Organic & Biomolecular Chemistry

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Cite this: *Org. Biomol. Chem.*, 2021, **19**, 8821

A glutamic acid-based traceless linker to address challenging chemical protein syntheses†

Riley J. Giesler,[a] Paul Spaltenstein,[a] Michael T. Jacobsen,[b] Weiliang Xu,[a] Mercedes Maqueda[c] and Michael S. Kay *[a]

Received 17th August 2021,
Accepted 21st September 2021

DOI: 10.1039/d1ob01611c

rsc.li/obc

Native chemical ligation (NCL) enables the total chemical synthesis of proteins. However, poor peptide segment solubility remains a frequently encountered challenge. Here we introduce a traceless linker that can be temporarily attached to Glu side chains to overcome this problem. This strategy employs a new tool, Fmoc-Glu(AlHx)-OH, which can be directly installed using standard Fmoc-based solid-phase peptide synthesis. The incorporated residue, Glu(AlHx), is stable to a wide range of chemical protein synthesis conditions and is removed through palladium-catalyzed transfer under aqueous conditions. General handling characteristics, such as efficient incorporation, stability and rapid removal were demonstrated through a model peptide modified with Glu(AlHx) and a Lys_6 solubilizing tag. Glu(AlHx) was incorporated into a highly insoluble peptide segment during the total synthesis of the bacteriocin AS-48. This challenging peptide was successfully synthesized and folded, and it has comparable antimicrobial activity to the native AS-48. We anticipate widespread use of this easy-to-use, robust linker for the preparation of challenging synthetic peptides and proteins.

1. Introduction

Chemical protein synthesis (CPS) provides an unparalleled route to proteins that cannot be readily accessed by other methods.[1–3] Proteins and peptides with unnatural chemical moieties (*e.g.*, biotin and fluorophore labels), post-translational modifications (*e.g.*, ubiquitylation, acetylation, glycosylation), and mirror-image D-amino acids have been synthesized by CPS. This process most commonly employs two primary techniques: solid-phase peptide synthesis (SPPS)[4] to synthesize peptide segments up to ~50 AAs and a chemoselective ligation method such as native chemical ligation (NCL)[1] (for other ligation strategies, see reviews).[3,5,6]

After SPPS, crude peptide segments are typically purified using reverse-phase HPLC (RP-HPLC) with an acidic aqueous/organic system using acetonitrile (ACN). The pure peptides can then be 'stitched' together through ligation. NCL requires a C-terminal peptide thioester and an N-terminal thiol (usually Cys) to generate a native amide bond between two segments. Unfortunately, Cys residues are relatively rare in proteins (1.4% abundance).[7] To circumvent this shortcoming, the Dawson and Danishefsky groups developed desulfurization methods to convert Cys to Ala.[2,8] These methods allows NCL junctions to be temporarily introduced at natural Ala (8.2% abundance)[7] residues, which are then desulfurized post-ligation. Similarly, selenocysteine can also be used to temporarily create a ligation junction (and converted back to Ala post-ligation), and has been extensively demonstrated by the Payne and Metanis groups.[9–11]

A major bottleneck in CPS is poor solubility of peptide segments. 'Difficult' peptide segments diminish purity and yield through on-resin aggregation, poor resolution during purification, and limited solubility during chemical reactions.[12,13] During SPPS, on-resin aggregation can be combatted through several tactics, such as lower resin loading densities and pseudoproline[14,15] or isoacyl[16–18] dipeptide building blocks. In-solution aggregation can also be disrupted through solvent or amide backbone manipulation, but another promising approach has been to introduce polar functional groups (generally cationic) into a difficult peptide segment (for a comprehensive review, see ref. 13). These solubilizing tags have been attached to peptides through N- and C-termini, residue side chains, and within the backbone.

For example, the Kent,[19] Tietze,[20] and Aimoto[21] groups have employed a solubilizing tag on peptide C-termini *via* a thioester linkage, which is then removed during NCL.

[a]*Department of Biochemistry, University of Utah School of Medicine, 15 North Medical Drive East, Room 4100, Salt Lake City, Utah 84112-5650, USA. E-mail: kay@biochem.utah.edu*
[b]*Department of Pediatrics, Division of Diabetes and Endocrinology, Stanford University, Palo Alto, CA 94304, USA*
[c]*Departamento de Microbiología, Universidad de Granada, Avda. Fuentenueva, s/n, 18071 Granada, Spain*

†Electronic supplementary information (ESI) available. See DOI: 10.1039/d1ob01611c

Unfortunately, the solubilizing benefits of these tags are not maintained throughout multiple ligations. The Liu group produced a generalized strategy for introducing solubilizing tags at any amino acid with their removable backbone modification.[22–26] Brik,[27–29] Aucagne,[30] and Yoshiya[31,32] developed Cys linkers based on acetomidomethyl (Acm)-, disulfide, and trityl-groups that form stable thioether or disulfide bonds. Yoshiya extended this concept through thiol-containing, premade Fmoc-Glu/Asp amino acids.[33]

These linkers have facilitated the syntheses of many difficult proteins and peptides. However, some require harsh removal conditions that may damage peptides, rely on less available amino acid residues, or are limited in the cargo that can be attached to the linker. To address these needs, we reported two generations of a Lys-based (4,4-dimethyl-2,6-dioxocyclohexylidene)ethyl (Dde)-like protecting group (Ddae and Ddap).[34,35] These linkers or 'helping hands' have been used to synthesize multiple proteins, such as the *E. coli* chaperonin GroES,[34] Shiga toxin subunit B,[35] and various insulin A-chains.[36,37]

While Ddae and Ddap are robust, they have a few inherent limitations. First, although Lys is a common amino acid (5.6%),[7] we have encountered numerous aggregation-prone peptide segments that are devoid of available Lys residues. In fact, a lack of Lys residues is often associated with poor peptide solubility (*e.g.*, a highly hydrophobic sequence). A lack of suitably positioned Lys residues could force the splitting of difficult peptides into smaller, more soluble pieces that require additional ligations and purifications. Also, the attachment of these Lys-based linkers negates the cationic charge of the native ε-amine of Lys and introduces a partly hydrophobic Dde-based linker. To overcome the loss of this native positive charge and generate a solubilizing effect, multiple cationic groups must be installed (*e.g.*, poly-Lys).[34]

For these reasons, we sought to create an additional 'helping hand' that would be orthogonal to our existing linkers and unlock new attachment points beyond Lys. Glu is also a common residue (6.6%)[7] and has several attributes that make it an attractive target as a new helping hand. First, the carboxylic acid side chain is easily derivatized and has been used to attach cargo to peptides.[33,38] Moreover, Glu sites, as compared with Asp sites, are less prone to cyclization byproducts such as aspartimide.[39] Finally, the carboxylic acid side chain of native Glu residues does not impart significant solubility enhancement under the acidic conditions commonly employed during RP-HPLC. Therefore, temporary modification at Glu residues is less likely to reduce solubility compared to Ddae or Ddap modification at Lys residues.

2. Results and discussion

2.1 Design and synthesis of a new Glu-based linker

To begin our search for a second helping hand design, we decided that a broadly applicable, semi-permanent linker should possess the following characteristics:

(1) Compatibility with standard Fmoc-SPPS conditions

(2) Stability to general CPS conditions (*e.g.*, acidic and neutral aqueous buffers with high denaturant concentrations)

(3) Selective attachment of the desired cargo to the peptide

(4) Numerous attachment points across a multitude of peptide sequences

(5) Removal conditions that are robust and traceless

(6) Scalable and straightforward linker synthesis.

Here, we introduce an allylic ester-based linker that satisfies all these requirements. The allylic ester (allyl) and its corresponding urethane derivative (for amine protection) allyloxycarbonyl (alloc) groups were first introduced by the Kunz group for their unique orthogonality to other traditional protecting approaches.[40,41] They can be removed under near-neutral conditions with a palladium-catalyzed transfer of the allyl group to various nucleophilic scavengers. These protecting groups are stable to both the basic conditions of Fmoc-removal and the subsequent trifluoroacetic acid (TFA) global deprotection and cleavage from resin. This exceptional degree of orthogonality has been exploited in many examples of peptide macrocyclization and glycopeptide synthesis.[42–47] Due to its facile removal, the allyl group has also found a multitude of uses in peptide chemistry. Albericio optimized the use of the alloc group to devise an α-amine protection scheme for SPPS in addition to the more popular Fmoc- and Boc-strategies.[48,49] Danishefsky created allyl- and alloc-based linkers to attach fixed guanidino moieties to study aggregation and optimize handling of a difficult peptide segment during their synthesis of glycosylated human erythropoietin.[38] This rich history encouraged us to pursue the allyl group as the foundation for our new linker, similar to our approach with Dde and our Lys-based helping hand. Unlike Ddae or Ddap linkers, the new Glu hand is synthesized as a pre-made Fmoc-protected amino acid. This new route allows direct incorporation of the linker through Fmoc-SPPS, eliminates the initial protecting group removal step required with Ddae or Ddap, and ensures complete linker attachment.

To increase linker accessibility, we set out to develop a scalable synthetic route to access Fmoc-Glu((*E*)-4-hydroxybut-2-en-1-yl 6-((1-(Dde)amino)hexanoate))-OH (Fmoc-Glu(AlHx)-OH) that uses commercially available reagents (Fig. 1). Overall, we prepared Fmoc-Glu(AlHx)-OH in four steps involving two purifications (see ESI for details). First, the amine of 6-aminohexanoic acid was protected with Dde-OH. This crude product was then activated and coupled with *trans*-2-butene-1,4-diol using Steglich esterification conditions.[50] Following purification, the resulting compound was coupled with the γ-carboxylate group of Fmoc-(L)-Glu-O*t*Bu. Finally, solvolysis of the α-*tert* butyl ester was effected by treatment with TFA to yield Fmoc-(L)-Glu(AlHx)-OH in 33% isolated (overall) yield.

2.2 Incorporation, stability, and removal of AlHx in a model peptide

In order to carefully characterize any undesired side reactions that may occur during CPS, Fmoc-Glu(AlHx)-OH was first incorporated in a short model peptide, C20-K9E. C20 is a

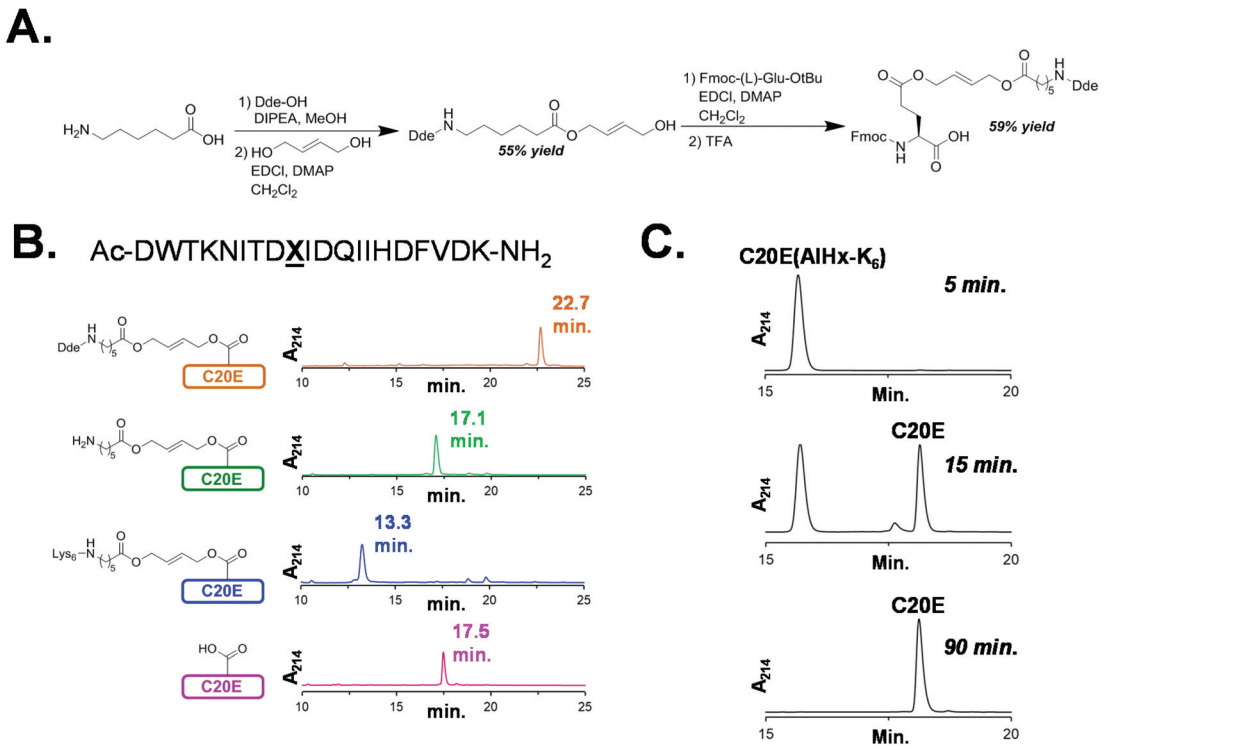

Fig. 1 Synthesis of Fmoc-Glu(AlHx)-OH and initial characterization in a model peptide. (A) The premade Fmoc-building block is synthesized in 4 steps with 2 purifications and 33% overall yield. (B) Comparison of the various forms of Glu(AlHx) in a modified model peptide, C20E. After on-resin Dde removal, the solubilizing Lys_6 tag can be added *via* standard Fmoc-SPPS. The AlHx-K_6 tag is removed in 6 M GuHCl with 25 mM $[Pd(allyl)Cl]_2$ and GSH. HPLC traces were generated using a gradient of 30–70% B over 25 min at 1 mL min^{-1}. (C) Removal of the AlHx-K_6 tag is complete in 90 min at pH 8 and 37 °C. HPLC traces were generated using a gradient of 10–90% B over 25 min at 1 mL min^{-1} (see ESI† for exact conditions).

20-residue peptide derived from the C-terminal heptad repeat region of the Ebola virus GP2 protein.[51] This peptide is a good model due to its solubility, diverse representation of amino acids, and use in benchmarking previous helping hand studies that incorporated Ddae or Ddap at Lys residues.[34,35] For this study, the native C20 sequence, Ac-DWTKNITDKIDQIIHDFVDK-NH_2, was modified to replace Lys9 with Glu to give C20-K9E (hereafter referred to as C20E), Ac-DWTKNITDEIDQIIHDFVDK-NH_2.

Fmoc-Glu(AlHx)-OH was cleanly incorporated into the peptide using standard Fmoc-SPPS coupling conditions (see ESI†). After the peptide N-terminus was acetylated, the Dde group was removed from the AlHx linker by iterative exposure to a 5% hydrazine solution in dimethylformamide (DMF) (3 × 5 min each). To address the potential alkene reduction by hydrazine, we added allyl alcohol (15% v/v) to the hydrazine solution, which has been shown to overcome this side reaction.[52,53] After Dde removal, a Lys_6 solubilizing tag was incorporated *via* Fmoc-SPPS, and the peptide was cleaved from resin using a standard TFA cleavage cocktail (95% TFA, 2.5% triisopropylsilane, and 2.5% water). Importantly, we observed no linker cleavage or side reactions during Fmoc-SPPS or acidic cleavage and global deprotection (ESI Fig. 8†). We purified C20E(AlHx-K_6) *via* RP-HPLC, and the lyophilized peptide was then used for a panel of stability and AlHx removal studies (ESI Fig. 9†).

To evaluate the general utility of AlHx, we next subjected C20E(AlHx-K_6) to a variety of common CPS conditions, most of which contain the denaturant guanidine hydrochloride (GuHCl). The peptide was dissolved at a concentration of 1 mM in the following solutions:

A: 0.1% TFA in 1/1 acetonitrile/water (HPLC buffer)

B: 6 M GuHCl, 100 mM $NaPO_4$, pH 3 (activation buffer)

C: 6 M GuHCl, 100 mM $NaPO_4$, pH 7 (ligation buffer)

D: 6 M GuHCl, 100 mM $NaPO_4$, 200 mM $MeONH_2$, pH 4 (Thz removal conditions)

E: 6 M GuHCl, 100 mM $NaPO_4$, 200 mM MPAA, 50 mM TCEP, pH 7 (NCL conditions).

We observed minimal (<5%) linker cleavage in any of these reaction conditions after 48 hours at room temperature, indicating a high level of stability of this new linker (ESI Fig. 10–14†).

After the AlHx linker was shown to be stable in various CPS conditions, we next sought to analyze and optimize the aqueous removal conditions. Ideally, we envisioned conditions that are compatible with high levels of denaturant in order to allow one-pot NCL and AlHx removal without additional HPLC purifications. In organic solvents and during SPPS, allyl removal is commonly performed using a Pd^0 complex, such as tetrakis(triphenylphosphine)palladium(0) ($Pd[P(Ph_3)]_3$), with silane as a reductant. Although similar solution conditions have been used in strong polar solvents like DMSO,[38] we chose

to follow recent reports from the Brik[54] and Okamoto[55] groups, both of which developed fully aqueous conditions for allyl removal from Glu and Asp side chains. We observed complete removal of AlHx-K_6 after 90 min in 6 M GuHCl, 100 mM Na_2PO_4, pH 8 at 37 °C using 25 mM $[Pd(allyl)Cl]_2$ and reduced glutathione (GSH), which function as the catalyst and scavenger for the AlHx-K_6, respectively (Fig. 1). After AlHx-K_6 removal, we added 60 mM 1,4-dithiothreitol (DTT) to precipitate the Pd.

2.3 Total synthesis of the cyclic bacteriocin AS-48 with AlHx

Our study with model peptide C20E demonstrated that Fmoc-Glu(AlHx)-OH can be readily incorporated and removed for CPS endeavors, and that Glu residues can potentially be targeted to improve the solubility and purification of peptide intermediates. Therefore, we set out to demonstrate the utility of Fmoc-Glu(AlHx)-OH incorporation for the synthesis of a challenging protein target with known solubility issues. To this end, we chose AS-48, a 70-residue circular bacteriocin that is a potent, broad-spectrum antimicrobial peptide produced by enterococci, such as *E. faecalis*.[56–58] Due to its cyclic and compact structure, which contains 5 α-helices and a hydrophobic core, AS-48 has a high degree of stability towards denaturants, heat, and protease degradation. Additionally, AS-48's activity and safety profile has continued to be extensively studied, making it a promising candidate for applications in the medical and veterinary fields, as well as a non-chemical food preservative.[59–61] Despite its small size, AS-48 has been documented by both the Tam and Bode groups as a difficult target to produce synthetically.[62,63] It has notable amphipathic character with a cluster of cationic residues in one section of the primary sequence and folded structure (Fig. 2). The synthetic difficulty stems from the inner region (Phe5-Ala45), which has a high percentage (~50%) of hydrophobic residues and no cationic residues (Lys, Arg, or His).

As evidence of this synthetic challenge, there are no published strategies that prepare AS-48 by NCL. For example, Tam attempted a 3-segment NCL strategy to break up this hydrophobic, aggregation-prone region.[62] This approach was abandoned due to the insolubility of peptide intermediates, and they instead opted for a butelase-mediated cyclization of the crude linear precursor of AS-48. While this strategy was successful, it required pre-folding of the linear precursor, an Asn residue for the enzymatic ligation site, and the temporary addition of a Lys_3 sequence to help with solubility (the Lys_3 is

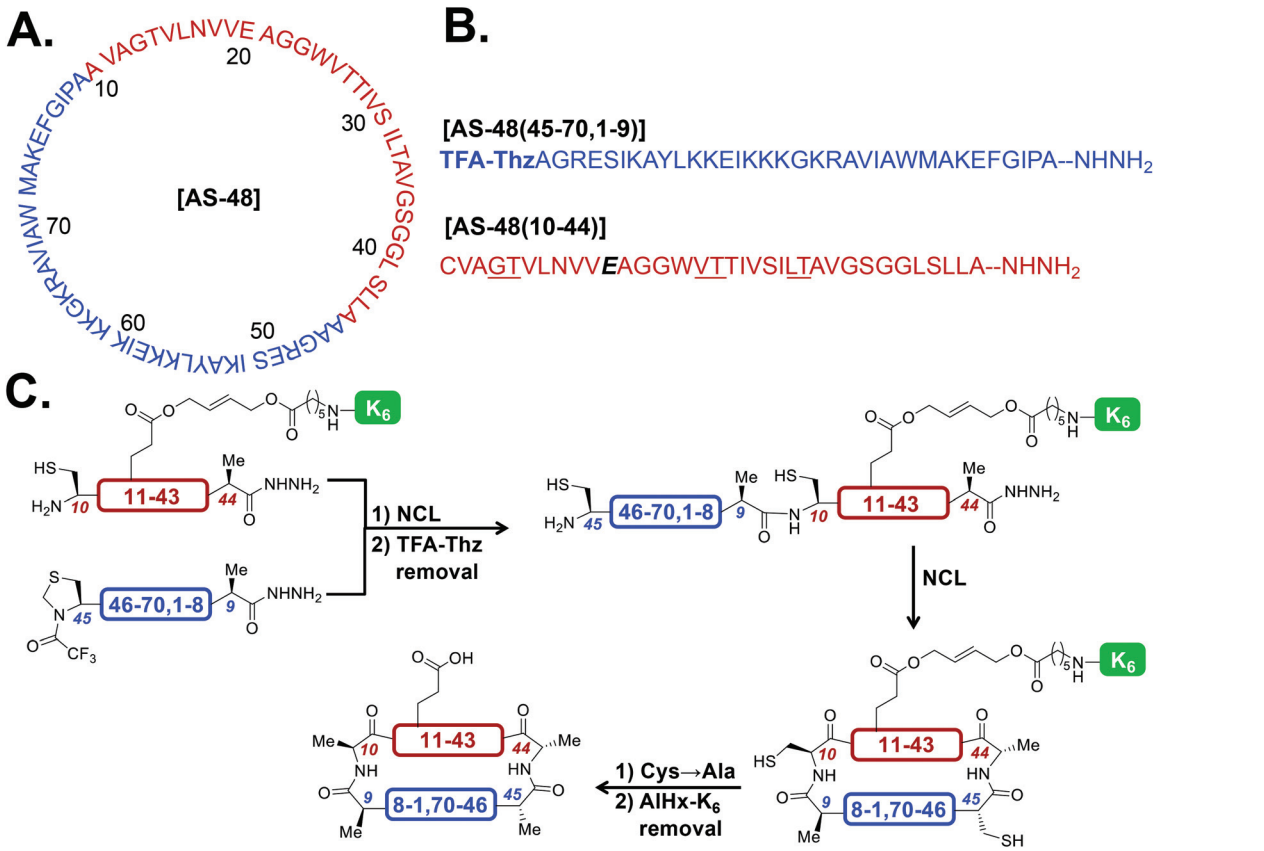

Fig. 2 Total synthesis of AS-48 with the help of Glu(AlHx). (A) The cyclic AS-48 sequence was broken into two 35-residue segments. Glu20 was chosen as the attachment point for AlHx due to its position in the hydrophobic, aggregation-prone segment of AS-48. (B) Ala10 and Ala45 were mutated to Cys, and Ala45 was protected with TFA-Thz to prevent intramolecular cyclization after activation of the hydrazide. AS-48(10–44) was synthesized and analyzed only after several pseudoprolines (underlined) and Glu(AlHx-K_6) (bolded and italicized) were added. (C) AS-48 synthetic scheme.

removed upon cyclization). Unfortunately, this method is also incompatible with high denaturant concentrations.

The Bode group employed α-ketoacid-hydroxylamine (KAHA) ligation to synthesize AS-48.[63] Although they were able to produce the bacteriocin using two peptide segments, homoserine mutations at Thr26 and Ser50 were required. They also noted challenges with the hydrophobic AS-48 region: initial attempts to ligate the segments in standard KAHA solvents (9 : 1 DMSO : H_2O) failed. Additional technical modifications, such as the use of cosolvents (added NMP) and heating (up to 95 °C) did not ameliorate the problem. Only after switching to a highly organic solvent system (1 : 1 HFIP : AcOH) did the authors observe the ligated product. Despite the advantages of the KAHA ligation strategy (*e.g.*, amenability to organic solvents for difficult peptide segments), NCL remains the most widely used ligation method in CPS, and we sought to demonstrate the broad utility of AlHx by synthesizing AS-48 *via* NCL.

First, we chose to mutate Ala10 and Ala45 to Cys in order to select favorable NCL ligation junctions. Due to the cyclic nature of AS-48, two NCL reactions are required. The first ligation will produce the full-length linear AS-48 sequence, and the second NCL will cyclize the peptide. We split AS-48 into two 35-residue segments that would be prepared as peptide hydrazides (Fig. 2). The hydrazide functions as an effective cryptothioester and can be converted *in situ* into the active thioester through $NaNO_2$ oxidation and subsequent thiolysis.[64] We chose Glu20 as the location for Glu(AlHx) insertion, as it is ideally placed in the middle of the hydrophobic region of AS-48. Moreover, this positioning does not break up the known hydrophobic region as reported in Tam and Bode's ligation strategies.[62,63] To prevent unwanted cyclization of AS-48 (45–70, 1–9) during the first NCL reaction, we chose to temporarily mask its Cys residue with a thiazolidine (Thz) derivative, TFA-Thz (Fig. 2).[65] This protected Cys can withstand hydrazide oxidation during treatment with $NaNO_2$ without affecting Thz oxidation (*e.g.*, Thz nitric oxide adduct).[64–66] After ligation, the trifluoroacetamide can be hydrolyzed using neutral to slightly basic aqueous conditions to yield Thz. The Cys residue can then be liberated using $MeONH_2$ for an efficient one-pot ligation-deprotection strategy.

With our strategy in hand, we first synthesized both segments with standard Fmoc-SPPS reagents and conditions. The hydrophilic segment, AS-48 (45–70, 1–9), was successfully synthesized and purified (Fig. 3, ESI Fig. 17 and 18†). As predicted, we were unable to analyze the native hydrophobic segment, AS-48 (10–44), due to its insolubility in a variety of aqueous and organic solvents, even in high levels of denaturant. Only after incorporating several pseudoprolines (Fig. 2) and the Glu(AlHx) linker with Lys_6 tag were we able to analyze this segment *via* LC/MS and RP-HPLC (ESI Fig. 15†). Interestingly, we observed that after removing the Dde from AlHx, but before the addition of the Lys_6 tag, we were able to achieve dissolution and analysis of AS-48 (10–44) in 50% ACN/H_2O with 0.1% TFA. After attachment of the Lys_6 tag to the AlHx linker with our standard SPPS coupling conditions, we easily dissolved AS-48 (10–44) in 10% ACN/H_2O with 0.1% TFA. AS-48 (10–44) was purified to homogeneity and lyophilized after RP-HPLC (Fig. 3, ESI Fig. 16†).

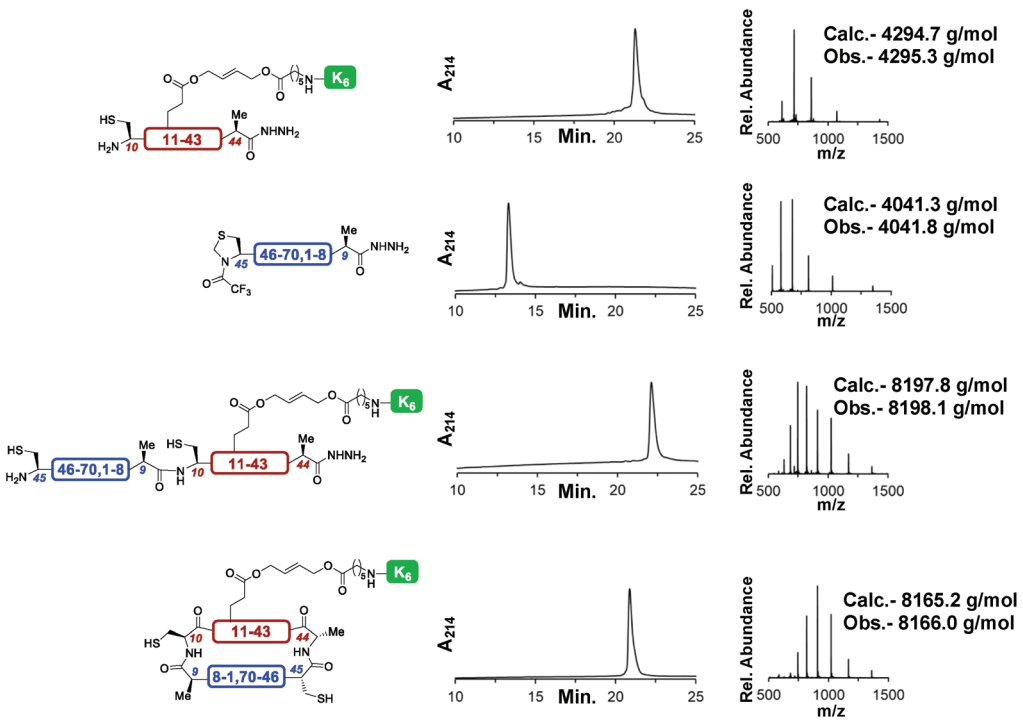

Fig. 3 Analytical HPLC and MS traces of purified peptides in AS-48 synthesis. HPLC traces were generated using a gradient of 10–90% B over 25 min at 1 mL min^{-1} and mass spectra (LC/MS) were generated using a gradient of 5–90% B over 7 min at 0.5 mL min^{-1} (see ESI† for exact conditions).

AS-48 (45–70, 1–9) was first converted into the corresponding thioester by sodium nitrite oxidation and MPAA thiolysis. We then added AS-48 (10–44) with AlHx-K_6 to a final concentration of 1.4 mM and 2 mM AS-48 (45–70, 1–9) in pH 7 ligation buffer. The ligation was complete in 1 hour; however, we began to observe trifluoroacetamide hydrolysis during the NCL reaction at neutral pH (ESI Fig. 19†). We opted to let the reaction continue overnight to remove the trifluoroacetamide under neutral conditions, rather than more basic conditions (pH 10).[65] After complete removal of the trifluoroacetyl group, we hydrolyzed the Thz ring by treatment with 200 mM $MeONH_2$ for 3 hours at pH 4 (ESI Fig. 21†). Importantly, we did not observe any premature cleavage of AlHx-K_6 under any of these conditions.[34,35]

After purification of the full-length AS-48 (45–70, 1–44 linear) by RP-HPLC (17.9% yield) (Fig. 3, ESI Fig. 22†), we performed the next NCL reaction to yield the cyclized peptide, AS-48 (1–70 cyclic). Due to the intramolecular nature of this ligation, we were able to completely cyclize the peptide in 1 hour at a relatively low concentration of 0.2 mM (ESI Fig. 23†). We purified AS-48 (1–70 cyclic) *via* RP-HPLC in 37% yield (Fig. 3, ESI Fig. 24†). To minimize losses associated with preparative RP-HPLC purification, we performed the subsequent desulfurization and AlHx-K_6 removal in one-pot as the final steps of this synthesis. First, we dissolved AS-48 (1–70 cyclic) in desulfurization buffer (1 mM) and then added VA-044, TCEP, and reduced GSH. After 4 hours at 37 °C, both Cys were fully converted to Ala (ESI Fig. 25†). We then removed the AlHx-K_6 linker by adding $[Pd(allyl)Cl]_2$ and fresh GSH (25 mM each) to successfully cleave the AlHx-K_6 linker in 1 hour at 37 °C (ESI Fig. 26†). Following RP-HPLC purification, we obtained the unfolded AS-48 in 14% isolated yield (Fig. 3, ESI Fig. 27†).

We folded 50 μM denatured AS-48 by overnight dialysis into 20 mM sodium phosphate at pH 6.5. The folded protein solution was analyzed using RP-HPLC, and LC/MS showed the correct mass of the natural protein (ESI Fig. 28†). Next, we obtained a circular dichroism (CD) spectrum of folded synthetic AS-48 at 25 μM in the folding buffer. As shown in Fig. 4, the synthetic AS-48 CD spectrum is in agreement with previously reported CD spectra and is indicative of proper AS-48 folding.[67–69]

As a final characterization, we verified the biological activity of our synthetic AS-48 by comparing it to a purified native (authentic) AS-48 sample (see ESI† for details). We first performed a spot-on-lawn assay against four Gram-positive bacteria susceptible to AS-48 and two Gram-negative AS-48 resistant strains (strains shown in Table 1). The results of the spot-on-lawn assay confirmed that the four Gram-positive bacteria selected for their sensitivity against native AS-48 were susceptible to the synthetic AS-48. As expected, we observed scarce activity against the two Gram-negative bacteria, whose inhi-

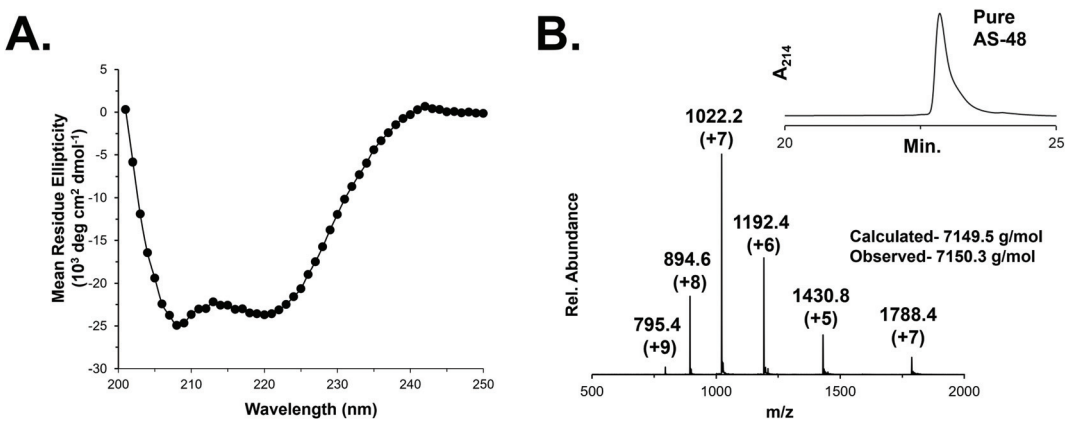

Fig. 4 Characterization of folded synthetic AS-48. (A) Circular dichroism spectrum of AS-48. (B) Analytical HPLC and MS trace of purified, folded AS-48. The HPLC trace was generated using a gradient of 10–90% B over 25 min at 1 mL min^{-1} and the mass spectrum (LC-MS) was generated using a gradient of 5–90% B over 7 min at 0.5 mL min^{-1} (see ESI† for exact conditions).

Table 1 Activity comparison of synthetic and authentic AS-48 samples

	Synthetic AS-48 MIC (μg mL^{-1})	Authentic AS-48 MIC (μg mL^{-1})
Gram-positive bacteria		
Listeria innocua CECT 4030	5.8	5.8
Enterococcus faecalis S-47	2.9	2.9
Staphylococcus aureus CECT 240	5.8	5.8
Arthrobacter sp.	0.36	0.36
Gram-negative bacteria		
Escherichia coli ATCC 25922	>93	>93
Pseudomonas putida	>93	>93

bition halos appeared as feeble cloudy zones. Once the activity of the synthetic AS-48 was confirmed, we compared the minimum inhibitory concentration (MIC) of both the synthetic and native bacteriocins (shown in Table 1), which show identical MIC values through two replicate studies across the 6-strain panel.

2.4. An alternative AlHx removal method

During our initial use of Glu(AlHx) in the total synthesis of AS-48, we noticed several characteristics of the existing Pd-catalyzed removal method that could be optimized. For instance, a large amount of $[Pd(allyl)Cl]_2$ (25 molar equivalents) is required, which can lead to yield loss due to peptide attachment to the Pd metal. Additionally, the DTT chelation step leads to partial Pd precipitation, although there is significant Pd remaining in solution that can precipitate later and complicate HPLC purifications.

We searched for other published examples of allyl removal and chose to examine a method from the Okamoto lab, who demonstrated efficient allyl and alloc removal on peptides in aqueous conditions.[55,70,71] In these studies, Pd was complexed with a water-soluble triphenyl phosphine ligand, triphenylphosphine-3,3′,3″-trisulfonic acid trisodium salt (TPPTS). To facilitate one-pot NCL/allyl or alloc removal, Okamoto utilized the already present MPAA to scavenge the allyl group and a final DTT treatment to deactivate the Pd metal without DTT-Pd precipitation. It is worth noting that they also reported rapid (10–20 min) allyl or alloc deprotection with only 2 equivalents of Pd.

With this new Pd method, we envisioned the use of a functionalized resin to chelate the Pd and remove it from solution. To this end, we employed thiourea-functionalized silica resin (THU-R) that is known to be an effective scavenger of Pd, and we verified that it can effectively remove Pd metal from ligation buffer. Serendipitously, we also observed that the THU-R could act as a dual scavenger to remove not only the Pd metal, but also the AlHx-K_6 linker from model peptides. Using our model peptide C20E(AlHx-K_6) (1 mM), we observed complete removal of the linker with 10 mM $Pd(TPPTS)_4$ in 15 min at 37 °C (Fig. 5). After 15 min, we added THU-R (10 equivalents) and incubated the solution for 5 min, centrifuged the sample, and collected the supernatant containing C20E. Compared to our original Pd removal protocol, this improved method is faster, requires less Pd, does not need thiol additives, and does not require a final DTT precipitation step. Because of these benefits, we envision this removal protocol will be the most convenient for most AlHx applications.

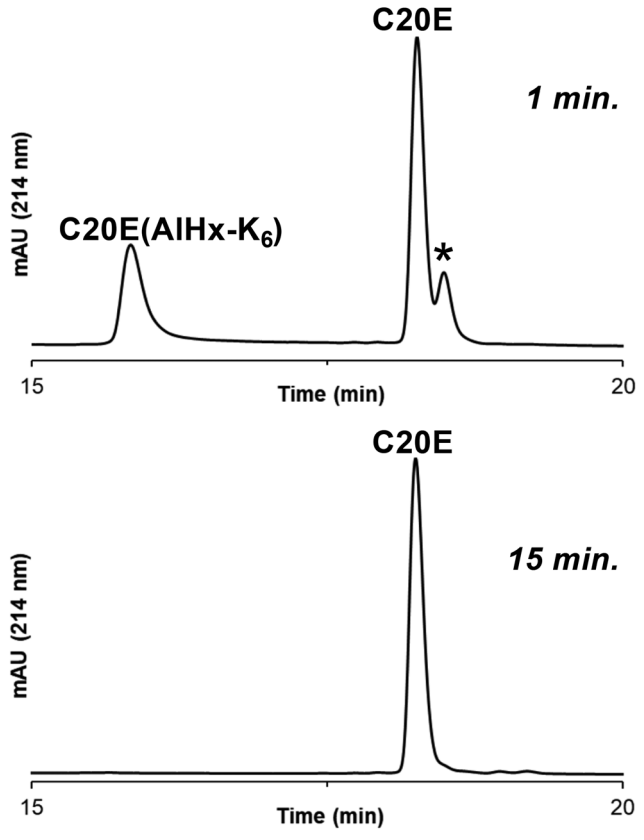

Fig. 5 Removal of the AlHx-K_6 tag with an alternative Pd-catalyzed method. In this protocol, 10 mM Pd complexed with 4 TPPTS ligands is used and the removal is complete in 15 minutes at 37 °C. RP-HPLC traces were generated using a gradient of 10–90% B over 25 min at 1 mL min^{-1} (see ESI† for exact conditions). In the top panel, the peak designated by (*) is likely a transient C20E species as it has the same mass as C20E. This peak was not further analyzed due to its disappearance as shown in the bottom panel.

3. Conclusions

In this work, we developed a novel, allyl-based Glu linker that is useful for the attachment of solubilizing tags during CPS. This modified Glu amino acid, Fmoc-Glu(AlHx)-OH, can be easily incorporated into peptides by standard SPPS conditions and removed in a traceless fashion by Pd-catalyzed reduction in fully aqueous conditions. Using Fmoc-Glu(AlHx)-OH, we demonstrated that AS-48 can be readily accessed by temporarily introducing a solubilizing poly-Lys tag to enable purification and ligation of a known difficult hydrophobic peptide segment. At the final stage of this AS-48 synthesis, the linker was easily removed, in one pot with desulfurization, to give the native sequence, which was successfully folded and demonstrated to be biologically equivalent to natural AS-48 through *in vitro* studies.

With the addition of Glu to our Lys helping hand toolkit, we anticipate greater coverage of peptide segments in CPS projects. Additionally, other solubilizing tools, such as Liu's reversible backbone modification group, can be used to enhance solubility in segments that lack Glu and Lys.[22,72] The combination of these solubilizing tags with other strategies (*e.g.*, pseudoprolines) will allow larger and difficult proteins to be synthesized that can address challenging biological questions.

To complement the expanding chemical tools available to peptide chemists, we have also designed the Automated

Ligator (Aligator),[73] a program that predicts the most efficient CPS strategies. We have now updated Aligator to consider the solubility enhancements provided by Glu(AlHx), or any other amino acid-based solubilizing strategy desired by the user. The improved Aligator is freely available on our GitHub page (https://github.com/kay-lab). Together, these tools expand the scope of CPS as larger and more complicated proteins that are unattainable by any other means become accessible by total chemical synthesis.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgements

The authors would like to thank Dr Debra Eckert, Dr Patrick Erickson, Dr Sarah Apple, and Zachary Cruz for experimental/material assistance. We also thank Dr Jack Skalicky and the U. of Utah Health Sciences NMR core for experimental assistance. Finally, we thank Dr Andrew Roberts and Dr Patrick Erickson for critical review of this manuscript. Funding for this work was provided by NIH grants P50-AI150464 and R01-AI076168 (to M. S. K.) and T32-GM122740 (to P. S.).

References

1 P. E. Dawson, T. W. Muir, I. Clark-Lewis and S. B. Kent, *Science*, 1994, **266**, 776–779.
2 L. Z. Yan and P. E. Dawson, *J. Am. Chem. Soc.*, 2001, **123**, 526–533.
3 V. Agouridas, O. El Mahdi, V. Diemer, M. Cargoet, J. M. Monbaliu and O. Melnyk, *Chem. Rev.*, 2019, **119**, 7328–7443.
4 R. B. Merrifield, *Biochemistry*, 1964, **3**, 1385–1390.
5 J. W. Bode, *Acc. Chem. Res.*, 2017, **50**, 2104–2115.
6 H. Liu and X. Li, *Acc. Chem. Res.*, 2018, **51**, 1643–1655.
7 C. UniProt, *Nucleic Acids Res.*, 2019, **47**, D506–D515.
8 Q. Wan and S. J. Danishefsky, *Angew. Chem., Int. Ed.*, 2007, **46**, 9248–9252.
9 T. S. Chisholm, S. S. Kulkarni, K. R. Hossain, F. Cornelius, R. J. Clarke and R. J. Payne, *J. Am. Chem. Soc.*, 2020, **142**, 1090–1100.
10 N. J. Mitchell, L. R. Malins, X. Liu, R. E. Thompson, B. Chan, L. Radom and R. J. Payne, *J. Am. Chem. Soc.*, 2015, **137**, 14011–14014.
11 P. S. Reddy, S. Dery and N. Metanis, *Angew. Chem., Int. Ed.*, 2016, **55**, 992–995.
12 M. Paradis-Bas, J. Tulla-Puche and F. Albericio, *Chem. Soc. Rev.*, 2016, **45**, 631–654.
13 R. J. Giesler, J. M. Fulcher, M. T. Jacobsen and M. S. Kay, in *Total Chemical Synthesis of Proteins*, 2021, pp. 185–209, DOI: 10.1002/9783527823567.ch7.
14 T. Haack and M. Mutter, *Tetrahedron Lett.*, 1992, **33**, 1589–1592.
15 T. Wöhr, F. Wahl, A. Nefzi, B. Rohwedder, T. Sato, X. Sun and M. Mutter, *J. Am. Chem. Soc.*, 1996, **118**, 9218–9227.
16 Y. Sohma, M. Sasaki, Y. Hayashi, T. Kimura and Y. Kiso, *Chem. Commun.*, 2004, 124–125, DOI: 10.1039/b312129a.
17 M. Mutter, A. Chandravarkar, C. Boyat, J. Lopez, S. Dos Santos, B. Mandal, R. Mimna, K. Murat, L. Patiny, L. Saucede and G. Tuchscherer, *Angew. Chem., Int. Ed.*, 2004, **43**, 4172–4178.
18 I. Coin, R. Dolling, E. Krause, M. Bienert, M. Beyermann, C. D. Sferdean and L. A. Carpino, *J. Org. Chem.*, 2006, **71**, 6171–6177.
19 E. C. Johnson and S. B. Kent, *Tetrahedron Lett.*, 2007, **48**, 1795–1799.
20 A. C. Baumruck, D. Tietze, L. K. Steinacker and A. A. Tietze, *Chem. Sci.*, 2018, **9**, 2365–2375.
21 T. Sato, Y. Saito and S. Aimoto, *J. Pept. Sci.*, 2005, **11**, 410–416.
22 J. S. Zheng, Y. He, C. Zuo, X. Y. Cai, S. Tang, Z. A. Wang, L. H. Zhang, C. L. Tian and L. Liu, *J. Am. Chem. Soc.*, 2016, **138**, 3553–3561.
23 C. Zuo, S. Tang, Y. Y. Si, Z. A. Wang, C. L. Tian and J. S. Zheng, *Org. Biomol. Chem.*, 2016, **14**, 5012–5018.
24 B. Zhang, Q. Deng, C. Zuo, B. Yan, X. X. Cao, T. F. Zhu, J. S. Zheng and L. Liu, *Angew. Chem., Int. Ed.*, 2019, **58**, 12231–12237.
25 X. Zhou, C. Zuo, W. Li, W. Shi, H. Wang, S. Chen, J. Du, G. Chen, W. Zhai, W. Zhao, Y. Wu, Y. Qi, L. Liu and Y. Gao, *Angew. Chem., Int. Ed.*, 2020, **59**, 15114–15118.
26 D.-L. Huang, C. Montigny, Y. Zheng, V. Beswick, Y. Li, X.-X. Cao, T. Barbot, C. Jaxel, J. Liang, M. Xue, C.-L. Tian, N. Jamin and J.-S. Zheng, *Angew. Chem., Int. Ed.*, 2020, **59**, 5178–5184.
27 S. K. Maity, G. Mann, M. Jbara, S. Laps, G. Kamnesky and A. Brik, *Org. Lett.*, 2016, **18**, 3026–3029.
28 S. K. Maity, M. Jbara, G. Mann, G. Kamnesky and A. Brik, *Nat. Protoc.*, 2017, **12**, 2293–2322.
29 S. Bondalapati, E. Eid, S. M. Mali, C. Wolberger and A. Brik, *Chem. Sci.*, 2017, **8**, 4027–4034.
30 S. A. Abboud, M. Doudeau, H. Bénédetti and V. Aucagne, *Chem. Sci.*, 2021, **12**, 3194–3201.
31 S. Tsuda, M. Mochizuki, H. Ishiba, K. Yoshizawa-Kumagaye, H. Nishio, S. Oishi and T. Yoshiya, *Angew. Chem., Int. Ed.*, 2018, **57**, 2105–2109.
32 S. Tsuda, S. Masuda and T. Yoshiya, *Org. Biomol. Chem.*, 2019, **17**, 1202–1205.
33 S. Tsuda, S. Masuda and T. Yoshiya, *ChemBioChem*, 2019, **20**, 2063–2069.
34 M. T. Jacobsen, M. E. Petersen, X. Ye, M. Galibert, G. H. Lorimer, V. Aucagne and M. S. Kay, *J. Am. Chem. Soc.*, 2016, **138**, 11775–11782.
35 J. M. Fulcher, M. E. Petersen, R. J. Giesler, Z. S. Cruz, D. M. Eckert, J. N. Francis, E. M. Kawamoto, M. T. Jacobsen and M. S. Kay, *Org. Biomol. Chem.*, 2019, **17**, 10237–10244.

36 M. M. Disotuar, M. E. Petersen, J. M. Nogueira, M. S. Kay and D. H. Chou, *Org. Biomol. Chem.*, 2019, **17**, 1703–1708.
37 N. Zheng, P. Karra, M. A. VandenBerg, J. H. Kim, M. J. Webber, W. L. Holland and D. H.-C. Chou, *J. Med. Chem.*, 2019, **62**, 11437–11443.
38 Z. Tan, S. Shang and S. J. Danishefsky, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 4297–4302.
39 R. Subirós-Funosas, A. El-Faham and F. Albericio, *Tetrahedron*, 2011, **67**, 8595–8606.
40 H. Kunz and H. Waldmann, *Angew. Chem., Int. Ed.*, 1984, **23**, 71–72.
41 H. Kunz and C. Unverzagt, *Angew. Chem., Int. Ed.*, 1984, **23**, 436–437.
42 M. Pelay-Gimeno, F. Albericio and J. Tulla-Puche, *Nat. Protoc.*, 2016, **11**, 1924–1947.
43 S. A. Kates, B. G. de la Torre, R. Eritja and F. Albericio, *Tetrahedron Lett.*, 1994, **35**, 1033–1034.
44 Y. Nakahara, S. Ando, M. Itakura, N. Kumabe, H. Hojo, Y. Ito and Y. Nakahara, *Tetrahedron Lett.*, 2000, **41**, 6489–6493.
45 Y. Asahina, T. Kawakami and H. Hojo, *Eur. J. Org. Chem.*, 2019, 1915–1920.
46 A. V. Vasco, M. Brode, Y. Mendez, O. Valdes, D. G. Rivera and L. A. Wessjohann, *Molecules*, 2020, **25**, 811.
47 T. Conroy, K. A. Jolliffe and R. J. Payne, *Org. Biomol. Chem.*, 2010, **8**, 3723–3733.
48 N. Thieriet, J. Alsina, E. Giralt, F. Guibé and F. Albericio, *Tetrahedron Lett.*, 1997, **38**, 7275–7278.
49 P. Gomez-Martinez, M. Dessolin, F. Guibé and F. Albericio, *J. Chem. Soc.*, 1999, 2871–2874, DOI: 10.1039/a906025a.
50 B. Neises and W. Steglich, *Angew. Chem., Int. Ed. Engl.*, 1978, **17**, 522–524.
51 W. Weissenhorn, A. Carfi, K. H. Lee, J. J. Skehel and D. C. Wiley, *Mol. Cell*, 1998, **2**, 605–616.
52 B. Rohwedder, Y. Mutti, P. Dumy and M. Mutter, *Tetrahedron Lett.*, 1998, **39**, 1175–1178.
53 M. Teixido, M. Altamura, L. Quartara, A. Giolitti, C. A. Maggi, E. Giralt and F. Albericio, *J. Comb. Chem.*, 2003, **5**, 760–768.
54 M. Jbara, E. Eid and A. Brik, *Org. Biomol. Chem.*, 2018, **16**, 4061–4064.
55 N. Kamo, G. Hayashi and A. Okamoto, *Chem. Commun.*, 2018, **54**, 4337–4340.
56 C. Gonzalez, G. M. Langdon, M. Bruix, A. Galvez, E. Valdivia, M. Maqueda and M. Rico, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 11221–11226.
57 M. Maqueda, A. Galvez, M. M. Bueno, M. J. Sanchez-Barrena, C. Gonzalez, A. Albert, M. Rico and E. Valdivia, *Curr. Protein Pept. Sci.*, 2004, **5**, 399–416.
58 M. Sanchez-Hidalgo, M. Montalban-Lopez, R. Cebrian, E. Valdivia, M. Martinez-Bueno and M. Maqueda, *Cell. Mol. Life Sci.*, 2011, **68**, 2845–2857.
59 M. Montalban-Lopez, M. Sanchez-Hidalgo, E. Valdivia, M. Martinez-Bueno and M. Maqueda, *Curr. Pharm. Biotechnol.*, 2011, **12**, 1205–1220.
60 T. del Castillo-Santaella, R. Cebrián, M. Maqueda, M. J. Gálvez-Ruiz and J. Maldonado-Valderrama, *Food Chem.*, 2018, **246**, 249–257.
61 M. J. Grande Burgos, R. P. Pulido, M. Del Carmen Lopez Aguayo, A. Galvez and R. Lucas, *Int. J. Mol. Sci.*, 2014, **15**, 22706–22727.
62 X. Hemu, Y. Qiu, G. K. Nguyen and J. P. Tam, *J. Am. Chem. Soc.*, 2016, **138**, 6968–6971.
63 F. Rohrbacher, A. Zwicky and J. W. Bode, *Chem. Sci.*, 2017, **8**, 4051–4055.
64 G. M. Fang, J. X. Wang and L. Liu, *Angew. Chem., Int. Ed.*, 2012, **51**, 10347–10350.
65 Y.-C. Huang, C.-C. Chen, S. Gao, Y.-H. Wang, H. Xiao, F. Wang, C.-L. Tian and Y.-M. Li, *Chem. – Eur. J.*, 2016, **22**, 7623–7628.
66 M. D. Simon, P. L. Heider, A. Adamo, A. A. Vinogradov, S. K. Mong, X. Li, T. Berger, R. L. Policarpo, C. Zhang, Y. Zou, X. Liao, A. M. Spokoyny, K. F. Jensen and B. L. Pentelute, *ChemBioChem*, 2014, **15**, 713–720.
67 R. Cebrian, M. Maqueda, J. L. Neira, E. Valdivia, M. Martinez-Bueno and M. Montalban-Lopez, *Appl. Environ. Microbiol.*, 2010, **76**, 7268–7276.
68 M. Montalban-Lopez, B. Spolaore, O. Pinato, M. Martinez-Bueno, E. Valdivia, M. Maqueda and A. Fontana, *FEBS Lett.*, 2008, **582**, 3237–3242.
69 M. Sanchez-Hidalgo, M. Martinez-Bueno, A. M. Fernandez-Escamilla, E. Valdivia, L. Serrano and M. Maqueda, *J. Antimicrob. Chemother.*, 2008, **61**, 1256–1265.
70 N. Kamo, G. Hayashi and A. Okamoto, *Angew. Chem., Int. Ed.*, 2018, **57**, 16533–16537.
71 N. Kamo, G. Hayashi and A. Okamoto, *Org. Lett.*, 2019, **21**, 8378–8382.
72 S. Tang, C. Zuo, D. L. Huang, X. Y. Cai, L. H. Zhang, C. L. Tian, J. S. Zheng and L. Liu, *Nat. Protoc.*, 2017, **12**, 2554–2569.
73 M. T. Jacobsen, P. W. Erickson and M. S. Kay, *Bioorg. Med. Chem.*, 2017, **25**, 4946–4952.