Cysteine: an Essential Inessential Amino Acid

George W. Preston

King's College London, 150 Stamford Street, London, SE1 9NH; e-mail: george.preston@kcl.ac.uk.

Cysteine was discovered indirectly in 1810 by Wollaston, who isolated a crystalline substance from urinary calculi (kidney stones) and called it "cystic oxide" (Vickery and Schmidt, 1931). We now know this substance as cystine, the oxidised dimeric form of cysteine, and the first amino acid to be discovered. As a building block of proteins, however, cysteine (in the form of cystine) evaded detection for several decades, probably because the early methods of amino acid analysis were incompatible with sulfur-containing analytes. Eventually, despite a few controversies and structural mis-assignments, it was confirmed that the substances derived from urinary calculi and proteins were one and the same. Modern biochemistry has uncovered diverse roles of cysteine, and peptide chemists have used this diversity to their advantage for the preparation of new and interesting structures. This short article will highlight a pair of recent publications that illustrate some interesting properties of cysteine's thiol group in chemistry and biology.

CYSTEINE AND PEPTIDE MACROCYCLISATION

The first example, from the group of Zigang Li at Peking University, highlights an interesting utilisation of the thiol group in chemical synthesis (Zhao *et al.*, 2016). Writing in *Journal of Peptide Science*, the authors describe the use of thiol-ene chemistry for the introduction of conformational constraints into peptides. Conformational constraints are a useful way of lowering the entropic cost of ligand binding, and they can also protect ligands against proteolysis *in vivo*. The constraints are often introduced as covalent intramolecular cross-links between side-chains of non-contiguous amino acid residues. A variety of methods have been used for this endeavour, including olefin metathesis, disulfide cross-linking and amide coupling (Lau *et al.*, 2015). A relatively recent addition to the list is the thiol-ene reaction, in which a photo-generated thiyl radical adds to an alkene to generate a thioether product.

In the context of peptide macrocyclisation strategies, the obvious choice of thiol is that belonging to cysteine, whilst the alkene tends to be incorporated *via* a non-proteinogenic amino acid (e.g., *α*-alkenylglycine). In an earlier report, Anseth's group at the University of Colorado, Boulder, described how a heptapeptide could be cyclised on-resin by cross-linking its *N*-terminal cysteine residue to an *N*-Alloc protecting group on a *C*-terminal lysine residue (Aimetti *et al.*, 2010). This report, along with several more recent ones, indicated that thiol-ene chemistry might be quite a versatile tool for peptide macrocyclisation. Zhao and colleagues' contribution has been to demonstrate the use of *α*-alkenylglycines (e.g., *α*-pentenylglycine) for on-resin intramolecular thiol-ene chemistry. The authors established their method using a model pentapeptide containing an *α*-alkenylglycine at the *N*-terminus (position *i*), cysteine at the *C*-terminus (position *i* + 4) and three internal alanine residues. Note that a cross-linked "*i*, *i* + 4" motif often features in designs for constrained peptides because it relates to the pattern of hydrogen bonds in an *α*-helix. The substrate for the intramolecular thiol-ene reaction was an *N*-acetylated peptidyl-resin, whose *S*-trityl protecting group had already been removed under mild conditions (Figure 1). By irradiating the peptidyl-resin with ultraviolet light (365 nm) in different solvents, and with different radical initiators, the authors were able to determine optimal conditions (solvent and photo-initiator) for the conversion of the linear substrate to the cyclic product.

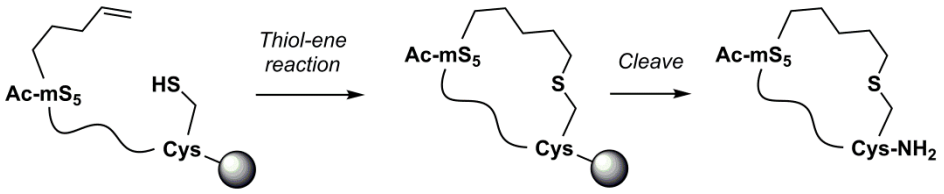

Figure 1. Generalised scheme illustrating the approach of Zhao *et al.* to on-resin peptide macrocyclisation.

Using the optimised conditions, the effects of two variables relating to the substrate's structure were also examined: the identities of residues *i* + 1, *i* + 2 and *i* + 3 (all of which had been alanine in the initial model), and the number of methylene groups separating the *α*-carbon and the terminal alkene in the *α*-alkenylglycine (two, three or four methylenes). Finally, the authors were able to demonstrate the applicability of their method to a much more complicated and biologically-relevant substrate (an eleven-residue estrogen-receptor binding peptide). It was shown that the product could bind its target with higher affinity than either the native ligand or a linear construct that had not been subjected to thiol-ene reaction conditions. Interestingly, however, the affinity of the linear construct was closer to that of the cyclic product than that of the native ligand. This important control experiment illustrates the potential complexity that can underlie apparently straightforward changes in binding affinity.

CYSTEINE AND METALLOTHIONEINS

The second example of cysteine's rich chemistry comes from the group of Eva Freisinger at the University of Zurich. Writing in *Biopolymers (Peptide Science)*, Tarasava *et al.* describe the effects of oxidants on the composition and structure of an isolated domain from a wheat metallothionein (Tarasava *et al.*, 2016). "Methallothionein" (MT) refers to a class of small intracellular metalloproteins with a remarkably high cysteine content and, in wheat (*Triticum aestivum*), there exists a MT called E_c-1 ("early cysteine-labelled protein"). The discovery of a wheat E_c protein was first reported in 1983 by Hanley-Bowdoin and Lane, who monitored the incorporation of ^{35}S-labelled cysteine in wheat germ. The finding of this early work was that an especially large amount of cysteine was committed to the synthesis of E_c during the first hour of embryo development (hence "early").

By virtue of the numerous cysteine residues (as thiolates), MTs are able to co-ordinate substantial numbers of transition metal ions, and have therefore been implicated in mechanisms of metal detoxification and/or homeostasis. Additionally, on account of their redox properties, it has been suggested that MTs might be scavengers of reactive oxygen species (Hassinen *et al.*, 2011). In wheat, the "holo" (i.e., metal-containing) form of E_c-1 can contain six bound zinc ions, which are distributed between two domains (β and γ). To investigate the behaviour of the γ-domain in isolation, Tarasava *et al.* prepared a 26-residue recombinant peptide corresponding to the "apo" (i.e., metal-free) form of the γ-domain (γ-E_c-1). This peptide contains six cysteine residues and is capable of binding up to two zinc ions. Thus, γ-E_c-1 served as minimal *in vitro* model for metal binding and intramolecular disulfide bond formation in MTs. Using UV-visible spectroscopy, the authors were able to monitor changes in the structure of "holo" γ-E_c-1 occurring upon addition of oxidants. Complete oxidation of "holo" γ-E_c-1 caused loss of both bound zinc ions and concomitant formation of three disulfide bonds. To elucidate the structural transitions associated with these chemical changes, the authors used tandem mass spectrometry to map disulfides, and circular dichroism spectroscopy to monitor the peptide's conformation. Partial oxidation of "holo" γ-E_c-1 was found to yield a heterogeneous mixture of species, each of which contained a variable number of zinc ions (zero or one) and a variable number of disulfides (one or two). On the other hand, fully oxidised "holo" γ-Ec-1 was found to be structurally homogenous, containing three specific disulfide bonds (Figure 2).

Figure 2. The pattern of disulfide bonds observed by Tarasava *et al.* in fully-oxidised γ-E_c-1.

Interestingly, this same specific pattern of disulfides was seen when "apo" γ-E_c-1 was oxidised fully. Given that the ordered arrangement of thiol groups around the zinc ions in "holo" γ-E_c-1, the specific pattern of disulfides formed by this structure is perhaps not surprising. More surprising, though, is the formation of this same product from "apo" γ-E_c-1, in which the authors had detected random coil behaviour using NMR spectroscopy. Finally, experiments in which the "holo", fully-oxidised and fully-reduced variants of γ-E_c-1 were digested with proteinase K revealed different sensitivities to proteolysis: the fully-reduced peptide was relatively sensitive, whereas the "holo" and fully-oxidised variants were relatively insensitive. The authors relate this finding to a potential effect of MT oxidation *in vivo*, namely that the formation of disulfide bonds might prevent an MT from being turned over unnecessarily when the concentration of metal ions is low.

REFERENCES

Aimetti A, Shoemaker R, Lin C-C, Anseth, K. On-resin peptide macrocyclization using thiol-ene click chemistry. *Chem. Commun.* 2010, 46, 4061-4063. PMID: 20379591

Hanley-Bowdoin L, Lane BG. A novel protein programmed by the mRNA conserved in dry wheat embryos. The principal site of cysteine incorporation during early germination. *Eur. J. Biochem.* 1983, 135, 9-15. PMID: 6688396

Hassinen VH, Tervahauta AI, Schat H, Karenlampi SO. Plant metallothioneins – metal chelators with ROS scavenging activity? *Plant Biol.* 2011, 13, 225-232. PMID: 21309968

Lau YH, de Andrade P, Wu Y, Spring DR. Peptide stapling techniques based on different macrocyclisation chemistries. *Chem. Soc. Rev.* 2015, 44, 91-102. PMID: 25199043

Tarasava K, Chesnov S, Freisinger, E. Oxidation of the N-terminal domain of the wheat metallothionein Ec-1 leads to the formation of three distinct disulfide bridges. *Biopolymers Peptide Sci.* 2016, 106, 295-308. PMID: 27061576

Vickery HB, Schmidt CLA. The history of the discovery of the amino acids. *Chem. Rev.* 1931, 9, 169-318.

Zhao B, Zhang Q, Li Z. Constructing thioether-tethered cyclic peptides via on-resin intra-molecular thiol-ene eeaction. *J. Peptide. Sci.* 2016, 22, 540-544. PMID: 27378284