A tale of two bioconjugations: pH controlled divergent reactivity of protein α-oxo aldehydes in competing α-oxo-Mannich and catalyst-free aldol ligations

Tessa Keenan[1], Richard J. Spears[1], Saeed Akkad[1], Clare S. Mahon[2], Natasha E. Hatton[1], Julia Walton[1], Amanda Noble[1], Nicholas D. Yates[1], Christoph G. Baumann[3], Alison Parkin[1], Nathalie Signoret[4], Martin A. Fascione[1]*

[1]Department of Chemistry, University of York, York, UK, YO10 5DD; [2]Department of Chemistry, Durham University, UK, DH1 3LE Durham; [3]Department of Biology, University of York, York, UK, YO10 5DD; [4]Hull York Medical School, University of York, York, UK, YO10 5DD

*Corresponding author, martin.fascione@york.ac.uk

ABSTRACT

Site-selective chemical methods for protein bioconjugation have revolutionised the fields of cell and chemical biology through the development of novel protein/enzyme probes bearing fluorescent, spectroscopic or even toxic cargos. Herein we report two new methods for the bioconjugation of α-oxo aldehyde handles within proteins using small molecule aniline and/or phenol probes. The 'α-oxo-Mannich' and 'catalyst-free aldol' ligations both compete for the electrophilic α-oxo aldehyde which displays pH divergent reactivity proceeding through the "Mannich" pathway at acidic pH to afford bifunctionalised bioconjugates, and the "catalyst-free aldol" pathway at neutral pH to afford monofunctionalised bioconjugates. We explore the substrate scope and utility of both these bioconjugations in the construction of neoglycoproteins, in the process formulating a mechanistic rationale for how both pathways intersect with each other at different reaction pH.

INTRODUCTION

Methods to site-selectively adorn biomolecules with small molecules is of major interest within the field of chemical biology as modification with functional moieties can vastly enhance their properties.[1-2] For example, bioconjugation of compounds such as polyethyleneglycol can improve the half-life of protein probes and therapeutics,[3-4] whilst ligation of fluorescent/spectroscopic probes has been utilised for *in vivo* imaging and tracking of proteins, carbohydrates and DNA.[5-9] Furthermore the ability to chemically generate proteins bearing 'mimics' of post-translational modifications such as glycosylation,[10-15] has armed biologists with tools to study modified proteins that may be otherwise challenging to obtain using standard recombinant expression techniques.

However, the development of new methods in this field has relied heavily on the modification of a small number of amino acid residues which are surface exposed or in low natural abundance,[16-26] or well-studied unnatural azide, alkyne or olefin handles.[27-30] Alternatively, non-proteinogenic aldehyde chemical handles which have also been incorporated into biomolecule scaffolds[31] are relatively underexplored in comparison. In particular, α-oxo aldehydes,[32] which can now be routinely

incorporated site-selectively at both *N*-terminal and internal positions within proteins[33-35] and offer great potential for the development of novel bioconjugations due to their uniquely electrophilic nature. Although recent studies have shown these α-oxo aldehydes are reactive in C-C ligations,[36-37] much of the work in this field has been focussed on the synthesis of heteroatom linked thiazolidine,[38] oxime and hydrazone bioconjugates from aldehyde precursors.[39-41] These transformations have principally employed anilines as organocatalysts leading to large increases in reaction rates at neutral pH or below.[42-46] This acceleration is facilitated by the ability of the aniline to rapidly react with the aldehyde under physiological conditions and form a more reactive Schiff base, which is then attacked by a reactive α-effect nucleophile resulting in the expulsion of the aniline catalyst in an example of *trans*-imination. Prior to this widespread application of anilines as organocatalysts in bioconjugations however, pioneering studies by Francis and co-workers demonstrated anilines could also be incorporated into protein bioconjugates by exploiting exposed tyrosine residues as nucleophiles to trap small molecule Schiff bases in a "three-component Mannich" reaction.[47] Although powerful in its ability to generate useful multicomponent products this Mannich transformation has been underutilised since its discovery and is potentially limited by lack of site-selectivity due to the abundance of exposed phenolic tyrosine residues within proteins. To overcome this issue, in this study we explored whether proteins bearing site-selectively installed α-oxo aldehyde handles could act as electrophilic scaffolds for novel α-oxo-Mannich bioconjugations wherein the aniline and the phenol components are both small molecules. Herein, not only do we describe the realisation of this approach in the development of multicomponent α-oxo-Mannich ligations on proteins, but also a divergence in the reactivity of α-oxo aldehydes in the presence of anilines and phenols at different pH (Fig. 1). Notably at acidic pH α-oxo aldehydes react as anticipated through a "Mannich" pathway to afford bifunctionalised bioconjugates, but at neutral pH we demonstrate phenols react directly with α-oxo

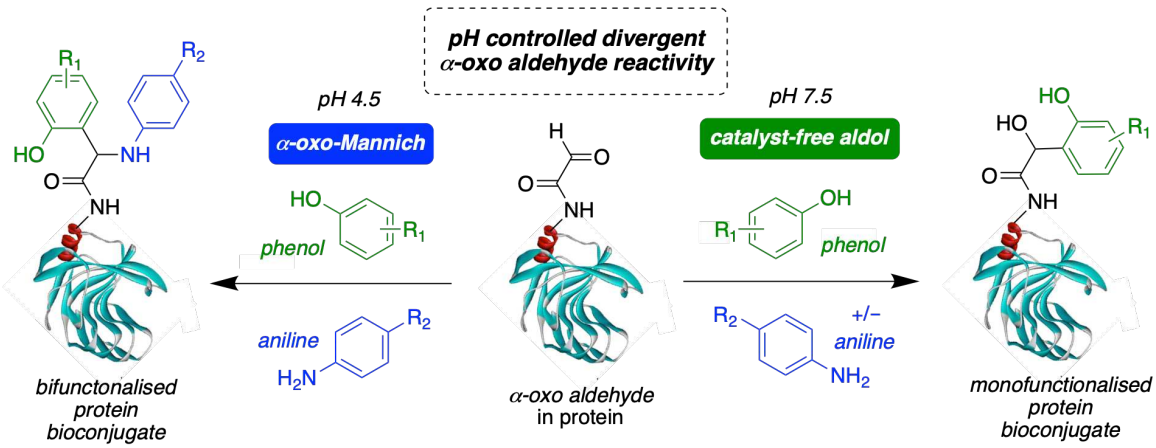


Fig. 1 pH dependant α-oxo-Mannich and catalyst-free aldol bioconjugation of protein α-oxo aldehydes using anilines and phenols.

aldehydes in a rapid "catalyst-free aldol" pathway to afford monofunctionalised bioconjugates. We use model peptides and proteins to explore the substrate scope and utility of both these bioconjugations and in the process formulate a mechanistic rationale for how both pathways compete and intersect with each other depending on reaction pH.

RESULTS

Peptide feasibility studies

We investigated the feasibility of the proposed α-oxo-Mannich bioconjugation using a model peptide α-oxo-aldehyde-VARLG **1** that lacked tyrosine residues. Reactions were assembled with α-oxo-aldehyde-VARLG **1**, commercially available aniline **2** and substituted 4-methoxy phenol **3** or 3,5-dimethoxy phenol **4** in phosphate buffer (PB) at pH 6.5 (Fig. 2a) to replicate conditions previously reported for Mannich reactions on tyrosine residues in proteins.[48-49] Relative conversion to the anticipated α-oxo-Mannich products **5** and **6** were assessed by LC-MS analysis. Pleasingly the anticipated α-oxo-Mannich products were observed in both reactions and these were validated by MS/MS analysis (Fig. S5-S6 and S8). However, only 12% relative conversion to α-oxo-Mannich product **5** was observed in the reaction with 4-methoxy phenol **3** with a significant quantity of starting material **1** remaining suggesting that the conditions required significant optimisation. Additionally, in the reaction with the more electron rich 3,5-dimethoxy phenol **4**, a significant amount of an unanticipated species **7** at 725.43 m/z was observed alongside α-oxo-Mannich product **6**. Interestingly, the *m/z* value of the unanticipated species **7** corresponds to the mass of α-oxo-aldehyde-VARLG **1** + the mass of phenol **4** (+1 Da for $[M+H]^+$). Subsequent MS/MS analysis of species **7** yielded a fragmentation pattern similar to that described for aldol-modified peptides,[37] with the modification identified at the *N*-terminus and a highly intense peak of species -18 Da, which upon fragmentation yielded x/y/z fragments corresponding to the "VARLG" peptide (Fig. 2b, Fig. S9-S10). As modifications of small molecule aldehydes[50] and other protein electrophiles[51] have previously been reported under aqueous conditions using phenol reagents proposed to act as enolate equivalents, we hypothesised that phenol **4** may be participating as a preformed enol equivalent in a direct aldol-type conjugation with α-oxo-aldehyde VARLG **1**. To validate this hypothesis, α-oxo-aldehyde VARLG **1** was incubated with different concentrations of phenol **4** in PB at pH 6.5 for 24 h in the absence of aniline and 95% relative conversion to the anticipated aldol product **7** was observed using only 1 molar equiv. of phenol, with complete conversion achieved using 5 molar equiv. (Fig. S13). We anticipated a 'catalyst-free aldol' bioconjugation of this type might be of significant interest in the field due to the simplicity of the phenol probe and were intrigued to note, perhaps unsurprisingly, that precedent for such reactivity already exists in the study of reactive α-oxo aldehydes *in vivo*. These include highly electrophilic

glyoxal, methyl glyoxal and 2-deoxyglucosone which are produced endogenously by glycolysis[52-53] or exogenously by the Maillard reaction between sugars and amino acids in foods and can react non-enzymatically with nucleophilic amino residues in proteins.[54] The resulting modified proteins are known as advanced glycation end products (AGEs) and their build up is associated with diabetic

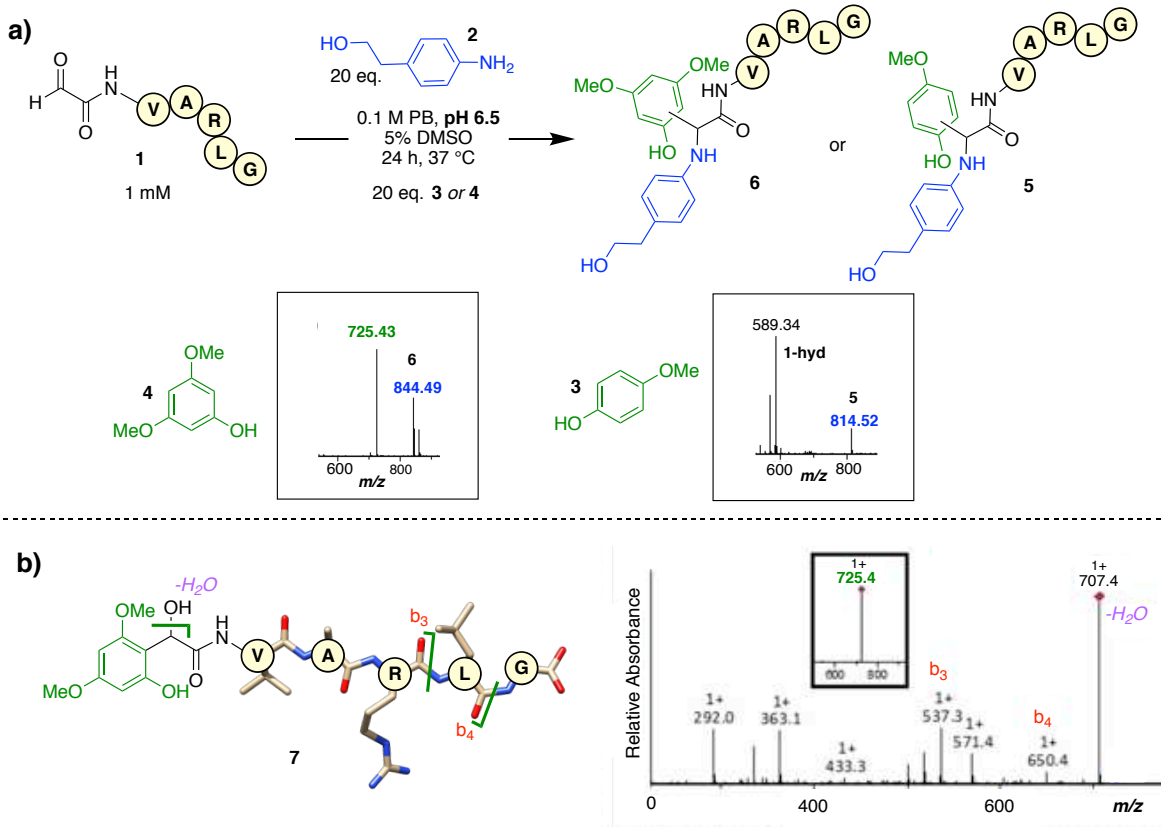


Fig. 2 a) Outline of reaction of α-oxo-aldehyde-VARLG **1** with aniline **2** and phenol **3** or **4**, formation of a mixture of stereoisomers is likely and therefore stereochemistry is omitted for clarity here and throughout. Calculated $[M+H]^+$ of product **5** = 814.44 Da. Calculated $[M+H]^+$ of product **6** = 844.45 Da. Inset: LC-MS spectra of the reaction products. b) MS^2 spectrum of phenol peptide product **7**. The connectivity between phenol **4** and aldehyde **1** is predicted on the basis of precedent established in the reaction between methyl glyoxal and catechins and used from hereon. PB = phosphate buffer.

complications and a number of other age-related pathologies,[55-57] however AGE formation can be reduced by polyphenols including catechins and theaflavins present in green and black teas,[58-60] which have been demonstrated to trap these reactive α-oxo-aldehydes in aldol-type reactions under physiological conditions[60-62] akin to those employed here. Notably the substitution was unequivocally proven to occur between a phenol and alkoxy substituent[54] (connectivity depicted in **7**, Fig. 2b, and adopted from herein), and reinforces the reactivity we observe here using 3,5-dimethoxy phenol **4** and the α-oxo aldehyde peptide **1**.

Controlling reaction pathways through pH

Having established that both an α-oxo-Mannich and a catalyst-free aldol bioconjugation were taking place simultaneously at the α-oxo aldehyde centre in pH 6.5 buffer when using both phenol **4** and

aniline **2** as reactants, we set out to determine whether the progress through each reaction pathway could be controlled by changing the pH. The three-component Mannich bioconjugation was previously reported to have an optimal reactivity observed at pH 5.5-6.5 and poor reactivity observed above pH 8,[47] with a cyclic transition state proposed at pH 6.5 which simultaneously activated both the electrophilic imine and the nucleophilic phenol.[47] On peptides however Mannich reactions have also been reported to proceed over a much wider pH range, with reactions proceeding at a pH as low as 2.[63] Conversely a phenol is more likely to react as a nucleophilic enolate/enol equivalent in a catalyst free-aldol bioconjugation at more basic pH.[50-51, 64] We therefore opted to incubate the α-oxo aldehyde-VARLG **1**, aniline **2** (20 equiv.), and phenol **4** (5 equiv.) for 24 h at 37 °C in buffer, at both pH 4.5 and pH 7.5 (Fig. 3). Notably at pH 4.5, the modified α-oxo-Mannich species **6** was now formed as the major product with the aldol species **7** the minor byproduct in a reversal of what was previously observed at pH 6.5 (Fig. 1a), indicating the anticipated preference for the α-oxo-Mannich pathway at more acidic pH. While at pH 7.5 the aldol product **7** was formed almost exclusively, with only trace α-oxo-Mannich product formed, indicating a preference for the aldol pathway at more basic pH. This observation was further supported when the catalyst-free aldol was again performed in the absence of aniline, with 60% relative conversion to aldol product **7** observed at pH 4.5 and complete conversion observed at pH 7.5 (Fig. S14). Thus screening the same reaction mixture at different pH had enabled us to identify two novel bioconjugations strategies- a modified α-oxo-Mannich ligation which has the potential to afford dual functionalised proteins in a single transformation, and a catalyst-free aldol ligation which has the potential to afford monofunctional bioconjugates at mild neutral pH.

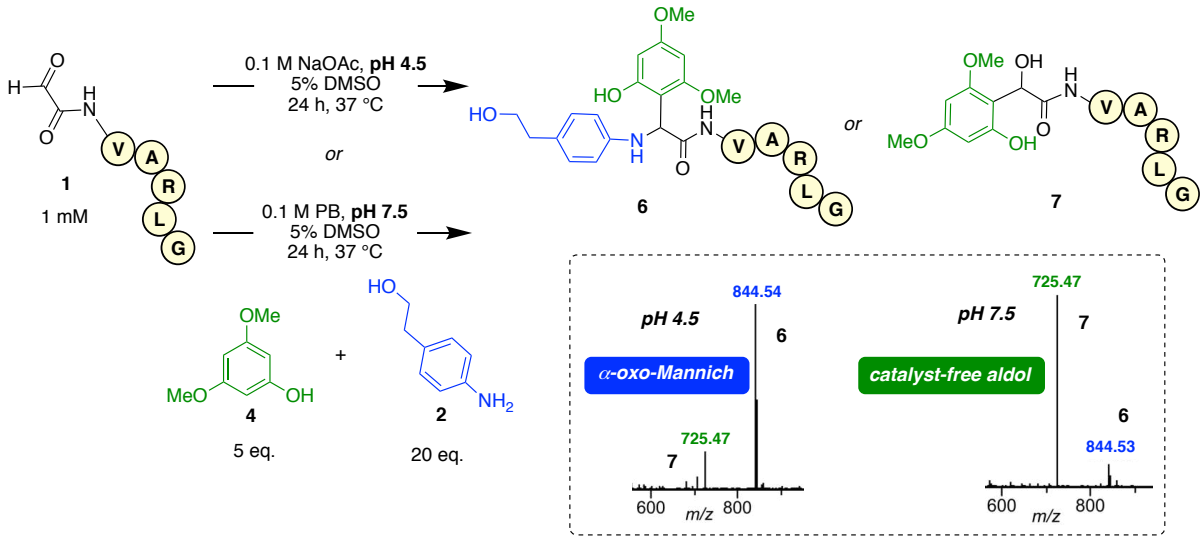


Fig. 3 Outline of reaction of α-oxo-aldehyde-VARLG **1** with aniline **2** and phenol **4** at pH 4.5 or pH 7.5. Inset: LC-MS spectra of the reaction products. Calculated $[M+H]^+$ of product **7** = 725.38 Da. Calculated $[M+H]^+$ of product **6** = 844.45 Da.

Screening differentially substituted anilines for the α-oxo-Mannich bioconjugation

Next we set out to optimise the conversion of the modified-α-oxo-Mannich at pH 4.5 using phenol **4** (Fig. 4) by screening a range of differentially substituted anilines **8-21**. Reactions were performed in duplicate using the α-oxo aldehyde peptide **1** and afforded either starting material **1** (orange in the bar chart), catalyst-free aldol product **7** (green), or α-oxo-Mannich product **22** (blue). Notably changing the nature of the aniline significantly affected the balance of the reaction at pH 4.5 and dictated whether the catalyst-free aldol or the α-oxo-Mannich reaction pathway was followed. For example, in the presence of electron poor anilines such as 4-nitroaniline **8** no α-oxo-Mannich product is observed with small amounts of unreacted peptide and aldol product dominating. This was also the case for anilines **9** and **10** which were substituted with electron withdrawing groups at the ortho position, wherein conversions to aldol product **22** were similar to that observed for the negative control [absence of an aniline (-)]. In contrast more basic anilines such as **16-21** which lack electron withdrawing groups and have pKa values in the 4-5.5 range afforded mostly α-oxo-Mannich product **22**, with 4-methoxyaniline **21** performing best with ~90% relative conversion. As previously noted anilines have been extensively used as nucleophilic organocatalysts in hydrazone/oxime bioconjugations of aldehydes in aqueous conditions,[39, 46] where the bioconjugations proceed via attack on a protonated aniline Schiff base intermediate[65-66] akin to that expected to form during an α-

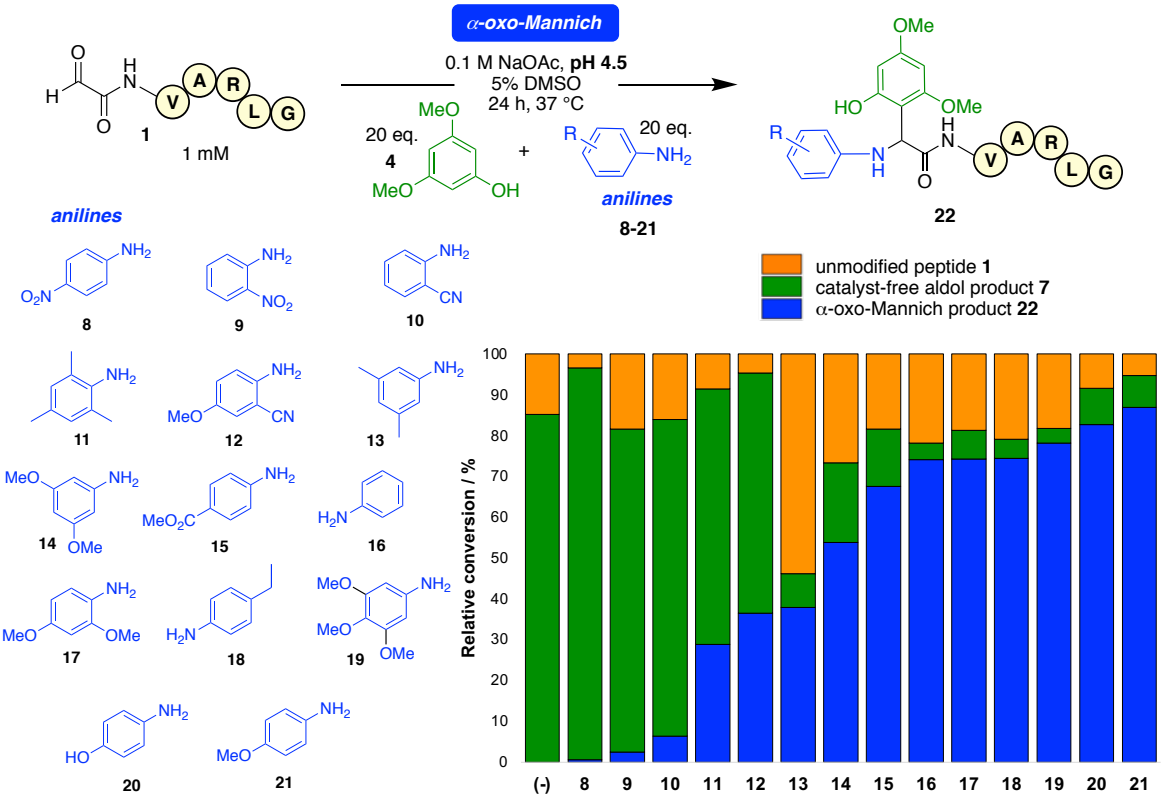


Fig. 4 Outline of the α-oxo-Mannich bioconjugation of α-oxo-aldehyde-VARLG **1** (orange in the bar chart) with anilines **8-21** and phenol **4** at pH 4.5, to afford α-oxo-Mannich product **22** (blue) or catalyst-free aldol product **7** (green).

oxo-Mannich reaction mechanism.[65] The most effective of these aniline organocatalysts studied often have a more basic pKa, closer to the pH of the reaction mixture, as this promotes protonation of the aniline Schiff base thus making it more electrophilic. It is therefore unsurprising that at pH 4.5, 4-ethylaniline **18** (pKa ~5.1), aniline **16** (pKa ~4.6),[65] and 4-methoxyaniline **21** (pKa 5.3)[65] afford mostly α-oxo-Mannich product. However, 2,4,6-trimethylaniline **11** (pKa ~4.4) affords little α-oxo-Mannich product with aldol product predominating, indicating the situation is likely governed by the interplay between both steric and electronic factors. Overall higher α-oxo-Mannich conversions were observed with electron rich anilines which presumably form more stable Schiff base intermediates. Indeed when comparing the equilibrium LCMS conversion to Schiff base intermediates upon incubation of the α-oxo peptide **1** with either 4-nitroaniline **8**, 2,4,6-trimethylaniline **11**, 3,5-dimethylaniline **13**, aniline **16** or 4-ethylaniline **18** at pH 4.5, a clear trend between higher conversion to the Schiff base and higher α-oxo-Mannich conversion was observed (Fig. S15). Notably an identical aniline screen at pH 7.5 afforded negligible α-oxo-Mannich conversions, and aldol conversions of >90% with all anilines (Fig. S16).

Site-selective *N-terminal* and *internal* α-oxo-Mannich bioconjugation for neoglycoprotein synthesis

Following optimisation studies on peptides we opted to utilise the α-oxo-Mannich procedure for the bioconjugations of proteins. Having established that 4-ethylaniline **18** afforded high conversion to the α-oxo-Mannich products we employed this aniline in combination with phenol **4** for the bioconjugation of the nanobody JVZ007, bearing an *N*-terminal α-oxo aldehyde easily accessed via rapid $NaIO_4$ oxidation of an *N*-terminal serine or threonine residue[37] or transamination of an N-terminal glycine.[67] JVZ007 binds specifically to the prostate specific membrane antigen (PSMA) overexpressed on prostate cancer cell surfaces,[68] and was successfully modified at pH 4.5 to afford the α-oxo-Mannich product in 68% conversion (Fig. S21), with 32% conversion to the aldol product **23**. As 4-methoxyaniline **21** afforded highest relative conversions to the α-oxo-Mannich product on the peptide we also synthesised a comparable para-derivatised aniline **24** bearing an α-D-mannose sugar. This sugar has potential therapeutic utility against urinary tract infections as an inhibitor of FimH, the type-one pilus subunit on the surface of uropathogenic *E. coli* responsible for adhesion to the urothelium.[69] The *N*-terminal α-oxo aldehyde JVZ007 could also be modified using this mannose derivatised aniline **24** to afford a neoglycoprotein **25**, affording 64% α-oxo-Mannich product and 36% conversion to the aldol product **23**(Fig. 5a and b).

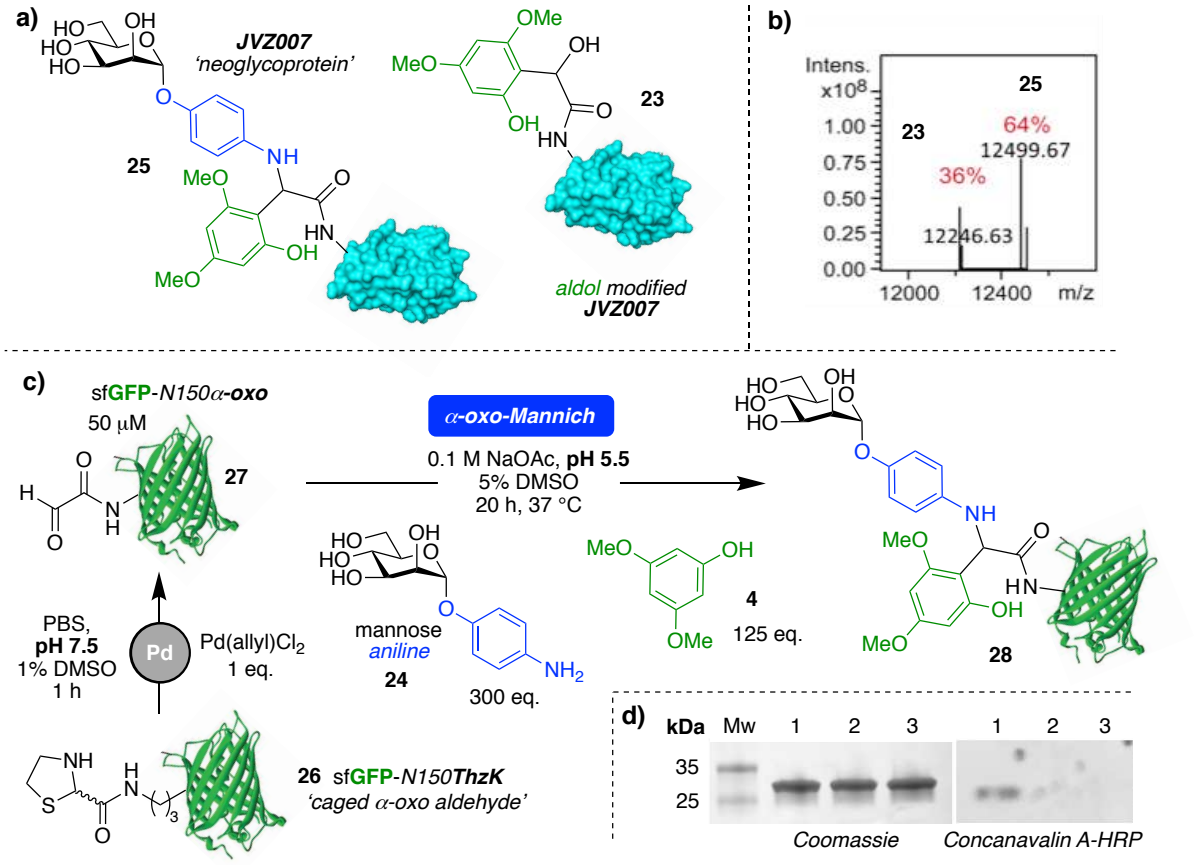


Fig. 5 a) Neoglycoprotein product **25** of the α-oxo-Mannich bioconjugation of JVZ007 (bearing an N-terminal α-oxo aldehyde) with phenol **4** and mannose aniline **24**. b) Deconvoluted LC-MS spectra showing the α-oxo-Mannich bioconjugation products. Aldol modified JVZ007 **23** calc. 12247, α-oxo-Mannich modified JVZ007 **25** calc. 12500. c) Outline of a two-step *internal* site-selective Pd-mediated decaging/α-oxo-Mannich bioconjugation to afford sfGFP neoglycoprotein **28**. Pd-mediated decaging of sfGFP-N150ThzK **26**, was followed by α-oxo-Mannich of sfGFP-N150α-oxo **27** (50 μM) with phenol **4** (125 equiv.), and mannose aniline **24** (300 equiv.) at pH 5.5 for 20 h. d) Left panel: SDS-PAGE analysis of the reaction. Mw) molecular weight ladder, 1) α-oxo-Mannich bioconjugation, 2) no aniline negative control, 3) no phenol negative control. Right panel: Concanavalin A lectin blot analysis of the reaction. 1) α-oxo-Mannich bioconjugation, 2) no aniline negative control, 3) no phenol negative control.

We next demonstrated that the α-oxo-Mannich could be also utilised to site-selectively modify an α-oxo aldehyde incorporated at an *internal* site within superfolder green fluorescent protein (sfGFP). The aldehyde was quantitatively installed at position 150 of sfGFP using a biocompatible Pd-mediated decaging (1 equiv. for 1 h at neutral pH) of an incorporated unnatural thiazolidine-lysine (ThzK) amino acid in sfGFP-*N*150*ThzK* **26** (Fig. 5c), recently developed in our lab.[35, 37] The resulting 'decaged' internal aldehyde within sfGFP-*N*150α-*oxoK* **27** was then functionalised via the α-oxo-Mannich using mannose aniline **24** and 3,5-dimethoxy phenol **4** to afford an internally modified mannose neoglycoprotein **28**. Although we observed marginally higher conversion to this α-oxo-Mannich neoglycoprotein **28** at pH 4.5 over pH 5.5, we also noted a 4-fold decrease in sfGFP fluorescence observed following incubation at pH 4.5 (Fig. S23) compared to pH 7.4. However incubation at pH 5.5 only resulted in a 1.6-fold

decrease in sfGFP fluorescence, indicating that the α-oxo-Mannich biconjugation is still a suitable method for yielding a site-selectively modified biologically functional sfGFP, despite its sensitivity to acidic pH.[70] The integrity and accessibility of the mannose aniline on the sfGFP neoglycoprotein surface was further confirmed in a lectin blot using a glycan binding Concanavalin A- Horseradish peroxidase (HRP) conjugate[71] (Fig. 5d right panel, lane 1), with no modification observed in the absence of the mannose aniline **24** (Fig. 5d, right panel, lane 2), or the 3,5-dimethoxy phenol **4** (Fig. 5d, right panel, lane 3).

Optimising mannose neoglycoprotein synthesis

Although the α-oxo-Mannich bioconjugation enabled the construction of neoglycoproteins on both JVZ007 and internally on sfGFP using mannose aniline **24**, competing formation of the catalyst free aldol product limited conversion to the α-oxo-Mannich product in 20 h reactions, particularly in the case of sfGFP (30% conversion to α-oxo-Mannich, 70% catalyst free aldol at pH 4.5, Fig. S25). Having established both the optimal aniline scaffold and pH for the α-oxo-Mannich bioconjugation, we therefore sought to increase reaction conversions through further optimisation of reaction conditions. As we anticipated the bioconjugation would proceed through attack on a protonated aniline Schiff base intermediate,[65-66] *vide supra*, we speculated that an increase in the concentration of aniline used in the reaction may also drive formation of the α-oxo-Mannich product over the catalyst free aldol product. To investigate the effect of increasing aniline concentration we utilised the Hydrophilic Acylated Surface Protein A (HASPA) from Leishmania, a highly immunogenic protein which is present in all human infective *Leishmania* parasites and a member of the HASP family of proteins which form the basis of a visceral leishmaniasis vaccine currently undergoing clinical trials in humans.[72] *N*-terminal α-oxo aldehyde HASPA[37] **29** was incubated at pH 4.5 for 20 h with 125 equiv. of 3,5-dimethoxy phenol **4** and an increasing concentration of mannose aniline **24** (75-1000 equiv., Fig. 6a). Using electrospray ionisation Fourier transform ion cyclotron resonance mass spectrometry (ESI-FTICR MS)[73] we were able to resolve MS peaks for α-oxo aldehyde HASPA starting material **29**, catalyst free aldol product **30**, and the α-oxo-Mannich HASPA product **31**, enabling accurate determination of relative conversion. In control reactions (Fig. 6b) lacking 3,5-dimethoxy phenol **4** or using a HASPA lacking an α-oxo aldehyde (non-oxidised) no modification occurs, while in the absence of the mannose aniline **24** the catalyst free aldol product **30** is predominantly observed. However, upon the inclusion of 75 equiv. of aniline **24** (Fig. 6c) a 53% relative conversion to α-oxo-Mannich product **31** is observed and increases to 58% when the aniline is doubled to 150 equiv., with only 38% relative conversion to catalyst free aldol product **30** observed. This trend continues as the mannose aniline **24** equiv. are increased further (Fig. 6d), with 60% relative to conversion to α-oxo-Mannich **31** and 38% conversion to catalyst free aldol product **30** observed at 300 equiv., and 75% relative conversion to α-oxo-

Mannich product **31** when using 1000 equiv. of aniline **24**. These results reinforce the notion that increased formation of the Schiff base affords higher relative conversion to the α-oxo-Mannich product over the catalyst free aldol product.

We also noted the effect of reaction time on product formation by comparing the relative conversions of α-oxo aldehyde HASPA **29** using 125 equiv. of 3,5-dimethoxy phenol **4** and 300 equiv. of mannose aniline **24** at pH 4.5 after 20 h and 6 h (Fig. S32-S33). Unexpectedly we observed a higher 90% relative conversion to α-oxo-Mannich product **31** (with ~8% catalyst free aldol product **30**) if the reaction was stopped at 6 h rather than after 20 h (60% α-oxo-Mannich product **31**; 34% catalyst free aldol product **30**). This trend was also conserved in the bioconjugation of the internal α-oxo aldehyde sfGFP **27** (Fig. S28) with higher conversions to α-oxo-Mannich product **28** observed when the reaction was stopped at 6 h, as opposed to 20 h. As the decrease in α-oxo-Mannich product between 6 h and 20 h was accompanied by an increase in catalyst free aldol product, we hypothesised that the α-oxo-Mannich product may be breaking down to the catalyst free aldol product down through the release of aniline

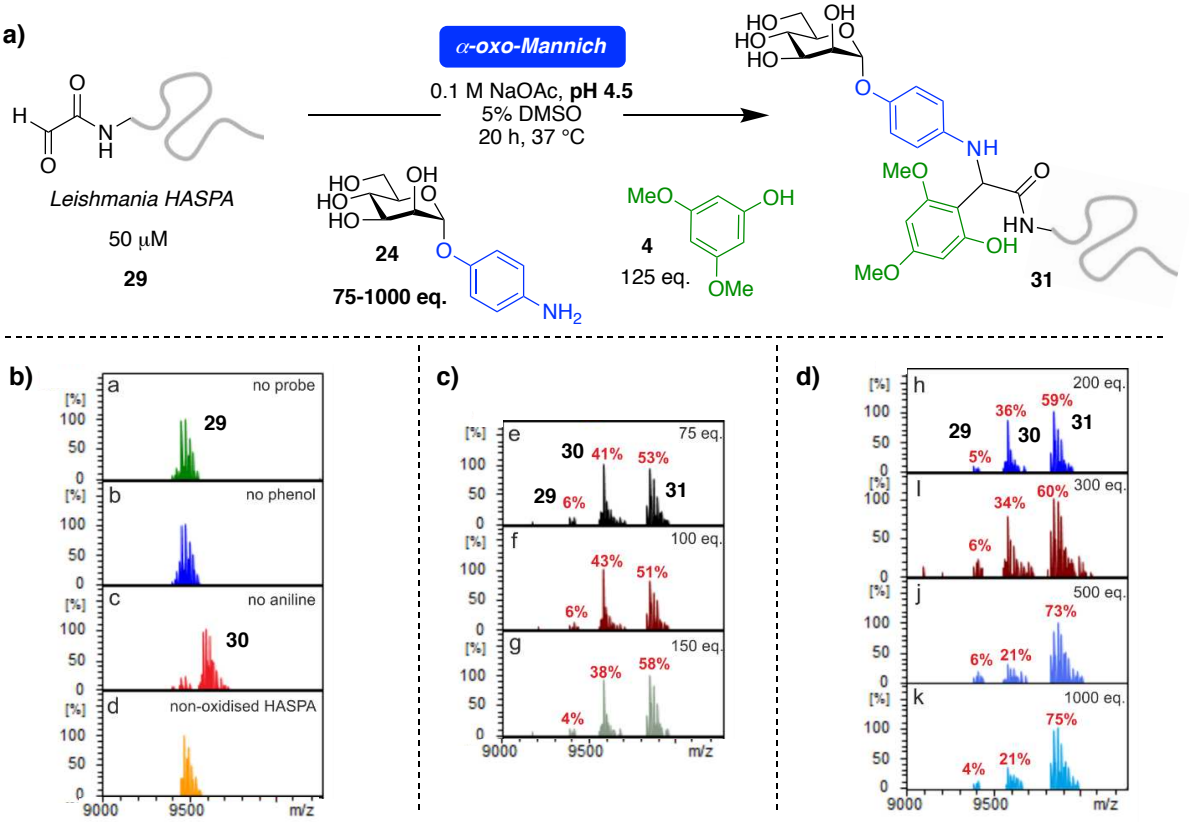


Fig. 6 Outline of the optimisation of the α-oxo-Mannich bioconjugation of Leishmania HASPA 33 (50 μM) at pH 4.5 for 20 h with phenol **4** (125 equiv.) and mannose aniline **24** (75-1000 equiv.) to afford HASPA neoglycoprotein **31**. Deconvoluted ESI-FTICR MS of reaction mixture with b) **a** no aniline or phenol probes, **b** no phenol **4**, **c** no mannose aniline **24**, **d** using non-oxidised Leishmania HASPA; c) **e** phenol **4** (125 equiv.) and mannose aniline **24** (75 equiv.), **f** phenol **4** (125 equiv.) and mannose aniline **24** (100 equiv.), **g** phenol **4** (125 equiv.) and mannose aniline **24** (150 equiv.); d) **h** phenol **4** (125 equiv.) and mannose aniline **24** (200 equiv.), **i** phenol **4** (125 equiv.) and mannose aniline **24** (300 equiv.), **j** phenol **4** (125 equiv.) and mannose aniline **24** (500 equiv.), **k** phenol **4** (125 equiv.) and mannose aniline **24** (1000 equiv.).

at pH 4.5. To validate this hypothesis we tested the stability of the α-oxo-Mannich products formed from 3,5-dimethoxy phenol **4** and 4-ethylaniline **18** on both the JVZ007 and HASPA proteins (50 μM protein concentration) in 0.1 M NaOAc at pH 4.5. Notably, in the absence of any excess aniline in the reaction mixture we observed complete breakdown of both the JVZ007 and HASPA α-oxo-Mannich products to the catalyst free aldol product within 6 h at acidic pH (Fig. S34-S36), indicating that to achieve optimal conversions to α-oxo-Mannich product it is preferable to use an excess of aniline probe over a shorter reaction time.

Optimising the catalyst-free aldol bioconjugation

Although competing formation and/or breakdown to the aldol product when attempting to perform multicomponent bioconjugations using the α-oxo-Mannich was problematic, this highlighted the potential utility of the catalyst-free aldol as a standalone bioconjugation capable of producing monofunctionalised bioconjugates. From our initial experiments with α-oxo aldehyde VARLG **1** at pH 6.5 using both 3,5-dimethoxy phenol **4** and 4-methoxy phenol **3** (Fig. 2) it was clear that the both compounds show significantly different reactivity towards ligation with the peptide. These results suggested that the substitution on the aromatic ring may impact the efficiency of the catalyst-free aldol bioconjugation. Therefore, to dissect the intricacies of the ligation, a further series of small molecule aromatic probes **32-37** were screened for their activity in modification of α-oxo aldehyde VARLG **1** (Figure 7). Using a 5-fold excess of probe over peptide at pH 7.5 we noted that when using electron rich aromatics which lacked a phenol substituent such as 1,3,5-trimethoxybenzene **32** and 1,3-dimethoxybenzene **35**, no bioconjugation occurred but when the phenol substituent(s) were reintroduced as in 3,5-dimethoxy phenol **4** or 1,3-resourcinol **35** complete relative conversion to the catalyst-free aldol VARLG product **38** was observed. This confirmed that the presence of the phenol group is essential for conversion, akin to the observations previously made in bioconjugations between electron rich aromatics and electrophilic selenocysteine residues,[51] and reinforcing the hypothesis that the phenol may act as an enol/enolate equivalents as has previously been reported in reaction with aldehydes in water.[50] When less electron rich dimethyl phenol **37** is used formation of the aldol product **38** is observed but at much lower relative conversion, and when an electron withdrawing ortho carboxylic acid is introduced in phenol **34** no conversion is observed. However on the introduction of a second phenol in to this scaffold in the case of diphenol **36** 49% conversion to the catalyst free aldol product **38** is observed suggesting that an intramolecular hydrogen bond

between the phenol and adjacent carbonyl in **34** may interfere with the bioconjugation, but can be compensated for by the introduction of another phenol elsewhere in the aromatic ring.

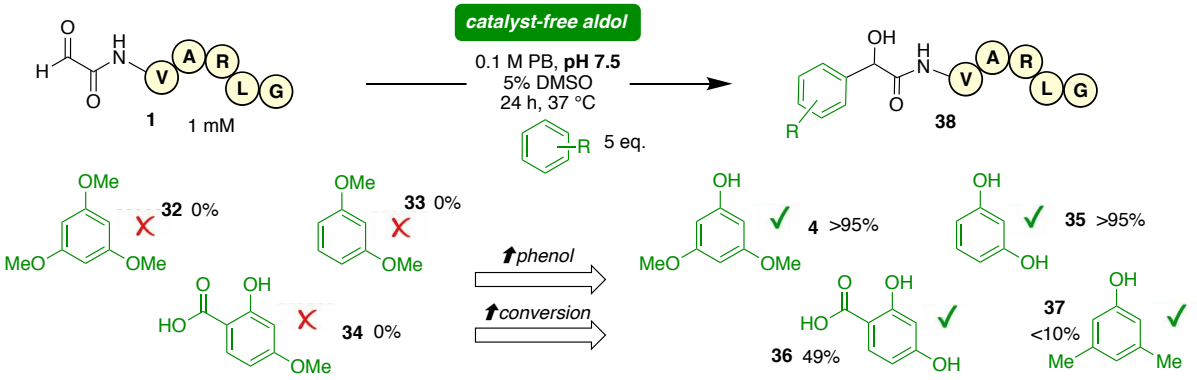


Fig. 7 Outline of the catalyst-free aldol reaction of α-oxo-aldehyde-VARLG **1** (1 mM) with aromatics **4, 32-37** (5 equiv.) at pH 7.5 for 24 h, demonstrating increased conversion to aldol product **38** upon the introduction of phenol substituents in the aromatic ring.

Catalyst-free aldol bioconjugation of proteins

Subsequently we demonstrated the catalyst-free aldol bioconjugation using 3,5-dimethoxy phenol probe **4** was also operative on proteins bearing α-oxo aldehyde residues, including myoglobin **39**,[67] Leishmania HASPA **29**, and JVZ007 **40** (Fig. 8a-c), with complete conversion to the aldol products observed within 1 h. However, to be of practical utility as a bioconjugation reaction the catalyst-free aldol would need to also be operative using modified phenol probes, thus enabling the introduction of new non-proteinogenic functionality into the protein target. We therefore synthesized a 1,3-resourcinol capped probe **43** containing a dansyl tag to enable the fluorescent labeling of a protein using the catalyst free aldol and screened increasing equivalents of this probe in the reaction with the HASPA protein **31** (Fig. 8d). SDS-PAGE and LC-MS analysis confirmed successful fluorescent modification of HASPA **29** when using >100 equiv. of probe thus demonstrating that simple phenol scaffolds can be used as the basis of complex functional probes for bioconjugation.

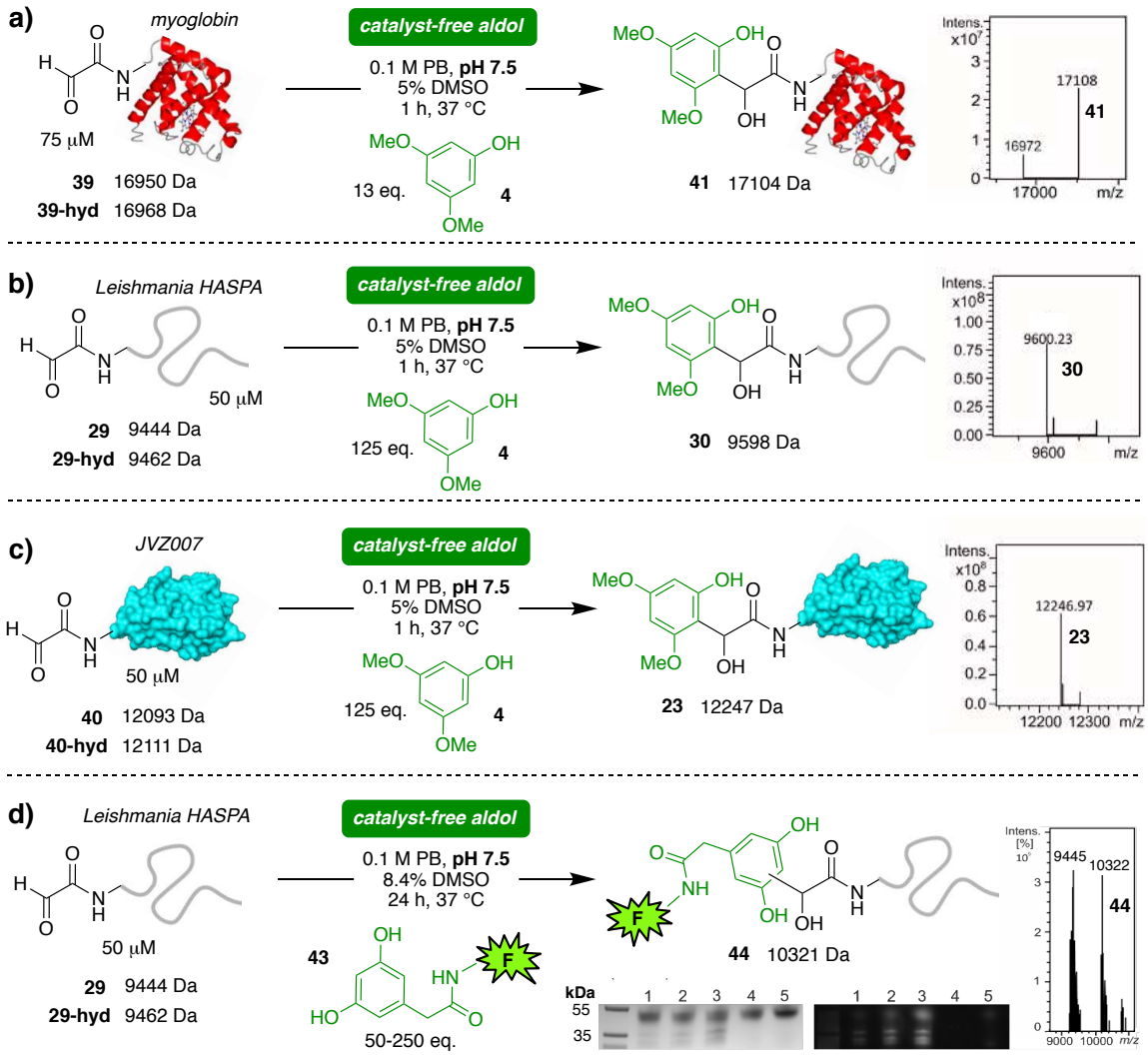


Fig. 8 Catalyst-free aldol reactions using phenol **4** on α-oxo aldehyde bearing proteins a) myoglobin **39**, b) Leishmania HASPA **29**, and c) JVZ007 **40**. Inset right: Deconvoluted LC-MS analysis of the respective reaction products. d) Catalyst-free aldol reaction of Leishmania HASPA **29** with fluorescent phenol **43**. Inset right: Deconvoluted LC-MS analysis of the reaction products. d) Inset below: SDS-PAGE analysis of the reaction with coomassie staining (left panel) and fluorescent visualization (right panel): 1) 50 equiv. **43**; 2) 100 equiv. **43**; 3) 250 equiv. **43**; 4) No probe **43** control; 5) Non-oxidised HASPA + **43** control. Substitution in **44** may occur *ortho/ortho* to both phenol substituents or *ortho/para*. HAPSA migrates aberrantly on SDS-PAGE due to its hydrophilicity[73] and migration is affected by bioconjugation.[37]

We also considered whether natural products containing phenols such as flavonoids, which have been shown to capture methyl glyoxal in water,[59-62] may also be capable of undergoing catalyst-free aldol ligation to proteins. We thus screened the plant secondary metabolite catechin **45** in reactions with the prostate cancer targeting nanobody **40** at pH 7.5 and observed full conversion to the aldol product **48** within 1 h (Fig. 9a). Additionally, we showed that the glucose bearing flavonoid phloridzin **47**, which is a competitive inhibitor of sodium-glucose cotransporter (SLGTs), was also capable of modifying the nanobody **40** to afford an aldol product **48** at pH 7.5 within 4 h (Fig. 9b). Once again the accessibility

of the glycan on the neoglycoprotein **48** surface was confirmed in a lectin blot using ConA-HRP (Fig. 9c). Notably SLGTs are functionally expressed in (prostate) cancer cells[75] to meet the increased metabolic demands for glucose in tumours[76] and therapeutic blockade with SLGT inhibitors can therefore lead to tumour necrosis.[77] However the therapeutic use of phlordizin is limited because as a potential treatment for type-2 diabetes it also non-selectivity targets SLGTs in the intestine and the kidney,[78] and additionally has poor oral bioavailability[79] due to glycoside hydrolysis by β-galactosidase

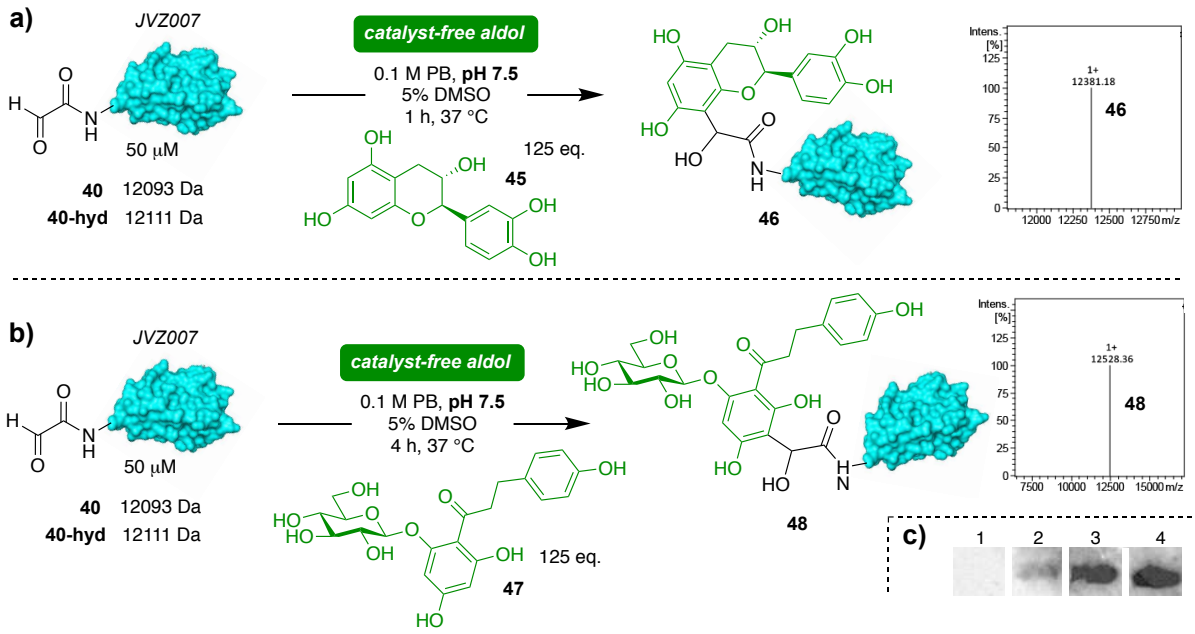


Fig. 9 Catalyst-free aldol bioconjugation of α-oxo aldehyde bearing JVZ007 **40** with a) catechin **45** and b) phloridzin **47**. Inset right: Deconvoluted LC-MS analysis of the respective reaction products. Catechin modified JVZ007 **48** calc. 12383; phloridzin modified JVZ007 **48** calc. 12529. Substitution in **46** and **48** is predicted on the basis of precedent established in the reaction between methyl glyoxal and catechins,[54] notably substitution in **48** may occur *para* or *ortho* to the glucose substituents. c) Concanavalin A lectin blot of 1) α-oxo JVZ007 **40**, 1.5 μg/mL; 2) phloridzin modified JVZ007 **48**, 0.5 μg/mL; 3) phloridzin modified JVZ007 **48**, 1 μg/mL; 4) phloridzin modified JVZ007 **48**, 1.5 μg/mL.

enzymes such as lactase. Our demonstration that phloridzin can easily undergo bioconjugation to a protein scaffold could therefore stimulate further studies into whether such constructs may have enhanced stability to glycosidases, and enhanced selectivity for inhibition of tumour SLGTs via antibody targeted delivery.

Mechanistic hypothesis and catalyst-free aldol bioconjugate stability

Based on the aforementioned experimental observations we formulated a mechanistic pathway (Fig. 10) for the formation of both the Mannich and aldol products and propose the α-oxo-Mannich bioconjugation initially proceeds through the formation of the protonated aniline Schiff base intermediate **49**. It is likely the equilibrium concentration of this intermediate in pH 4.5 buffer is governed by both the concentration of aniline present in the reaction mixture (see Fig. 6), and the

electron donating ability of the aniline (see Fig. 4). The Schiff base **49** then undergoes attack by the phenol to generate the α-oxo-Mannich product **50**. The competing catalyst-free aldol bioconjugation can also take place at pH 4.5 (and pH 7.5) and we propose the product **51** results from a direct aldol-type attack of the phenol on the α-oxo aldehyde **52**, wherein the phenol substituent is essential for conversion (see Fig. 7). We also observed that the α-oxo-Mannich product **50** undergoes breakdown to the catalyst-free aldol product at pH 4.5 within 6 h and hypothesise this occurs via the intermediate carbocation **53** which can be trapped by water to afford product **51**, or alternatively excess phenol to afford **54** which was identified as a minor byproduct. To confirm that the α-oxo-Mannich product **50** was formed via the Schiff base **49** and not by aniline attack on carbocation **53** following an initial catalyst-free aldol reaction, we also incubated purified VARLG-aldol product **7** with mannose aniline **24** in pH 4.5 buffer for 20 h, and observed no conversion from aldol to α-oxo-Mannich product over the course of the reaction (Fig. S49), unequivocally demonstrating that Mannich product formation does not result from initial aldol product formation. This may be because under the conditions of the reaction the (protonated) aniline nucleophile is unable to outcompete water in attack on carbocation **53**. In the absence of aniline the catalyst-free aldol reaction proceeds at pH 4.5 but more efficiently at pH 7.5 (Fig. S14) to afford the aldol bioconjugate within as little as 1 h. However, when we tested the stability of the aldol bioconjugates by re-incubating purified JVZ007-aldol bioconjugate **23** in buffer at 50 μM concentration, we were surprised to observe that while the conjugate was stable at pH 4.5, at pH 7.5 breakdown to the JVZ007 α-oxo aldehyde starting material **40** (~32%) occurred within 6 h (Fig. S51). As it is well established that the redox potential of phenols with adjacent proton acceptors are lower than that of simple phenols,[80] we hypothesised this breakdown (depicted in Fig 10 as **51** to **53**) may result from an initial one-electron oxidation of the electron rich phenol component of the aldol product facilitated by intramolecular H-bonding interactions at pH 7.5. Indeed the ubiquitous enzyme complex Photosystem II contains a Tyr_Z-His190 hydrogen bonding pair in its reaction centre which undergoes facile proton-coupled electron transfer (PCET) resulting in the generation of a neutral phenolic radical,[81] and a number of biomimetic models of this system have been constructed[82-83] which resemble elements of the aldol products described here. Notably these models have been shown to undergo facile concerted one-electron two-proton transfer processes. Therefore to explore whether a one-electron oxidation may trigger breakdown we re-screened the stability of JVZ007-aldol bioconjugate **23** at pH 7.5 in the presence of antioxidant sodium ascorbate.[84] Whilst breakdown of the aldol bioconjugate **23** to JVZ007 α-oxo aldehyde **40** was still observed in the presence of 1 mM ascorbate, at a higher concentration of 6 mM the aldol bioconjugate **23** appeared to show reduced breakdown over 6h (Fig. S52). Although electrochemical, spectroscopic and DFT studies will be

undoubtedly be required to unequivocally prove the hypothesised mechanism, this result tentatively suggests the aldol product may be unstable at low concentrations as a result of phenol oxidation.

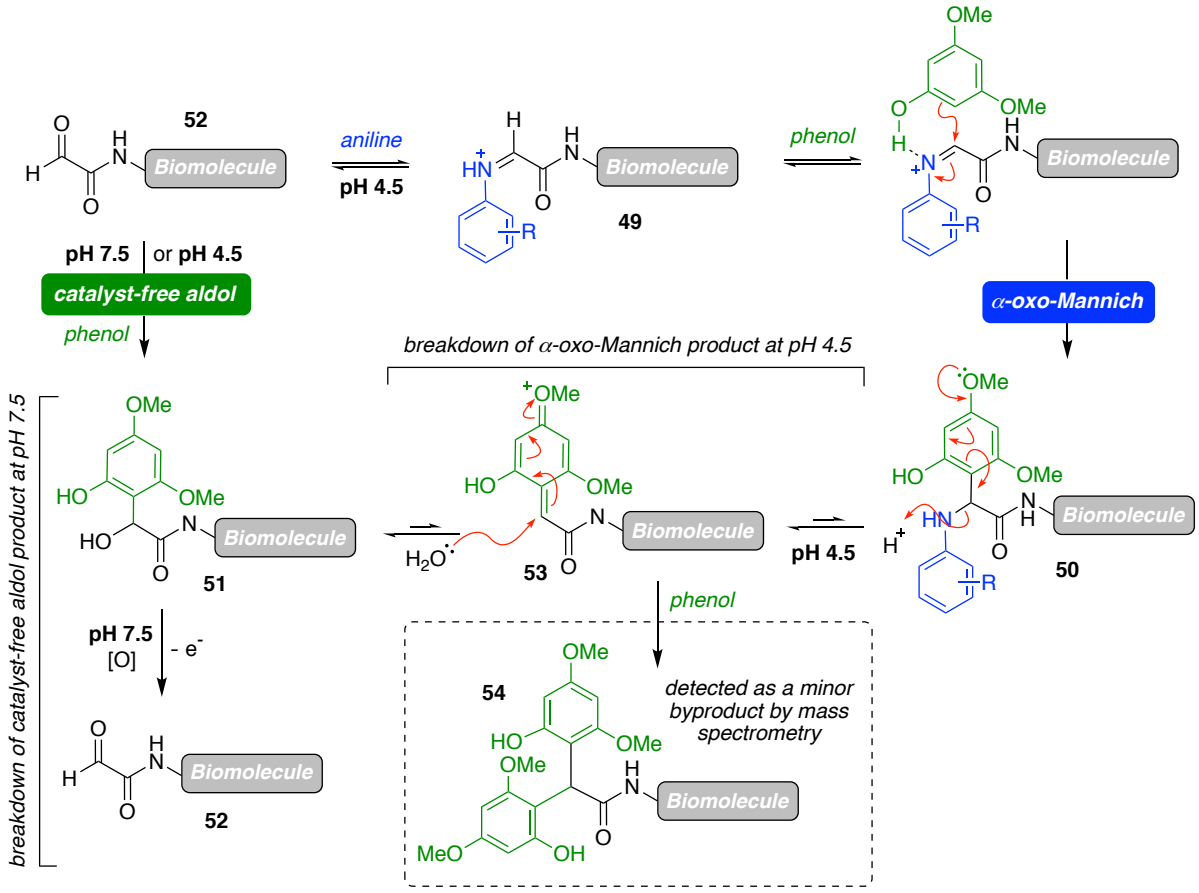


Fig. 10 Proposed mechanistic rationale for formation and breakdown of both the α-oxo-Mannich product **50** and the catalyst-free aldol product **51**.

CONCLUSIONS

In conclusion we have developed two novel bioconjugation reactions using an α-oxo aldehyde handle that can be site-selectively installed into biomolecules. In the presence of both phenol and aniline nucleophiles the α-oxo-Mannich and catalyst-free aldol bioconjugations compete with each other for this unique electrophile, but through judicious choice of pH, substrates, and optimisation of reaction conditions formation of the multicomponent α-oxo-Mannich products can be favoured. We also demonstrated that the Mannich products breakdown to afford the catalyst-free aldol product at both neutral and acidic pH. However, the rate of this breakdown reaction is likely governed by both the electron donating ability of the phenol and the basicity of the aniline, potentially providing a tuneable scaffold for the controllable release of small molecule cargo from a protein delivery system in future studies.

In the absence of aniline, phenols react cleanly and rapidly with α-oxo aldehyde containing proteins at pH 7.5, akin to reactivity previously observed in the sequestration of AGE precursors glyoxal and methyl glyoxal by green/black tea polyphenols. Intriguingly however, when our catalyst-free aldol bioconjugate was incubated in the absence of the electron rich phenol **4**, we observed breakdown to the α-oxo aldehyde starting material at pH 7.5, but none at pH 4.5. This process may be a result of one-electron oxidation of the phenol and highlights a need to revisit the study of tea polyphenol sequestrations of reactive α-oxo aldehydes *in vitro* and *in vivo*, particularly the stability of the conjugates in the absence of polyphenol. Additionally, the instability observed for both the aldol and α-oxo-Mannich bioconjugates may limit the *in vivo* utility of the conjugates due to potential premature breakdown in the bloodstream at physiological pH, which in the case of antibody-drug conjugates can lead to toxicity.[85] However, it is likely the nature of the phenol component within the catalyst-free aldol bioconjugate could have a direct effect on its stability, suggesting that like the α-oxo-Mannich, if pH stability could be increased then the catalyst-free aldol bioconjugation may afford a tuneable scaffold which could potentially be leveraged in the future for oxidative release of small molecules from proteins. For example cancer cells have higher levels of reactive oxygen species due to increased metabolic activity and a number other oncogenic processes,[86] and thus could conceivably provide an ideal environment for selective release of phloridzin, or other SLGT inhibitors from the prostate cancer targeting nanobody bioconjugate **48** constructed here.

METHODS

General procedures and materials

All solvents were dried prior to use according to standard methods, with the exception of solvents used for flash chromatography purposes, where GPR-grade solvents were used. All commercially available reagents were used as received. Analytical grade reagents were supplied by Sigma-Aldrich, Fisher Scientific, VWR International, Carbosynth, Alfa aesar and TCI. All solution-phase reactions were carried out under a dry nitrogen atmosphere using oven-dried glassware unless otherwise stated. All concentrations were performed in vacuo unless otherwise stated. Thin layer chromatography was carried out on Merck silica gel 60 F254 pre-coated aluminium foil sheets and these were visualised using UV light (254 nm) or charred following immersion in 5% sulphuric acid in methanol. Hydrophilic acylated surface protein A (HASPA) G1S mutant was prepared exactly as previously described.[37] sfGFP(N150Thz) was prepared exactly as previously described.[35] Myoglobin from equine heart (M1882) was purchased from Sigma-Aldrich and used without further purification.

NMR

1H and 13C NMR spectra were measured at 500 MHz and 126 MHz respectively using a Bruker 500-MR spectrometer at the University York Centre for Magnetic Resonance, using Me4Si as an internal standard when using chloroform d. Multiplicities are given as singlet (s), doublet (d) doublet of doublets (dd), doublet of doublet of doublets (ddd) or multiplet (m). Resonances were assigned using HH-COSY and CH-HSQC. All NMR chemical shifts (δ) were recorded in pp and coupling constants (J) are reported in Hz. Topspin 4.0.6 was primarily used for processing the spectral data.

FTIR and optical rotations

Fourier transform infrared (FTIR) spectra were recorded on a PerkinElmer UATR 2 spectrometer using the attenuated total reflectance (ATR) technique. (ESI) Optical rotations were measured using a Bellingham and Stanley ADP 450 Automatic Digital Peltier Controlled Polarimeter equipped with a 589 nm LED. Concentration is denoted as c and was calculated as grams per 100 millilitres (g / 100 mL) whereas the solvent is indicated in parenthesis (c, solvent).

Mass spectrometry

Small-molecule high resolution mass spectrometry (HRMS) data were obtained at room temperature on a Bruker Daltonics microTOF mass spectrometer coupled to an Agilent 1200 series LC system.

High Performance Liquid Chromatography-Electrospray Ionisation Mass Spectrometry (LC-MS) was accomplished using a Dionex UltiMate® 3000 LC system (ThermoScientific) equipped with an UltiMate® 3000 Diode Array Detector (probing 250-400 nm) in line with a Bruker HCTultra ETD II system (Bruker Daltonics), using Chromeleon® 6.80 SR12 software (ThermoScientific), Compass 1.3 for esquire HCT Build 581.3, esquireControl version 6.2, Build 62.24 software (Bruker Daltonics), and Bruker compass HyStar 3.2-SR2, HyStar version 3.2, Build 44 software (Bruker Daltonics) at The University York Centre of Excellence in Mass Spectrometry (CoEMS). All mass spectrometry was conducted in positive ion mode unless stated otherwise. Data analysis was performed using ESI Compass 1.3 DataAnalysis, Version 4.1 software (Bruker Daltonics). Prior to analysis by LC-MS, peptide or protein ligation mixture was diluted 1:3 in water and then further diluted 1:1 in acetonitrile with 1 % (v/v) formic acid. Peptide samples were analysed using an Accucore™ C18 2.6 µm column (50 x 2.1 mm) (ThermoScientific). Water with 0.1 % (v/v) formic acid (solvent A) and acetonitrile with 0.1 % (v/v) formic acid (solvent B) were used as the mobile phase at a flow rate of 0.3 mL/min at room temperature (RT). A multi-step gradient of 6.5 min was programmed as follows: 90% A for 0.5 min, followed by a linear gradient to 95% B over 3.5 min, followed by 95% B for an additional 0.5 min. A linear gradient to 95% A was used to re-equilibrate the column Under these conditions all peptides typically eluted between 2-5 min. Protein samples were analysed without the use of a column at RT.

Water with 0.1 % (v/v) formic acid (solvent A) and acetonitrile with 0.1 % (v/v) formic acid (solvent B) were used as the mobile phase at a 1:1 ratio over the course of 3 min as follows: 0.05 mL/min to 0.25 mL/min for 1 min, 0.025 mL/min for 1 min, followed by 1.0 mL/min for 1 min. Under these conditions, all proteins typically eluted between 0.1-1.5 min.

Protein electrospray ionisation (ESI) mass spectra were obtained on a Bruker Solarix XR 9.4 T FTICR mass spectrometer. Samples were desalted and analysed at a final concentration of 0.3-10 µM in 50:50:1 (v/v) H2O:MeCN:FA. Mass spectrometry data analysis was performed using ESI Compass 1.3 DataAnalysis, Version 4.4 software (Bruker Daltonics).

Determination of bioconjugation relative conversion

For peptide bioconjugations, conversion from starting material to the desired product (relative conversion, %) was calculated by analysing the peak intensities of starting material and product species obtained after LC-MS analysis. For protein bioconjugations mass spectra were first deconvoluted, then conversion from starting material to the desired product was calculated by analysing the peak intensities of the starting material and product species.

Solid Phase Peptide Synthesis (SPPS)

Peptides were synthesised via manual solid phase peptide synthesis (SPPS) using an in-situ neutralisation/HCTU activation procedure for Fmoc chemistry on an H-Gly-2- ClTrt resin (Sigma) using Fmoc protected amino acids as described below:

Preloaded resin preparation. The preloaded 2-chlorotrityl resin was weighed out into a 2 mL SPPS cartridge fitted with a PTFE stopcock, swollen in DMF for 30 min and then filtered.

Amino acid coupling. DIPEA (11.0 eq.) was added to a solution of amino acid (5.0 eq.) and HCTU (5.0 eq.) dissolved in the minimum volume of DMF and the solution added to the resin. The reaction mixture was gently agitated by rotation for 1 h, and the resin filtered off and washed with DMF (3 × 2 min with rotation).

Fmoc deprotection. A solution of 20% piperidine in DMF was added to the resin and gently agitated by rotation for 2 minutes. The resin was filtered off and repeated four more times, followed by washes with DMF (5 × 2 min with rotation).

Cleavage and Isolation. Resins containing full synthesised peptides were washed with DCM (3 × 2 min with rotation) and MeOH (3 × 2 min with rotation). The resin was dried on a vacuum manifold and further dried on a high vacuum line overnight. A solution of cleavage cocktail 95:2.5:2.5 (v/v) TFA:H2O:triisopropylsilane was then added to the resin, and the resulting mixture was gently agitated by rotation for 60 min. The reaction mixture was drained into ice-cold Et2O and centrifuged at 6000

rpm at 4 °C until pelleted (ca. 5-10 min). The supernatant was carefully decanted and subsequently resuspended, centrifuged and supernatant decanted three more times. The precipitated peptide pellet was then either dissolved 10% MeCN or in 10% aq. AcOH and lyophilised. The desired peptide was then further purified via size-exclusion chromatography (Sephadex LH-20 in water), and fractions containing pure, desired peptide were lyophilised and stored at -20 °C until required.

SUPPORTING INFORMATION

Supporting information figures, bioconjugation experimental protocols and characterization, protein purification and characterization, small molecule synthesis and characterization, and lectin blots.

AUTHOR INFORMATION

Corresponding Author Email: martin.fascione@york.ac.uk

ACKNOWLEDGEMENTS

We thank E. Bergstrom and The York Centre of Excellence in Mass Spectrometry. The York Centre of Excellence in Mass Spectrometry was created thanks to a major capital investment through Science City York, supported by Yorkshire Forward with funds from the Northern Way Initiative, and subsequent support from EPSRC (EP/K039660/1; EP/M028127/1). This work was supported by The University of York, the EPSRC (PhD award [1944852] to N. Hatton) and EP/P030653/1, EP/S013741/1, EP/V044303/1) and the BBSRC (BB/M02847X/1). C. Mahon is a grateful recipient of a Marie Sklodowska-Curie Global Fellowship from the European Union's Horizon 2020 Research and Innovation Programme under the Marie Sklodowska-Curie grant agreement No. 702927. pBAD-sfGFP 150TAG was a gift from R. Mehl (Addgene plasmid # 85483) under MTA, pEVOL PylRS WT was a kind gift from E. Lemke (EMBL Heidelberg) under MTA.

REFERENCES

1. Boutureira, O.; Bernardes, G. J. L., Advances in Chemical Protein Modification. *Chem. Rev.* **2015,** *115* (5), 2174-2195.
2. Spicer, C. D.; Davis, B. G., Selective chemical protein modification. *Nat. Commun.* **2014,** *5* (1), 4740.

3. Gupta, V.; Bhavanasi, S.; Quadir, M.; Singh, K.; Ghosh, G.; Vasamreddy, K.; Ghosh, A.; Siahaan, T. J.; Banerjee, S.; Banerjee, S. K., Protein PEGylation for cancer therapy: bench to bedside. *J. Cell Commun. Signal.* **2019,** *13* (3), 319-330.
4. Fee, C. J., Size comparison between proteins PEGylated with branched and linear poly(ethylene glycol) molecules. *Biotechnol. Bioeng.* **2007,** *98* (4), 725-731.
5. Agarwal, P.; Beahm, B. J.; Shieh, P.; Bertozzi, C. R., Systemic Fluorescence Imaging of Zebrafish Glycans with Bioorthogonal Chemistry. *Angew. Chem. Int. Ed.* **2015,** *54* (39), 11504-11510.
6. Baskin, J. M.; Dehnert, K. W.; Laughlin, S. T.; Amacher, S. L.; Bertozzi, C. R., Visualizing enveloping layer glycans during zebrafish early embryogenesis. *Proc. Nat. Acad. Sci. U.S.A.* **2010,** *107* (23), 10360.
7. Baskin, J. M.; Prescher, J. A.; Laughlin, S. T.; Agard, N. J.; Chang, P. V.; Miller, I. A.; Lo, A.; Codelli, J. A.; Bertozzi, C. R., Copper-free click chemistry for dynamic in vivo imaging. *Proc. Nat. Acad. Sci. U.S.A.* **2007,** *104* (43), 16793.
8. Saxon, E.; Bertozzi, C. R., Cell Surface Engineering by a Modified Staudinger Reaction. *Science* **2000,** *287* (5460), 2007.
9. Trads, J. B.; Tørring, T.; Gothelf, K. V., Site-Selective Conjugation of Native Proteins with DNA. *Acc. Chem. Res.* **2017,** *50* (6), 1367-1374.
10. Bilyard, M. K.; Bailey, H. J.; Raich, L.; Gafitescu, M. A.; Machida, T.; Iglésias-Fernández, J.; Lee, S. S.; Spicer, C. D.; Rovira, C.; Yue, W. W. *et al.,* Palladium-mediated enzyme activation suggests multiphase initiation of glycogenesis. *Nature* **2018,** *563* (7730), 235-240.
11. Raj, R.; Lercher, L.; Mohammed, S.; Davis, B. G., Synthetic Nucleosomes Reveal that GlcNAcylation Modulates Direct Interaction with the FACT Complex. *Angew. Chem. Int. Ed.* **2016,** *55* (31), 8918-8922.
12. Parsons, T. B.; Struwe, W. B.; Gault, J.; Yamamoto, K.; Taylor, T. A.; Raj, R.; Wals, K.; Mohammed, S.; Robinson, C. V.; Benesch, J. L. P. *et al.,* Optimal Synthetic Glycosylation of a Therapeutic Antibody. *Angew. Chem. Int. Ed.* **2016,** *55* (7), 2361-2367.
13. van Kasteren, S. I.; Kramer, H. B.; Jensen, H. H.; Campbell, S. J.; Kirkpatrick, J.; Oldham, N. J.; Anthony, D. C.; Davis, B. G., Expanding the diversity of chemical protein modification allows post-translational mimicry. *Nature* **2007,** *446* (7139), 1105-1109.
14. Macmillan, D.; Bill, R. M.; Sage, K. A.; Fern, D.; Flitsch, S. L., Selective in vitro glycosylation of recombinant proteins: semi-synthesis of novel homogeneous glycoforms of human erythropoietin. *Chem. Biol.* **2001,** *8* (2), 133-145.
15. Schlatzer, T.; Kriegesmann, J.; Schröder, H.; Trobe, M.; Lembacher-Fadum, C.; Santner, S.; Kravchuk, A. V.; Becker, C. F. W.; Breinbauer, R., Labeling and Natural Post-Translational Modification of Peptides and Proteins via Chemoselective Pd-Catalyzed Prenylation of Cysteine. *J. Am. Chem. Soc.* **2019,** *141* (37), 14931-14937.
16. Weng, Y.; Song, C.; Chiang, C.-W.; Lei, A., Single electron transfer-based peptide/protein bioconjugations driven by biocompatible energy input. *Commun. Chem.* **2020,** *3* (1), 171.
17. Josephson, B.; Fehl, C.; Isenegger, P. G.; Nadal, S.; Wright, T. H.; Poh, A. W. J.; Bower, B. J.; Giltrap, A. M.; Chen, L.; Batchelor-McAuley, C. *et al.*, Light-driven post-translational installation of reactive protein side chains. *Nature* **2020,** *585* (7826), 530-537.
18. Lobba, M. J.; Fellmann, C.; Marmelstein, A. M.; Maza, J. C.; Kissman, E. N.; Robinson, S. A.; Staahl, B. T.; Urnes, C.; Lew, R. J.; Mogilevsky, C. S. *et al.,* Site-Specific Bioconjugation through Enzyme-Catalyzed Tyrosine–Cysteine Bond Formation. *ACS Cent. Sci.* **2020,** *6* (9), 1564-1571.
19. Kim, J.; Li, B. X.; Huang, R. Y. C.; Qiao, J. X.; Ewing, W. R.; MacMillan, D. W. C., Site-Selective Functionalization of Methionine Residues via Photoredox Catalysis. *J. Am. Chem. Soc.* **2020,** *142* (51), 21260-21266.
20. Dhanjee, H. H.; Saebi, A.; Buslov, I.; Loftis, A. R.; Buchwald, S. L.; Pentelute, B. L., Protein–Protein Cross-Coupling via Palladium–Protein Oxidative Addition Complexes from Cysteine Residues. *J. Am. Chem. Soc.* **2020,** *142* (20), 9124-9129.

21. Zhang, C.; Vinogradova, E. V.; Spokoyny, A. M.; Buchwald, S. L.; Pentelute, B. L., Arylation Chemistry for Bioconjugation. *Angew. Chem. Int. Ed.* **2019,** *58* (15), 4810-4839.
22. Taylor, M. T.; Nelson, J. E.; Suero, M. G.; Gaunt, M. J., A protein functionalization platform based on selective reactions at methionine residues. *Nature* **2018,** *562* (7728), 563-568.
23. Bloom, S.; Liu, C.; Kölmel, D. K.; Qiao, J. X.; Zhang, Y.; Poss, M. A.; Ewing, W. R.; MacMillan, D. W. C., Decarboxylative alkylation for site-selective bioconjugation of native proteins via oxidation potentials. *Nat. Chem.* **2018,** *10* (2), 205-211.
24. Lin, S.; Yang, X.; Jia, S.; Weeks, A. M.; Hornsby, M.; Lee, P. S.; Nichiporuk, R. V.; Iavarone, A. T.; Wells, J. A.; Toste, F. D. *et al.,* Redox-based reagents for chemoselective methionine bioconjugation. *Science* **2017,** *355* (6325), 597.
25. Zhang, C.; Welborn, M.; Zhu, T.; Yang, N. J.; Santos, M. S.; Van Voorhis, T.; Pentelute, B. L., π-Clamp-mediated cysteine conjugation. *Nat. Chem.* **2016,** *8* (2), 120-128.
26. Rosen, C. B.; Francis, M. B., Targeting the N terminus for site-selective protein modification. *Nat. Chem. Biol.* **2017,** *13* (7), 697-705.
27. Dadová, J.; Galan, S. R. G.; Davis, B. G., Synthesis of modified proteins via functionalization of dehydroalanine. *Curr. Opin. Chem. Biol.* **2018,** *46*, 71-81.
28. Bhushan, B.; Lin, Y. A.; Bak, M.; Phanumartwiwath, A.; Yang, N.; Bilyard, M. K.; Tanaka, T.; Hudson, K. L.; Lercher, L.; Stegmann, M. *et al.,* Genetic Incorporation of Olefin Cross-Metathesis Reaction Tags for Protein Modification. *J. Am. Chem. Soc.* **2018,** *140* (44), 14599-14603.
29. McKay, Craig S.; Finn, M. G., Click Chemistry in Complex Mixtures: Bioorthogonal Bioconjugation. *Chem. Biol.* **2014,** *21* (9), 1075-1101.
30. Blackman, M. L.; Royzen, M.; Fox, J. M., Tetrazine Ligation: Fast Bioconjugation Based on Inverse-Electron-Demand Diels−Alder Reactivity. *J. Am. Chem. Soc.* **2008,** *130* (41), 13518-13519.
31. Spears, R. J.; Fascione, M. A., Site-selective incorporation and ligation of protein aldehydes. *Org. Biomol. Chem.* **2016,** *14* (32), 7622-7638.
32. El-Mahdi, O.; Melnyk, O., α-Oxo Aldehyde or Glyoxylyl Group Chemistry in Peptide Bioconjugation. *Bioconjug. Chem.* **2013,** *24* (5), 735-765.
33. Bi, X.; Pasunooti, K. K.; Lescar, J.; Liu, C.-F., Thiazolidine-Masked α-Oxo Aldehyde Functionality for Peptide and Protein Modification. *Bioconjug.Chem.* **2017,** *28* (2), 325-329.
34. Brabham, R. L.; Keenan, T.; Husken, A.; Bilsborrow, J.; McBerney, R.; Kumar, V.; Turnbull, W. B.; Fascione, M. A., Rapid sodium periodate cleavage of an unnatural amino acid enables unmasking of a highly reactive α-oxo aldehyde for protein bioconjugation. *Org. Biomol. Chem.* **2020,** *18* (21), 4000-4003.
35. Brabham, R. L.; Spears, R. J.; Walton, J.; Tyagi, S.; Lemke, E. A.; Fascione, M. A., Palladium-unleashed proteins: gentle aldehyde decaging for site-selective protein modification. *Chem. Commun.* **2018,** *54* (12), 1501-1504.
36. Mahesh, S.; Adebomi, V.; Muneeswaran, Z. P.; Raj, M., Bioinspired Nitroalkylation for Selective Protein Modification and Peptide Stapling. *Angew. Chem. Int. Ed.* **2020,** *59* (7), 2793-2801.
37. Spears, R. J.; Brabham, R. L.; Budhadev, D.; Keenan, T.; McKenna, S.; Walton, J.; Brannigan, J. A.; Brzozowski, A. M.; Wilkinson, A. J.; Plevin, M. *et al.,* Site-selective C–C modification of proteins at neutral pH using organocatalyst-mediated cross aldol ligations. *Chem. Sci.* **2018,** *9* (25), 5585-5593.
38. Zhang, L.; Tam, J. P., Thiazolidine Formation as a General and Site-Specific Conjugation Method for Synthetic Peptides and Proteins. *Analytical Biochemistry* **1996,** *233* (1), 87-93.
39. Kölmel, D. K.; Kool, E. T., Oximes and Hydrazones in Bioconjugation: Mechanism and Catalysis. *Chem. Rev.* **2017,** *117* (15), 10358-10376.
40. Larsen, D.; Pittelkow, M.; Karmakar, S.; Kool, E. T., New Organocatalyst Scaffolds with High Activity in Promoting Hydrazone and Oxime Formation at Neutral pH. *Org. Lett.* **2015,** *17* (2), 274-277.
41. Larsen, D.; Kietrys, A. M.; Clark, S. A.; Park, H. S.; Ekebergh, A.; Kool, E. T., Exceptionally rapid oxime and hydrazone formation promoted by catalytic amine buffers with low toxicity. *Chem. Sci.* **2018,** *9* (23), 5252-5259.

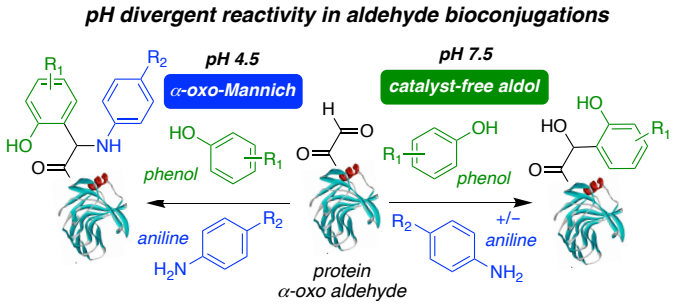
42. Trausel, F.; Maity, C.; Poolman, J. M.; Kouwenberg, D. S. J.; Versluis, F.; van Esch, J. H.; Eelkema, R., Chemical signal activation of an organocatalyst enables control over soft material formation. *Nat. Commun.* **2017,** *8* (1), 879.
43. Zhou, Y.; Piergentili, I.; Hong, J.; van der Helm, M. P.; Macchione, M.; Li, Y.; Eelkema, R.; Luo, S., Indoline Catalyzed Acylhydrazone/Oxime Condensation under Neutral Aqueous Conditions. *Org. Lett.* **2020,** *22* (15), 6035-6040.
44. Trausel, F.; Fan, B.; van Rossum, S. A. P.; van Esch, J. H.; Eelkema, R., Aniline Catalysed Hydrazone Formation Reactions Show a Large Variation in Reaction Rates and Catalytic Effects. *Adv. Synth. Cat.* **2018,** *360* (13), 2571-2576.
45. Drienovská, I.; Mayer, C.; Dulson, C.; Roelfes, G., A designer enzyme for hydrazone and oxime formation featuring an unnatural catalytic aniline residue. *Nat. Chem.* **2018,** *10* (9), 946-952.
46. Dirksen, A.; Hackeng, T. M.; Dawson, P. E., Nucleophilic Catalysis of Oxime Ligation. *Angew. Chem. Int. Ed.* **2006,** *45* (45), 7581-7584.
47. Joshi, N. S.; Whitaker, L. R.; Francis, M. B., A Three-Component Mannich-Type Reaction for Selective Tyrosine Bioconjugation. *J. Am. Chem. Soc.* **2004,** *126* (49), 15942-15943.
48. Joshi, N. S.; Whitaker, L. R.; Francis, M. B., A three-component Mannich-type reaction for selective tyrosine bioconjugation. *Journal of the American Chemical Society* **2004,** *126* (49), 15942-15943.
49. McFarland, J. M.; Joshi, N. S.; Francis, M. B., Characterization of a Three-Component Coupling Reaction on Proteins by Isotopic Labeling and Nuclear Magnetic Resonance Spectroscopy. *J. Am. Chem. Soc.* **2008,** *130* (24), 7639-7644.
50. Saimoto, H.; Yoshida, K.; Murakami, T.; Morimoto, M.; Sashiwa, H.; Shigemasa, Y., Effect of Calcium Reagents on Aldol Reactions of Phenolic Enolates with Aldehydes in Alcohol. *J. Org. Chem.* **1996,** *61* (20), 6768-6769.
51. Cohen, D. T.; Zhang, C.; Fadzen, C. M.; Mijalis, A. J.; Hie, L.; Johnson, K. D.; Shriver, Z.; Plante, O.; Miller, S. J.; Buchwald, S. L. *et al.,* A chemoselective strategy for late-stage functionalization of complex small molecules with polypeptides and proteins. *Nat. Chem.* **2019,** *11* (1), 78-85.
52. Szent-Györgyi, A.; Együd, L. G.; McLaughlin, J. A., Keto-Aldehydes and Cell Division. *Science* **1967,** *155* (3762), 539.
53. Kalapos, M. P., Methylglyoxal in living organisms: Chemistry, biochemistry, toxicology and biological implications. *Toxicol. Lett.* **1999,** *110* (3), 145-175.
54. Wang, Y.; Ho, C.-T., Flavour chemistry of methylglyoxal and glyoxal. *Chem. Soc. Rev.* **2012,** *41* (11), 4140-4149.
55. Goldin, A.; Beckman Joshua, A.; Schmidt Ann, M.; Creager Mark, A., Advanced Glycation End Products. *Circulation* **2006,** *114* (6), 597-605.
56. Rhee, S. Y.; Kim, Y. S., The Role of Advanced Glycation End Products in Diabetic Vascular Complications. *Diabetes Metab. J.* **2018,** *42* (3), 188-195.
57. Chaudhuri, J.; Bains, Y.; Guha, S.; Kahn, A.; Hall, D.; Bose, N.; Gugliucci, A.; Kapahi, P., The Role of Advanced Glycation End Products in Aging and Metabolic Diseases: Bridging Association and Causality. *Cell Metab.* **2018,** *28* (3), 337-352.
58. Lambert, J. D.; Sang, S.; Yang, C. S., Biotransformation of green tea polyphenols and the biological activities of those metabolites. *Mol. Pharm.* **2007,** *4* (6), 819-25.
59. Li, X.; Zheng, T.; Sang, S.; Lv, L., Quercetin inhibits advanced glycation end product formation by trapping methylglyoxal and glyoxal. *J. Agric. Food Chem.* **2014,** *62* (50), 12152-8.
60. Lv, L.; Shao, X.; Chen, H.; Ho, C. T.; Sang, S., Genistein inhibits advanced glycation end product formation by trapping methylglyoxal. *Chem. Res. Toxicol.* **2011,** *24* (4), 579-86.
61. Lo, C.-Y.; Hsiao, W.-T.; Chen, X.-Y., Efficiency of Trapping Methylglyoxal by Phenols and Phenolic Acids. *J. Food Sci.* **2011,** *76* (3), H90-H96.
62. Lo, C.-Y.; Li, S.; Tan, D.; Pan, M.-H.; Sang, S.; Ho, C.-T., Trapping reactions of reactive carbonyl species with tea polyphenols in simulated physiological conditions. *Mol. Nutr. Food Res.* **2006,** *50* (12), 1118-1128.

63. Minakawa, M.; Guo, H.-M.; Tanaka, F., Imines that React with Phenols in Water over a Wide pH Range. *The Journal of Organic Chemistry* **2008,** *73* (21), 8669-8672.
64. Casiraghi, G.; Salerno, G.; Sartori, G., A Novel One-Step Synthesis of 2-Hydroxyphenacyl Alcohols. *Synthesis* **1975,** *1975* (03), 186-187.
65. Cordes, E. H.; Jencks, W. P., Nucleophilic Catalysis of Semicarbazone Formation by Anilines. *J. Am. Chem. Soc.* **1962,** *84* (5), 826-831.
66. List, B.; Pojarliev, P.; Biller, W. T.; Martin, H. J., The Proline-Catalyzed Direct Asymmetric Three-Component Mannich Reaction: Scope, Optimization, and Application to the Highly Enantioselective Synthesis of 1,2-Amino Alcohols. *J. Am. Chem. Soc.* **2002,** *124* (5), 827-833.
67. Gilmore, J. M.; Scheck, R. A.; Esser-Kahn, A. P.; Joshi, N. S.; Francis, M. B., N-Terminal Protein Modification through a Biomimetic Transamination Reaction. *Angew. Chem. Int. Ed.* **2006,** *45* (32), 5307-5311.
68. Chatalic, K. L. S.; Veldhoven-Zweistra, J.; Bolkestein, M.; Hoeben, S.; Koning, G. A.; Boerman, O. C.; de Jong, M.; van Weerden, W. M., A novel 111In-labeled anti-PSMA nanobody for targeted SPECT/CT imaging of prostate cancer. *J. Nuc. Med.* **2015**.
69. Hatton, N. E.; Baumann, C. G.; Fascione, M. A., Developments in mannose-based treatments for Uropathogenic Escherichia coli induced urinary tract infections. *ChemBioChem* **2021,** *22* (4), 613-629.
70. Haupts, U.; Maiti, S.; Schwille, P.; Webb, W. W., Dynamics of fluorescence fluctuations in green fluorescent protein observed by fluorescence correlation spectroscopy. *Proc. Nat. Acad. Sci.* **1998,** *95* (23), 13573.
71. Peumans, W. J.; Damme, E. J. M. V., Plant Lectins: Versatile Proteins with Important Perspectives in Biotechnology. *Biotechnol. Genet. Eng. Rev.* **1998,** *15* (1), 199-228.
72. Osman, M.; Mistry, A.; Keding, A.; Gabe, R.; Cook, E.; Forrester, S.; Wiggins, R.; Di Marco, S.; Colloca, S.; Siani, L. *et al.,* A third generation vaccine for human visceral leishmaniasis and post kala azar dermal leishmaniasis: First-in-human trial of ChAd63-KH. *PLOS Neglected Tropical Diseases* **2017,** *11* (5), e0005527.
73. Senko, M. W.; Hendrickson, C. L.; Paša-Tolić, L.; Marto, J. A.; White, F. M.; Guan, S.; Marshall, A. G., Electrospray Ionization Fourier Transform Ion Cyclotron Resonance at 9.4 T. *Rapid Communications in Mass Spectrometry* **1996,** *10* (14), 1824-1828.
74. McKean, P. G.; Delahay, R.; Pimenta, P. F. P.; Smith, D. F., Characterisation of a second protein encoded by the differentially regulated LmcDNA16 gene family of Leishmania major. *Mol. Biochem. Parasitol.* **1997,** *85* (2), 221-231.
75. Scafoglio, C.; Hirayama, B. A.; Kepe, V.; Liu, J.; Ghezzi, C.; Satyamurthy, N.; Moatamed, N. A.; Huang, J.; Koepsell, H.; Barrio, J. R. *et al.,* Functional expression of sodium-glucose transporters in cancer. *Proc. Nat. Acad. Sci. U.S.A.* **2015,** *112* (30), E4111.
76. Gillies, R. J.; Robey, I.; Gatenby, R. A., Causes and Consequences of Increased Glucose Metabolism of Cancers. *J. Nuc. Med.* **2008,** *49* (Suppl 2), 24S.
77. Villani, L. A.; Smith, B. K.; Marcinko, K.; Ford, R. J.; Broadfield, L. A.; Green, A. E.; Houde, V. P.; Muti, P.; Tsakiridis, T.; Steinberg, G. R., The diabetes medication Canagliflozin reduces cancer cell proliferation by inhibiting mitochondrial complex-I supported respiration. *Mol. Metab.* **2016,** *5* (10), 1048-1056.
78. White, J. R., Apple Trees to Sodium Glucose Co-Transporter Inhibitors: A Review of SGLT2 Inhibition. *Clin. Diabetes* **2010,** *28* (1), 5.
79. Crespy, V.; Aprikian, O.; Morand, C.; Besson, C.; Manach, C.; Demigné, C.; Rémésy, C., Bioavailability of Phloretin and Phloridzin in Rats. *J. Nutr.* **2001,** *131* (12), 3227-3230.
80. Rhile, I. J.; Markle, T. F.; Nagao, H.; DiPasquale, A. G.; Lam, O. P.; Lockwood, M. A.; Rotter, K.; Mayer, J. M., Concerted Proton−Electron Transfer in the Oxidation of Hydrogen-Bonded Phenols. *J. Am. Chem. Soc.* **2006,** *128* (18), 6075-6088.

687 81. Hammarström, L.; Styring, S., Proton-coupled electron transfer of tyrosines in Photosystem II
688 and model systems for artificial photosynthesis: the role of a redox-active link between catalyst and
689 photosensitizer. *Energy Environ. Sci.* **2011,** *4* (7), 2379-2388.
690 82. Huynh, M. T.; Mora, S. J.; Villalba, M.; Tejeda-Ferrari, M. E.; Liddell, P. A.; Cherry, B. R.; Teillout,
691 A.-L.; Machan, C. W.; Kubiak, C. P.; Gust, D. *et al.,* Concerted One-Electron Two-Proton Transfer
692 Processes in Models Inspired by the Tyr-His Couple of Photosystem II. *ACS Cent. Sci.* **2017,** *3* (5), 372-
693 380.
694 83. Costentin, C.; Robert, M.; Savéant, J.-M.; Tard, C., Inserting a Hydrogen-Bond Relay between
695 Proton Exchanging Sites in Proton-Coupled Electron Transfers. *Angew. Chem. Int. Ed.* **2010,** *49* (22),
696 3803-3806.
697 84. Niki, E., Action of ascorbic acid as a scavenger of active and stable oxygen radicals. *Am. J. Clin.*
698 *Nutr.* **1991,** *54* (6 Suppl), 1119s-1124s.
699 85. Ducry, L.; Stump, B., Antibody–Drug Conjugates: Linking Cytotoxic Payloads to Monoclonal
700 Antibodies. *Bioconjugate Chemistry* **2010,** *21* (1), 5-13.
701 86. Storz, P., Reactive oxygen species in tumor progression. *Front. Biosci.* **2005,** *10* (2), 1881-1896.

702

703 **SYNOPSIS TOC**



704