Going round in circles with N→S acyl transfer

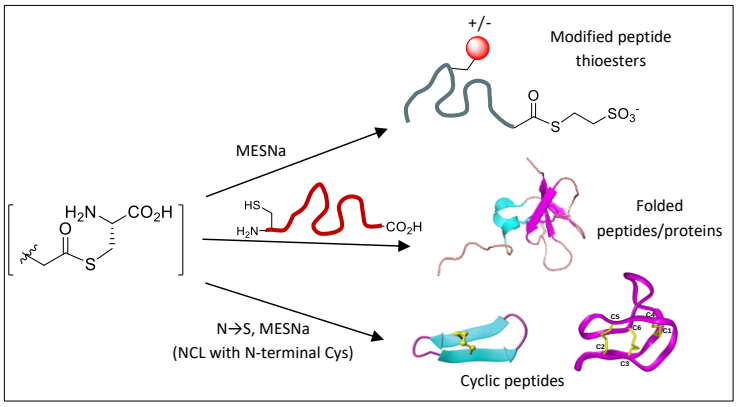
***Derek Macmillan**

[a] Department of Chemistry, University College London, 20 Gordon Street, London, WC1H 0AJ, UK.

* indicates the main/corresponding author.

d.macmillan@ucl.ac.uk

Click here to insert a dedication.



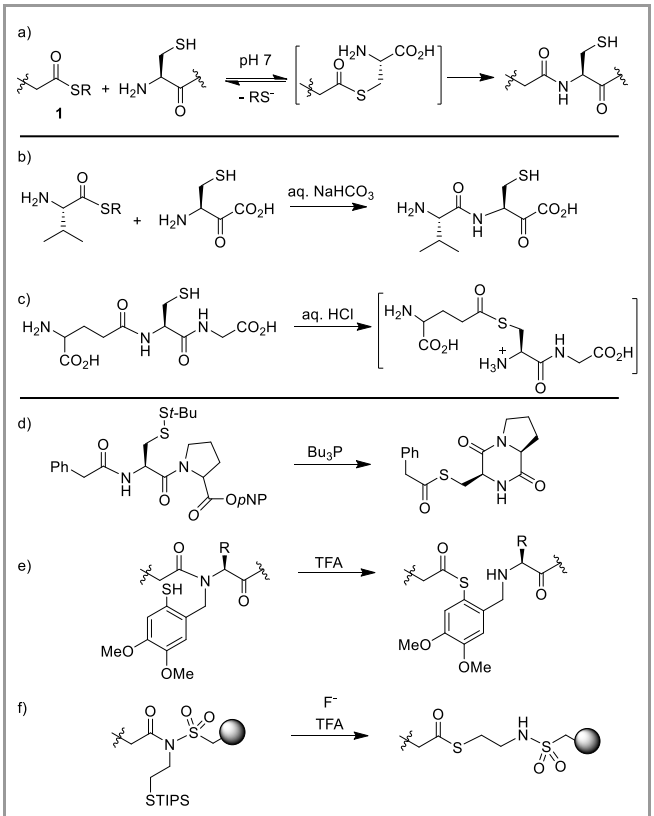
Received:
Accepted:
Published online:
DOI:

Abstract It is not highly sophisticated, yet the N→S acyl transfer reaction of a native peptide sequence potentially fills an important technology gap. While several routes to synthetic peptide thioesters exist, only one is routinely applicable for biologically derived samples. Using the naturally occurring amino acid cysteine as the sole activator for N→S acyl transfer we have demonstrated transformation of synthetic and biologically derived precursors into thioesters for use in Native Chemical Ligation, providing a viable alternative for biological samples. Further refinement will be key to realizing the full potential of this intriguing process, and increase the number of applications in peptide engineering and therapeutics.

Key words Thioester, Native Chemical Ligation, cyclic peptides, acyl transfer

1 Introduction

The thioester has attracted the interest of scientists for decades and due to its ubiquitous presence in cellular metabolism it is considered highly relevant to prebiotic chemistry and the origins of life.[1] The peptide C_α-thioester (**1**) also sits centrally as a key tool for protein synthesis and semi-synthesis using Native Chemical Ligation (NCL, Scheme 1a).[2] The early observations of Wieland (Scheme 1b),[3] the advent of efficient automated solid phase-peptide assembly protocols,[4] and the development of NCL by the Kent group have brought small proteins under the command of the synthetic organic chemist,[5] and it is the ability to routinely engineer proteins at atomic level that has allowed their structure and function to be dissected in unprecedented detail. As with NCL, the reverse process involving N→S acyl transfer (Scheme 1c) had been of interest, and subject to considerable study, over half a century ago.[6] However it was not until much later, after the observation of N→S acyl transfer mediated by protein splicing elements called inteins,[7] that the compatibility of this transformation with more complex protein components was explored in detail. Pioneering research by Zanotti,[8] Vorherr and Aimoto,[9, 10] Melnyk (Scheme 1d, 1e, and 1f respectively),[11] and others had shown that N→S acyl shift could be initiated to form thioesters under acidic or mildly basic conditions, highlighting the potential of this process for peptide thioester synthesis.[9, 12]



Scheme 1 a) Mechanism of Native Chemical ligation (NCL) b) An NCL-like active ester condensation between a valine thioester and cysteine and c) The reverse process studied in the context of glutathione. Both NCL and N→S acyl transfer were investigated in the 1950's.[3, 6] d)-f) Pioneering thioester syntheses, utilising N→S acyl transfer developed in subsequent years.[8, 9, 11] *p*-NP = *para*-nitrophenyl, TFA = trifluoroacetic acid, TIPS = triisopropylsilyl.

Further study of inteins demonstrated that the seemingly unfavourable N→S acyl shift was achieved through a combination of highly conserved catalytic residues and a distorted amide bond geometry,[13, 14] and although the precise details of amide bond geometry in an active intein remain unclear,[15] the notion that the scissile peptide bond requires distortion or weakening (ground state destabilization) through ancillary structural or stereoelectronic effects has served as the working model for the most recent synthetic developments.

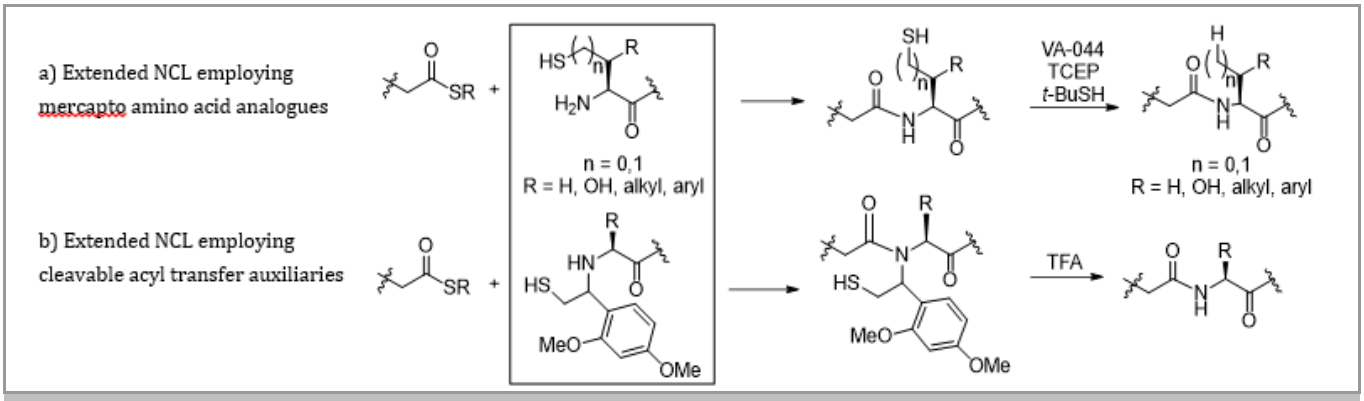
In this short account we discuss our own unexpected exploration of N→S acyl transfer in native peptide sequences. It was prompted by the observation of amide bond cleavage at Xaa-Cys motifs, with accompanying thioester formation, upon heating in aqueous mercaptopropionic acid (MPA). Rather than recounting the details of our findings to date we hope to highlight some key advantages of employing this simple methodology, and its relevance to NCL. At the same time we emphasise some of the challenges that remain and suggest how they could be overcome in the near future.

Since its introduction,[16] practitioners of NCL have sought to improve and simplify the process. Initially this was based around finding efficient methods for the synthesis of the required thioester and cysteinyl components[14, 17, 18, 19] but as the methodology developed the emphasis shifted towards performing multiple ligation reactions sequentially in one pot, and minimising handling steps.[20] These streamlining processes typically comprise one pot kinetically controlled ligations, [21, 22] combined with *in-situ* removal of protecting groups[22, 23] and desulfurisation.[24] The desire to conduct *in-situ* desulfurisation was prompted by the rapid growth, and effective use, of mercapto amino acid derivatives capable of performing NCL-reactions,[25] yet are reverted to non-cysteinyl amino acids upon reduction under mild conditions (Scheme 2a).[26] This advance circumvents the often stated limitation that NCL requires cysteine, a relatively rare amino acid, at the ligation junction. The issue was also previously tackled through the use of removable acyl transfer auxiliaries (Scheme 2b)[27] but these appendages were often criticized for their inability to perform NCL efficiently across ligation junctions other than Gly-Gly. In our own laboratory we found auxiliaries to work effectively across several Xaa-Gly and Gly-Xaa junctions so long as typical NCL thiol additives were removed from the reaction.[28] Nevertheless this perceived difficulty, along with their poor atom economy and known potential to reverse amide bond formation (Scheme 1e),[9, 29] has fuelled the search for alternatives even though recent developments that combine their relative ease of synthesis with improved performance across an increased variety of ligation junctions holds promise.[30]

Despite the growing number of options available to the synthetic chemist, NCL practitioners, who are often biologists, may seek a less sophisticated approach, one that does not require specific expertise or specialized apparatus. Even chemists can become frustrated by the poor availability and high cost of amino acid analogues or acyl transfer auxiliaries. Furthermore their added substitution (relative to cysteine) means they don't "perform" like cysteine in NCL reactions and their behaviours can be laced with idiosyncrasy, or necessitate the use of protecting groups to prevent off-target side-reactions. Meanwhile conservative point mutations are often well tolerated by proteins, and so the strategic substitution of non-catalytic residues in order to facilitate NCL is also commonplace and consequently, introduction of additional "rogue" cysteine residues (which can perhaps be capped at a later stage) may still present the more attractive option. Such corner-cutting often means that chemical yield or absolute replication of a natural structure are compromised in order to progress from A to something that closely resembles B in a timely and cost-effective manner, but the "cost" of such compromise ultimately depends on the experimental question being asked, and may be negligible. If a favourable outcome is achieved then it is clear that, on balance, practical considerations far outweigh perfection. Indeed, prior to the development of 4-mercaptophenyl acetic acid (MPAA) as an NCL catalyst,[31] many researchers would more routinely employ sodium 2-mercaptoethanesulfonate (MESNa) as an NCL additive in place of the superior thiophenol because of its low odour, high water solubility, reduced requirement for handling in an inert atmosphere, despite its relatively poor reaction kinetics.[32] So, in the right setting, simplicity can trump sophistication and efficiency, setting the scene for our investigations concerning N→S acyl transfer.

2 N→*S* acyl transfer in "normal" peptide sequences.

While streamlined NCL protocols that minimise sample loss resulting from repeated purification steps are crucial, production of the required thioester component has remained the subject of continuous innovation and development too. Our entry into this area began out of an interest in the compatibility, or lack thereof, between thioester synthesis and glycopeptides.

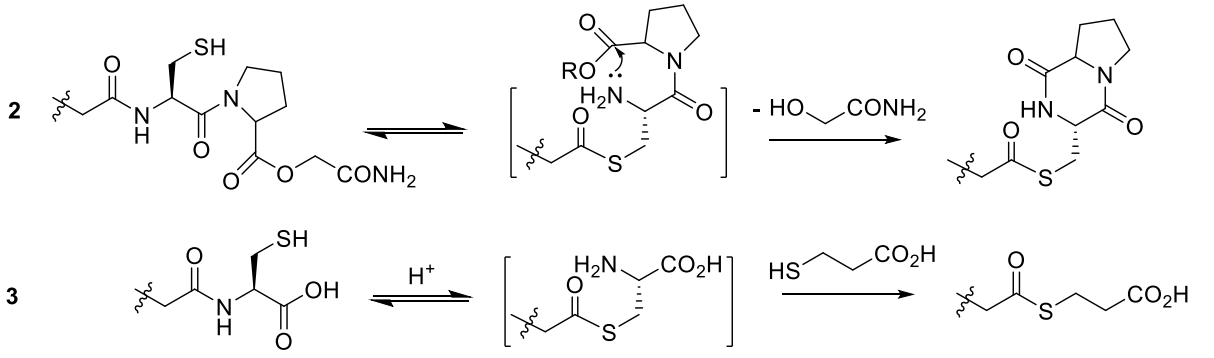


Scheme 2 Common extensions of the NCL methodology, designed to abrogate the need for cysteine at the N-terminus of one peptide component

The superior quality and efficiency of *tert*-butyloxycarbonyl (Boc)-based thioester synthesis had been well documented[33] but the harsh cleavage protocol (usually anhydrous HF) rendered Boc-based approaches incompatible with certain glycosides, and the high toxicity of HF combined with the requirement for specialist apparatus proved unattractive. This obstacle has largely been overcome by the recent introduction of elegant non-HF based cleavage protocols[34] but 9-fluorenylmethoxycabonyl (Fmoc)-based glycopeptide syntheses currently remain the most common.[35] While exceptions exist,[36] Fmoc thioester syntheses are usually indirect since the thioester linkage itself is not compatible with piperidine mediated Fmoc cleavage,[37] and the thioester is introduced upon cleavage from the solid support or post-synthesis. We initially experimented with the sulphonamide "safety-catch" approach which produced several long (>30 residues) glycopeptide thioesters cleanly and reliably,[38] but we also often grappled with the chemistry when thioester recovery was poor. Although helpful methods for monitoring resin activation were emerging[39] we craved an approach that was relatively easy to initiate, monitor and cleave, so turned towards post-synthesis N→S acyl transfer reactions. The highest yielding peptide syntheses, in general, are those that furnish products as C- terminal carboxylates (e.g. using Wang resin) and carboxamides (e.g. using Rink amide resin) and post-synthesis thioester formation allowed precursors to be prepared in this way, so we attempted the method introduced by Zanotti (Scheme 1d) but optimised for peptide thioester synthesis by Aimoto and co-workers (Scheme 3).[40] We reasoned that, even if thioester formation turned out to be problematic we would at least have ample material with which to optimise the process. Using the cysteinyl prolyl ester (CPE, **2**) the unfavourable N→S acyl transfer is driven by sequestering the liberated α-amino group through rapid intramolecular diketopiperazine (DKP) formation in the presence of the proximal proline ester. Although peptide synthesis was conducted on Rink amide resin the resin-loading protocol is rather protracted. The Xaa-Cys motif (where Xaa is potentially any amino acid residue) must be prepared separately and loaded as a dipeptide in order to prevent spontaneous DKP formation upon Fmoc deprotection from a resin bound Cys-Pro ester. Nevertheless, precursor peptides were easily assembled and purified in good isolated yields. However, upon exposure of a model peptide to typical reaction conditions (phosphate buffer; pH7, MESNa) little thioester was obtained unless an additional internal Cys residue was protected. Enlightened by a further report of how an unreactive thioester precursor, similar to **2**, could be salvaged employing 40% aqueous 3-mercaptopropionic acid (MPA) with microwave heating (60-80 °C),[41] we subjected our model peptide to these conditions, resulting in cleavage of the peptide across an internal Gly-Cys motif (**3**) in addition to the terminal Cys-Pro-ester.[42]

Despite the literature precedent we were still surprised by this observation because productive thioester formation from native peptide sequences in this manner had not been described previously. Furthermore, subjecting whole proteins (Bovine serum albumin and erythropoietin) to aqueous MPA at elevated temperature resulted in significant fragmentation across multiple Xaa-Cys sites with accompanying thioester formation. This observation was interesting because it hinted at potential application to biological samples even though native Xaa-Cys motifs were not considered sufficiently primed for fragmentation via N→S acyl transfer without the aid of an intein.

To explore the process further model reactions were conducted and monitored by HPLC and/or ^{13}C NMR spectroscopy.[10, 43] The commercial availability of ^{13}C-1 labelled glycine allowed the effect of varying reaction conditions to be assessed rapidly using Gly-Cys terminated precursors and the overall picture that emerged was even more intriguing. Not only did different Xaa-Cys motifs display varying aptitudes for thioester formation, but identical motifs (e.g Gly-Cys) were not equally reactive, depending on their sequence context (Figure 1).[44] In model studies several short peptide precursors containing particularly reactive C-terminal Xaa-Cys motifs (where Xaa = Gly, His or Cys) formed thioesters cleanly at 40-50 °C over a period of 24-48 h. In contrast, when the penultimate residue (Xaa) was a β-branched amino acid, thioester formation occurred slowly under forcing conditions (>60 °C), and was often accompanied by significant hydrolysis. Moreover, when a favourable Xaa-Cys motif was placed *within* a peptide sequence, rather than at the terminus, the reaction rate was reduced further. Surprisingly, even very small changes to the C-terminal Cys residue, when presented as a C-terminal ester or amide rather than carboxylate, caused the reaction rate to drop by approximately half under typical reaction conditions. The reason for the increased reactivity of the C-terminal Cys carboxylate may be due to the zwitterionic nature of the intermediate which increases the basicity of the liberated α-amino group and presents a barrier to reforming the starting material at acidic pH.[45] Recently the enhanced reactivity of thioester precursors adorned with a C-terminal carboxyl group



Scheme 3 Peptide Cα-thioester formation enabled by the cysteinyl prolyl ester (**2**)[40] and by a single cysteine residue (**3**).[42]

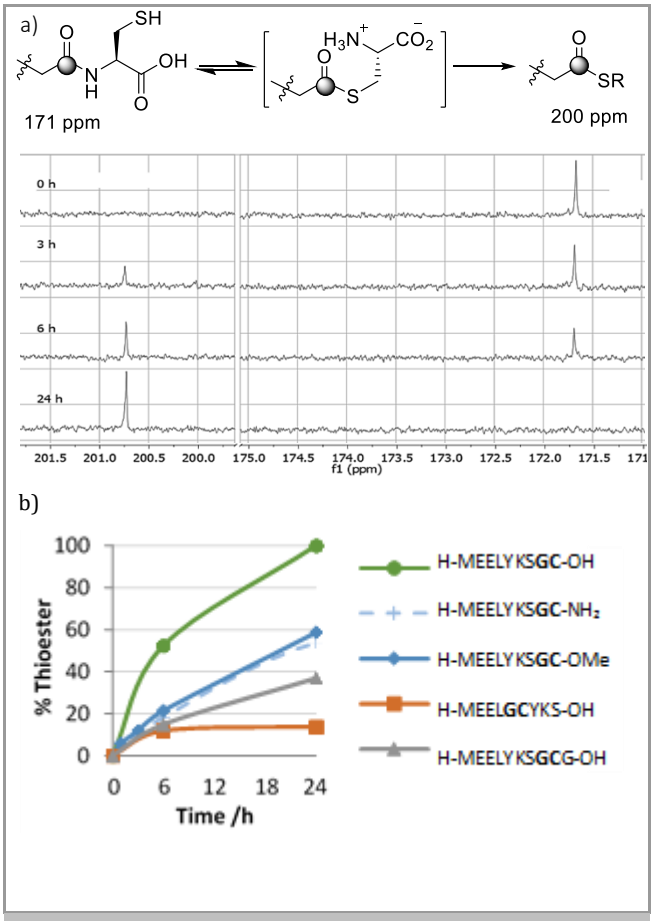


Figure 1 "Some are more equal than others". a) Following thioester formation using ^{13}C-1 labelled glycine and ^{13}C NMR spectroscopy. b) The effect of sequence context on the reactivity of a Gly-Cys motif as monitored by ^{13}C NMR at 60 °C in 10% w/v (~0.7 M) MESNa.

was also observed in the case of *N*-alkylated cysteine derivatives. In this case intramolecular catalysis by the carboxyl group is proposed since the scissile amide bond in *N*-alkyl Cys derivatives are non-planar and the amide nitrogen is considerably more basic.[46]

At first glance these findings could appear to impose significant limitations on the N→S acyl transfer reaction in that thioester formation may be restricted to a few favourable terminal Xaa-Cys motifs. Furthermore slow reactions cannot be easily accelerated by increased heating or heating time, because this is often met with a loss of selectivity, peptide precipitation and thioester hydrolysis. However the notion that simple peptide sequences adorned with a single C-terminal cysteine could undergo thioester formation with a high degree of inherent selectivity, and be compatible with chemically fragile post-translational modifications such as glycosyation[44, 47, 48] and phosphorylation[48] (Figure 2), led us to investigate further.

3 Reduced reactivity of internal Xaa-Cys motifs as an advantage in head-to-tail peptide cyclisation.

The difference in reactivity between identical yet differentially disposed (*terminal* vs *internal*) Xaa-Cys motifs additionally suggested that, if the thioester precursor contained an N-terminal cysteine, this feature could drive peptide cyclisation reactions. Head-to-tail peptide cyclisation to afford an "inert" product is an ideal application since, in forming a cyclic product, interference from side-reactions such as thioester hydrolysis are also minimised when cyclisation is rapid.[45] Although acyl transfer reactions utilising cysteine as the acyl transfer facilitator are often described as sluggish, we have found that several short peptides cyclise completely in 6-24 h at 40-50 °C in aqueous buffer.[45] Furthermore the reaction conditions are capable of facilitating effective cyclisation at Ser-Cys, Arg-Cys, Phe-Cys, and Leu-Cys motifs. Although we have demonstrated the cyclisation reaction countless times using short (usually anti-microbial) peptide substrates it is still remarkable that a linear peptide, in all its complexity, will undergo a selective amide bond cleavage followed by selective peptide coupling in water and in the absence of enzymes or chemical coupling reagents (Figure 3).[49] We have also applied head-to-tail

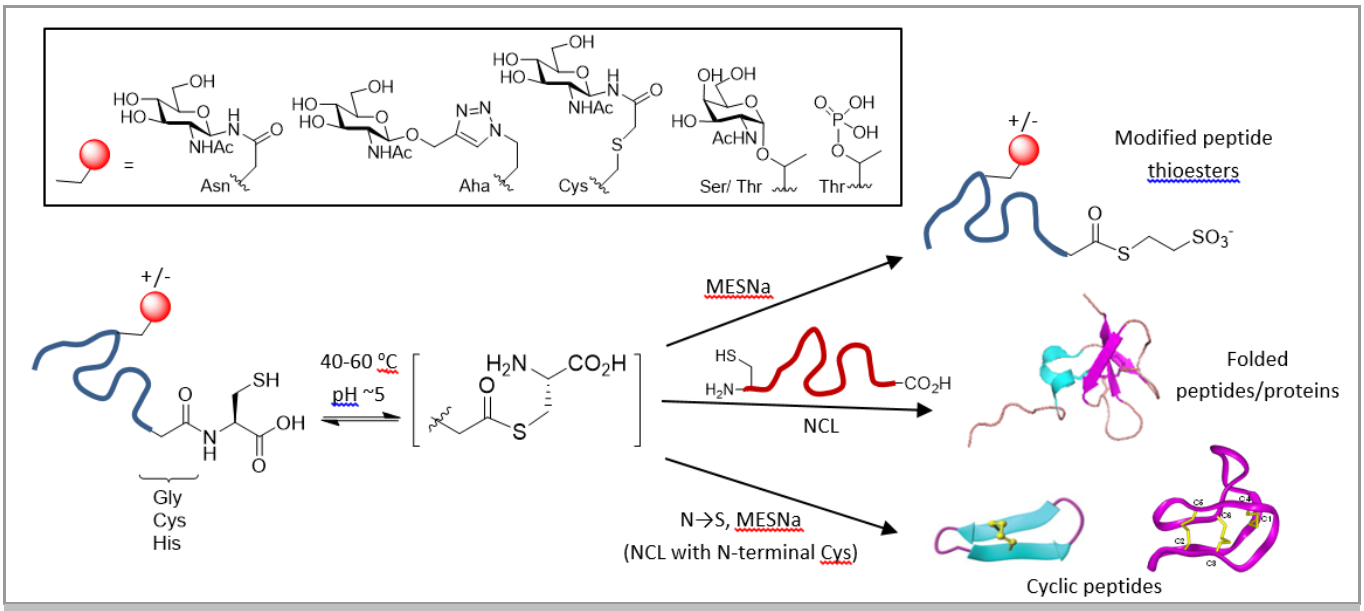


Figure 2 A selection of experiments performed on Xaa-Cys terminated peptides. Reaction with MESNa at 40-60 °C is sufficient to convert linear precursors into thioesters and cyclic peptides. For intermolecular NCL reactions the MESNa thioester is usually isolated first. Aha =azidohomoalanine.

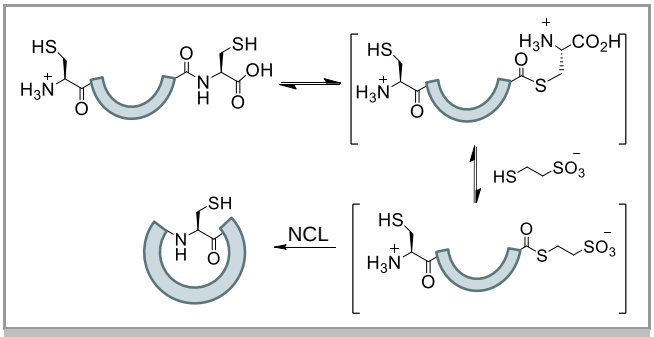


Figure 3 Likely progression from linear precursor to head-to-tail cyclized peptide via an initial N→S acyl transfer.

cyclisation to some short peptides (7-10 residues) that are usually constrained by a disulphide bond between the C and N-termini and found, where tested, that the bioactivity of the peptide was maintained. Although not examined in detail, it is likely that the products are also considerably more stable metabolically.

It is expected that, as the length of an unfolded peptide increases, the efficiency of the cyclisation reaction decreases but this remains an interesting challenge to address. Successful and efficient cyclisation using N→S acyl transfer in one-pot is the outcome of a considerable balancing act, which depends on the ability of each player (thioester formation, cyclisation and hydrolysis) to perform. Cysteine promoted thioester formation is favoured at lower pH yet the cyclisation is favoured at neutral to mildly basic pH and so the reaction conditions are usually not ideal for either process. However, if the appropriate balance cannot be achieved then the potential remains to perform the cyclization in steps. Performing the reaction at lower pH allows isolation of the thioester intermediate and then cyclisation, under regular NCL conditions, can be conducted after purification. In native peptide sequences the efficiency of cyclisation of longer peptides is most compromised by competing hydrolysis during prolonged heating of the intermediate thioester when the cyclisation step is slow. Nevertheless peptides up to approximately 30 amino acid residues have been successfully cyclised to date (see section 4).

4 Reduced reactivity of internal Xaa-Cys motifs as an advantage in modification and cyclisation of biologically produced precursors.

It had not escaped our attention that, since cysteine is a naturally occurring amino acid, our method could be translated to biologically derived precursors and so complement intein-mediated approaches. While there is no shortage of available methods for the production of synthetic thioesters, inteins stand alone as the sole facilitators of N→S acyl transfer in recombinant proteins, yet there are countless thioesters that have proved recalcitrant to intein-mediated thiolysis. This is usually because the precursors are unfolded and so the inactive intein fusion is

cted into inclusion bodies, or the fusion protein possesses ...vourable extein sequences adjacent to the splice junction. Alternatively, the linkage can be too reactive and undergo significant hydrolysis inside the cell. Consequently, as a very doable alternative, cysteine promoted N→S acyl shift should be a welcome addition to the protein semi-synthetic toolbox. The obvious advantages of employing biologically produced precursors, in general, are that water can employed as the solvent for chain assembly as well as cyclisation, the peptide chain can be produced sustainably at scale using renewable materials, and does not require excess protected amino acids or coupling reagents.

A potential disadvantage of employing a biological precursor is that we cannot readily distinguish between cysteine residues in an expressed polypeptide. However the inherent differences in reactivity between *terminal* and *internal* Xaa-Cys motifs towards N→S acyl shift suggested that under non-forcing conditions it should be possible to achieve selectivity in thioester formation and this was indeed found to be the case.[50] Using Ubiquitin as an example (approximately 10 kDa as a His-tagged fusion protein purified from *E. coli*) we first showed that "well behaved" proteins (soluble, folded, and free of additional cysteine residues) would undergo smooth C-terminal thioester formation. Thiolysis was more rapid at 60 °C but reactions on recombinant proteins are normally conducted at 50°C or below because, in earlier unpublished experiments, we already knew that C-terminal thioester formation needed to be conducted below 50 °C with Green Fluorescent Protein (GFP) in order to prevent loss of sample fluorescence. We then also went on to show that recombinant samples containing as many as 6 additional cysteine residues would readily undergo selective N→S acyl transfer.

Constrained, often cysteine rich, cyclic peptides are gaining significant momentum as therapeutic lead structures.[51] This is primarily due to their attractive blend of potency, selectivity and extraordinary chemical and metabolic stability, which is conferred upon them by the presence of a head-to-tail cyclic backbone and a network of intramolecular disulphide bonds.[52] Furthermore, regions of these peptides can be replaced with alternative peptide sequences without loss of structure or stability and this "grafting" process can be used to protect otherwise fragile bioactive peptides from metabolic degradation. Although several chemical methods exist for their production, the reliance on solid-phase assembly of the linear precursor compromises their commercial viability. Because these cysteine rich peptides are all relatively short (14-40 amino acid residues) the synthetic genes which are codon optimised for bacterial expression can be readily purchased from several sources at low cost and so little molecular cloning expertise is required. Of the existing methods available to cyclise unprotected peptides inteins,[32, 53] butelase[54] or sortase[55] appear the most attractive. However inteins often fail to produce short cysteine rich peptides efficiently unless produced in yeast,[56] Butelase has yet to be cloned and still relies on isolation of the active enzyme and Sortase A generally requires that its recognition consensus sequence is incorporated into the cyclic product. So, evaluation of new methods for the cyclisation of short cysteine rich linear precursors at low cost, with low environmental impact, could be considered a worthy pursuit.

To this end we over expressed common cyclic peptide scaffolds including prototypical cyclotide Kalata B1 (KB1) and sunflower trypsin inhibitor (SFTI-1) as linear precursors in *E. coli*, (Figure 4) and only four subsequent processing steps (protein isolation, affinity tag cleavage, cyclisation and oxidation) were required in order to produce SFTI-1 and

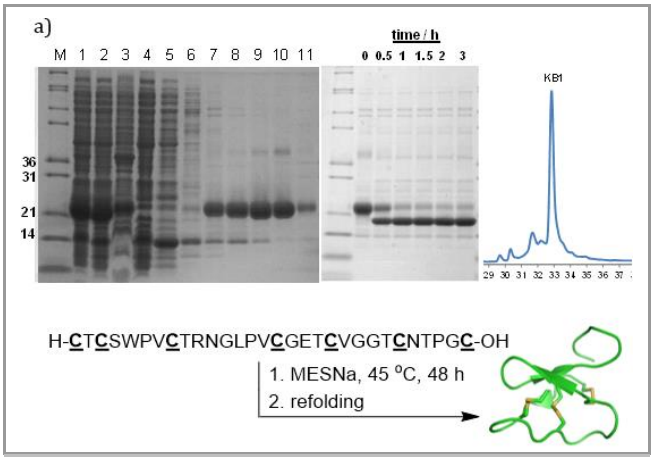


Figure 4 a) *left*: Purification of linear Kalata B1-Thioredoxin fusion (Trx-KB1) from *E. coli* using an Ni^{2+} affinity column. M = molecular weight markers. Lane 1 = whole cell lysate, lane 2 = soluble fraction, lane 3 = insoluble fraction, lane 4 = column flow-through, lanes 5-6 = 5-20 mM imidazole washes, lanes 7-11 = Eluted fractions (40-500 mM imidazole). *middle:* TEV protease digestion of the fusion protein shows accumulation of Trx (released linear KB1 is not visible on the gel). *right*: KB1 is isolated by preparative reverse phase HPLC. b) Cyclization of the isolated linear precursor and oxidative refolding in 0.1 M NH_4HCO_3 pH 8.5/iPrOH (1:1) produce correctly folded kalata B1.

cyclotide analogues in sufficient quantities for structural confirmation and bioactivity assays. Generally, each thioester precursor was fused to the C-terminus of polyhistidine-tagged thioredoxin (Trx) which facilitated stable expression and easy visualisation using SDS-PAGE, which would otherwise not be trivial owing to their small size. A single step purification on a Ni^{2+} column followed by exposure to Tobacco Etch Virus (TEV) protease allowed the linear precursor to be isolated by preparative HPLC. From 1 L of bacterial cell culture 60 mg (unoptimised) of Trx-KB1 fusion protein was obtained and after exposure to TEV protease the linear KB1 precursor was isolated in 23% yield. Kalata B1 was finally produced following backbone cyclisation (45 °C for 48 h) and oxidative refolding. The correct cyclotide fold was confirmed by NMR analysis of the final product and comparison with an authentic sample. The ease with which this peptide, containing 6 additional Cys residues, underwent cyclisation is noteworthy and may have benefitted from a "thia-zip" type cyclisation mechanism where cyclisation is aided by successive intramolecular thiolactone forming reactions which ultimately bring the termini into close proximity.[57] An identical process was adopted to produce variants of 14 residue SFTI-1 which were ultimately tested as inhibitors of Kalikrein 5 (KLK5), a protease involved in skin barrier homeostasis.[45, 58]

Overall, rather than constituting a severe limitation, the differing responses of Xaa-Cys motifs toward thiolysis and the reduced reactivity at internal Cys residues has in fact proved very beneficial on several occasions. A rather general strategy has been adopted for producing SFTI and cyclotide analogues from linear precursors and it is likely that this approach could be extended to other classes of cyclic peptides and combined with unnatural amino acids using an expanded genetic code,[59] in order to introduce novel functionality. Furthermore, when the thioester is only generated fleetingly, as in a cyclisation process, hydrolysis can become less problematic. This phenomenon has allowed us to tackle a significant technical obstacle associated with cyclisation at scale. Although linear peptide precursors can be generated sustainably using bacteria, the cyclisation usually employs MESNa at high concentrations (0.7 M) and large reaction volumes are laborious to purify by HPLC as they cannot be easily concentrated owing to the high MESNa concentration. However, in cyclisation reactions MESNa can often be replaced with 3-mercaptopropanol (3-mPrOH), a relatively poor-performing thiol when applied to linear thioesters, owing to competing *S→O* acyl shift resulting in ester products. Fortuitously 3-mPrOH can be employed cleanly in several cyclisation reactions and, in contrast to MESNa, can be easily extracted from the reaction mixture into ethyl acetate and potentially be reused.[45] This advance means that reactions conducted at scale can be concentrated over 200-fold prior to purification, resulting in a streamlined process for large scale recombinant cyclic peptide production.

5 Hydrazinolysis of Xaa-Cys motifs and the acyl hydrazide as a stable thioester equivalent

In our experience to date, there are few peptides meeting the criteria for a favourable reaction (good aqueous solubility with a reactive C-terminal Xaa-Cys motif) that subsequently fail to furnish a thioester product. The efficiency of the reaction, as evidenced by the yield of isolated product, and clarity of the HPLC trace, is however dependent on a number of additional factors. Short peptides containing Asp-Gly sequences are particularly prone to aspartimide formation if reactions are conducted at elevated temperature (50-60 °C) for as long as 48 h. Poor reaction yields tend to be observed when thioester formation is especially slow (>48 h to reach >50% conversion), and usually when the peptide is especially prone to aggregation upon heating or has low aqueous solubility at the start. Most poorly soluble peptides can be dissolved using chaotropes such as 6 M guanidinium hydrochloride or 6-8 M urea but, while peptides remain soluble in the denaturant, progress can remain painfully slow. Also, urea should be avoided since peptide carbamylation can occur under the reaction conditions. The problems are exacerbated by oxidation of the excess thiol additive, required for transthioesterification, over prolonged reaction times. This is difficult to observe when MESNa is employed owing to its high polarity but the symptoms are clearly visible as the reaction will grind to a halt. Progress can be resumed by the addition of fresh reagents, which is easy to do with larger proteins under non-denaturing conditions using a centrifugal filter with an appropriate molecular weight cut-off. However this is more challenging with shorter peptides and so we have also sought improved methods that reduce the need for prolonged heating at acidic pH.

One potential solution is to use the acyl hydrazide (**4,** Figure 5 a) as a stable thioester precursor.[60] An increasing number of researchers are employing the acyl hydrazide as a thioester precursor because of its stability, relative ease of synthesis and greater flexibility in terms of synthetic strategy and tactics.[23, 61] Once formed, **4** can be converted to a thioester via diazotisation and thiolysis of an acyl azide intermediate (**5**). Indeed, the thioester need not be isolated at all and the diazotisation and subsequent NCL can be conducted in one pot, minimising the opportunity for thioester hydrolysis. Although we have not studied hydrazinolysis in combination with N→S acyl transfer

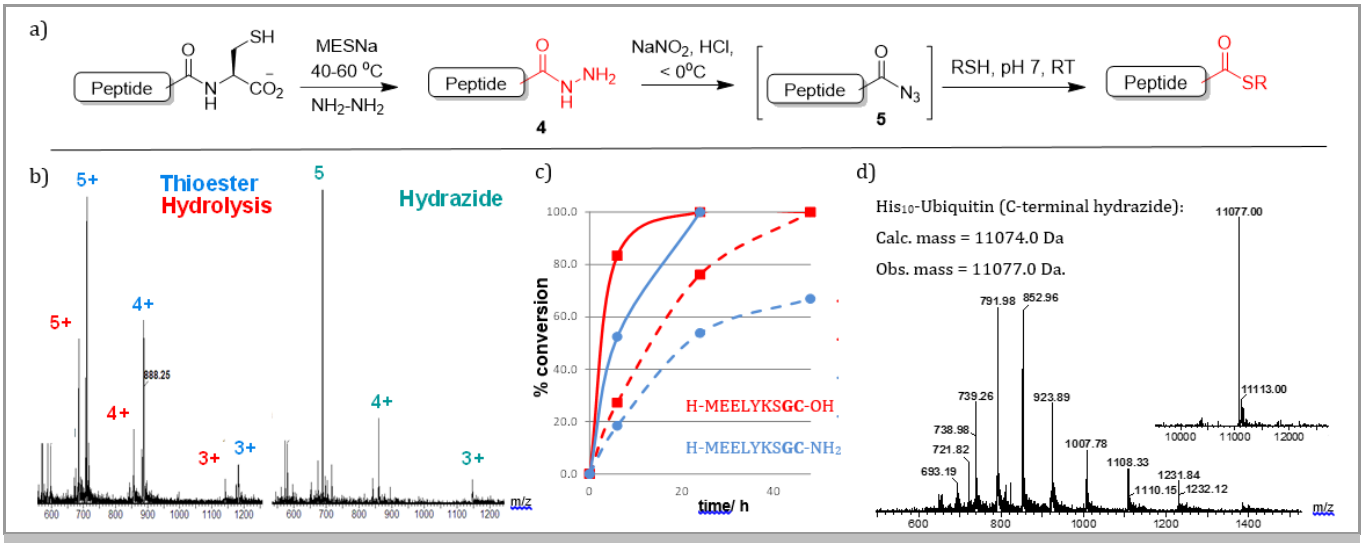


Figure 5 a) In the presence of Hydrazine, added as hydrazinium acetate, the thioester can be intercepted and undergo quantitative hydrazinolysis with minimal hydrolysis. The hydrazide (**4**) can be converted back to a thioester rapidly at low temperature via diazotisation. b) LC-MS analysis of a model peptide (Hepcidin 1-25) upon thiolysis with MESNa (left) and hydrazinolysis (right) for 48 h at 60 °C. c) ^{13}C NMR monitoring of hydrazinolysis (solid lines) relative to thioester formation (dashed lines). d) Recombinant His$_{10}$-ubiquitin modified as a C-terminal hydrazide after exposure to hydrazinium acetate (40 °C, 48 h).

extensively, a number of advantageous features were quick to emerge. First, the hydrazide product is stable and does not hydrolyse over a prolonged reaction time. We applied C-terminal hydrazinolysis to peptides that failed to form thioesters efficiently under typical reaction conditions and found that the results were much improved (Figure 5b). Second, the rate of hydrazinolysis is significantly faster when compared to thioester formation, allowing us to perform reactions at lower temperatures (40-50 °C). The faster reaction may be a consequence of the greater nucleophilicity of hydrazine relative to the thiol of MESNa, but also the inability of the acyl hydrazide to exchange with the liberated cysteine residue once formed. Interestingly no reaction (other than peptide degradation) is observed when MESNa is absent from the reaction mixture suggesting that hydrazine either intercepts the MESNa thioester rather than the *S*-peptide intermediate, or MESNa performs an additional protective role. Notably the final reaction pH is generally raised upon the addition of hydrazinium acetate to nearer pH 7 and C-terminal carboxylates still undergo hydrazinolysis faster than C- terminal carboxamides (Figure 5c). In summary, thioesters that were poorly accessible using typical reaction conditions can alternatively be formed rapidly from hydrazides under denaturing conditions below 0°C, so avoiding the conditions that can drive peptide precipitation (prolonged heating at acidic pH).

Like others, we have utilized synthetic acyl hydrazides (produced via N→S acyl shift) for the assembly of synthetic peptides using NCL, but have additionally investigated hydrazinolysis of biological samples. As with biological thioester synthesis, the protein hydrazide (Figure 5d) is also accessible by intercepting intein fusions and backbone oxoesters with hydrazine.[62] The latter can be introduced in response to an evolved mutant pyrrolysyl-tRNA synthetase (PylRS) from *M. barkeri* , engineered to incorporate α-hydroxy acid analogues of pyrrolysine in a bacterial system.[63] Our straightforward method of hydrazide production may well prove attractive for C-terminal protein labelling or conjugation, particularly where the more common approaches fail.

6 Rapid thioester formation via an N→Se acyl shift

A banal extension to increase the reactivity of the Xaa-Cys system was to examine selenocysteine in place of cysteine as an acyl transfer facilitator, and so we were first to investigate the use of selenocysteine (Sec, U) for this purpose.[64] The superior nucleophilicity and leaving group properties of selenols (in NCL) had already been widely reported and utilized advantageously through elegant prior work by Raines,[65] Hondal,[66] Van der Donk,[67] and Hilvert.[68] This significant precedent allowed us to easily assemble peptides furnished with a C-terminal Sec residue, the only significant challenge being removal of Se protecting groups (usually *p*-methoxybenzyl) to liberate the free selenol for N→Se transfer.

With nitropyridyl selenosulfides (**6**, Figure 6) in hand we

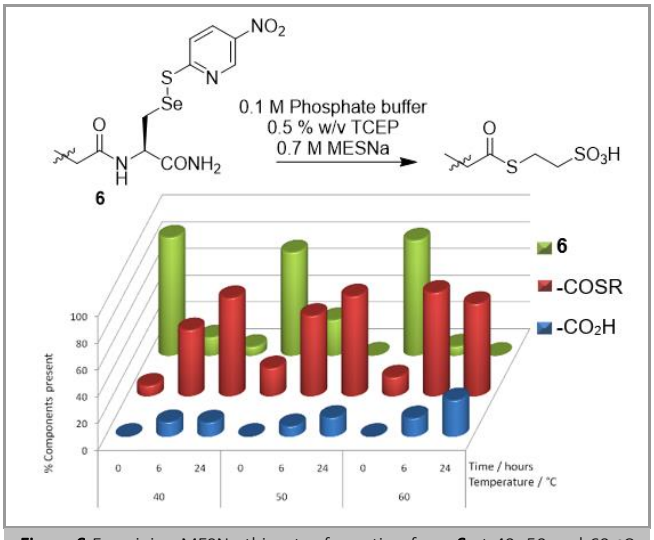


Figure 6 Examining MESNa thioester formation from **6** at 40, 50 and 60 °C over 24 h. Considering the use of the carboxamide terminated precursor these reactions are relatively fast with almost complete reaction within nearly 6 h at 60°C.

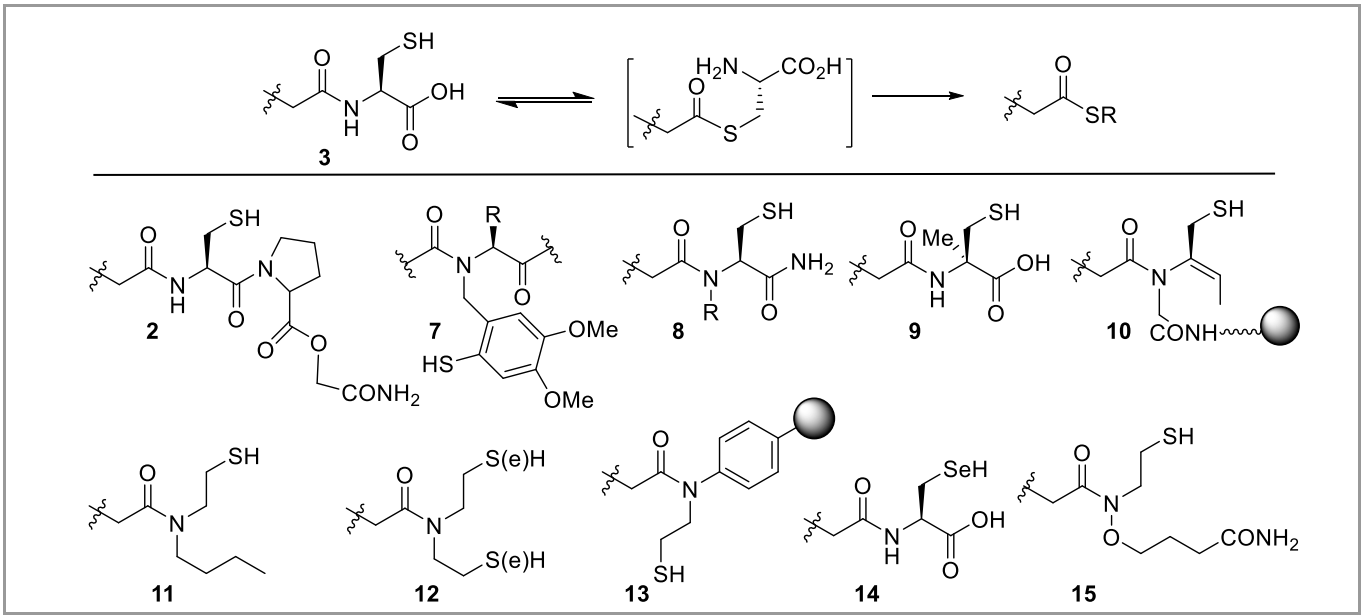
were able to show that, under reducing conditions (in the presence of *tris*-carboxyethyl phosphine (TCEP)), thioester formation occurred rapidly. Reactions performed at 60 °C were essentially complete in less than 6 h. Efficient thioester formation was also observed at 40 °C although in all reactions slightly more hydrolysis was observed to occur relative to the corresponding (Xaa-Cys) terminated peptide. Despite the very promising results in model systems (short peptides devoid of additional Sec or Cys residues) the reaction was poor when applied in a more "realistic" setting. This was likely due to the strong tendency for Sec to form less reactive diselenide and selenosulfide bonds, both of which could be observed by LC-MS analysis of reaction mixtures.[69] The ability to balance reducing conditions with minimal deselenization[70] was also difficult to achieve but recent research by the Hondal group, where selenol species are liberated from nitropyridyl selenosulfides through ascorbolysis may yet provide a way forward.[71]

Several other groups have also explored the synthesis and use of selenoesters in ligation processes[72] [73] In a notable example Melnyk and co-workers showed that selenocysteine (present within peptide sequences as Xaa-Sec) was able to undergo N→Se acyl transfer in the context of transamidation reactions,[74] confirming our findings that Xaa-Sec motifs, present in whatever context, were significantly more reactive than the corresponding Xaa-Cys motif. This research showed that Sec containing peptides were capable of undergoing transamidation at 37 °C, albeit requiring a specialised selenophosphine (TCEP=Se) reagent to prevent deselenization, as well as catalysis by MPAA. We and others [75] others have confirmed that transamidation via internal or terminal Xaa-Cys motifs, outwith the context of intramolecular lactamization, is extremely inefficient at pH 7, even at elevated temperature. We previously observed what was presumed to be a similar reaction of Xaa-Cys terminated peptides in the form of peptide concatenation (as a minor nuisance), when performing cyclisation reactions at higher concentration.

7 Future outlook and conclusion

Up to this point our research has been mainly focussed on examining N→S acyl transfer activity in the context of native peptide (Xaa-Cys) sequences, comparing their differing aptitudes for N→S acyl transfer as well as the relative reactivity of multiple "identical" Xaa-Cys motifs. We have also attempted to accelerate N→S acyl transfer through evaluation of various alternative water soluble thiols, under a range of reaction conditions, and in the presence of new additives but thus far MESNa has produced the most stable thioesters. Currently utilizing N→S acyl transfer in combination with hydrazinolysis shows most promise for delivering thioesters and ligation products from native peptides under the mildest conditions. We have operated under a self-imposed limitation that our best process should be directly transferrable to biologically derived samples thus retaining the requirement for Cys (or Sec) in unmodified form. The development of acyl transfer catalysts that have increased solubility (relative to MPAA) at lower pH have also been of much interest to us, particularly in the context of cyclisation reactions, since these reactive intermediates could potentially enhance such a process despite low stability of the corresponding linear thioester towards hydrolysis. [76] In countless favourable cases selective thiolysis of Xaa-Cys terminated peptides is sufficient to afford valuable peptide thioesters and cyclic peptides and, as a possible alternative we also investigated employing Xaa-Sec motifs. Although using selenocysteine to facilitate the chemistry is attractive at first sight both in chemical and biological terms, the capricious nature of this doppelganger (in our hands at least) means that it is likely to remain a muse for specialists.

From the research undertaken in this burgeoning area over the past 5-10 years it is clear that several scaffolds other than Cys (**3**) can function effectively in N→S acyl shift reactions (Scheme 4) including the CPE (**2**), acyl transfer auxiliaries (**7**),[9, 10, 77] *N*-alkyl cysteine derivatives (**8**),[75, 78] α-Methyl cysteine (**9**),[75] enamide **10**,[79] thioethylbutylamides (TEBA, **11**),[80] bis-(2-sulfanyl/selenylethyl)amides (bis-S(e)EAlide, **12**),[82]



Scheme 4 N→S acyl transfer promoted by a single Cys residue and some N→S acyl transfer devices that have been developed to date.[19, 81]

sulfanylethylanilides (SEAlide, **13**),[72, 83] selenocysteine (**14**),[64, 74]and N-Sulfanylethyl aminooxybutyramide (SEAoxy, **15**).[84] The majority of these possess greater acyl transfer activity than cysteine under similar reaction conditions,[19, 85] although most similarly require acidic pH and/or elevated temperature for optimal performance. In some cases a real "witches brew" of additives, including a high concentration of TCEP, which is known to potentially damage protein samples,[86] and an inert atmosphere are also required in order to initiate the chemistry. Meanwhile, we have found that in several reactions of Xaa-Cys terminated peptides that TCEP can be omitted or replaced with sodium ascorbate.[87] The most widely adopted method for thioester synthesis via an N→S acyl shift will ultimately be a compromise between the most affordable/user friendly, and thioester yield. Only these features will likely be competitive with more common and reliable approaches.[18] A key consideration that is often overlooked is the difficulty in loading the first amino acid and on those grounds **9** and **15** appear particularly attractive since the peptide is assembled using typical resins, coupling reagents and protocols.

Despite the greater reactivity of **7**-**15** relative to **3** there is currently little scope to combine them with the benefits of routine biological production. The desire to utilise biological precursors will not be universally shared but only once this conceptual and practical leap is made may the power of thioester synthesis through N→S acyl transfer be fully realised. In this respect it may be challenging to introduce N-alkylated amino acids because their slow incorporation by the translation machinery can lead to termination of the synthesis and a truncated product.[88] Consequently α-Methyl cysteine terminated peptides could be particularly attractive for this purpose. The CPE (**2**) has already been incorporated successfully into short linear and cyclic peptide sequences using an engineered in vitro-transcription-translation system.[89] The encoded CPE was used to effect head to tail cyclisation of short peptides containing additional non-canonical amino acids. To the best of our knowledge there are no reports of its extension to "regular" proteins. Using genetically encoded cysteine analogues at the C-terminus of proteins is likely to produce recombinant proteins that can be processed more reliably, with tuneable initiation of N→S acyl transfer in the test-tube or inside cells. The process would be complementary to already extremely powerful intein-mediated methods and provide genuine contingency for the countless instances where the most widely available intein-fusions fail to produce a thioester product. Peptides under study may neither need to be expressed as fusions nor be fully folded, and those also containing an N-terminal cysteine residue may cyclise spontaneously inside cells.

Peptides can also be cyclised via alternative modes such as head-to-sidechain and tail-to-sidechain to form novel branched topologies,[90] and consequently the use of N→S acyl transfer in concert with unnatural amino acid mutagenesis has the potential to take acyl transfer reactions into this exciting new arena, particularly because inteins are incapable of performing N→S acyl transfer at the sidechain of amino acids. Melnyk and co-workers have already shown how the bis-SEALide (**12**), when appended to the side chain of Asp and Glu to afford **16** (Figure 7a) can facilitate sidechain activation in synthetic peptides, affording tail-to-sidechain cyclised products.[91] This process could be highly transformative if such devices can be incorporated into biologically produced proteins. The products of such a process should more generally enable site-selective carboxyl activation in recombinant proteins that can be targeted with nucleophilic payloads (e.g. fluorescent label, PEG, or carbohydrates that are available as hydrazine or aminooxy derivatives, or through NCL). Fascinatingly most of the tools required to enable the genetic encoding of branched dipeptides such as **18**-**20** (Figure 7b) are already available and the fact that dipeptide **21** and further analogues have already been encoded, enabling NCL at the sidechain of Lys residues,[92] sets a significant precedent. The ability to conduct sidechain activation and cyclisation reactions under native-like conditions, and/or with the aid of reversible tethers, may ultimately bring the challenging peptide scaffolds such as the lariat peptide (Figure 7c) under the command of the synthetic chemist.[93]

In conclusion, our unanticipated diversion from glycopeptide synthesis took us on a new adventure in peptide thioester chemistry. With further refinement our approach surely presents a number of opportunities for applications relevant to peptide and protein engineering. N→S acyl transfer in native peptide sequences may not be the most effective means to produce thioesters but in countless examples to date it is has proved "good enough", while its obvious simplicity renders it amenable to chemical and biological samples alike. Furthermore, the ease with which biological protein production can be scaled could render this approach very attractive, trumping total chemical synthesis, for the production of various cysteine rich cyclic peptides, which are a rapidly growing class of therapeutics.

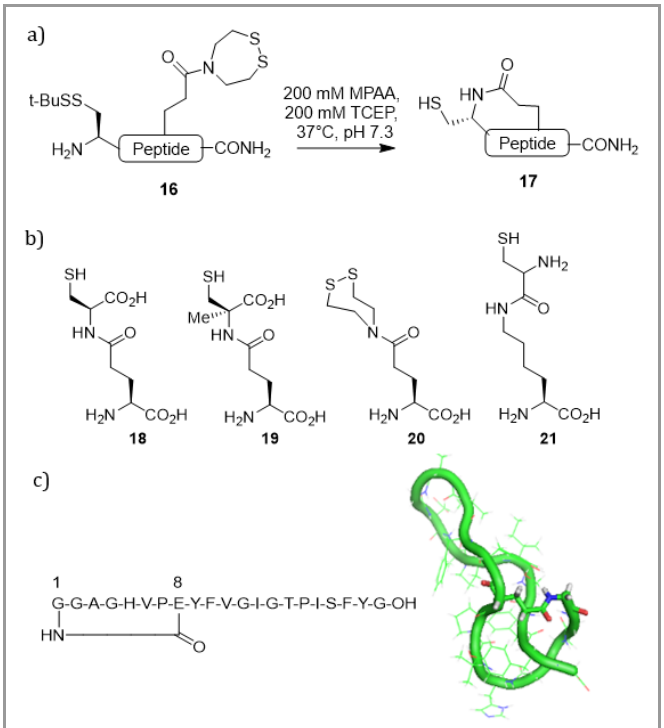


Figure 7 a) tail-to-sidechain peptide cyclisation enabled by the *bis*-SEAlide appended to Asp or Glu. b) Dipeptides that could potentially be genetically encoded to facilitate sidechain N→S acyl shift (**18**-**20**) alongside Nε-cysteinyl lysine (**21**) that has already been genetically encoded to enable sidechain ubiquitination via NCL. c) Antimicrobial lariat peptide microcin J25, a potential target for synthesis via sidechain N→S acyl transfer.

In future the combination of biological peptide production with access to new topologies, and using optimised methods that are compatible with existing NCL streamlining practices could ultimately see peptide research N→S acyl shift in the right direction.

References

(1) (a) de Duve, C. *American Scientist* **1995**, 83, (5), 428-437. (b) Bracher, P. J.; Snyder, P. W.; Bohall, B. R.; Whitesides, G. M. *Origins of Life and Evolution of Biospheres* **2011,** 41, (5), 399-412.

(2) Dawson, P.; Muir, T.; Clark-Lewis, I.; Kent, S. *Science* **1994,** 266, (5186), 776-779.

(3) Wieland, T.; Bokelmann, E.; Bauer, L.; Lang, H. U.; Lau, H. *Justus Liebigs Annalen der Chemie* **1953,** 583, (1), 129-149.

(4) Merrifield, R. B. *J. Am. Chem. Soc.* **1963,** 85, (14), 2149-2154.

(5) Kent, S. B. H. *Chem. Soc. Rev.* **2009,** 38, (2), 338-351.

(6) Martin, R. B.; Lowey, S.; Elson, E. L.; Edsall, J. T. *J. Am. Chem. Soc.* **1959,** 81, (19), 5089-5095.

(7) Shah, N. H.; Muir, T. W. *Chem. Sci.* **2014,** 5, (2), 446-461.

(8) Zanotti, G.; Pinnen, F.; Lucente, G. *Tetrahedron Lett.* **1985,** 26, (44), 5481-5484.

(9) Kawakami, T.; Sumida, M.; Nakamura, K. i.; Vorherr, T.; Aimoto, S. *Tetrahedron Lett.* **2005,** 46, (50), 8805-8807.

(10) Ken'ichiroh, N.; Megumi, S.; Toru, K.; Thomas, V.; Saburo, A. *Bull. Chem. Soc. Jpn.* **2006,** 79, (11), 1773-1780.

(11) Ollivier, N.; Behr, J.-B.; El-Mahdi, O.; Blanpain, A.; Melnyk, O. *Org. Lett.* **2005,** 7, (13), 2647-2650.

(12) Nakamura, K. i.; Sumida, M.; Kawakami, T.; Vorherr, T.; Aimoto, S. *Bull. Chem. Soc. Jpn* **2006,** 79, (11), 1773–1780; Ollivier, N.; Behr, J.-B.; El-Mahdi, Q.; Blanpain, A.; Melnyk, O. *Organic Letters* **2005,** 7, (13), 2647-2650.

(13) (a) Liu, Z.; Frutos, S.; Bick, M. J.; Vila-Perelló, M.; Debelouchina, G. T.; Darst, S. A.; Muir, T. W. *Proc. Natl. Acad. Sci. USA* **2014,** 111, (23), 8422-8427. (b) Romanelli, A.; Shekhtman, A.; Cowburn, D.; Muir, T. W. *Proc. Natl. Acad. Sci. USA* **2004,** 101, (17), 6397-6402. (c) Binschik, J.; Mootz, H. D. *Angew. Chem. Int. Ed.* **2013,** 52, (15), 4260-4264.

(14) Volkmann, G.; Mootz, H. *Cell. Mol. Life Sci.* **2013,** 70, (7), 1185-1206.

(15) (a) Eryilmaz, E.; Shah, N. H.; Muir, T. W.; Cowburn, D. *J. Biol. Chem.* **2014,** 289, (21), 14506-14511. (b) Mills, K. V.; Johnson, M. A.; Perler, F. B. *J. Biol. Chem.* **2014,** 289, (21), 14498-14505. (c) Callahan, B. P.; Topilina, N. I.; Stanger, M. J.; Van Roey, P.; Belfort, M. *Nat. Struct. Mol. Biol.* **2011,** 18, (5), 630-633.

(16) Kent, S. *Peptide Sci* **2010,** 94, (4), iv-ix; Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. *Science* **1994,** 266, (5186), 776-779; Kent, S. B. H. *Journal of Peptide Science* **2015,** 21, (3), 136-138.

(17) (a) Zheng, J.-S.; Tang, S.; Huang, Y.-C.; Liu, L. *Accounts Chem. Res.* **2013,** 46, (11), 2475-2484;. (b) Ling, J. J.; Policarpo, R. L.; Rabideau, A. E.; Liao, X.; Pentelute, B. L. *J. Am. Chem. Soc.* **2012,** 134, (26), 10749-10752. (c) Unverzagt, C.; Kajihara, Y. *Chem. Soc. Rev.* **2013,** 42, (10), 4408-4420. (d) Schnölzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S. *Int. J. Pept. Res. Ther.* **2007,** 13, (1), 31-44.

(18) Mende, F.; Seitz, O. *Angew. Chem. Int. Ed.* **2011,** 50, (6), 1232-1240.

(19) Melnyk, O.; Agouridas, V. *Curr. Opin. Chem. Biol.* **2014,** 22, 137-145.

(20) Bondalapati, S.; Jbara, M.; Brik, A. *Nat. Chem.* **2016,** 8, (5), 407-418.

(21) (a) Lee, J.; Kwon, Y.; Pentelute, B. L.; Bang, D. *Bioconjugate* Chem. **2011,** 22, (8), 1645-1649. (b) Bang, D.; Pentelute, B. L.; Kent, S. B. H. *Angew. Chem. Int. Ed.* **2006,** 45, (24), 3985-3988.

(22) Bang, D.; Kent, S. B. H. *Angew. Chem. Int. Ed.* **2004,** 43, (19), 2534-2538.

(23) Li, J.; Li, Y.; He, Q.; Li, Y.; Li, H.; Liu, L. *Org. Biomol. Chem.* **2014,** 12, (29), 5435-5441.

(24) Thompson, R. E.; Liu, X.; Alonso-García, N.; Pereira, P. J. B.; Jolliffe, K. A.; Payne, R. J. *J. Am. Chem. Soc.* **2014,** 136, (23), 8161-8164. (b) Moyal, T.; Hemantha, H. P.; Siman, P.; Refua, M.; Brik, A. *Chem. Sci.* **2013,** 4, (6), 2496-2501.

(25) Malins, L. R.; Payne, R. J. *Curr. Opin. Chem. Biol.* **2014,** 22, 70-78.

(26) (a) Kan, C.; Danishefsky, S. J. *Tetrahedron* **2009,** 65, (45), 9047-906. (b) Wan, Q.; Danishefsky, S. J. *Angew. Chem. Int. Ed.* **2007,** 46, (48), 9248-9252. (c) Yan, L. Z.; Dawson, P. E. *J. Am. Chem. Soc.* **2001,** 123, (4), 526-533.

(27) Offer, J. *Peptide Sci.* **2010,** 94, (4), 530-541.

(28) Ackrill, T.; Anderson, D. W.; Macmillan, D. *Peptide Sci.* **2010,** 94, (4), 495-503.

(29) Wu, B.; Chen, J.; Warren, J. D.; Chen, G.; Hua, Z.; Danishefsky, S. J. *Angew. Chem. Int. Ed.* **2006,** 45, (25), 4116-4125.

(30) Loibl, S. F.; Harpaz, Z.; Seitz, O. *Angew. Chem. Int. Ed.* **2015,** 54, (50), 15055-15059.

(31) Johnson, E. C. B.; Kent, S. B. H. *J. Am. Chem. Soc.* **2006,** 128, (20), 6640-6646.

(32) Xu, M.-Q.; Evans Jr, T. C. *Methods* **2001,** 24, (3), 257-277.

(33) (a) Aimoto, S. *Peptide Sci.* **1999,** 51, (4), 247-265. (b) Hackeng, T. M.; Griffin, J. H.; Dawson, P. E. *Proc. Natl. Acad. Sci. USA* **1999,** 96, (18), 10068-10073.

(34) (a) Raz, R.; Burlina, F.; Ismail, M.; Downward, J.; Li, J.; Smerdon, S. J.; Quibell, M.; White, P. D.; Offer, J. *Angew. Chem. Int. Ed.* **2016,** 55, (42), 13174-13179. (b) Gates, Z. P.; Dhayalan, B.; Kent, S. B. H. *Chem. Commun.* **2016,** 52, (97), 13979-13982.

(35) Behrendt, R.; White, P.; Offer, J. *J. Peptide Sci.* **2016,** 22, (1), 4-27.

(36) (a) Clippingdale, A. B.; Barrow, C. J.; Wade, J. D. *J. Peptide Sci.* **2000,** 6, (5), 225-234. (b) Hasegawa, K.; Sha, Y. L.; Bang, J. K.; Kawakami, T.; Akaji, K.; Aimoto, S. *Lett. Pep. Sci.* **2001,** 8, (3), 277-284. (c) Raz, R.; Rademann, J. *Org. Lett.* **2011,** 13, (7), 1606-1609.

(37) Blanco-Canosa, J. B.; Dawson, P. E. *Angew. Chem. Int. Ed.* **2008,** 47, (36), 6851-6855.

(38) Shin, Y.; Winans, K. A.; Backes, B. J.; Kent, S. B. H.; Ellman, J. A.; Bertozzi, C. R. *J. Am. Chem. Soc.* **1999,** 121, (50), 11684-11689.

(39) (a) Mezzato, S.; Schaffrath, M.; Unverzagt, C. *Angew. Chem. Int. Ed.* **2005,** 44, (11), 1650-1654. (b) Burlina, F.; Morris, C.; Behrendt, R.; White, P.; Offer, J. *Chem. Commun.* **2012,** 48, (20), 2579-2581.

(40) Kawakami, T.; Aimoto, S. *Tetrahedron* **2009,** 65, (19), 3871-3877.

(41) Nagaike, F.; Onuma, Y.; Kanazawa, C.; Hojo, H.; Ueki, A.; Nakahara, Y.; Nakahara, Y. *Org. Lett.* **2006,** 8, (20), 4465-4468.

(42) Kang, J.; Richardson, J. P.; Macmillan, D. *Chem. Commun.* **2009**, (4), 407-409.

(43) Kang, J.; Reynolds, N. L.; Tyrrell, C.; Dorin, J. R.; Macmillan, D. *Org. Biomol. Chem.* **2009,** 7, (23), 4918-4923.

(44) Richardson, J. P.; Chan, C.-H.; Blanc, J.; Saadi, M.; Macmillan, D. *Org. Biomol. Chem.* **2010,** 8, (6), 1351-1360.

(45) Cowper, B.; Shariff, L.; Chen, W.; Gibson, S. M.; Di, W.-L.; Macmillan, D. *Org. Biomol. Chem.* **2015,** 13, (27), 7469-7476.

(46) Asahina, Y.; Nabeshima, K.; Hojo, H. *Tetrahedron Lett.* **2015,** 56, (11), 1370-1373.

(47) Premdjee, B.; Adams, A. L.; Macmillan, D. *Bioorg. Med. Chem. Lett.* **2011,** 21, (17), 4973-4975.

(48) 48. Masania, J.; Li, J.; Smerdon, S. J.; Macmillan, D. *Org. Biomol. Chem.* **2010,** 8, (22), 5113-5119.

(49) Macmillan, D.; De Cecco, M.; Reynolds, N. L.; Santos, L. F. A.; Barran, P. E.; Dorin, J. R. *ChemBioChem* **2011,** 12, (14), 2133-2136.

(50) Adams, A. L.; Cowper, B.; Morgan, R. E.; Premdjee, B.; Caddick, S.; Macmillan, D. *Angew. Chem. Int. Ed.* **2013,** 52, (49), 13062-13066.

(51) Northfield, S. E.; Wang, C. K.; Schroeder, C. I.; Durek, T.; Kan, M.-W.; Swedberg, J. E.; Craik, D. J. *Eur. J. Med. Chem.* **2014,** 77, 248-257.

(52) (a) Bhardwaj, G.; Mulligan, V. K.; Bahl, C. D.; Gilmore, J. M.; Harvey, P. J.; Cheneval, O.; Buchko, G. W.; Pulavarti, S. V. S. R. K.; Kaas, Q.; Eletsky, A.; Huang, P.-S.; Johnsen, W. A.; Greisen, P., Jr.; Rocklin, G. J.; Song, Y.; Linsky, T. W.; Watkins, A.; Rettie, S. A.; Xu, X.; Carter, L. P.; Bonneau, R.; Olson, J. M.; Coutsias, E.; Correnti, C. E.; Szyperski, T.; Craik, D. J.; Baker, D. *Nature* **2016,** 538, (7625), 329-335. (b) Clark, Richard J.; Daly, Norelle L.; Craik, David J. *Biochem. J.* **2006,** 394, 85-93. (c) Poth, A. G.; Chan, L. Y.; Craik, D. J. *Peptide Sci.* **2013,** 100, (5), 480-491.

(53) (a) Tavassoli, A.; Benkovic, S. J. *Nat. Protocols* **2007,** 2, (5), 1126-1133. (b) Tarasava, K.; Freisinger, E. *Protein Eng. Des. Sel.* **2014,** 27,

(12), 481-488. (c) Aboye, T. L.; Camarero, J. A. *J. Biol. Chem.* **2012,** 287, (32), 27026-27032.
(54) (a) Nguyen, G. K. T.; Wang, S.; Qiu, Y.; Hemu, X.; Lian, Y.; Tam, J. P. *Nat. Chem. Biol.* **2014,** 10, (9), 732-738. (b) Nguyen, G. K. T.; Qiu, Y.; Cao, Y.; Hemu, X.; Liu, C.-F.; Tam, J. P. *Nat. Protocols* **2016,** 11, (10), 1977-1988.
(55) (a) Wu, Z.; Guo, X.; Guo, Z. *Chem. Commun.* **2011,** 47, (32), 9218-9220. (b) Antos, J. M.; Popp, M. W.-L.; Ernst, R.; Chew, G.-L.; Spooner, E.; Ploegh, H. L. *J. Biol. Chem.* **2009,** 284, (23), 16028-16036. (c) Jia, X.; Kwon, S.; Wang, C.-I. A.; Huang, Y.-H.; Chan, L. Y.; Tan, C. C.; Rosengren, K. J.; Mulvenna, J. P.; Schroeder, C. I.; Craik, D. J. *J. Biol. Chem.* **2014,** 289, (10), 6627-6638.
(56) Jagadish, K.; Gould, A.; Borra, R.; Majumder, S.; Mushtaq, Z.; Shekhtman, A.; Camarero, J. A. *Angew. Chem. Int. Ed.* **2015,** 54, (29), 8390-8394.
(57) Tam, J. P.; Wong, C. T. T. *J. Biol. Chem.* **2012,** 287, (32), 27020-27025.
(58) *Chen*, W.; Kinsler, V. A.; Macmillan, D.; Di, W.-L. *PLOS ONE* **2016,** 11, (11), e0166268.
(59) (a) Xie, J.; Schultz, P. G. *Nat Rev Mol Cell Biol* **2006,** 7, (10), 775-782. (b) Davis, L.; Chin, J. W. *Nat Rev Mol Cell Biol* **2012,** 13, (3), 168-182.
(60) (a) 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, (33), 7645-7649. (b) Wang, P.; Layfield, R.; Landon, M.; Mayer, R. J.; Ramage, R. *Tetrahedron Lett.* **1998,** 39, (47), 8711-8714.
(61) (a) Reif, A.; Siebenhaar, S.; Tröster, A.; Schmälzlein, M.; Lechner, C.; Velisetty, P.; Gottwald, K.; Pöhner, C.; Boos, I.; Schubert, V.; Rose-John, S.; Unverzagt, C. *Angew. Chem. Int. Ed.* **2014,** 53, (45), 12125-12131. (b) Murakami, M.; Kiuchi, T.; Nishihara, M.; Tezuka, K.; Okamoto, R.; Izumi, M.; Kajihara, Y. *Science Advances* **2016,** 2, (1). (c) Siman, P.; Karthikeyan, S. V.; Nikolov, M.; Fischle, W.; Brik, A. *Angew. Chem. Int. Ed.* **2013,** 52, (31), 8059-8063. (d) Bello, C.; Wang, S.; Meng, L.; Moremen, K. W.; Becker, C. F. W. *Angew. Chem. Int. Ed.* **2015,** 54, (26), 7711-7715.
(62) Thom, J.; Anderson, D.; McGregor, J.; Cotton, G. *Bioconjugate Chem.* **2011,** 22, (6), 1017-1020.
(63) Li, Y.-M.; Yang, M.-Y.; Huang, Y.-C.; Li, Y.-T.; Chen, P. R.; Liu, L. *ACS Chem. Biol.* **2012,** 7, (6), 1015-1022.
(64) Adams, A. L.; Macmillan, D. *J. Peptide Sci.* **2013,** 19, (2), 65-73.
(65) Hondal, R. J.; Nilsson, B. L.; Raines, R. T. *J. Am. Chem. Soc.* **2001,** 123, (21), 5140-5141.
(66) Harris, K. M.; Flemer, S.; Hondal, R. J. *J. Peptide Sci.* **2007,** 13, (2), 81-93.
(67) *Gieselman*, M. D.; Xie, L.; van der Donk, W. A. *Organic Lett.* **2001,** 3, (9), 1331-1334.
(68) Roelfes, G.; Hilvert, D. *Angew. Chem. Int. Ed.* **2003,** 42, (20), 2275-2277.
(69) Muttenthaler, M.; Alewood, P. F. *J. Peptide Sci.* **2008,** 14, (12), 1223-1239.
(70) Metanis, N.; Keinan, E.; Dawson, P. E. *Angew. Chem. Int. Ed.* **2010,** 49, (39), 7049-7053.
(71) Ste. Marie, E. J.; Ruggles, E. L.; Hondal, R. J. *J. Peptide Sci.* **2016,** 22, (9), 571-576.
(72) Raibaut, L.; Cargoet, M.; Ollivier, N.; Chang, Y. M.; Drobecq, H.; Boll, E.; Desmet, R.; Monbaliu, J.-C. M.; Melnyk, O. *Chem. Sci.* **2016,** 7, (4), 2657-2665.
(73) Mitchell, N. J.; Malins, L. R.; Liu, X.; Thompson, R. E.; Chan, B.; Radom, L.; Payne, R. J. *J. Am. Chem. Soc.* **2015,** 137, (44), 14011-14014.
(74) Ollivier, N.; Blanpain, A.; Boll, E.; Raibaut, L.; Drobecq, H.; Melnyk, O. *Organic Letters* **2014,** 16, (15), 4032-4035.
(75) Burlina, F.; Papageorgiou, G.; Morris, C.; White, P. D.; Offer, J. *Chem. Sci.* **2014,** 5, (2), 766-770.
(76) Cowper, B.; Sze, T. M.; Premdjee, B.; Bongat White, A. F.; Hacking, A.; Macmillan, D. *Chem. Commun.* **2015,** 51, (15), 3208-3210.
(77) 77. Terrier, V. P.; Adihou, H.; Arnould, M.; Delmas, A. F.; Aucagne, V. *Chem. Sci.* **2016,** 7, (1), 339-345.
(78) (a) Hojo, H.; Onuma, Y.; Akimoto, Y.; Nakahara, Y.; Nakahara, Y. *Tetrahedron Lett.* **2007,** 48, (1), 25-28. (b) Hojo, H. *Org. Biomol. Chem.* **2016,** 14, (27), 6368-6374. (c) Erlich, L. A.; Kumar, K. S. A.; Haj-Yahya, M.; Dawson, P. E.; Brik, A. *Org. Biomol. Chem.* **2010,** 8, (10), 2392-2396.
(79) Zheng, J.-S.; Chen, X.; Tang, S.; Chang, H.-N.; Wang, F.-L.; Zuo, C. *Org. Lett.* **2014,** 16, (18), 4908-4911.
(80) (a) Hemu, X.; Taichi, M.; Qiu, Y.; Liu, D.-X.; Tam, J. P. *Peptide Sci.* **2013,** 100, (5), 492-501. (b) Taichi, M.; Hemu, X.; Qiu, Y.; Tam, J. P. *Org. Lett.* **2013,** 15, (11), 2620-2623.
(81) (a) Kang, J.; Macmillan, D. *Org. Biomol. Chem.* **2010,** 8, (9), 1993-2002. (b) Macmillan, D.; Adams, A.; Premdjee, B. *Isr. J. Chem.* **2011,** 51, (8-9), 885-899. (c) Tailhades, J.; Patil, N. A.; Hossain, M. A.; Wade, J. D. *J. Peptide Sci.* **2015,** 21, (3), 139-147.
(82) (a) Ollivier, N.; Dheur, J.; Mhidia, R.; Blanpain, A.; Melnyk, O. *Org. Lett.* **2010,** 12, (22), 5238-5241. (b) Hou, W.; Zhang, X.; Li, F.; Liu, C.-F. *Org. Lett.* **2011,** 13, (3), 386-389. (c) Raibaut, L.; Adihou, H.; Desmet, R.; Delmas, A. F.; Aucagne, V.; Melnyk, O. *Chem. Sci.* **2013,** 4, (10), 4061-4066.
(83) (a) Tsuda, S.; Shigenaga, A.; Bando, K.; Otaka, A. *Organic Letters* **2009,** 11, (4), 823-826. (b) Sato, K.; Shigenaga, A.; Tsuji, K.; Tsuda, S.; Sumikawa, Y.; Sakamoto, K.; Otaka, A. *ChemBioChem* **2011,** 12, (12), 1840-1844.
(84) Tsuda, S.; Mochizuki, M.; Sakamoto, K.; Denda, M.; Nishio, H.; Otaka, A.; Yoshiya, T. *Org. Lett.* **2016,** 18, (22), 5940-5943.
(85) Raibaut, L.; Ollivier, N.; Melnyk, O. *Chem. Soc. Rev.* **2012,** 41, (21), 7001-7015.
(86) Liu, P.; O'Mara, B. W.; Warrack, B. M.; Wu, W.; Huang, Y.; Zhang, Y.; Zhao, R.; Lin, M.; Ackerman, M. S.; Hocknell, P. K.; Chen, G.; Tao, L.; Rieble, S.; Wang, J.; Wang-Iverson, D. B.; Tymiak, A. A.; Grace, M. J.; Russell, R. J. *J. Am. Soc. Mass Spectrom.* **2010,** 21, (5), 837-844.
(87) Rohde, H.; Schmalisch, J.; Harpaz, Z.; Diezmann, F.; Seitz, O. *ChemBioChem* **2011,** 12, (9), 1396-1400.
(88) 88. Pavlov, M. Y.; Watts, R. E.; Tan, Z.; Cornish, V. W.; Ehrenberg, M.; Forster, A. C. *Proc. Natl. Acad. Sci. USA* **2009,** 106, (1), 50-54.
(89) Kawakami, T.; Ohta, A.; Ohuchi, M.; Ashigai, H.; Murakami, H.; Suga, H. *Nat. Chem. Biol.* **2009,** 5, (12), 888-890.
(90) Hegemann, J. D.; Zimmermann, M.; Xie, X.; Marahiel, M. A. *Accounts Chem. Res.* **2015,** 48, (7), 1909-1919.
(91) Raibaut, L.; Drobecq, H.; Melnyk, O. *Org. Lett.* **2015,** 17, (14), 3636-3639.
(92) Nguyen, D. P.; Elliott, T.; Holt, M.; Muir, T. W.; Chin, J. W. *J. Am. Chem. Soc.* **2011,** 133, (30), 11418-11421.
(93) Al Toma, R. S.; Kuthning, A.; Exner, M. P.; Denisiuk, A.; Ziegler, J.; Budisa, N.; Süssmuth, R. D. *ChemBioChem* **2015,** 16, (3), 503-509.

Biosketches

Derek Macmillan completed his BSc (1995), and then PhD (1999) under the supervision of Sabine Flitsch (1999) at the University of Edinburgh. After completing 2 years as a Post-doc in the laboratory of Professor Carolyn Bertozzi at the University of California, Berkeley he returned to the UK and was awarded a Royal Society University Research Fellowship.

In 2005 he relocated to The Department of Chemistry at UCL where projects continue to explore the use of chemistry to understand biological systems, particularly through the assembly of post-translationally modified peptides and proteins using semi-synthesis.