantitative conversion of the compound to the pyruvate condensation adduct within a short time (Figure S2). Altogether, these experiments indicated that rapid and quantitative intracellular modification of AmmF by pyruvate impedes the desired *in vivo* peptide cyclization process. Attempts to minimize this modification by altering the expression conditions[66] failed to result in a detectable increase of the unmodified AmmF-containing protein.

***In Vitro* Macrocyclization of AmmF-Containing Constructs.** These findings prompted us to explore the possibility to execute the desired AmmF-mediated peptide macrocyclization upon chemical unmasking of its modified 1,3-aminothiol functionality. While not directly applicable for intracellular synthesis of cyclic peptides, this approach would provide a means to trigger the formation of peptide macrocycles in a time-controlled manner *in vitro*. Although deprotection of cysteine-derived thiazolidines has been achieved under rather harsh, protein-denaturing conditions (0.2 M methoxyamine, 6 M guanidinium chloride),[67,68] we discovered that the AmmF dihydrobenzothiazine adduct can be readily converted to the free 1,3-aminothiol counterpart by simple exposure of the protein to slightly acidic pH, under which conditions the GyrA intein remains functional. Gratifyingly, incubation of the AmmF-containing constructs at pH 5.0 and in the presence of sodium 2-sulfanylethanesulfonate (MESNA) resulted in the formation of small molecular weight products with masses corresponding to the expected CBD-fused peptide macrocycle in each case, as determined by MALDI-TOF MS (Figure S3). Notably, identical results could be achieved in the absence of any thiol catalyst, resulting in the clean formation of the desired macrocyclic lactams as the only product (Figure 4A).

Exposure of these macrocyclic products to iodoacetamide resulted in complete conversion to the S-alkylated adduct (+57, Figure S4), demonstrating the occurrence of the desired S → N acyl transfer rearrangement (c′ → d′ step, Figure 3A) to give rise to the hydrolytically stable side chain-to-C-terminus amide linkage. Although the extent of intein cleavage in these reactions was not elevated (15−28%, Figure 4B), AmmF-mediated cyclization was found to be largely independent of the target peptide length, enabling the formation of peptide macrocycles spanning from 4 to 12 amino acid residues with equal efficiency. Suspecting that steric clashes between the CBD and intein protein may disfavor attack of AmmF thiol group onto the downstream thioester (Figure 3A), an 8-mer construct in which the CBD was replaced with a smaller N-terminal tail (Met-Gly) was also tested (entry 7, Table 1). Incubation of the corresponding protein precursor at pH 5.0 resulted in the formation of the desired peptide macrocycle, whose cyclic structure was further confirmed by MS/MS analysis (Figure 4C). Importantly, the construct was determined to undergo 64% intein cleavage (vs 28% for the CBD-fused counterpart), with the macrocycle being the only detectable product, thereby demonstrating the possibility to

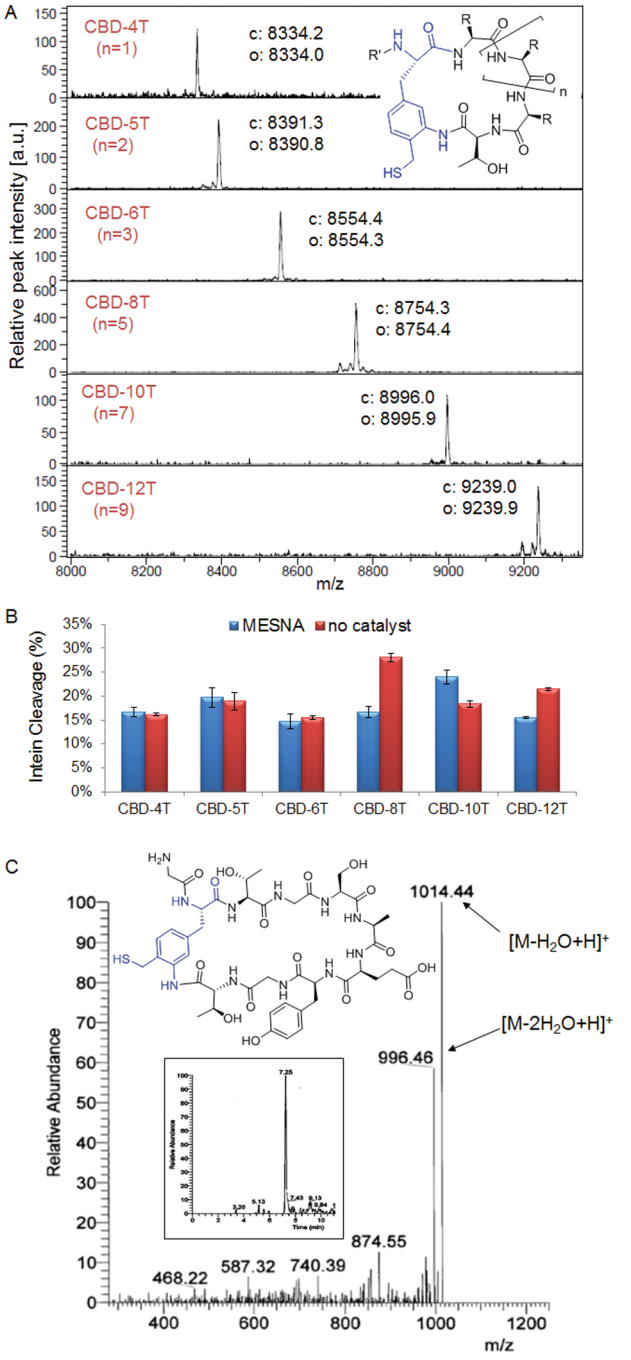


Figure 4. pH-induced cyclization of AmmF-containing constructs *in vitro*. (A) MALDI-TOF MS spectra of the macrocyclic peptides generated under catalyst-free conditions. The calculated (c) and observed (o) *m*/*z* values corresponding to the $[M + H]^+$ adduct of the macrocylic product are indicated. R′ = chitin binding domain. (B) Extent of intein cleavage for constructs CBD-4T(AmmF) through CBD-12T(AmmF) at pH 5.0 after 24 h as determined by SDS-PAGE. (C) Chemical structure, extracted-ion chromatogram (inset), and MS/MS spectrum of the macrocyclic peptide obtained from pH-induced cyclization of the MG-8T(AmmF) construct. Calculated *m*/*z* values for the $[M - H_2O + H]^+$ and $[M - 2H_2O + H]^+$ species are 1014.07 and 996.05, respectively.

DOI: 10.1021/acschembio.5b00119
ACS Chem. Biol. 2015, 10, 1805−1816

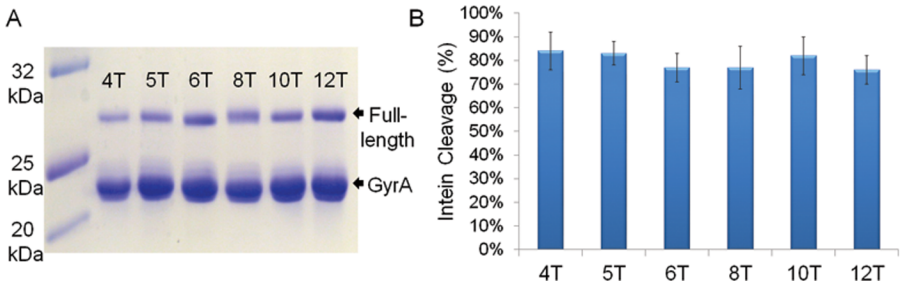


Figure 5. MeaF-mediated intein cleavage. (A) SDS-PAGE analysis of the precursor proteins CBD-4T(MeaF) through CBD-12T(MeaF) after expression for 12 h in *E. coli*. Bands corresponding to full-length protein and cleaved GyrA intein are labeled. (B) The extent of intein cleavage as measured by gel band densitometric analysis.

dramatically increase the yields of these macrocyclization reactions via the use of shorter (or more flexible) N-terminal tails.

Design and Synthesis of 1,2-Aminothiol Unnatural Amino Acid for *in Vivo* Macrocycle Formation. Having demonstrated the possibility to generate macrocyclic peptides *in vitro* by means of AmmF, we next sought an alternative aminothiol UAA to enable this process to take place *in vivo*. To this end, a second phenylalanine derivative was designed, namely, 3-(2-mercapto-ethyl)amino-Phe (MeaF, Figure 1), which was expected to exhibit superior performance compared to that of AmmF based on the following arguments. First, the secondary amino group of MeaF was expected to disfavor condensation with pyruvate in cells, thereby preserving the side chain aminothiol functionality required for cyclization. Second, the more acidic and less sterically hindered sulphydryl group in MeaF compared to that in AmmF (i.e., pK_a of ~8.5[69] vs ~9.5[70]) would favor nucleophilic attack onto the downstream intein thioester, in particular under the near-neutral conditions of the intracellular milieu. Third, the thiolactone formed during the initial transthiosterification step was envisioned to be less strained in the presence of MeaF vs AmmF (b″ vs c′, Figure 3). Finally, MeaF 1,2-aminothiol moiety was expected to remain capable of undergoing an S → N acyl shift rearrangement to yield the desired N-alkyl macrolactam, as suggested by peptide ligation studies with other aminothiol compounds.[71]

As summarized in Scheme 1B, MeaF was prepared starting from 3-nitro-benzyl bromide via alkylation with diethylacetamidomalonate and nitro group reduction by hydrogenation with Pd/C catalyst. Reductive amination of intermediate **8** with α-chloroacetaldehyde followed by substitution with triphenylmethylmercaptan resulted in the protected intermediate **10**. Deprotection with TFA followed by decarboxylation with concentrated HCl resulted in the desired amino acid MeaF in 38% isolated yield over six steps. This route could be scaled up to the gram scale (1.5 g final product) with no significant reduction in the overall yield.

AARS for Ribosomal Incorporation of MeaF. The same approach described above for AmmF was applied to identify a suitable AARS for amber stop codon suppression with MeaF. Initial screening of the panel of engineered *Mj* AARSs showed that both pAmF-RS and AzF-RS can recognize and incorporate MeaF into the reporter YFP protein (Figure 2B). At the same time, O2beY-RS emerged as the most promising enzyme for this purpose based on the approximately 20% higher fluorescent signal in the assay (Figure 2B). This synthetase, which carries a Ala32Gly mutation in the OpgY-RS background, was engineered by our group to accommodate tyrosine derivatives containing larger alkyl substituents at the *para* position.[50] Considering the much lower incorporation of MeaF obtained with OpgY-RS, this substitution is clearly beneficial also toward recognition of *meta*-substituted aromatic UAAs such as MeaF. Comparison of the expression yield of this protein (23 mg/L culture) with that of wild-type YFP (45 mg/L culture) indicated a suppression efficiency as high as 52%, which is an about 16% higher efficiency than that achieved with AmmF and AzF-RS (Figure 2C).

***In Vivo* Cyclization of MeaF-Containing Precursor Proteins.** To compare side-by-side the performance of MeaF vs AmmF toward promoting peptide macrocyclization *in vivo*, the same six constructs with a 4-mer to 12-mer target sequence (entries 1−6, Table 1) were expressed in *E. coli* in the presence of MeaF and the O2beY-RS-based suppressor system. Notably, SDS-PAGE analysis of these proteins revealed a significantly larger degree of intein cleavage (75−85%, Figure 5) as compared to that of the related AmmF-containing counterparts (20−25%) under identical expression conditions. Since the two sets of constructs differ only by the side chain aminothiol functionality in the installed UAA, these results strongly suggested that the ribosomally incorporated MeaF had remained available for nucleophilic attack on the intein thioester.

To examine the outcome of these *in vivo* reactions, the cleaved N-terminal fragments were isolated from the cell lysates using chitin-coated resin beads. To our delight, MALDI-TOF MS analysis of these samples revealed the occurrence of a macrocyclic product (cyc, Figure 6), corresponding to either the thiolactone or the isobaric lactam (species b″ and c″, respectively, in Figure 3B), as the only or largely predominant product for the constructs with 6−12 amino acid-long target sequences. In contrast, the 4-mer and 5-mer constructs showed only MS signals consistent with the acyclic, hydrolyzed product (Figure S5). Although this species could arise from hydrolysis of the full-length intein-containing precursor protein directly, the higher degree of intein cleavage observed with the MeaF-containing vs AmmF-containing constructs (Figure 5) suggests that it more likely derives from hydrolysis of the thiolactone intermediate formed after MeaF-mediated transthioesterification. Importantly, no adducts were observed for any of the constructs, indicating that the side chain functionality of MeaF did not react with pyruvate or other intracellular metabolites.

To assess the occurrence and extent of S → N acyl shift rearrangement in the 6-mer to 12-mer macrocyclic products, the corresponding cell lysate samples were first treated with iodoacetamide under alkaline conditions (pH 8) prior to chitin-affinity purification. Through this procedure, the macrolactam is converted to its corresponding S-carboxyamidomethyl adduct (m, Figure S6), whereas the thiolactone is hydrolyzed and

DOI: 10.1021/acschembio.5b00119
ACS Chem. Biol. 2015, 10, 1805−1816

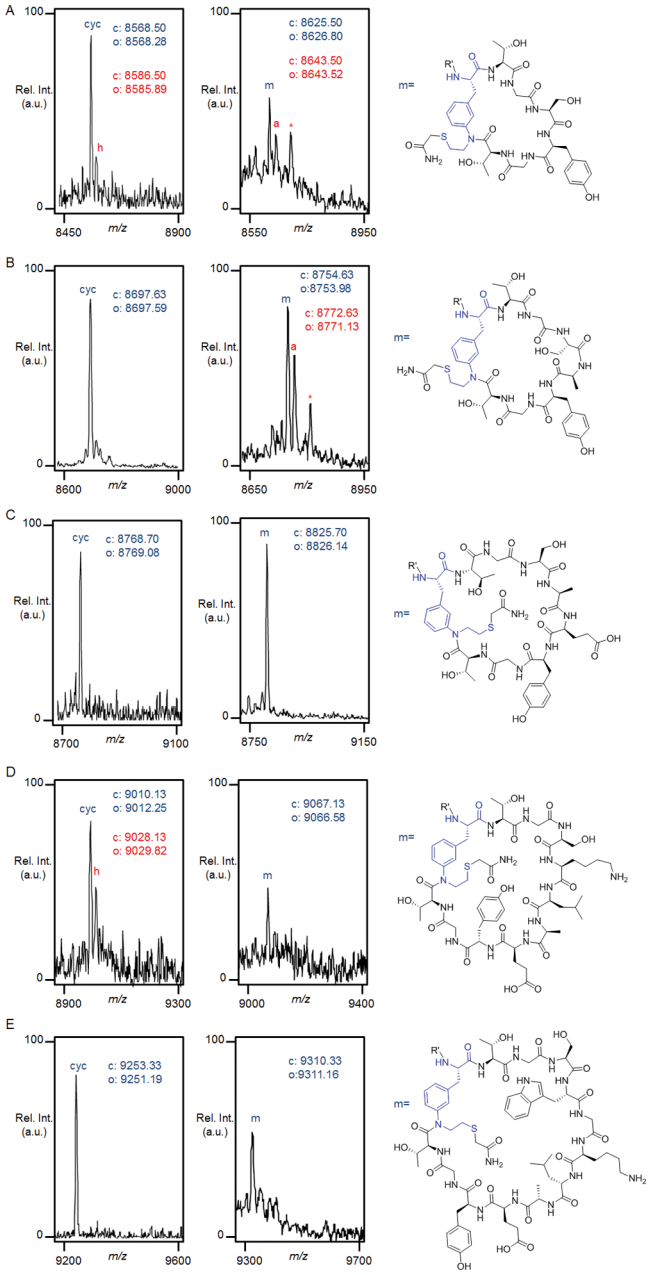


Figure 6. MALDI-TOF spectra of the macrocyclization products obtained from the MeaF-containing constructs as isolated from *E. coli* cells after 12 h (left panel) and 24 h (right panel) using pH 5.0 and 8.0 buffers containing iodoacetamide, respectively. (A) CBD-6T(MeaF); (B) CBD-7T(MeaF); (C) CBD-8T(MeaF); (D) CBD-10T(MeaF); and (E) CBD-12T(MeaF). See also Figures S5 and S6. cyc, macrocyclic product in either thiolactone or lactam form; h, hydrolyzed product; m, macrolactam (S-carboxyamidomethyl adduct); a, acyclic side product (S-carboxyamidomethyl adduct); *, acetylated a. The calculated (c) and observed (o) m/z values corresponding to the $[M + H]^+$ adducts are indicated and color-coded.

alkylated, thus appearing as the S-alkylated adduct of the linear acyclic peptide (a). These experiments showed that the 6-mer-, 10-mer-, and 12-mer-based macrocycles were present exclusively or mostly (~90% for 12-mer) in the thiolactone form. As an exception, the 8-mer-based macrocycle was found to have undergone partial rearrangement, being present in about 40% as the desired lactam product, with the remaining 60% as thiolactone as estimated based on the intensity of the corresponding MS signals (Figure S6C) and assuming the two species share similar ionization properties. Collectively, these results demonstrated the ability of MeaF to efficiently mediate the first step (i.e., intramolecular transthioesterification) of the envisioned macrocyclization process within 12 h. The partial rearrangement of the 8-mer-based construct also suggested that the second step (S,N acyl shift) is comparatively much slower. This finding is consistent with the slow kinetics of the S,N acyl transfer observed for 1,2-aminothiol moieties containing secondary amines in the context of peptide ligation studies.[71]

Given the remarkable stability *in vivo* of the thiolactone intermediate for the 6-mer to 12-mer constructs, we reasoned that longer culture times could result in the intracellular formation of the desired macrolactam product. Accordingly, cell cultures were grown for 24 h prior to isolation and analysis of the spliced products using the procedure described above. Gratifyingly, these experiments showed a dramatic increase in the S,N acyl transfer rearrangement for all the constructs. Indeed, the 6-mer-based macrocycle was found to consist for the most part of the desired lactam product (~60%, Figure 6A). Even better results were achieved with the 8-mer-, 10-mer-, and 12mer-based macrocycles, which were isolated exclusively in the lactam form (Figure 6C–E). These results were corroborated by parallel tests in which the peptides were isolated via chitin affinity immediately after cell lysis under alkaline conditions (pH 8.0). Under these conditions, MS analysis revealed a macrocycle/hydrolyzed product ratio (corresponding to the lactam/thiolactone ratio) comparable or identical to the m/a ratio obtained after incubation with iodoacetamide, further confirming the occurrence of the S,N acyl transfer step inside the cells.

An intriguing aspect emerging from these studies concerns the effect of the target peptide length on the outcome of MeaF-mediation peptide cyclization *in vivo*. Our results indicate that 4-mer and 5-mer sequences undergo MeaF-induced transthioesterification, but the resulting thiolactones are apparently susceptible to hydrolysis before the irreversible S,N acyl transfer process can take place. In contrast, both of these steps occur efficiently in the context of longer peptide sequences (6–12 amino acids). Possibly, a higher ring strain in the thiolactone intermediate and/or slower kinetics for the subsequent ring contraction step may be at the basis of the higher susceptibility of the shorter constructs toward hydrolysis. In the future, it will be interesting to determine whether alternative aminothiol UAAs and/or alterations within the peptide sequence (e.g., at the Intein-1 site) can provide access also to these smaller rings. Another implication of the results above is that 7-mer peptide sequences should also be amenable to cyclization. To test this hypothesis, the construct corresponding to entry 8 in Table 1 was prepared and expressed. As anticipated, this construct underwent extensive cyclization *in vivo* (~75–80% intein cleavage), resulting in the formation of the desired CBD-fused peptide macrocycle as the major product (~60:40 lactam/thiolactone ratio, Figure 6B). Altogether, these studies demonstrated the functionality of the methodology outlined in Figure 3B to orchestrate the spontaneous, intracellular formation of side chain-to-tail peptide macrocycles in *E. coli*.

Generation of a Streptavidin-Binding Macrocyclic Peptide in Living Cells. To demonstrate the production of a bioactive macrocyclic peptide in living cells, an additional construct (entry 9, Table 1) was designed based on a streptavidin-binding head-to-tail cyclopeptide reported by

DOI: 10.1021/acschembio.5b00119
ACS Chem. Biol. 2015, 10, 1805–1816

Benkovic and co-workers.[72] In this case, the precursor protein consists of a short N-terminal tail (Met-Gly) followed by MeaF and a 11-mer peptide sequence encompassing a histidine−proline−glutamine (HPQ) motif known to interact with streptavidin.[73,74] After expression of this construct in *E. coli*, cell lysates were passed over streptavidin-coated beads to isolate any streptavidin-bound material. Remarkably, analysis of the eluate revealed the presence of the desired macrocyclic peptide (*cyclo*(Strep1)) as the only product, as shown in the MS spectrum in Figure 7A. The lack of detectable amounts of the

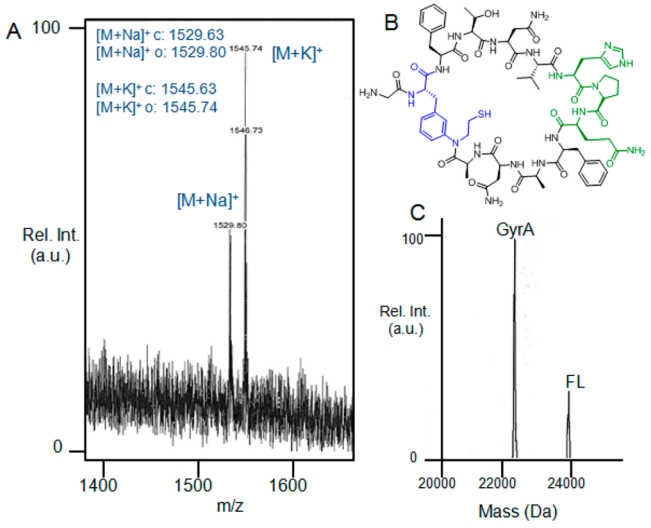


Figure 7. Isolation of streptavidin-binding peptide macrocycle from *E. coli*. (A) MALDI-TOF spectrum and (B) chemical structure of the *cyclo*(Strep1) as isolated from *E. coli* cells using streptavidin-coated beads. The HPQ motif is highlighted in green. (C) LC-MS spectrum of the Ni-affinity purified full-length precursor protein (FL) and spliced GyrA intein (GyrA) illustrating the extent of intein cleavage occurring *in vivo*.

acyclic byproduct, which would be also captured by this procedure as determined via control experiments, indicated that MeaF-mediated cyclization had occurred with high efficiency, that is, outcompeting hydrolysis of the full-length construct and/or of the thiolactone intermediate. Furthermore, the precursor protein was determined to have undergone >75% cleavage (Figure 7C), a value comparable to that observed for target sequences of comparable length in the CBD-fused format. On the basis of the expression level of the isolated GyrA protein and extent of *in vivo* cyclization, the yield of the Strep1-derived cyclopeptide was estimated to be around 0.85 mg L^{-1}.

Isolation of Improved Streptavidin-Binding Peptide from Macrocyclic Peptide Library. To begin to explore the potential utility of this methodology toward the discovery of macrocyclic peptides with novel or improved function, *cyclo*(Strep1) was subjected to mutagenesis in search for cyclopeptide variants with enhanced affinity toward streptavidin. To this end, we availed insights from the available crystal structure of streptavidin bound to a HPQ-containing peptide isolated by phage display.[73] Inspection of this complex revealed that, whereas the key HPQ motif inserts into the biotin binding site of streptavidin, the amino acid residues flanking this region also lie within short distance from the protein surface (<6 Å, Figure S7). In addition, these residues are likely to influence the β-turn conformation of the HPQ sequence, thereby affecting the affinity of the peptide ligand for the protein. Accordingly, a library of cyclic peptides was generated by randomizing the positions flanking the HPQ tripeptide in the Strep1 target sequence (V18, F22) using the degenerate codon NDC (12 codons/12 amino acids). In the corresponding precursor proteins, a FLAG peptide (sequence: DYKDDDDK) was placed upstream of the UAA and the semirandomized target sequence in order to obtain macrocyclic peptides with an N-terminal affinity tag useful for detection purposes. As a reference, a construct encoding for a FLAG-tagged version of *cyclo*(Strep1) was also prepared (entry 10, Table 1).

To screen the library, an ELISA-like assay was implemented and validated using a synthetic biotinylated FLAG peptide. As schematically illustrated in Figure 8A, this assay relies on capturing the streptavidin-binding peptide on a streptavidin-coated plate, followed by colorimetric detection and quantification of the amount of bound peptide by means of an anti-FLAG antibody−horseradish peroxidase (HRP) conjugate. After expression and cell lysis in a 96-well plate format, 480 recombinants from the Flag-Strep1(V18NDC/F22NDC) library were screened using this assay. This experiment led to the identification of 10 variants displaying a significantly enhanced response (>1.5-fold) as compared to that of the reference Flag-*cyclo*(Strep1) peptide (Figure 8B). Rescreening of these hits yielded two most promising variants, Flag-Strep2 and Flag-Strep3, which exhibited >2-fold higher signal in the immunoassay over that of the parent cyclic peptide (Figure S8). Upon sequencing, Flag-Strep2 was found to carry a single mutation (F22S), whereas both of the HRP flanking positions were mutated in Flag-Strep3 (V18Y, F22D) (entries 11 and 12, Table 1). Interestingly, many of the other improved variants contain a Tyr or Phe residue in position 18 (4/10) and a polar residue (H, D, S, R) in position 22 (6/10), thus defining a consensus for these sites. Efficient production *in vitro* of the two best cyclic peptides, Flag-*cyclo*(Strep2) and Flag-*cyclo*(Strep3), confirmed that both sequences undergo MeaF-mediated cyclization, leading to the expected macrocycles as the only product (Figure S9). These cyclic peptides, along with the parent cyclic peptide Flag-*cyclo*(Strep1), were then tested in the assay to measure their relative affinity for streptavidin (Figure 8C). From the resulting dose−response curve, the parent cyclic peptide was found to bind streptavidin with an equilibrium dissociation constant (K_D) of 7.7 μM. Importantly, both of the evolved macrocyclic peptides showed significantly (2- to 7-fold) improved affinity for this protein, with measured K_D values of 4.2 μM (Flag-*cyclo*(Strep2)) and 1.1 μM (Flag-*cyclo*(Strep3)). Altogether, these studies provides a proof-of-principle demonstration of the utility of the present methodology in aiding the discovery of macrocyclic binders against a target protein of interest.

Conclusions. In summary, we have developed a novel and versatile methodology useful for generating peptide macrocycles from ribosomally produced polypeptides *in vitro* and *in vivo*. This approach leverages the ability of two genetically encodable aminothiol amino acids, AmmF and MeaF, to induce a side chain → C-end peptide cyclization via an intein-mediated intramolecular transthioesterification followed by ring contraction through an S,N acyl transfer rearrangement. Whereas the AmmF-based strategy (Figure 3A) provides a means to generate macrocyclic peptides *in vitro* in a pH-controlled manner, the MeaF-based strategy (Figure 3B) offers the opportunity to achieve the spontaneous cyclization of peptide sequences of variable length (6-mer to 12-mer) and composition directly in living bacterial cells. As such, it

DOI: 10.1021/acschembio.5b00119
ACS Chem. Biol. 2015, 10, 1805−1816

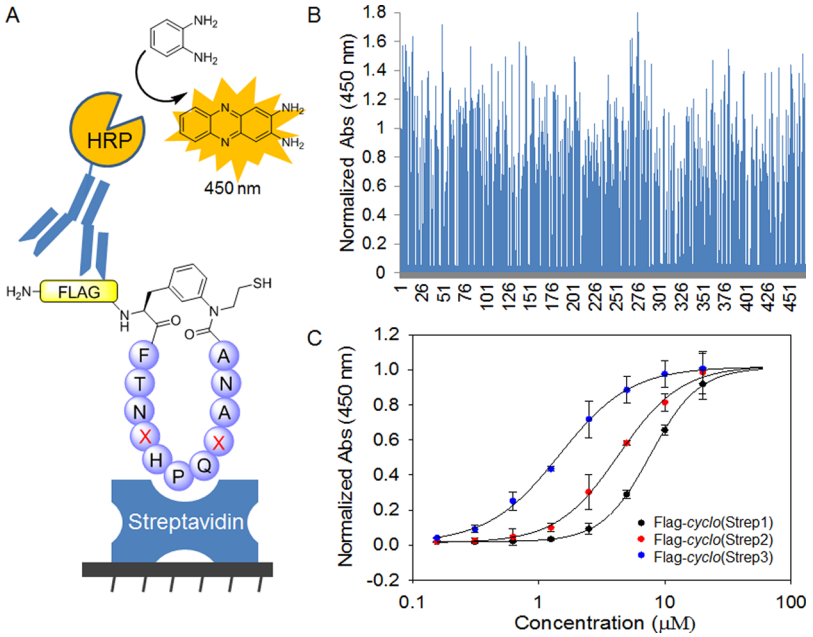


Figure 8. Selection of improved streptavidin-targeting cyclic peptides. (A) Schematic representation of the *in vitro* assay used to screen the Strep1-derived library and measure K_D for peptide binding to streptavidin. HRP, horseradish peroxidase. (B) Results from the primary screening of the Flag-Strep1(V18NDC/F22NDC) library in 96-well format. The absorbance values are normalized to that of the reference cyclic peptide Flag-*cyclo*(Strep1). (C) Streptavidin binding curves for isolated Flag-*cyclo*(Strep1) (K_D = 7.7 ± 0.6 μM), Flag-*cyclo*(Strep2) (K_D = 4.2 ± 0.1 μM), and Flag-*cyclo*(Strep3) (K_D = 1.1 ± 0.2 μM).

provides a new approach, complementary in scope to those previously reported,[49,50] to direct the biosynthesis of cyclic peptides of completely arbitrary sequence. Compared to SICLOPPS, for example, the present method offers the versatility to access the desired peptide macrocycle in a broader range of molecular arrangements (i.e., cyclic, lariat, or C-terminally fused to a carrier protein) according to the nature of the N-terminal tail. In this study, this structural feature was leveraged to couple in-cell library generation with a high-throughput *in vitro* assay to achieve the affinity maturation of a streptavidin-targeting cyclic peptide. Beyond that, we expect this methodology to make possible the screening of genetically encoded cyclopeptide libraries in a wide range of other formats, which include not only intracellular selection[42−44,48] or reporter systems[45] but also two-hybrid[75] or surface display[76] systems. Whereas the molecular size of the macrocycles currently accessible with this method (i.e., 900−1600 Da) seems to be appropriate for targeting extended biomolecular interfaces characterizing protein−protein interactions,[28−31,77] future studies will explore whether this strategy can be further evolved to provide access to smaller peptide ring structures.

METHODS

Cloning and Plasmid Construction. Genes encoding for the precursor proteins were placed under an ITPG-inducible T7 promoter in pET22 vectors (Novagen). Construction of the plasmids for expression of constructs CBD-4T through CBD-12T was described previously.[52] Construct 9 of Table 1 was prepared following an identical procedure. Construct 10 of Table 1 was prepared by stepwise assembly PCR using FLAG-HPQ-GyrA_for1/3, FLAG-HPQ-GyrA_-for2/3, and FLAG-HPQ-GyrA_for3/3 (Table S1) as forward primers, T7_Term_long_rev as reverse primer, and pET22b_CBD-12T as template. The PCR product (0.68 Kbp) was cloned into the *Nde*I/*Xho*I cassette of pET22b(+) vector. The FLAG-Strep1-derived library was prepared through randomization of positions V18 and F22 in pET22b_FLAG-Strep1 with NDC degenete codons (= 12 codons encoding C, D, F, G, H, I, L, N, R, S, V, and Y). The corresponding plasmid library was prepared by overlap extension PCR using primers HPQ_SOE(A)_for, HPQ_SOE(A)_rev, HPQ(NDC)_SOE(B)_for, and HPQ(NDC)_SOE(B)_rev for fragment generation and then primers HPQ_SOE(A)_for and T7m _Term_long_rev for amplification of the final gene (1.35 kb), which was then cloned into the *Xba*I/*Xho*I cassette of pET22b(+) vector. The library was transformed into *E. coli* DH5α cells, and a plasmid stock was generated by pooling >1000 cfu.

Synthesis of AmmF and MeaF. Detailed synthetic procedures and characterization data for AmmF and MeaF are provided as Supporting Information.

Fluorescence-Based Assay. *E. coli* BL21(DE3) cells were cotransformed with a pET22_YFP(stop) plasmid[50] encoding MetGly-(amber stop)YFP-His_6 and a pEVOL plasmid encoding the appropriate aminoacyl-tRNA synthetase. Cells were then grown in LB media containing ampicillin (50 mg L^{-1}) and chloramphenicol (26 mg L^{-1}) at 37 °C overnight. The overnight cultures were used to inoculate 96-deep well plates containing minimal M9 media. At an OD_{600} of 0.6, cell cultures were induced by adding arabinose (0.06%), IPTG (0.2 mM), and either AmmF or MeaF at a final concentration of 2 mM. After overnight growth at 27 °C, the cell cultures were diluted (1:1) with phosphate buffer (50 mM, 150 mM NaCl, pH 7.5), and fluorescence intensity (λ_{ex} = 514 nm; λ_{em} = 527 nm) was determined using a Tecan Infinite 1000 plate reader. Cell cultures containing no unnatural amino acid were included as controls. Each sample was measured in triplicate.

Protein Expression and Purification. Proteins were expressed in *E. coli* BL21(DE3) cells cotransformed with the plasmid encoding the biosynthetic precursor and pEVOL_AzF or pEVOL_O2beY vectors. After overnight growth, cells were used to inoculate M9 medium (0.4 L) supplemented with ampicillin (50 mg L^{-1}), chloramphenicol (34 mg L^{-1}), and 1% glycerol. At an OD_{600} of 0.6, protein expression was induced by adding L-arabinose (0.05%), IPTG (0.25 mM), and AmmF or MeaF (2 mM). Cultures were grown for an additional 12−24 h at 27 °C and harvested by centrifugation at 3400*g*. Frozen cells were resuspended in Tris buffer (50 mM; pH 7.4) containing 300 mM NaCl and 20 mM imidazole and lysed by sonication. Protein purification by Ni-NTA affinity chromatography was carried out

DOI: 10.1021/acschembio.5b00119
ACS Chem. Biol. 2015, 10, 1805−1816

using a Tris buffer (50 mM, pH 7.4, NaCl 150 mM) containing 50 and 300 mM imidazole for protein loading and elution, respectively. Protein samples were concentrated in potassium phosphate buffer (50 mM, NaCl 150 mM, pH 7.5) and stored at −80 °C.

***In Vitro* Macrocyclization Reactions.** Reactions were carried out incubating the AmmF-containing precursor protein (50 μM) in potassium phosphate buffer (50 mM, NaCl 150 mM, pH 5.0) in the presence of TCEP (10 mM). For the thiol-catalyzed reactions, 10 mM MESNA was also added. After 1.5 h, the pH of the solution was adjusted to 5.0. Intein cleavage was monitored and quantified by SDS-PAGE, and densitometry analysis of the gel bands using ImageJ software. The low MW products (8−10 kDa) of the reactions were analyzed by MALDI-TOF using a Brüker Autoflex II mass spectrometer.

***In Vivo* Macrocyclization Reactions.** The AmmF-containing precursor proteins were expressed in *E. coli* BL21(DE3) cells (50 mL culture) as described above. After 12 and 24 h of induction, cells were harvested by centrifugation (3400*g*), resuspended in 800 μL of potassium phosphate buffer (50 mM, NaCl 150 mM) at either pH 5.0 or 8.0, and lysed via sonication. The samples at pH 5.0 were used for detection of the thioester intermediate, and they were processed immediately. The samples at pH 8.0 were incubated with iodoacetamide (50 mM) for 3 h at RT prior to further processing. For protein isolation, the cell lysate samples were clarified by centrifugation (20 100*g*) and passed over chitin-coated beads (New England Biolabs); the beads were washed with lysis buffer, and the chitin-bound proteins were eluted with 70% acetonitrile in water (1 mL) followed by MALDI-TOF MS analysis.

Isolation of Streptavidin-Binding Peptide. The Strep1(MeaF) protein construct was expressed in *E. coli* BL21(DE3) cells (50 mL culture) for 24 h at 27 °C as described above. After harvesting, cells were resuspended in 800 μL of potassium phosphate buffer (50 mM, NaCl 150 mM, 7.5), lysed via sonication, and centrifuged at 20 100*g*. The clarified cell lysate was passed over streptavidin beads (Pierce), and the immobilized peptide was eluted with 70% acetonitrile in water and analyzed by MALDI-TOF MS.

Library Screening. Recombinants from the FLAG-Strep1-(V18NDC/F22NDC) library were expressed in *E. coli* BL21(DE3) cells in 96 deep well plates (1 mL culture) for 24 h at 27 °C. After harvesting, cells were resuspended in 400 μL of TBS (Tris 50 mM, NaCl 150 mM, 7.4) and incubated for 1 h at 37 °C in the presence of lysosyme, DNase, and $MgCl_2$ to lyse the cells, followed by centrifugation at 3400*g*. The clarified cell lysate (200 μL) was transferred to NeutrAvidin coated 96-well plates (Pierce) and incubated for 2 h at RT. The plates were washed three times with 200 μL of TBS + 0.5% Tween 20. Next, a solution of anti-FLAG-antibody−HRP conjugate (100 μL, 1:2500 dilution) was added to each well and incubated for 1 h. The plates were washed three times with 200 μL of TBS + 0.5% Tween 20. 100 μL of SigmaFast OPD solution was added to each well, and the absorbance at 450 nm was measured after 20 min. The signals were normalized to the signal corresponding to the cells containing the FLAG-Strep1 cyclic peptide. The 10 most active variants (>1.5 rel. act.) were re-expressed and screened again in triplicate using the same procedure (Figure S8).

K_D Determination for Streptavidin Binding. The cyclic peptides Flag-*cyclo*(Strep1), Flag-*cyclo*(Strep2), and Flag-*cyclo*(Strep3) were obtained via cyclization of the corresponding full-length precursor proteins in TBS in the presence of TCEP (20 mM) and thiophenol (10 mM). The full-length precursor proteins were obtained via expression in *E. coli* BL21(DE3) cells for 7 h at 22 °C followed by Ni-affinity purification as described above. In each case, the macrocyclization reaction proceeded quantitatively, yielding the cyclic peptide as the only product, as determined by SDS-PAGE and MS (Figure S9). A solution of each of the macrocyclic peptide (50 μL) was added to NeutrAvidin-coated plates at varying concentrations (0.156−20 μM) and assayed in duplicate as described above. K_D values were calculated with SigmaPlot via fitting of the dose−response curves using a 1:1 binding model.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Synthetic procedures, additional MS spectra, sequences of oligonucleotide primers, crystal structure of streptavidin in complex with a disulfide bridged peptide, SDS-PAGE gel, and results from rescreening the 10 most active variants isolated from primary screening of the Flag-Strep1(V18NDC/F22NDC) library. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acschembio.5b00119.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: fasan@chem.rochester.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported in part by National Science Foundation grant no. CHE-1112342 and in part by National Institute of Health grant no. R21 CA187502. N.T.J. and L.J.P. acknowledges the NSF REU program for financial support. MS instrumentation was supported by the NSF grant nos. CHE-0840410 and CHE-0946653. We also thank Dr. Kestutis Bendinskas (SUNY Oswego) for providing access to the Bruker Autoflex II MALDI-TOF mass spectrometer.

■ REFERENCES

(1) Driggers, E. M., Hale, S. P., Lee, J., and Terrett, N. K. (2008) The exploration of macrocycles for drug discovery–an underexploited structural class. *Nat. Rev. Drug Discovery 7*, 608−624.

(2) Marsault, E., and Peterson, M. L. (2011) Macrocycles are great cycles: applications, opportunities, and challenges of synthetic macrocycles in drug discovery. *J. Med. Chem. 54*, 1961−2004.

(3) Al-Obeidi, F., Castrucci, A. M., Hadley, M. E., and Hruby, V. J. (1989) Potent and prolonged acting cyclic lactam analogues of alpha-melanotropin: design based on molecular dynamics. *J. Med. Chem. 32*, 2555−2561.

(4) Tang, Y. Q., Yuan, J., Osapay, G., Osapay, K., Tran, D., Miller, C. J., Ouellette, A. J., and Selsted, M. E. (1999) A cyclic antimicrobial peptide produced in primate leukocytes by the ligation of two truncated alpha-defensins. *Science 286*, 498−502.

(5) Dechantsreiter, M. A., Planker, E., Matha, B., Lohof, E., Holzemann, G., Jonczyk, A., Goodman, S. L., and Kessler, H. (1999) N-Methylated cyclic RGD peptides as highly active and selective alpha(V)beta(3) integrin antagonists. *J. Med. Chem. 42*, 3033−3040.

(6) Graciani, N. R., Tsang, K. Y., McCutchen, S. L., and Kelly, J. W. (1994) Amino acids that specify structure through hydrophobic clustering and histidine-aromatic interactions lead to biologically active peptidomimetics. *Bioorg. Med. Chem. 2*, 999−1006.

(7) Fasan, R., Dias, R. L., Moehle, K., Zerbe, O., Vrijbloed, J. W., Obrecht, D., and Robinson, J. A. (2004) Using a beta-hairpin to mimic an alpha-helix: cyclic peptidomimetic inhibitors of the p53−HDM2 protein−protein interaction. *Angew. Chem., Int. Ed. 43*, 2109−2112.

(8) Dias, R. L., Fasan, R., Moehle, K., Renard, A., Obrecht, D., and Robinson, J. A. (2006) Protein ligand design: from phage display to synthetic protein epitope mimetics in human antibody Fc-binding peptidomimetics. *J. Am. Chem. Soc. 128*, 2726−27232.

(9) Cardoso, R. M., Brunel, F. M., Ferguson, S., Zwick, M., Burton, D. R., Dawson, P. E., and Wilson, I. A. (2007) Structural basis of enhanced binding of extended and helically constrained peptide epitopes of the broadly neutralizing HIV-1 antibody 4E10. *J. Mol. Biol. 365*, 1533−1544.

DOI: 10.1021/acschembio.5b00119
ACS Chem. Biol. 2015, 10, 1805−1816

(10) Henchey, L. K., Porter, J. R., Ghosh, I., and Arora, P. S. (2010) High specificity in protein recognition by hydrogen-bond-surrogate alpha-helices: selective inhibition of the p53/MDM2 complex. *ChemBioChem 11*, 2104−2107.
(11) Satoh, T., Li, S., Friedman, T. M., Wiaderkiewicz, R., Korngold, R., and Huang, Z. (1996) Synthetic peptides derived from the fourth domain of CD4 antagonize off function and inhibit T cell activation. *Biochem. Biophys. Res. Commun. 224*, 438−443.
(12) Fairlie, D. P., Tyndall, J. D. A., Reid, R. C., Wong, A. K., Abbenante, G., Scanlon, M. J., March, D. R., Bergman, D. A., Chai, C. L. L., and Burkett, B. A. (2000) Conformational selection of inhibitors and substrates by proteolytic enzymes: Implications for drug design and polypeptide processing. *J. Med. Chem. 43*, 1271−1281.
(13) Wang, D., Liao, W., and Arora, P. S. (2005) Enhanced metabolic stability and protein-binding properties of artificial alpha helices derived from a hydrogen-bond surrogate: application to Bcl-xL. *Angew. Chem., Int. Ed. 44*, 6525−6529.
(14) Gudmundsson, O. S., Vander Velde, D. G., Jois, S. D., Bak, A., Siahaan, T. J., and Borchardt, R. T. (1999) The effect of conformation of the acyloxyalkoxy-based cyclic prodrugs of opioid peptides on their membrane permeability. *J. Pept. Res. 53*, 403−413.
(15) Walensky, L. D., Kung, A. L., Escher, I., Malia, T. J., Barbuto, S., Wright, R. D., Wagner, G., Verdine, G. L., and Korsmeyer, S. J. (2004) Activation of apoptosis in vivo by a hydrocarbon-stapled BH3 helix. *Science 305*, 1466−1470.
(16) Rezai, T., Yu, B., Millhauser, G. L., Jacobson, M. P., and Lokey, R. S. (2006) Testing the conformational hypothesis of passive membrane permeability using synthetic cyclic peptide diastereomers. *J. Am. Chem. Soc. 128*, 2510−2511.
(17) Finking, R., and Marahiel, M. A. (2004) Biosynthesis of nonribosomal peptides. *Annu. Rev. Microbiol. 58*, 453−488.
(18) Fischbach, M. A., and Walsh, C. T. (2006) Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: logic, machinery, and mechanisms. *Chem. Rev. 106*, 3468−3496.
(19) Arnison, P. G., Bibb, M. J., Bierbaum, G., Bowers, A. A., Bugni, T. S., Bulaj, G., Camarero, J. A., Campopiano, D. J., Challis, G. L., Clardy, J., Cotter, P. D., Craik, D. J., Dawson, M., Dittmann, E., Donadio, S., Dorrestein, P. C., Entian, K. D., Fischbach, M. A., Garavelli, J. S., Goransson, U., Gruber, C. W., Haft, D. H., Hemscheidt, T. K., Hertweck, C., Hill, C., Horswill, A. R., Jaspars, M., Kelly, W. L., Klinman, J. P., Kuipers, O. P., Link, A. J., Liu, W., Marahiel, M. A., Mitchell, D. A., Moll, G. N., Moore, B. S., Muller, R., Nair, S. K., Nes, I. F., Norris, G. E., Olivera, B. M., Onaka, H., Patchett, M. L., Piel, J., Reaney, M. J., Rebuffat, S., Ross, R. P., Sahl, H. G., Schmidt, E. W., Selsted, M. E., Severinov, K., Shen, B., Sivonen, K., Smith, L., Stein, T., Sussmuth, R. D., Tagg, J. R., Tang, G. L., Truman, A. W., Vederas, J. C., Walsh, C. T., Walton, J. D., Wenzel, S. C., Willey, J. M., and van der Donk, W. A. (2013) Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Nat. Prod. Rep. 30*, 108−160.
(20) Knerr, P. J., and van der Donk, W. A. (2012) Discovery, biosynthesis, and engineering of lantipeptides. *Annu. Rev. Biochem. 81*, 479−505.
(21) Craik, D. J., and Conibear, A. C. (2011) The chemistry of cyclotides. *J. Org. Chem. 76*, 4805−4817.
(22) Donia, M. S., Ravel, J., and Schmidt, E. W. (2008) A global assembly line for cyanobactins. *Nat. Chem. Biol. 4*, 341−343.
(23) White, C. J., and Yudin, A. K. (2011) Contemporary strategies for peptide macrocyclization. *Nat. Chem. 3*, 509−524.
(24) Frost, J. R., Smith, J. M., and Fasan, R. (2013) Design, synthesis, and diversification of ribosomally derived peptide macrocycles. *Curr. Opin. Struct. Biol. 23*, 571−580.
(25) Smith, J. M., Frost, J. R., and Fasan, R. (2013) Emerging strategies to access peptide macrocycles from genetically encoded polypeptides. *J. Org. Chem. 78*, 3525−3531.
(26) Passioura, T., Katoh, T., Goto, Y., and Suga, H. (2014) Selection-based discovery of druglike macrocyclic peptides. *Annu. Rev. Biochem. 83*, 727−752.
(27) Aboye, T. L., and Camarero, J. A. (2012) Biological synthesis of circular polypeptides. *J. Biol. Chem. 287*, 27026−27032.
(28) Wrighton, N. C., Farrell, F. X., Chang, R., Kashyap, A. K., Barbone, F. P., Mulcahy, L. S., Johnson, D. L., Barrett, R. W., Jolliffe, L. K., and Dower, W. J. (1996) Small peptides as potent mimetics of the protein hormone erythropoietin. *Science 273*, 458−464.
(29) DeLano, W. L., Ultsch, M. H., de Vos, A. M., and Wells, J. A. (2000) Convergent solutions to binding at a protein−protein interface. *Science 287*, 1279−1283.
(30) Livnah, O., Stura, E. A., Johnson, D. L., Middleton, S. A., Mulcahy, L. S., Wrighton, N. C., Dower, W. J., Jolliffe, L. K., and Wilson, I. A. (1996) Functional mimicry of a protein hormone by a peptide agonist: the EPO receptor complex at 2.8 Å. *Science*, 464−471.
(31) Skelton, N. J., Chen, Y. M., Dubree, N., Quan, C., Jackson, D. Y., Cochran, A., Zobel, K., Deshayes, K., Baca, M., Pisabarro, M. T., and Lowman, H. B. (2001) Structure−function analysis of a phage display-derived peptide that binds to insulin-like growth factor binding protein I. *Biochemistry 40*, 8487−8498.
(32) Pasqualini, R., Koivunen, E., and Ruoslahti, E. (1995) A peptide isolated from phage display libraries is a structural and functional mimic of an RGD-binding site on integrins. *J. Cell Biol. 130*, 1189−1196.
(33) Heinis, C., Rutherford, T., Freund, S., and Winter, G. (2009) Phage-encoded combinatorial chemical libraries based on bicyclic peptides. *Nat. Chem. Biol. 5*, 502−507.
(34) Angelini, A., Cendron, L., Chen, S. Y., Touati, J., Winter, G., Zanotti, G., and Heinis, C. (2012) Bicyclic peptide inhibitor reveals large contact interface with a protease target. *ACS Chem. Biol. 7*, 817−821.
(35) Jafari, M. R., Deng, L., Kitov, P. I., Ng, S., Matochko, W. L., Tjhung, K. F., Zeberoff, A., Elias, A., Klassen, J. S., and Derda, R. (2014) Discovery of light-responsive ligands through screening of a light-responsive genetically encoded library. *ACS Chem. Biol. 9*, 443−450.
(36) Yamagishi, Y., Shoji, I., Miyagawa, S., Kawakami, T., Katoh, T., Goto, Y., and Suga, H. (2011) Natural product-like macrocyclic N-methyl-peptide inhibitors against a ubiquitin ligase uncovered from a ribosome-expressed de novo library. *Chem. Biol. 18*, 1562−1570.
(37) Hayashi, Y., Morimoto, J., and Suga, H. (2012) In vitro selection of anti-Akt2 thioether-macrocyclic peptides leading to isoform-selective inhibitors. *ACS Chem. Biol. 7*, 607−613.
(38) Morimoto, J., Hayashi, Y., and Suga, H. (2012) Discovery of macrocyclic peptides armed with a mechanism-based warhead: isoform-selective inhibition of human deacetylase SIRT2. *Angew. Chem., Int. Ed. 51*, 3423−3427.
(39) Schlippe, Y. V. G., Hartman, M. C. T., Josephson, K., and Szostak, J. W. (2012) In vitro selection of highly modified cyclic peptides that act as tight binding inhibitors. *J. Am. Chem. Soc. 134*, 10469−10477.
(40) Hofmann, F. T., Szostak, J. W., and Seebeck, F. P. (2012) In vitro selection of functional lantipeptides. *J. Am. Chem. Soc. 134*, 8038−8041.
(41) Kawakami, T., Ishizawa, T., Fujino, T., Reid, P. C., Suga, H., and Murakami, H. (2013) In vitro selection of multiple libraries created by genetic code reprogramming to discover macrocyclic peptides that antagonize VEGFR2 activity in living cells. *ACS Chem. Biol. 8*, 1205−1214.
(42) Horswill, A. R., Savinov, S. N., and Benkovic, S. J. (2004) A systematic method for identifying small-molecule modulators of protein-protein interactions. *Proc. Natl. Acad. Sci. U.S.A. 101*, 15591−15596.
(43) Naumann, T. A., Tavassoli, A., and Benkovic, S. J. (2008) Genetic selection of cyclic peptide Dam methyltransferase inhibitors. *ChemBioChem 9*, 194−197.
(44) Young, T. S., Young, D. D., Ahmad, I., Louis, J. M., Benkovic, S. J., and Schultz, P. G. (2011) Evolution of cyclic peptide protease inhibitors. *Proc. Natl. Acad. Sci. U.S.A. 108*, 11052−11056.
(45) Cheng, L., Naumann, T. A., Horswill, A. R., Hong, S. J., Venters, B. J., Tomsho, J. W., Benkovic, S. J., and Keiler, K. C. (2007)

DOI: 10.1021/acschembio.5b00119
ACS Chem. Biol. 2015, 10, 1805−1816

Discovery of antibacterial cyclic peptides that inhibit the ClpXP protease. *Protein Sci. 16*, 1535−1542.
(46) Tavassoli, A., and Benkovic, S. J. (2005) Genetically selected cyclic-peptide inhibitors of AICAR transformylase homodimerization. *Angew. Chem., Int. Ed. 44*, 2760−2763.
(47) Tavassoli, A., Lu, Q., Gam, J., Pan, H., Benkovic, S. J., and Cohen, S. N. (2008) Inhibition of HIV budding by a genetically selected cyclic peptide targeting the Gag−TSG101 interaction. *ACS Chem. Biol. 3*, 757−764.
(48) Kritzer, J. A., Hamamichi, S., McCaffery, J. M., Santagata, S., Naumann, T. A., Caldwell, K. A., Caldwell, G. A., and Lindquist, S. (2009) Rapid selection of cyclic peptides that reduce alpha-synuclein toxicity in yeast and animal models. *Nat. Chem. Biol. 5*, 655−663.
(49) Scott, C. P., Abel-Santos, E., Wall, M., Wahnon, D. C., and Benkovic, S. J. (1999) Production of cyclic peptides and proteins in vivo. *Proc. Natl. Acad. Sci. U.S.A. 96*, 13638−13643.
(50) Bionda, N., Cryan, A. L., and Fasan, R. (2014) Bioinspired strategy for the ribosomal synthesis of thioether-bridged macrocyclic peptides in bacteria. *ACS Chem. Biol. 9*, 2008−2013.
(51) Smith, J. M., Vitali, F., Archer, S. A., and Fasan, R. (2011) Modular assembly of macrocyclic organo-peptide hybrids using synthetic and genetically encoded precursors. *Angew. Chem., Int. Ed. 50*, 5075−5080.
(52) Satyanarayana, M., Vitali, F., Frost, J. R., and Fasan, R. (2012) Diverse organo-peptide macrocycles via a fast and catalyst-free oxime/intein-mediated dual ligation. *Chem. Commun. 48*, 1461−1463.
(53) Frost, J. R., Vitali, F., Jacob, N. T., Brown, M. D., and Fasan, R. (2013) Macrocyclization of organo-peptide hybrids through a dual bio-orthogonal ligation: insights from structure−reactivity Studies. *ChemBioChem 14*, 147−160.
(54) Smith, J. M., Hill, N. C., Krasniak, P. J., and Fasan, R. (2014) Synthesis of bicyclic organo−peptide hybrids via oxime/intein-mediated macrocyclization followed by disulfide bond formation. *Org. Biomol. Chem. 12*, 1135−1142.
(55) Wang, L., and Schultz, P. G. (2004) Expanding the genetic code. *Angew. Chem., Int. Ed. 44*, 34−66.
(56) Stokes, A. L., Miyake-Stoner, S. J., Peeler, J. C., Nguyen, D. P., Hammer, R. P., and Mehl, R. A. (2009) Enhancing the utility of unnatural amino acid synthetases by manipulating broad substrate specificity. *Mol. BioSyst. 5*, 1032−1038.
(57) Young, D. D., Young, T. S., Jahnz, M., Ahmad, I., Spraggon, G., and Schultz, P. G. (2011) An evolved aminoacyl-tRNA synthetase with atypical polysubstrate specificity. *Biochemistry 50*, 1894−1900.
(58) Seyedsayamdost, M. R., Xie, J., Chan, C. T., Schultz, P. G., and Stubbe, J. (2007) Site-specific insertion of 3-aminotyrosine into subunit alpha2 of *E. coli* ribonucleotide reductase: direct evidence for involvement of Y730 and Y731 in radical propagation. *J. Am. Chem. Soc. 129*, 15060−15071.
(59) Santoro, S. W., Wang, L., Herberich, B., King, D. S., and Schultz, P. G. (2002) An efficient system for the evolution of aminoacyl-tRNA synthetase specificity. *Nat. Biotechnol. 20*, 1044−1048.
(60) Chin, J. W., Santoro, S. W., Martin, A. B., King, D. S., Wang, L., and Schultz, P. G. (2002) Addition of p-azido-L-phenylalanine to the genetic code of *Escherichia coli*. *J. Am. Chem. Soc. 124*, 9026−9027.
(61) Wang, L., Zhang, Z., Brock, A., and Schultz, P. G. (2003) Addition of the keto functional group to the genetic code of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A. 100*, 56−61.
(62) Deiters, A., and Schultz, P. G. (2005) In vivo incorporation of an alkyne into proteins in *Escherichia coli*. *Bioorg. Med. Chem. Lett. 15*, 1521−1524.
(63) Telenti, A., Southworth, M., Alcaide, F., Daugelat, S., Jacobs, W. R., Jr., and Perler, F. B. (1997) The *Mycobacterium xenopi* GyrA protein splicing element: characterization of a minimal intein. *J. Bacteriol. 179*, 6378−6382.
(64) Young, T. S., Ahmad, I., Yin, J. A., and Schultz, P. G. (2010) An enhanced system for unnatural amino acid mutagenesis in *E. coli*. *J. Mol. Biol. 395*, 361−374.
(65) Bennett, B. D., Kimball, E. H., Gao, M., Osterhout, R., Van Dien, S. J., and Rabinowitz, J. D. (2009) Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nat. Chem. Biol. 5*, 593−599.
(66) Gentle, I. E., De Souza, D. P., and Baca, M. (2004) Direct production of proteins with N-terminal cysteine for site-specific conjugation. *Bioconjugate Chem. 15*, 658−663.
(67) Villain, M., Vizzavona, J., and Rose, K. (2001) Covalent capture: a new tool for the purification of synthetic and recombinant polypeptides. *Chem. Biol. 8*, 673−679.
(68) Nguyen, D. P., Elliott, T., Holt, M., Muir, T. W., and Chin, J. W. (2011) Genetically encoded 1,2-aminothiols facilitate rapid and site-specific protein labeling via a bio-orthogonal cyanobenzothiazole condensation. *J. Am. Chem. Soc. 133*, 11418−11421.
(69) Perrin, D. D. (1965) *Dissociation Constants of Organic Bases in Aqueous Solution*, Butterworth, London.
(70) Kreevoy, M. M., Harper, E. T., Duvall, R. E., Wilgus, H. S., and Ditsch, L. T. (1960) Inductive effects on the acid dissociation constants of mercaptans. *J. Am. Chem. Soc. 82*, 4899−4902.
(71) Canne, L. E., Bark, S. J., and Kent, S. B. H. (1996) Extending the applicability of native chemical ligation. *J. Am. Chem. Soc. 118*, 5891−5896.
(72) Naumann, T. A., Savinov, S. N., and Benkovic, S. J. (2005) Engineering an affinity tag for genetically encoded cyclic peptides. *Biotechnol. Bioeng. 92*, 820−830.
(73) Katz, B. A. (1995) Binding to protein targets of peptidic leads discovered by phage display: crystal structures of streptavidin-bound linear and cyclic peptide ligands containing the HPQ sequence. *Biochemistry 34*, 15421−15429.
(74) Giebel, L. B., Cass, R. T., Milligan, D. L., Young, D. C., Arze, R., and Johnson, C. R. (1995) Screening of cyclic peptide phage libraries identifies ligands that bind streptavidin with high affinities. *Biochemistry 34*, 15430−15435.
(75) Dove, S. L., and Hochschild, A. (2004) A bacterial two-hybrid system based on transcription activation. *Methods Mol. Biol. 261*, 231−246.
(76) Lofblom, J. (2011) Bacterial display in combinatorial protein engineering. *Biotechnol. J. 6*, 1115−1129.
(77) Fasan, R., Dias, R. L., Moehle, K., Zerbe, O., Obrecht, D., Mittl, P. R., Grutter, M. G., and Robinson, J. A. (2006) Structure−activity studies in a family of beta-hairpin protein epitope mimetic inhibitors of the p53−HDM2 protein−protein interaction. *ChemBioChem 7*, 515−526.

DOI: 10.1021/acschembio.5b00119
ACS Chem. Biol. 2015, 10, 1805−1816

Ribosomal Synthesis of Macrocyclic Peptides in Vitro and in Vivo Mediated by Genetically Encoded Amino-thiol Unnatural Amino Acids

John R. Frost, Nicholas T. Jacob, Louis J. Papa, Andrew E. Owens and Rudi Fasan*

Department of Chemistry, University of Rochester, Hutchinson Hall, Rochester NY, 14627

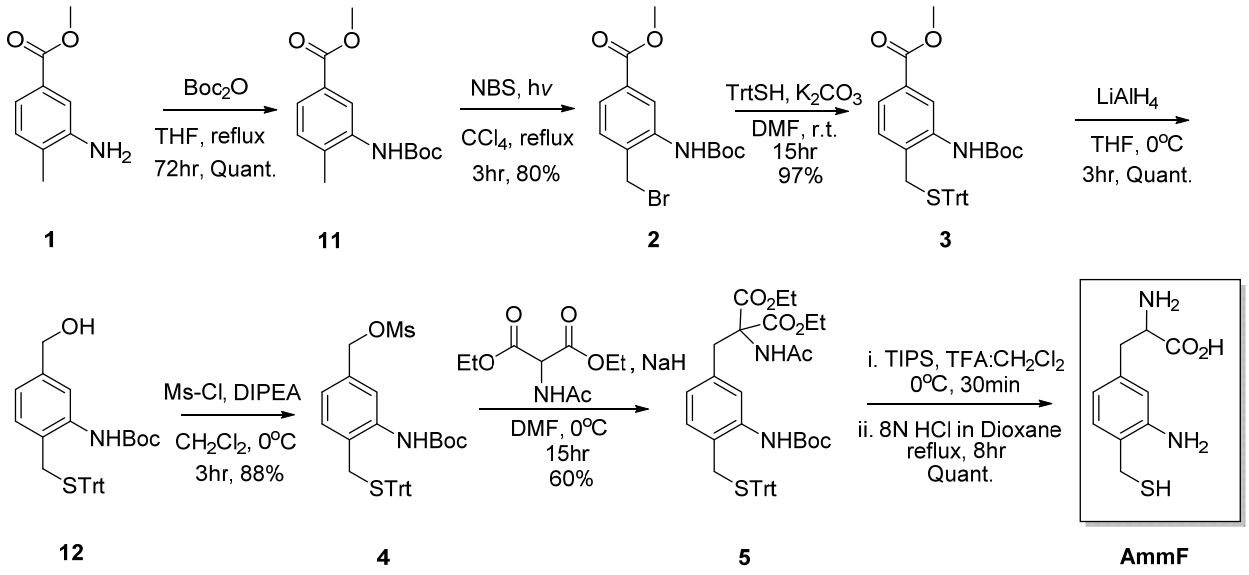
* corresponding author: fasan@chem.rochester.edu

Table of Contents

GENERAL INFORMATION.

Chemical reagents and solvents were purchased from Sigma–Aldrich, Acros Organics, and Fluka. Silica gel chromatography purifications were carried out by using AMD Silica Gel 60 230–400 mesh. ^{1}H and ^{13}C NMR spectra were recorded on Bruker Avance spectrometers by using solvent peaks as reference. LC-MS analyses were performed on a Thermo Scientific LTQ Velos ESI/ion-trap mass spectrometer coupled to an Accela U-HPLC. MALDI-TOF spectra were acquired on a Bruker Autoflex III MALDI-TOF spectrometer by using a stainless steel MALDI plate and sinapinic acid as matrix.

SYNTHETIC PROCEDURES



Scheme S1: Synthesis of AmmF

Synthesis of *N*-Boc-*S*-trityl-3-amino-4-(mercaptomethyl)benzoate methyl ester (3): Methyl 3-amino-4-methylbenzoate (**1**) (9.7 g, 58.7 mmol) and di-*tert*-butyl dicarbonate (17 mL, 74 mmol, 1.2 equiv) were dissolved in anhydrous THF (200 mL), and the mixture was heated under reflux for 72 h. The solvent was removed by rotary evaporation to afford a pink–white solid that was suspended in ice-cold hexanes (30 mL) and filtered to afford methyl 3-((*tert*-butoxycarbonyl)amino)-4-methylbenzoate (**11**) as a white solid (99 % yield). ^{1}H NMR ($CDCl_3$, 500 MHz): δ = 8.45 (s, 1 H), 7.69 (d, *J*=7.9 Hz, 1 H), 7.21 (d, *J*=7.9 Hz, 1 H), 6.29 (s, 1 H), 3.90

(s, 3 H), 2.30 (s, 3 H), 1.55 ppm (d, *J*=11.2 Hz, 9 H); ^{13}C NMR ($CDCl_3$, 126 MHz): δ = 166.9, 152.8, 136.4, 132.6, 130.3, 128.9, 124.9, 121.8, 80.8, 52.0, 28.3, 17.9 ppm; MS (ESI): calcd for $C_{14}H_{19}NO_4$: 288.31 $[M+Na]^+$; found: 288.20.

Compound **11** (6.63 g, 25 mmol) was dissolved in carbon tetrachloride (100 mL), and *N*-bromosuccinamide (4.89 g, 27.5 mmol, 1.1 equiv) was added. The reaction vessel was equipped with a reflux condenser and irradiated with UV light for 3 h, then cooled to room temperature and the mixture was filtered. The filtrate was dissolved in CH_2Cl_2 (100 mL), washed with saturated K_2CO_3 (aq) and brine, then dried over anhydrous $MgSO_4$. Volatiles were removed to afford methyl 4-(bromomethyl)-3-((*tert*-butoxycarbonyl)amino)benzoate (**2**) (6.7 g, 78 % yield) as an orange–white solid. ^{1}H NMR ($CDCl_3$, 500 MHz): δ = 8.47 (s, 1 H), 7.73 (dd, *J*=8.0, 1.7 Hz, 1 H), 7.36 (d, *J*=8.0 Hz, 1 H), 6.75 (s, 1 H), 4.50 (s, 2 H), 3.91 (s, 3 H), 1.55 ppm (s, 9 H); ^{13}C NMR ($CDCl_3$, 126 MHz): δ 28.2, 29.9, 52.3, 81.3, 123.8, 125.1, 130.0, 131.5, 131.7, 136.9, 152.6, 166.2 ppm; MS (ESI): calcd for $C_{14}H_{18}BrNO_4$: 345.20 $[M+H]^+$; found: 345.24.

Compound **2** (6.7 g, 19.59 mmol), tritylmercaptan (6.5 g, 23.5 mmol, 1.2 equiv) and potassium carbonate (3.25 g, 23.5 mmol, 1.2 equiv) were dissolved in dry DMF (100 mL), and the mixture was stirred under argon at room temperature for 15 h, concentrated by rotary evaporation, then dissolved in CH_2Cl_2. The solution was washed with ice-cold H_2O, with saturated $NaHCO_3$ (aq), and finally with brine. The organic layer was dried over anhydrous $MgSO_4$ and filtered, and volatiles were removed to afford methyl 3-((*tert*-butoxycarbonyl)amino)-4-((tritylthio)methyl)benzoate **(3)** as a golden yellow solid (10.24 g, 97 % crude yield). This material was carried forward without further purification. ^{1}H NMR ($CDCl_3$, 500 MHz): δ = 8.41 (s, 1 H), 7.65 (d, *J*=7.9 Hz, 1 H), 7.48 (d, *J*=8.0 Hz, 5 H), 7.34 (t, *J*=7.8 Hz, 6 H), 7.25 (t, *J*=7.3 Hz, 5 H), 7.18 (d, *J*=8.0 Hz, 1 H), 6.72 (s, 1 H), 3.88 (s, 3 H), 3.21 (s, 2 H), 1.56 ppm (d, *J*=2.5 Hz, 9 H); ^{13}C NMR ($CDCl_3$, 126 MHz): δ = 166.7, 152.8, 144.1, 136.9, 130.7, 129.3, 128.2, 126.9, 124.9, 123.1, 80.8, 67.4, 52.1, 34.1, 28.4 ppm. MS (ESI): calcd for $C_{33}H_{33}NO_4S$: 562.69 $[M+Na]^+$; found: 562.74.

Synthesis of *N*-Boc-*S*-trityl-3-amino-4-(mercaptomethyl)benzyl mesylate (4): Compound **3** (20.32 g, 48 mmol) was dissolved in anhydrous THF (400 mL), then the solution was cooled to 0 °C. A solution of $LiAlH_4$ in THF (1 M, 52.8 mL, 52.8 mmol, 1.1 equiv) was added slowly. The

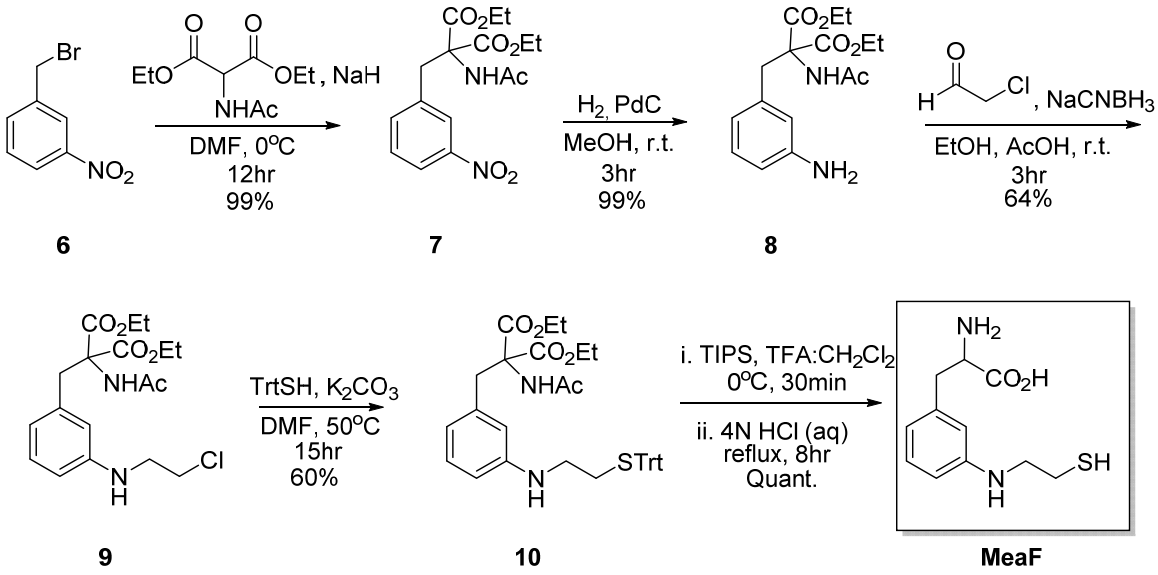
reaction mixture was stirred under argon at 0 °C for 3 h, the reaction was quenched by slow addition of cold H_2O (3 mL) and 4N NaOH (aq) (1 mL) at 0 °C, then the mixture was stirred for 10 min at room temperature. The mixture was concentrated under reduced pressure, suspended in EtOAc/sat. $NaHCO_3$ (10:1, 330 mL), then filtered through celite. The filtrate was washed once with saturated $NaHCO_3$, then with brine. The organic layer was dried with anhydrous $MgSO_4$, and volatiles were removed to afford a yellow solid, which was purified by flash column chromatography (silica gel, hexanes/EtOAc 7:3) to afford *N*-Boc-*S*-trityl-3-amino-4-(mercaptomethyl)benzyl alcohol (**12**) as a yellow oil (18 g, 95 % yield). ^{1}H NMR ($CDCl_3$, 500 MHz): δ 7.78 (s, 1 H), 7.49 (d, *J*=7.3 Hz, 5 H), 7.34 (t, *J*=7.7 Hz, 5 H), 7.26 (t, *J*=3.0 Hz, 5 H), 7.13 (d, *J*=7.8 Hz, 1 H), 7.01 (d, *J*=7.8 Hz, 1 H), 6.73 (s, 1 H), 4.63 (s, 2 H), 3.17 (s, 2 H), 1.54 ppm (s, 9 H); ^{13}C NMR ($CDCl_3$, 126 MHz): δ 153.1, 144.3, 141.5, 136.8, 130.9, 129.3, 128.2, 126.9, 124.5, 122.2, 120.4, 80.5, 67.2, 65.1, 33.9, 28.4 ppm; MS (ESI): calcd for $C_{32}H_{33}NO_3S$: 534.68 $[M+Na]^+$; found: 535.64.

Compound **12** (9.3 g, 18.19 mmol) was dissolved in anhydrous CH_2Cl_2 (100 mL), and the solution was cooled to 0 °C. Methanesulfonylchloride (1.8 mL, 23.66 mmol, 1.3 equiv) and *N*,*N*-diisopropylethylamine (DIPEA; 4.2 mL, 23.66 mmol, 1.3 equiv) were added, and the reaction mixture was stirred under argon at 0 °C for 2 h. The mixture was then dissolved in CH_2Cl_2, washed twice with saturated $NaHCO_3$ (aq), and then once with brine. The organic layer was dried over anhydrous $MgSO_4$, and volatiles were removed to afford compound **4** as a yellow solid (9.42 g, 88 % yield). The material was carried forward without further purification. ^{1}H NMR ($CDCl_3$, 500 MHz): δ 7.88 (s, 1 H), 7.49 (d, *J*=7.3 Hz, 5 H), 7.34 (t, *J*=7.7 Hz, 5 H), 7.26 (d, *J*=14.6 Hz, 5 H), 7.16 (d, *J*=7.8 Hz, 1 H), 7.04 (d, *J*=9.5 Hz, 1 H), 6.75 (s, 1 H), 5.18 (s, 2 H), 3.17 (s, 2 H), 2.90 (s, 3 H), 1.54 ppm (s, 9 H); ^{13}C NMR ($CDCl_3$, 126 MHz): δ 152.8, 144.1, 137.3, 133.7, 131.3, 129.3, 128.2, 126.9, 126.3, 123.8, 121.9, 80.8, 71.3, 67.3, 38.4, 33.9, 28.4 ppm; MS (ESI): calcd for $C_{33}H_{35}NO_5S_2$: 612.76 $[M+Na]^+$; found: 612.04.

Synthesis of diethyl 2-acetamido-2-(3-((*tert*-butoxycarbonyl)amino)-4 ((tritylthio)methyl)benzyl)malonate (5): To a dry, argon-filled round-bottom flask was added compound **4** (9.42g) and diethylacetamidomalonate (4.52g, 20.8mmol, 1.3eq). This mixture was dissolved in 100 mL anhydrous dimethylformamide and then cooled to 0°C. Sodium hydride (60% in mineral oil dispersion) (0.84 g, 20.8 mmol, 1.3 eq) was then added and the reaction

mixture stirred under argon for 15 hours at 0ºC. Upon completion, the reaction was concentrated to 10 mL and extracted with 350 mL dichloromethane. The organic layer was washed twice with cold sat. sodium bicarbonate, and once with brine, dried over magnesium sulfate, and volatiles were removed under reduced pressure to afford 12.40g of a crude yellow solid. Material purified via flash chromatography (8:2→1:1 Hex:EtOAc) to afford compound **5** as a yellow oil that solidifies to a white solid over time (5.4 g, 60% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ= 7.52 (s, 1H), 7.48 (d, *J*= 8 Hz, 6H), 7.32 (t, J= 7.5 Hz, 6H), 7.23 (t, *J* = 7.5 Hz, 3H), 7.01 (d, *J*= 8 Hz, 1H), 6.68 (s, 1H), 6.63 (d, *J*= 7.5 Hz, 1H), 6.57 (s, 1H), 4.26 (q, *J*= 7.5 Hz, 4H), 3.58 (s, 2H), 3.12 (s, 2H), 2.06 (s, 3H), 1.52 (s, 9H), 1.27 ppm (t, *J* = 7 Hz, 6H); ^{13}C NMR (126 MHz, $CDCl_3$) δ= 169.24, 167.41, 152.76, 144.23, 136.74, 135.59, 130.53, 129.32, 128.16, 126.85, 125.25, 123.91, 123.37, 80.20, 67.15, 62.66, 37.57, 33.89, 28.36, 22.95, 13.95 ppm; MS (ESI): calcd for $C_{41}H_{46}N_2O_7S$: 733.89 $[M+Na]^+$; found: 733.22.

Synthesis of 2-amino-3-(3-amino-4-(mercaptomethyl)phenyl)propanoic acid (AmmF): Compound **5** (5.4 g, 7.6 mmol, 1 eq) was added to a dry, argon-filled 250mL round-bottom flask and dissolved in 70 mL anhydrous dichloromethane. To the solution triisopropylsilane (3.88 mL, 19 mmol, 2.5 eq) was added and the reaction mixture was cooled in a ice bath. Trifloroacetic acid (18mL) was slowly added via a syringe and the reaction was left stirring under argon at 0ºC for 30 minutes. The reaction was then warmed to room temperature and volatiles were removed under reduced pressure. To this product was added 30 mL 8 M hydrochloric acid in dioxane and the reaction mixture was brought to reflux via heating mantle and allowed to stir under argon for 9 hours. Upon completion, the reaction was dried under reduced pressure to yield a brown solid, which was washed exaustively with cold hexanes to remove triphenylmethane and triisopropylsilane, yielding the dihydrochloride salt of AmmF (2.27 g , 100%). ^{1}H NMR (500 MHz, D_2O) δ 7.54 (d, *J*= 10 Hz, 1H), 7.41-7.30 (m, 2H), 4.31 (t, *J*= 5 Hz, 1H), 3.88 (s, 2H), 3.36-3.27 ppm (m, 2H); ^{13}C NMR (126 MHz, D_2O) δ= 171.57, 135.46, 134.41, 131.42, 130.60, 128.73, 124.78, 54.29, 35.01, 23.12 ppm; MS (ESI): calcd for $C_{10}H_{14}N_2O_2S$: 227.08 $[M+H]^+$; found: 226.98.



Scheme S2: Synthesis of MeaF

Synthesis of diethyl 2-acetamido-2-(3-nitrobenzyl)malonate (7): NaH (60% dispersion in mineral oil) (1.11 g, 27.7 mmol, 1.2eq) was added to a dry, argon-filled round bottom flask and dissolved in 150 mL anhydrous DMF. The flask was cooled to 0° C and to the solution was added diethyl acetomidomalonate (5.53 g, 25.5 mmol, 1.1 eq). After five minutes, 1-(bromomethyl)-3-nitrobenzene **(6)** (5 g, 23.14 mmol, 1 eq) was added. The reaction proceeded under argon at 0° C for 16 hours. The product was separated by extraction with 150 mL ethyl acetate and washed twice with dH_2O (150 mL), once with saturated $NaHCO_3$ (150 mL) and once with brine (150 mL). The organic layer was dried over $MgSO_4$ and volatiles were removed by rotary evaporation to yield **7** as a white solid (8 g, 98% yield). ^{1}H NMR (400MHz, $CDCl_3$) δ = 8.12 (d, *J*= 8 Hz, 1H), 7.90 (s, 1H), 7.46, (t, *J*= 7.6 Hz, 1H), 7.36 (d, *J*= 7.2 Hz, 1H), 6.61 (s, 1H), 4.27 (q, *J*= 8 Hz, 4H), 3.78 (s, 2H), 2.39 (s, 3H), 1.48 ppm (t, *J*= 6.8 Hz, 6H); MS (ESI): calcd for $C_{16}H_{20}N_2O_7$: 375.34 $[M+Na]^+$; found: 375.85.

Synthesis of diethyl 2-acetamido-2-(3-aminobenzyl)malonate (8): A dry 500 mL flask was purged with argon and compound **7** (5 g , 14.2 mmol, 1 eq) was added. To this, 1 g Pd/C was added and the flask was capped with a septa. 115 mL degassed methanol was added via syringe while purging with argon. The reaction mixture was sparged with argon for 10min clearing the needle with methanol as needed. Once purged, the reaction was sparged with H_2, again clearing the needle as needed. Reaction continued under H_2 for 3 h at r.t. The flask was sparged with

argon to clear out the residual hydrogen gas, and the flask was washed with ethyl acetate and filtered over celite. Deionized H_2O was used to quench Pd/C and the solvent was removed by rotary evaporation to yield a off-white solid (4.45 g, 97%). ^{1}H NMR (500MHz, $CDCl_3$) δ = 8.12 (d, *J*=8.5 Hz, 1H), 7.90 (s, 1H), 7.45 (t, *J*=7.5 Hz, 1H), 7.36 (d, *J*=7.5 Hz, 1H), 6.55 (s, 1H), 4.29 (q, *J*=7 Hz, 4H), 3.78 (s, 2H), 2.06 (s, 1H), 1.32 ppm (t, *J*=7 Hz, 6H); ^{13}C NMR (126MHz, $CDCl_3$) δ = 169.42, 167.07, 148.17, 137.45, 136.14, 129.24, 124.59, 122.35, 66.98, 63.06, 37.35, 23.00, 14.00 ppm; MS (ESI): calcd for $C_{16}H_{22}N_2O_5$: 345.36 $[M+Na]^+$; found: 345.19.

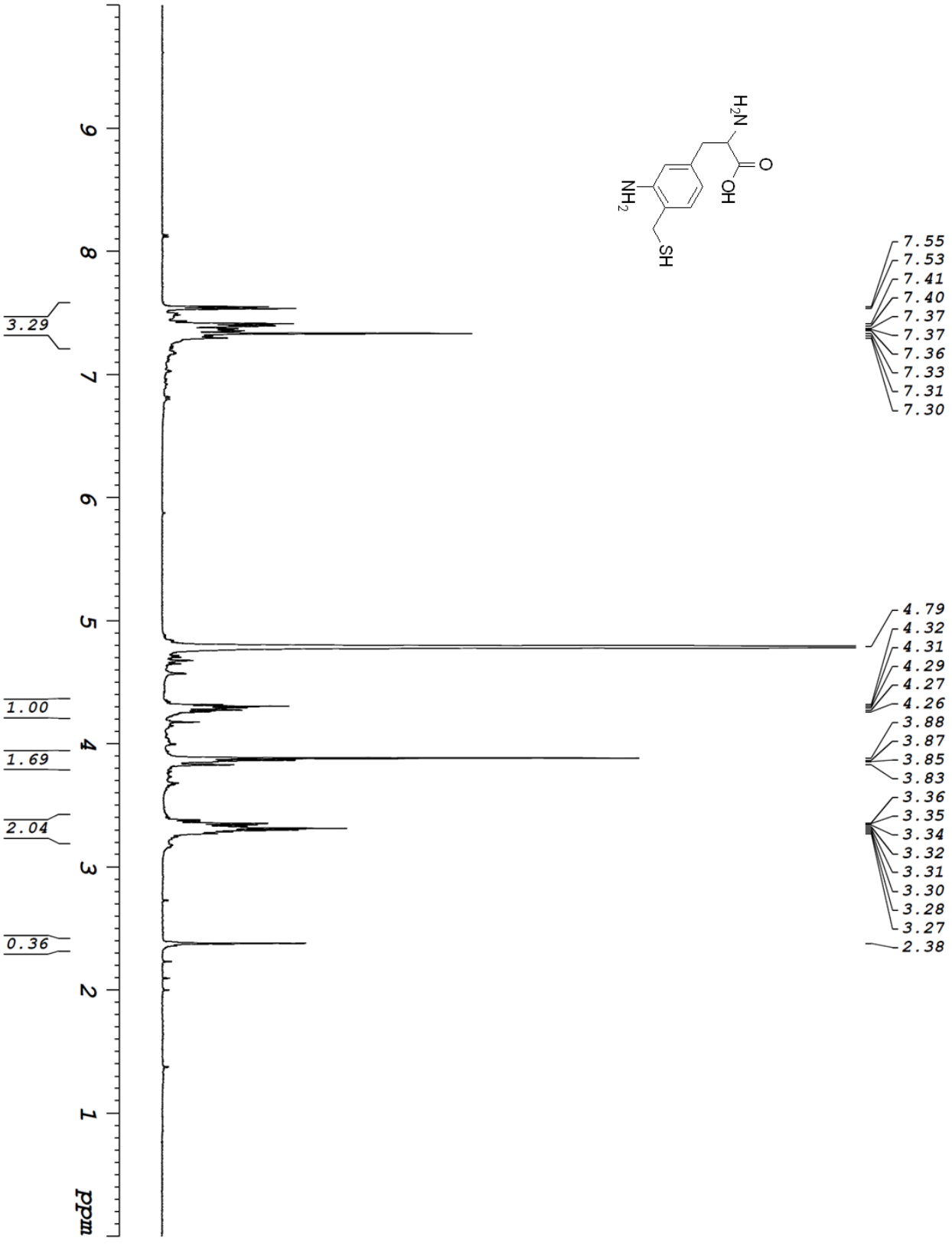
Synthesis of diethyl 2-acetamido-2-(3-((2-chloroethyl)amino)benzyl)malonate (9): Compound **8** (4.5 g, 14 mmol, 1eq) was added to a round bottom flask and dissolved in 90 mL ethanol. $NaCNBH_3$ (0.97 g, 15.4 mmol, 1.1 eq) was added and the mixture was sonicated to aid solubility. Chloroacetaldehyde (2.7 mL, 15.4 mmol, 1.1 eq) was added followed by glacial acetic acid (0.81 mL, 14 mmol, 1 eq). The reaction was run at r.t. under argon for 4 h then the solvent was removed by rotary evaporation. The crude product was re-suspended in DCM and washed with H_2O, then with sat. $NaHCO_3$ and then with brine. The resulting organic layer was dried over $MgSO_4$ and the solvent was removed by rotary evaporation. The oil was purified by flash chromatography on silica gel using a solvent gradient of 8:1 to 7:3 hexanes:ethyl acetate to yield compound **9** as a yellow oil (3.46 g, 9 mmol, 64.2% yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ = 6.91 (t, *J*=8 Hz, 1H), 6.42 (d, *J*=8 Hz, 1H), 6.21 (d, *J*=7.6 Hz, 1H), 6.19 (s, 1H), 4.18 - 4.06 (m,4H), 3.52 (t, *J*= 6.4 Hz, 2H), 3.41 - 3.29 (m, 4H), 1.91 (s, 3H), 1.16 ppm (t, *J*=7.2 Hz, 6H); MS (ESI): calcd for $C_{18}H_{25}ClN_2O_5$: 407.86 $[M+Na]^+$; found: 407.33.

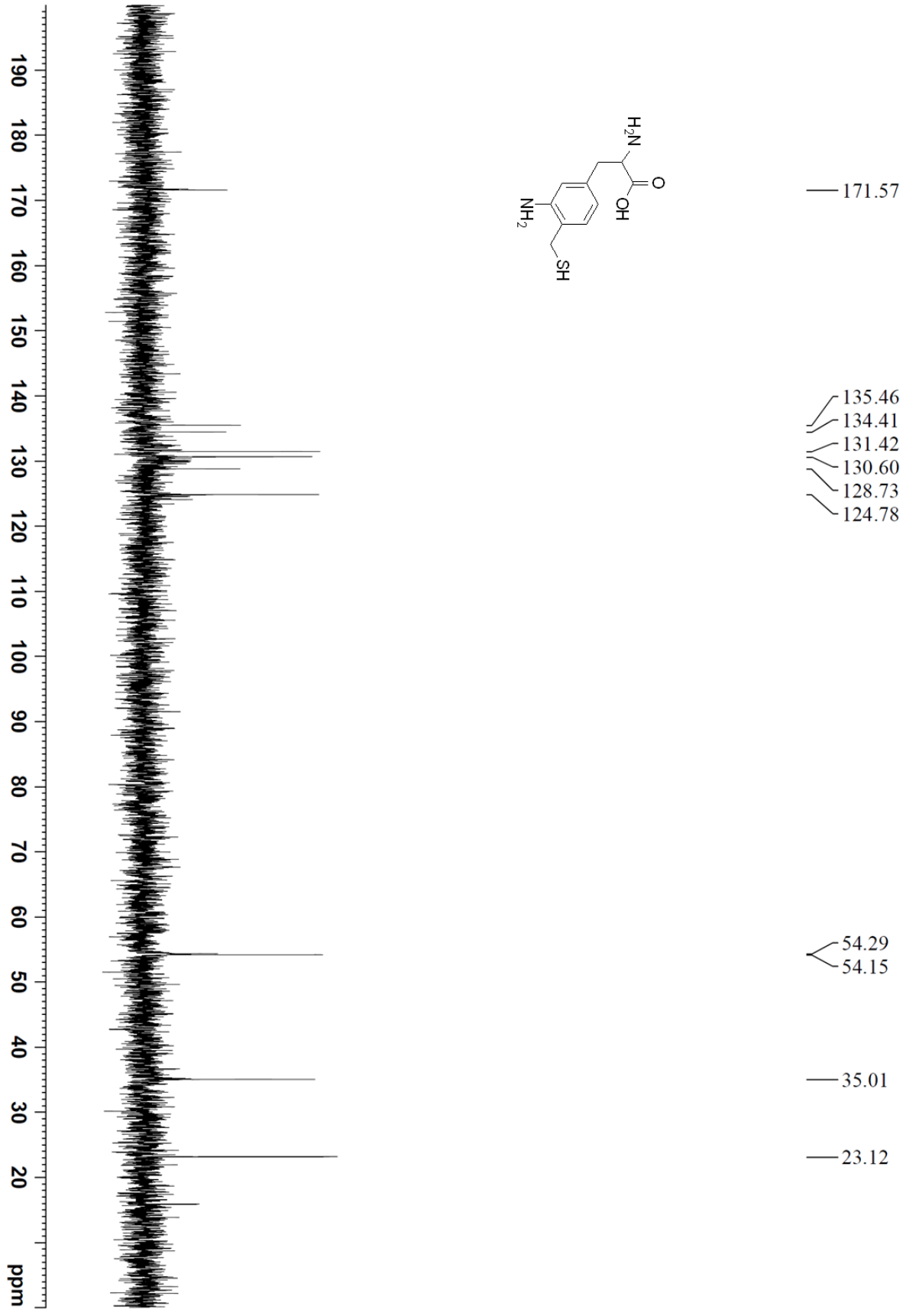
Synthesis of diethyl 2-acetamido-2-(3-((2-(tritylthio)ethyl)amino)benzyl)malonate (10): Compound **9** (3.46 g, 9 mmol, 1 eq) was dissolved in 85 mL anhydrous DMF. To this was added triphenyl methylmercaptan (2.73 g, 9.9 mmol, 1.1 eq) followed by potassium carbonate (1.5 g, 10.8 mmol, 1.2 eq). The reaction was brought to 50ºC in an oil bath under argon and left stirring for 12 h. The reaction was dilute in EtOAc and washed twice with dH_2O, once with sat. $NaHCO_3$ and then with brine. The organic layer was dried over $MgSO_4$ and solvent removed by rotary evaporation. The crude product purified on silica gel by flash chromatography with a of 8:1→7:3 Hex:EtOAc gradient to yield yellow oil (3.38 g, 5.4 mmol, 60% yield). ^{1}H NMR (500MHz, $CDCl_3$) δ = 7.39 (d, *J*=7.5 Hz, 6H), 7.26 (t, *J*=7.2 Hz, 3H), 7.22 (d, *J*=7.1 Hz, 6H), 7.19 (s, 1H), 7.00 (t, *J*=7.8 Hz, 1H), 6.51 (s, 1H), 6.32 (d, *J*=7.1 Hz, 2H), 6.13 (s, 1H), 4.25 (p, *J*=7.2 Hz, 4H),

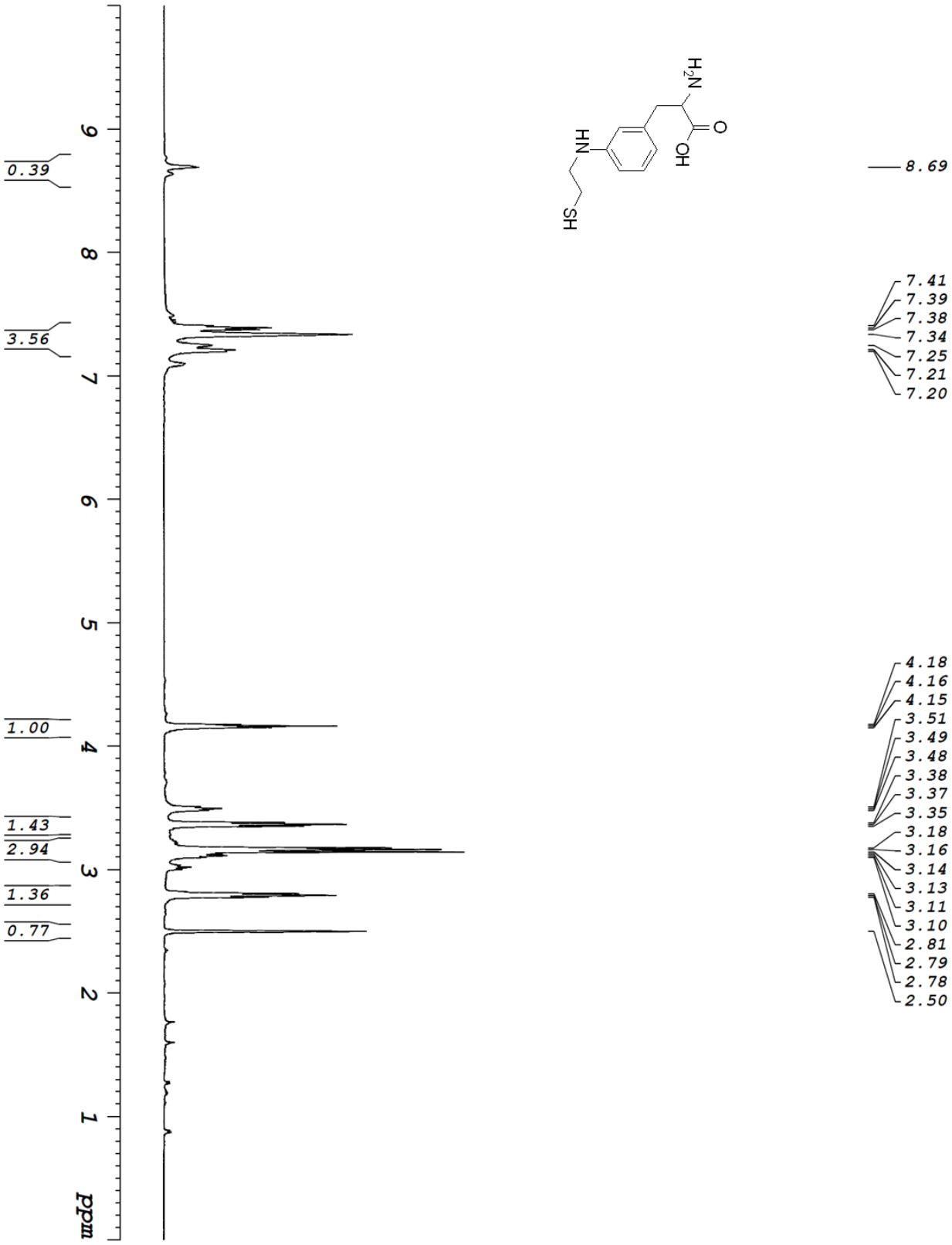
3.53 (s, 2H), 2.98 (t, *J*=6.6 Hz, 2H), 2.47 (t, *J*=6.5 Hz, 2H), 1.96 (s, 3H), 1.27 ppm (t, *J*=7.1 Hz, 6H); ^{13}C NMR (126MHz, $CDCl_3$) δ = 168.89, 167.53, 147.45, 144.65, 136.22, 129.51, 129.07, 127.89, 127.85, 126.72, 118.92, 114.40, 111.85, 77.25, 77.20, 77.00, 76.74, 67.14, 66.75, 62.51, 53.39, 42.42, 37.94, 31.62, 23.02, 13.99 ppm; MS (ESI): calcd for $C_{37}H_{40}N_2O_5S$: 647.80 $[M+Na]^+$; found: 647.91.

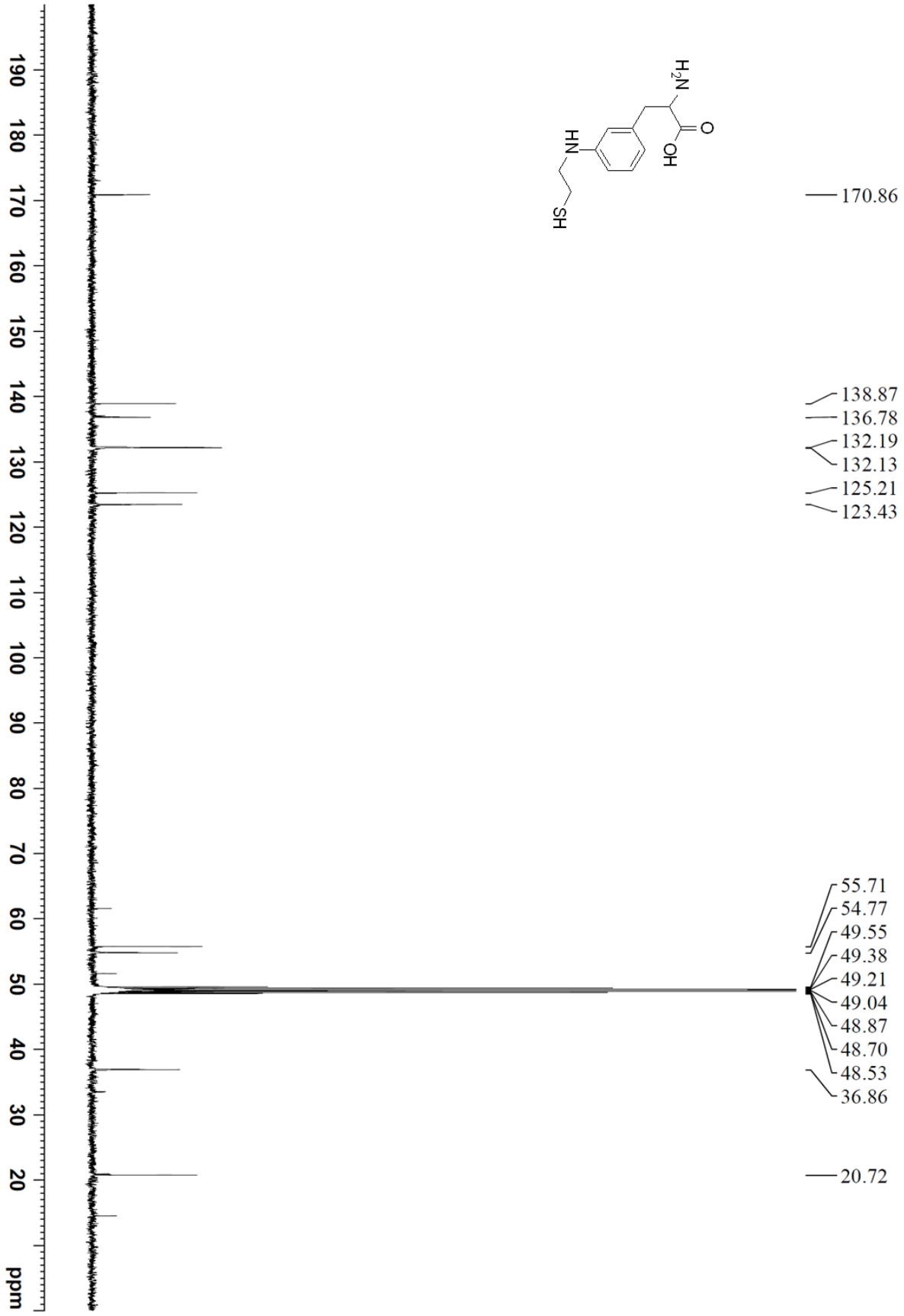
Synthesis of 2-amino-3-(3-((2-mercaptoethyl)amino)phenyl)propanoic acid (MeaF): Compound **5** (3.38 g, 5.4 mmol, 1 eq) was dissolved in 12 mL anhydrous DCM and the reaction mixture was purged with argon and cooled in a ice bath. The solution was added with triisopropylsilane (1.23 mL, 6.1 mmol, 1.1eq) and then with 8 mL TFA dropwise via a syringe. The reaction was stirred for 20 minutes at 0^o C, then warmed up to room temperature, and let reach completion. Volatiles were removed by rotary evaporation and dried under high vacuum overnight. The dried product was washed several times with cold hexanes then dissolved in 35 mL of 4N HCl (aq) and heated at reflux overnight. The solvent was removed by rotary evaporation to yield a orange-brown solid. This solid was washed several times in 20% diethyl ether in hexanes. The purified product was placed under high-vacuum to remove additional solvent and to yield a light-brown solid MeaF as a mixture of monomer and dimer. (1.5 g, 5.4 mmol, 99.9%). ^{1}H NMR (500 MHz, D_6-DMSO) δ= 8.69 (s, 1H), 7.41-7.2 (m, *3H*), 4.16 (t, *J*= 10 Hz, 1H), 3.37 (t, *J*= 5 Hz, 2H), 3.18-3.10 (m, 2H), 2.79 (t, *J*= 10 Hz, 2H), 2.50 (s, 1H); ^{13}C NMR (126 MHz, MeOD) δ= 170.86, 138.87, 136.78, 132.19, 132.13, 125.21, 123.43, 55.71, 54.77, 36.86, 20.72 ppm; MS (ESI): calcd for $C_{11}H_{16}N_2O_2S$: 241.32 $[M+H]^+$; found: 241.61.

Small molecule study for pyruvate modification. To demonstrate the reactivity of the AmmF amino-thiol side chain with pyruvate under physiological conditions, methyl 3-amino-4-(mercaptomethyl)benzoate (4 mM) was incubated with pyruvate (5 mM) and TCEP (15 mM) in phosphate saline buffer (50 mM phosphate, 150 mM NaCl, pH 7) at 23°C. Complete conversion of methyl 3-amino-4-(mercaptomethyl)benzoate to a single product occurred after 16 h as determined by HPLC. LC-MS analysis confirmed the identity of the product as 7-(methoxycarbonyl)-2-methyl-1,4-dihydro-2H-benzo[d][1,3]thiazine-2-carboxylic acid (**Figure S2**).


7.55
7.53
7.41
7.40
7.37
7.37
7.36
7.33
7.31
7.30
4.79
4.32
4.31
4.29
4.27
4.26
3.88
3.87
3.85
3.83
3.36
3.35
3.34
3.32
3.31
3.30
3.28
3.27
2.38
3.29
1.00
1.69
2.04
0.36
ppm


190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
ppm
H2N
O
OH
NH2
SH
171.57
135.46
134.41
131.42
130.60
128.73
124.78
54.29
54.15
35.01
23.12

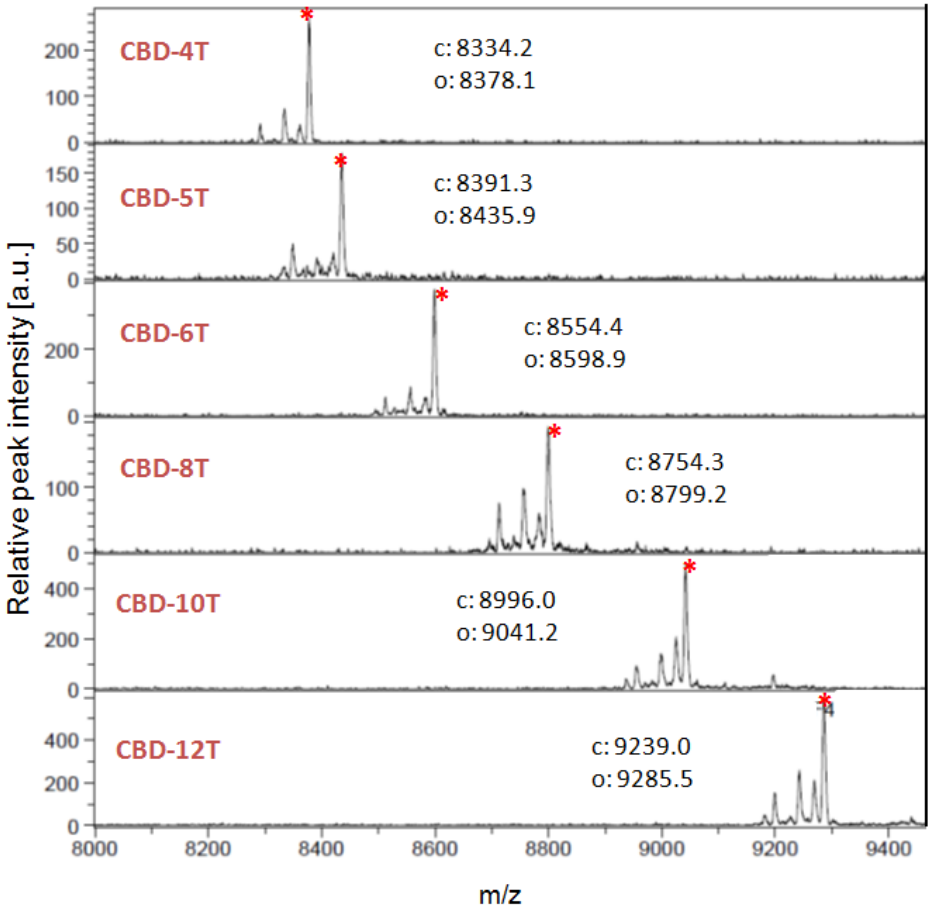

0.39
3.56
1.00
1.43
2.94
1.36
0.77
9
8
7
6
5
4
3
2
1
ppm
8.69
7.41
7.39
7.38
7.34
7.25
7.21
7.20
4.18
4.16
4.15
3.51
3.49
3.48
3.38
3.37
3.35
3.18
3.16
3.14
3.13
3.11
3.10
2.81
2.79
2.78
2.50


170.86
138.87
136.78
132.19
132.13
125.21
123.43
55.71
54.77
49.55
49.38
49.21
49.04
48.87
48.70
48.53
36.86
20.72
ppm

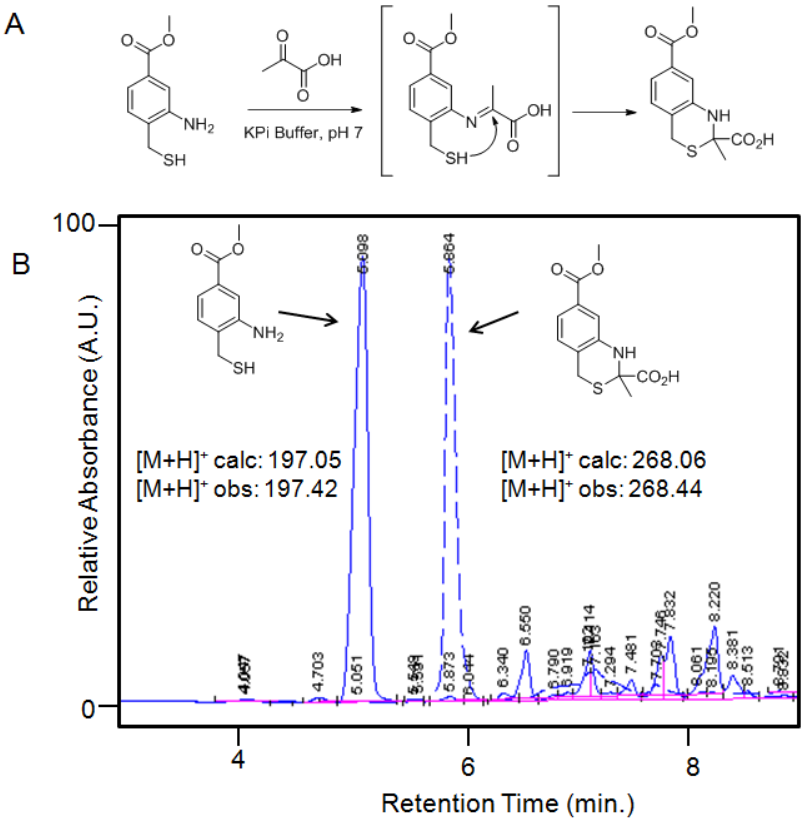
Supplementary Table S1. Sequences of oligonucleotide primers.

Primer	**Sequence**
FLAG-HPQ-GyrA_for1/3	5'- AGCGGCTAGTTCACCAACGTTCACCCGCAG TTCGCAAACGCGTGCATCACGGGAGATGCA -3'
FLAG-HPQ-GyrA_for2/3	5'- GACGATGACGATAAAGGCAGCAGCGGCTAG TTCACCAACGTTC -3'
FLAG-HPQ-GyrA_for3/3	5'- CACTGCATATGGATTACAAAGACGATGACG ATAAAGGCAGCAG -3'
T7_Term_long_rev	5'- GCTAGTTATTGCTCAGCGGTGGC -3'
HPQ_SOE(A)_for	5'- CAACGCCAATCAGCAACGAC -3'
HPQ_SOE(A)_rev	5'- GTTGGTGAACTAGCCGCTGC -3'
HPQ(NDC)_SOE(B)_for	5'- GCAGCGGCTAGTTCACCAACNDCCACCCGC AGNDCGCAAACG -3'
HPQ(NDC)_SOE(B)_rev	5'-GCTAGTTATTGCTCAGCGGTGGCAGCAGCC AACTC -3'

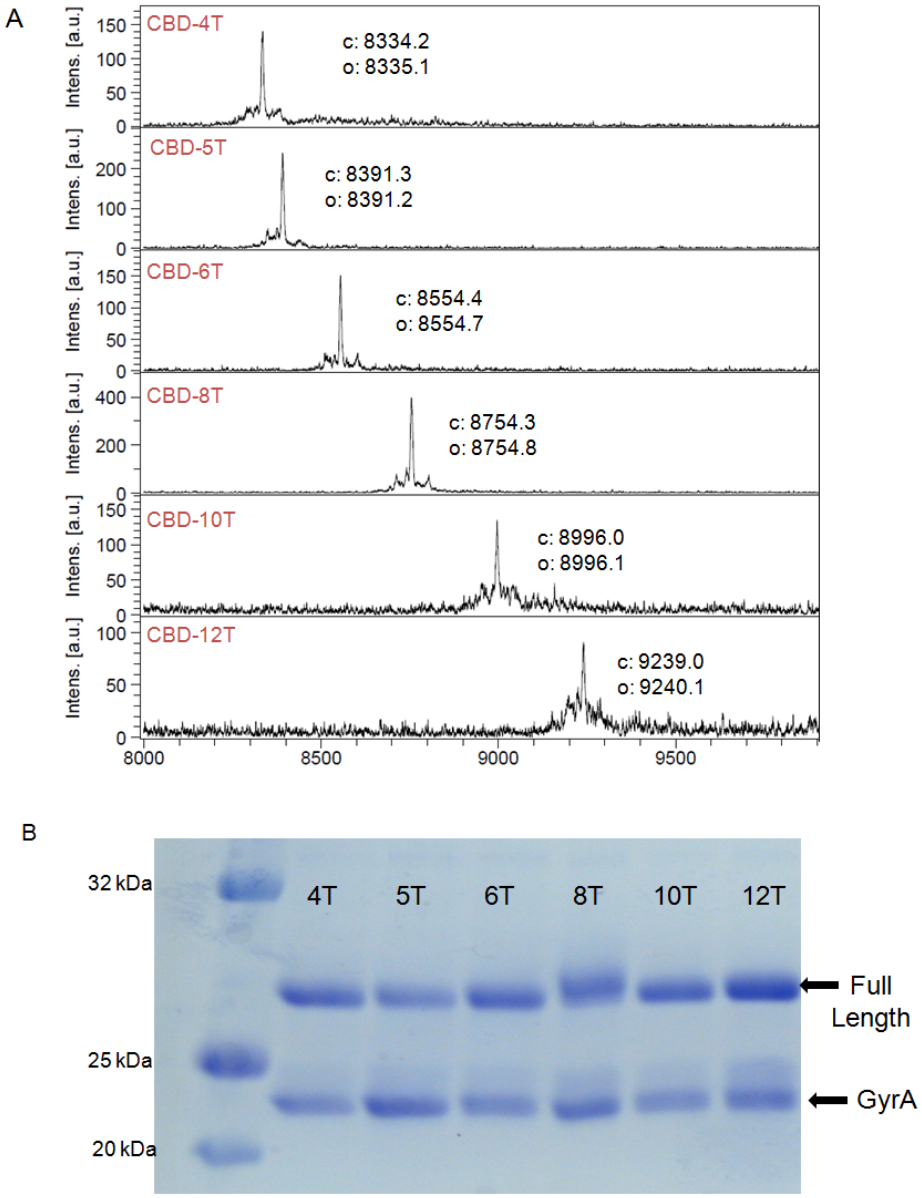
Supplementary Figure S1. MALDI-TOF MS spectra of the purified AmmF-containing constructs CBD-4T(AmmF) through CBD-12T(AmmF) after reaction with thiophenol. The calculated ("c") *m/z* value corresponding to the expected $[M+H]^+$ adduct of the macrocylic product is indicated. In each case, the major product (labeled with a star) corresponds to a +26 Da adduct of the linear, hydrolyzed product.



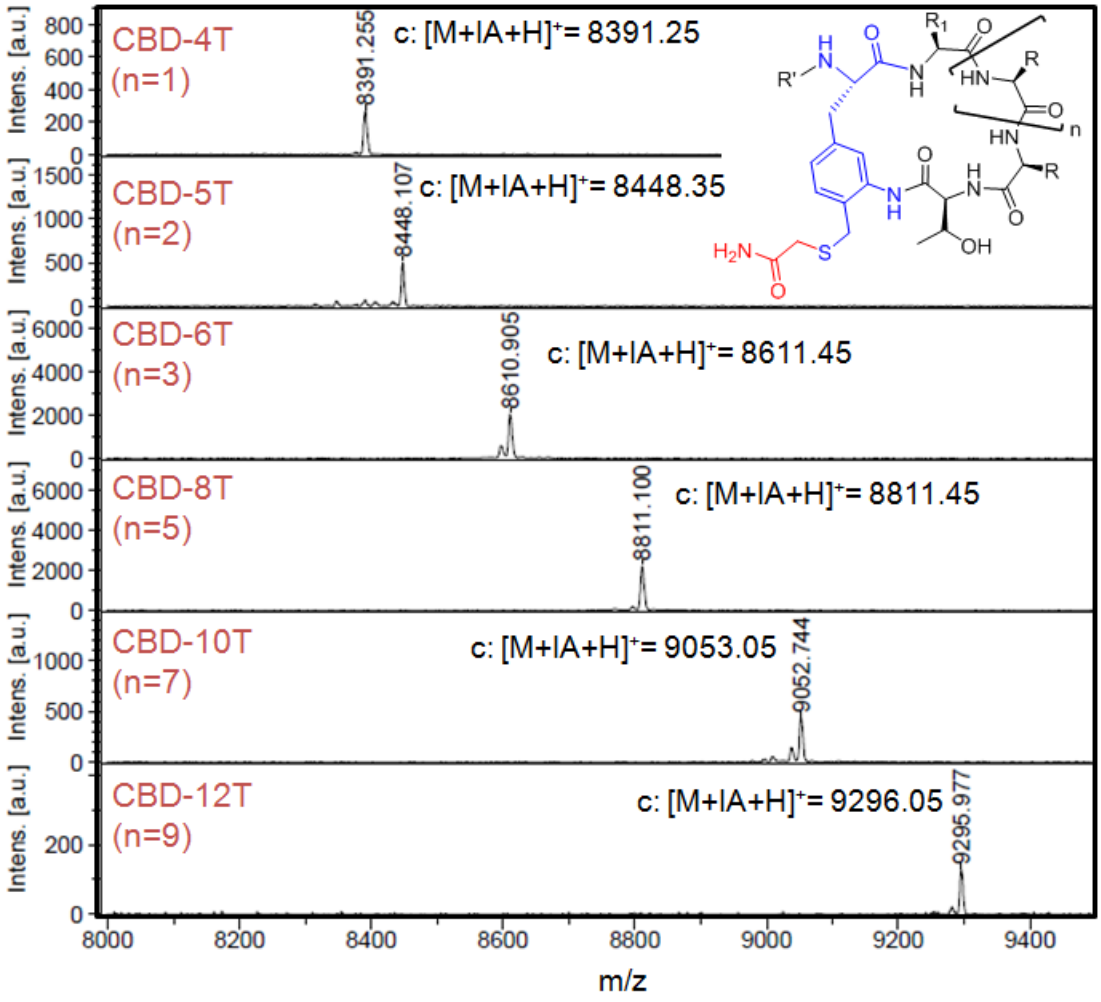
Supplementary Figure S2. Modification of *ortho*-amino-mercaptomethyl-aryl moiety with pyruvate. (A) Scheme of the reaction between the AmmF surrogate, methyl 3-amino-4-(mercaptomethyl)benzoate, with pyruvate to give the benzothiazine adduct. (B) Overlay of HPLC traces corresponding to 3-amino-4-(mercaptomethyl)benzoate before (solid line) and after the reaction with pyruvate (dashed line). Calculated ("calc") and observed ("obs") *m/z* values corresponding to the $[M+H]^+$ adducts of the two species are indicated.



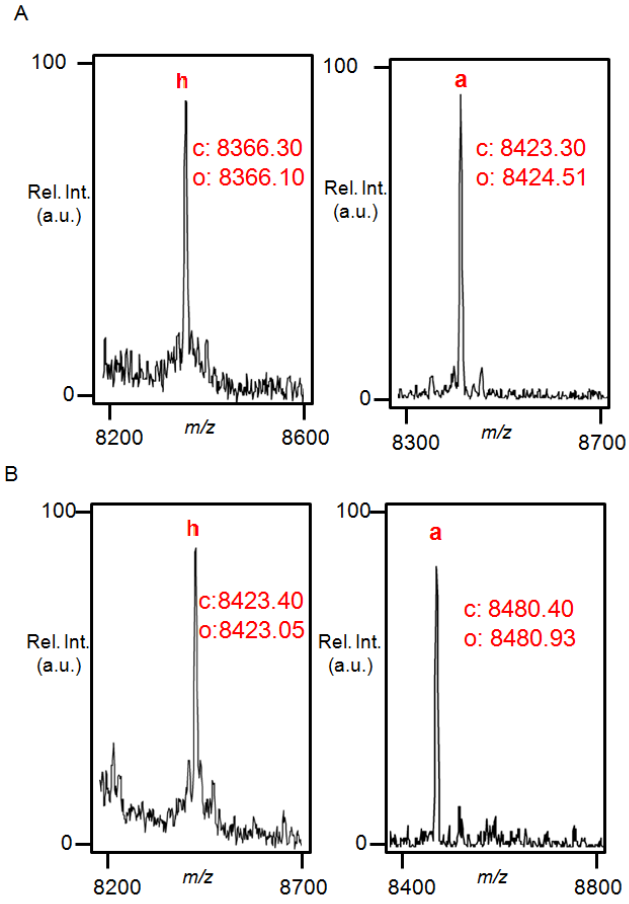
Supplementary Figure S3. In vitro cyclization of AmmF-containing precursor proteins. (A) MALDI-TOF MS spectra of the CBD-fused macrocyclic peptides generated upon incubation of the proteins at pH 5.0 in the presence of MESNA (10 mM). The calculated ("c") and observed ("o") *m/z* values corresponding to the $[M+H]^+$ adduct of the macrocylic product are indicated. (B) SDS-PAGE analysis of these reactions after 24 hours. Bands corresponding to the full-length precursor protein ("full-length") and cleaved intein ("GyrA") are labeled.



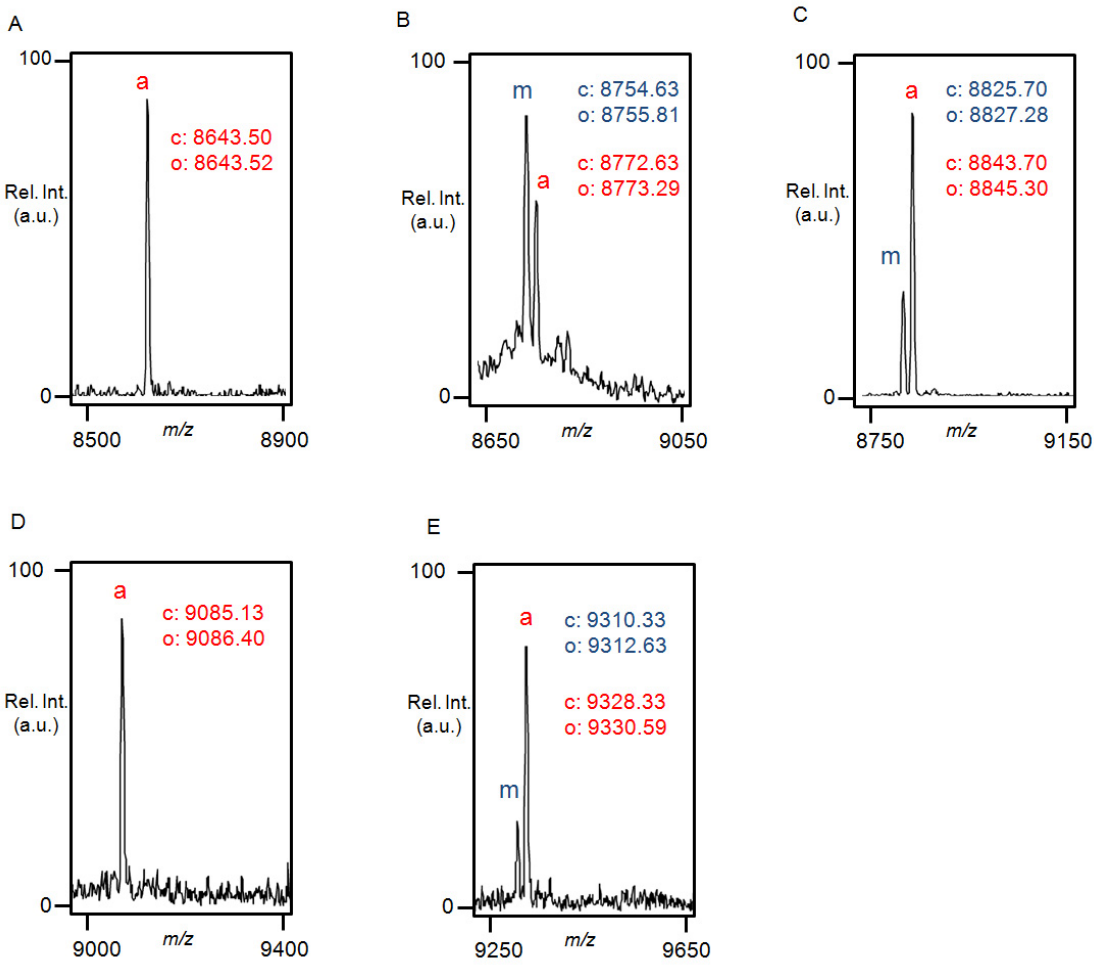
Supplementary Figure S4. MALDI-TOF MS spectra of the macrocyclic S-carboxyamidomethyl adducts generated from the catalyst-free (i.e. thiol-free) cyclization of the purified AmmF-containing constructs, followed by alkylation with iodoacetamide. The calculated ("c") *m/z* values corresponding to the $[M+H]^+$ adduct of the of the macrocyclic products are indicated. R' = chitin binding domain.



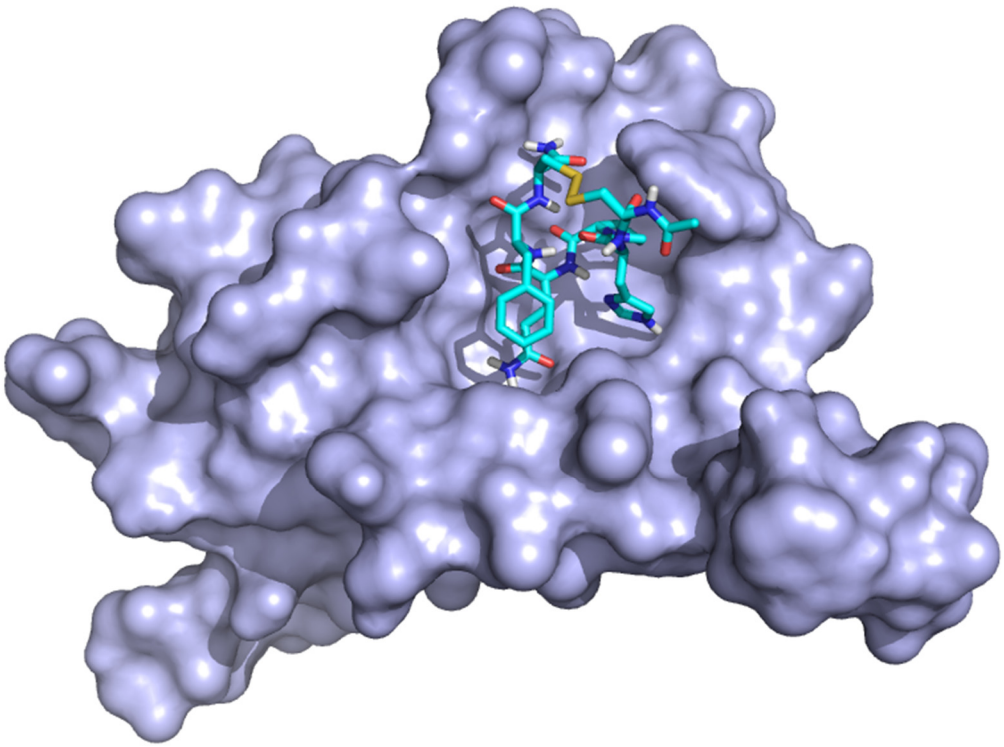
Supplementary Figure S5. MALDI-TOF spectra of the products obtained from the MeaF-containing constructs CBD-4T(MeaF) (A) and CBD-5T(MeaF) (B) as isolated from *E. coli* cells after 12 hours expression using pH 5.0 buffer (*left panel*) and pH 8.0 buffer containing iodoacetamide (*right panel*). "h" = hydrolyzed product, "a" = acyclic side product (S-carboxyamidomethyl adduct). The calculated ("c") and observed ("o") *m/z* values corresponding to the $[M+H]^+$ adduct are indicated.



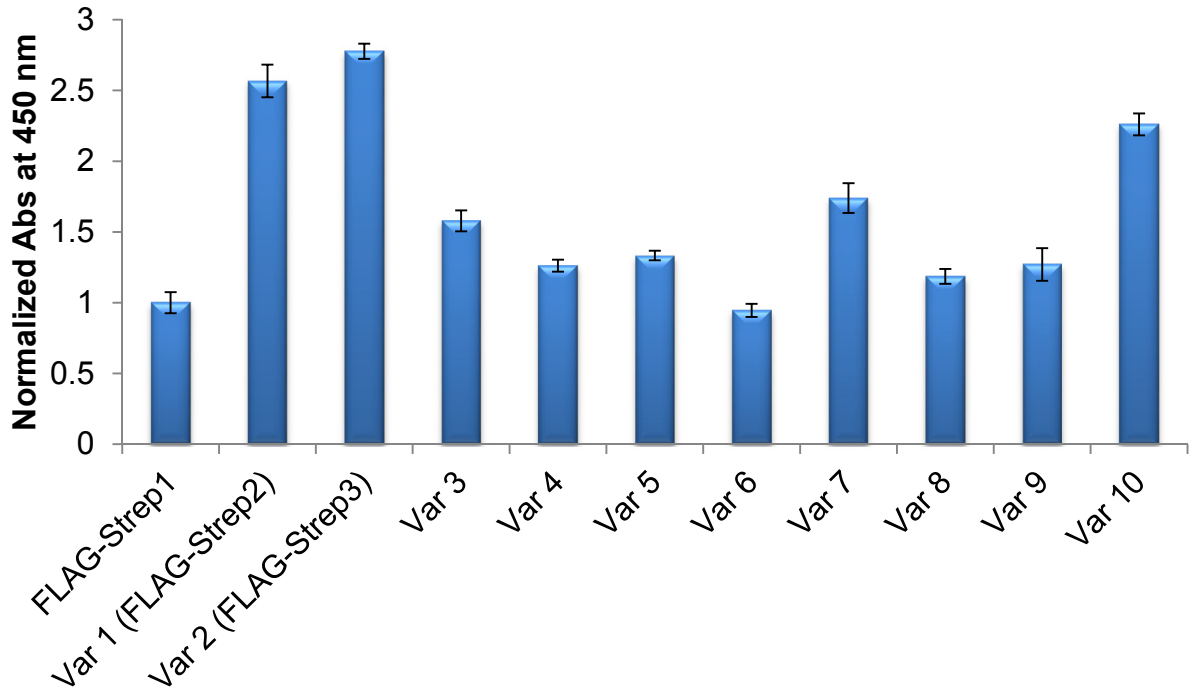
Supplementary Figure S6. MALDI-TOF spectra of the products obtained from the MeaF-containing constructs as isolated from *E. coli* cells after 12-hour expression using pH 8.0 buffer containing iodoacetamide. (A) CBD-6T(MeaF); (B) CBD-7T(MeaF); (C) CBD-8T(MeaF); (D) CBD-10T(MeaF); and (E) CBD-12T(MeaF). "m" = macrolactam (S-carboxyamidomethyl adduct), "a" = acyclic side product (S-carboxyamidomethyl adduct). The calculated ("c") and observed ("o") *m/z* values corresponding to the $[M+H]^+$ adducts are indicated and color coded.



Supplementary Figure S7. Crystal structure of streptavidin (blue, surface model) in complex with a disulfide bridged peptide, *cyclo*-(CHPQFC)-NH_2 (cyan, stick model), isolated by phage display (pdb 1SLD; Katz, B. A. *Biochemistry* 1995, 34, 15421). The α carbon atoms of Cys1 and Phe5 in the peptide are within 6 Å from the closest amino acid residue on the protein surface.



Supplementary Figure S8: Results from the re-screening of the 10 most active variants isolated from primary screening of the Flag-Strep1(V18NDC/F22NDC) library. Data were obtained using the streptavidin binding assay described in the text and in **Figure 8A**. The absorbance values are normalized to that of the reference cyclic peptide Flag-*cyclo*(Strep1). Mean and standard deviation values are calculated from samples analyzed in triplicate.



Supplementary Figure S9: MALDI-TOF spectra of the macrocyclic peptides Flag-*cyclo*(Strep1) (A), Flag-*cyclo*(Strep2) (B), and Flag-*cyclo*(Strep3) (C), as obtained after *in vitro* cyclization of precursor protein Flag-Strep1(MeaF), Flag-Strep2(MeaF), and Flag-Strep3(MeaF), respectively. The calculated ("c") and observed ("o") *m/z* values corresponding to the $[M+H]^+$ and $[M+Na]^+$ adducts are indicated.