Post-translational mutagenesis for installation of natural and unnatural amino acid side chains into recombinant proteins

Tom H Wright & Benjamin G Davis

Department of Chemistry, University of Oxford, Oxford, UK. Correspondence should be addressed to B.G.D. (ben.davis@chem.ox.ac.uk).

Published online 29 September 2017; doi:10.1038/nprot.2017.087

Methods for installing natural and unnatural amino acids and their modifications into proteins in a benign and precise manner are highly sought-after in protein science. Here we describe a protocol for 'post-translational mutagenesis' that enables the programmed installation of protein side chains through the use of rapid, mild and operationally simple free-radical chemistry performed on recombinantly expressed and purified proteins. By introduction of protein dehydroalanine (Dha) residues (in this instance, from a unique cysteine residue introduced by site-directed mutagenesis) as free-radical trapping 'tags' for downstream modification, exquisite control over the site of subsequent modification is achieved. Using readily available alkyl halide precursors and simple borohydride salts, alkyl radicals can be generated in aqueous solution. These alkyl radicals react rapidly with protein-bound Dha residues to yield functionalized protein products with new carbon–carbon bonds. Once the Dha is installed, the introduction of the desired functionality is limited only by the requirement for polarity matching of the generated radical with the Dha 'acceptor', the solubility of the alkyl halide precursors in aqueous solution and the kinetics of the reaction itself. For example, methylated derivatives of lysine, arginine and glutamine are readily accessible. Furthermore, as the side chains are constructed chemically, many unnatural modifications can also be directly introduced as part of the side chain, including isotope reporters (19F, 13C) that can be used in biophysical experiments such as protein NMR. From a suitable cysteine mutant of the target protein, the entire procedure for this chemical post-translational mutation takes 2 d and is readily performed by nonchemists.

INTRODUCTION

Post-translational modifications (PTMs) expand the chemical diversity available for protein function beyond ribosomally processed amino acids[1]. The installation, removal and recognition of protein PTMs appear to underlie many of the central processes in biology. Yet the precise functional consequences of many PTMs are only poorly understood, as the availability of homogeneous samples of site-specifically modified proteins is still limited. A straightforward protocol would be to install these modified residues in a manner analogous to the pre-expression site-directed gene mutagenesis of canonical amino acids. This would enable the use of, for example, intact, more relevant, fully structured modified proteins as substrates in biological assays that currently use only peptide substrates (such as in *in vitro* binding assays, structural analyses or proteomic 'pull-down' technologies to identify interacting partners).

Chemical total or semisynthesis of modified proteins and codon reassignment approaches have emerged as useful tools for access to proteins incorporating site-specific, defined PTMs and unnatural, 'designer' amino acids. The stepwise chemical synthesis of proteins enables precise control of the PTM site and chemical identity via linear assembly of the final polypeptide chain. However, as the routine synthesis of peptides is still limited to 40–50 aa, assembly of whole proteins involves the synthesis of multiple peptide fragments, followed by chemical ligation and final refolding[2,3]. These are nontrivial procedures involving a substantial investment in time and expertise not available to all laboratories.

Purely biological routes to modified proteins offer the advantage of requiring less chemical expertise. Amber codon suppression enables the programmed introduction of noncanonical amino acids into proteins in response to a reassigned stop codon, via evolution of an orthogonal aminoacyl-tRNA synthetase/tRNA pair specific to the amino acid of interest[4,5]. However, despite impressive progress, the range of biologically relevant PTMs that have been incorporated remains limited. This is probably due to the difficulty in generating sufficient selectivity or tolerance in synthetase enzymes for either very small (methylation) or very large (glycosylation) modifications in amino acid substrates. Furthermore, expression yields and fidelity of the mutant proteins generated this way can show strong dependence on the amino acid incorporated, the synthetase system used for incorporation and even the site and sequence of the target protein itself[6].

A third route to modified proteins, the 'tag and modify' approach, combines the genetic introduction of a uniquely reactive 'tag' amino acid and its subsequent chemical modification[7]. Tags may be reactive natural amino acids such as cysteine or unnatural handles with reactivities (e.g., azide, alkyne)[8] not typically found in nature. The use of chemistry that is selective for the introduced tag results in site-specific chemoselective modification, avoiding the production of heterogeneous mixtures. Dha, in particular, is a readily accessible protein electrophile that possesses chemical reactivity distinct from that of the canonical amino acids[9,10]. Dha can be installed in a site-specific manner into polypeptides through alkylation/elimination of unique cysteine or selenocysteine residues, by chemical elimination of phosphoserine or phenyl- or alkyl-modified selenocysteine under basic conditions or through the use of enzyme-mediated side-chain elimination. Our group and others[11–13] have demonstrated the use of Dha as a tag for the introduction of PTMs and their mimics via operationally simple thia-Michael (carbon–sulfur β,γ bond formation) chemistry using thiol-functionalized reagents[14].

We have recently reported a new approach, termed 'post-translational mutagenesis'[15,16], that extends and generalizes the

'tag-and-modify' approach by using the carbon–carbon (β–γ) bond-forming reaction of alkyl radicals with Dha-tagged proteins to chemically 'construct' or 'bolt-on' the side chain of interest. This exploits the radical acceptor reactivity of Dha. As the β,γ carbon–carbon bond is naturally present in the side chains of many amino acids, this chemistry enables facile access to these residues and their PTMs in a programmed manner. Notably, as our strategy relies on *in vitro* chemistry to effect the key chemical mutation, it is not constrained by the 'evolutionary filter' of the cell's translational machinery, allowing installation of side chains that have not yet been incorporated by codon reassignment (such as methylated arginine and glutamine). In total, this approach allowed programmed installation of >25 different natural, modified and unnatural side chains from a single Dha 'tag' residue into a variety of protein scaffolds[15].

Overview of the procedure

The protocol presented here describes two facile and chemoselective chemical modification reactions that, starting from a free cysteine residue, together enable a formal chemical mutation to the desired PTM or designer amino acid (**Fig. 1**). The first step describes the conversion of cysteine to Dha via a mild one-pot alkylation and elimination procedure (Steps 1–8). The resulting Dha-tagged protein is then used as a substrate for modification with alkyl halide reagents, with sodium borohydride mediating the radical alkylation that installs the desired side chain (Steps 9–21). The polarity of the radical generated from the corresponding halide is a key consideration for selection of such 'side-chain' donors, with 'nucleophilic'-type radical reagents better matched to the electrophilic Dha center. Notably, tertiary, secondary and primary alkyl halides are all readily coupled, including those with unprotected, biologically relevant hydroxyl and amine functionality[15]. In this protocol, our approach is illustrated by two examples—installation of dimethyl-lysine at site 9 of histone H3 and installation of the phosphatase-resistant difluoromethylene analog of phosphoserine (Pfa)[17], at site 10 of histone H3 (**Fig. 1**). Both of these sites are biologically relevant and are naturally occupied by either modified or unmodified variants of K9 and S10, respectively.

Applications of the method

The downstream uses of the thus-modified proteins are varied and include both applied and fundamental applications. Lack of access to the requisite modified proteins as substrates or reagents has greatly hindered *in vitro* biochemical assays of PTM function, including biophysical studies such as protein NMR, which are further enabled by the ready incorporation of the isotopic labels demonstrated here. Although cellular studies of chemically modified protein conjugates are limited by the requirement for techniques for localizing proteins to the cellular compartment of interest (such as microinjection), we have shown the use of modified proteins as affinity reagents for proteome-wide assay of PTM binding partners[15]. In a biomedical context, we have previously demonstrated the modification of a single-chain mammalian antibody, including modification with complex glycosylated structures[15], highlighting potential applications in antibody-mediated cellular and *in vivo* imaging[18] and antibody–drug conjugates. Fundamentally, by extending the 'side-chain space' accessible to protein function, we envision wide-ranging future applications of our method, including in relatively unexplored areas such as modulation of enzyme catalysis by unnatural amino acids.

Limitations of the method

Our protocol is, of course, not without limitations. The most severe of these is the requirement for Dha as an orthogonal 'handle' for modification. Although it provides the requisite chemoselectivity[15] for the subsequent radical modification, the requirement for introduction of Dha limits the method to protein targets in which this residue can be installed in a site-specific manner. Interest in Dha as a useful 'tag' for chemical modification has led to a number of methods for its site-specific installation into expressed proteins[11,19,20]; although the chemical method described here is the most straightforward, it is not necessarily the only or the ideal route to consider. For both installation of Dha from cysteine and its subsequent radical modification, a degree of reagent (and hence solvent) accessibility is required for the targeted site, which can be determined empirically. That said, accessibility may be increased even for buried sites by the use of denaturing (or partially denaturing) conditions; if the latter approach is adopted, then secondary structure determination (e.g., through circular dichroism) and/or a protein refolding protocol, if necessary, should be considered. For proteins containing many free cysteine residues, for example, alternate routes to Dha based on ribosomal incorporation of noncanonical amino acids and subsequent chemical modification[19] may be used to circumvent the requirement for chemical modification of a specific cysteine.

In conjunction with other approaches, we believe that protocols for such benign protein side-chain manipulation by chemical mutation will enable a new approach to direct 'protein editing' to supplement and extend genetic manipulations.

Experimental design

Choice of modification strategy. Our protocol involves recombinant expression of a suitably designed cysteine mutant of the protein of interest, followed by two facile and chemoselective chemical modification reactions that together enable a formal 'chemical mutation' of the desired PTM or designer amino acid (**Fig. 1**). The protocol is dependent on the introduction of the noncanonical amino acid Dha as a 'tag' or uniquely 'radical-reactive' handle that guides the subsequent key chemical modification that constructs the target side chain. Dha can be introduced into proteins via a number of routes, typically from free cysteine[10], phosphoserine[19] or phenyselenocysteine[11,20]. Of these routes, one of the most straightforward is genetic incorporation of cysteine at the desired site of modification (using standard site-directed mutagenesis), followed by one-pot chemical bisalkylation and elimination to yield Dha. Undesired native cysteines can be transformed by mutagenesis to near-isosteric Ser if needed; although this is a conservative mutation, this procedure will require validation of the physical and biological effects of such a change in the resulting construct. Chemical conversion of phosphoserine and phenylselenocysteine to Dha requires highly basic or oxidative conditions that may not be compatible with sensitive protein substrates. Furthermore, as these latter precursors are nonribosomal amino acids, additional genetic engineering steps are required to incorporate these residues by codon reassignment, which can markedly affect yields of expressed protein. However, in cases in which a target protein contains essential free cysteines, the use of alternative precursor residues might, in principle, enable chemoselective (or potentially sequential) Dha formation. Notably, the key modification of Dha by alkyl radicals is possible in the presence of free cysteine residues[15], potentially enabling further subsequent

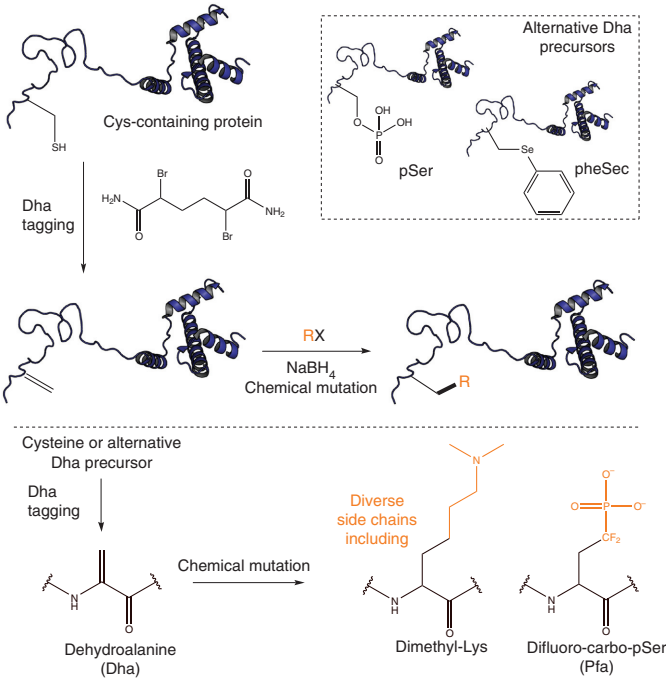

Figure 1 | General reaction scheme for the two-step chemical mutation from Cys→Dha→R. The reaction scheme is exemplified here by the installation of dimethyllysine and difluoro-carbo-pSer (Pfa). The structures of the modified residues are shown, with the segment derived from the alkyl iodide side chain donor highlighted in orange. The box highlights phosphoserine and phenylselenocysteine, alternative Dha precursors that can be installed biosynthetically via stop-codon suppression techniques.

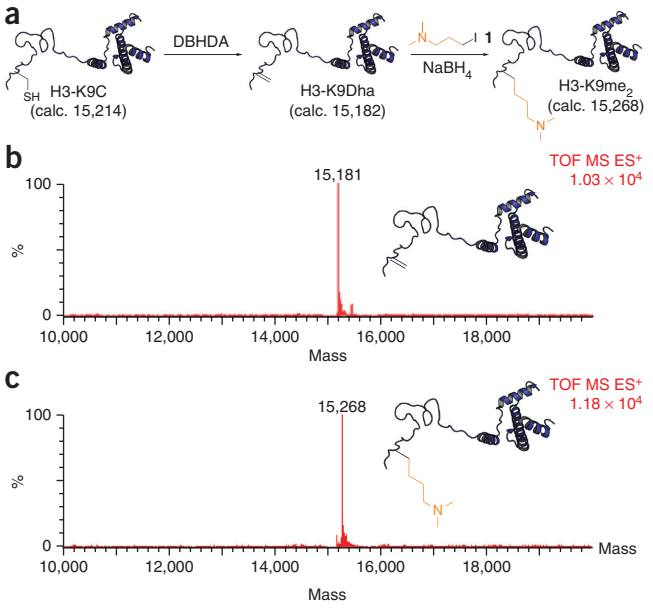

Figure 2 | Installation of dimethyllysine by reaction of histone–Dha. Dimethyllysine is installed from Dha with 3-(iodopropyl)dimethylamine hydroiodide (**1**) mediated by sodium borohydride. (**a**) Two-step reaction scheme. (**b**,**c**) Mass spectra showing deconvoluted spectra of H3-K9Dha (calc. 15,182 Da, obs. 15,181 Da) (**b**) and H3-K9me$_2$ (calc. 15,268 Da, obs. 15,268 Da) (**c**). Full spectra, including multiply charged ion series, can be found in **Supplementary Figure 1**. calc., calculated; DBHDA, 2,5-dibromohexanediacetamide; obs., observed.

chemical mutations. In this work, cysteine introduction to histone proteins is performed by site-directed mutagenesis according to the manufacturer's instructions. Expression and purification of histone proteins is a robust procedure outlined elsewhere[21].

Relative reactivity: protein. Both Dha formation and subsequent radical modification have been demonstrated on a number of protein sites and scaffolds, under both 'native' and denaturing conditions[11]. When using folded proteins as substrates for chemical reactions, a critical variable that can affect rates and even the possibility of reaction is the reactivity of the targeted amino acid in the context of the protein tertiary structure. Solvent accessibility provides one crude estimate that can correlate with reactivity (a form of 'reactive accessibility'). In the case of multiple cysteines, relative reactivity can be gauged (for example, by titration with Ellman's reagent or iodoacetamide and tryptic peptide mapping) to enable regioselective Dha formation on folded proteins. Undesired reactive cysteines can be 'mutated out' (e.g., to Ser) to further guide the chemoselectivity of the Dha introduction[22]. Alternatively, other methods of Dha introduction may be considered (see above). For both of the reactions discussed in this protocol, if a particularly hindered site is targeted, partial or full denaturation of the protein may simplify chemical modification, in which case consideration should be given to suitable procedures for refolding the protein following reaction.

Relative reactivity: alkyl halide. In this protocol, the installation of dimethyllysine, via radical generation from 3-(iodopropyl) dimethylamine hydroiodide, and the installation of Pfa, via radical generation from (bromodifluoromethyl)phosphonic acid, are highlighted. Reactivity of alkyl halide 'side-chain donor' reagents follows the expected trend based on the radical mechanism: tertiary >secondary >primary. Primary alkyl halides are still competent in the reaction and, depending on aqueous solubility, may be very reactive. Iodides are preferred, although in certain cases the bromide congener may prove useful if the iodide is very reactive and results in high levels of di-alkylation side products (see below). In general, salts are preferred because of their greater water solubility, although a number of commercially available short-chain alkyl iodides (e.g., isopropyl, tert-butyl iodide) can be used successfully as neat liquids[15,19]. As radical polarity effects may dominate reactivity, the most useful donors generate nucleophilic radicals following halogen abstraction, which can be productively trapped by protein Dha residues[15]. For a given alkyl halide, reactivity can be readily determined by initial 'titration' (in a series of reactions) of the molar equivalents of the reagent in the presence of fixed concentrations of protein–Dha and sodium borohydride, and adjusted accordingly for full conversion to the desired monoalkylation product.

Reaction scale. The post-translational mutagenesis procedure can be performed on a variety of scales, with our laboratory having successfully performed the key radical C–C bond formation on individual protein samples ranging from 50 μg to 10 mg. In all cases examined, a higher working concentration of protein results in a more efficient reaction. The examples reported in this protocol make use of 200 μg of histone–Dha in each reaction, with a working concentration of ~1 mg/ml (66 μM for the H3-Dha used).

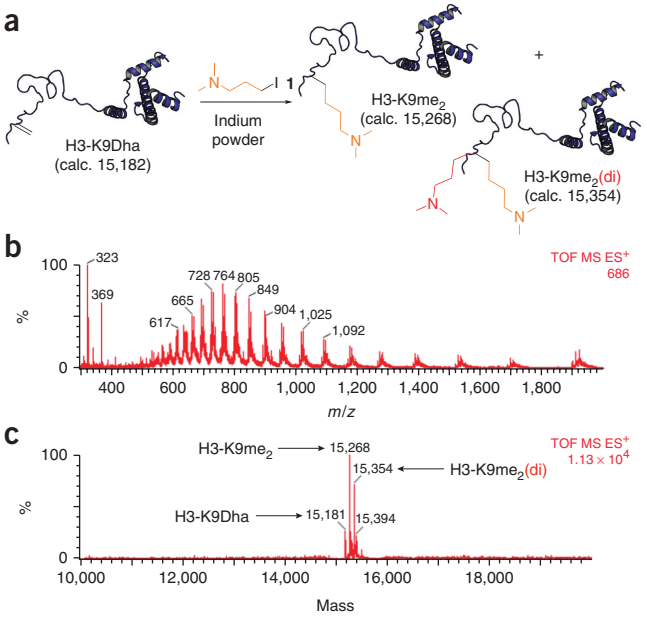

Figure 3 | Unsuccessful installation of dimethyllysine by reaction of histone–Dha. Dimethyllysine is installed with 3-(iodopropyl)dimethylamine hydroiodide (**1**) mediated by indium powder. Similar results are obtained with zinc powder and related metal-initiated systems. In addition to the desired $K9me_2$ product, a di-alkylated side product is also observed. See 'Experimental design' and 'Troubleshooting' sections for details. (**a**) Reaction scheme. (**b**,**c**) Mass spectra showing representative multiply charged ion series (**b**) and deconvoluted spectrum (**c**) of H3-$K9me_2$ (calc. 15,268 Da, obs. 15,268 Da) and H3-$K9me_2$(di) (calc. 15,354 Da, obs. 15,354 Da).

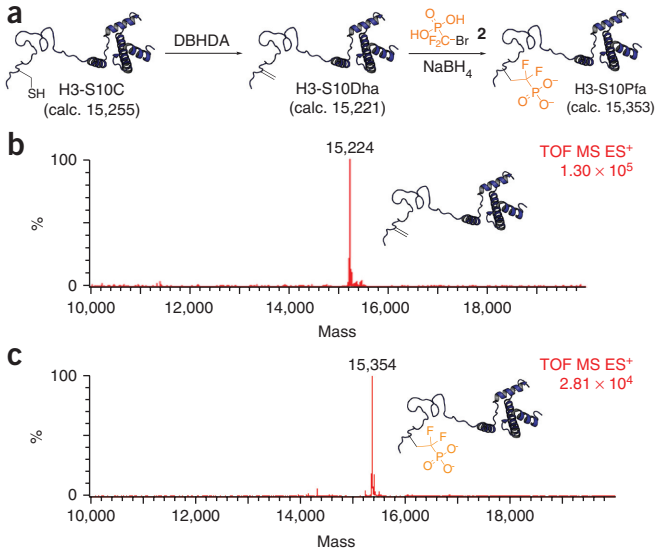

Figure 4 | Installation of difluoro-carbo-pSer (Pfa) by reaction of histone–Dha. Pfa is installed from Dha with (bromodifluoromethyl)phosphonic acid (**2**) mediated by sodium borohydride. (**a**) Two-step reaction scheme. (**b**,**c**) Mass spectra showing representative deconvoluted spectra of (**b**) H3-S10Dha (calc. 15,221 Da, obs. 15,224 Da) and (**c**) H3-S10Pfa (calc. 15,353 Da, obs. 15,354 Da). Full spectra, including multiply charged ion series can be found in **Supplementary Figure 2**.

Reaction monitoring. We recommend the use of LC–MS for monitoring of reactions on protein substrates. SDS–PAGE analysis typically does not have the requisite resolution for cleanly distinguishing the <300 Da mass shifts typical of side-chain modification during chemical mutation. We typically use a liquid chromatography–electrospray–time-of-flight (LC–ES–TOF) MS instrument in positive mode equipped with a short protein desalting column (ProSwift or Chromolith). Protein content can be observed by UV-visible plus total-ion-count (TIC) chromatograms. It is important that to be truly representative during such monitoring, all protein peaks are utilized (integrated to observe protein ion series) for subsequent deconvolution methods yielding total protein masses (**Figs. 2–4**). Typically, variation of ionizability is small between product and reactant proteins, thereby allowing near-quantitative use of ion count chromatograms; such quantification is readily verified by plotting concentration versus TIC.

Side reactions. Dha is readily formed from free cysteines by bis-alkylation with the dibromide compound 2,5-dibromohexanediacetamide (DBHDA). As mildly basic conditions (pH 8.0 or higher) promote this reaction, alkylation at nontarget nucleophilic residues, such as lysines, is a potentially possible but rarely observed side reaction that could be avoided, if seen, by optimization of buffer pH or reagent amount for cysteine selectivity. More generally, temperature, reagent concentration and reaction time can all be optimized to effectively eliminate competing side reactions. Radical modification is largely insensitive to pH, enabling the use of neutral or slightly acidic conditions that disfavor possible alkylation side reactions with the alkyl halides used as radical precursors. However, radical-mediated dual modification ('di-alkylation') at Dha is a competing side reaction that is inherent to the reaction mechanism and may be observed as a 'second addition' of the expected side-chain mass (**Fig. 3**; ref. 15), especially with methods proposed to generate radicals via metals (e.g., Zn(0) or In(0)). Use of borohydride and, if needed, titration of the side-chain precursor reagent concentration while fixing the molar equivalents of borohydride relative to protein should reveal conditions in which monoaddition is the sole product. Under anaerobic conditions, the only other side reaction potentially observed is reduction of Dha to alanine; although rare, this may be controlled, typically through temperature variation (e.g., cooling). Under aerobic conditions, oxidative protein cleavage at the introduced Dha is observed as a side reaction[15]. Although it may prove possible to perform size-exclusion chromatography or HPLC to remove unwanted cleaved fragments, yields of the desired protein are inevitably lowered. We recommend the use of a positive-pressure inert atmosphere glovebox or similar low-oxygen environment.

MATERIALS

REAGENTS

▲ **CRITICAL** Full procedures for all noncommercially available reagents can be found in the supplementary materials of Wright *et al.*[15]: http://www.sciencemag.org/cgi/content/full/science.aag1465/DC1.

- Protein of interest containing a unique cysteine. In this protocol, we use cysteine-modified histone H3 as an example, which is produced and purified as described in ref. 15.
- DL-DTT (Thermo Fisher Scientific, cat. no. R0861)

- 2,5-Dibromohexane diacetamide ((DBHDA) is readily synthesized from adipic acid (Sigma-Aldrich, cat. no. A26357)[14]. This reagent is also available to the community through our laboratory and can be stored at 15 °C for many years as a bench-stable solid.)
  ! CAUTION Toxicity information for this compound is not yet available: assume that it is toxic. Always handle the compound with gloves and work in an appropriately ventilated environment.
- Sodium borohydride (Sigma-Aldrich, cat. no. 213462)
- 3-(Iodopropyl)dimethylamine hydroiodide (readily synthesized in a single step from commercial starting materials per Wright *et al.*[15]. Solid crystals or powder thus obtained can be stored at 15 °C for many years as a bench-stable solid.)
  ! CAUTION Toxicity information for this compound is not yet available: assume that it is toxic. Always handle the compound with gloves and work in an appropriately ventilated environment.
- (Bromodifluoromethyl)phosphonic acid (readily synthesized by refluxing diethyl (bromodifluoromethyl)phosphonate (Sigma-Aldrich, cat. no. 411361) overnight in concentrated HCl (Sigma-Aldrich, cat. no. 320331), per Wright *et al.*[15]. The oil thus obtained should be stored at −20 °C and should be stable for months at this temperature.)
  ! CAUTION Toxicity information for this compound is not yet available: assume that it is toxic. Always handle it with gloves and work in an appropriately ventilated environment.
- Pierce formic acid, LC–MS grade (Thermo Fisher Scientific, cat. no. 85178)
- Water (deionized water: (>15 MΩ/cm resistance), filtered through a 0.2-µM disc filter)
- Acetonitrile, Optima, (LC–MS grade (Fisher chemical, cat. no. A955)
- *N*, *N*-Dimethylformamide (DMF; Sigma-Aldrich, cat. no. 437573)

EQUIPMENT

- PD MiniTrap G-25 (GE Healthcare Life Sciences, cat. no. 28-9180-07)
- Vivaspin 500 centrifugal concentrator, MWCO 5 kDa (GE Healthcare Life Sciences, cat. no. 28932223)
- SpinTrap G-25 desalting column (GE Healthcare Life Sciences, cat. no. 28-9180-04)
- Positive-pressure inert gas glovebox (We use a Belle Technology glovebox (http://www.belletechnology.co.uk/glovebox.php) equipped with the BASF R3-11G catalyst)
- LC–MS system (Waters LCT Classic coupled to a Shimadzu 20 Series HPLC system using a Thermo Proswift (250 × 4.6 mm × 5 µm) column)
- Standard biological pipette
- Eppendorf tubes
- Benchtop laboratory centrifuge
- Eppendorf cooler
- Size-exclusion column

REAGENT SETUP

Dha reaction buffer Dha reaction buffer is 50 mM sodium phosphate buffer (pH 8.0) supplemented with 3 M guanidine hydrochloride. Once made up, this buffer can be stored and used for up to a month at room temperature (15–25 °C).

Chemical mutation buffer The radical chemical mutation reaction is largely insensitive to the buffer selected. For work with histone substrates, we have found 500 mM ammonium acetate (pH 6.0) supplemented with 5 M guanidine hydrochloride to be a highly effective choice. Once made up, this buffer can be stored and used for up to a month at room temperature.

DBHDA Prepare a stock solution (0.5 M) of DBHDA by dissolving 100 mg in 610 µl of DMF. Stock solutions of DBHDA should be freshly prepared for each experiment **! CAUTION** Toxicity information for this compound is not yet available: assume that it is toxic. Always handle it with gloves and work in an appropriately ventilated environment.

3-(Iodopropyl)dimethylamine hydroiodide (1) Stock solutions (0.8 M) of this reagent should be freshly prepared on the day of the experiment **! CAUTION** Toxicity information for this compound is not yet available: assume that it is toxic. Always handle it with gloves and work in an appropriately ventilated environment.

(Bromodifluoromethyl)phosphonic acid (2) Stock solutions (1 M) of this reagent should be freshly prepared on the day of the experiment **! CAUTION** Toxicity information for this compound is not yet available: assume that it is toxic. Always handle it with gloves and work in an appropriately ventilated environment.

LC–MS solvent A LC–MS solvent A is 0.1% (vol/vol) formic acid in water.

LC–MS solvent B LC–MS solvent B is 0.1% (vol/vol) formic acid and 95% (vol/vol) acetonitrile in water. **▲ CRITICAL** For best results, freshly prepare all HPLC solvents and purge the LC–MS system before analysis of protein samples.

EQUIPMENT SETUP

Liquid chromatography Water (solvent A) and 95% (vol/vol) acetonitrile (solvent B), each containing 0.1% (vol/vol) formic acid (Reagent Setup), are used as mobile phase for gradient elution. The gradient program is as shown below.

Time interval (min)	% (vol/vol) Solvent B
0	5
5	5
15	100
18	100

Mass spectrometry analysis Mass spectrometry is performed using a Waters LCT Classic instrument, with parameters as displayed in the table below.

Method parameter	Value
Capillary voltage	3,000 V
Cone voltage	250 V

Nitrogen is used as the desolvation gas and nebulizer at a total flow of 600 liters/h.

Using Masslynx (Waters), mass spectra for the total protein content of the TIC chromatogram are combined to display and subsequently deconvolute the multiply charged ion series for all proteins present.

Spectra are calibrated by matching at least 17 of the multiply charged ion series of equine myoglobin run under the same gradient conditions.

PROCEDURE

Dha formation on histone substrates ● TIMING 5 h

1| Prepare a protein substrate solution by dissolving 12 mg of lyophilized powder of the relevant histone cysteine mutant (for dimethyllysine modification, H3-K9C; for 'Pfa' modification, H3-S10C) in 500 µl of Dha reaction buffer. Add 30 mg of DTT and vortex or sonicate the protein–DTT solution until all material is dissolved.

2| Incubate the solution for 30 min at room temperature to ensure complete reduction of the cysteine residue.

3| Carefully desalt the reduced protein by passing the protein solution through a PD MiniTrap G-25 column pre-equilibrated with Dha reaction buffer. After elution, the protein solution should be 1 ml, with a concentration of ~10 mg/ml.
▲ CRITICAL STEP This desalting step is essential to remove excess DTT. Excess DTT will react with and inhibit the Dha formed in the next step.

4| Initiate the formation of Dha by adding 100 µl of the DBHDA stock solution to the reduced protein. Gently shake the reaction at 25 °C and monitor every 30 min by removing a 5-µl aliquot using a standard biological pipette and dissolving it in 95 µl of Milli-Q H_2O (+0.1% (vol/vol) formic acid) for LC–MS.

5| When no protein starting material is visible by LC–MS within 30 min to 1 h, raise the temperature to 37 °C to promote the elimination reaction. Continue to monitor the reaction every 30 min by LC–MS.
? TROUBLESHOOTING

6| When the reaction is judged complete by LC–MS (See **Figs. 2** and **4** for examples of histone-Dha LC–MS spectra), concentrate the reaction mixture to 500 µl using a Vivaspin 500 concentrator (MWCO 5 kDa).

7| Desalt the sample by passing the concentrated solution through a PD MiniTrap G-25 column pre-equilibrated with chemical mutation buffer. Elute the sample with 1 ml of chemical mutation buffer to provide the Dha-tagged protein, ready for the next step.

8| Check the concentration of the protein by measuring the absorbance at 280 nm. With careful desalting, concentrations >80% of the starting material should be obtained.
■ **PAUSE POINT** The protein can be stored at –20 °C for up to 6 months.

Chemical mutation of histone proteins ● TIMING 8–24 h

9| Prepare the reagents for overnight 'degassing' in the glovebox. It is most convenient to prepare Dha-tagged protein samples for overnight incubation in the glovebox, for reaction the following morning. Prepare five or more Eppendorf tubes (in order to 'titrate' reagent equivalents for optimal reaction) containing 200-µl samples of Dha–protein solution at a concentration of 1 mg/ml (for a 15-kDa protein, ~65 µM), as determined by A280 measurement or bicinchoninic acid (BCA) assay.

10| Prepare a stock solution of the alkyl bromide or iodide 'side-chain donor' reagent. Dissolve the reagent in deionized water to a concentration of ~0.8 M for 3-(iodopropyl)dimethylamine hydroiodide (**1**) or 1 M for (bromodifluoromethyl)phosphonic acid (**2**). Vortex, if necessary, to obtain a homogeneous solution. Adjust the pH of the stock solution to ~6.0.

11| Carefully weigh out five or more (depending on the number of reactions required) 1-mg samples of solid sodium borohydride into 2-ml Eppendorf tubes. Quickly spin the tubes in a benchtop microcentrifuge (2200*g*, 30 s, room temperature) to ensure that the solid contents are at the bottom of the tubes. The tubes containing solid borohydride will be the final 'reaction vessel', so ensure legible labeling of these tubes before placing them into the glovebox.

12| Combine the protein, reagent and borohydride tubes on a sample rack and port into the glovebox, using rapid port pre-equilibration per the manufacturer's instructions (typically 10 min). It is essential that the caps of all Eppendorf tubes remain open throughout the degassing procedure. In the case of sensitive proteins, storage in an Eppendorf cooler kept at 4 °C within the glovebox is recommended. Note that this will slow the diffusion of O_2 from the protein solution, and hence extra time should be allotted for sample degassing. Robust or denatured protein solutions may be stored at room temperature overnight, within the glovebox.
▲ **CRITICAL STEP** It is essential that the protein solution be sufficiently degassed before chemical mutation is performed. On the basis of our experience, for a 200-µl buffered solution, this takes a minimum of 6 h at room temperature. Allow longer degassing times for larger-volume solutions and for proteins that require storage at reduced temperature. For example, for protein solutions stored at 4 °C within the glovebox, 12 h of degassing may be more appropriate.

13| Working within the glovebox, use a pipette to ascertain the volume of the protein and stock solutions, as the positive-pressure flow of inert gas within the glovebox will cause some evaporation (typically ~20 µl). 'Top up' the solutions as necessary using extra buffer or deionized water. Place the Eppendorf tubes containing solid sodium borohydride into an Eppendorf cooler kept at 4 °C (if available—room temperature reaction is also sufficient).

14| When performing a chemical mutation reaction for the first time on a given substrate/reagent pair, 'titrate' the amount of reagent required for optimal modification and protein recovery. Three protein samples are usually sufficient. Calculate 100, 500 and 2,000 molar equivalents of the bromide or iodide reagent relative to protein and determine the volume of reagent stock solution required.

15| Carefully pipette the required amounts of reagent stock into the relevant protein sample, and gently mix the resultant solution by pipette.

16| Initiate the reaction by pipetting the entire volume of the protein/reagent solution directly into the Eppendorf tube containing solid sodium borohydride and gently shake the tube. Effervescence, or foaming, is observed due to the hydrolysis of borohydride in an acidic medium.

17| Repeat Steps 15 and 16 for as many reactions as are being performed in parallel.

18| Incubate the reaction mixtures at 4 °C with the cap open. Shaking is not necessary but may help to maintain a homogeneous solution.

19| Following 30 min of reaction time, cap the tubes and remove them from the glovebox.

20| Analyze the reaction outcome by LC–MS. Aliquots of the protein reaction mixture can be directly used for LC–MS analysis. Dilute the reaction mixture 20-fold with H_2O containing 0.1% (vol/vol) formic acid and inject a small volume (e.g., 10 µl) onto the LC–MS system.
? TROUBLESHOOTING

21| (Optional) Depending on the downstream application envisioned and the behavior of the target protein, desalt the sample to remove excess reagent by loading the sample onto a SpinTrap G-25 desalting column, pre-equilibrated with the desired storage buffer. Elute the sample according to the manufacturer's instructions. For further downstream applications, such as biological assays, an additional size-exclusion column may be necessary for efficient sample desalting.

? TROUBLESHOOTING
Troubleshooting advice can be found in **Table 1**.

TABLE 1 | Troubleshooting table.

Step	Problem	Possible reason	Solution
5	No or partial conversion to histone–Dha observed by LC–MS	Incomplete reaction	(i) Add another 20 µl of the DBHDA stock solution and incubate the reaction mixture for 1 h at 37 °C before re-analysis by LC–MS (ii) Check whether the pH of the Dha buffer is >8.0 (iii) Reattempt the reaction with a fresh batch of DBHDA
20	No conversion to the desired product is observed by LC–MS, or protein products with a MW less than the starting protein are present	Reaction inhibition and oxidative protein cleavage due to the presence of O_2	(i) Repeat the reaction with an extended period of degassing. Ensure that all reaction components are thoroughly degassed (ii) Check glovebox function according to the manufacturer's instructions
	Some, but incomplete, conversion to the desired product	Incomplete reaction	(i) Repeat the reaction with fresh protein–Dha, using a higher number of molar equivalents for the bromo/iodo reagent (Note: Reduction of Dha to Ala will occur even if no conversion to the chemical mutation product is observed; therefore, it is not recommended to reattempt reaction on 'used' Dha.) (ii) Repeat the reaction with a higher number of molar equivalents of sodium borohydride
	n additions of the side-chain reagent to protein observed by LC–MS	Radical di-alkylation at Dha, or alkylation of protein nucleophiles	(i) Repeat the reaction with fresh protein Dha, using a reduced number of molar equivalents for the bromo/iodo reagent (ii) If >2 modifications are observed, check the reaction buffer pH and attempt the reaction at pH 6.0 or lower

● TIMING
Dha formation on histone substrates
Steps 1 and 2, preparation and incubation of protein substrate solution: 30 min
Step 3, desalting of the reduced protein: 15 min
Steps 4 and 5, formation of Dha: variable, typically 2–3 h
Steps 6–8, concentration and desalting of reaction mixture, and measurement of the protein: 1 h

Chemical mutation of histone substrates
Steps 9–12, preparation of reagents: 30 min; degassing of samples: 6 h to overnight
Steps 13–17, initiation of chemical mutation reaction: 10 min
Step 18, incubation: 30 min
Steps 19–21, LC–MS analysis and optional desalting: 30 min

ANTICIPATED RESULTS

Following the two reactions described above, proteins should be observed by LC–MS as a single species with >90% conversion to the desired chemical mutation (**Figs. 2** and **4**; **Supplementary Figs. 1** and **2**), as judged by calculation from peak intensities in the deconvoluted spectrum. No peaks below the expected starting material mass should be observed; if such peaks are visible, the protein may have undergone oxidative cleavage[15]. SDS–PAGE analysis can also be used to assess protein fragmentation if oxidative cleavage is suspected. Proteolytic digest and subsequent peptide mapping by LC–MS/MS should confirm the desired site of modification and that there are no additional, undesired modifications at other residues[15]. An unsuccessful reaction is shown in **Figure 3**; attempted installation of H3-K9me$_2$ using indium powder[15] in place of $NaBH_4$ resulted in substantial di-alkylation products observable in the mass spectrum.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS T.H.W. was supported by the Rutherford Foundation. B.G.D. was supported by a Royal Society Wolfson Research Merit Award.

AUTHOR CONTRIBUTIONS B.G.D. and T.H.W. designed the research. T.H.W. performed the experiments. T.H.W. and B.G.D. wrote the paper.

COMPETING FINANCIAL INTERESTS The authors declare competing financial interests: details are available in the online version of the paper.

1. Walsh, C.T., Garneau-Tsodikova, S. & Gatto, G.J. Protein posttranslational modifications: the chemistry of proteome diversifications. *Angew. Chem. Int. Ed. Engl.* **44**, 7342–7372 (2005).
2. Dawson, P.E., Muir, T.W., Clark-Lewis, I. & Kent, S.B. Synthesis of proteins by native chemical ligation. *Science* **266**, 776–776 (1994).
3. Malins, L.R. & Payne, R.J. Recent extensions to native chemical ligation for the chemical synthesis of peptides and proteins. *Curr. Opin. Chem. Biol.* **22**, 70–78 (2014).
4. Chin, J.W. Expanding and reprogramming the genetic code of cells and animals. *Annu. Rev. Biochem.* **83**, 379–408 (2014).
5. Wang, L., Brock, A., Herberich, B. & Schultz, P.G. Expanding the genetic code of *Escherichia coli*. *Science* **292**, 498–500 (2001).
6. Dumas, A., Lercher, L., Spicer, C.D. & Davis, B.G. Designing logical codon reassignment - expanding the chemistry in biology. *Chem. Sci.* **6**, 50–69 (2015).
7. Chalker, J.M., Bernardes, G.J.L. & Davis, B.G. A 'tag-and-modify' approach to site-selective protein modification. *Acc. Chem. Res.* **44**, 730–741 (2011).
8. van Kasteren, S.I. *et al.* Expanding the diversity of chemical protein modification allows post-translational mimicry. *Nature* **446**, 1105–1109 (2007).
9. Bernardes, G.J., 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).
10. Chalker, J.M. *et al.* Methods for converting cysteine to dehydroalanine on peptides and proteins. *Chem. Sci.* **2**, 1666–1676 (2011).
11. Levengood, M.R. & van der Donk, W.A. Dehydroalanine-containing peptides: preparation from phenylselenocysteine and utility in convergent ligation strategies. *Nat. Protoc.* **1**, 3001–3010 (2006).
12. Wang, J., Schiller, S.M. & Schultz, P.G. A biosynthetic route to dehydroalanine-containing proteins. *Angew. Chem. Int. Ed. Engl.* **46**, 6849–6851 (2007).
13. Guo, J., Wang, J., Lee, J.S. & Schultz, P.G. Site-specific incorporation of methyl- and acetyl-lysine analogues into recombinant proteins. *Angew. Chem. Int. Ed. Engl.* **47**, 6399–6401 (2008).
14. Chalker, J.M., Lercher, L., Rose, N.R., Schofield, C.J. & Davis, B.G. Conversion of cysteine into dehydroalanine enables access to synthetic histones bearing diverse post-translational modifications. *Angew. Chem. Int. Ed. Engl.* **124**, 1871–1875 (2012).
15. Wright, T.H. *et al.* Posttranslational mutagenesis: a chemical strategy for exploring protein side-chain diversity. *Science* **354**, aag1465 (2016).
16. Wright, T.H., Vallée, M.R.J. & Davis, B.G. From chemical mutagenesis to post-expression mutagenesis: a 50 year odyssey. *Angew. Chem. Int. Ed. Engl.* **55**, 5896–5903 (2016).
17. Zheng, W. *et al.* Cellular stability of serotonin N-acetyltransferase conferred by phosphonodifluoromethylene alanine (Pfa) substitution for Ser-205. *J. Biol. Chem.* **280**, 10462–10467 (2005).
18. Gunnoo, S.B. *et al.* Creation of a gated antibody as a conditionally functional synthetic protein. *Nat. Commun.* **5**, 4388 (2014).
19. Yang, A. *et al.* A chemical biology route to site-specific authentic protein modifications. *Science* **354**, 623–626 (2016).
20. Wang, J., Schiller, S.M. & Schultz, P.G. A biosynthetic route to dehydroalanine-containing proteins. *Angew. Chem. Int. Ed. Engl.* **119**, 6973–6975 (2007).
21. Luger, K., Rechsteiner, T.J. & Richmond, T.J. Expression and purificationof recombinant histones and nucleosome reconstitution. *Methods Mol. Biol.* **119**, 1–16 (1999).
22. Chooi, K.P. *et al.* Synthetic phosphorylation of p38α recapitulates protein kinase activity. *J. Am. Chem. Soc.* **136**, 1698–1701 (2014).