Light-Driven Posttranslational Installation of Reactive Protein Side Chains

Brian Josephson[1,5], Charlie Fehl[1,5]γ, Patrick G. Isenegger[1,5], Simon Nadal[1], Tom H. Wright[1]ϕ, Adeline W.J. Poh[1], Ben J. Bower[1], Andrew M. Giltrap[1], Lifu Chen[2], Christopher Batchelor-McAuley[2], Grace Roper[1], Oluwatobi Arisa[1], Jeroen B. I. Sap[1], Akane Kawamura,[1] Andrew Baldwin[1], Shabaz Mohammed[1,3], Richard G. Compton[2], Veronique Gouverneur[1,*] and Benjamin G. Davis[1,4]*

[1] Department of Chemistry, Chemistry Research Laboratory, Mansfield Road, Oxford, OX1 3TA, UK

[2] Department of Chemistry, Physical & Theoretical Chemistry Laboratory, South Parks Road, Oxford, OX1 3QZ, UK

[3] Department of Biochemistry, New Biochemistry Building, South Parks Road, Oxford, OX1 3QU, UK

[4] The Rosalind Franklin Institute, Oxfordshire, OX11 0FA, UK

[5] These authors contributed equally.

*To whom correspondence should be addressed: veronique.gouverneur@chem.ox.ac.uk, ben.davis@chem.ox.ac.uk, ben.davis@rfi.ac.uk

γ Current address: Department of Chemistry, Wayne State University, USA

ϕDepartment of Molecular Biology, Massachusetts General Hospital, USA

Summary Paragraph 276 words

Post-translational modifications (PTMs) greatly expand the structures and functions of proteins in nature.5,6 While synthetic protein functionalization strategies now allow mimicry of PTMs,7,8 as well as the formation of unnatural protein variants with diverse potential functions ranging from drug carrying,9 to tracking, imaging,10 and partner crosslinking11, the range of functional groups that can be introduced is still limited. Here, we describe the visible-light-driven installation of sidechains at dehydroalanine residues in proteins by the formation of carbon radicals which allow C-C bond formation in water. Control of reaction redox allows site-selective modification with good conversions and reduced protein damage. *In situ* generation of boronic acid catechol-ester (BACED) derivatives generates $RH_2C•$ radicals that form the native (βCH_2-γCH_2) linkage of natural residues and PTMs, while *in situ* potentiation of pyridylsulfonyl (pySOOF) derivatives by Fe(II) generates $RF_2C•$ radicals that form equivalent (βCH_2-γCF_2) linkages bearing difluoromethylene labels. These reactions are chemically-tolerant, incorporating an unprecedented range of functionality (>50 unique residues/sidechains) into diverse protein scaffolds and sites. Initiation can be applied chemoselectively in the presence of sensitive groups in the radical precursors, enabling installation of previously incompatible sidechains. This provides access to new function and reactivity in proteins, used here to (a) install radical precursors for homolytic *on-protein* radical-generation; (b) study enzyme function with natural, unnatural, and CF_2-labeled post-translationally modified protein substrates via simultaneous sensing of both chemo- and stereo-selectivity; and (c) create access to generalized ‘alkylator-proteins’ with a spectrum of heterolytic covalent-bond-forming activity (reacting diversely with small molecules at one extreme or selectively with protein targets through good mimicry at the other).

The resulting post-translational access to new reactions and chemical groups on proteins should prove useful in both revealing and creating protein function.

Introduction ~550 words

Methods that use the translational machinery of the cell provide some powerful advantages for installing select modifications into proteins, but can be limited in scope and efficiency.[12-14] Fed unnatural amino acid precursors can be degraded or may not be tolerated during biosynthesis; this is especially the case for those with reactive side-chains.[15] Post-translational functionalization[1-4,16] offers an alternative strategy that, through its late-stage use, could be potentially broader in scope. In principle, it is only limited by the compatibility of the reaction conditions used with the protein substrate and its context.

In one version[1-4,17] of post-translational functionalization, a readily-generated dehydroalanine (Dha) residue is used in proteins as a singly-occupied molecular orbital (SOMO) acceptor ('radical acceptor' or 'SOMO-phile') that is highly reactive towards several C• species thereby allowing selective β,γ-C-C bond formation to introduce new side-chains in a 'scarless / traceless' manner (**Extended Data Figure 1**). However, incompatibilities of side-chain/C• radical precursors and the reagents that generate them (e.g. single electron transfer (SET) from metals or BH_4^-)[3,4] currently limit scope. Nonetheless, such homolytic 1e- chemistry has potential advantages over typical heterolytic 2e- reagents. The intrinsic challenges[18] of biomolecule modification include: water-compatibility; requirement for a 'benign-ness'; and low (or non-) reactivity towards a plethora of biogenic acids, amines, alcohols, and thiols (ready 2e- reactants) present in most biological environments (**Extended Data Figure 1a**). By contrast, water and native proteins are less reactive to most[19] C• radicals –suitably placed SOMOphiles such as Dha can therefore allow more general chemo- and site-selectivity in certain 1e- chemistries (**Extended Data Figure 1b**).

Other methods for SET (and hence C• initiation, oxidative or reductive) exist. Catalytic protein methods bring clear advantages[20] over prior super-stoichiometric methods (which can drive unwanted side-reactions, **Extended Data Figure 1c-e**). Furthermore, if regulated by a relatively benign, potentially tissue-penetrating, trigger such as light,[21] it could allow additional layers of e.g. temporal, spatial and even kinetic control to complement those of $1e^-$ chemo-selectivity. Light-stimulated outer-sphere ET has seen a resurgence in applications to small molecules.[22,23] However, its use in site-selective, biomolecule modification has been more limited. Leading examples (**Extended Data Figure 1d,e** and **Supplementary Discussion 1**) have largely been restricted to peptides[24-26] sometimes requiring mixed organic solvents[26] and/or ET systems that sit towards the extremes of redox 'windows' and resulting side-reactions have been noted.[25] Moreover, dependence on certain precursor moieties, such as α-C-carboxyl[24] or β-C-H,[25] that cannot be re- / pre-positioned, can limit the site of reaction and/or lead to lower site-selectivity due to abundance. These methods have therefore yet to reach their full potential in protein chemistry.

Here, we show that a three-fold combination of: (i) ET at benign, moderate redox potentials using (ii) side-chain C• precursors 'redox-matched' with low, even substoichiometric, amounts of photocatalyst, triggered by (iii) light of appropriate flux, allows the generation and use of both *off-protein* and *on-protein* radicals to modify proteins via C-C bond formation (**Figure 1**).

Results ~2400 words

Photocatalytic C• Protein Modification

Exploitation of C• could use either *off*- or *on*- protein radicals with *reductive*- or *oxidative*- initiation. A pre-positioned *on-protein* SOMO-phile, such as Dha, allows the flexibility of *off-protein* C• generation by either method (**Figures 1, ED Figure 2**). Initial scoping with photocatalysts covering a wide redox spectrum (**ED Figure 2A**) under varying (aerobic/anaerobic, pH, co-solvent, redox mediators, light flux) aqueous reaction conditions avoided use of organic co-solvents or extremes of pH as these prove incompatible with many full-length proteins (see **Supplementary Discussion 1, Supplementary Tables 1-3** and **ED Figure 3a,b**). These revealed: a) an effectiveness of catechol beyond hydrogen atom transfer[27] (as noted previously for organosilicates[28]) in *oxidative* activation of alkyl organoboronates and b) relative ineffectiveness of alkylhalides as *reductive* precursors. Both suggested avenues for improvement.

Potentiation of (and activation of previously unreactive) alkyl boronates by catechol using low $*E_{ox}/*E_{red}$ catalysts was surprising, since $*E_{ox\text{-}catalyst} > E_{ox\text{-}substrate}$ typically determines reactivity. Primary C–B bonds ($\geq$ +1.5 V) were previously inaccessible; however, with catechol even challenging substrate $PhCH_2CH_2–BF_3K$ (E_{ox} > +1.6 V) proved reactive not only with **Cat3** ($*E_{ox}$ = +1.32 V), **Cat4** ($Ru(bpz)_3$, $*E_{ox}$ = +1.45V) but also much weaker catalysts **Cat1** ($Ru(bpy)_3$, $*E_{ox}$ = +0.77 V), **Cat2** ($Ru(bpm)_3$, $*E_{ox}$ = +0.99 V). These milder catalysts valuably gave enhanced conversions: e.g. >90% H3-homohomoPhe9 (H3-**1h**, see **Extended Data Figure 8** for side-chain glossary) from H3-Dha9.

Catechol's action was inconsistent with HAT[27] (**ED Figure 3a** and **Supplementary Table 4**). We tested three mechanistic possibilities (**ED Figure 4**

and **Supplementary Discussion 2**): mediated electron transfer; catalyst modification or substrate modification.[29,30] First, from analogues and potential redox mediators (**ED Figure 4a-c**) only aromatic 1,2-vicinal diols potentiated. Second, the Ru-complex[31] from ligand-exchange catecholo-Ru(bpy)$_2$-**Cat6** (**ED Figure 4d**) displayed no activity. Third, however, use of pre-formed boronic acid catechol esters revealed efficient conversion, even without exogenous catechol (**ED Figure 4e,f**). Moreover, cyclic voltammetry (**ED Figure 4g-l** and **ED Figure 2b inset**) revealed concentration-dependent, shifted boronate E_{ox} that brought these C• precursor substrates within range of mild catalysts **Cat1**, **Cat2** (see also **Supplementary Discussion 3**). Together, these data were consistent with *in situ* boronic acid catechol ester derivatives (termed here **BACED** reagents, **Figure 1, ED Figure 2c**) allowing efficient *oxidative* side-chain C• generation via turnover (even at 2mol%, **Supplementary Table 5**) of catechol in aqueous medium and lowered photocatalyst loadings (e.g. 100-equiv.→25 mol% of **Cat1**, **Supplementary Table 5**) to rare[20] substoichiometry.

Next, the failure in scoping of simple alkylhalides as *reductive* C•-precursors suggested substrate modulation. α-Fluoro substitution alters C•-radical stability and reactivity,[32] and increases addition reactivity in water.[33] H→F variation might therefore give enhanced RF_2C•/RFHC• as near-size[34] RH_2C• equivalents with the added potential to generate near-'zero-size' H→F labels in protein side-chains. From various potential RF_2C• precursors,[35] scoping (**Supplementary Discussion 4**) revealed pySOO–CF_2R pyridyl-sulfone-fluorides (pySOOF) reagents with suitable E_{red}[36] and enhanced reactivity (in 15 min). However, initially observed products were instead consistent with partial, unwanted oxidative (instead of reductive) quenching of resulting intermediate protein-α-C• to hemiaminal/imine (e.g. only 58% conversion

to H3-DfeGly9 from H3-Dha using pySOO-CF_2H, **ED Figure 2d**). Various reductive additives and HAT sources failed or led to other unwanted side reaction (**ED Figure 5**) but led us to consider direct ET from metals (**ED Figure 2d** and **ED Figure 6**); Zn(II), Mn(II) gave no change, Ni(II), Cu(II) inhibited reaction but Fe(II) sources (preferably iron(II)sulfate) drove good conversion (>92% to H3-DfeGly9, **ED Figure 6**). Notably, consistent with prior reactivity profiles,[32] mono-fluoro pySOOF reagent pySOO-CFH_2 failed to react, even under optimized conditions (**ED Figure 7**) yet pySOOF reagents bearing additional carboxylate or acetamide groups (pySOO-CHFCOOH and pySOO-CHFC(O)NH_2 generating sidechains **2s** and **2u**, respectively, **ED Figure 8**) displayed high reactivities (with only 10-25 equiv. pySOOF, **ED Figure 7**). Together, these and other (**ED Figure 7**) data were consistent with dual reductive quenching and potentiation by Fe(II) allowing efficient reductive side-chain C• generation and, again, rare substoichiometry, not only of photocatalyst but also Fe(II) (e.g. 25 mol% **Cat1**, 50 mol% $FeSO_4$, 10 equiv. pySOOF, 66% conversion, **Supplementary Tables 8-22** and **ED Figure 6**).

Optimization of BACED and pySOOF Reagents

Together, these mechanistic studies revealed mild, efficient complementary pathways using BACED or pySOOF reagents (**Figure 1**), potentiated by catechol or Fe(II), respectively. The wide availability of boronic acids (directly as BACED reagents) or pyridylsulfones (allowing ready synthesis of pySOOF reagents, see **Supplementary Methods**) enabled rapid, broad-scope optimization and application to introduce native and difluorinated amino-acid sidechains and/or those bearing post-translational modifications into proteins (**ED Figure 8**). Reaction times were shortened by investigating light flux in a variable-intensity photoreactor (**ED Figure**

3d); use of 50W at ~450nm shortened reaction times (4h → typically <1h for BACED; → < 15min for pySOOF, **Supplementary Tables 6,18** and **ED Figures 3d, 7**). Photocatalyst matching also enhanced efficiency: e.g. **Cat1** proved superior for secondary and benzyl BACED and pySOOF reagents (**ED Figure 3c, 6**) whilst **Cat2** better matched more difficult primary BACED reagents.

These allowed us to optimize efficient light-controlled protein modification to the following set of general conditions for BACED: protein 1 mg.mL^{-1}; 50W/450nm light; 4˚C-to-RT; 100-1500 equiv. BACED-precursor reagent; 10 equiv. **Cat1** or **Cat2**; 100 equiv. catechol; <6ppm O_2; pH 6.0 buffer (500 mM NH_4OAc or PBS ± 3M Gdn•HCl) and for pySOOF: protein 1 mg.mL^{-1}; 50W/450nm light; RT; 2-5 equiv. pySOOF-precursor reagent; 0.4-4 equiv. **Cat1**; 50-100 equiv. $FeSO_4$; <6ppm O_2; pH 6.0 buffer (500 mM NH_4OAc or various other buffers). Both reactions proved readily scalable to 5 mg protein levels (see **Supplementary Methods**). In all cases, control reactions under essentially identical conditions but in the absence of light or photocatalyst or catechol (for BACED) failed (**Supplementary Tables 7,16** and **ED Figure 7**). Sustained irradiation was also necessary, precluding chain mechanisms (**ED Figures 3e, 7**). It is striking that under these optimized conditions, even using C•-precursor as limiting reactant (instead of protein) good yields were still obtained: 0.5 equiv of pySOO-CF_2H gave 40% conversion (= 80% yield, **ED Figure 6**).

Diverse Side-Chains Inserted into Proteins

In this way, a diverse range of side-chains (**ED Figure 8**) were readily inserted through catalytic, light-driven, C•-mediated, C–C-bond-formation into a range of representative protein scaffolds [e.g. mixed disordered/alpha-helix-rich H3

and H4, the enzyme PanC (mixed alpha-beta 2/3-layered sandwich), Npβ (pentapeptide beta-sheet repeat), periplasmic protein AcrA (alpha-helix coiled-coil), single-chain-antibody cAbLys3 (beta-sheet CDRs)] and at different sites within scaffolds (e.g. H3-Dha4, -Dha9, -Dha18, and -Dha27) (**Figure 1**).

Many side-chains (**ED Figure 8**) were previously incompatible with some prior methods for protein modification,[3,4] including olefins (**1e,1f**), azides (**1r,2b,2ac**), halogens (**1s,1t,1y,1v,1w,2b,2ad**), sulfoxide/sulfones/sulfates (**2y,2z,2aa,2ae,2af**), esters (**1p,1q,2m,2n,2r**) and amides (peptidic (**1x**), biotinyl (**1y**,**2ag**), and also acyllysine derivatives (**1m,1n,1y,2k,2l,2m,2n,2w,2ag**).

For example, iodo-amides (precursors for acyllysines) are unstable with respect to cyclization[4,37,38] under current *reductive* methods, yet here, BACED and pySOOF reagents allowed ready, intact installation. Reductively-sensitive[3] esters (**1p,1q,2m,2n,2r**) or varied oxidation states of methionine (**2x,2y,2z**) were tolerated. Excellent compatibility and chemoselectivity was also seen for olefins **1e**, **1f** (well-known radical-'traps'[39]) and known photoredox-sensitive azides[40] (e.g. into azidonorleucine (Anl) **1r**,**2ab** (γCF_2-azidonorvaline) or **2ac** (γCF_2-Anl)) allowing potential for subsequent redox 'uncaging' or other azide reactions.

Such was chemoselectivity, that we next tested use of C•-precursors bearing *two* competing (geminal or distal) groups for potential initiation. First, remarkably, halo-pySOOF reagents pySOO–CF_2–Hal (**Figure 2a**) could be chemoselectively activated to generate pySOO-F_2C• radicals that allowed installation of mono- and difluoro-pySOOF sidechains (**2ae,2af**) into proteins (**Figure 2a**). Second, BACED and pySOOF reagents bearing two, *distal* initiator groups allowed installation of a diverse range side-chain halides (**Figure 2b**, F,Cl,Br,I), either with short alkyl linkers to protein backbone (**1s,1t,1u,2b,2ad**) or more complex substrate-mimetic side-

chains (**1v,1w**) – these were installed without activation of the radical precursor halide.

On-Protein Heterolytic Reactivity

Halides (**1s,1t,1u,2b,2ad**,**1v,1w**) are potentially powerful general 2e- electrophilic, side-chains and afforded a novel, on-protein heterolytic reaction platform for potential conjugation with off-protein nucleophiles. By varying pH, nucleophile concentration and halide, it proved possible to selectively facilitate intermolecular nucleophilic substitution at C-Hal (**Figure 2b**) thereby creating C–S, C–P and C–N bonds (creating e.g. methyllysine PTMs and N_3- giving Anl) and, even, direct Finkelstein-type, halogen-exchange (Br→Cl or I→Cl), allowing further tuning of electrophile reactivity. As such, this is therefore a direct strategic inversion of common practice in the field of protein conjugation (using prevalent *nucleophiles* in proteins Cys,Lys, etc. to target *off-protein* electrophiles).

On-Protein Homolytic Reactivity

Mono- and difluoro-pySOOF sidechains (**2ae,2af**) suggested immediate potential for generating *on-protein* side-chain radicals (**Figure 2a**). In agreement with reactivity of off-protein pySOOF, on-protein initiation of mono-fluoro-pySOOF sidechain **2ae** failed. However, when, histone-H3-pySOOF9 (containing di-fluoro-pySOOF sidechain **2af** at site 9) was activated in the absence of a co-reactant, reduced product (H3-DfeGly9) was cleanly generated (**Figure 2a** and **Supplementary Tables 23,24**), consistent with on-protein C•-initiation and Cγ–H bond-formation.

To test the scope of this on-protein C•-radical we targeted more challenging, Cγ–X bonds. First, Cγ–Cδ bonds were formed using various off-protein, C=C-radical-acceptors (**Figure 2a, bottom right** and **Supplementary Tables 25-34**): mono-/di-substituted polarized acceptors as well as challenging, tri-substituted and less-polarized, albeit with lower conversions. Indeed, such was the Cγ–Cδ-bond-formation efficiency that it allowed on-protein C–C-bond-forming side-chain *oligomerization/polymerization* [Cγ–Cδ–Cε– etc] – this could observed directly by intact-protein MS (e.g. mono/double/triple/quadruple side-chain addition unit growth for *N*-acetyldehydroalanine, **Figure 2a, inset to bottom right**). Furthermore, oligomerization could be modulated by monomer, Fe(II) and/or photocatalyst loadings (e.g. 25→125equiv. *N*-acetyldehydroalanine and 2→4equiv. of **Cat1** shifted oligomer (mono-:di-:tri-:tetra-) from 15:65:25:0→0:37:45:18) as well as by higher monomer reactivity (e.g. acrylamide gave up to side-chain-hexamer, see **Supplementary Table 34**). This demonstrated site-selective, side-chain C–C-synthesis even up to 14 carbons in length in a protein.

Second, we explored Cγ–heteroatom bonds: stable nitroxide allowed Cγ–Oδ-bond-formation (>90%, **Figure 2a, middle left, Supplementary Table 35**) and Se–Se-bond cleavage even allowed installation of difluorophenyl-SeMet (>90%, **Figure 2a, bottom centre** and **Supplementary Table 36**) through Cγ–Seδ-bond-formation.

Finally, we explored the potential of such *on-protein* radical chemistry in covalent, protein-protein-hetero-dimer formation, a challenging and rare application that to date has not, to our knowledge, been achieved by controlled C(sp3)-C(sp3) bond-formation. When H3-pySOOF9 (~15kDa) was initiated under photocatalytic conditions in the presence of equimolar, Dha-containing-protein histone-eH3.1-Dha9 (~18kDa) a cross-linked protein (~ 33 kDa) adduct was immediately observed

(**Figure 2a, bottom left** and **Supplementary Table 37**), consistent with heteromeric-cross-linking.

Probing of Post-translational Enzymes

Suitable side-chains allow potential PTM-mimicry (**Figure 3**), which was tested through installation of acetyl- (AcLys/KAc,**1m**) and benzoyl-lysine (BzLys/KBz,**1n**) side-chains (using BACED) as well as H→F-labeled lysine-analogues (K[γF_2]Ac,**2k** and K[γF_2],**2f**, using pySOOF). First, human histone eH3.1 proteins eH3.1-KAc18 and eH3.1-KBz18 enabled timecourse studies of NAD+-dependent deacylase Sirt2 that not only confirmed suggested[41] Sirt2 activity now on full-length proteins, but also revealed a strong, previously undetermined substrate KAc>KBz selectivity by Sirt2 in full-length protein for the first time [$k_{cat}/K_{M,app}$(eH3.1-K18Ac):$k_{cat}/K_{M,app}$(eH3.1-K18Bz) > 100:1] (**Figure 3a**).

As well as these constitutionally-native side chains, we also tested H→F labeled side-chains K[γF_2]Ac **2k** and K[γF_2] **2f** sidechains; the centrally-placed γ-carbon-F_2 label allowed *in situ* reporting of the modification state of these sidechains, in three ways (**ED Figure 9,10**): chemical/modification-state (e.g. +/- Ac), stereochemical-state (e.g. D-vs-L processing) or assembly-state (protein mono-/mulit-mer). Sidechain identity eH3.1-K18-*vs*-K18Ac was sensitively distinguished by ^{19}F-NMR-shift despite 5-bond distance from γ-carbon-F_2-label to site-of-change (**2f**→**2k**, δ_F -98.0→-99.4, **Figure 3b**), as were other sidechain variations H3.1-K9→KAc9→K_{me3}9, **2f**→**2k**→**2j**, H3.1-M9(**2x**), H3.1-E9(**2u**) (**ED Figure 9**). The diverse scope of further sidechains (**ED Figure 8**) would also allow such monitoring for heteroatom variation [N→O-'deaza-oxo'-H3.1-K_OAc(**2r**)] or precisely assaying sidechain Met-oxidation state [H3.1-M27→M_{ox}27→M_{ox2}27, **2x**→**2y**→**2z**].

Moreover, through additional correlated simulation of ^{19}F-NMR-multiplicity (**Figure 3b**), γ-F_2 label allowed *simultaneous in situ* on-protein reporting on both modification-state (KAc→K at Nε-site, 5-bonds-'*down'* sidechain) and, by virtue of highly-sensitive CF_2-diastereotopicity, stereo-chemical state (and hence processing-selectivity L-*vs*-D at Cα-site, 3-bonds-'*up'* sidechain). This α-γ-Nε-sensitivity along full-length of sidechain revealed HDAC-Sirt2-deacylation displays α-L>α-D selectivity preference at eH3.1-KAc18 of >14 ($\Delta\Delta G^{ø}$ >6.6 kJ mol^{-1}, **ED Figure 10**), despite six 6-bonds distance to Nε. To our knowledge, such simultaneous, real-time determinations of substrate- *and* stereo-selectivity in intact proteins has not been previously possible.

Finally, the sensitivity of γ-F_2-label could even be applied to monitor differential folding and assembly states in a single protein: full step-wise assembly[42] of H3-DfeGly9 histone into octamer [unfolded-H3-monomer→folded-H3-monomer→$(H3)_2$•$(H4)_2$-hetero-tetramer→$(H3)_2$•$(H4)_2$•$(H2A)_2$•$(H2B)_2$-hetero-octamer] even at μg scales (**ED Figure 10**).

Alkylator Mimic Proteins trap Buried PPIs

The electrophilic, halide side-chains included those with sidechain lengths well-matched to Lys [bromonorleucine (Bnl,**1t**), bromohomonorleucine (Bhn,**1u**) iodonorleucine (Inl,**1s**) **ED Figure 8**] that enabled design of 'protein alkylators' with potential context-dependent reactivity. If designed correctly these would remain inactive under typical conditions in a biological mixture but then display enhanced alkylative reactivity in a 'guided' manner by virtue of solvent exclusion, effective molarity[43,44] and proper mimicry when suitably engaged at a protein-protein-interface (PPI). Such a system would require a critical balance in electrophilic reactivity and

native shape fidelity (**ED Figure 11a**) that has been presciently highlighted as a key goal in protein science (see **Supplementary Discussion 5**).[18]

Site-selective insertion of minimally-sized, alkylhalide-side-chains Bnl(**1t**), Bhn(**1u**), Inl(**1s**) into proteins has not been previously possible. By more-closely mimicking the binding of Lys sidechains they allowed potential for probing, artefact-free, even buried PPIs. We tested potential buried[45] and transient (substrate•enzyme) PPIs using Bhn(**1u**) at three sites (4, 9, 27) normally occupied by Lys in C-terminally-FLAG-HA-tagged-histone-eH3.1. When incubated with a partner enzyme that processes (and so binds) Lys, Lys-demethylase-KDM4A, crosslinking exclusive to Bhn-containing eH3.1-Bhn4,eH3.1-Bhn9,eH3.1-Bhn27 was observed – but not with WT-histone-eH3.1 (**Figure 3c**). The Lys-'guided' nature of this cross-linking was consistent with zero-cross-linking from incubations with other proteins: neither Cys-rich serum albumin nor known nucleosomal binding-partner histone-H4 – H3•H4 PPI does not involve Lys4,9,27,[46] whereas only the H3•KDM4A PPI does.[45] This seeming PPI-selective reaction was further confirmed by MS/MS analysis (**Figure 3c** and **ED Figure 11c**) of crosslinking to KDM4A-Cys234,Cys306[45] located in the buried Zn-binding domain in H3•KDM4A PPI as well as by real-time fluorescent monitoring of Zn-ejection (**Figure 3c** and **ED Figure 11b**)[47] by eH3.1-Bhn9 but not WT-eH3.1. Moreover, when incubated with human-cell (HeLa) nuclear lysate, eH3.1-Bhn9 showed an ability to enhance capture of interaction partners via eH3.1-adduct formation (**ED Figure 11e**).

The unusual reactivity of these alkylator proteins was further illustrated by observation of an inter-protein Williamson-type-(-C–O–C-)ether formation that, to our knowledge, is wholly unprecedented[16] (**ED Figure 11c,d**) and has a typical rate seemingly too low ($k_{app} < 10^{-4}$ $M.s^{-1}$)[48,49] to allow effective cross-linking (see

Supplementary Discussion 6). This observation of *inter*-protein (**ED Figure 11c,d**) H3-tail•H3-tail Cβ–O–CH_2-Bhn4-trapping may suggest transient interactions; JmjN-domain found next to catalytic-JmjC-domain in KDM4A has recently been implicated as a functionally-essential dimerization motif in cellular models.[50] Together with suggested observations of KDM4A homo-multimers *in vitro*[51] this could suggest a model (**Extended Figure 11d**) where KDM4A multimers sample pre-existing H3•H3-dimers, and suggests the further potential of precisely-mimicking residues such as Bhn to suggest new speculative mechanistic models.

Discussion ~500 words

Our strategy for photoredox-initiation of C•-precursors displays chemoselectivity precise enough for broad installation of previously difficult, even reactive, side-chains. Limits to our process remain: H_3C• and Ar• precursors failed, perhaps due to higher redox barriers to initiation or rapid termination that outstrips addition; the pySOOF system however does allow use of HF_2C•. Notably, we have also observed first examples of the site-selective creation of *quaternary* sites in proteins through dual C•-addition (e.g. dual Bn• to give 4-benzyl,3-phenyl-Aib, BFab at ~80% (**ED Figure 3a**). Although, not the focus of this work, the ability to vary conditions to effect formation of such centres using longer-lived C• radicals with more potent photocatalysts (here **Cat3** > **Cat1**) to quench terminate an *on-protein* α-C• is an exciting additional observation. It suggests future design of *gem*,*gem*-α,α-bisalkylated motifs known[52] to control stability and conformation.

Our consequent discovery of new cross-links using previously inaccessible unnatural amino acids in proteins (e.g. Bhn) suggests that precise mimicry of residues in PPIs may drive new chemistries and hence selectivities. Whilst current cross-linking methods non-specifically 'fix' (e.g formaldehyde, *bis*-esters) and hence trap complexes that are longer-lived or more-favoured, our results here suggest the future application of crosslinks could allow exploitation of reactivity enhanced by effective molarity (EM, see **Supplementary Discussion 6**) to precisely cross-react minimal, size-matched sidechains (e.g. Bhn) even in transient, reactive interactions. Importantly, in this scenario it is the *relative* rate enhancement caused by environment that is important rather than the inherent rate of reactivity *per se*. We speculate therefore that better cross-linking selectivities may now be designed around new amino acids with surprising chemistries (e.g. Williamson-ether-

formation) with counterintuitively *lower* (not higher) reactivities that work in only rare but more information-rich and relevant contexts (e.g. insertion into precisely-matching PPIs that drive EM). Moreover, this has enabled complex, activity-based protein inhibition through the modification of conserved active site residues (e.g. covalent inhibition of KDM4A by eH3.1-Bhn9). In this way, one may consider future protein analogues – 'protein covalent inhibitors' – that act akin to emerging, targeted small-molecule covalent-inhibitors[53] yet with enhanced potency and selectivity that exploits PPIs.

Our observed requirement for continual irradiation in our light-driven approach enables on-off temporal control (**ED Figure 3e**, **7**). This suggests not only chemical precision in individual proteins but potential future use of defined spatiotemporal irradiation[21,54] to control states of protein ensembles over time and site (e.g. tissues). When also paired with *in cellulo* generation of Dha[55] the chemical side-chain versatility that we have now shown, could drastically improve our ability to probe complex biological systems with atomic precision using the tissue-penetrating trigger of light. To this future end we have also already shown that various biogenic catecholamines (**ED Figure 3f**) function in the BACED manifold. This, along with endogenous presence of Fe(II) for pySOOF reactivity, suggests promise for *in vivo* reactivity.

Data Availability Statement

Key raw MS data and primary numerical data for graphical plots are deposited in the open-access depository ORA-data (https://ora.ox.ac.uk/objects/uuid:2a618e7e-551b-4360-a2de-237453d49a31) and all raw data is available upon request. MS/MS raw data files have been uploaded to the PRIDE repository (accession number PXD019565). The following additional databases were used (with links): MaxQuant contaminants database (https://www.maxquant.org); Uniprot Human Database (https://www.uniprot.org/proteomes/UP000005640)

Author Contributions

C.F., B.J., P.G.I., T.H.W. V.G. and B.G.D. conceived and designed the experiments. T.H.W. designed and performed initial experiments exploring the oxidative initiation pathway. C.F. performed initial experiments exploring reductive pathways. C.F. optimized the initial photochemical boronate reaction. C.F. designed the high-flux visible light photoreactor. P.G.I. designed and performed all experiments for the use of pySOOF reagents.

C.F., B.J., B.J.B., O.A., P.G.I., A.M.G. and A.P. synthesized and characterized BACED substrates and catalysts; B.J., P.G.I. and B.J.B. expressed and generated protein starting materials.

B.J., C.F., P.G.I., A.M.G. and A.P. explored the scope of BACED substrate side chains and proteins; B.J. optimized additions of BACED reagents to proteins and explored additional protein scope. P.G.I. designed and performed all experiments exploring the pySOOF reagents with iron/$Ru(bpy)_3$ and so optimized the corresponding photochemical reaction; P.G.I. designed, synthesized and characterized pySOOF reagents and explored the scope of substrate side chains

and proteins. A.P. and A.M.G. synthesized additional pySOOF and tested them on proteins. J.B.I.S. synthesized additional pySOOF reagents.

B.J. developed the on-protein substitution of side-chain alkyl halides with small molecules; B.J. compared the reductive and oxidative initiation of model substrates, explored side reactions and effects of catalyst as well as methods for recycling of side-chain substrates materials from protein reactions.

B.J. and A.M.G. conducted UV/Vis-HPLC analysis of reaction products.

L.C. and C.B.M. made the electrochemical measurements on the basis of which they suggested mechanistic interpretation along with RGC.

P.G.I. designed and conducted all of the on-protein radical reactions and applications.

B.J. developed and conducted the on-protein MS enzymatic deacylation assays. B.J. P.G.I. designed and conducted Sirt2 deacetylation ^{19}F-NMR (and other) enzyme-tracking reactions; B.J., P.G.I., A.J.B., and B.G.D. analyzed the corresponding data.

B.J. and P.G.I. conducted and characterized ^{19}F-NMR tracking of histone octamer reconstitution.

G.R. and B.J. developed and conducted the zinc ejection assays.

B.J. and C.F. designed and performed protein cross-linking experiments; S.N. and S.M. performed LC-MS/MS experiments and cross-linked product analyses. B.J. developed and conducted lysate crosslinking immunoprecipitation experiments and analysis, S.N. provided HeLa nuclear extracts.

C.F., B.J., S.M., S.N., P.G.I., T.H.W., B.J.B., A.P., L.C., C.B.M., G.R., A.K., A.J.B. and B.G.D. collected and/or analyzed data.

C.F., B.J., P.G.I. and B.G.D. wrote the paper.

All authors read and commented on the paper.

Acknowledgements

This research has received funding from the EPSRC (EP/V011359/1), UK Catalysis Hub (EPSRC Portfolio Grant EP/K014668/1, B.G.D., C.F.), the Swiss National Science Foundation (P2BSP2_178609, P.G.I.), BBSRC (BB/P026311/1, B.G.D., V.G., P.G.I.), Oxford-GSK-Crick Chemical Biology Centre for Doctoral Training Programme (EPSRC, GSK to G.R.) via the EPSRC Systems Approaches to Biomedical Science DTC (EP/R512333/1), Oxford Clarendon Scholarship (to B.J.), Rutherford Foundation (to T.W), UCB (to B.B.), Brunei Government Scholarship (to. A.P.) and EU H2020 under the Grant Agreement no. 721902 (to O.A.). We thank Ms Svenja Hester for experimental support in mass spectrometry, Tim Mollner, Mateusz Imiołek for providing small molecule substrates, Sarah Faulkner for providing protein, the Chemistry Department workshop for construction of the photoreactor and Dr Christopher am Ende, Dr Will Stockdale and Prof Mathilda Moomersteeg for useful discussions.

Competing Interests

A patent is being filed that might afford authors royalties were it to be licensed.

References

1 Chalker, J. M. & Davis, B. G. Chemical mutagenesis: selective post-expression interconversion of protein amino acid residues. *Current Opinion in Chemical Biology* **14**, 781-789, doi:doi.org/10.1016/j.cbpa.2010.10.007 (2010).

2 Wright, T. H., Vallée, M. R. J. & Davis, B. G. From Chemical Mutagenesis to Post-Expression Mutagenesis: A 50 Year Odyssey. *Angewandte Chemie International Edition* **55**, 5896-5903, doi:doi:10.1002/anie.201509310 (2016).

3 Wright, T. H. *et al.* Posttranslational mutagenesis: A chemical strategy for exploring protein side-chain diversity. *Science*, aag1465, doi:10.1126/science.aag1465 (2016).

4 Yang, A. *et al.* A chemical biology route to site-specific authentic protein modifications. *Science*, aah4428, doi:10.1126/science.aah4428 (2016).

5 Walsh, C. T., Garneau-Tsodikova, S. & Gatto Jr., G. J. Protein Posttranslational Modifications: The Chemistry of Proteome Diversifications. *Angewandte Chemie International Edition* **44**, 7342-7372, doi:doi:10.1002/anie.200501023 (2005).

6 Deribe, Y. L., Pawson, T. & Dikic, I. Post-translational modifications in signal integration. *Nature Structural and Molecular Biology* **17**, 666, doi:10.1038/nsmb.1842 (2010).

7 Howard, C. J., Yu, R. R., Gardner, M. L., Shimko, J. C. & Ottesen, J. J. Chemical and biological tools for the preparation of modified histone proteins. *Topics in Current Chemistry* **363**, 193-226, doi:10.1007/128_2015_629 (2015).

8 Yang, A., Cho, K. & Park, H.-S. Chemical biology approaches for studying posttranslational modifications. *RNA Biology* **15**, 427-440, doi:10.1080/15476286.2017.1360468 (2017).

9 Ducry, L. & Stump, B. Antibody−Drug Conjugates: Linking Cytotoxic Payloads to Monoclonal Antibodies. *Bioconjugate Chemistry* **21**, 5-13, doi:10.1021/bc9002019 (2010).

10 Hinner, M. J. & Johnsson, K. How to obtain labeled proteins and what to do with them. *Current Opinion in Biotechnology* **21**, 766-776, doi:10.1016/j.copbio.2010.09.011 (2010).

11 Leitner, A. *et al.* Probing Native Protein Structures by Chemical Cross-linking, Mass Spectrometry, and Bioinformatics. *Molecular and Cellular Proteomics* **9**, 1634, doi:10.1074/mcp.R000001-MCP201 (2010).

12 Wang, L., Brock, A., Herberich, B. & Schultz, P. G. Expanding the Genetic Code of <em>Escherichia coli</em&gt. *Science* **292**, 498, doi:10.1126/science.1060077 (2001).

13 Dumas, A., Lercher, L., Spicer, C. D. & Davis, B. G. Designing logical codon reassignment - Expanding the chemistry in biology. *Chemical Science* **6**, 50-69, doi:10.1039/c4sc01534g (2015).

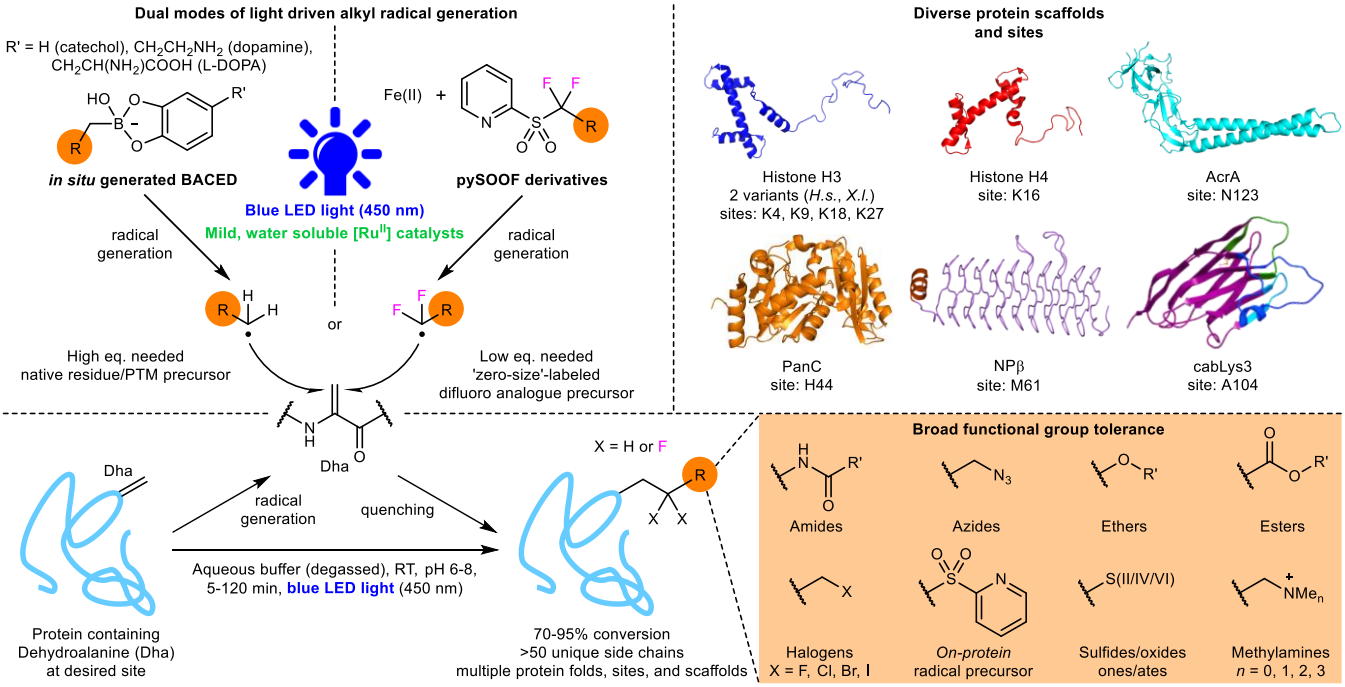
14 Chin, J. W. Expanding and reprogramming the genetic code. *Nature* **550**, 53-60, doi:10.1038/nature24031 (2017).

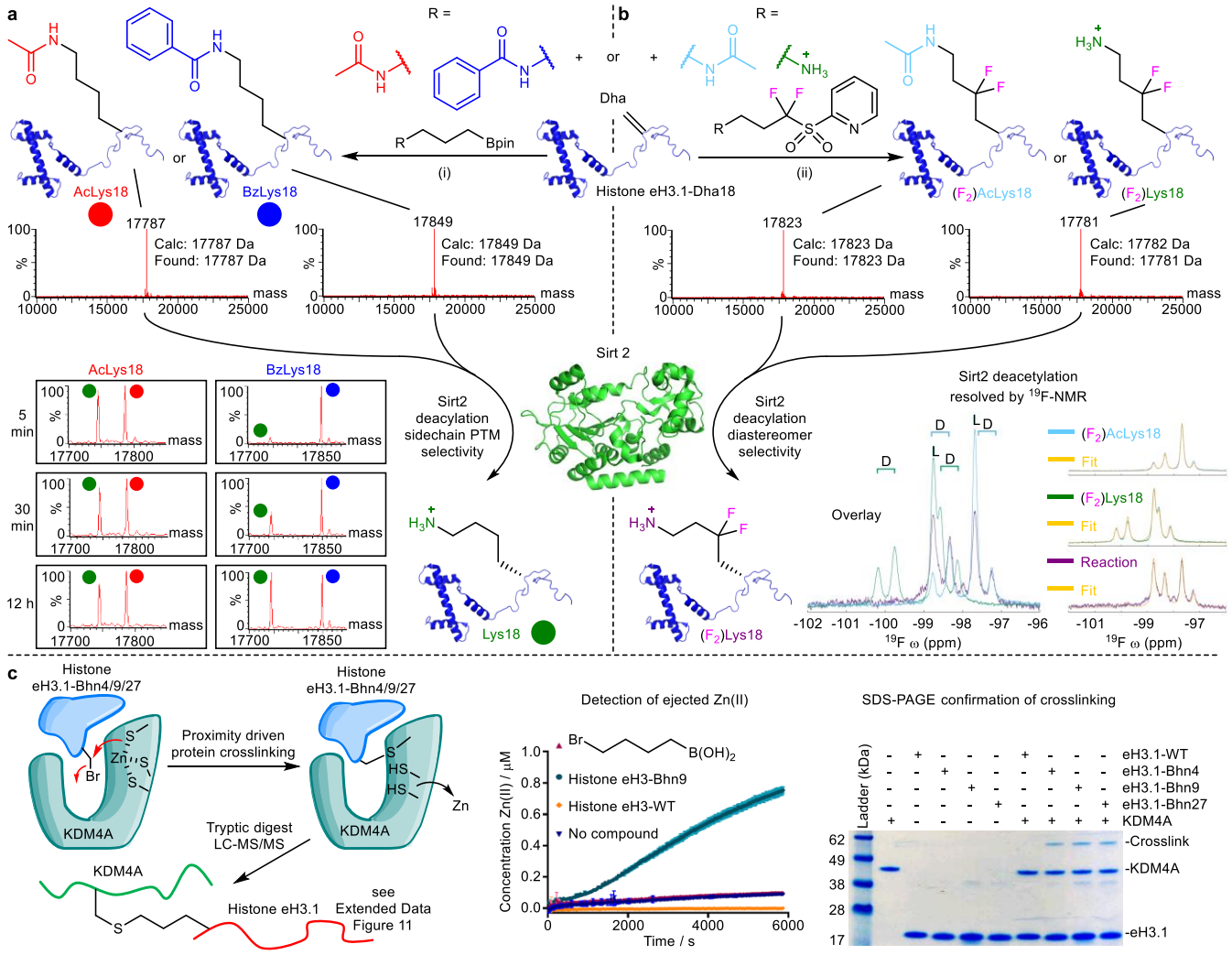
15 Klemes, Y., Etlinger, J. D. & Goldberg, A. L. Properties of abnormal proteins degraded rapidly in reticulocytes. Intracellular aggregation of the globin

molecules prior to hydrolysis. *Journal of Biological Chemistry* **256**, 8436-8444 (1981).
16 Tamura, T. & Hamachi, I. Chemistry for Covalent Modification of Endogenous/Native Proteins: From Test Tubes to Complex Biological Systems. *Journal of the American Chemical Society* **141**, 2782-2799, doi:10.1021/jacs.8b11747 (2019).
17 Wright, T. H. & Davis, B. G. Post-translational mutagenesis for installation of natural and unnatural amino acid side chains into recombinant proteins. *Nature Protocols* **12**, 2243-2250, doi:10.1038/nprot.2017.087 (2017).
18 Sletten, E. M. & Bertozzi, C. R. Bioorthogonal Chemistry: Fishing for Selectivity in a Sea of Functionality. *Angewandte Chemie International Edition* **48**, 6974-6998, doi:doi:10.1002/anie.200900942 (2009).
19 Imiolek, M. *et al.* Selective Radical Trifluoromethylation of Native Residues in Proteins. *Journal of the American Chemical Society* **140**, 1568-1571, doi:10.1021/jacs.7b10230 (2018).
20 Isenegger, P. G. & Davis, B. G. Concepts of Catalysis in Site-Selective Protein Modifications. *Journal of the American Chemical Society* **141**, 8005-8013, doi:10.1021/jacs.8b13187 (2019).
21 Lim, R. K. V. & Lin, Q. Photoinducible Bioorthogonal Chemistry: A Spatiotemporally Controllable Tool to Visualize and Perturb Proteins in Live Cells. *Accounts of Chemical Research* **44**, 828-839, doi:10.1021/ar200021p (2011).
22 Twilton, J. *et al.* The merger of transition metal and photocatalysis. *Nature Reviews Chemistry* **1**, 0052, doi:10.1038/s41570-017-0052 (2017).
23 Prier, C. K., Rankic, D. A. & MacMillan, D. W. C. Visible Light Photoredox Catalysis with Transition Metal Complexes: Applications in Organic Synthesis. *Chemical Reviews* **113**, 5322-5363, doi:10.1021/cr300503r (2013).
24 Bloom, S. *et al.* Decarboxylative alkylation for site-selective bioconjugation of native proteins via oxidation potentials. *Nat Chem* **10**, 205, doi:10.1038/nchem.2888 (2017).
25 Yu, Y. *et al.* Chemoselective Peptide Modification via Photocatalytic Tryptophan β-Position Conjugation. *Journal of the American Chemical Society* **140**, 6797-6800, doi:10.1021/jacs.8b03973 (2018).
26 deBruijn, A. D. & Roelfes, G. Chemical Modification of Dehydrated Amino Acids in Natural Antimicrobial Peptides by Photoredox Catalysis. *Chemistry – A European Journal* **24**, 11314-11318, doi:10.1002/chem.201803144 (2018).
27 Povie, G. *et al.* Catechols as Sources of Hydrogen Atoms in Radical Deiodination and Related Reactions. *Angewandte Chemie International Edition* **55**, 11221-11225, doi:10.1002/anie.201604950 (2016).
28 Matsui, J. K., Lang, S. B., Heitz, D. R. & Molander, G. A. Photoredox-Mediated Routes to Radicals: The Value of Catalytic Radical Generation in Synthetic Methods Development. *ACS Catalysis* **7**, 2563-2575, doi:10.1021/acscatal.7b00094 (2017).
29 During the course of this work the beneficial effects of catechols were also independently noted in small-molecule systems. See the following reference.
30 Robole, Z. M., Rahn, K. L., Lampkin, B. J., Anand, R. K. & VanVeller, B. Tuning the Electrochemical Redox Potentials of Catechol with Boronic Acid Derivatives. *The Journal of Organic Chemistry* **84**, 2346-2350, doi:10.1021/acs.joc.8b03087 (2019).

31 Ghosh, T. *et al.* Mixed-ligand complexes of ruthenium(II) containing new photoactive or electroactive ligands: synthesis, spectral characterization and DNA interactions. *JBIC Journal of Biological Inorganic Chemistry* **10**, 496, doi:10.1007/s00775-005-0660-6 (2005).
32 Dolbier, W. R. Structure, Reactivity, and Chemistry of Fluoroalkyl Radicals. *Chemical Reviews* **96**, 1557-1584, doi:10.1021/cr941142c (1996).
33 Zhang, L., Dolbier, W. R., Sheeller, B. & Ingold, K. U. Absolute Rate Constants of Alkene Addition Reactions of a Fluorinated Radical in Water. *Journal of the American Chemical Society* **124**, 6362-6366, doi:10.1021/ja0256010 (2002).
34 O'Hagan, D. Understanding organofluorine chemistry. An introduction to the C–F bond. *Chemical Society Reviews* **37**, 308-319, doi:10.1039/B711844A (2008).
35 Lemos, A., Lemaire, C. & Luxen, A. Progress in Difluoroalkylation of Organic Substrates by Visible Light Photoredox Catalysis. *Adv Synth Catal* **361**, 1500-1537, doi:10.1002/adsc.201801121 (2019).
36 Rong, J. *et al.* Radical Fluoroalkylation of Isocyanides with Fluorinated Sulfones by Visible-Light Photoredox Catalysis. *Angewandte Chemie International Edition* **55**, 2743-2747, doi:10.1002/anie.201510533 (2016).
37 Berlicki, L., Obojska, A., Forlani, G. & Kafarski, P. Design, Synthesis, and Activity of Analogues of Phosphinothricin as Inhibitors of Glutamine Synthetase. *J. Med. Chem.* **48**, 6340-6349, doi:10.1021/jm050474e (2005).
38 In our hands the final step to such precursors typically yields <1% of unstable product cf literature reports of <4% (see prior citation). The corresponding boronate is readily prepared and used in excellent yield.
39 Griller, D. & Ingold, K. U. Free-radical clocks. *Accounts of Chemical Research* **13**, 317-323, doi:10.1021/ar50153a004 (1980).
40 Chen, Y., Kamlet, A. S., Steinman, J. B. & Liu, D. R. A biomolecule-compatible visible-light-induced azide reduction from a DNA-encoded reaction-discovery system. *Nat Chem* **3**, 146-153, doi:10.1038/nchem.932 (2011).
41 Huang, H. *et al.* Lysine benzoylation is a histone mark regulated by SIRT2. *Nature Communications* **9**, 3374, doi:10.1038/s41467-018-05567-w (2018).
42 Dyer, P. N. *et al.* Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods Enzymol* **375**, 23-44, doi:10.1016/s0076-6879(03)75002-2 (2004).
43 Page, M. I. & Jencks, W. P. Entropic contributions to rate accelerations in enzymic and intramolecular reactions and the chelate effect. *Proceedings of the National Academy of Sciences of the United States of America* **68**, 1678-1683, doi:10.1073/pnas.68.8.1678 (1971).
44 Krishnamurthy, V. M., Semetey, V., Bracher, P. J., Shen, N. & Whitesides, G. M. Dependence of Effective Molarity on Linker Length for an Intramolecular Protein−Ligand System. *Journal of the American Chemical Society* **129**, 1312-1320, doi:10.1021/ja066780e (2007).
45 Ng, S. S. *et al.* Crystal structures of histone demethylase JMJD2A reveal basis for substrate specificity. *Nature* **448**, 87-91, doi:10.1038/nature05971 (2007).
46 English, C. M., Adkins, M. W., Carson, J. J., Churchill, M. E. & Tyler, J. K. Structural basis for the histone chaperone activity of Asf1. *Cell* **127**, 495-508, doi:10.1016/j.cell.2006.08.047 (2006).

47 Meeusen, J. W., Tomasiewicz, H., Nowakowski, A. & Petering, D. H. TSQ (6-methoxy-8-p-toluenesulfonamido-quinoline), a common fluorescent sensor for cellular zinc, images zinc proteins. *Inorg Chem* **50**, 7563-7573, doi:10.1021/ic200478q (2011).
48 Freedman, H. H. & Dubois, R. A. An improved Williamson ether synthesis using phase transfer catalysis. *Tetrahedron Letters* **16**, 3251-3254, doi:10.1016/S0040-4039(00)91417-5 (1975).
49 Mandal, S. *et al.* A review on the advancement of ether synthesis from organic solvent to water. *RSC Advances* **6**, 69605-69614, doi:10.1039/C6RA12914E (2016).
50 Levin, M., Stark, M. & Assaraf, Y. G. The JmjN domain as a dimerization interface and a targeted inhibitor of KDM4 demethylase activity. *Oncotarget* **9**, 16861-16882 (2018).
51 Shin, S. & Janknecht, R. Diversity within the JMJD2 histone demethylase family. *Biochemical and Biophysical Research Communications* **353**, 973-977 (2007).
52 Karle, I. L. & Balaram, P. Structural characteristics of .alpha.-helical peptide molecules containing Aib residues. *Biochemistry* **29**, 6747-6756, doi:10.1021/bi00481a001 (1990).
53 Lonsdale, R. & Ward, R. A. Structure-based design of targeted covalent inhibitors. *Chemical Society Reviews* **47**, 3816-3830, doi:10.1039/C7CS00220C (2018).
54 Angerani, S. & Winssinger, N. Visible Light Photoredox Catalysis Using Ruthenium Complexes in Chemical Biology. *Chemistry – A European Journal* **25**, 6661-6672, doi:10.1002/chem.201806024 (2019).
55 Yang, B. *et al.* Genetically Introducing Biochemically Reactive Amino Acids Dehydroalanine and Dehydrobutyrine in Proteins. *Journal of the American Chemical Society* **141**, 7698-7703, doi:10.1021/jacs.9b02611 (2019).


Dual modes of light driven alkyl radical generation
R' = H (catechol), CH2CH2NH2 (dopamine), CH2CH(NH2)COOH (L-DOPA)
in situ generated BACED
Fe(II) +
pySOOF derivatives
Blue LED light (450 nm)
Mild, water soluble [Ru^II] catalysts
radical generation
radical generation
or
High eq. needed
native residue/PTM precursor
Low eq. needed
'zero-size'-labeled
difluoro analogue precursor
Dha
Dha
radical generation
quenching
Aqueous buffer (degassed), RT, pH 6-8,
5-120 min, blue LED light (450 nm)
Protein containing
Dehydroalanine (Dha)
at desired site
X = H or F
70-95% conversion
>50 unique side chains
multiple protein folds, sites, and scaffolds
Diverse protein scaffolds
and sites
Histone H3
2 variants (H.s., X.l.)
sites: K4, K9, K18, K27
Histone H4
site: K16
AcrA
site: N123
PanC
site: H44
NPβ
site: M61
cabLys3
site: A104
Broad functional group tolerance
Amides
Azides
Ethers
Esters
Halogens
X = F, Cl, Br, I
On-protein
radical precursor
Sulfides/oxides
ones/ates
Methylamines
n = 0, 1, 2, 3


a
b
R =
or
Dha
Bpin
(i)
(ii)
AcLys18
BzLys18
Histone eH3.1-Dha18
(F2)AcLys18
(F2)Lys18
17787
Calc: 17787 Da
Found: 17787 Da
17849
Calc: 17849 Da
Found: 17849 Da
17823
Calc: 17823 Da
Found: 17823 Da
17781
Calc: 17782 Da
Found: 17781 Da
mass
Sirt 2
Sirt2 deacylation sidechain PTM selectivity
Sirt2 deacylation diastereomer selectivity
Sirt2 deacetylation resolved by 19F-NMR
5 min
30 min
12 h
Lys18
Overlay
Fit
Reaction
19F ω (ppm)

c
Histone eH3.1-Bhn4/9/27
Proximity driven protein crosslinking
KDM4A
Tryptic digest LC-MS/MS
Histone eH3.1
see Extended Data Figure 11
Detection of ejected Zn(II)
Histone eH3-Bhn9
Histone eH3-WT
No compound
Concentration Zn(II) / μM
Time / s
SDS-PAGE confirmation of crosslinking
Ladder (kDa)
eH3.1-WT
eH3.1-Bhn4
eH3.1-Bhn9
eH3.1-Bhn27
KDM4A
-Crosslink
-KDM4A
-eH3.1

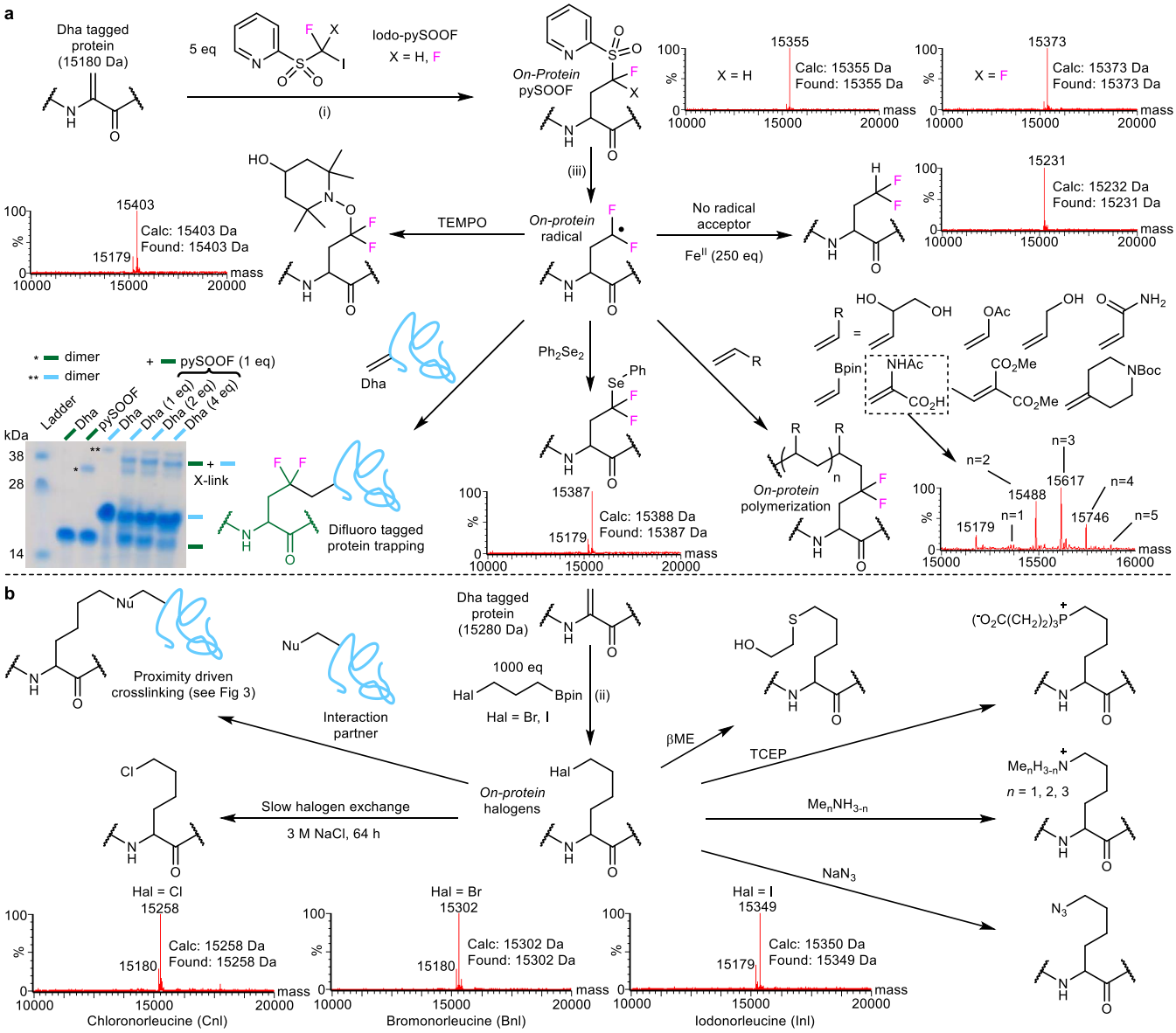

a
Dha tagged protein (15180 Da)
5 eq
Iodo-pySOOF
X = H, F
(i)
On-Protein pySOOF
X = H
Calc: 15355 Da
Found: 15355 Da
X = F
Calc: 15373 Da
Found: 15373 Da
(iii)
TEMPO
On-protein radical
No radical acceptor
Fe^II (250 eq)
Calc: 15403 Da
Found: 15403 Da
Calc: 15232 Da
Found: 15231 Da
Dha
Ph2Se2
dimer
dimer
pySOOF (1 eq)
Ladder
Dha
pySOOF
Dha
Dha (1 eq)
Dha (2 eq)
Dha (4 eq)
kDa
38
28
14
X-link
Difluoro tagged protein trapping
Calc: 15388 Da
Found: 15387 Da
On-protein polymerization
n=1
n=2
n=3
n=4
n=5
b
Proximity driven crosslinking (see Fig 3)
Interaction partner
Dha tagged protein (15280 Da)
1000 eq
Hal = Br, I
(ii)
On-protein halogens
βME
TCEP
Slow halogen exchange
3 M NaCl, 64 h
MenNH3-n
n = 1, 2, 3
NaN3
Hal = Cl
Calc: 15258 Da
Found: 15258 Da
Chloronorleucine (Cnl)
Hal = Br
Calc: 15302 Da
Found: 15302 Da
Bromonorleucine (Bnl)
Hal = I
Calc: 15350 Da
Found: 15349 Da
Iodonorleucine (Inl)

Figure legends

Figure 1. Site-Selective, Light-Driven Post-Translational Protein Editing. The excitation of mild, water soluble, protein compatible [Ru^{II}] photocatalysts with Blue Light enables dual modes of alkyl radical generation from complementary radical precursors with low oxidative exposure of protein substrates (E_{ox} < +1.0 V). *In situ* generated Boronic Acid Catechol Ester Derivatives (**BACED**) release less stable alkyl radicals, which when used in higher abundance (100-2000 eq), react selectively with the SOMOphile dehydroalanine (Dha) to install constitutively traceless, side-chain residues in natural, unnatural and post-translationally-modified (PTM) form. Difluoropyridylsulfone (**pySOOF**) derivatives release stabilized RF_2C• radicals, requiring even lower substrate concentrations (2-5 eq) that also efficiently react with Dha to install corresponding side-chains containing 'zero-size' difluoro labels (purple) at the gamma carbon ($C\gamma H_2 \rightarrow C\gamma F_2$). These low-$E_{ox}$, mild conditions allow not only application to diverse protein scaffolds but are also tolerated by reactive, side-chain groups allowing the direct, site-selective insertion of unprecedented chemical functionality (bottom right, >50 unique native or difluoro-labelled sidechains in proteins – see **ED Figure 8**).

160 words

Figure 2. *On-Protein* Homolytic and Heterolytic Reactivity via Installation of Radical-Precursor and Electrophile Side-Chains. a) Utilization of iodo-**pySOOF** allowed reductive installation of *on-protein* **pySOOF** side-chain that is itself a protein-radical precursor. Both mono- and difluoro- **pySOOF** sidechains were installed [Reagents-Conditions: **(i)** Histone-H3-Dha9 (66μM), Iodo-pySOOF (5 eq), $FeSO_4$•$7H_2O$ (20eq), $Ru(bpy)_3Cl_2$ (0.4eq), NH_4OAc (500mM, pH 6, 3M GdnHCl), 50W Blue-LED, RT, 15min]. Intact-protein-LC-MS (**right, top**).. After activation under standard conditions **(iii)** the resulting *on-protein* radical allowed further protein functionalization via *on-protein* homolytic bond-forming modes: polymerized with various radical acceptors via C–C-bond-formation (**right, bottom**); C–C-trapped by another Dha-containing protein – C–C-bond-forming protein-protein-crosslinking (**left, bottom**); quenched with stable-O•-nitroxide-radical to form C–O bonds (**left, middle**); used to cleave diselenide $(SePh)_2$ to form C–Se bonds (**centre, bottom**); or reduced (C–H-bond-formation) to DfeGly (**right, middle**). [Typical reagents-conditions: Histone-H3-pySOOF9 (66μM), substrate (10-250eq), $FeSO_4$•$7H_2O$ (0-25eq), $Ru(bpy)_3Cl_2$ 1-5eq), NH_4OAc (500mM, pH 6, 3M GdnHCl), 50W Blue LED, RT, 15min, see **Supplementary Tables 23-37** for details, residual-Dha = 15179 Da]. **b)** Utilization of alkylhalide-functionalized **BACED** reagent allowed oxidative installation with C-Halogen(C-Hal)-bond unperturbed. This installs *on-protein* alkylhalide electrophile side-chains (yellow highlight) [Reagents-Conditions: **(ii)** Histone H3-Dha9 (66 μM), alkylboronic acid pinacol ester (1000 eq), catechol (100 eq), $Ru(bpm)_3Cl_2$ (10 eq), NH_4OAc (500mM, pH 6, 3M GdnHCl), 50 W Blue LED, RT, 1-3 h)]. This provided a further reaction platform for *on-protein* heterolytic bond-forming modes. These *on-protein* alkylhalide electrophiles were reacted through substitution with various small molecule P, S, N, Hal nucleophiles at higher concentrations (TCEP = tris(2-carboxyethyl)phosphine, βME = betamercaptoethanol), allowing diverse C–P, C–S,

C–N, C–Hal bond-formation (residual Dha = 15179/80 Da). The ability to install a range of alkylhalide side-chains (e.g., chloro-(Cnl), bromo-(Bnl), iodo-(Inl) norleucines, intact-protein-LC-MS, **left, bottom**) also allowed protein-protein-crosslinking with interaction partners (see **Figure 3c**).

295 words

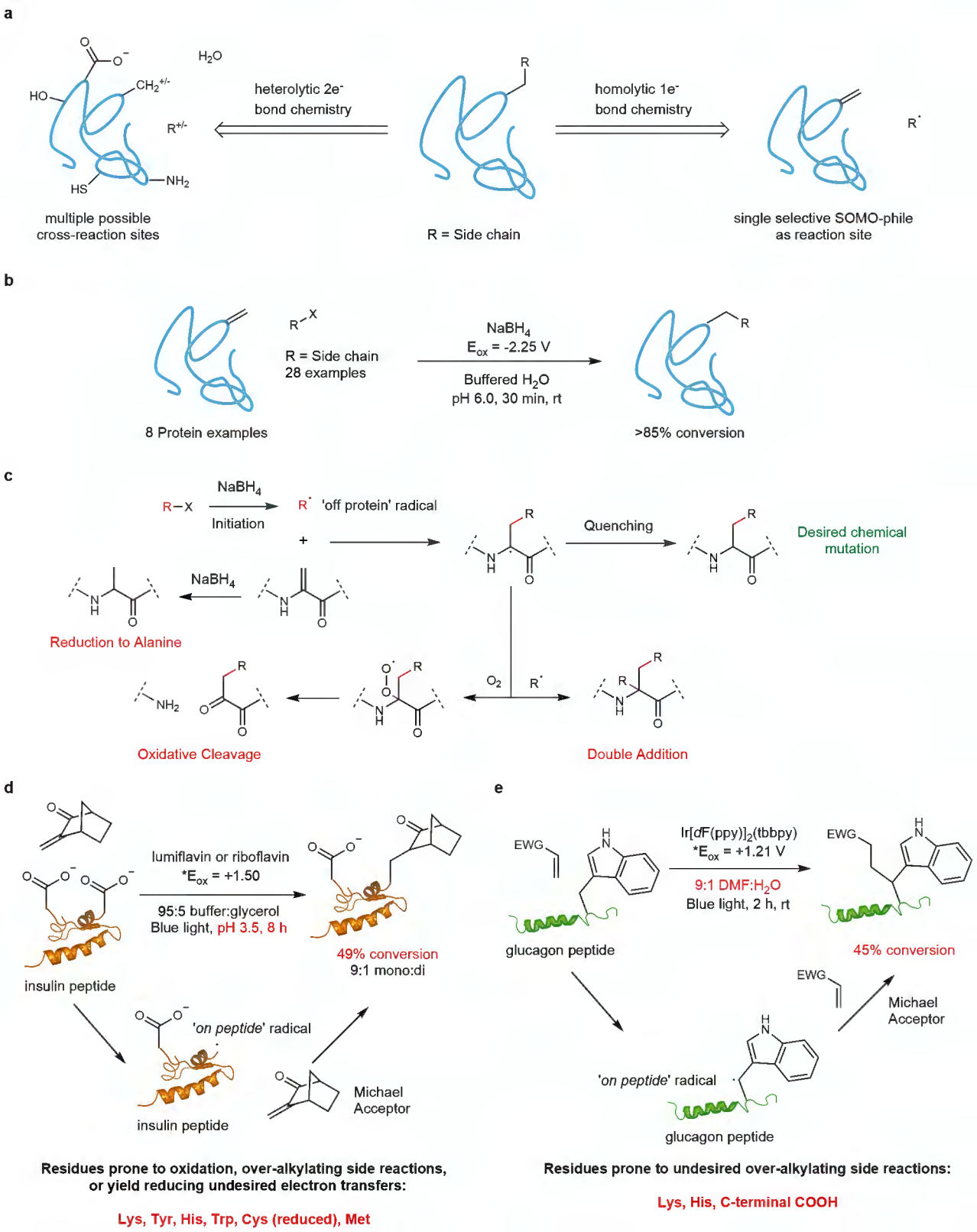
Figure 3: Insertion of Native, Difluoro-labeled and Electrophile-containing Sidechains into Proteins Allows Insight into Enzymes that Post-translationally Modify Them. a) Sirt2 enzyme displays different deacylation rates (via intact-protein-LC-MS monitoring) towards installed acetyl-(**red**) and benzoyl-(**blue**) lysine on Histone-eH3.1-K18 proteins thereby confirming, in this experiment, both plasticity but also selectivity towards acyl groups. [Reagents-Conditions for Installation: **(i)** Histone-H3-Dha9(66µM), pinacol alkylboronate(250eq), catechol(100eq), $Ru(bpm)_3Cl_2$ (10eq), NH_4OAc (500mM, pH 6, 3M GdnHCl), 50W Blue-LED, RT, 1h)]. **b)** Activity of Sirt2-deacetylation can be directly and site-specifically monitored via ^{19}F-NMR using a $C\gamma F_2$-difluoro-tag on gamma-carbon of installed Lys(**green**) and AcLys(**red**) sidechains [Reagents-Conditions: **(ii)** Histone-H3-Dha9(66µM), alkyl-pySOOF(50eq), $FeSO_4 \bullet 7H_2O$(50eq), $Ru(bpy)_3Cl_2$(2eq), NH_4OAc(500mM, pH 6, 3M GdnHCl), 50W Blue-LED, RT, 15min). Met_{ox} = 15838Da]. Although four-bonds-distant from site of PTM, $C\gamma F_2$-labels display sufficient sensitivity to chemical environment (δ_F perturbation) to allow direct simultaneous monitoring of Sirt2's chemo- and stereo-selectivity during processing (**ED Figure 9**). **c)** Crosslinking between KDM4A and Histone-eH3.1-Bhn4/9/27 traps KDM4A-Zn-binding-cysteines near active site. Coomassie-Blue-SDS-PAGE **(centre),** tryptic-LC-MS/MS (**bottom,left**, see also **ED Figure 3c**) and Zn(II)-ejection (**right**) confirm crosslinking between KDM4A and Histone-eH3.1-Bhn9 (see also **ED Figure 11c**) [Zn(II)-ejection rates: eH3-Bhn9 = 9.27±0.025 nM/min, eH3-WT = 0.09±0.006 nM/min, **1u**-precursor = 0.805±0.010 nM/min, no-compound = 0.87±0.028 nM/min, N=3 independent experiments. Data plotted is average +/- standard deviation (N=3 technical replicates), $p < 0.0001$ 1-way ANOVA]. See also **ED Figure 11** for further alkylator protein experiments.

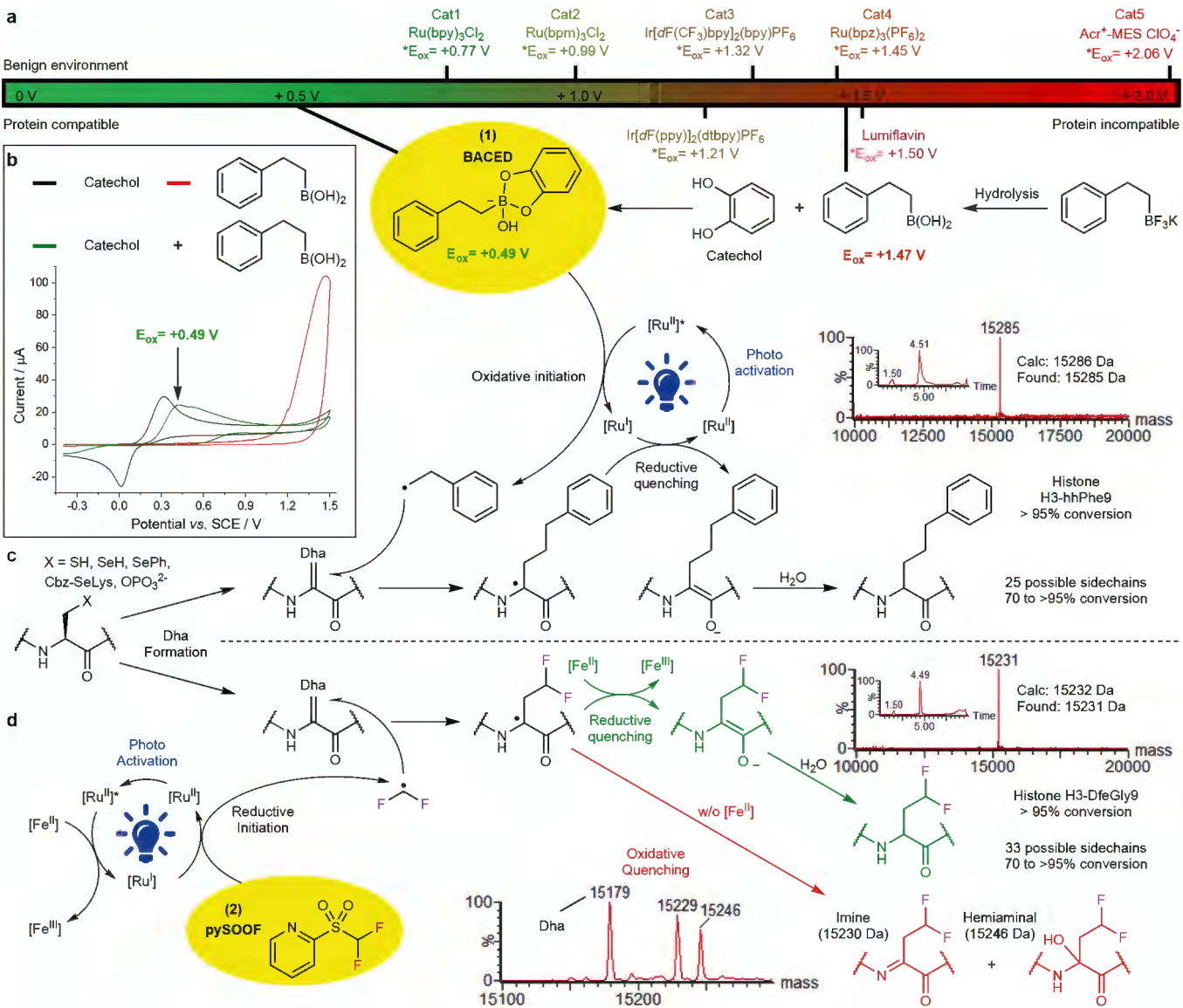
235 words

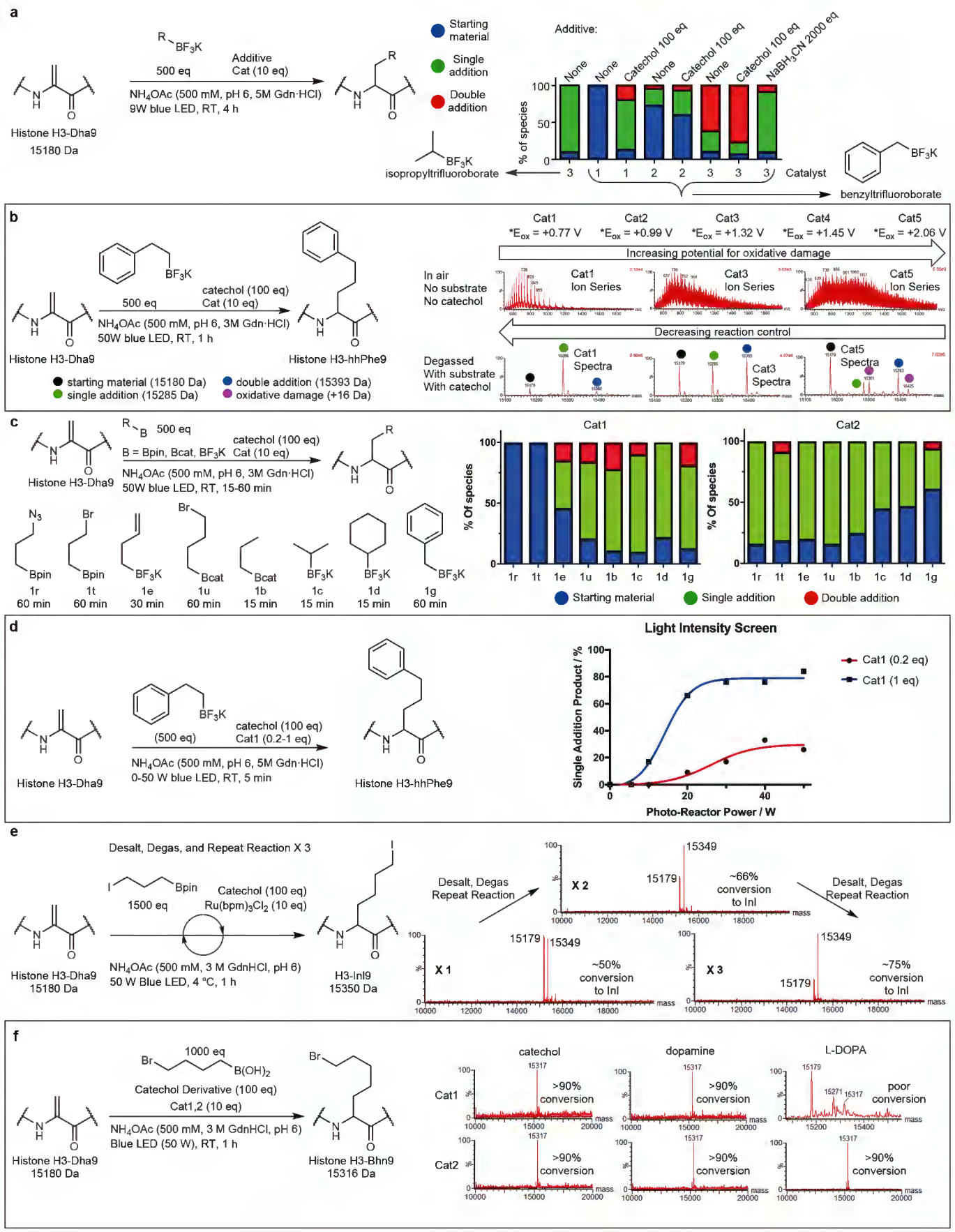
Online Methods

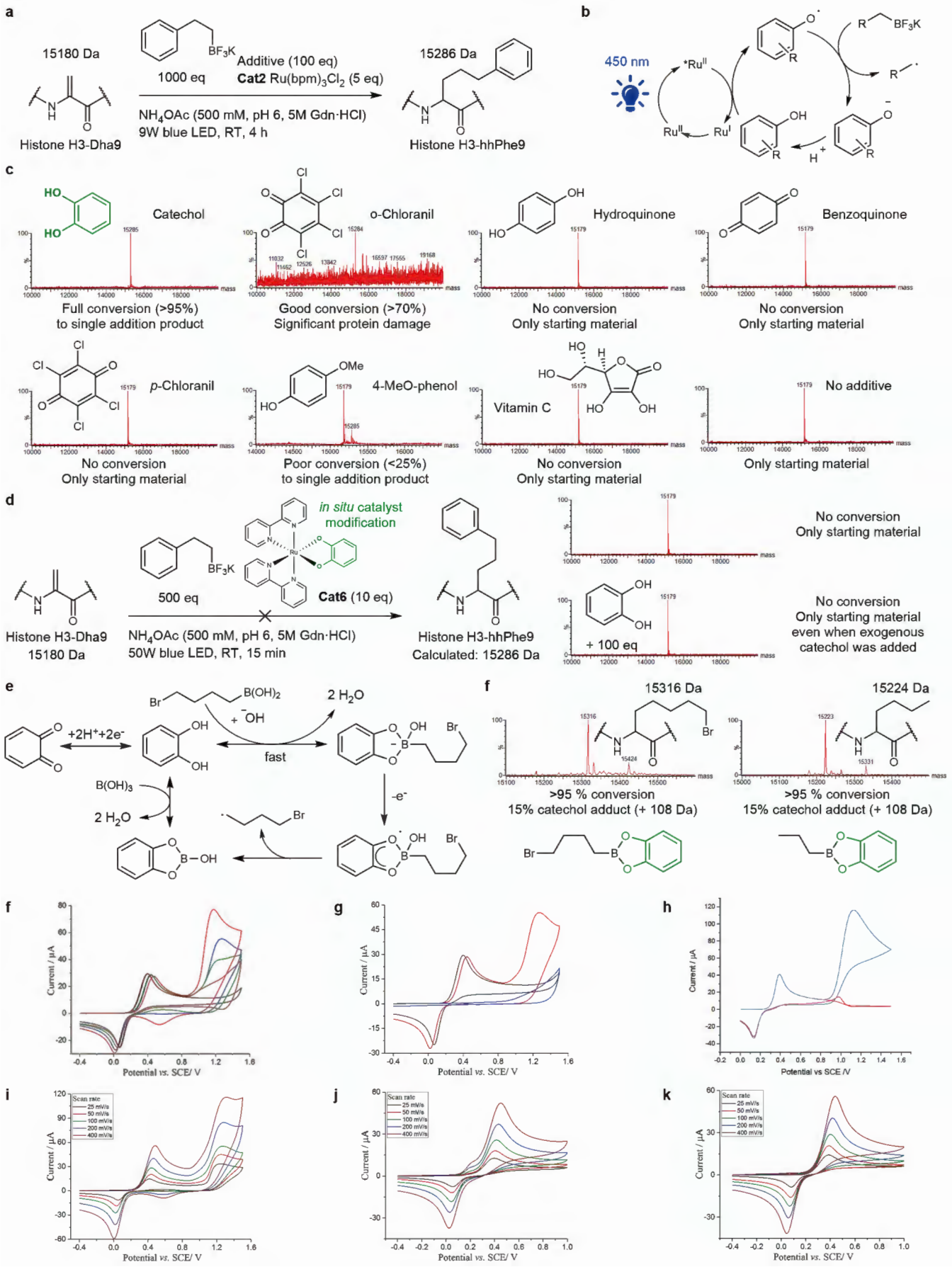
General Experiment Protocol for BACED and pySOOF Reagents

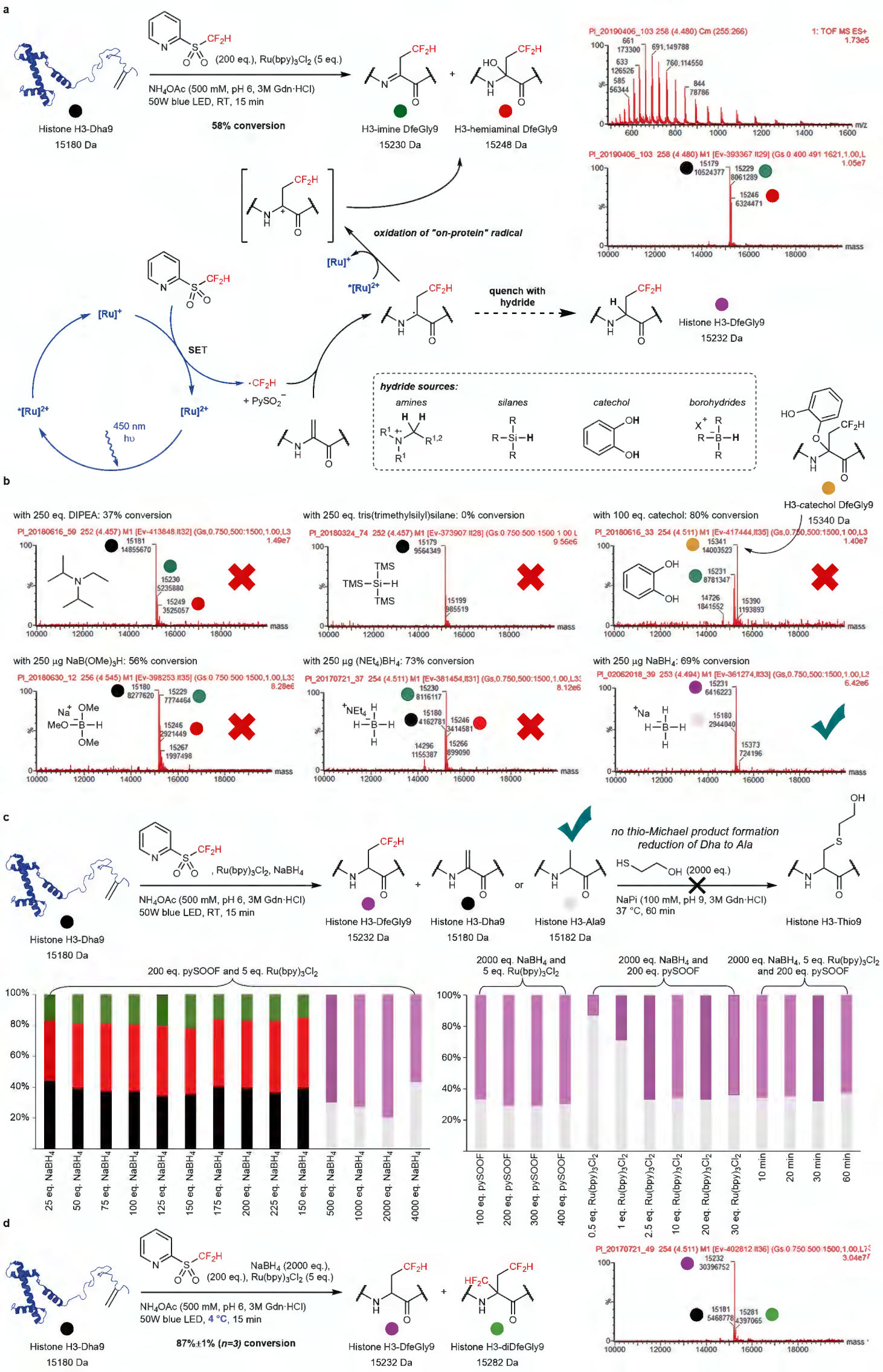
For all protein modification reactions, all aqueous buffers, solvents and protein stocks were degassed for at least 8 h in a glovebox (<6 ppm O_2). Then, inside the glovebox a glass vial was charged with the Dha-tagged protein of choice and diluted to the protein concentration of choice with the desired reaction buffer, followed by the sequential addition of an aliquot of catalyst, additive and radical precursor from stocks prepared fresh in buffer (in some cases, small amounts of cosolvent were needed in stock preparation and are explicitly stated when relevant). Afterwards, the reaction mixtures were mixed either by shaking or thoroughly by pipette, capped and transferred out from the glovebox for irradiation with blue LED (50 W) for the stated reaction time. After the reaction an aliquot of the crude mixture was diluted 25-fold for mass spectrometric analysis (2 µL in 48 µL water + 0.1% formic acid) and conversions calculated relative to total ion counts. For selected examples (see **Supplementary Methods**) high (>85%) protein recovery was observed after purification by PD SpinTrap G-25 (GE Healthcare) desalting columns and tracking overall protein absorbance. *Note:* Most reaction optimization and BACED/pySOOF reagent screening reactions were performed on model protein substrate *X.l.* Histone H3-Dha9 in denaturing buffer (500 mM NH_4OAc, 3 M Gdn·HCl, pH 6.0) at a final concentration of 1 mg/mL (66 µM) in volumes of 50-200 µL. Short reaction times (<20 min) did not result in significant temperature increases but longer reaction times (>20 min) could have their temperatures controlled by submerging the reaction vials in a glass beaker filled with water at the desired temperature.

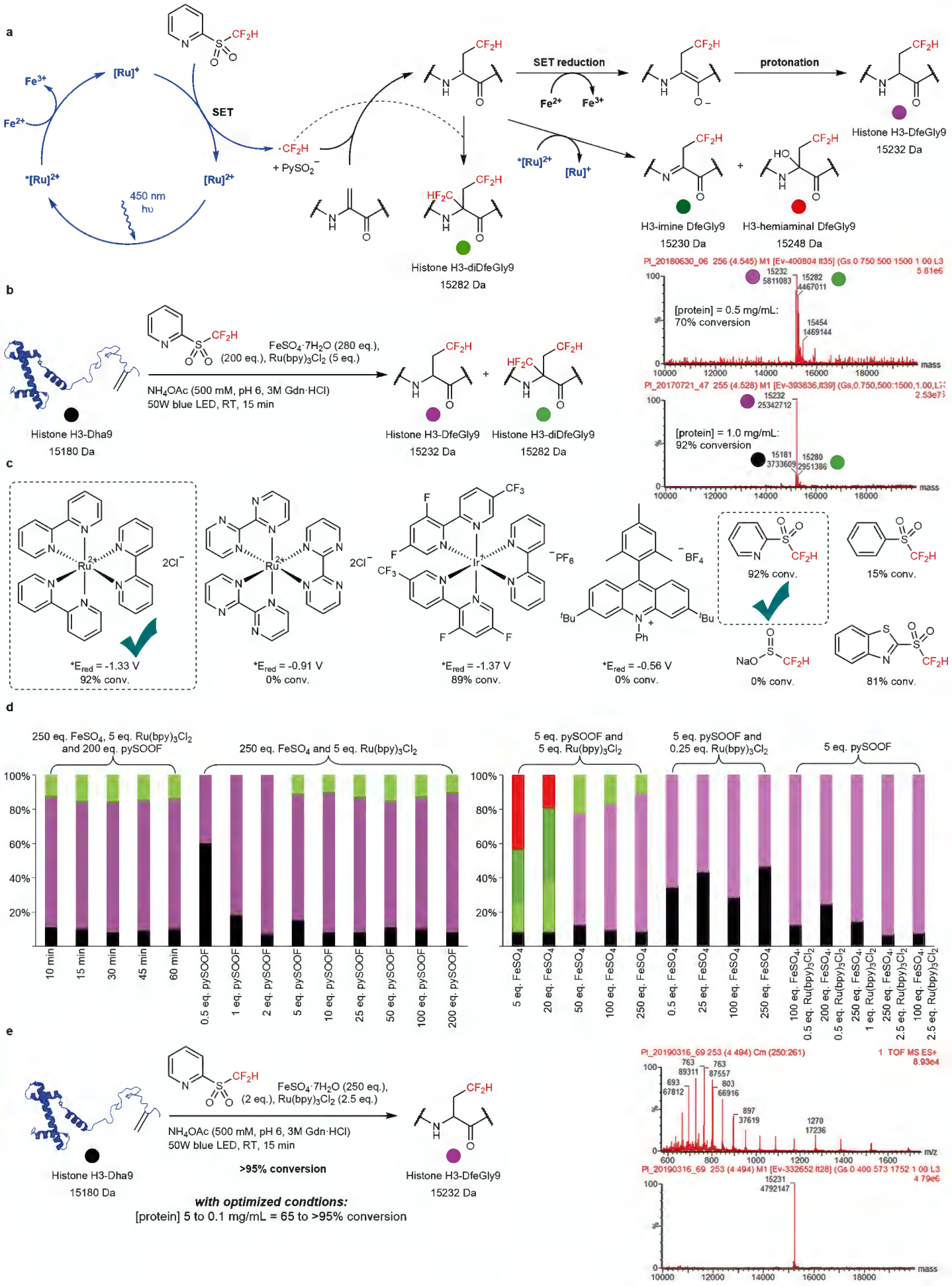

a
heterolytic 2e⁻ bond chemistry
homolytic 1e⁻ bond chemistry
multiple possible cross-reaction sites
R = Side chain
single selective SOMO-phile as reaction site
b
R = Side chain
28 examples
NaBH4
Eox = -2.25 V
Buffered H2O
pH 6.0, 30 min, rt
8 Protein examples
>85% conversion
c
NaBH4
Initiation
'off protein' radical
Quenching
Desired chemical mutation
Reduction to Alanine
Oxidative Cleavage
Double Addition
d
lumiflavin or riboflavin
*Eox = +1.50
95:5 buffer:glycerol
Blue light, pH 3.5, 8 h
49% conversion
9:1 mono:di
insulin peptide
'on peptide' radical
Michael Acceptor
Residues prone to oxidation, over-alkylating side reactions, or yield reducing undesired electron transfers:
Lys, Tyr, His, Trp, Cys (reduced), Met
e
Ir[dF(ppy)]2(tbbpy)
*Eox = +1.21 V
9:1 DMF:H2O
Blue light, 2 h, rt
45% conversion
glucagon peptide
'on peptide' radical
Michael Acceptor
Residues prone to undesired over-alkylating side reactions:
Lys, His, C-terminal COOH

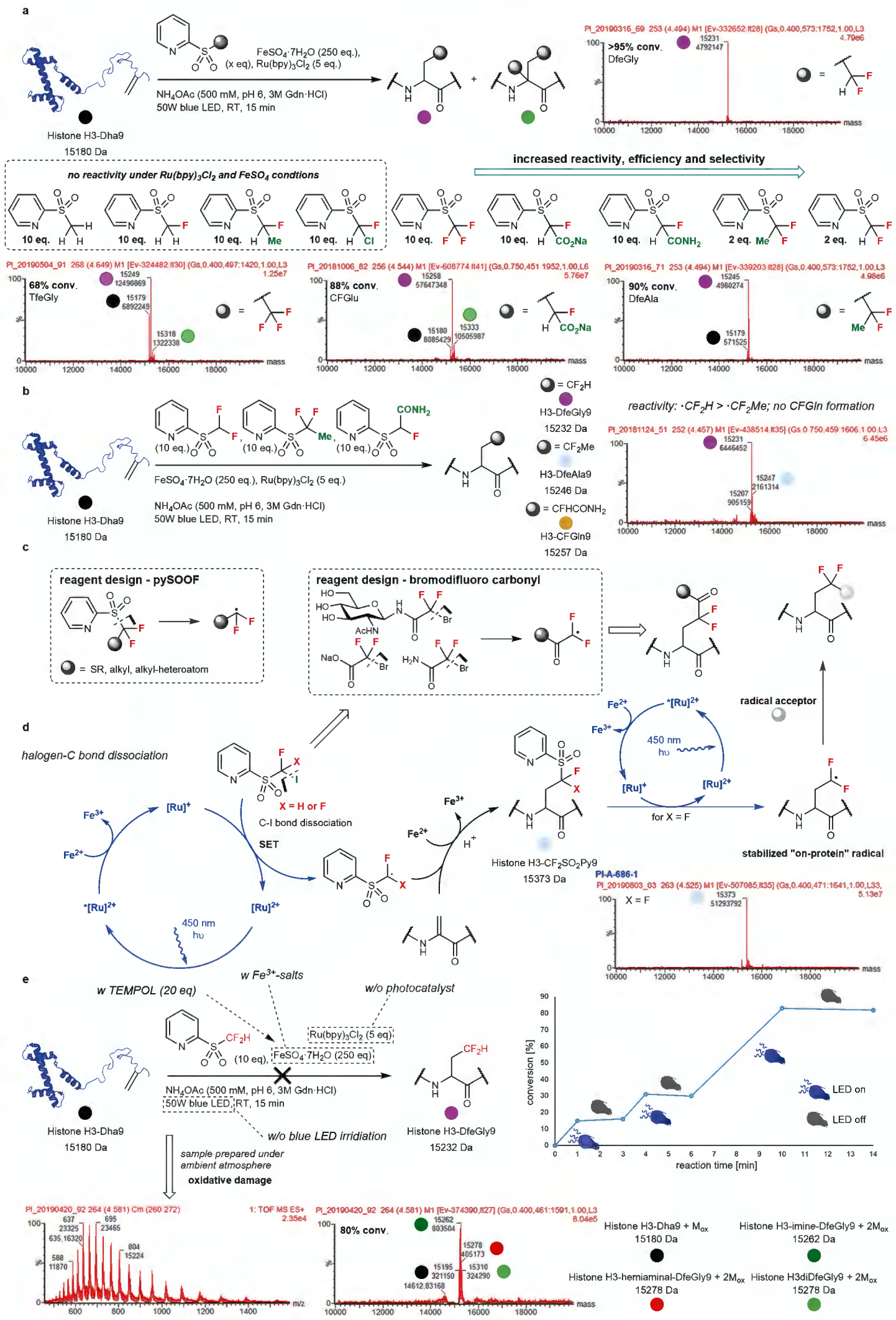

a
Cat1
Ru(bpy)3Cl2
*Eox= +0.77 V
Cat2
Ru(bpm)3Cl2
*Eox= +0.99 V
Cat3
Ir[dF(CF3)bpy]2(bpy)PF6
*Eox= +1.32 V
Cat4
Ru(bpz)3(PF6)2
*Eox= +1.45 V
Cat5
Acr+-MES ClO4-
*Eox= +2.06 V
Benign environment
0 V
+ 0.5 V
+ 1.0 V
Protein compatible
Protein incompatible
Ir[dF(ppy)]2(dtbpy)PF6
*Eox= +1.21 V
Lumiflavin
*Eox= +1.50 V
(1)
BACED
Eox= +0.49 V
Catechol
Eox= +1.47 V
Hydrolysis
b
Catechol
Catechol +
Eox= +0.49 V
Current / µA
Potential vs. SCE / V
Oxidative initiation
Photo activation
Reductive quenching
Histone
H3-hhPhe9
> 95% conversion
15285
Calc: 15286 Da
Found: 15285 Da
c
X = SH, SeH, SePh,
Cbz-SeLys, OPO3 2-
Dha
Dha Formation
H2O
25 possible sidechains
70 to >95% conversion
d
Photo Activation
Reductive Initiation
Oxidative Quenching
w/o [FeII]
(2)
pySOOF
15231
Calc: 15232 Da
Found: 15231 Da
Histone H3-DfeGly9
> 95% conversion
33 possible sidechains
70 to >95% conversion
15179
15229
15246
Imine
(15230 Da)
Hemiaminal
(15246 Da)

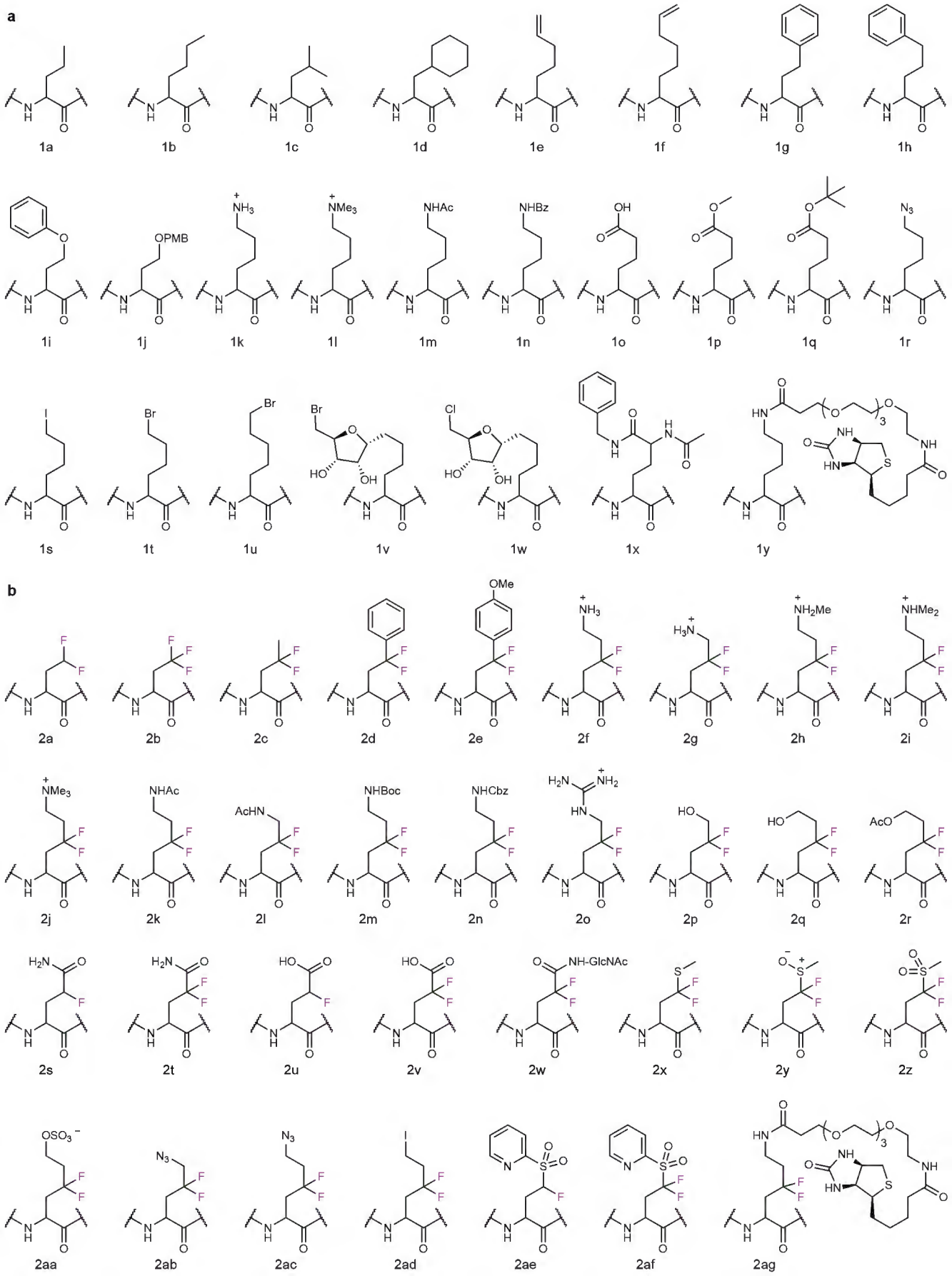

a
Histone H3-Dha9
15180 Da
500 eq
Additive
Cat (10 eq)
NH4OAc (500 mM, pH 6, 5M Gdn·HCl)
9W blue LED, RT, 4 h
Starting material
Single addition
Double addition
isopropyltrifluoroborate
Additive:
None
None
Catechol 100 eq
None
Catechol 100 eq
None
Catechol 100 eq
NaBH3CN 2000 eq
% of species
Catalyst
benzyltrifluoroborate
b
500 eq
catechol (100 eq)
Cat (10 eq)
NH4OAc (500 mM, pH 6, 3M Gdn·HCl)
50W blue LED, RT, 1 h
Histone H3-Dha9
Histone H3-hhPhe9
starting material (15180 Da)
single addition (15285 Da)
double addition (15393 Da)
oxidative damage (+16 Da)
Cat1 *Eox = +0.77 V
Cat2 *Eox = +0.99 V
Cat3 *Eox = +1.32 V
Cat4 *Eox = +1.45 V
Cat5 *Eox = +2.06 V
Increasing potential for oxidative damage
In air
No substrate
No catechol
Cat1 Ion Series
Cat3 Ion Series
Cat5 Ion Series
Decreasing reaction control
Degassed
With substrate
With catechol
Cat1 Spectra
Cat3 Spectra
Cat5 Spectra
c
B = Bpin, Bcat, BF3K
500 eq
catechol (100 eq)
Cat (10 eq)
NH4OAc (500 mM, pH 6, 3M Gdn·HCl)
50W blue LED, RT, 15-60 min
Histone H3-Dha9
1r 60 min
1t 60 min
1e 30 min
1u 60 min
1b 15 min
1c 15 min
1d 15 min
1g 60 min
Cat1
Cat2
% Of species
Starting material
Single addition
Double addition
d
(500 eq)
catechol (100 eq)
Cat1 (0.2-1 eq)
NH4OAc (500 mM, pH 6, 5M Gdn·HCl)
0-50 W blue LED, RT, 5 min
Histone H3-Dha9
Histone H3-hhPhe9
Light Intensity Screen
Single Addition Product / %
Photo-Reactor Power / W
Cat1 (0.2 eq)
Cat1 (1 eq)
e
Desalt, Degas, and Repeat Reaction X 3
1500 eq
Catechol (100 eq)
Ru(bpm)3Cl2 (10 eq)
NH4OAc (500 mM, 3 M GdnHCl, pH 6)
50 W Blue LED, 4 °C, 1 h
Histone H3-Dha9
15180 Da
H3-InI9
15350 Da
Desalt, Degas
Repeat Reaction
X 2
15179
15349
~66%
conversion
to InI
X 1
~50%
conversion
to InI
X 3
~75%
conversion
to InI
mass
f
1000 eq
Catechol Derivative (100 eq)
Cat1,2 (10 eq)
NH4OAc (500 mM, 3 M GdnHCl, pH 6)
Blue LED (50 W), RT, 1 h
Histone H3-Dha9
15180 Da
Histone H3-Bhn9
15316 Da
catechol
dopamine
L-DOPA
Cat1
Cat2
>90%
conversion
poor
conversion
15317
15179
15271

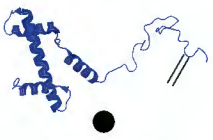

a
15180 Da
1000 eq
Additive (100 eq)
Cat2 Ru(bpm)3Cl2 (5 eq)
NH4OAc (500 mM, pH 6, 5M Gdn·HCl)
9W blue LED, RT, 4 h
15286 Da
Histone H3-Dha9
Histone H3-hhPhe9
b
450 nm
*Ru^II
Ru^II
Ru^I
H+
c
Catechol
Full conversion (>95%)
to single addition product
o-Chloranil
Good conversion (>70%)
Significant protein damage
Hydroquinone
No conversion
Only starting material
Benzoquinone
No conversion
Only starting material
p-Chloranil
No conversion
Only starting material
4-MeO-phenol
Poor conversion (<25%)
to single addition product
Vitamin C
No conversion
Only starting material
No additive
Only starting material
d
in situ catalyst modification
500 eq
Cat6 (10 eq)
Histone H3-Dha9
15180 Da
NH4OAc (500 mM, pH 6, 5M Gdn·HCl)
50W blue LED, RT, 15 min
Histone H3-hhPhe9
Calculated: 15286 Da
No conversion
Only starting material
+ 100 eq
No conversion
Only starting material
even when exogenous
catechol was added
e
+2H+ +2e-
2 H2O
fast
B(OH)3
2 H2O
-e-
f
15316 Da
>95 % conversion
15% catechol adduct (+ 108 Da)
15224 Da
>95 % conversion
15% catechol adduct (+ 108 Da)
f
Current / µA
Potential vs. SCE/ V
g
Current / µA
Potential vs. SCE/ V
h
Current / µA
Potential vs SCE /V
i
Scan rate
25 mV/s
50 mV/s
100 mV/s
200 mV/s
400 mV/s
Current / µA
Potential vs. SCE/ V
j
Scan rate
25 mV/s
50 mV/s
100 mV/s
200 mV/s
400 mV/s
Current / µA
Potential vs. SCE/ V
k
Scan rate
25 mV/s
50 mV/s
100 mV/s
200 mV/s
400 mV/s
Current / µA
Potential vs. SCE/ V


a
Histone H3-Dha9
15180 Da
(200 eq.), Ru(bpy)3Cl2 (5 eq.)
NH4OAc (500 mM, pH 6, 3M Gdn·HCl)
50W blue LED, RT, 15 min
58% conversion
H3-imine DfeGly9
15230 Da
H3-hemiaminal DfeGly9
15248 Da
oxidation of "on-protein" radical
[Ru]+
*[Ru]2+
SET
·CF2H
+ PySO2−
*[Ru]2+
[Ru]2+
450 nm
hυ
quench with hydride
Histone H3-DfeGly9
15232 Da
hydride sources:
amines
silanes
catechol
borohydrides
H3-catechol DfeGly9
15340 Da
b
with 250 eq. DIPEA: 37% conversion
with 250 eq. tris(trimethylsilyl)silane: 0% conversion
with 100 eq. catechol: 80% conversion
with 250 μg NaB(OMe)3H: 56% conversion
with 250 μg (NEt4)BH4: 73% conversion
with 250 μg NaBH4: 69% conversion
c
Histone H3-Dha9
15180 Da
, Ru(bpy)3Cl2, NaBH4
NH4OAc (500 mM, pH 6, 3M Gdn·HCl)
50W blue LED, RT, 15 min
Histone H3-DfeGly9
15232 Da
Histone H3-Dha9
15180 Da
or
Histone H3-Ala9
15182 Da
no thio-Michael product formation
reduction of Dha to Ala
(2000 eq.)
NaPi (100 mM, pH 9, 3M Gdn·HCl)
37 °C, 60 min
Histone H3-Thio9
200 eq. pySOOF and 5 eq. Ru(bpy)3Cl2
2000 eq. NaBH4 and 5 eq. Ru(bpy)3Cl2
2000 eq. NaBH4 and 200 eq. pySOOF
2000 eq. NaBH4, 5 eq. Ru(bpy)3Cl2 and 200 eq. pySOOF
25 eq. NaBH4
50 eq. NaBH4
75 eq. NaBH4
100 eq. NaBH4
125 eq. NaBH4
150 eq. NaBH4
175 eq. NaBH4
200 eq. NaBH4
225 eq. NaBH4
150 eq. NaBH4
500 eq. NaBH4
1000 eq. NaBH4
2000 eq. NaBH4
4000 eq. NaBH4
100 eq. pySOOF
200 eq. pySOOF
300 eq. pySOOF
400 eq. pySOOF
0.5 eq. Ru(bpy)3Cl2
1 eq. Ru(bpy)3Cl2
2.5 eq. Ru(bpy)3Cl2
10 eq. Ru(bpy)3Cl2
20 eq. Ru(bpy)3Cl2
30 eq. Ru(bpy)3Cl2
10 min
20 min
30 min
60 min
d
Histone H3-Dha9
15180 Da
NaBH4 (2000 eq.),
(200 eq.), Ru(bpy)3Cl2 (5 eq.)
NH4OAc (500 mM, pH 6, 3M Gdn·HCl)
50W blue LED, 4 °C, 15 min
87%±1% (n=3) conversion
Histone H3-DfeGly9
15232 Da
Histone H3-diDfeGly9
15282 Da

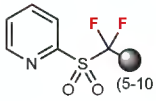

a
Fe3+
Fe2+
[Ru]+
SET
*[Ru]2+
[Ru]2+
450 nm
hυ
·CF2H
+ PySO2−
SET reduction
Fe2+
Fe3+
protonation
*[Ru]2+
[Ru]+
Histone H3-DfeGly9
15232 Da
H3-imine DfeGly9
15230 Da
H3-hemiaminal DfeGly9
15248 Da
Histone H3-diDfeGly9
15282 Da
b
FeSO4·7H2O (280 eq.),
(200 eq.), Ru(bpy)3Cl2 (5 eq.)
NH4OAc (500 mM, pH 6, 3M Gdn·HCl)
50W blue LED, RT, 15 min
Histone H3-Dha9
15180 Da
Histone H3-DfeGly9
15232 Da
Histone H3-diDfeGly9
15282 Da
[protein] = 0.5 mg/mL:
70% conversion
[protein] = 1.0 mg/mL:
92% conversion
mass
c
*Ered = -1.33 V
92% conv.
*Ered = -0.91 V
0% conv.
*Ered = -1.37 V
89% conv.
*Ered = -0.56 V
0% conv.
92% conv.
15% conv.
0% conv.
81% conv.
d
250 eq. FeSO4, 5 eq. Ru(bpy)3Cl2
and 200 eq. pySOOF
250 eq. FeSO4 and 5 eq. Ru(bpy)3Cl2
5 eq. pySOOF and
5 eq. Ru(bpy)3Cl2
5 eq. pySOOF and
0.25 eq. Ru(bpy)3Cl2
5 eq. pySOOF
10 min
15 min
30 min
45 min
60 min
0.5 eq. pySOOF
1 eq. pySOOF
2 eq. pySOOF
5 eq. pySOOF
10 eq. pySOOF
25 eq. pySOOF
50 eq. pySOOF
100 eq. pySOOF
200 eq. pySOOF
5 eq. FeSO4
20 eq. FeSO4
50 eq. FeSO4
100 eq. FeSO4
250 eq. FeSO4
0.5 eq. FeSO4, 0.5 eq. Ru(bpy)3Cl2
1 eq. Ru(bpy)3Cl2
2.5 eq. Ru(bpy)3Cl2
e
FeSO4·7H2O (250 eq.),
(2 eq.), Ru(bpy)3Cl2 (2.5 eq.)
NH4OAc (500 mM, pH 6, 3M Gdn·HCl)
50W blue LED, RT, 15 min
>95% conversion
Histone H3-Dha9
15180 Da
Histone H3-DfeGly9
15232 Da
with optimized condtions:
[protein] 5 to 0.1 mg/mL = 65 to >95% conversion
m/z
mass


a
Histone H3-Dha9
15180 Da
FeSO4·7H2O (250 eq.),
(x eq), Ru(bpy)3Cl2 (5 eq.)
NH4OAc (500 mM, pH 6, 3M Gdn·HCl)
50W blue LED, RT, 15 min
>95% conv.
DfeGly
no reactivity under Ru(bpy)3Cl2 and FeSO4 condtions
increased reactivity, efficiency and selectivity
10 eq. H H
10 eq. H F
10 eq. H Me
10 eq. H Cl
10 eq. F F
10 eq. H CO2Na
10 eq. H CONH2
2 eq. Me F
2 eq. H F
68% conv.
TfeGly
88% conv.
CFGlu
90% conv.
DfeAla
b
(10 eq.)
FeSO4·7H2O (250 eq.), Ru(bpy)3Cl2 (5 eq.)
NH4OAc (500 mM, pH 6, 3M Gdn·HCl)
50W blue LED, RT, 15 min
Histone H3-Dha9
15180 Da
= CF2H
H3-DfeGly9
15232 Da
= CF2Me
H3-DfeAla9
15246 Da
= CFHCONH2
H3-CFGln9
15257 Da
reactivity: ·CF2H > ·CF2Me; no CFGln formation
c
reagent design - pySOOF
= SR, alkyl, alkyl-heteroatom
reagent design - bromodifluoro carbonyl
d
halogen-C bond dissociation
X = H or F
C-I bond dissociation
SET
[Ru]+
[Ru]2+
*[Ru]2+
Fe2+
Fe3+
450 nm
hυ
for X = F
radical acceptor
stabilized "on-protein" radical
Histone H3-CF2SO2Py9
15373 Da
PI-A-686-1
X = F
e
w TEMPOL (20 eq)
w Fe3+-salts
w/o photocatalyst
Ru(bpy)3Cl2 (5 eq)
(10 eq), FeSO4·7H2O (250 eq)
NH4OAc (500 mM, pH 6, 3M Gdn·HCl)
50W blue LED, RT, 15 min
w/o blue LED irridiation
Histone H3-Dha9
15180 Da
Histone H3-DfeGly9
15232 Da
sample prepared under
ambient atmosphere
oxidative damage
conversion [%]
reaction time [min]
LED on
LED off
Histone H3-Dha9 + Mox
15180 Da
Histone H3-imine-DfeGly9 + 2Mox
15262 Da
Histone H3-hemiaminal-DfeGly9 + 2Mox
15278 Da
Histone H3diDfeGly9 + 2Mox
15278 Da


a
1a
1b
1c
1d
1e
1f
1g
1h
1i
OPMB
1j
+NH3
1k
+NMe3
1l
NHAc
1m
NHBz
1n
OH
1o
1p
1q
N3
1r
I
1s
Br
1t
Br
1u
Br
HO
OH
1v
Cl
HO
OH
1w
1x
HN
3
NH
HN
S
1y
b
2a
2b
2c
2d
OMe
2e
+NH3
2f
H3N+
2g
+NH2Me
2h
+NHMe2
2i
+NMe3
2j
NHAc
2k
AcHN
2l
NHBoc
2m
NHCbz
2n
H2N
+NH2
HN
2o
HO
2p
HO
2q
AcO
2r
H2N
2s
H2N
2t
HO
2u
HO
2v
NH-GlcNAc
2w
S
2x
O
S+
2y
O
S
2z
OSO3−
2aa
N3
2ab
N3
2ac
I
2ad
N
S
2ae
N
S
2af
HN
3
NH
HN
S
2ag

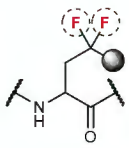


Histone H3-Dha9
15180 Da
3-5 mg protein



$FeSO_4 \cdot 7H_2O$ (100 eq.),
(5-10 eq), $Ru(bpy)_3Cl_2$ (2.5-5 eq.)

NH_4OAc (500 mM, pH 6, 3M Gdn·HCl)
50W blue LED, RT, 15 min



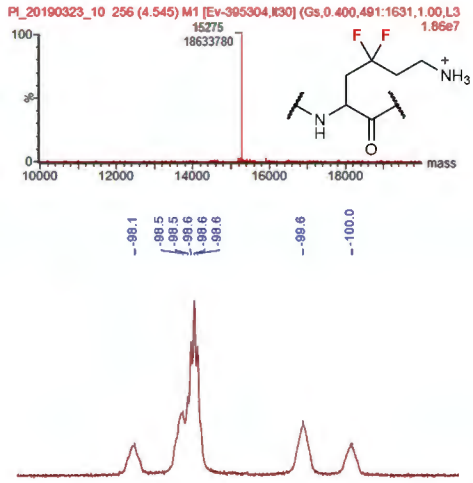
^{19}F-NMR active label

1) EDTA (5 mg) or DTT (10 mg)
vortexed for 30 s
2) miniTrap and midiTrap purification
in deuterated buffer

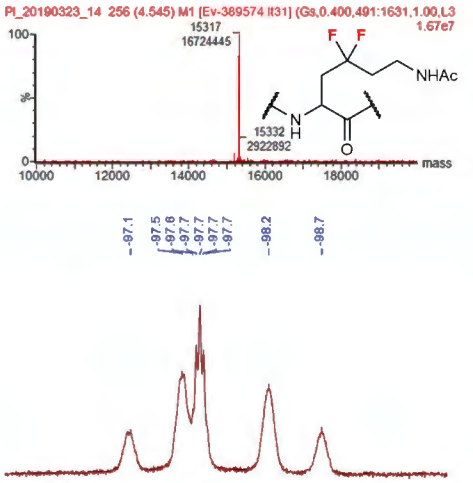
3) vivaspin column to 0.5 mL

^{19}F-NMR

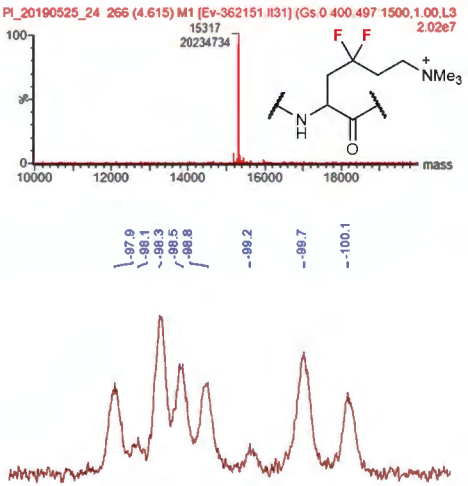
Histone H3-CF_2Lys9
>95% conv.
4.96 mg/mL (96% yield)
in NH_4OAc (250 mM, pH 7 in D_2O)



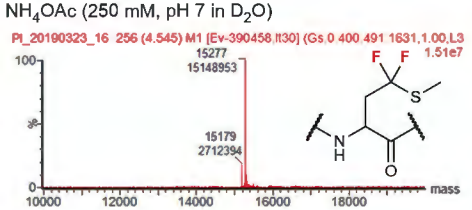
Histone H3-CF_2LysAc9
>95% conv.
4.82 mg/mL (95% yield)
in NH_4OAc (250 mM, pH 7 in D_2O)

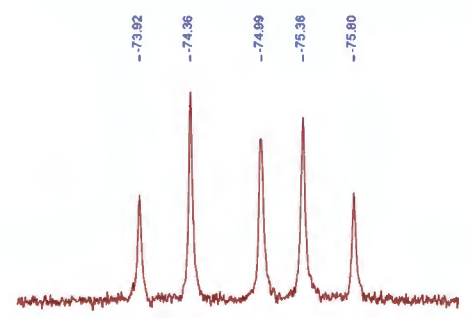


Histone H3-CF_2Lys(Me)$_3$9
>95% conv.
2 mg/mL (93% yield)
NH_4OAc (250 mM, pH 7 in D_2O)

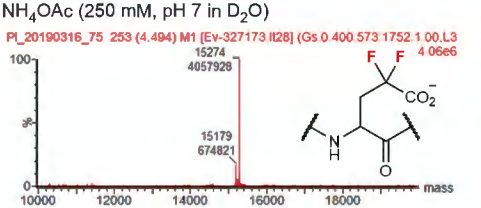


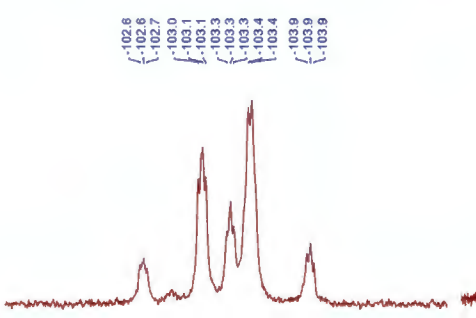
Histone H3-CF_2Met9
86% conv.
4.82 mg/mL (90% yield)
NH_4OAc (250 mM, pH 7 in D_2O)



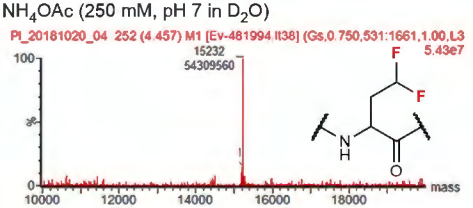


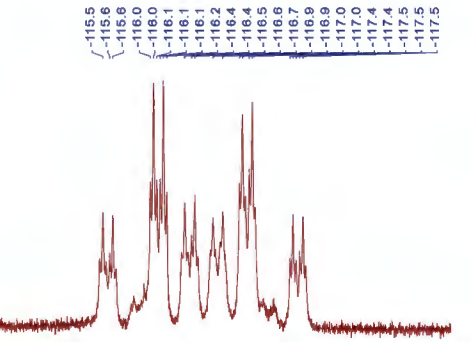
Histone H3-CF_2Glu9
86% conv.
4.22 mg/mL (91% yield)
NH_4OAc (250 mM, pH 7 in D_2O)

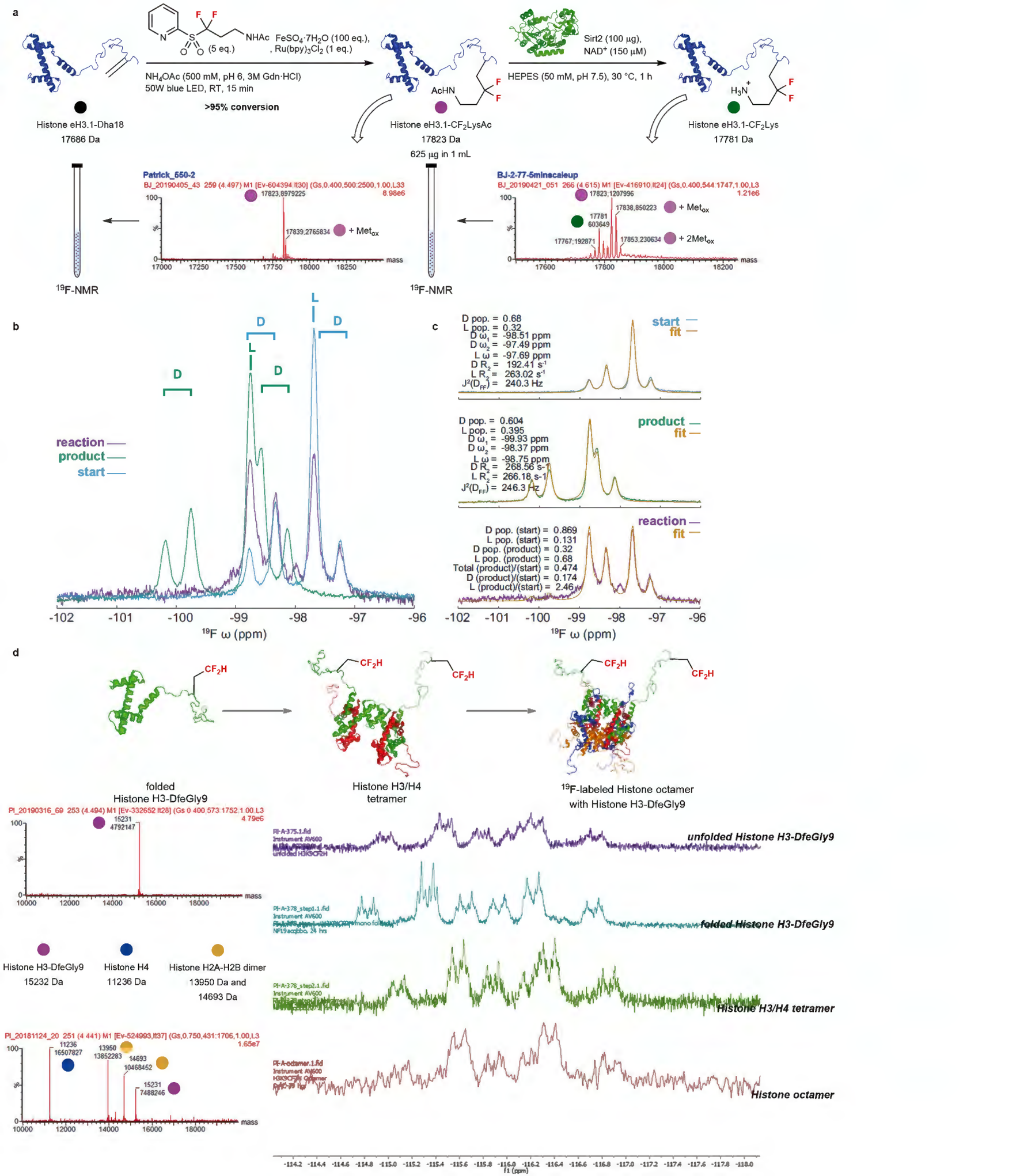


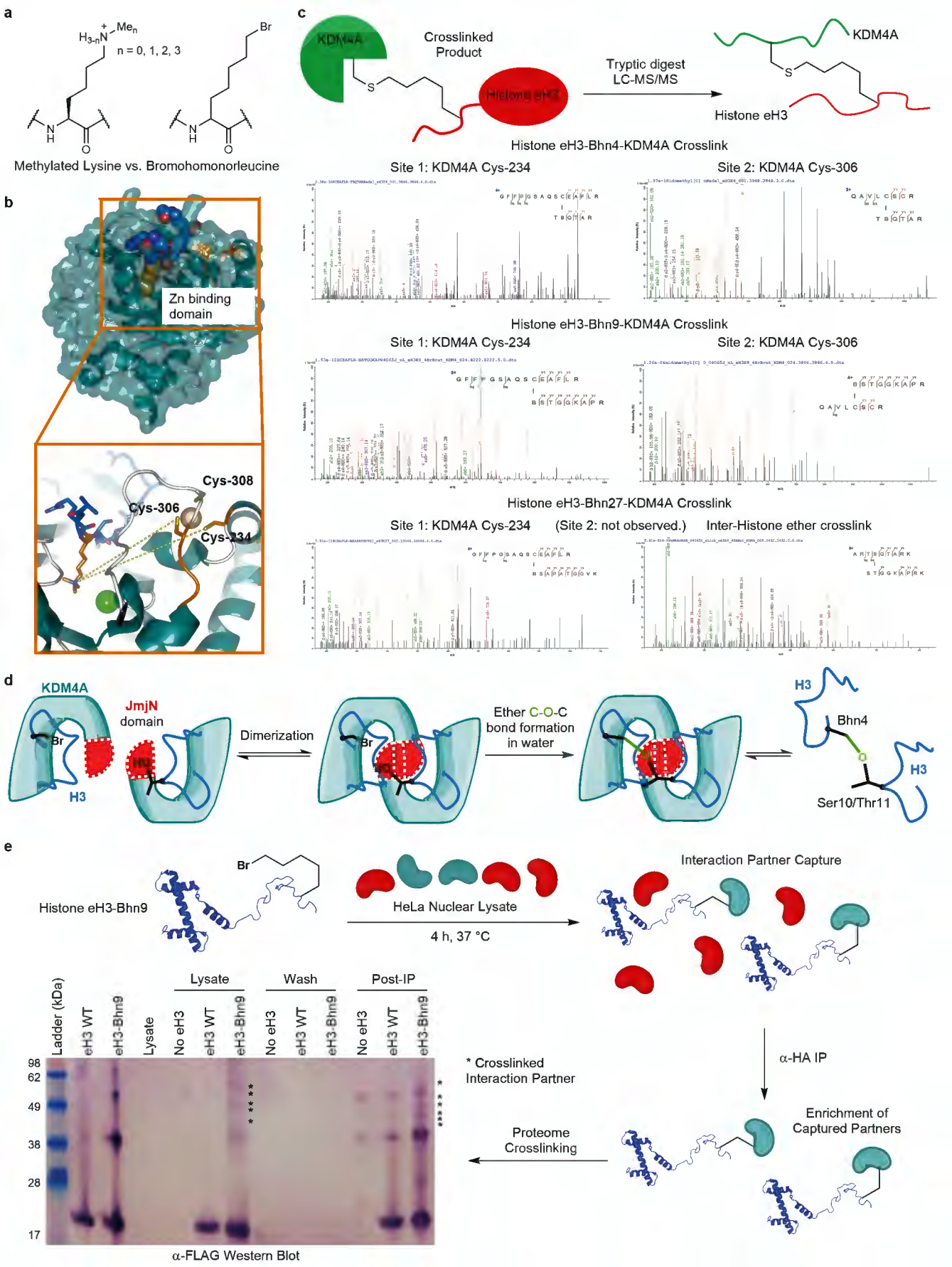


Histone H3-DfeGly9
>95% conv.
5.92 mg/mL (95% yield)
NH_4OAc (250 mM, pH 7 in D_2O)






a
(5 eq.)
NHAc
FeSO4·7H2O (100 eq.),
, Ru(bpy)3Cl2 (1 eq.)
NH4OAc (500 mM, pH 6, 3M Gdn·HCl)
50W blue LED, RT, 15 min
>95% conversion
Sirt2 (100 μg),
NAD+ (150 μM)
HEPES (50 mM, pH 7.5), 30 °C, 1 h
Histone eH3.1-Dha18
17686 Da
Histone eH3.1-CF2LysAc
17823 Da
625 μg in 1 mL
Histone eH3.1-CF2Lys
17781 Da
19F-NMR
19F-NMR
Patrick_650-2
BJ_20190405_43 259 (4.497) M1 [Ev-604394,It30] (Gs,0.400,500:2500,1.00,L33
6.98e6
17823,8979225
17839,2765834
+ Metox
mass
BJ-2-77-5minscaleup
BJ_20190421_051 266 (4.615) M1 [Ev-416910,It24] (Gs,0.400,544:1747,1.00,L3
1.21e6
17823,1207996
17838,850223
+ Metox
17781
603649
17767,192871
17853,230634
+ 2Metox
mass
b
reaction
product
start
D
L
D
L
D
D
-102 -101 -100 -99 -98 -97 -96
19F ω (ppm)
c
D pop. = 0.68
L pop. = 0.32
D ω1 = -98.51 ppm
D ω2 = -97.49 ppm
L ω = -97.69 ppm
D R2 = 192.41 s-1
L R2 = 263.02 s-1
J2(DFF) = 240.3 Hz
start
fit
D pop. = 0.604
L pop. = 0.395
D ω1 = -99.93 ppm
D ω2 = -98.37 ppm
L ω = -98.75 ppm
D R2 = 268.56 s-1
L R2 = 266.18 s-1
J2(DFF) = 246.3 Hz
product
fit
D pop. (start) = 0.869
L pop. (start) = 0.131
D pop. (product) = 0.32
L pop. (product) = 0.68
Total (product)/(start) = 0.474
D (product)/(start) = 0.174
L (product)/(start) = 2.46
reaction
fit
-102 -101 -100 -99 -98 -97 -96
19F ω (ppm)
d
CF2H
CF2H
CF2H
CF2H
CF2H
folded
Histone H3-DfeGly9
Histone H3/H4
tetramer
19F-labeled Histone octamer
with Histone H3-DfeGly9
PI_20190316_69 253 (4.494) M1 [Ev-332652 It28] (Gs 0 400 573 1752 1.00 L3
4 79e6
15231
4792147
mass
unfolded Histone H3-DfeGly9
folded Histone H3-DfeGly9
Histone H3/H4 tetramer
Histone octamer
Histone H3-DfeGly9
15232 Da
Histone H4
11236 Da
Histone H2A-H2B dimer
13950 Da and
14693 Da
PI_20181124_20 251 (4 441) M1 [Ev-524993,It37] (Gs,0.750,431:1706,1.00,L3
1.65e7
11236
16507827
13950
13852283
14693
10468452
15231
7488246
mass
-114.2 -114.4 -114.6 -114.8 -115.0 -115.2 -115.4 -115.6 -115.8 -116.0 -116.2 -116.4 -116.6 -116.8 -117.0 -117.2 -117.4 -117.6 -117.8 -118.0
f1 (ppm)


a
n = 0, 1, 2, 3
Methylated Lysine vs. Bromohomonorleucine
b
Zn binding domain
Cys-308
Cys-306
Cys-234
c
KDM4A
Crosslinked Product
Histone eH3
Tryptic digest LC-MS/MS
KDM4A
Histone eH3
Histone eH3-Bhn4-KDM4A Crosslink
Site 1: KDM4A Cys-234
Site 2: KDM4A Cys-306
Histone eH3-Bhn9-KDM4A Crosslink
Site 1: KDM4A Cys-234
Site 2: KDM4A Cys-306
Histone eH3-Bhn27-KDM4A Crosslink
Site 1: KDM4A Cys-234
(Site 2: not observed.)
Inter-Histone ether crosslink
d
KDM4A
JmjN domain
Br
H3
Dimerization
Ether C-O-C bond formation in water
H3
Bhn4
Ser10/Thr11
e
Histone eH3-Bhn9
Br
HeLa Nuclear Lysate
4 h, 37 °C
Interaction Partner Capture
α-HA IP
Enrichment of Captured Partners
Proteome Crosslinking
* Crosslinked Interaction Partner
Ladder (kDa)
eH3 WT
eH3-Bhn9
Lysate
No eH3
eH3 WT
eH3-Bhn9
Lysate
Wash
Post-IP
98
62
49
38
28
17
α-FLAG Western Blot

ED Figure Legends

Extended Data Figure 1 | Overview of approach to radical sidechain installation and relevant previous literature. a, Retrosynthetic analysis highlights the chemoselective advantages of using C· reagents paired with the radical acceptor Dha (right) over the typical heterolytic $2e^-$ reagents (left) for the site-specific modification of proteins. **b,** Our group's previous work on radical addition to Dha[3,17] and **c,** the mechanism highlighting unwanted side reactions (red). **d,f,** Summaries of previous work by others[24,25] using photocatalysts for '*on-peptide*' radical generation for site-selective peptide modification, highlighting their limitations and potential for undesired side reactions (red).

Extended Data Figure 2 | Complementary Strategies for Mild Protein-Compatible Photoredox Reactions. a, Oxidative half potential (E_{ox}) spectrum for catalyst compatibility with protein-based chemistry with relevant catalysts found in literature (**bottom and top**) and tested in this work (**top**). **b-c,** *In situ* formation of **BACED** reagents (for side-chains **1**, yellow highlight) advantageously allows [Ru^{II}]-catalyzed, low-E_{ox} activation (as compared to other derivatives) to RC**H_2**• radicals that then react with Dha in proteins to install side-chains. Independent and mixed voltammetric responses of 1 mM catechol and 12 mM phenethylboronic acid on GC in PBS, pH 7.10 (**inset c**). See also **ED Figure 4** and **Supplemental Discussions 2,3** on mechanism for more detailed electrochemical experiments. Intact protein LC-MS (**bottom right**, chromatogram and *m*/*z*) shows homohomophenylalanine (**1h**) installation into Histone H3 protein. **d,** [Ru^{II}]-catalyzed activation of **pySOOF** reagents to RC**F_2**• radicals that then react with Dha in proteins to install 'zero-size'-labelled side-chains. Added [Fe^{II}] drives unprecedented efficiency (2-5 equiv of precursor) by suppressing oxidation by [Ru^{II}]* to imine (and hydrate) that suggests a key role as a reductant (readily-available in Biology) that quenches the alpha-C• radical adduct generated during the reaction. Intact protein LC-MS shows difluoroethylglycine (DfeGly, **2a**) installation into Histone H3 protein is successful with [Fe^{II}] (full conversion, **top right**, chromatogram and *m*/*z*) yet not without iron (poor conversion to unwanted side products, **bottom centre**), see also **ED Figures 5-7** for further details. For the full reaction scope of all side-chains (types **1** and **2**) edited into proteins, including those allowing previously inaccessible *on-protein* reactivity see **ED Figure 8** and also **Figure 2** and **Figure 3**).

Extended Data Figure 3 | Investigation and optimization of BACED chemistry. a, A 100% stacked bar chart (*n*=1, with single data values represented by the y-axis span of the corresponding bars) showing the results of initial studies into the oxidation of benzyltrifluoroborate with different catalyst strengths and additives to achieve selective single addition. Catechol increased reactivity for all catalysts, leading to the emergence of reactivity with the weakest **Cat1** and almost complete conversion to the double addition product with **Cat3**, while $NaCNBH_3$ with **Cat3** successfully quenched the alpha-carbon radical promoting single addition. **b,** Trends of increasing oxidative damage and decreasing reaction control with higher $*E_{ox}$ catalysts with representative LC-MS ion series and spectra (see **Supplementary Methods** for details). **c,** A 100% Stacked bar chart (*n*=1, with single data values represented by the y-axis span of the corresponding bars) for comparing **Cat1** and **Cat2** reaction conversions to single addition product and unwanted alpha carbon radical quenching (via either double addition or catechol quenching) for different boronate substrates, arranged with higher E_{ox} primary boronate substrates on the left and increasingly stabilized radical precursors on the right. The trends clearly show the utility or necessity of using the stronger Cat2 (*Eox = +0.99 V) for the primary boronate substrates, and the increased efficiency of using the weaker Cat1 (*Eox = +0.77 V) for the more stabilized radical precursors (secondary and benzyl substrates), likely due to a slow, controlled release of stable radicals, ensuring efficient addition to Dha instead of self-quenching or overalkylation. **d,** A light intensity screen shows that increasing light intensity (450 Blue LED, 0-50 W) allows for high conversions to the desired single addition product with shorter reaction times and less equivalents of **Cat1** (**Supplementary Table 5**, *n*=1, line of best fit overlayed). **e,** Reaction scheme and LC-MS spectra for the installation of Iodonorleucine (Inl). While Inl installation had poor initial conversion (~50%), the only other species present after the reaction was the starting material (Dha), allowing successive reactions to push conversions to a more reasonable ~75%. **f,** A screen of catechol derivatives finds that the naturally occurring catecholamines dopamine and L-DOPA could efficiently substitute for the role of catechol.

Extended Data Figure 4 | Mechanistic investigation into the role of catechol in BACED reactions. a, Reaction scheme screening different quinone derivatives for their influence on the oxidation of potassium phenethyltrifluoroborate and subsequent addition to Dha. **b,** Potential mechanism of catechol derivatives acting as redox mediators to bridge electron transfer between catalyst and substrate. **c,** LC-MS results of the quinone derivative screening (described in **a** and **b**) ruled out the mechanism in **b** as only 1,2-diols showed significant activity, with only catechol avoiding protein degradation. **d,** A potential mechanism of *in situ* catalyst modification with catechol, creating catecholo-Ru(bpy)$_2$ (**Cat6**)[31] was ruled out as it did not promot alkylation with or without the addition of exogenous catechol. **e,** The *in situ* formation of a reactive boronic acid catechol ester is suggested. **f,** 1-propylboronic acid catechol ester and 4-bromobutylboronic acid catechol ester successfully added without the addition of exogenous catechol. This suggests that the formation of the catechol ester lowers the E_{ox} value of the substrate to a range accessible by **Cat1** ($^*E_{ox}$ = +0.77). **g,** Voltammetric response of 1 mM catechol in the presence of increasing concentrations of 4-bromobutylboronic acid (black, 0 mM; brown, 3 mM; green, 6 mM; blue, 12 mM; red, 24 mM) at a glassy carbon macroelectrode in 50 mM phosphate buffer (pH 6) recorded at 100 mVs^{-1}. **h,** Voltammetric response of catechol only (1 mM, black), 4-bromobutylboronic acid (12 mM, blue), and 4-bromobutylboronic acid in the presence of catechol (12 mM and 1 mM, respectively, red) recorded at 100 mVs^{-1}. **i,** A simulated voltammetric response for the oxidation of 4-bromobutylboronic acid (12 mM) in the presence of catechol (1 mM) at 100 mV s^{-1}, following the simplified mechanism outlined in **e**. Simulation highlights the importance of the oxidation of the boronic acid ester being catalytic and leading to the reformation of catechol (See **Supplementary Discussion 2,3**). The rate of decomposition of the radical ester has been set at either 1x10^4 or 0 s^{-1} (blue and red respectively). If the oxidation of the boronic acid ester is not catalytic, no peak is predicted to be voltammetrically observable at > 1V vs SCE. **j,** Voltammetric response of 4-bromobutylboronic acid (12 mM) in the presence of catechol (1 mM) in 50 mM phosphate buffer (pH 6) recorded as a function of scan rate (25−400 mV s^{-1}). **k,** Voltammetric response of the preformed 4-bromobutylboronic acid catechol ester (1 mM) and **l,** catechol (1 mM) in 50 mM phosphate buffer (pH 6) recorded as a function of scan rate (25−400 mV s^{-1}).

Extended Date Figure 5 | Initial experiments without iron using various hydride sources and optimization study with sodium borohydride for pySOOF. a, The formation of oxidative-derived side products H3-imine-DfeGly9 and H3-hemiaminal-DfeGly9 were detected in an initial, additive-free, light-driven protein modification reaction with pySOOF and $Ru(bpy)_3Cl_2$. Based on this observation, a mechanism was postulated, where the on-protein radical intermediate was oxidized by the excited-state of the photocatalyst. To avoid this reaction pathway, the use of hydride donors such as silanes, tertiary amines or borohydrides were postulated to favour the formation of DfeGly-modified histone at site 9. **b,** Inferior results were observed when 250 eq. of DIPEA, TTMS, and catechol (100 eq.) were used (with 200 eq. pySOOF and 5 eq. $Ru(bpy)_3$) (**Supplementary Table 8, 11, 12**). For catechol, excellent conversions were observed, however a catechol-DfeGly-modified histone was detected as a major side product. Borohydrides such as $Na(OMe)_3BH$ (0.25 mg) or NEt_4BH_4 (0.25 mg) showed promising reactivity, but only the oxidation-derived side products were formed. Only with sodium borohydride, the desired DfeGly-modified histone was formed in moderate conversion (**Supplementary Table 8**). **c**, Reduction of Dha to Ala by $NaBH_4$ was identified as potential limitation for the $NaBH_4$ mediated reaction. To test this undesired pathway, a crude reaction mixture was buffer exchanged to NaPi (100 mM, pH 9 3M GdnHCl) and incubated with a large excess of 2-mercaptoethanol. After 1 hour at 37 °C, no formation of the corresponding thio-Michael protein adduct was detected, proofing Ala formation. The 100% stacked bar graphs (*n*=1, with single data values represented by the y-axis span of the corresponding bars) summarized the results of the optimization studies into the $NaBH_4$-mediated photochemical reaction with different equivalent of $NaBH_4$, photocatalyst, pySOOF and reaction times (**Supplementary Table 9-10**). <500 eq. $NaBH_4$ resulted in formation of undesired oxidative side products. Increasing reagent concentration, photocatalyst loading or prolonged reaction times did not improve the reaction (The results are represented in a 100% stacked bar chart (*n*=1) with single data values represented by the y-axis span of the corresponding bars). **d,** Lower temperature increased the conversion, most likely by slowing down the reduction of Dha by $NaBH_4$.

Extended Date Figure 6 | Optimization study of the iron(II) mediated protein modification reaction with pySOOF. a, The unique role of iron(II) for the photochemical modification reaction with pySOOF as a reductive quencher for the photoredox-cycle and single-electron reductant of the on-protein radical intermediate forming the enolate intermediate. However, side products such as H3-imine-DfeGly9, H3-hemiaminal-DfeGly9 or H3-diDfeGly9 could be generated because of inefficient quenching of the on-protein radical intermediate. **b**, In an initial experiment with 200 eq. pySOOF, 280 eq. $FeSO_4$, 5 eq. $Ru(bpy)_3Cl_2$ and 66 µM Histone H3-Dha9 with a protein concentration of 0.5 mg/mL, 70% conversion to a mixture of H3-DfeGly9 and H3-diDfeGly (57:43) was observed. The formation of the mono-addition product was favoured at higher protein concentration (1 mg/mL) and the conversion was increased to 92% (**Supplementary Table 13**). **c,** For the iron(II)-mediated reaction, Various metalo- and organophotocatalysts with different $*E_{red}$ values (-0.56 to -1.37 V) and radical precursors were tested and $Ru(bpy)_3$ and pySOOF were identified as the best combination (Supplementary Table 17). **d,** The 100% stacked bar charts (*n*=1, single data values represented by the y-axis span of the corresponding bars) summarize the results of the optimization studies into the $FeSO_4$-mediated photochemical reaction with different reaction times, equivalents of pySOOF and $FeSO_4$, and catalytic amounts of iron(II) and photocatalyst, respectively (**Supplementary Table 15-10, 18, 20**). Fast reaction times and high efficiency with low equivalents of pySOOF were found. Albeit in cases with high levels of reactivity with <100 eq. $FeSO_4$, only oxidation-derived products were created, indicating the dual role of iron(II) for the single-electron reduction of the on-protein radical intermediate. **e,** With the optimized conditions in hand, good to excellent conversions to Histone H3-DfeGly9 were obtained at various protein concentrations (0.1-5 mg/mL).

Extended Date Figure 7 | Investigations on pySOOF reagent reactivity and on-protein mechanism. a, Effect of substituents on the alpha centre of the created radical on reactivity, efficiency and selectivity (mono vs. double addition). No reactivity was observed for non- and mono-fluorinated precursors with an H, Me or Cl substituent. With an increase in the electron-withdrawing effect of the additional substituent, radical generation was observed and the resulting products were formed. Moreover, the highest levels of reactivity were observed for $\bullet CF_2R$. **b**, An equimolar mixture of three pySOOF substrates were applied to the protein modification reaction to assess the reactivity order between $\bullet CF_2H$, $\bullet CF_2Me$ and $\bullet CFHCONH_2$. During this competition experiment, the following reactivity trend was identified: $\bullet CF_2H > \bullet CF_2Me$ and no product formation was observed for $\bullet CFHCONH_2$. **c,** Based on the reactivity study of various mono- and difluoro-pySOOF reagents, a suitable radical precursor was designed to allow the efficient generation of $\bullet CF_2R$ for the installation of difluorinated- amino acid residues or PTMs. **d,** With iodo-pySOOF, homolytic bond cleavage between iodine and carbon centre was induced and the pySOOF unit was installed on the Dha-tagged histone (**Supplementary Table 22-23**). After further photoredox-activation, a captodative effect stabilized on-protein radical was formed what allowed further on-protein chemical reactions by trapping the radical with various acceptors (**Supplementary Table 24-37**). Moreover, the homolytic iodo-carbon bond cleavage inspired the design of bromo-difluoro carbonyl based radical precursors for the installation of CF_2Gln or CF_2Glu derived amino acids. **e,** Mechanistic studies proved a radical mechanism as no product formation was observed with TEMPOL. Furthermore, a photocatalytic process was proved as no conversions were detected without photocatalyst nor blue LED irradiation (**Supplementary Table 16**). On-off experiments showed only product formation during the irradiation period, excluding a radical chain mechanism. Finally, good reactivity (80% conv.) was observed when the sample was prepared under ambient atmosphere, however, high levels of oxidative damage (Met oxidation) and oxidative-derived DfeGly-modified Histones were formed.

Extended Data Figure 8 | Substrate Scopes for BACED and pySOOF. a, A comprehensive list of all sidechains installed into proteins with the BACED reaction manifold. General conditions: protein 1 $mg.mL^{-1}$; 50W/450nm light; 4˚C-to-RT; 100-1500 equiv. BACED-precursor reagent; 10 equiv. **Cat1** or **Cat2**; 100 equiv. catechol; <6ppm O_2; pH 6.0 buffer (500 mM NH_4OAc or PBS ± 3M Gdn•HCl). **b,** A comprehensive list of all sidechains installed into proteins with the pySOOF reaction manifold. General conditions: protein 1 $mg.mL^{-1}$; 50W/450nm light; RT; 2-5 equiv. pySOOF-precursor reagent; 0.4-4 equiv. **Cat1**; 50-100 equiv. $FeSO_4$; <6ppm O_2; pH 6.0 buffer (500 mM NH_4OAc or various other buffers).

Extended Date Figure 9 | Upscaling of the protein modification with pySOOF and ^{19}F-NMR analysis. The scalability of protein modification reaction with pySOOF was studied and high conversions were observed for pySOOF-Lys, LysAc, Lys(Me)$_3$, Met, Glu and DfeGly using 4-6 mg of Dha-tagged Histone. After the photochemical reaction, the crude mixture was vortexed with either EDTA or DTT followed by buffer exchange to deuterated buffer (NH_4OAc, 250 mM, pH 7 in D_2O with TFA as internal standard) using desalting columns. After purification, excellent yields were determined by checking the protein concentration by nanodrop. Spectra were recorded on Bruker NMR (AV600) and every single PTM showed characteristic peaks with unique chemical shifts.

Extended Date Figure 10 | Application of difluorinated amino acid labelled proteins in ^{19}F-NMR studies. a, Milligram-scale chemical mutagenesis to generate the substrate, Histone eH3.1-CF_2LysAc18, to investigate the deacetylation by Sirt2 by comparing the ^{19}F-NMR analysis of the starting material and after enzymatic reaction. The desired difluoro-labelled histone was formed in excellent conversion, with low levels of Met oxidation. After the deacetylation, Histone eH3.1-CF_2LysAc18 and Histone eH3.1-CF_2Lys were detected by LC-MS, proofing that the fluorinated histone was accepted as substrate for Sirt2. **b-c,** ^{19}F-NMR spectra were acquired for 'start' and 'product' states, as well as when the solution containing 'start' was added to the enzyme Sirt2 (reaction). The spectra of the start and product states consisted of an AB quartet (D), and an apparent singlet state (L). The spectrum of the reaction contained the singlet state of the product, and the AB quartet of the start, suggesting that the singlet state almost completely reacted with Sirt2, and the species that gives rise to the AB quartet has not. To quantify this, the 'start' and 'product' spectra were simulated to determine relevant resonance frequencies and coupling constants, and the 'reaction' spectrum was simulated, taking the species identified in 'start' and 'product', but with scaled intensities. **d**, Histone H3-DfeGly9 sample had its ^{19}F-NMR spectra recorded as an unfolded protein species in unfolding buffer (7 M GdnHCl, 10 mM Tris, 1 mM EDTA, 10 mM DTT, 1 mM Benzamidine, pH 7.5, 50% D_2O, 0.1 µL trifluoroethanol internal standard) folded species in Tris buffer (150 mM NaCl, 10 mM Tris, 1 mM EDTA, 2 mM βME, pH 7.5, 50% D_2O, 2 mg/mL protein, 0.75 mL, 1 µL of 0.1% trifluoroethanol internal standard), reconstituted Histone H3-H4 tetramer in refolding buffer (2.5 mg/mL tetramer, 1 mL buffer, 50% D_2O, 0.1 µL trifluoroethanol internal standard), and reconstituted Histone H2A-H2B-H3-H4 octamer in refolding buffer (1.6 mg/mL octamer, 0.5 mL buffer, 50% D_2O, 0.1 µL trifluoroethanol internal standard). All the recorded ^{19}F-NMR spectra were compared and small changes in the chemical shifts were observed. Interestingly, the biggest changes were detected between the unfolded and folded histone H3 and after tetramer formation, potentially due to changes in tumbling rate. LC-MS analysis of the octamer product confirmed the presence of all four histone types.

Extended Data Figure 11 | Effective molarity driven Protein-Protein crosslinking with electrophile containing sidechains. a, A comparison of the connectivity of a modified lysine sidechain (via methylation) and bromohomonorleucine (Bhn, **1u**) used for crosslinking. **b,** Structure of Histone eH3.1 mimic peptide binding to KDM4A. Identified cysteines undergoing crosslink (Cys-306 and Cys-234) are highlighted in orange and their distance to H3-K9 highlighted. **c,** Workflow of the crosslinking reaction products and their identification by LC-MS/MS. Representative spectra are shown for each crosslinking site (top: eH3.1-Bhn4, middle: eH3.1-Bhn9, bottom: eH3.1-Bhn27) and captured cysteine residue. The bottom right spectra shows an inter-histone crosslink via ether formation as found in in-solution digest. **d,** Unprecedented Williamson C–O–C bond ether formation in an inter-molecular fashion between H3 proteins (Bhn4 in one linked to hydroxyl in another) is driven by effective molarity, possibly suggesting a transient dimer model for KDM4A function. **e,** The Histone eH3.1-Bhn9 alkylator protein was incubated with HeLa nuclear lysate to capture interaction partners via promixity-driven crosslinking. After an enrichment via the HA-tag (on Histone eH3.1), an α-FLAG western blot reveals multiple higher MW bands corresponding to the mass of the histone plus that of the captured interaction partner. No higher MW bands were seen in conditions lacking Bhn.