This is a repository copy of *Palladium-unleashed proteins: gentle aldehyde decaging for site-selective protein modification*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/128868/

Version: Published Version

**Article:**

**Reuse**
This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here: https://creativecommons.org/licenses/

**Takedown**
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
Palladium-unleashed proteins: gentle aldehyde decaging for site-selective protein modification†

Robin L. Brabham, a Richard J. Spears, a Julia Walton, a Swati Tyagi, b Edward A. Lemke b and Martin A. Fascione a *

Protein bioconjugation frequently makes use of aldehydes as reactive handles, with methods for their installation being highly valued. Here a new, powerful strategy to unmask a reactive protein aldehyde is presented. A genetically encoded caged glyoxyl aldehyde, situated in solvent-accessible locations, can be rapidly decaged through treatment with just one equivalent of allylpalladium(II) chloride dimer at physiological pH. The protein aldehyde can undergo subsequent oxime ligation for site-selective protein modification. Quick yet mild conditions, orthogonality and powerful exposed reactivity make this strategy of great potential in protein modification.

Aldehydes are a powerful yet underutilised tool for bioorthogonal chemistry, where the high electrophilicity coupled with good stability and low abundance in nature make them an attractive handle for protein modification. Bioorthogonal reactions involving aldehydes have been developed to take advantage of the unique reactivity of this functional group and an impressive array of bioconjugates are synthetically accessible, including antibody-drug conjugates, protein–protein conjugates and labelled live cells.

Access to such aldehydes, however, can be an impediment to their usage, with incorporation methods either requiring enzyme recognition sequences, location at a protein terminus, or both. Use of formylglycine-generating enzyme (FGE), for example, will only form an aldehyde on the side chain of a cysteine in a CXPXR sequence (Fig. 1a). Some strategies are less flexible for aldehyde positioning: periodate-mediated oxidative cleavage of serine or threonine residues or transamination of glycine residues occurs only at such N-terminal residues, whilst tubulin tyrosine ligase (“Tub”) requires a Tub tag on a protein C-terminus to append tyrosine derivatives such as m-formyl-L-tyrosine 1 (Fig. S1, ESI†).

The technique of unnatural amino acid (UAA) mutagenesis has become a standard tool in chemical biology. Use of the pyrrolysine (Pyl) tRNA CUA/pyrrolyl-tRNA synthetase (RS) pair from several species of archaeal methanogens for amber stop codon (TAG) suppression has allowed access to proteins containing a wide range of non-canonical functionality, including alkenes, alkynes, azides, and aryl halides, with generally excellent levels of site specificity; indeed, UAA mutagenesis has become a widely utilised tool in chemical biology. Notably, the aldehyde-containing UAA m-formyl-L-phenylalanine 2 (Fig. S1, ESI†) has been genetically encoded using an engineered pylRS variant from Methanosarcina mazei. The wild-type Methanosarcina barkeri pyrrolysine-tRNA-RS pair genetically encodes 2-thiazolidine derivative ThzK, with the methyl ester 3 used as a suitable precursor for incorporation into proteins (Fig. 1c).

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c7cc07740h

Fig. 1 Selected methods for the installation of aldehydes in proteins.
has been used as a caged pseudo-N-terminal cysteine for protein CBT condensation ligations, showing the potential for a 2-thiazolidine to be used to smuggle an aldehyde group into a protein. Notably, a thiazolidine will decage to yield a glyoxyl group, a highly reactive aldehyde to the extent that more reliable, reproducible and standardised modification methodologies have been established for this protein aldehyde than any other. As reactive carbonyls have been documented as forming undesired adducts with 1,2-aminothiols in biological media, the use of a protection/deprotection strategy avoids such side reactions which may suppress genetic incorporation or conjugation yields. Indeed, UAA mutagenesis has been shown to exhibit excellent synergy with decaging strategies, where photodecaging and metal-mediated decaging reactions have expanded the paradigm of functionality which can be genetically encoded. In this work a rapid yet mild palladium glyoxyl-decaging strategy is presented which reveals protein aldehydes from surface-exposed ThzK residues in a protein without the need for enzyme recognition sequences, using just a single equivalent of palladium (Fig. 1c).

The new aldehydes are subsequently shown to be amenable to site-selective modification.

Green fluorescent protein (GFP) and superfolder green fluorescent protein (sfGFP) are highly useful test systems for protein modification due to ease of visualisation and highly optimised expression systems for use with UAA mutagenesis. Two GFP mutants containing amber stop codons at surface-exposed sites were selected: sfGFP(N150TAG) and GFP(Y39TAG). An advantage of using GFP is the facile confirmation and assaying of successful stop codon suppression through green fluorescence of harvested cell pellets. Genetic encoding of ThzK has previously made use of the *M. barkeri* pyrrolysine tRNA-RS pair, although the promiscuity of the corresponding *M. mazei* pair is generally sufficiently similar to the extent that either pair would be likely to genetically encode ThzK. Protected ThzK was synthesised in four straightforward steps with a cumulative yield of ca. 45% (Scheme S1, ESI†) following literature precedent. Separate saponification is unnecessary as this amino acid can be delivered into protein expression systems in

---

**Fig. 2** (a) Reagents screened to decage protein thiazolidine GFP(Y39ThzK) 10 and sfGFP(N150ThzK) 11 to afford protein aldehyde/hydrate 12/13 (calculated masses shown). Table shows decaging conditions and conversions for protein thiazolidine 10. (b) Complete decaging of 10 (upper) using 16 to form 12 (lower) as the aldehyde (experimental masses shown). (c) Complete decaging of 11 (upper) using 16 to form 13 (lower) as the hydrate (experimental masses shown).
a mildly basic stock solution, exposing the C-terminus as needed for translation. Using an adapted general method for amber stop codon suppression in GFP expression, both mutants were individually expressed with the supplementation of 3 at 1.6 mM in growth media and the green fluorescence of the harvested cell pellets confirmed the presence of full-length protein, demonstrating that ThzK can be encoded by the M. mazei pyrrolysine tRNA-RS pair. Following nickel affinity purification of the cell lysate, both GFP(Y39ThzK) and sfGFP(N150ThzK) could be isolated and their purity confirmed by SDS-PAGE and ESI-FTICR-MS (Fig. S2 and S3, ESI†).

Following successful ThzK incorporation, 10 and 11 were used as test beds for novel decaging strategies to yield GFP(Y39GlyoxylK) and sfGFP(N150GlyoxylK) complexes are a mixture of Pd(0) and Pd(II) reagents for unmasking of glyoxyl aldehydes. Palladium complexes available palladium sources as potentially biologically compatible reagents for unmasking of glyoxyl aldehydes. Palladium complexes could be isolated and their purity confirmed by SDS-PAGE and ESI-FTICR-MS (Fig. S2 and S3, ESI†).

As further confirmation of the exposed aldehyde reactivity, aniline-catalysed oxime ligation was performed upon protein glyoxyl species 12 and 13 using an aminooxy biotin probe to afford biotinylated proteins 19 and 20 respectively (Fig. 3a). Pleasingly, complete ligation was observed with both proteins within 24 h, with a Western blot confirming the incorporation of the biotin probe in 19 (Fig. 3b). Further oxime ligation was performed using the fluorescent aminooxy dansyl probe with 12 and 13 to afford dansylated proteins 22 and 23 respectively. Again full conversion was observed and through denatured protein in-gel fluorescence the presence of a dansyl group in 23 could be unequivocally visualised (Fig. 3c).

Irrespective of differences in protein and aldehyde/hydration distribution, the aldehydes uncaged by the method reported here can be modified to completion following established protocols.

![Image](https://example.com/image.png)

**Fig. 3** (a) Oxime ligation of protein aldehydes 12/13 using biotin probe 18 or dansyl probe 21 (calculated masses shown). (b) Complete formation of biotinylated GFP 19 and biotinylated sfGFP 20 confirmed by MS (experimental masses shown) and accompanying Coomassie-stained SDS-PAGE (upper) and Western blot (lower). (c) Complete formation of dansyl proteins 22/23 confirmed by MS and accompanying Coomassie-stained SDS-PAGE (upper) & denatured protein in-gel fluorescence (lower).
In summary, a new way to uncap a genetically encoded glyoxyl aldehyde precursor at physiological pH has been demonstrated using stoichiometric Pd(II), facilitating access to internally-modified proteins through aldehyde ligations without the need for an enzyme recognition sequence and hence minimising structural perturbations. This method requires only short reaction times under gentle conditions and the resulting aldehyde can be modified in completion. It is hoped that this latest addition to the chemical biologist’s toolbox will open up opportunities for creating exciting new bioconjugates, achieving a greater understanding of complex biological systems.

We thank Dr Ed Bergstrom and the CoEMS for support with protein MS. This work was supported by the University of York (R. L. B.) and an EPSRC DTG studentship (EP/M506680/1, R. J. S.).

Conflicts of interest

R. L. B., R. J. S., & M. A. F. are authors on PCT/GB/2017/052896 application which in one claim covers Pd-mediated decaging.

References