

# Chemoselective umpolung of thiols to episulfoniums for cysteine bioconjugation

Received: 12 January 2023

Accepted: 27 October 2023

Published online: 20 December 2023

Check for updates

Philipp Hartmann<sup>1,2</sup>, Kostiantyn Bohdan<sup>1,2</sup>, Moritz Hommrich<sup>1,2</sup>, Fabio Juliá<sup>1</sup>, Lara Vogelsang<sup>3</sup>, Jürgen Eirich<sup>4</sup>, Rene Zangl<sup>5</sup>, Christophe Farès<sup>1</sup>, Julia Beatrice Jacobs<sup>1</sup>, Dwaipayan Mukhopadhyay<sup>6</sup>, Johanna Marie Mengeler<sup>1</sup>, Alessandro Vetere<sup>1</sup>, Marie Sophie Sterling<sup>1</sup>, Heike Hinrichs<sup>1</sup>, Stefan Becker<sup>6</sup>, Nina Morgner<sup>5</sup>, Wolfgang Schrader<sup>1</sup>, Iris Finkemeier<sup>4</sup>, Karl-Josef Dietz<sup>3</sup>, Christian Griesinger<sup>6</sup> & Tobias Ritter<sup>1</sup>✉

Cysteine conjugation is an important tool in protein research and relies on fast, mild and chemoselective reactions. Cysteiny l thiols can either be modified with prefunctionalized electrophiles, or converted into electrophiles themselves for functionalization with selected nucleophiles in an independent step. Here we report a bioconjugation strategy that uses a vinyl thianthrenium salt to transform cysteine into a highly reactive electrophilic episulfonium intermediate in situ, to enable conjugation with a diverse set of bioorthogonal nucleophiles in a single step. The reactivity profile can connect several nucleophiles to biomolecules through a short and stable ethylene linker, ideal for introduction of infrared labels, post-translational modifications or NMR probes. In the absence of reactive exogenous nucleophiles, nucleophilic amino acids can react with the episulfonium intermediate for native peptide stapling and protein–protein ligation. Ready synthetic access to isotopologues of vinyl thianthrenium salts enables applications in quantitative proteomics. Such diverse applications demonstrate the utility of vinyl-thianthrenium-based bioconjugation as a fast, selective and broadly applicable tool for chemical biology.

Efficient functionalization of proteins is challenging because it requires fast and chemoselective reactions that produce stable conjugates under biocompatible conditions. Introduction of a functional group at a specific position in the protein is achieved with methods that modify amino acid residues with low natural abundance<sup>1</sup> such as cysteine (Cys)<sup>2</sup>, methionine<sup>3,4</sup>, tryptophan<sup>5,6</sup> and tyrosine<sup>7</sup>. Due to its high intrinsic reactivity and easy incorporation via site-directed mutagenesis, Cys is often the target of choice for site-selective bioconjugations under physiological conditions<sup>8</sup>. A common approach exploits the reactivity between Cys and electrophiles, such as Michael acceptors<sup>9–11</sup>,

$\alpha$ -halocarbonyls<sup>12</sup>, organometallic reagents<sup>13</sup>, disulfides<sup>14</sup>, hypervalent iodine reagents<sup>15,16</sup>, phosphoramidates<sup>17</sup> or perfluoroarenes<sup>18</sup>, that carry a payload of interest. Many methods allow for fast and efficient Cys functionalization. However, each functionalization to introduce a specific payload requires synthesis of a new derivative of the reagent if not commercially available.

An alternative strategy entails Cys transformation into an electrophilic linchpin for subsequent follow-up modifications. For example, reagents that contain multiple electrophilic positions can react with Cys and retain an electrophilic site<sup>19–22</sup>, Cys can be converted to form

<sup>1</sup>Max-Planck-Institut für Kohlenforschung, Mülheim an der Ruhr, Germany. <sup>2</sup>Institute of Organic Chemistry, RWTH Aachen University, Aachen, Germany.

<sup>3</sup>Biochemistry and Physiology of Plants, Faculty of Biology, Bielefeld University, Bielefeld, Germany. <sup>4</sup>Institute of Plant Biology and Biotechnology, University of Münster, Münster, Germany. <sup>5</sup>Institute of Physical and Theoretical Chemistry, Goethe University Frankfurt/Main, Frankfurt/Main, Germany.

<sup>6</sup>Max Planck Institute for Multidisciplinary Sciences, Göttingen, Germany. ✉e-mail: [ritter@kofo.mpg.de](mailto:ritter@kofo.mpg.de)

dehydroalanine by elimination<sup>23–25</sup>, or be transformed into electrophilic palladium complexes<sup>26</sup>. In an additional, subsequent step, nucleophiles are added to the initially formed electrophilic products to form the desired bioconjugates. The inherent advantage of this approach is that multiple bioconjugates can be formed based on a single introduced linchpin. However, the two-step sequence used in current methods is necessary to convert Cys into a linchpin before the nucleophile can be added to avoid cross-reactivity. Otherwise, the reagent required to furnish the linchpin would be quenched by the excess nucleophile. It is therefore necessary for the formed electrophilic intermediates to persist in solution, and that these intermediates can only be functionalized with strong nucleophiles in the next step because their stability results in attenuated reactivity.

We envisioned that a Cys-selective reagent that, upon conjugation, results in formation of a more reactive electrophile *in situ* would enable reactions with various nucleophiles that are not reactive enough to be used with current methods. Our group has previously reported the synthesis and reactivity of several thianthrenium-based compounds, for example, aryl<sup>27</sup> and trifluoromethyl<sup>28</sup> thianthrenium salts, which can access reactivity that goes beyond what has been possible with conventional halides or pseudohalides<sup>29,30</sup>. Based on the promising and distinct reactivity that the thianthrenium substituent can impart, we hypothesized that vinyl thianthrenium salts can be used to achieve versatile functionalization of biomolecules with different complexity at biocompatible conditions, and provide the opportunity to install a stable linker that is shorter than other methods allow.

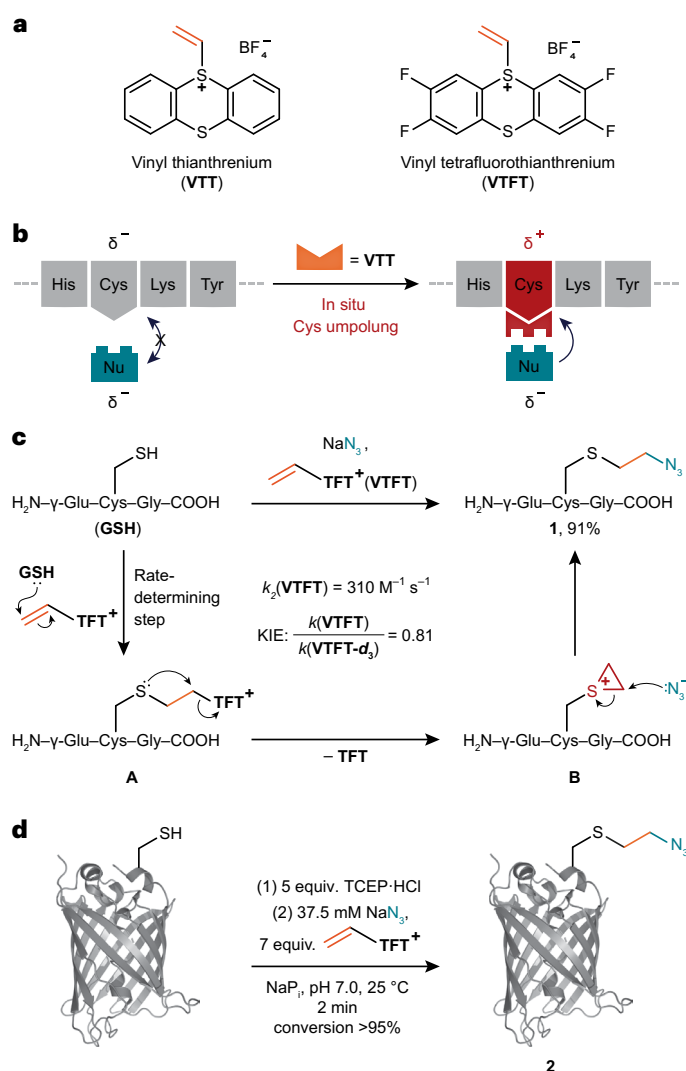
## Results and discussion

### Kinetic profile of the bioconjugation

Vinyl thianthrenium tetrafluoroborate (**VTT**) and vinyl tetrafluorothianthrenium tetrafluoroborate (**VTFT**) (Fig. 1a) are bench-stable, water-soluble vinylating reagents, which can be accessed in one step from ethylene gas<sup>31</sup>. The fundamental difference in reactivity of **VTT** and its tetrafluoro derivative **VTFT** when compared to conventional Michael acceptors such as maleimides lies in the identity and reactivity of the cationic intermediates formed after chemoselective addition to Cys to enable a one-step bioconjugation with otherwise unreactive nucleophiles present in the mixture (Fig. 1b). The cationic thianthrenium moiety in **VTT** and **VTFT** makes the reagents reactive Michael acceptors, while its ability as a leaving group enables its ensuing displacement from alkylsulfonium intermediate **A** to yield episulfonium intermediate **B** (Fig. 1c), a reactive species for reactions with various nucleophiles.

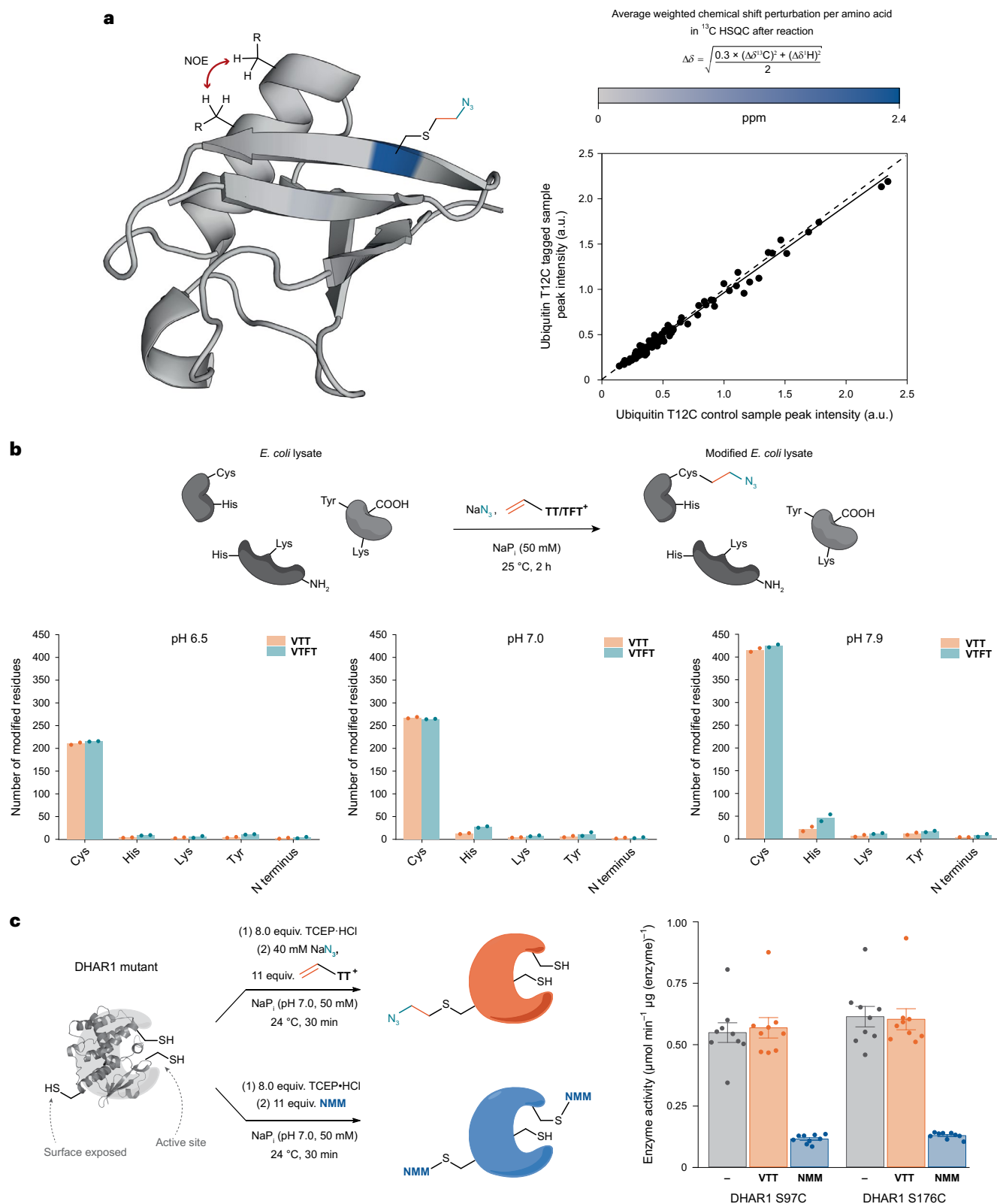
For example, *in situ* formation of **B** in the presence of sodium azide results in formation of the azidated tripeptide glutathione (**GSH**) **1** in 91% yield, without any observed side reactivity. Formation of **B** as the key intermediate is supported by NMR studies: when the reaction is performed in a deuterated buffer, the formed product has an equal distribution of deuterium over the ethylene bridge, consistent with the presence of a cyclic, symmetric intermediate prior to addition of azide (Supplementary Fig. 208). Kinetic studies revealed reaction orders of 0 for azide and 1.0 for **VTFT** and **GSH**, respectively, in line with episulfonium ring-opening after the rate-determining step. To distinguish between rate-determining Cys addition to **VTFT** and episulfonium formation, we performed a competition experiment between **VTFT** and **VTFT-d<sub>3</sub>** and observed an inverse secondary kinetic isotope effect (KIE) of 0.81 (Supplementary Figs. 201 and 202). The KIE value is consistent with a change of hybridization from *sp*<sup>2</sup> to *sp*<sup>3</sup> at the reacting carbon in the rate-limiting step and excludes rate-limiting episulfonium formation<sup>32</sup>. The data are in agreement with rate-limiting addition of Cys to **VTFT**, followed by fast episulfonium formation and ensuing rapid, irreversible nucleophilic opening with azide.

The reaction between **GSH**, **VTT** and NaN<sub>3</sub> in sodium phosphate buffer (NaP<sub>i</sub>) at pH 7.4 has a second-order rate constant of  $38.1 \pm 0.1 \text{ M}^{-1} \text{ s}^{-1}$ , and proceeds faster than the click reactions commonly used for chemical ligation<sup>33</sup> or Cys alkylation with iodoacetamide<sup>34</sup>.



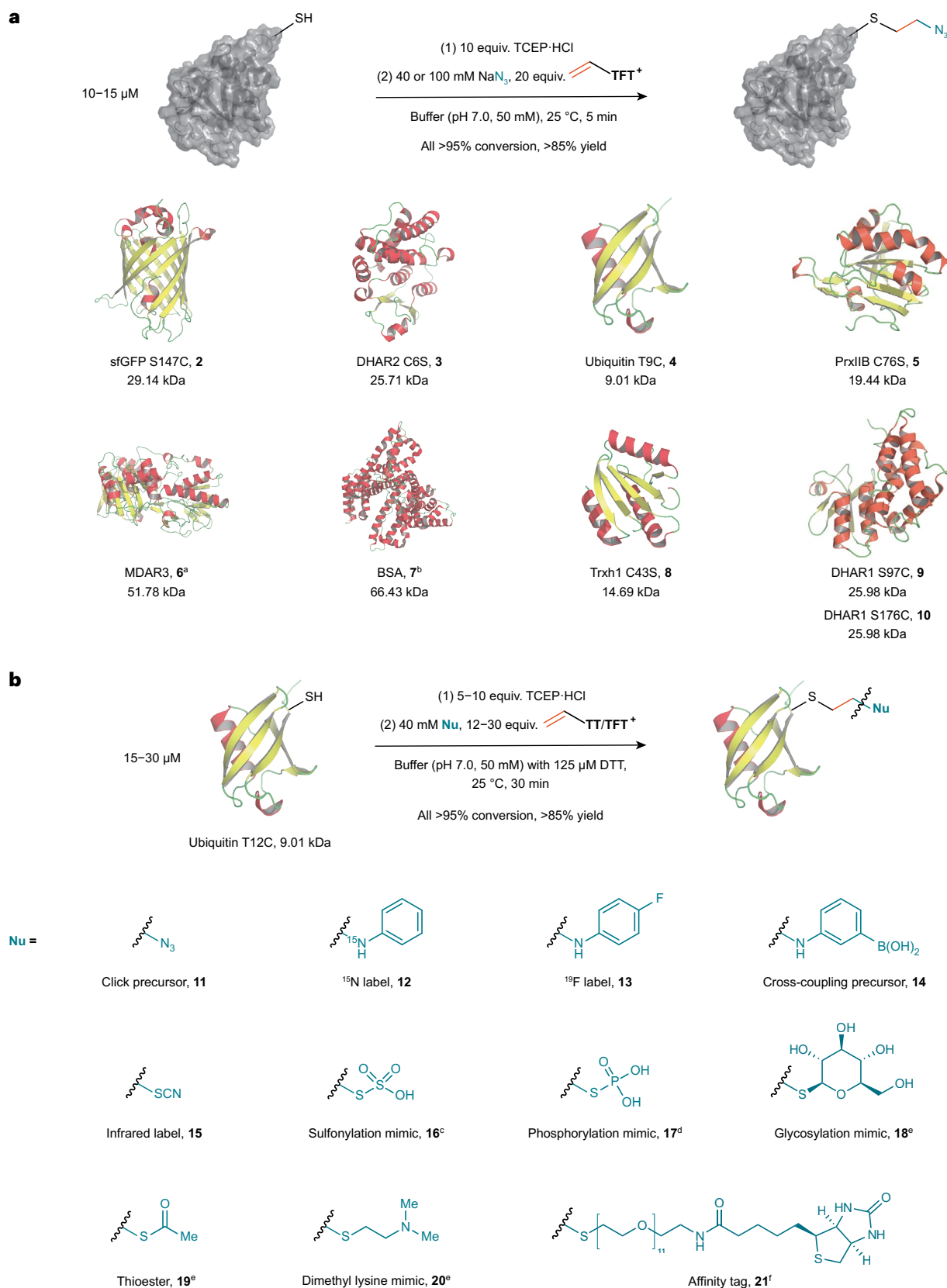
**Fig. 1 | Biofunctionalization with vinyl thianthrenium reagents.** **a**, The vinyl thianthrenium reagents **VTT** and **VTFT** utilized in this work. **b**, Strategy for the *in situ* activation of Cys based on site-selective linchpin formation in the presence of other amino acids and exogenous nucleophiles. **c**, Proposed mechanism for the umpolung of thiols based on the KIE determined via a competition experiment between **VTFT** and **VTFT-d<sub>3</sub>**, and the second-order rate constant for the functionalization of **GSH** with **VTFT**. **d**, Bioconjugation of sfGFP S147C with **VTFT** and sodium azide. The following publicly available protein structure was used: sfGFP (PDB:2B3P). His, histidine; Lys, Lysine; Tyr, tyrosine; Nu, nucleophile.

**VTFT** functionalizes Cys even faster with a determined second-order rate constant of  $310 \pm 8 \text{ M}^{-1} \text{ s}^{-1}$ , which is the same order of magnitude as the conjugation with alkylmaleimides<sup>10</sup>. The difference in rate between **VTT** and **VTFT** is due to the electronic effects arising from the four fluoride substituents on tetrafluorothianthrene (**TFT**)<sup>35</sup>. Due to the high reactivity and chemoselectivity of both **VTT** and **VTFT**, the conjugation reactions were readily extended to proteins. For example, full conversion of the superfolder green fluorescent protein (sfGFP) S147C mutant to the azidated product **2** was observed within 120 s under biocompatible reaction conditions at pH 7.0 (Fig. 1d and Supplementary Figs. 118 and 119). To ensure that the redox-active cysteinyl thiols for protein ligation are not present in higher oxidation states<sup>36</sup>, a small excess of reducing tris(2-carboxyethyl)phosphine (TCEP) was added prior to addition of vinyl thianthrenium. No reported method allows for introduction of alkyl azides into proteins from azide at such high rates and mild conditions in a single step.



**Fig. 2 | Selectivity of bioconjugation with vinyl thianthrenium salts. a**, Three-dimensional representation of the average  $^1\text{H}$  and  $^{13}\text{C}$  NMR shift perturbations of each amino acid after functionalization of  $^{13}\text{C}$ ,  $^{15}\text{N}$ -labelled ubiquitin T12C and the NOESY correlation plot of scaled cross-peak heights before and after functionalization of ubiquitin T12C. The linear regression of the experimental data is indicated as a solid line. **b**, Mean number of modified residues in *E. coli* lysate resulting from reaction with **VTT**/**VTFT** and sodium azide at different pH values. Data are derived from two technical replicates. TCEP was used for 1 h

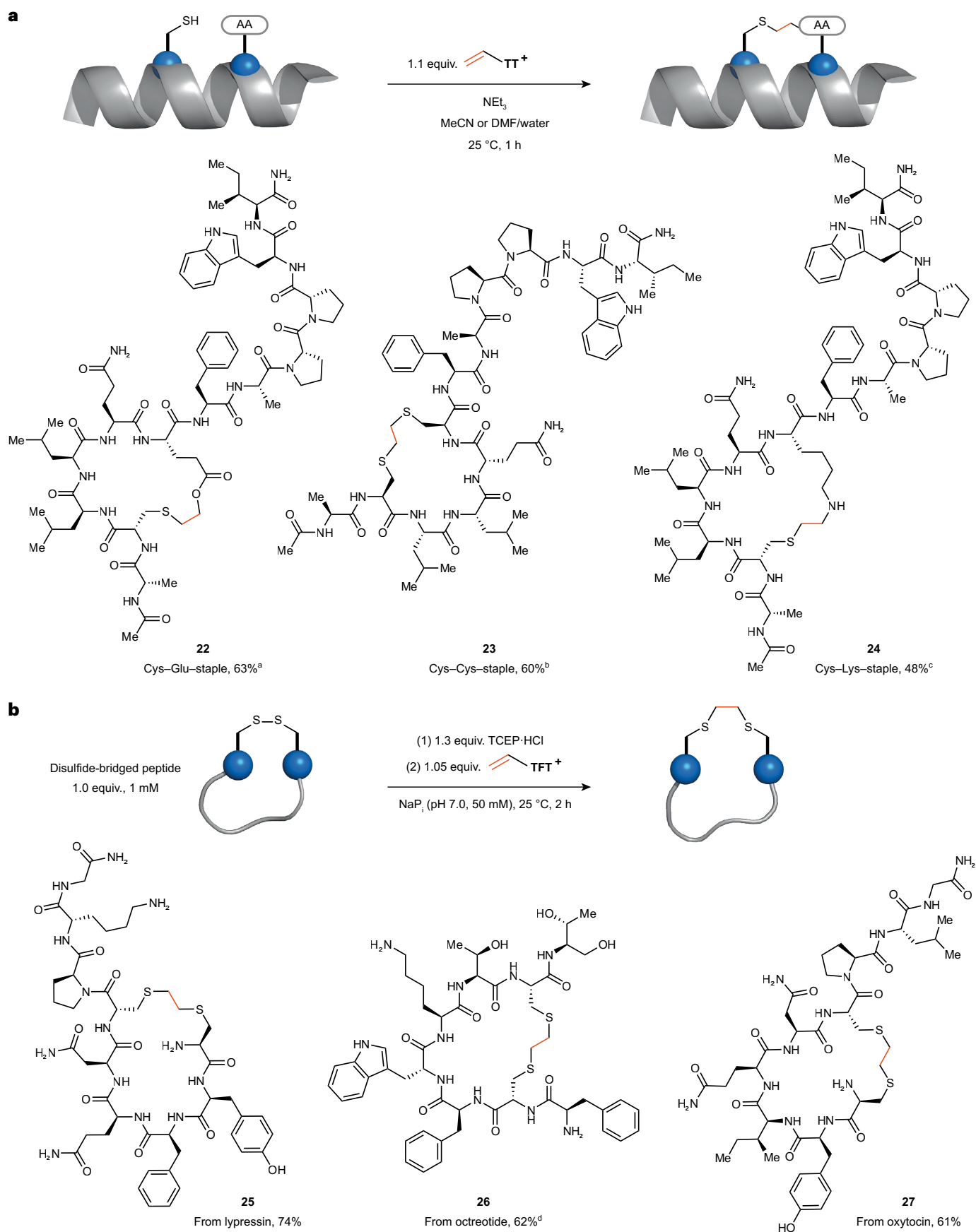
pre-reduction at 37 °C. **c**, Reactivity of DHAR1 mutants with **VTT** and *N*-methylmaleimide (**NMM**) and enzyme assay subsequent to the modification. The data are expressed as mean  $\pm$  s.e. of nine independent experiments. The statistical analysis was performed via a two-sided analysis of variance, followed by post hoc Tukey test with  $P < 0.05$  (see Supplementary Tables 11 and 12 for exact  $P$  values). The following publicly available protein structures were used: ubiquitin (PDB:1D3Z) and DHAR1 (PDB:SEL8).



**Fig. 3 | Scope of vinyl thianthrenium salts as bioconjugation reagents.**

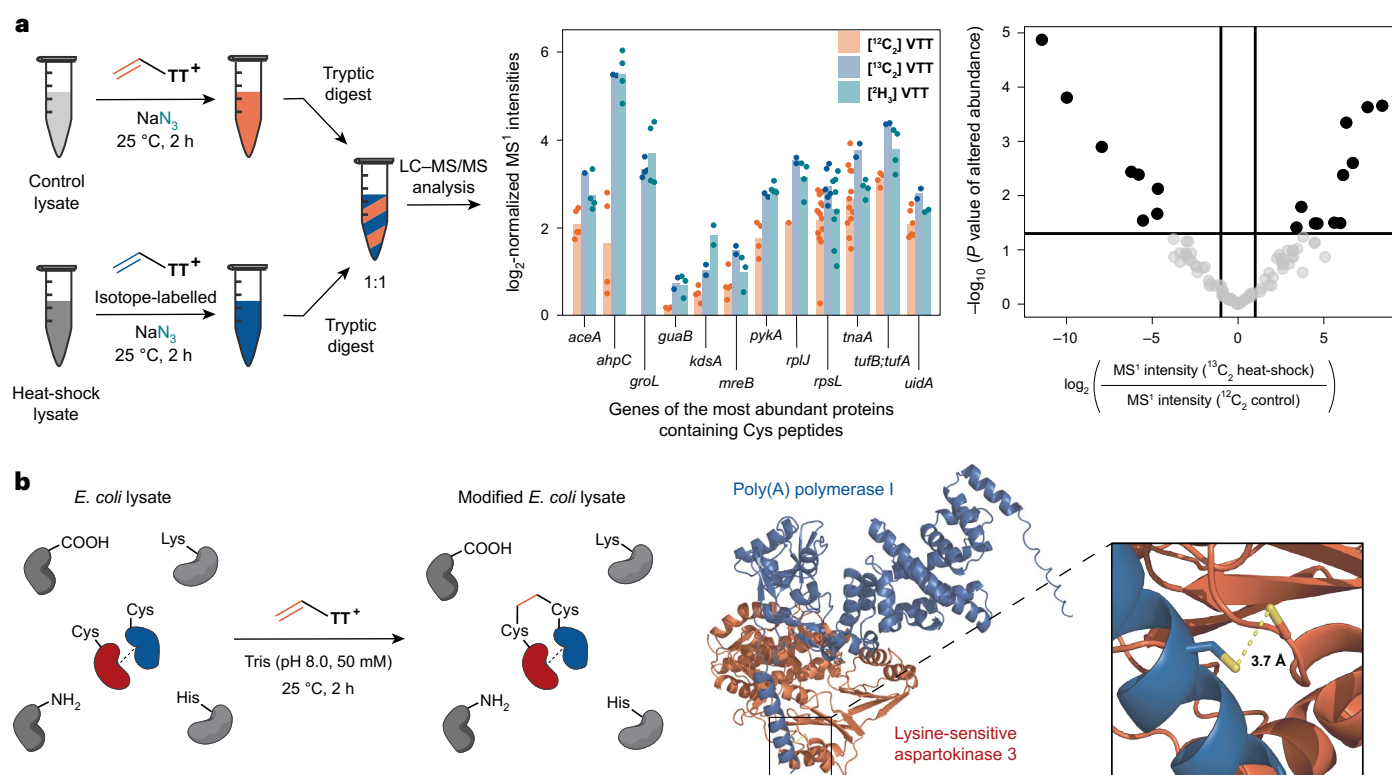
**a**, Scope of proteins used for conjugation reactions with  $\text{NaN}_3$ . **b**, Scope of functionalities introduced with  $\text{VTFT/VTFT}^+$ . <sup>a</sup>40 equiv.  $\text{VTFT}$ , yield not determined; <sup>b</sup>yield not determined; <sup>c</sup>10 mM Nu; <sup>d</sup>3 mM Nu, 35 equiv.  $\text{VTFT}$ ; <sup>e</sup>40 mM sodium iodide as nucleophile with subsequent addition of 40 mM S-centred nucleophiles after 3 min; <sup>f</sup>ubiquitin T9C, HEPES (8.0, 50 mM), 40 mM sodium

iodide as nucleophile with subsequent addition of 20 mM S-centred nucleophile after 2 min, yield, 74%. The following publicly available protein structures were used: sfGFP (PDB:2B3P), DHAR2 (PDB:5L0L), ubiquitin (PDB:1UBQ), Trxh1 (PDB:1XFL), DHAR1 (PDB:5EL8), BSA (PDB:3V03). MDAR, monodehydroascorbate reductase; PrxIIB, type 2 peroxidase; Trxh1, thioredoxin-*h1*; DTT, dithiothreitol.



**Fig. 4 | Utilization of vinyl thianthreniums for stapling and macrocyclization of peptides. a**, Stapling of nucleophilic amino acids in native peptides with **VTT**.

<sup>a</sup>MeCN:H<sub>2</sub>O 9:1, 2.2 equiv. NEt<sub>3</sub>. <sup>b</sup>DMF:H<sub>2</sub>O 1:1, 5.0 equiv. NEt<sub>3</sub>. <sup>c</sup>MeCN:H<sub>2</sub>O 9:1, 3.0 equiv. NEt<sub>3</sub>. **b**, Macrocyclization of reduced disulfide bonds with **VTFT**. <sup>d</sup>NaPi (pH 6.5, 50 mM). AA, amino acid.



**Fig. 5 | Applications of VTT in proteomics. a**, Application of VTT isotopologues in quantitative proteomics. Reaction with control and heat-shock lysates from *E. coli* cultures incubated at 37 °C or 43 °C, respectively, prior to lysis; mean of MS<sup>1</sup> intensities from Cys-containing peptides functionalized with isotopologues of VTT and their respective protein origins. Data are derived from two technical replicates; volcano-plot with auxiliary lines resulting from two-sided linear models for microarray data (LIMMA)-based differential expression analysis<sup>60</sup> with  $P < 0.05$  derived from MS<sup>1</sup> intensities of identified peptides based on two

technical replicates. All  $P$  values are associated with empirical Bayes moderated  $t$ -statistics without adjustments for multiple comparisons. The horizontal auxiliary line indicates the limit of significance based on  $P < 0.05$ . The vertical auxiliary lines indicate the values at which the protein populations have at least halved or doubled. **b**, Identified and simulated PPI between poly(A) polymerase I and lysine-sensitive aspartokinase 3 via VTT induced cross-linking in an *E. coli* lysate. MS<sup>1</sup>, ion spectra of intact peptides.

### Site selectivity and site specificity of the reaction

Chemical modifications of biomolecules are often utilized to study native structures and enzymatic activity. Therefore, it is important that bioconjugation is site selective and the protein structure not significantly altered<sup>37</sup>. We used protein NMR as a non-invasive method to inquire about the chemoselectivity of the thianthrene-based transformation and the potential resulting structural perturbation in the tertiary structure of the protein<sup>38</sup>. Heteronuclear single quantum coherence (HSQC) NMR experiments of <sup>13</sup>C, <sup>15</sup>N-labelled ubiquitin T12C prior to and after reaction with VTT and sodium azide showed an average weighted chemical shift perturbation of 2.40 ppm for Cys in the <sup>13</sup>C NMR resonances. All other amino acids only displayed changes of up to a maximum of 0.11 ppm in the <sup>13</sup>C NMR resonances, and amino acids at a distance larger than 16 Å relative to the Cys sulfur atom only showed perturbations up to 0.01 ppm. The data are consistent with high chemoselectivity for Cys modification (Fig. 2a). In addition, nuclear Overhauser effect spectroscopy (NOESY) experiments of unlabelled protein before and after conjugation were carried out. The obtained data indicate less than 3% change in interproton distances and therefore minimal structural changes and preservation of the tertiary protein structure (Fig. 2a).

When the reaction was performed on an entire *Escherichia coli* (*E. coli*) cell lysate containing various native proteins, we observed 90% chemoselectivity for Cys modification for VTT and 85% for VTFT at pH 7.0, as determined by liquid chromatography–tandem mass spectrometry (LC–MS/MS) analysis of the tryptic digest (Fig. 2b). Even when the pH was changed to 6.5 and 7.9, respectively, the high selectivity

was retained in all cases, which is noteworthy given the presence of a variety of native proteins with different microenvironments and numerous competing amino acid residues, including hyper-reactive residues<sup>12,39</sup>. Previously reported Cys-selective reagents for proteomics such as iodoacetamide or tetrafluoroalkyl benziodoxoles reach Cys coverages of 0.5% and 1.6%, respectively<sup>40,41</sup>. At pH 7.9, VTFT labels 2.7% of all available cysteines even at concentrations of only 0.2 mM (Supplementary Table 2).

Site-specific Cys modifications are often challenging due to insufficient discrimination between different Cys residues. For example, maleimides are usually not able to distinguish between multiple Cys residues on the same protein and therefore cannot usually be utilized to modify proteins that contain catalytically active cysteinyl thiols<sup>8</sup>. To evaluate a potential site specificity of VTT, we introduced surface-exposed Cys residues into dehydroascorbate reductase 1 (DHAR1), which already features two catalytic Cys residues in the active site. Functionalization of both S97C and S176C mutants with an excess of VTT led to a single modification of the introduced surface-exposed Cys and preservation of the enzyme activity. In contrast, conjugation of both enzymes with the equimolar concentration of *N*-methylmaleimide led to functionalization of both surface-exposed and internal, catalytic Cys, with a concomitant substantial decrease in catalytic activity (Fig. 2c). When no surface-exposed Cys was introduced into the protein, VTT did not interfere with the catalytic Cys residues of DHAR1. However, maleimide reagents with different substituents readily react with the catalytic Cys residues of the native protein (Supplementary Figs. 141–143).

Although steric repulsion might prevent **VTT** from reacting with the catalytic Cys residues, the unspecific reactivity of different arylmaleimide derivatives with DHAR1 suggests that the size of the reagent might not be the decisive factor for the site specificity. We speculate that the site specificity of vinyl thianthrenium reagents arises from the cationic charge of the thianthrene moiety which leads to repulsion in hydrophobic areas or environments with high concentrations of positively charged amino acids. These results bode well for potential applications of **VTT** for site-specific functionalization of proteins or enzymes that contain several and even catalytic Cys residues.

### Episulfonium enables versatile bioconjugation

Based on their promising reactivity and selectivity, we evaluated a diverse set of recombinant and native Cys-containing proteins of various sizes and functions for azidoethylation (Fig. 3a). In each case, the LC-MS analysis of the product shows >95% conversion in all the cases and formation of the desired species in yields of at least 85% in all the quantifiable cases. Although close to the protein surface as a consequence of the short linker, the furnished azide can readily be utilized in copper-catalysed or strain-promoted click reactions (Supplementary Figs. 97–100).

In contrast to the previously reported methods for cysteine bis-alkylation<sup>42</sup>, the episulfonium intermediate, such as **B**, does not eliminate to form dehydroalanine. Instead, given the high reactivity of **B** for nucleophilic ring opening, bioconjugation through opening of **B** with other bioorthogonal nucleophiles, less reactive than azide, is possible and efficient (Fig. 3b). For example, anilines can be used to install NMR-active nuclei<sup>43</sup> (**12**, **13**) or bioorthogonal functionalities<sup>44</sup> (**14**). Introduction of sulfur-centred nucleophiles gives access to thiocyanate (**15**) as an infrared label<sup>45</sup>, thiosulfate (**16**) as a mimic for sulfation<sup>46</sup> and thiophosphate as mimic for phosphorylation<sup>47</sup> (**17**). Iodide can be utilized to form primary alkyl iodide intermediates (Supplementary Figs. 211, 213 and 214), which can then be functionalized with thiol-based nucleophiles to introduce mimics for glycosylation<sup>48</sup> (**18**), reversible modifications in the form of thioesters<sup>49</sup> (**19**), lysine methylation<sup>50</sup> (**20**) or to attach an affinity tag for purification (**21**). The episulfonium approach enables bioconjugation with a broad set of different nucleophiles as weak as anilines in a convenient one-pot process, which no other method appears to allow. No hydrolysis, that is, episulfonium opening with the solvent water, was detected for the various nucleophiles shown in Fig. 3b, except for the anilines and **21**, when up to 15% of the water addition by-product was observed as determined by LC-MS (Supplementary Table 5).

To investigate potential side reactivity that might go undetected in the presence of excess exogenous nucleophiles, we performed control reactions without any additives except for reductant and buffer. When sfGFP S147C, DHAR1 S97C and DHAR1 S176C were modified with **VTT** in the absence of additional nucleophiles, LC-MS measurements revealed intramolecular ring opening of the episulfonium intermediates and phosphate adducts from the phosphate buffer in addition to the already observed hydrolysis side reaction. Because intramolecular cross-linking can be accompanied by structural alteration and loss of function<sup>51</sup>, the proteins were subjected to activity assays to probe such scenarios. All activity assays of the modified proteins, however, showed no statistically relevant alteration of protein function compared to the non-modified proteins (Supplementary Figs. 162 and 167). We speculate that the formation of the covalent linkage between two nucleophilic amino acids only takes place if the two residues are already located in close proximity to each other and therefore does not influence the tertiary protein structure and the function of the protein upon intramolecular reaction. Dose-dependence experiments with sodium azide revealed that the presence of 40 mM of exogenous nucleophile is sufficient to ensure the formation of the desired product (Supplementary Fig. 110). The fact that different nucleophiles can be installed only four

atoms remote from the amide backbone allows for modulation of the protein in close proximity to its surface.

### Peptide stapling with vinyl thianthrenium

We hypothesized that the intramolecular reactivity of the episulfonium intermediate with nucleophilic amino acids can be utilized for the synthesis of macrocyclic, stapled peptides. Peptide stapling can induce and stabilize helicity in peptides, which increases metabolic stability<sup>52</sup> – a major goal in the development of peptide-based therapeutics<sup>53</sup>. Stapling of natural amino acids is attractive because it allows the use of native peptides to avoid the synthesis of unnatural amino acids and peptides<sup>54</sup>. We used **VTT** to introduce the short, stable ethylene linker into peptides containing Cys and another nucleophilic amino acid in an *i, i + 4* relationship<sup>55</sup>. Cys–Lys, Cys–Cys and Cys–Glu linkages were introduced to form stapled peptides in 48–63% isolated yield (Fig. 4a). This strategy provides an excellent method for stapling of three different pairs of natural amino acids, including non-activated carboxylic acids, via a very small hydrocarbon-based linker. Vinyl thianthrenium salts can also be used to irreversibly transform labile disulfide bonds in pharmaceutically relevant peptides into thioethers and provide immediate access to stable disulfide mimetics under physiological conditions. Macrocyclic analogues of oxytocin, octreotide and lypressin were obtained directly from the native peptides in 61–74% isolated yield (Fig. 4b).

### Quantitative proteomics and protein–protein cross-linking

The scalability of and easy synthetic access to **VTT** and **VTFT** from ethylene allows for a straightforward synthesis of isotopically enriched reagents for advanced applications in mass spectrometry (MS). For example, MS-based quantitative proteomics relies on isotope-enriched reagents and is used to quantify proteins<sup>56</sup> or redox states<sup>57</sup> in complex native systems. We accomplished the synthesis of [<sup>2</sup>H<sub>3</sub>]VTT and [<sup>13</sup>C<sub>2</sub>]VTT isotopologues on a decagram scale in a single step. To evaluate whether the reagents can serve to read out and quantify cellular stress responses, we performed a heat-shock experiment with *E. coli* as a model organism. We treated *E. coli* lysates derived from cell cultures that were incubated at 37 °C and 43 °C with unlabelled and labelled **VTT**, respectively, in the presence of azide. Equimolar mixtures of unlabelled and labelled tryptic digests of the lysates were then analysed via LC-MS/MS. The 1:1 mixture of <sup>12</sup>C- and <sup>13</sup>C-modified lysates revealed 244 Cys-containing peptides, of which 20 showed a statistically significant change in abundance caused by the elevated temperature (Fig. 5a). The same experiment with <sup>2</sup>H<sub>3</sub>-labelled **VTT** resulted in a Pearson correlation coefficient of 0.76 when compared to the results with <sup>13</sup>C-labelled **VTT**. The reliability of the experimental set-up was further supported by data obtained for the *groL* gene that codes for a heat-shock protein<sup>58</sup> (Fig. 5a). Heat-shock proteins are expressed to ensure proper folding of other proteins under extreme conditions, and the corresponding peptides were only found in the lysates incubated at 43 °C. If desired, the introduced azide can be utilized to attach an additional affinity handle after the initial functionalization step to increase sensitivity of the whole set-up<sup>15,16,40</sup>. Hence, vinyl thianthrenium salts can be used for isotope-coded affinity tagging strategy with adjustable sensitivity induced by the choice of nucleophile.

As a salient feature of **VTT** is its ability to introduce a short two-carbon linker, which may identify protein–protein interactions (PPIs) that are not detectable with other reagents<sup>59</sup>. To test this hypothesis, we treated *E. coli* lysates with **VTT** in the absence of additional exogenous nucleophiles. Trypsin digestion and LC-MS/MS analysis led to the identification of a previously unappreciated PPI, discovered through a Cys–Cys cross-link between lysine-sensitive aspartokinase 3 and poly(A) polymerase I. Based on the short length of the ethylene linkage and the identity and the position of the cross-linked amino acids, a protein-docking simulation with geometrical restraints from 2 to 6 Å was performed. The calculation indicates a PPI between the two

proteins where both Cys residues are located 3.7 Å apart from each other (Fig. 5b). The same PPI was detected in similar abundance when the heat-shock lysate was modified with isotope-labelled VTT derivatives. Currently, no other Cys–Cys cross-linking reagents appear to be available that would furnish such a short linkage and, hence, be able to link two residues so proximal to one another. Moreover, there appears to be no comparable system that provides a quantifiable tool for applications in Cys tagging and cross-linking based on a single reagent.

## Conclusions

Cys conjugation is a state-of-the-art tactic to introduce modifications into biomolecules, but access to a variety of modifications currently requires substantial synthetic effort to prepare new reagents or is complicated by the low reactivity of introduced linchpins. Vinyl thianthrenium salts unite versatility and site selectivity, and lead to a common, reactive, cationic species that can be diversified to a large battery of conjugates with a short two-carbon linker. The single set of a few globally applicable and practical thianthrenium salts, including their readily accessible isotopologues, permits user-friendly applications in quantitative derivatization and analyses of single peptides, proteins and complex protein mixtures in different important disciplines in chemical biology.

## Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41557-023-01388-7>.

## References

- Fodje, M. N. & Al-Karadaghi, S. Occurrence, conformational features and amino acid propensities for the  $\pi$ -helix. *Protein Eng. Des. Sel.* **15**, 353–358 (2002).
- Gunnoo, S. B. & Madder, A. Chemical protein modification through cysteine. *ChemBioChem* **17**, 529–553 (2016).
- Lin, S. et al. Redox-based reagents for chemoselective methionine bioconjugation. *Science* **355**, 597–602 (2017).
- Taylor, M. T., Nelson, J. E., Suero, M. G. & Gaunt, M. J. A protein functionalization platform based on selective reactions at methionine residues. *Nature* **562**, 563–568 (2018).
- Seki, Y. et al. Transition metal-free tryptophan-selective bioconjugation of proteins. *J. Am. Chem. Soc.* **138**, 10798–10801 (2016).
- Tower, S. J., Hetcher, W. J., Myers, T. E., Kuehl, N. J. & Taylor, M. T. Selective modification of tryptophan residues in peptides and proteins using a biomimetic electron transfer process. *J. Am. Chem. Soc.* **142**, 9112–9118 (2020).
- Li, B. X. et al. Site-selective tyrosine bioconjugation via photoredox catalysis for native-to-bioorthogonal protein transformation. *Nat. Chem.* **13**, 902–908 (2021).
- Stephanopoulos, N. & Francis, M. B. Choosing an effective protein bioconjugation strategy. *Nat. Chem. Biol.* **7**, 876–884 (2011).
- Bernardim, B. et al. Stoichiometric and irreversible cysteine-selective protein modification using carbonylacrylic reagents. *Nat. Commun.* **7**, 13128 (2016).
- Ravasco, J. M. J. M., Faustino, H., Trindade, A. & Gois, P. M. P. Bioconjugation with maleimides: a useful tool for chemical biology. *Chem. Eur. J.* **25**, 43–59 (2019).
- De Geyter, E. et al. 5-Hydroxy-pyrrolone based building blocks as maleimide alternatives for protein bioconjugation and single-site multi-functionalization. *Chem. Sci.* **12**, 5246–5252 (2021).
- Weerapana, E. et al. Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature* **468**, 790–795 (2010).
- Vinogradova, E. V., Zhang, C., Spokoiny, A. M., Pentelute, B. L. & Buchwald, S. L. Organometallic palladium reagents for cysteine bioconjugation. *Nature* **526**, 687–691 (2015).
- Bernardes, G. J. L. et al. From disulfide- to thioether-linked glycoproteins. *Angew. Chem. Int. Ed.* **47**, 2244–2247 (2008).
- Abegg, D. et al. Proteome-wide profiling of targets of cysteine reactive small molecules by using ethynyl benziodoxolone reagents. *Angew. Chem. Int. Ed.* **54**, 10852–10857 (2015).
- Tessier, R. et al. Ethynylation of cysteine residues: from peptides to proteins in vitro and in living cells. *Angew. Chem. Int. Ed.* **59**, 10961–10970 (2020).
- Kasper, M.-A. et al. Cysteine-selective phosphoramidate electrophiles for modular protein bioconjugations. *Angew. Chem. Int. Ed.* **58**, 11625–11630 (2019).
- Zhang, C. et al.  $\pi$ -Clamp-mediated cysteine conjugation. *Nat. Chem.* **8**, 120–128 (2016).
- Smith, M. E. B. et al. Protein modification, bioconjugation, and disulfide bridging using bromomaleimides. *J. Am. Chem. Soc.* **132**, 1960–1965 (2010).
- Zhang, Y. et al. Cysteine-specific protein multi-functionalization and disulfide bridging using 3-bromo-5-methylene pyrrolones. *Nat. Commun.* **11**, 1015 (2020).
- Wall, A. et al. One-pot thiol–amine bioconjugation to maleimides: simultaneous stabilisation and dual functionalisation. *Chem. Sci.* **11**, 11455–11460 (2020).
- Laserna, V. et al. Protein conjugation by electrophilic alkynylation using 5-(alkynyl)dibenzothiophenium triflates. *Bioconjugate Chem.* **32**, 1570–1575 (2021).
- Bernardes, G. J. L., Chalker, J. M., Errey, J. C. & Davis, B. G. Facile conversion of cysteine and alkyl cysteines to dehydroalanine on protein surfaces: versatile and switchable access to functionalized proteins. *J. Am. Chem. Soc.* **130**, 5052–5053 (2008).
- Dadová, J., Galan, S. R. G. & Davis, B. G. Synthesis of modified proteins via functionalization of dehydroalanine. *Curr. Opin. Chem. Biol.* **46**, 71–81 (2018).
- Josephson, B. et al. Light-driven post-translational installation of reactive protein side chains. *Nature* **585**, 530–537 (2020).
- Gazvoda, M. et al. Palladium-mediated incorporation of carboranes into small molecules, peptides, and proteins. *J. Am. Chem. Soc.* **144**, 7852–7860 (2022).
- Berger, F. et al. Site-selective and versatile aromatic C–H functionalization by thianthrenation. *Nature* **567**, 223–228 (2019).
- Jia, H., Häring, A. P., Berger, F., Zhang, L. & Ritter, T. Trifluoromethyl thianthrenium triflate: a readily available trifluoromethylating reagent with formal  $\text{CF}_3^+$ ,  $\text{CF}_3^{\cdot}$ , and  $\text{CF}_3^-$  reactivity. *J. Am. Chem. Soc.* **143**, 7623–7628 (2021).
- Zhao, D., Petzold, R., Yan, J., Muri, D. & Ritter, T. Tritiation of aryl thianthrenium salts with a molecular palladium catalyst. *Nature* **600**, 444–449 (2021).
- Li, J. et al. Photoredox catalysis with aryl sulfonium salts enables site-selective late-stage fluorination. *Nat. Chem.* **12**, 56–62 (2020).
- Juliá, F., Yan, J., Paulus, F. & Ritter, T. Vinyl thianthrenium tetrafluoroborate: a practical and versatile vinylating reagent made from ethylene. *J. Am. Chem. Soc.* **143**, 12992–12998 (2021).
- Gomez-Gallego, M. & Sierra, M. A. Kinetic isotope effects in the study of organometallic reaction mechanisms. *Chem. Rev.* **111**, 4857–4963 (2011).
- Saito, F., Noda, H. & Bode, J. W. Critical evaluation and rate constants of chemoselective ligation reactions for stoichiometric conjugations in water. *ACS Chem. Biol.* **10**, 1026–1033 (2015).
- Motiwala, H. F., Kuo, Y.-H., Stinger, B. L., Palfey, B. A. & Martin, B. R. Tunable heteroaromatic sulfones enhance in-cell cysteine profiling. *J. Am. Chem. Soc.* **142**, 1801–1810 (2020).



35. Smart, B. E. Fluorine substituent effects (on bioactivity). *J. Fluor. Chem.* **109**, 3–11 (2001).
36. Paulsen, C. E. & Carroll, K. S. Cysteine-mediated redox signaling: chemistry, biology, and tools for discovery. *Chem. Rev.* **113**, 4633–4679 (2013).
37. Danielson, M. A. & Falke, J. J. Use of <sup>19</sup>F NMR to probe protein structure and conformational changes. *Annu. Rev. Biophys.* **25**, 163–195 (1996).
38. Guntert, P. Structure calculation of biological macromolecules from NMR data. *Q. Rev. Biophys.* **31**, 145–237 (1998).
39. Hacker, S. M. et al. Global profiling of lysine reactivity and ligandability in the human proteome. *Nat. Chem.* **9**, 1181–1190 (2017).
40. Abegg, D. et al. Chemoproteomic profiling by cysteine fluoroalkylation reveals Myrocin G as an inhibitor of the nonhomologous end joining DNA repair pathway. *J. Am. Chem. Soc.* **143**, 20332–20342 (2021).
41. Suttapitugsakul, S., Xiao, H. P., Smeekens, J. & Wu, R. H. Evaluation and optimization of reduction and alkylation methods to maximize peptide identification with MS-based proteomics. *Mol. Biosyst.* **13**, 2574–2582 (2017).
42. Chalker, J. M. et al. Methods for converting cysteine to dehydroalanine on peptides and proteins. *Chem. Sci.* **2**, 1666–1676 (2011).
43. Marsh, E. N. G. & Suzuki, Y. Using <sup>19</sup>F NMR to probe biological interactions of proteins and peptides. *ACS Chem. Biol.* **9**, 1242–1250 (2014).
44. Akgun, B. & Hall, D. G. Boronic acids as bioorthogonal probes for site-selective labeling of proteins. *Angew. Chem. Int. Ed.* **57**, 13028–13044 (2018).
45. Fafarman, A. T., Webb, L. J., Chuang, J. I. & Boxer, S. G. Site-specific conversion of cysteine thiols into thiocyanate creates an IR probe for electric fields in proteins. *J. Am. Chem. Soc.* **128**, 13356–13357 (2006).
46. Stone, M. J., Chuang, S., Hou, X., Shoham, M. & Zhu, J. Z. Tyrosine sulfation: an increasingly recognised post-translational modification of secreted proteins. *New Biotechnol.* **25**, 299–317 (2009).
47. Ardito, F., Giuliani, M., Perrone, D., Troiano, G. & Lo Muzio, L. The crucial role of protein phosphorylation in cell signaling and its use as targeted therapy. *Int. J. Mol. Med.* **40**, 271–280 (2017).
48. Moremen, K. W., Tiemeyer, M. & Nairn, A. V. Vertebrate protein glycosylation: diversity, synthesis and function. *Nat. Rev. Mol. Cell Biol.* **13**, 448–462 (2012).
49. Chen, J., Zhao, M., Feng, F., Sizovs, A. & Wang, J. Tunable thioesters as ‘reduction’ responsive functionality for traceless reversible protein PEGylation. *J. Am. Chem. Soc.* **135**, 10938–10941 (2013).
50. Lanouette, S., Mongeon, V., Figeys, D. & Couture, J.-F. The functional diversity of protein lysine methylation. *Mol. Syst. Biol.* **10**, 724 (2014).
51. Migneault, I., Dartiguenave, C., Bertrand, M. J. & Waldron, K. C. Glutaraldehyde: behavior in aqueous solution, reaction with proteins, and application to enzyme crosslinking. *BioTechniques* **37**, 790–802 (2004).
52. Schafmeister, C. E., Po, J. & Verdine, G. L. An all-hydrocarbon cross-linking system for enhancing the helicity and metabolic stability of peptides. *J. Am. Chem. Soc.* **122**, 5891–5892 (2000).
53. Verdine, G. L. & Hilinski, G. J. Stapled peptides for intracellular drug targets. *Methods Enzymol.* **503**, 3–33 (2012).
54. Li, X., Chen, S., Zhang, W.-D. & Hu, H.-G. Stapled helical peptides bearing different anchoring residues. *Chem. Rev.* **120**, 10079–10144 (2020).
55. Jo, H. et al. Development of  $\alpha$ -helical calpain probes by mimicking a natural protein–protein interaction. *J. Am. Chem. Soc.* **134**, 17704–17713 (2012).
56. Gygi, S. P. et al. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **17**, 994–999 (1999).
57. Leichert, L. I. et al. Quantifying changes in the thiol redox proteome upon oxidative stress in vivo. *Proc. Natl Acad. Sci. USA* **105**, 8197–8202 (2008).
58. Hemmingsen, S. M. et al. Homologous plant and bacterial proteins chaperone oligomeric protein assembly. *Nature* **333**, 330–334 (1988).
59. Sinz, A. Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein–protein interactions. *Mass Spectrom. Rev.* **25**, 663–682 (2006).
60. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023

## Methods

### General protein modification protocol

At 20–25 °C, the protein starting material in NaP<sub>i</sub>, HEPES or Bis-Tris buffer (pH 7.0, 50 mM) was added to a 1.5 ml Eppendorf tube. Next, the protein solution was diluted with buffer and TCEP (5–10 equiv.) in ultra-high-quality (UHQ)-H<sub>2</sub>O was added. The mixture was vortexed for 1 s, transferred into a Thermocycler preheated at 37 °C, and incubated at 37 °C for 1 h at 400 r.p.m. Next, a nucleophile stock solution in the respective buffer (pH 7.0, 50 mM) was added to the mixture at 20–25 °C to obtain a nucleophile concentration of 3–100 mM and a protein concentration of 15–30 μM. The nucleophile concentration varies depending on the chosen nucleophile (Supplementary Information). Next, a vinyl thianthrenium salt (12–30 equiv.) was added as a stock solution in dimethylformamide (DMF). The amount of DMF should not exceed 1% v/v of the whole mixture. The reaction mixture was vortexed for 1 s, transferred into a Thermocycler preheated at 25 °C, and incubated at 25 °C at 400 r.p.m. for either 5 min (VTFT) or 30 min (VTT).

### Protein modification protocol for thiol nucleophiles

At 20–25 °C, the protein starting material in NaP<sub>i</sub>, HEPES or Bis-Tris buffer (pH 7.0, 50 mM) was added to a 1.5 ml Eppendorf tube. Next, the protein solution was diluted with buffer and TCEP (5–10 equiv.) in UHQ-H<sub>2</sub>O was added. The mixture was vortexed for 1 s, transferred into a Thermocycler preheated at 37 °C and incubated at 37 °C for 1 h at 400 r.p.m. Next, a sodium iodide stock solution in the respective buffer (pH 7.0, 50 mM) was added to the mixture at 20–25 °C to obtain an iodide concentration of 40 mM and a protein concentration of 15–30 μM. Next, 15–30 equiv. of VTFT was added as a stock solution in DMF. The amount of DMF should not exceed 1% v/v of the whole mixture. The reaction mixture was vortexed for 1 s, transferred into a Thermocycler preheated at 25 °C, and incubated at 25 °C at 400 r.p.m. for exactly 3 min. Immediately afterwards, a thiol stock solution in the respective buffer (pH 7.0 or 8.0, 50 mM) was added to the mixture to obtain a nucleophile concentration of 20–50 mM. The concentration varies depending on the chosen thiol (Supplementary Information). The mixture was incubated at 25 °C at 400 r.p.m. for 30 min.

### Peptide stapling

At 20–25 °C, a screwcap vial (20 ml) was charged with peptide **S10** (10.0 mg, 7.00 μmol, 1.00 equiv.) and a Teflon-coated magnetic stirring bar. MeCN (4.5 ml) and UHQ-H<sub>2</sub>O (500 μl) were added and the resulting solution was stirred for 3 min. Then, NEt<sub>3</sub> (2.15 μl, 1.56 mg, 15.4 μmol, 2.20 equiv.) was added and the mixture was stirred again for 3 min. Next, a 1 M solution of VTT in DMF (7.70 μl, 7.70 μmol, 1.10 equiv.) was added and the solution was stirred for 1 h. The reaction mixture was filtered and the filtrate was purified by HPLC. Fractions containing the product were collected and lyophilized to afford the stapled peptide **22** as a colourless powder (6.4 mg, 4.4 μmol, 63% yield).

### Disulfide rebridging

At 20–25 °C, a round-bottom flask (25 ml) was charged with Lypressin (10.5 mg, 9.94 μmol, 1.00 equiv.) and a Teflon-coated magnetic stirring bar. Then, 5.0 ml of sodium phosphate buffer (pH 7.0, 100 mM) and 4.1 ml of UHQ-H<sub>2</sub>O were added to the flask and the starting material was dissolved while stirring at 300 r.p.m. Next, a freshly prepared 20 mM solution of TCEP·HCl in UHQ-H<sub>2</sub>O (0.65 ml, 13 μmol, 1.3 equiv.) was added to the flask and the mixture was stirred at 300 r.p.m. for 1.5 h at 25 °C. A 50 mM solution of VTFT in DMF (0.21 ml, 10 μmol, 1.1 equiv.) was added dropwise over 1 min and the mixture was stirred for 15 min. The reaction mixture was filtered and the filtrate was purified by HPLC. Fractions containing the product were collected and lyophilized to afford the stapled peptide as a colourless powder (9.7 mg, 7.4 μmol, 74% yield).

### Lysate labelling for quantitative proteomics

At 20–25 °C, 6.0 μl of non-stressed control or heat-shock lysate (8.6–8.7 mg ml<sup>-1</sup>, 7.8 × 10<sup>2</sup> μM Cys, 5.0 nmol Cys, 1.0 equiv.) in NaP<sub>i</sub> buffer (pH 7.0, 50 mM) was added to a 1.5 ml Eppendorf tube and diluted with 0.13 ml of NaP<sub>i</sub> buffer (pH 7.0, 50 mM). Subsequently, 5.0 μl of a TCEP stock solution (10 mM, 50 nmol, 13 μg, 10 equiv.) in UHQ-H<sub>2</sub>O was added to the mixture. The mixture was vortexed for 1 s, transferred into a Thermocycler preheated at 37 °C and incubated at 37 °C for 1 h at 400 r.p.m. Next, a 0.30 M sodium azide stock solution (30 μl, 9.0 μmol, 0.59 mg, 1.8 × 10<sup>3</sup> equiv.) in NaP<sub>i</sub> buffer (pH 7.0, 50 mM) was added to the mixture at 20–25 °C (nucleophile concentration = 38 mM), followed by addition of 0.5 μl of a VTT ([<sup>13</sup>C<sub>2</sub>]VTT or [<sup>2</sup>H<sub>3</sub>]VTT for the heat-shock lysate) stock solution (80 mM, 0.04 μmol, 0.01 mg, 8 equiv.) in DMF. The reaction mixture was vortexed for 1 s, transferred into a Thermocycler preheated at 25 °C and incubated at 25 °C at 400 r.p.m. for 120 min. Subsequently, 22 μl of a β-mercaptoethanol stock solution (0.15 M, 3.3 μmol, 6.6 × 10<sup>2</sup> equiv.) in UHQ-H<sub>2</sub>O was added, and the mixture was incubated at 25 °C for 30 min.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

Raw SDS-PAGE, enzyme assay, protein LC-MS, LC-MS/MS and NMR data are deposited in the repositories: ownCloud (<https://owncloud.gwdg.de/index.php/s/uf0bv5a6HLflfcW>), jPOSTrepo (<https://repository.jpostdb.org/preview/169290337964832fb33c1bc>), BioMagRes-Bank (<https://bmr.io/>; entry IDs 51721, 51725) and Zenodo (<https://doi.org/10.5281/zenodo.7472436>). Source data are provided with this paper.

### Acknowledgements

We thank S. Lin and E. de Pedro Beato for helpful discussions and P. Münstermann, N. Sauerborn, C. Heidgen and G. Breitenbruch for HPLC purification. P.H., K.B., M.H., F.J., J.B.J., C.F., J.M.M., A.V., H.H., M.S.S., W.S. and T.R. thank the Max-Planck-Institut für Kohlenforschung for funding. D.M., S.B. and C.G. thank the Max-Planck-Institut für Multidisziplinäre Naturwissenschaften for funding. L.V., J.E., N.M., I.F. and K.-J.D. thank the Deutsche Forschungsgemeinschaft (project IDs DI346/23-1, 469950637, 426191805 and INST 211/1038-1) for funding. R.Z. thanks the State of Hesse in the LOEWE Schwerpunkt GLUE for funding.

### Author contributions

F.J. initiated the project. P.H. and F.J. developed the bioconjugation. P.H. and J.M.M. synthesized the vinyl thianthrenium reagents. K.B. performed kinetic studies. P.H., K.B. and M.H. performed the reactions on proteins and peptides. L.V., S.B. and K.-J.D. selected, synthesized and purified proteins. P.H. and L.V. carried out the enzyme assays. J.E., R.Z., A.V., M.S.S. and H.H. obtained the MS data. P.H., J.E., R.Z., A.V., M.S.S., H.H., N.M., W.S. and I.F. analysed the MS data. C.F., J.B.J. and D.M. performed the protein NMR experiments. C.F., J.B.J., D.M. and C.G. interpreted the protein NMR data. P.H., K.B., M.H. and T.R. wrote the manuscript. T.R. directed the project.

### Funding

Open access funding provided by Max Planck Society.

### Competing interests

F.J. and T.R. may benefit from potential royalty income related to vinyl-thianthrenium-based reagents. The other authors declare no competing interests.

**Additional information**

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41557-023-01388-7>.

**Correspondence and requests for materials** should be addressed to Tobias Ritter.

**Peer review information** *Nature Chemistry* thanks Annemieke Madder, Guilhem Chaubet and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

## Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | n/a                                 | Confirmed                                                                                                                                                                                                                                                                                      |
|-------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement                                                                                                                                    |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly                                                                                                                                    |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided<br><i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>                                                               |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A description of all covariates tested                                                                                                                                                                                                                     |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons                                                                                                                                                   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. $F$ , $t$ , $r$ ) with confidence intervals, effect sizes, degrees of freedom and $P$ value noted<br><i>Give <math>P</math> values as exact values whenever suitable.</i>                            |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings                                                                                                                                                                      |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes                                                                                                                                                |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Estimates of effect sizes (e.g. Cohen's $d$ , Pearson's $r$ ), indicating how they were calculated                                                                                                                                                         |

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

- |                 |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|-----------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Data collection | Data was collected utilizing the software that is provided by the instrument vendors and details can be found in the Methods section.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| Data analysis   | Data was analyzed with Unidec (open source, version 5.0.4), Mestrenova (Mestrelab research, version 14.2.3-29241), LabSolutions (Shimadzu Corporations, version V5.97 & V5.114), Microsoft Excel 16 (Microsoft, version 16.0.5369.1000), xiSEARCH (open source, version 1.7.6.7), MaxQuant (partially open source, version 2.1.3.0), Haddock 2.4 (open source, v2.4-2022.08), Origin 2019b (OriginLab Corporation, version 9.65), UVProbe (Shimadzu Corporations, version 2.50), PyMOL (Schroedinger, Inc., version 2.5.4 & 2.4), Maestro (Schroedinger, Inc., version 12.6.149), Xcalibur (Thermo Fisher Scientific Inc., version V4.1.31.9), MassLib (MSP Kofel, version V9.4-415), R (open source, 4.1.1), AlphaFold 2 (open source, version 2.0), Cary UV Workstation (Agilent Technologies, version: 1.0.1284.0), TopSpin (Bruker, version 3.6.2 & 3.6.3), NMRPipe (open source, version 11.0 Rev 2022.193.18.59), NMRFAM-SPARKY (open source, version 1.470), Image Lab (Bio-Rad Laboratories, Inc., version 6.1.0 build 7), LabSolutions Insight Explore (Shimadzu Corporations, version 3.8 SP4) |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Raw protein LC-MS, raw protein MS/MS, and raw protein NMR data are deposited in the repositories: ownCloud (<https://owncloud.gwdg.de/index.php/s/uf0bv5a6HLlfcW>), jPOSTrepo (<https://repository.jpostdb.org/preview/169290337964832fb33c1bc>, access key: 1252), BioMagRes-Bank (accession numbers: 51721, 51725), and Zenodo (<https://doi.org/10.5281/zenodo.7472436>). The following publicly available protein structures were used: sfGFP (PDB ID 2B3P), DHAR2 (PDB ID 5L0L), Ubiquitin (PDB ID 1D3Z & 1UBQ), Trxh1 (PDB ID 1XFL), and DHAR1 (PDB ID 5EL8), BSA (PDB ID 3V03).

## Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

### Reporting on sex and gender

*Use the terms sex (biological attribute) and gender (shaped by social and cultural circumstances) carefully in order to avoid confusing both terms. Indicate if findings apply to only one sex or gender; describe whether sex and gender were considered in study design whether sex and/or gender was determined based on self-reporting or assigned and methods used. Provide in the source data disaggregated sex and gender data where this information has been collected, and consent has been obtained for sharing of individual-level data; provide overall numbers in this Reporting Summary. Please state if this information has not been collected. Report sex- and gender-based analyses where performed, justify reasons for lack of sex- and gender-based analysis.*

### Population characteristics

*Describe the covariate-relevant population characteristics of the human research participants (e.g. age, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."*

### Recruitment

*Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.*

### Ethics oversight

*Identify the organization(s) that approved the study protocol.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

### Sample size

Enzyme assays were conducted with n=2,3,6,9,18. The exact number of replicates is reported in the supplementary information for each experiment. LC-MS/MS analysis was carried out as technical replicates with n=2. Chemical synthesis and SDS-PAGE analysis were carried out as single experiments. Protein modifications were carried out as single experiments and the reported values for conversion and yield should be regarded as semi-quantitative. Synthetic reactions and protein modifications were carried out independently for multiple times and proved to be reliable and consistent.

### Data exclusions

No data was excluded.

### Replication

Enzyme assays were conducted with n=2,3,6,9,18. The exact number of replicates is reported in the supplementary information for each experiment. LC-MS/MS analysis was carried out as technical replicates with n=2. Chemical synthesis was carried out as single experiments. Protein modifications were carried out as single experiments and the reported values for conversion and yield should be regarded as semi-quantitative. Synthetic reactions and protein modifications were carried out independently for multiple times and proved to be reliable and consistent. All carried out reactions with proteins and small molecules could be repeated. The general replication of the described methods were successful.

### Randomization

Not applicable

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

- | n/a                                 | Involvement in the study                               |
|-------------------------------------|--------------------------------------------------------|
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Antibodies         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Eukaryotic cell lines         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data                 |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern  |

### Methods

- | n/a                                 | Involvement in the study                        |
|-------------------------------------|-------------------------------------------------|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq               |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |

## Antibodies

### Antibodies used

His-Tag (H-3; Lot # H2621; mouse monoclonal IgG1; 200 µg/ml; sc-8036; Santa Cruz Biotechnology), anti mouse IgG (peroxidase antibody produced in rabbit; buffered solution; Lot # 089M4797V; A9044-2ML; Sigma Aldrich)

### Validation

<https://www.scbt.com/de/p/his-probe-antibody-h-3>  
<https://www.sigmaaldrich.com/DE/en/product/sigma/a9044>