JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

A Helping Hand to Overcome Solubility Challenges in Chemical Protein Synthesis

Michael T. Jacobsen,[†] Mark E. Petersen,[†] Xiang Ye,[‡] Mathieu Galibert,[§] George H. Lorimer,[‡] Vincent Aucagne,*[,§] and Michael S. Kay*[,†]

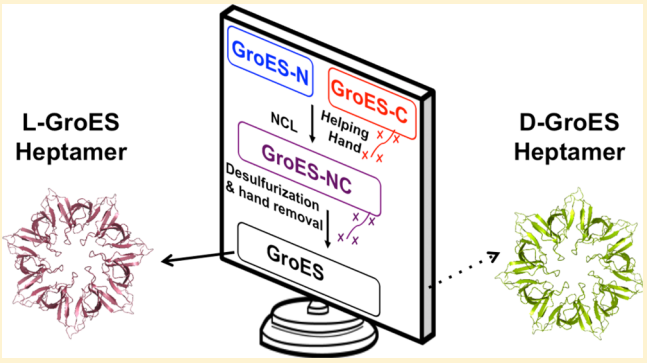
[†]Department of Biochemistry, University of Utah School of Medicine, 15 North Medical Drive East, Room 4100, Salt Lake City, Utah 84112-5650, United States

[‡]Department of Chemistry & Biochemistry, 8051 Regents Drive, University of Maryland, College Park, Maryland 20742-4454, United States

[§]Centre de Biophysique Moléculaire, CNRS UPR4301, Rue Charles Sadron, Orléans CEDEX 2 45071, France

Supporting Information

ABSTRACT: Although native chemical ligation (NCL) and related chemoselective ligation approaches provide an elegant method to stitch together unprotected peptides, the handling and purification of insoluble and aggregation-prone peptides and assembly intermediates create a bottleneck to routinely preparing large proteins by completely synthetic means. In this work, we introduce a new general tool, Fmoc-Ddae-OH, *N*-Fmoc-1-(4,4-dimethyl-2,6-dioxocyclo-hexylidene)-3-[2-(2-aminoethoxy)ethoxy]-propan-1-ol, a heterobifunctional traceless linker for temporarily attaching highly solubilizing peptide sequences ("helping hands") onto insoluble peptides. This tool is implemented in three simple and nearly quantitative steps: (i) on-resin incorporation of the linker at a Lys residue ε-amine, (ii) Fmoc-SPPS elongation of a desired solubilizing sequence, and (iii) in-solution removal of the solubilizing sequence using mild aqueous hydrazine to cleave the Ddae linker after NCL-based assembly. Successful introduction and removal of a Lys_6 helping hand is first demonstrated in two model systems (Ebola virus C20 peptide and the 70-residue ribosomal protein L31). It is then applied to the challenging chemical synthesis of the 97-residue co-chaperonin GroES, which contains a highly insoluble C-terminal segment that is rescued by a helping hand. Importantly, the Ddae linker can be cleaved in one pot following NCL or desulfurization. The purity, structure, and chaperone activity of synthetic L-GroES were validated with respect to a recombinant control. Additionally, the helping hand enabled synthesis of D-GroES, which was inactive in a heterochiral mixture with recombinant GroEL, providing additional insight into chaperone specificity. Ultimately, this simple, robust, and easy-to-use tool is expected to be broadly applicable for the synthesis of challenging peptides and proteins.



INTRODUCTION

The development of native chemical ligation (NCL)[1,2] by Kent's group was a seminal discovery for the total chemical synthesis of peptides and proteins. Using NCL, peptides containing N-terminal Cys and C-terminal thioesters can be chemoselectively ligated to generate larger synthetic products.[2,3] This synthetic strategy, solid-phase peptide synthesis (SPPS) to prepare peptide segments followed by NCL or another ligation method[4−6] to stitch the peptides together, has been very successful for producing synthetic proteins to probe biological problems that would be difficult to assess using molecular biology. Notable achievements include innovative studies on ubiquitination pathways,[7−10] extraordinary efforts to prepare well-defined glycoforms of erythropoietin,[11−18] and total syntheses of mirror-image proteins not accessible by recombinant means,[19−21] among many other interesting targets (recently reviewed in refs 22 and 23). However, total chemical syntheses of these protein targets can be unexpectedly challenging, due to difficulties in handling (dissolving, purifying, and reacting) poorly soluble and aggregation-prone sequences. In particular, certain hydrophobic peptide segments, prominent in membrane proteins, remain especially challenging.[24−26] Furthermore, it would be highly beneficial to deploy a tool for routinely increasing peptide/protein concentrations to the millimolar range in order to generally accelerate chemoselective ligation reactions.[27,28]

To address this challenge, many groups have devised chemistries for temporarily improving peptide segment solubility. One of the most prominent examples is the thioester poly-Arg tag, introduced by the Kent[29] and Aimoto[30] groups. In this method, the peptide segment of interest is directly

Received: June 4, 2016
Published: August 17, 2016

ACS Publications © 2016 American Chemical Society

DOI: 10.1021/jacs.6b05719
J. Am. Chem. Soc. 2016, 138, 11775−11782

synthesized and purified with a highly solubilizing peptide sequence (Arg_n, where n = 5−9) introduced on the thioester leaving group.

This method has been used in several cases[31−33] to endow peptides with greatly improved solubility and handling properties. However, the solubilizing tag is eliminated during the ligation reaction (via *in situ* trans-thioesterification with a better thiol leaving group), so the solubility enhancement cannot be retained during or beyond the ligation step. Furthermore, these tags are difficult to install by Fmoc-SPPS due to the instability of thioesters to piperidine-mediated Fmoc deprotection conditions,[34] although an indirect approach using *ortho*-aminoanilide cryptothioesters has been described.[35] A less commonly used route for introducing semipermanent C-terminal solubilizing tags entails C-terminal base-labile linkers (e.g., esters), introduced either by Boc[36] (glycolic acid linker) or Fmoc-SPPS[37−39] (4-hydroxymethylbenzoic acid linker). An approach using the acid-labile phenylacetamido (PAM) linker has also been described.[40] Alternatively, linkers containing substrate sequences cleaved by HIV protease[32] and carboxypeptidase[41] have also been used. Thus far, these base-, acid-, and enzyme-sensitive linkers have been limited to peptide termini.

Some versatile approaches to this problem have recently been developed. Danishefsky's group designed an allyl-based protection method for introducing semipermanent solubilizing groups by preparing custom Fmoc-Glu and Fmoc-Lys building blocks linked to solubilizing guanidinium moieties through ester and carbamate groups, respectively.[42] Liu's group developed a photolabile linker for incorporating semipermanent solubilizing sequences at Fmoc-Gln building blocks.[43] The same group also introduced an elegant two-step method for selectively removing semipermanent solubilizing groups at Fmoc-Gly building blocks.[44] Here, a TFA-stable derivative of the *N*-(2-hydroxy-4-methoxy benzyl) (Hmb) group[45] is converted *in situ* into a TFA-labile Hmb during the NCL reaction.[46,47] Building on this work, Liu's group recently published a method to introduce removable backbone modifications (RBMs) via a salicylaldehyde derivative.[48]

A New Helping Hand Approach. Unfortunately, no single method for introducing tailor-made semipermanent solubilizing sequences has been widely accepted in the field for multiple reasons, including complex reagent synthesis, potentially damaging cleavage conditions, slow kinetics of incorporation or elimination, limited placement within peptide sequences, or lability to NCL conditions. We desire a general chemical tool for routinely introducing semipermanent solubilizing sequences ("helping hands") onto difficult peptides. Accordingly, we had five requirements:

(1) Compatibility and facile incorporation by Fmoc-SPPS
(2) Compatibility with NCL and other chemical conditions typically employed in assembly of peptide segments (including desulfurization and oxidative hydrazide activation)
(3) Site-selective attachment of the designed solubilizing sequence
(4) Ability to be incorporated at several sites in nearly all peptide segments of the protein target (i.e., not limited to N- or C-termini)
(5) Simple and selective removal of the linker to generate native peptide after the solubilizing application is complete

In response to this challenge, we describe here synthesis and applications of Fmoc-Ddae-OH, *N*-Fmoc-1-(4,4-dimethyl-2,6-dioxocyclo-hexylidene)-3-[2-(2-aminoethoxy)ethoxy]-propan-1-ol, a new heterobifunctional traceless linker for attaching solubilizing groups to difficult peptide sequences (Figure 1a).

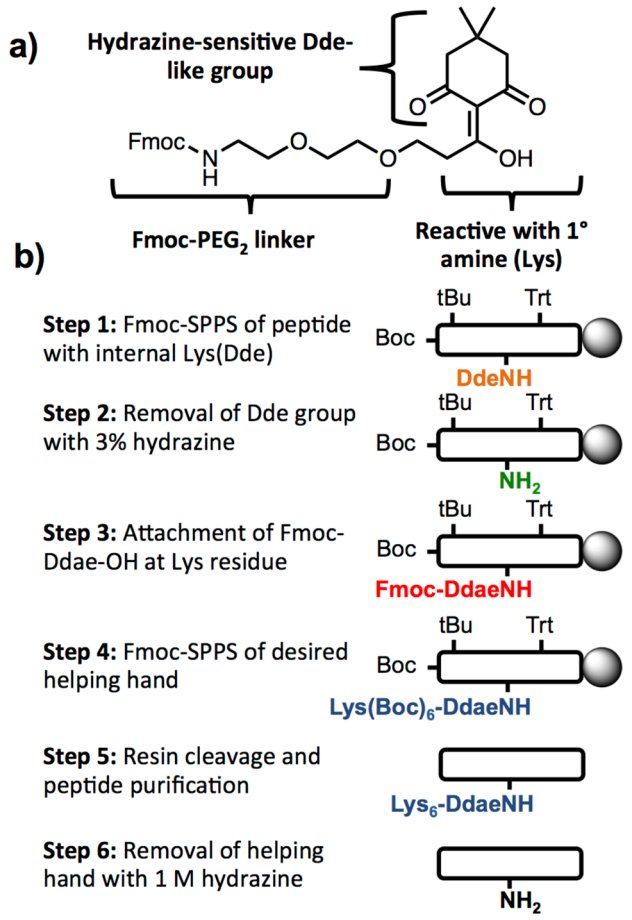


Figure 1. Fmoc-Ddae-OH linker. (a) Key properties of the new linker. (b) Steps for installing a temporary solubilizing sequence (helping hand) using the new linker.

With this simple-to-prepare tool, we can attach and later remove highly solubilizing peptide sequences on any peptide segment containing a Lys residue via on-resin Fmoc-SPPS elongation and aqueous hydrazine treatment, respectively. To validate the versatility and robustness of our tool, we applied it in three different synthetic contexts: a model 20-residue peptide and two proteins of 70 and 97 residues requiring NCL-based assembly. In particular, we prepared both L- and D- versions of the synthetically challenging heptameric chaperone protein GroES.

RESULTS AND DISCUSSION

Design and Synthesis of Fmoc-Ddae-OH Linker. The inspiration for our linker comes from Bycroft's pioneering work[49−51] on the (4,4-dimethyl-2,6-dioxocyclohexylidene)ethyl (Dde) amine protecting group. The enamine bond is selectively cleaved by treatment with bis-α-nucleophiles such as hydrazine or hydroxylamine[52] (as well as $NaBH_4$)[53] but is stable to Fmoc elongation and deprotection conditions. Due to its robustness and specific cleavage conditions, a number of Dde-based linkers have been developed for various applications, including SPPS,[54]

DOI: 10.1021/jacs.6b05719
J. Am. Chem. Soc. 2016, 138, 11775−11782

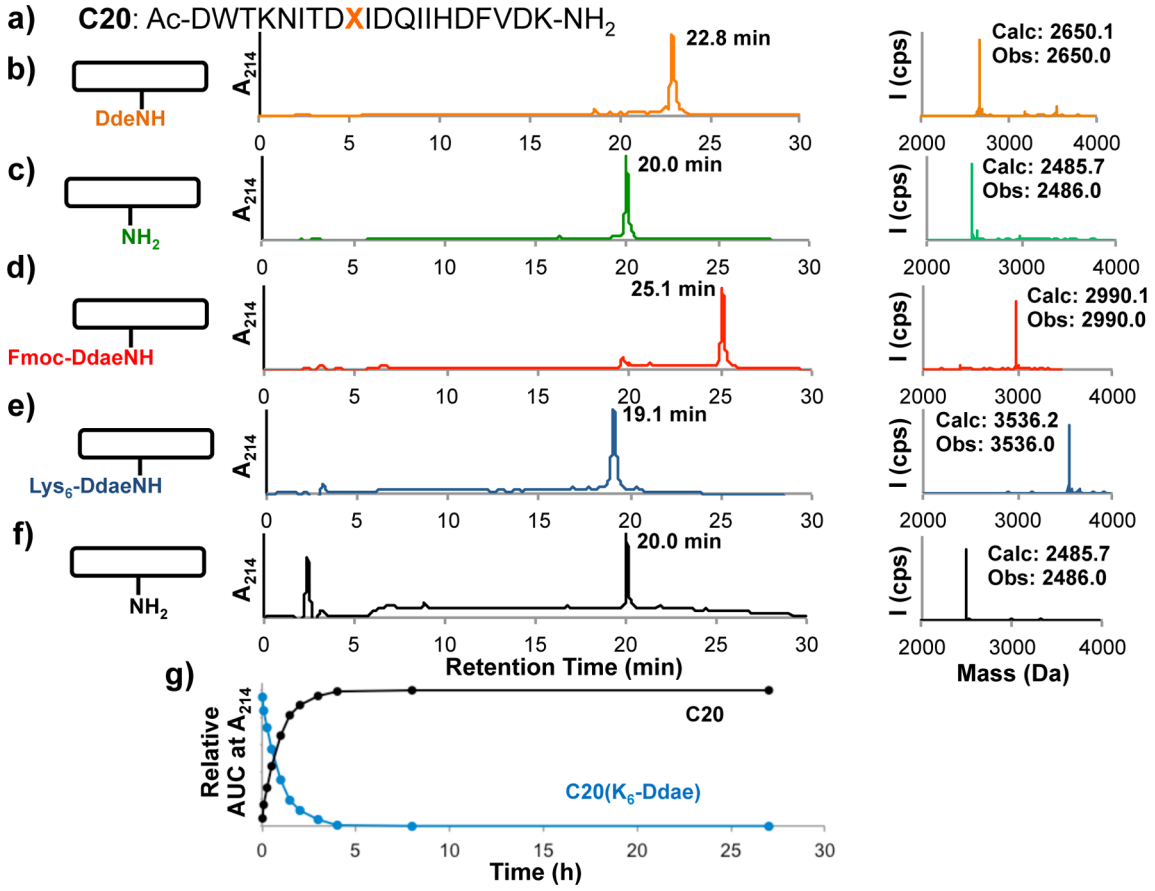


Figure 2. Incorporation and removal of Lys_6-Ddae in C20 peptide. (a) C20 target amino acid sequence. X indicates position of linker incorporation at Lys residue. (b–f) HPLC traces and deconvoluted MS for C20(Dde) (b), C20 (c), C20(Fmoc-Ddae) (d), C20(Lys_6-Ddae) (e), and purified C20 after linker cleavage (f). (g) Time course of linker cleavage using 1 M hydrazine.

oligosaccharide synthesis,[55] polyamine synthesis,[49,56,57] affinity purification,[58] and protein labeling.[59]

We prepared Fmoc-Ddae-OH (Figure 1a) in one step, by reaction of commercially available *N*-Fmoc-amido-PEG_2-acid with dimedone (Figures S1–S4). The general strategy for implementing the new linker is shown in Figure 1b. Here, a "difficult" peptide is prepared by standard Fmoc-SPPS, incorporating a single internal Lys residue with an orthogonal protecting group (Dde in this case), while the N-terminus is protected by a piperidine-stable group (Boc in this case) (Figure 1b, step 1). Upon elongation of the linear peptide chain but before cleavage from the resin, the Dde group is selectively removed with 3% hydrazine in DMF to reveal a single primary amine in the solid-supported peptide (Figure 1b, step 2).

Fmoc-Ddae-OH can then be directly reacted (i.e., no other additives are needed, and excess reagent can be recycled) with the primary amine to introduce a fresh *N*-Fmoc-protected amine linked to the peptide sequence through a hydrophilic (PEG_2) Dde-derived cleavable moiety (Figure 1b, step 3). Next, Fmoc-SPPS can be used to build a solubilizing peptide sequence (helping hand) (Figure 1b, step 4) on the Lys side chain. Note that we previously observed partial instability under aqueous conditions of a Dde-based linker on the N-terminus, while Dde linked on Lys side chain was stable over 24 h.[60] Cleavage from resin and HPLC purification can then be performed to obtain the purified peptide segment equipped with a solubilizing tag (Figure 1b, step 5).

After NCL-based protein assembly, cleavage of the linker and restoration of native peptide structure can then be performed by addition of aqueous hydrazine (Figure 1b, step 6) and mild conditions similar to those used for acetate deprotection during glycopeptide and glycoprotein synthesis.[61,62]

Ddae Linker Introduction, Stability, and Cleavage in a Model Peptide. A model peptide, Ebola virus C20 (Ac-DWTKNITDKIDQIIHDFVDK-NH_2), was used to characterize incorporation and removal of the Ddae linker with a Lys_6 helping hand. The C20 peptide is derived from the C-terminal heptad repeat region of the Ebola virus GP2 protein.[63,64]

We first synthesized C20 (Figure 2a) including a central Lys(Dde) (Figure 2b and S15–S17), then removed the Dde group using hydrazine (Figures 2c and S18–S20). Fmoc-Ddae-OH (200 mM in NMP) was then added to the peptide resin for 7 h to generate C20(Fmoc-Ddae) (Figure 2d, time course studies show nearly complete reaction after 2 h, Figures S21–S23). On-resin Fmoc-SPPS was next performed to extend a Lys_6 helping hand, followed by cleavage from resin using standard TFA conditions to generate C20(K_6-Ddae) (Figures 2e and S24–S26). The cleanliness of the crude products (Figures 2b–e and S27) highlights the efficiency of the multistep process.

We next evaluated removal of the helping hand in solution via hydrazine treatment by conducting a time-course study to measure the conversion of C20(K_6-Ddae) to C20. Purified C20(K_6-Ddae) was dissolved in 6 M GuHCl, 100 mM phosphate buffer, pH 7.5. Cleavage of the K_6-Ddae group

DOI: 10.1021/jacs.6b05719
J. Am. Chem. Soc. 2016, 138, 11775–11782

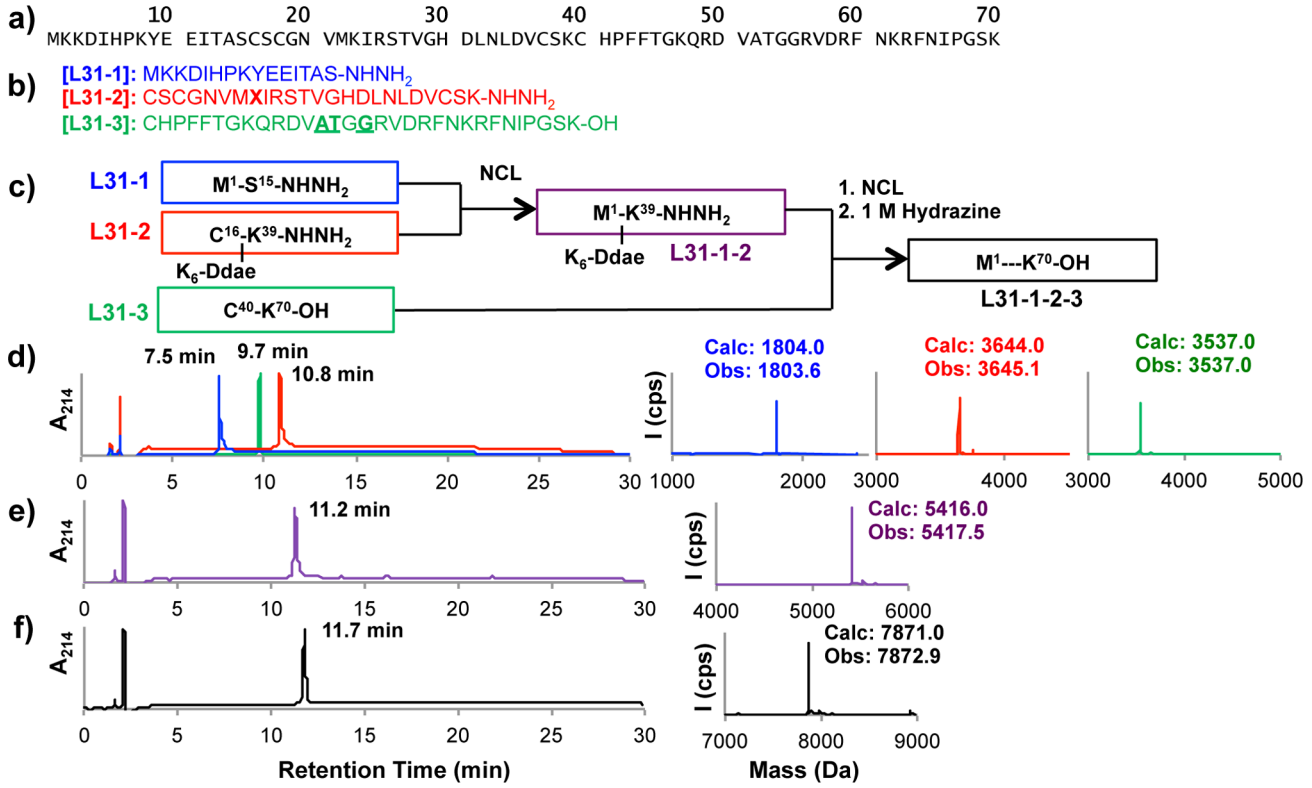


Figure 3. Synthesis of *E. coli* L31 protein. (a) L31 target amino acid sequence. (b) Amino acid sequences of peptide segments; X indicates position of linker incorporation at Lys residue. Pseudoprolines and (Dmb)Gly are indicated in bold and underlined text. (c) Synthetic strategy for L31. (d) HPLC traces and deconvoluted MS for initial peptide segments L31-1 (blue), L31-2 (red), and L31-3 (green). (e) HPLC trace and deconvoluted MS of first purified ligation product L31-1-2. (f) HPLC trace and deconvoluted MS of final, full-length target, L31-1-2-3.

was triggered by addition of a hydrazine solution in phosphate buffer (final 1 M hydrazine, pH 7.5) (see section S8 of the Supporting Information for specific details on preparing the hydrazine solution). Clean elimination of the (K_6-Ddae) group was achieved with a $t_{1/2}$ of ~1 h (Figures 2f and S28–31). This C20 peptide, prepared using a helping hand approach, possessed matched HPLC and MS traces with C20 prepared without Fmoc-Ddae linker (compare Figures 2c and S18–20 with Figures 2f and S31–S32).

We also examined the stability of C20(K_6-Ddae) under several different conditions employed in modern chemical peptide/protein synthesis including acidic HPLC buffer (Figure S33), denaturing conditions at low (Figure S34 and S36) and neutral pH (Figure S35), and a typical NCL buffer (Figure S37). No significant Ddae cleavage was observed under these standard conditions (Figure S39). In contrast, we did observe partial sensitivity of the linker (~20% cleavage after 24 h) to thiazolidine ring-opening conditions[65−67] (Figures S38 and S39). However, we tested relatively harsh conditions, 200 mM $MeONH_2$ for 24 h, while most reports employ either lower $MeONH_2$ concentrations (20–50 mM) or shorter incubation times (3–5 h).

Our studies with the C20 model peptide show that the Fmoc-Ddae-OH linker can be used to introduce solubilizing sequences at Lys side chains, while quick and clean removal of the linker is achieved with 1 M hydrazine. Importantly, the linker is stable to several buffers commonly used in chemical peptide/protein synthesis.

NCL-Based Synthesis of Ribosomal Protein L31 Using a Helping Hand. We next pursued the synthesis of a full-length protein target via NCL. Here, we selected the L31 protein from *E. coli*. L31 (Uniprot ID: P0A7M9, RL31_ECOLI) is a 70-residue protein within the large ribosomal subunit.[68,69]

The synthesis strategy for preparing L31 entailed three peptide segments (L31-1, -2, and -3), including Lys23 in segment 2 modified with a Lys_6 helping hand (Figure 3a–c). All three starting peptide segments were synthesized and purified (Figures 3d and S40–S51). L31-1 and L31-2 were prepared as peptide hydrazides for subsequent *in situ* activation/thiolysis for NCL,[70−73] while L31-3 was prepared with a C-terminal carboxylic acid, in order to match the natural protein (Figure 3c).

The first ligation reaction, L31-1 to L31-2 at the Ser15–Cys16 junction, was complete in 1 h (Figure S52). The ligation product L31-1-2 was then purified by HPLC, confirming that the helping hand was stable in an NCL reaction (Figures 3e and S53–S55). We next ligated L31-1-2 to L31-3 to directly evaluate the linker's stability under oxidative hydrazide activation conditions. However, during preliminary studies of this reaction, we encountered significant lactam formation, which can be a challenge with Lys thioesters.[74] In this case, we overcame the problem by activating L31-1-2 hydrazide in batches that were sequentially added to L31-3 in the ligation reaction (see Supporting Information for details). In addition to demonstrating that the second ligation was successful (Figures S56–S57), we were able to purify both the MESNa thioester of L31-1-2 (Figures S58–S60) and the full-length product with helping hand intact (Figures S61–S63). Critically, the linker was stable under peptide hydrazide activation conditions (20 mM $NaNO_2$, pH 3, −20 °C, 20 min).[70−73]

DOI: 10.1021/jacs.6b05719
J. Am. Chem. Soc. 2016, 138, 11775–11782

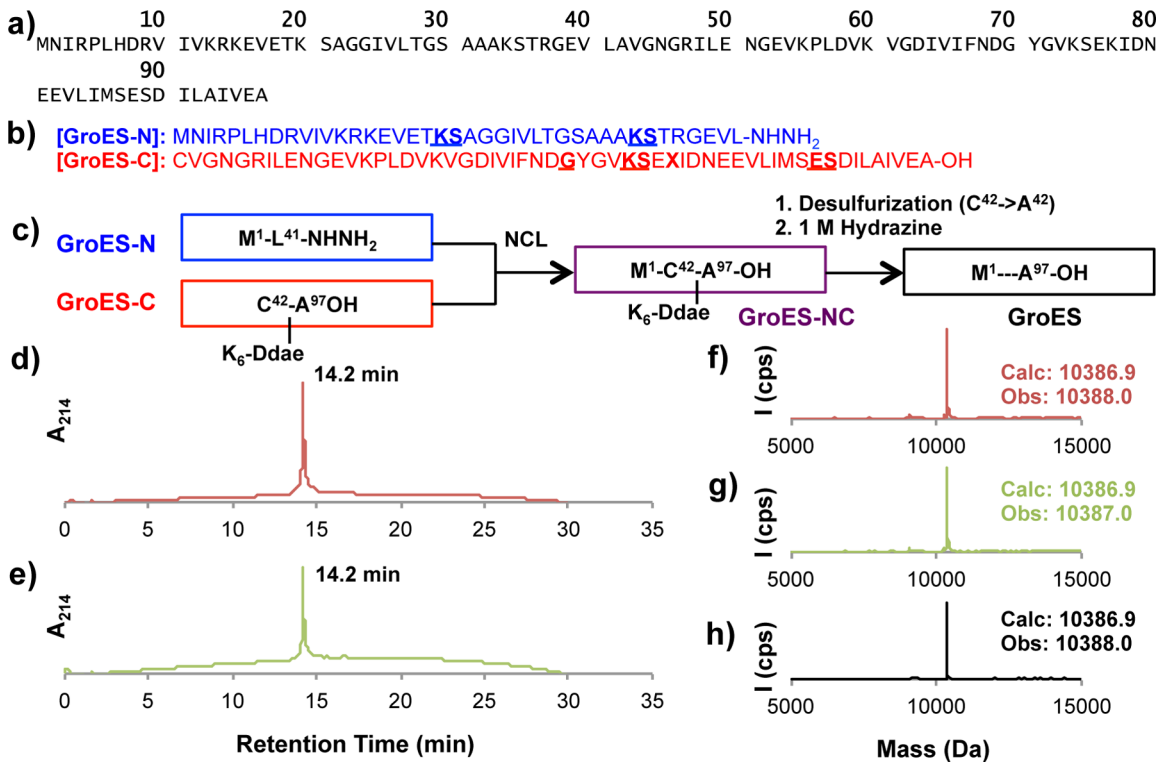


Figure 4. Synthesis of *E. coli* GroES protein. (a) GroES target amino acid sequence. (b) Amino acid sequences of peptide segments; X indicates position of linker incorporation at Lys residue. Pseudoprolines and (Dmb)Gly are indicated in bold and underlined text. (c) Synthetic strategy for GroES. (d) HPLC trace of purified, folded synthetic L-GroES (red). (e) HPLC trace of purified, folded synthetic D-GroES (green). (f–h) Deconvoluted MS for folded synthetic L- (red), D- (green), and recombinant GroES (black), respectively.

We next showed that the helping hand could be removed in one pot after ligation of the full-length product by treatment with 1 M hydrazine, pH 7.5, for 2.5 h (Figure S64). We generated clean and correct full-length product, without helping hand, using this straightforward one-pot ligation and helping hand removal strategy (Figures 3f and S65–S67).

Synthesis Strategy and Initial Efforts to Synthesize GroES. Having validated the helping hand approach in Ebola virus C20 peptide and *E. coli* L31 protein, we pursued a more ambitious and biologically relevant protein target: the *E. coli* co-chaperonin GroES. Our interest in synthesizing GroES was motivated by our previous work showing that the GroEL/ES chaperone was capable of folding both natural (L-) and mirror-image (D-) forms of a synthetic substrate protein, DapA.[75] This result demonstrated that GroEL/ES recognizes substrate proteins via nonspecific (i.e., nonstereoselective) hydrophobic interactions.

We were interested to probe this interaction more deeply. On this note, crystal structures of GroEL/ES complexes have shown considerable structural heterogeneity in the interaction of the GroES mobile loop with GroEL.[76] Another group reported that substitution of six positions in the GroES mobile loop with Ala residues did not perturb GroEL complementation in *E. coli* cells.[77] Additionally, Gierasch's lab showed by NMR that both L- and D-peptides could bind to GroEL.[78] These data suggest a high degree of plasticity in the GroEL/ES interface. Based on this information and our previous data, we hypothesized that a heterochiral complex (L-GroEL/D-GroES or D-GroEL/L-GroES) may possess chaperone activity. To test this hypothesis, we pursued the chemical synthesis of D-GroES (97 residues) to be tested in complex with recombinant (L-) GroEL (548 residues). We also pursued the chemical synthesis of L-GroES as a positive control to validate the activity of our synthetic material.

Total chemical synthesis of *E. coli* GroES (Uniprot ID: P0A6F9, CH10_ECOLI) was reported in 1991 by an ambitious preparation of the entire protein via Boc-SPPS; however, no analytical HPLC or MS data were provided on the synthetic material.[79] We reasoned that modern advances in peptide synthesis and ligation technologies made this protein well-suited for a multisegment synthetic approach.

Our retrosynthetic strategy involved two peptide segments, with a single ligation junction at position Leu41–Ala42 (Figure 4a,b). Following ligation of the GroES-N and GroES-C peptides, desulfurization (Cys → Ala) at position 42 would generate full-length native GroES. Unfortunately, our efforts to prepare GroES in this manner failed due to severe handling difficulties with the GroES-C peptide. Specifically, we were unable to dissolve a significant amount of peptide in any usable solvent (e.g., water/acetonitrile + 0.1% TFA, 6 M GuHCl + 5% AcOH, DMF, and ammonium acetate/water/acetonitrile at pH 6), and we were also unable to isolate a clean product by RP-HPLC (Figure S68).

These challenges with GroES-C forced us to rethink our synthesis strategy. We considered introducing an additional ligation site C-terminal to Leu41 in order to break up this difficult segment. However, the absence of any suitable Cys or Ala residues in this region (residues 43–97) complicated this prospect. Potentially, a non-natural thiolated residue (reviewed in refs 28, 80, and 81) could be used in this region. However, this would require multistep chemical syntheses of both L- and D-versions of a specialty thiolated amino acid, as well as extra handling steps (additional ligation and desulfurization) in the GroES assembly. Furthermore, we found that a shorter version

DOI: 10.1021/jacs.6b05719
J. Am. Chem. Soc. 2016, 138, 11775–11782

of GroES-C (peptide 61−97) was even less soluble and could not be dissolved for HPLC analysis.

Synthesis of L- and D-GroES Using a Helping Hand. Based on these challenges, we incorporated a Lys_6 helping hand in the GroES-C peptide (at Lys77, Figure 4b,c). The four starting peptide segments for generating L- and D-GroES were then synthesized, with L- and D-GroES-N being prepared as peptide hydrazides and L-GroES-C and D-GroES-C being prepared as C-terminal acids to match the natural protein. Standard Fmoc building blocks and SPPS conditions were used in the syntheses, with a few exceptions. First, several Xaa−Ser($\Psi^{Me,Me}$Pro) pseudoproline dipeptides were utilized in all four peptides to enhance synthesis quality by reducing peptide aggregation during SPPS (underlined, Figure 4b). The noncommercially available D-pseudoproline dipeptides were synthesized according to Mutter's original protocol, with slight modifications (Figures S5−S14).[82,83] Additionally, achiral (Dmb)Gly was used at Gly70 to prevent piperidine-mediated aspartimide formation at the Asp69−Gly70 site. Lys(Dde) was used at position Lys77 to install the helping hand, while Boc-Cys(Trt) was incorporated at position 42 to protect the GroES-C N-terminus. Upon completion of the linear GroES-C peptide sequences, the Dde group was removed on-resin, and Fmoc-Ddae-OH was then coupled, followed by Fmoc-SPPS to generate a Lys_6 helping hand. All four peptides (L- and D-GroES-N and L- and D-GroES-C) were then cleaved, purified by preparative HPLC, and characterized by HPLC and LC/MS (Figures S69−S76 and S84−S91).

Peptide hydrazides L- and D-GroES-N were activated via $NaNO_2$ oxidation and 4-mercaptophenylacetic acid (MPAA) thiolysis to generate thioesters for reaction with their partner Cys-peptides L- and D-GroES-C, respectively. Gratifyingly, ligations were complete in less than 4 h, showing no noticeable traces of products arising from premature Ddae cleavage or related undesirable reactions. The ligation products were purified by HPLC and characterized by analytical HPLC and LC/MS (Figures S77−S80 and S92−S95). We obtained isolated yields of 24 and 21% for L- and D-GroES-NC based on limiting GroES-C peptides.

We next proceeded to the final synthetic step: one-pot desulfurization and helping hand removal to generate full-length, native L- and D-GroES. Peptides L- and D-GroES-NC(Cys42, K_6-Ddae77) were desulfurized[84] using the free radical method[85] to convert the ligation junction Cys42 into Ala42 (Figures S81 and S96). The desulfurization was complete in less than 4 h, as determined by LC/MS. Importantly, the Ddae linker was stable under these conditions. The crude mixtures were then treated with 1 M hydrazine, pH 7.5, for 3 h to cleave the Ddae linker, remove the helping hand, and generate native GroES sequences (Figures S82 and S97).

At this stage, the crude reaction mixture reaction could be injected over RP-HPLC to isolate pure GroES. However, we found that this led to a large yield loss. Instead, taking advantage of the clean one-pot process, we found that GroES in the final reaction could be directly refolded by overnight dialysis (3000 MWCO) into 25 mM Tris, 50 mM NaCl, pH 7.5. The folded heptameric GroES was then purified by size-exclusion chromatography (SEC, Figures S83 and S98) and analyzed under denaturing conditions by HPLC and LC/MS (Figure 4d−g). L- and D-GroES showed clean products with identical retention times on RP-HPLC (Figure 4d,e) and matched MS to recombinant GroES purified in the same way (pooled from SEC, Figures S99 and 4f−h). We obtained isolated yields of 24 and 23% for folded L- and D-GroES, respectively.

Characterization and Activity Testing of Synthetic L- and D-GroES. With the heptameric proteins (recombinant, synthetic L, and synthetic D) in hand, we performed circular dichroism (CD) spectroscopy to confirm the correctly folded structures. CD (Figure 5a) demonstrated that our synthetic L-GroES possessed structure that mimicked the recombinant control and that synthetic D-GroES showed the inverted spectrum expected for a mirror-image protein.

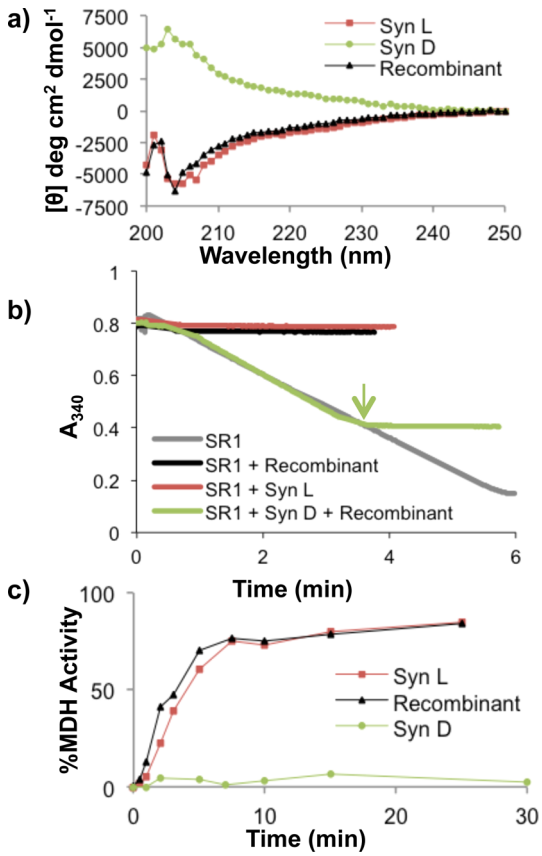


Figure 5. Structure and activity of synthetic GroES. (a) Circular dichroism spectra of synthetic L- (red), D- (green), and recombinant (black) GroES. (b) SR1 activity assay. (c) MDH refolding assay. Arrow in b shows that additional recombinant GroES can still bind to SR1 in the presence of D-GroES.

We then performed enzymatic assays to evaluate the chaperone activity of our synthetic GroES proteins. Here, we first sought to demonstrate that our synthetic L-GroES was fully active compared to recombinant protein in assisting protein folding when combined with recombinant GroEL. Second, we tested if D-GroES was capable of forming an active heterochiral complex with recombinant GroEL to assist in protein folding.

We used two quantitative folding assays to evaluate GroES activity. We first tested an SR1 assay to evaluate binding of GroES to GroEL. In this assay, functional GroES binds to a single-ring mutant of GroEL (termed SR1) to form a stable, dead-end complex due to the absence of the trans GroEL ring. When saturated with GroES, the ATPase activity of SR1 is completely inhibited. In the absence of GroES, the ATPase of SR1 is uninhibited. In the second assay, we directly measured GroEL-/ES-mediated refolding of client protein malate

DOI: 10.1021/jacs.6b05719
J. Am. Chem. Soc. 2016, 138, 11775−11782

dehydrogenase (MDH) by GroEL, using synthetic D-GroES with synthetic and recombinant L-GroES as positive controls.

In both the SR1 and MDH assays, synthetic L-GroES performed similarly to recombinant GroES, validating the high quality of our synthetic material (Figure 5b,c). However, counter to our expectations, D-GroES was completely non-functional in both the SR1 and MDH assays (Figure 5b,c). Based on these results, we conclude that heterochiral GroEL/ES is not functional in protein folding, and unlike substrate recognition/refolding,[75] there is an element of stereoselectivity to the interaction between GroEL and GroES.

CONCLUSIONS

In this work, we introduce a general new tool, Fmoc-Ddae-OH, for dealing with poorly soluble and aggregation-prone peptides. Fmoc-Ddae-OH is easy to synthesize and then incorporate into peptide segments for introducing desired solubilizing sequences by Fmoc-SPPS. This new linker can be incorporated at any Lys residue within a peptide segment and is stable under a wide range of conditions allowing handling, purification, and storage of the resulting peptides, as well as NCL-based assembly including hydrazide activation and desulfurization. Cleavage can be triggered by mild treatment with aqueous hydrazine in sodium phosphate buffer at near-neutral pH. Importantly, the Ddae linker can be cleaved in one pot after NCL and desulfurization, and before a folding step.

We thus conclude that beyond NCL applications this new reagent will be generally applicable for installing semi-permanent solubilizing groups in difficult peptides. Moreover, considering the compatibility of its parent Dde protecting group, this strategy is likely also compatible with Boc-SPPS.

We capitalized on these benefits to prepare highly pure forms of both L- and D-GroES heptamers by a NCL strategy that could not be implemented without the help of semipermanent solubilizing tags. Using these proteins, we show that heterochiral L-GroEL/D-GroES is incapable of folding substrate proteins. Thus, although substrate recognition by GroEL/ES is ambidextrous,[75] the composition of the folding machine is not.

Current work is ongoing to test this new linker in other difficult NCL-based chemical protein synthesis projects and exploit the possibility to screen for a tailor-made solubilizing tag[86,87] adapted to any difficult segment.[88] We aim to expand its application beyond the introduction of solubilizing tags (e.g., to immobilize a peptide for solid phase ligations, introduce affinity purification tags, or temporarily link protein subunits in order to facilitate difficult folding processes).

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b05719.

Synthesis and characterization of Fmoc-Ddae-OH and D-pseudoproline dipeptides, methods, synthesis information for L31, L-GroES, and D-GroES, and analysis of GroES species (PDF)

AUTHOR INFORMATION

Corresponding Authors

*aucagne@cnrs-orleans.fr
*kay@biochem.utah.edu

Present Address

M.T.J.: Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, United Kingdom.

Notes

The authors declare the following competing financial interest(s): M.S.K. is a Scientific Director, consultant, and equity holder of the D-Peptide Research Division of Navigen, which is commercializing D-peptide inhibitors of viral entry.

ACKNOWLEDGMENTS

We greatly thank Dr. Agnès Delmas for fruitful discussions and advice. We thank Rob Marquardt for his mass spectrometry advice and training, and Hervé Meudal, Dr. Cyril Colas, and the CBM and FR2708 federation NMR and HRMS platforms for analysis of Fmoc-Ddae-OH. Funding for this work was provided by NIH Predoctoral Training Grant F31CA171677 (to M.T.J.), NIH T32 Cardiovascular Training Grant HL007576 (to M.E.P.), NIH Grants AI076168 and AI102347 (to M.S.K.), French National Funding Agency ANR-15-CE07-022 EasyMiniProt grant (to V.A.), and la Ligue contre le cancer for a postdoctoral fellowship (to M.G.).

REFERENCES

(1) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. *Science* **1994**, *266*, 776.
(2) Kent, S. B. *Chem. Soc. Rev.* **2009**, *38*, 338.
(3) Hackenberger, C. P.; Schwarzer, D. *Angew. Chem., Int. Ed.* **2008**, *47*, 10030.
(4) Lee, C. L.; Li, X. *Curr. Opin. Chem. Biol.* **2014**, *22*, 108.
(5) Pattabiraman, V. R.; Bode, J. W. *Nature* **2011**, *480*, 471.
(6) Harmand, T. J.; Murar, C. E.; Bode, J. W. *Curr. Opin. Chem. Biol.* **2014**, *22*, 115.
(7) Hemantha, H. P.; Brik, A. *Bioorg. Med. Chem.* **2013**, *21*, 3411.
(8) Spasser, L.; Brik, A. *Angew. Chem., Int. Ed.* **2012**, *51*, 6840.
(9) Ajish Kumar, K. S.; Haj-Yahya, M.; Olschewski, D.; Lashuel, H. A.; Brik, A. *Angew. Chem., Int. Ed.* **2009**, *48*, 8090.
(10) Kumar, K. S.; Spasser, L.; Erlich, L. A.; Bavikar, S. N.; Brik, A. *Angew. Chem., Int. Ed.* **2010**, *49*, 9126.
(11) Wang, P.; Dong, S.; Shieh, J. H.; Peguero, E.; Hendrickson, R.; Moore, M. A.; Danishefsky, S. J. *Science* **2013**, *342*, 1357.
(12) Wang, P.; Dong, S.; Brailsford, J. A.; Iyer, K.; Townsend, S. D.; Zhang, Q.; Hendrickson, R. C.; Shieh, J.; Moore, M. A. S.; Danishefsky, S. J. *Angew. Chem., Int. Ed.* **2012**, *51*, 11576.
(13) Hirano, K.; Macmillan, D.; Tezuka, K.; Tsuji, T.; Kajihara, Y. *Angew. Chem., Int. Ed.* **2009**, *48*, 9557.
(14) Murakami, M.; Okamoto, R.; Izumi, M.; Kajihara, Y. *Angew. Chem., Int. Ed.* **2012**, *51*, 3567.
(15) Wilson, R. M.; Dong, S.; Wang, P.; Danishefsky, S. J. *Angew. Chem., Int. Ed.* **2013**, *52*, 7646.
(16) Kent, S. B. *Angew. Chem., Int. Ed.* **2013**, *52*, 11988.
(17) Chen, S. Y.; Cressman, S.; Mao, F.; Shao, H.; Low, D. W.; Beilan, H. S.; Cagle, E. N.; Carnevali, M.; Gueriguian, V.; Keogh, P. J.; Porter, H.; Stratton, S. M.; Wiedeke, M. C.; Savatski, L.; Adamson, J. W.; Bozzini, C. E.; Kung, A.; Kent, S. B.; Bradburne, J. A.; Kochendoerfer, G. G. *Chem. Biol.* **2005**, *12*, 371.
(18) Kochendoerfer, G. G.; Chen, S. Y.; Mao, F.; Cressman, S.; Traviglia, S.; Shao, H.; Hunter, C. L.; Low, D. W.; Cagle, E. N.; Carnevali, M.; et al. *Science* **2003**, *299*, 884.
(19) Milton, R. C.; Milton, S. C.; Kent, S. B. *Science* **1992**, *256*, 1445.
(20) Zhao, L.; Lu, W. *Curr. Opin. Chem. Biol.* **2014**, *22*, 56.
(21) Zawadzke, L. E.; Berg, J. M. *Proteins: Struct., Funct., Genet.* **1993**, *16*, 301.
(22) Verzele, D.; Madder, A. *ChemBioChem* **2013**, *14*, 1032.
(23) Raibaut, L.; El Mahdi, O.; Melnyk, O. *Top. Curr. Chem.* **2014**, *363*, 103.
(24) Zuo, C.; Tang, S.; Zheng, J.-S. *J. Pept. Sci.* **2015**, *21*, 540.

DOI: 10.1021/jacs.6b05719
J. Am. Chem. Soc. 2016, 138, 11775−11782

(25) Shen, F.; Huang, Y.-C.; Tang, S.; Chen, Y.-X.; Liu, L. *Isr. J. Chem.* **2011**, *51*, 940.
(26) Olschewski, D.; Becker, C. F. *Mol. BioSyst.* **2008**, *4*, 733.
(27) Saito, F.; Noda, H.; Bode, J. W. *ACS Chem. Biol.* **2015**, *10*, 1026.
(28) Malins, L. R.; Payne, R. J. *Curr. Opin. Chem. Biol.* **2014**, *22*, 70.
(29) Johnson, E. C.; Kent, S. B. *Tetrahedron Lett.* **2007**, *48*, 1795.
(30) Sato, T.; Saito, Y.; Aimoto, S. *J. Pept. Sci.* **2005**, *11*, 410.
(31) Boerema, D. J.; Tereshko, V. A.; Kent, S. B. *Biopolymers* **2008**, *90*, 278.
(32) Johnson, E. C.; Malito, E.; Shen, Y.; Rich, D.; Tang, W. J.; Kent, S. B. *J. Am. Chem. Soc.* **2007**, *129*, 11480.
(33) Lahiri, S.; Brehs, M.; Olschewski, D.; Becker, C. F. *Angew. Chem., Int. Ed.* **2011**, *50*, 3988.
(34) Li, X.; Kawakami, T.; Aimoto, S. *Tetrahedron Lett.* **1998**, *39*, 8669.
(35) Wang, J. X.; Fang, G. M.; He, Y.; Qu, D. L.; Yu, M.; Hong, Z. Y.; Liu, L. *Angew. Chem., Int. Ed.* **2015**, *54*, 2194.
(36) Englebretsen, D. R.; Alewood, P. F. *Tetrahedron Lett.* **1996**, *37*, 8431.
(37) Hossain, M. A.; Belgi, A.; Lin, F.; Zhang, S.; Shabanpoor, F.; Chan, L.; Belyea, C.; Truong, H. T.; Blair, A. R.; Andrikopoulos, S.; Tregear, G. W.; Wade, J. D. *Bioconjugate Chem.* **2009**, *20*, 1390.
(38) Yang, S. H.; Wojnar, J. M.; Harris, P. W.; DeVries, A. L.; Evans, C. W.; Brimble, M. A. *Org. Biomol. Chem.* **2013**, *11*, 4935.
(39) Harris, P. W.; Brimble, M. A. *Biopolymers* **2010**, *94*, 542.
(40) Peigneur, S.; Paolini-Bertrand, M.; Gaertner, H.; Biass, D.; Violette, A.; Stocklin, R.; Favreau, P.; Tytgat, J.; Hartley, O. *J. Biol. Chem.* **2014**, *289*, 35341.
(41) Chemuru, S.; Kodali, R.; Wetzel, R. *Biopolymers* **2014**, *102*, 206.
(42) Tan, Z.; Shang, S.; Danishefsky, S. J. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 4297.
(43) Huang, Y. C.; Li, Y. M.; Chen, Y.; Pan, M.; Li, Y. T.; Yu, L.; Guo, Q. X.; Liu, L. *Angew. Chem., Int. Ed.* **2013**, *52*, 4858.
(44) Zheng, J. S.; Yu, M.; Qi, Y. K.; Tang, S.; Shen, F.; Wang, Z. P.; Xiao, L.; Zhang, L.; Tian, C. L.; Liu, L. *J. Am. Chem. Soc.* **2014**, *136*, 3695.
(45) Johnson, T.; Quibell, M.; Sheppard, R. C. *J. Pept. Sci.* **1995**, *1*, 11.
(46) Wahlström, K.; Planstedt, O.; Undén, A. *Tetrahedron Lett.* **2008**, *49*, 3921.
(47) Wahlström, K.; Planstedt, O.; Undén, A. *Tetrahedron Lett.* **2008**, *49*, 3779.
(48) Zheng, J. S.; He, Y.; Zuo, C.; Cai, X. Y.; Tang, S.; Wang, Z. A.; Zhang, L. H.; Tian, C. L.; Liu, L. *J. Am. Chem. Soc.* **2016**, *138*, 3553.
(49) Chhabra, S. R.; Khan, A. N.; Bycroft, B. W. *Tetrahedron Lett.* **1998**, *39*, 3585.
(50) Chhabra, S. R.; Hothi, B.; Evans, D. J.; White, P. D.; Bycroft, B. W.; Chan, W. C. *Tetrahedron Lett.* **1998**, *39*, 1603.
(51) Bycroft, B. W.; Chan, W. C.; Chhabra, S. R.; Hone, N. D. *J. Chem. Soc., Chem. Commun.* **1993**, 778.
(52) Díaz-Mochón, J. J.; Bialy, L.; Bradley, M. *Org. Lett.* **2004**, *6*, 1127.
(53) Turner, R. A.; Hauksson, N. E.; Gipe, J. H.; Lokey, R. S. *Org. Lett.* **2013**, *15*, 5012.
(54) Chhabra, S. R.; Parekh, H.; Khan, A. N.; Bycroft, B. W.; Kellam, B. *Tetrahedron Lett.* **2001**, *42*, 2189.
(55) Drinnan, N.; West, M. L.; Broadhurst, M.; Kellam, B.; Toth, I. *Tetrahedron Lett.* **2001**, *42*, 1159.
(56) Chhabra, S. R.; Khan, A. N.; Bycroft, B. W. *Tetrahedron Lett.* **2000**, *41*, 1095.
(57) Chhabra, S. R.; Khan, A. N.; Bycroft, B. W. *Tetrahedron Lett.* **2000**, *41*, 1099.
(58) Kellam, B.; Chan, W. C.; Chhabra, S. R.; Bycroft, B. W. *Tetrahedron Lett.* **1997**, *38*, 5391.
(59) Yang, Y.; Verhelst, S. H. L. *Chem. Commun.* **2013**, *49*, 5366.
(60) Galibert, M.; Piller, V.; Piller, F.; Aucagne, V.; Delmas, A. F. *Chemical Science* **2015**, *6*, 3617.
(61) Tan, Z.; Shang, S.; Halkina, T.; Yuan, Y.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2009**, *131*, 5424.
(62) Shin, Y.; Winans, K. A.; Backes, B. J.; Kent, S. B. H.; Ellman, J. A.; Bertozzi, C. R. *J. Am. Chem. Soc.* **1999**, *121*, 11684.
(63) Weissenhorn, W.; Carfi, A.; Lee, K. H.; Skehel, J. J.; Wiley, D. C. *Mol. Cell* **1998**, *2*, 605.
(64) Malashkevich, V. N.; Schneider, B. J.; McNally, M. L.; Milhollen, M. A.; Pang, J. X.; Kim, P. S. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 2662.
(65) Bang, D.; Kent, S. B. *Angew. Chem., Int. Ed.* **2004**, *43*, 2534.
(66) Dang, B.; Kubota, T.; Correa, A. M.; Bezanilla, F.; Kent, S. B. *Angew. Chem., Int. Ed.* **2014**, *53*, 8970.
(67) Villain, M.; Vizzavona, J.; Rose, K. *Chem. Biol.* **2001**, *8*, 673.
(68) Arnold, R. J.; Reilly, J. P. *Anal. Biochem.* **1999**, *269*, 105.
(69) Plunkett, G., 3rd; Burland, V.; Daniels, D. L.; Blattner, F. R. *Nucleic Acids Res.* **1993**, *21*, 3391.
(70) Zheng, J. S.; Tang, S.; Guo, Y.; Chang, H. N.; Liu, L. *ChemBioChem* **2012**, *13*, 542.
(71) Fang, G. M.; Wang, J. X.; Liu, L. *Angew. Chem., Int. Ed.* **2012**, *51*, 10347.
(72) Fang, G. M.; Li, Y. M.; Shen, F.; Huang, Y. C.; Li, J. B.; Lin, Y.; Cui, H. K.; Liu, L. *Angew. Chem., Int. Ed.* **2011**, *50*, 7645.
(73) Zheng, J. S.; Tang, S.; Qi, Y. K.; Wang, Z. P.; Liu, L. *Nat. Protoc.* **2013**, *8*, 2483.
(74) Siman, P.; Karthikeyan, S. V.; Nikolov, M.; Fischle, W.; Brik, A. *Angew. Chem., Int. Ed.* **2013**, *52*, 8059.
(75) Weinstock, M. T.; Jacobsen, M. T.; Kay, M. S. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 11679.
(76) Fei, X.; Ye, X.; LaRonde, N. A.; Lorimer, G. H. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 12775.
(77) Nojima, T.; Ikegami, T.; Taguchi, H.; Yoshida, M. *J. Mol. Biol.* **2012**, *422*, 291.
(78) Wang, Z.; Feng, H.; Landry, S. J.; Maxwell, J.; Gierasch, L. M. *Biochemistry* **1999**, *38*, 12537.
(79) Mascagni, P.; Tonolo, M.; Ball, H.; Lim, M.; Ellis, R. J.; Coates, A. *FEBS Lett.* **1991**, *286*, 201.
(80) He, Q.-Q.; Fang, G.-M.; Liu, L. *Chin. Chem. Lett.* **2013**, *24*, 265.
(81) Wong, C. T. T.; Tung, C. L.; Li, X. *Mol. BioSyst.* **2013**, *9*, 826.
(82) Haack, T.; Mutter, M. *Tetrahedron Lett.* **1992**, *33*, 1589.
(83) Wöhr, T.; Wahl, F.; Nefzi, A.; Rohwedder, B.; Sato, T.; Sun, X.; Mutter, M. *J. Am. Chem. Soc.* **1996**, *118*, 9218.
(84) Yan, L. Z.; Dawson, P. E. *J. Am. Chem. Soc.* **2001**, *123*, 526.
(85) Wan, Q.; Danishefsky, S. J. *Angew. Chem., Int. Ed.* **2007**, *46*, 9248.
(86) Paradis-Bas, M.; Albert-Soriano, M.; Tulla-Puche, J.; Albericio, F. *Org. Biomol. Chem.* **2014**, *12*, 7194.
(87) Schlapschy, M.; Binder, U.; Borger, C.; Theobald, I.; Wachinger, K.; Kisling, S.; Haller, D.; Skerra, A. *Protein Eng., Des. Sel.* **2013**, *26*, 489.
(88) Paradis-Bas, M.; Tulla-Puche, J.; Albericio, F. *Chem. Soc. Rev.* **2016**, *45*, 631.

DOI: 10.1021/jacs.6b05719
J. Am. Chem. Soc. 2016, 138, 11775−11782