Site-selective installation of an electrophilic handle on proteins for bioconjugation

Bernadette Lee [a,d], Shuang Sun [a,d], Ester Jiménez-Moreno [a], André A. Neves [b], Gonçalo J.L. Bernardes [a,c,*]

[a] Department of Chemistry, University of Cambridge, Lensfield Road, CB2 1EW Cambridge, UK
[b] Li Ka Shing Centre, Cancer Research UK Cambridge Institute, Robinson Way, CB2 0RE Cambridge, UK
[c] Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Avenida Professor Egas Moniz, 1649-028 Lisboa, Portugal

ABSTRACT

Site-selective protein modification strategies can be used to insert non-natural functional groups into protein structures. Herein, we report on the use of the bis-electrophile 3-bromo-2-bromomethyl-1-propene as a reagent to introduce an electrophilic handle at cysteine residues under mild conditions. This method is demonstrated on a variety of proteins containing a solvent-exposed cysteine residue, including an anti-HER2 nanobody. Chemically distinct protein conjugates are then efficiently formed through further reaction of the electrophilic site with various nucleophiles, including thiols and amines. The resulting chemically-defined conjugates are highly stable in the presence of glutathione or human plasma and retain both the structure and function of the native protein.

© 2018 Published by Elsevier Ltd.

1. Introduction

Site-selective methods for protein modification[1] are important for conjugating useful synthetic molecules such as fluorescent probes[2–4] and drugs[5–7] to proteins in a well-defined manner. Techniques to achieve selective protein labelling include both the modification of natural amino acids and the introduction of unnatural amino acids via genetic expansion to install bioorthogonal chemical handles.[8–10] *In vitro*, cysteine is often targeted for protein bioconjugation due to its low natural abundance and high nucleophilicity.[11–14] A high degree of chemoselectivity can be achieved through careful control of the reaction conditions to enable cysteine to react preferentially over other nucleophilic residues such as histidine and lysine.[15–17] For proteins lacking a cysteine residue, site-directed mutagenesis can be used to introduce a single cysteine at a user-defined site within the protein sequence,[18,19] allowing this method to be applicable to a wide range of proteins.

Cysteines can be converted to suitable orthogonal handles for further derivatization to provide access to a diverse range of proteins functionalized at specific sites.[9,20,21] An electrophilic motif, which is not found naturally in proteins, may be incorporated and subsequently conjugated using various nucleophiles.[22,23] For example, a cysteine can be selectively converted to dehydroalanine,[24] which then serves as a handle for further functionalization via thiol-[25,26] and aza-Michael addition[27] or carbon-based radical mechanisms.[28] An alternative method to incorporate an electrophilic handle is by using bis-electrophilic reagents such as dibromomaleimide[29] or dibromopyridazinedione.[30] However, these linkers have been shown to be unstable in the presence of excess thiols, making them only useful for the temporary modification of cysteine residues. Moreover, these linkers could be used to form ubiquitin-protein conjugates.[31]

With this in mind, our group has previously shown that it is possible to introduce an electrophilic handle at a cysteine residue via the homobifunctional electrophile 3,3-bis(bromomethyl)oxetane (Fig. 1a)[32] under relatively harsh reaction conditions, up to 37 °C and pH 11.0. However, unlike the dibromomaleimide and dibromopyridazinedione reagents, the conjugates formed from the site-selective bis-alkylation using the oxetane reagent are stable in the presence of biological thiols such as glutathione (GSH). In order to take advantage of the increased S_N2 reactivity exhibited by allylic systems, in this work we investigated the use of 3-bromo-2-bromomethyl-1-propene **1** for alkylating proteins at cysteine residues (Fig. 1b). This electrophilic handle was then further modified by nucleophiles to yield chemically-defined protein conjugates. The high reactivity of this alkylation reagent for

* Corresponding author at: Department of Chemistry, University of Cambridge, Lensfield Road, CB2 1EW Cambridge, UK.
E-mail addresses: gb453@cam.ac.uk, gbernardes@medicina.ulisboa.pt (G.J.L. Bernardes).
[d] These authors equally contributed to this work.
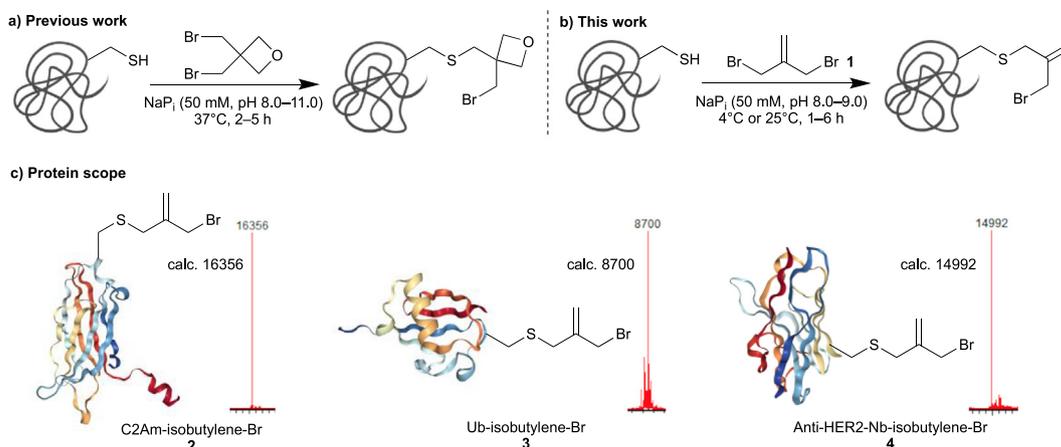
Fig. 1. Incorporating an electrophilic tag into proteins at cysteine. a) Previous work to introduce an electrophilic handle containing an oxetane motif; b) This work, where an isobutylene-Br group was introduced under mild reaction conditions; c) Protein scope of the reaction with C2Am, Ub-K63C and anti-HER2 nanobody, containing an engineered solvent-exposed cysteine residue.

disulfide re-bridging has been demonstrated in previous work by our group.[33]

2. Results and discussion

2.1. Installing the electrophilic handle on proteins

We began investigating the viability of the cysteine alkylation reaction using proteins containing a single engineered cysteine residue. The engineered versions of the C2A domain of synaptotagmin-I (C2Am)[34] and ubiquitin (Ub-K63C)[35] were chosen to target the solvent-exposed cysteine residues that we envisioned would be highly reactive with **1**. Indeed, the reaction of C2Am with 100 equivalents of **1** at pH 9.0 proceeded efficiently at room temperature. Analysis of the reaction by Liquid Chromatography-Electrospray Ionisation-Mass Spectrometry (LC-ESI-MS) after 1 h showed the complete conversion (>95%) of C2Am to the desired alkylation product **2** (Figs. 1c and S1). Similarly, incubation of Ub-K63C with 50 equivalents of **1** at pH 8.0 at 4 °C afforded the alkylation product **3** in good conversion after 6 h (Figs. 1c and S10). The chemoselectivity of the reaction was confirmed using Ellman's test: the protein conjugates remained unchanged while the unmodified proteins readily reacted with Ellman's reagent (Figs. S3 and S12), indicating that the cysteine residue had been modified selectively.

After obtaining these preliminary results, we proceeded to test **1** against an anti-HER2 nanobody containing both an engineered, unpaired cysteine residue and an internal disulfide. Based on previous studies demonstrating the ability to modify similar nanobodies without affecting the internal disulfide,[36] we were hopeful that under the same mild reaction conditions used for C2Am and Ub, site-selective modification of the anti-HER2 nanobody would be possible. While minimal reaction was observed using the same conditions that were used to achieve full conversion on C2Am (100 equivalents of **1**, 25 °C, pH 9.0), an increase to 1000 equivalents of **1** resulted in the single conjugate **4** after 1 h, as verified by LC-ESI-MS (Figs. 1c and S19).

Interestingly, during optimisation of the alkylation step for Ub-K63C, we observed that the alkylation product degraded to a product at 8618 Da (Supplementary Data S17). This reaction occurred relatively rapidly at room temperature. However, it was partially suppressed both by conducting the reaction at 4 °C and by keeping reaction times relatively short. We propose that this product forms from the reaction of a lysine[37] on Ub-K63C with the second S_N2

reactive site on **1** via an intramolecular "stapling" reaction. Due to the proximity effect, similar intramolecular reactions between an electrophilic, unnatural amino acid and nucleophilic residues such as cysteine and lysine have been reported.[38–40] Despite this, Ub-K63C conjugates with multiple linkers incorporated were not observed, indicating that the intermolecular reaction between **1** and Ub-K63C proceeds selectively at cysteine.

2.2. Conjugating with nucleophiles

Having determined that the alkylation step reliably introduces **1**, the isobutylene-Br handle, into proteins at cysteine residues, we proceeded to react the single conjugates (**2–4**) with a range of thiol nucleophiles.[32] Treatment of **2** with β-mercaptoethanol (BME), thiophenol (PhSH) and β-D-thioglucose sodium salt (βGluSNa) at room temperature afforded the desired products **2a-2c** after 30 min to 1 h (Figs. 2 and S4–6). These results indicated that the second electrophilic site on **1** was preserved following the alkylation step. Under similarly mild conditions, reacting **3** and **4** with BME, PhSH, and βGluSNa afforded the chemically-defined conjugates **3a–3c** (Figs. S13–15) and **4a–4c** (Figs. S22–24) further highlighting the high reactivity of the allylic system. However, in the reactions with **3** and **4**, analysis by LC-ESI-MS showed that a small amount of the single conjugate was converted to the unmodified protein.

Further tests using **2c** showed that the isobutylene linker in the conjugated proteins was highly stable when incubated with the endogenous thiol GSH[41] or human plasma for 24 h at 37 °C (Figs. S8 and S9). This demonstrates the possible use of this conjugation method to produce homogenous conjugates for *in vivo* applications.

We also investigated the reaction of **2–4** with amines, using benzylamine as a representative amine nucleophile capable of reacting under mildly basic conditions and in the presence of disulfides on proteins (pH 8.0–9.0). Owing to the reduced reactivity of amine groups as compared to the thiol nucleophiles, reaction times of up to 6 h were required to obtain the conversion to the conjugated products **2d–4d** (Figs. 2, S7, S16 and S25). Some hydrolysis of the bromide was observed in all cases but was least pronounced when using **2**. Some "stapling" was also observed with **3**. While the nucleophile scope can be expanded to encompass amines, these findings indicate that the inherent reactivity of the linker makes it prone to hydrolysis, which becomes more noticeable when less reactive nucleophiles are used. This result also lends support to
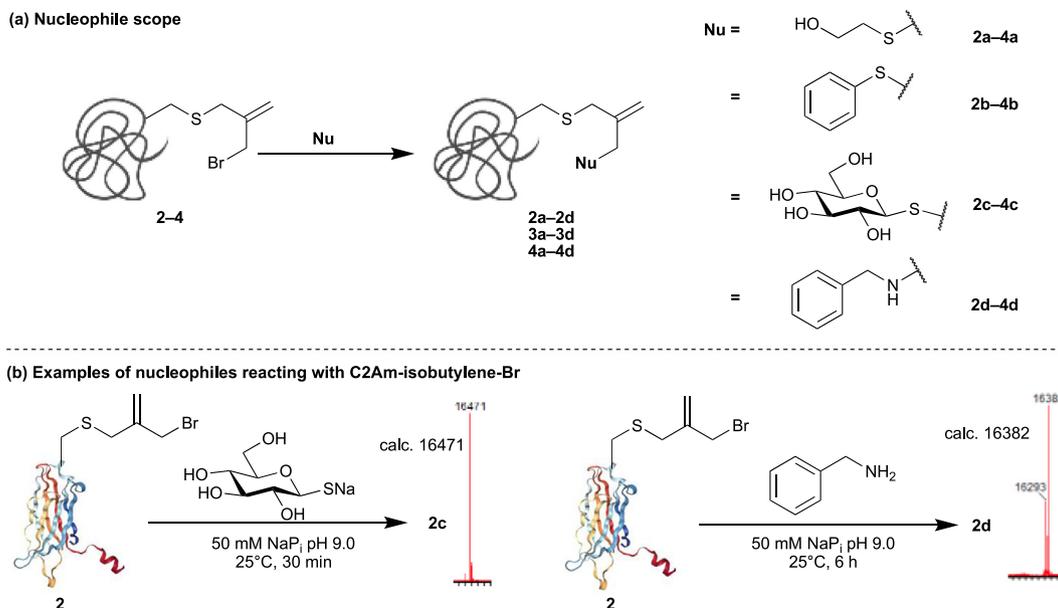
Fig. 2. Reaction of nucleophiles with the electrophilic handle incorporated into C2Am, Ub-K63C and anti-HER2 nanobody. a) Range of nucleophiles tested, which included both thiol and amine nucleophiles; b) Representative examples of the reactions with C2Am, highlighting the differences in reactivity between the thiol and amine nucleophiles.

our hypothesis that an appropriately positioned lysine group in Ub-K63C reacts with the linker to form a "stapled" product.

With the C2Am-Glc conjugate **2c** in hand, we wanted to investigate the impact of this conjugation method on the protein's structure and biological function. The highly similar circular dichroism (CD) spectra indicated that the conjugate **2c** maintained its secondary structure (Fig. 3a). Additionally, the surface plasmon resonance (SPR) analysis showed that the C2Am-Glc conjugate **2c** retained a strong binding activity against phosphatidylserine (PS), proving to be only slightly lower when comparing to the native form of C2Am.

3. Conclusions

We developed a site-selective method for installing a small electrophilic isobutylene handle onto proteins at solvent-exposed cysteine residues through the reaction with a homobifunctional linker. This conjugation proceeds rapidly with complete chemoselectivity under mild reaction conditions. Optimisation of the reac-

tion conditions allowed the reaction to proceed efficiently with the range of proteins tested. The incorporated electrophilic site proved to be highly reactive, enabling further conjugation of the modified proteins with either thiol or amine nucleophiles. Moreover, the native structure and biological function of the proteins were retained in the conjugates. Based on this proof-of-concept study, we believe that this operationally simple and efficient site-selective conjugation method will find wide use in the field of protein engineering for producing small molecule conjugated proteins.

4. Experimental

4.1. General methods

Liquid chromatography electrospray ionisation mass spectrometry (LC-ESI-MS) was performed on a Waters SQ Detector 2 mass spectrometer coupled to an Acquity UPLC system (Acquity UPLC BEH300 C4 column, 1.7 μm, 2.1 mm × 50 mm). Water with 0.1% formic acid (solvent A) and 70% acetonitrile and 29% water with
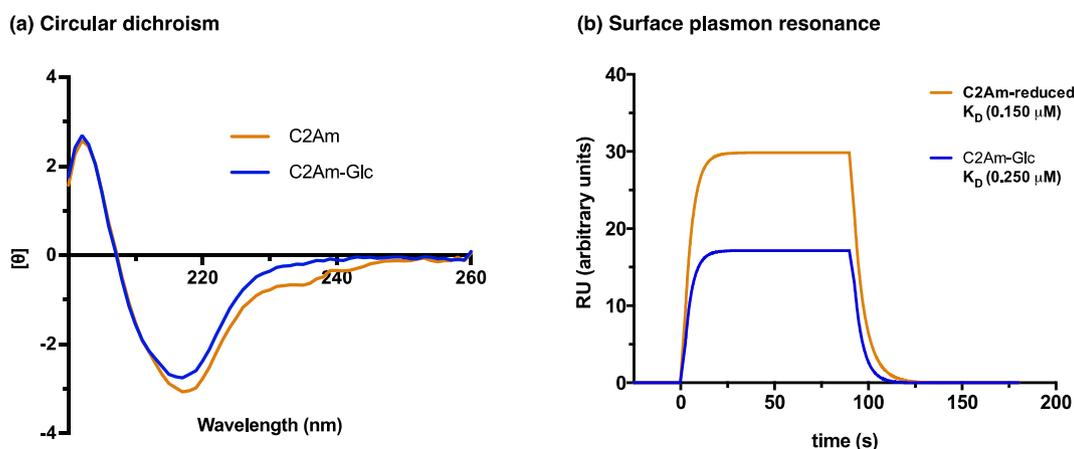
Fig. 3. a) Circular dichroism (CD) spectroscopy of C2Am, yellow, and C2Am-Glc conjugate **2c**, blue, average of 3 scans; b) Biacore surface plasmon resonance (SPR) analysis of C2Am, yellow, and C2Am-Glc conjugate **2c**, blue.

0.075% formic acid (solvent B), were used as the mobile phase at a flow rate of 0.2 mL min^{-1}. The gradient was programmed as follows: From 72% A to 100% B for 13 min then 100% B for 3 min and 72% A for 4 min. The electrospray source was operated with a capillary voltage of 3.0 kV and a cone voltage of 20 V. Nitrogen was used as the desolvation gas at a total flow of 800 L·h^{-1}. Total mass spectra were reconstructed from the ion series using the MaxEnt algorithm preinstalled on MassLynx software (v. 4.1 from Waters) according to the manufacturer's instructions.

4.2. Electrophilic handle installation

4.2.1. Procedure for C2Am
A 50 μL aliquot of a solution of C2Am (20 μM in 50 mM NaP$_i$ pH 9.0, 1 nmol) was added to 49.5 μL of 50 mM NaP$_i$ pH 9.0. To this was added 0.5 μL of a solution of tris(2–carboxyethyl)phosphine hydrochloride (TCEP·HCl) (100 mM in H$_2$O, 50 nmol, 50 eq.). The reaction mixture was vortexed briefly, then shaken at 25 °C for 45 min. 2.5 μL of a solution of 3-bromo-2-bromomethyl-1-propene (40 mM in DMF, 100 nmol, 100 eq.) was then added and the resulting mixture was vortexed briefly. After 1 h of additional shaking at 25 °C, a 10 μL aliquot was analysed by LC-ESI-MS. The expected conversion to C2Am-isobutylene-Br was observed (calculated mass: 16356 Da; observed mass: 16356 Da). Removal of small molecules and buffer exchange into 50 mM NaP$_i$ pH 9.0 or 50 mM NaP$_i$ pH 11.0 were effected by Vivaspin 500 protein concentrators (5k MWCO, GE Healthcare). Purified samples were stored at −20 °C.

4.2.2. Procedure for Ub-K63C
A 25 μL aliquot of a solution of Ub-K63C (100 μM in 50 mM NaP$_i$ pH 8.0, 2.5 nmol) was added to 68.8 μL of 50 mM NaP$_i$ pH 8.0. To this was added 6.25 μL of a solution of 3-bromo-2-bromomethyl-1-propene (20 mM in DMF, 125 nmol, 50 eq.). The reaction mixture was vortexed briefly, then incubated at 4 °C for 6 h. At the end of the reaction, a 10 μL aliquot was analysed by LC-ESI-MS. The major product formed was Ub-isobutylene-Br (calculated mass: 8700 Da; observed mass: 8700 Da).

4.2.3. Procedure for anti-HER2 nanobody
A 13 μL aliquot of a solution of anti-HER2 nanobody (77 μM in PBS pH 7.4, 1 nmol) was added to 82 μL of 50 mM NaP$_i$ pH 9.0. To this was added 5 μL of a solution of Tris(2-carboxyethyl)phosphine hydrochloride (TCEP·HCl) (10 mM in H$_2$O, 50 nmol, 50 eq.). The reaction mixture was vortexed briefly, then shaken at 25 °C for 30 min. 10 μL of a solution of 3-bromo-2-bromomethyl-1-propene (100 mM in DMF, 1000 nmol, 1000 eq.) was then added and the resulting mixture was vortexed briefly. After 1 h of additional shaking at 25 °C, a 10 μL aliquot was analysed by LC-ESI-MS. The expected conversion to anti-HER2-nanobody-isobutylene (Nb-isobutylene) was observed (calculated mass: 14992 Da; observed mass: 14992 Da).

4.3. Tests with nucleophiles

4.3.1. General procedure for tests with thiol nucleophiles
To a 10 μM solution of the protein conjugate was added 100–1000 eq. of the appropriate thiol nucleophile. The reaction mixture was vortexed briefly, then shaken at 25 °C for 30 min to 1 h. A 10 μL aliquot of the reaction mixture was then analysed by LC-ESI-MS to determine if the doubly conjugated product had been formed. Further details on the tests with thiol nucleophiles can be found in the Supplementary Data (S4–S6, S13–S15, S22–S24).

4.3.2. General procedure for tests with benzylamine
To a 10 μM solution of the protein conjugate was added 1000 eq. of benzylamine. The reaction mixture was vortexed briefly, then shaken at 25 °C for 4–6 h. A 10 μL aliquot of the reaction mixture was then analysed by LC-ESI-MS to determine if the doubly conjugated product had been formed. Further details on the tests with benzylamine can be found in the Supplementary Data (S7, S16, S25).

4.4. Stability tests of the conjugates

To a 10 μL aliquot of a 10 μM solution of 2c was added either GSH (1000 eq.) or reconstituted human plasma (0.5 μL). The reaction mixture was vortexed briefly, then shaken at 37 °C for 24 h. The reaction mixture was then analysed by LC-ESI-MS. In both instances, 2c was detected unaltered (calculated mass: 16471 Da; observed mass: 16471 Da, see Supplementary Data S8 and S9), thus indicating that the doubly conjugated product was fully stable under the reaction conditions.

4.5. Circular dichroism spectroscopy

Protein concentrations were determined by nanodrop. The final concentration of the protein samples was 0.2 mg/mL in 1X PBS pH 7.4 buffer. CD measurements were performed on an Aviv Model 410 spectrometer, which was routinely calibrated with (1S)-(+)-10-camphorsulfonic acid. Spectra were recorded at 298 K with a 0.1 cm quartz cell over the wavelength range 260–200 nm at 50 nm·min^{-1}, with a bandwidth of 1.0 nm, the response time of 1 s, resolution step width of 1 nm and sensitivity of 20–50 Mdeg. Each spectrum represents the average of 3 scans.

4.6. Surface plasmon resonance analysis

Surface Plasmon Resonance Analysis of C2Am and C2Am-Glc. SPR experiments were performed using a Biacore T200 instrument (GE Healthcare). Small lamellar vesicle (SMVs) liposomes were immobilised on to the flow cells of L1 sensor chips (500–1000 RU) as described before.[34] The immobilisation was performed by injecting a solution of SMVs (0.5 mM) at 10 μL/min in 20 mM HEPES, 150 mM sodium chloride, pH 7.4 (HBS buffer). The flow cells were then primed with HBS buffer containing 2 mM calcium chloride which was also used as running buffer and dilution buffer. Regeneration of the surfaces was performed using HBS buffer containing 20 mM EDTA. Post immobilization, the chip was left to stabilize with a constant flow (30 μL/min) of running buffer. The chip surface was conditioned by injecting 3 × injections of running buffer followed by 3 × injections of regeneration buffer. Control surfaces were prepared by preparing SMVs that did not contain the binding ligand of C2Am (phosphatidyl serine).[34] For determination of binding kinetics, serial dilutions of C2Am and C2Am-Glc (0–0.3 μM) were injected over immobilized SMVs at a constant flow rate (30 μL/min) at 37 °C. Protein solutions were kept at 10 °C at all times. In all experiments, data were zero adjusted and the reference cell subtracted. Data evaluations were performed using Biacore T200 Evaluation 3.0 software (GE Healthcare).

A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.bmc.2018.02.028.

References

1. Krall N, da Cruz FP, Boutureira O, Bernardes GJL. *Nat Chem*. 2015;8:103–113.
2. Lang K, Davis L, Torres-Kolbus J, Chou C, Deiters A, Chin JW. *Nat Chem*. 2012;4:298–304.
3. Toseland CP. *J Chem Biol*. 2013;6:85–95.
4. Zhang G, Zheng S, Liu H, Chen PR. *Chem Soc Rev*. 2015;44:3405–3417.
5. Chari RVJ, Miller ML, Widdison WC. *Angew Chem Int Ed*. 2014;53:3796–3827.
6. Chudasama V, Maruani A, Caddick S. *Nat Chem*. 2016;8:114–119.
7. Beck A, Goetsch L, Dumontet C, Corvaïa N. *Nat Rev Drug Discov*. 2017;16:315–337.
8. Sletten EM, Bertozzi CR. *Angew Chem Int Ed*. 2009;48:6974–6998.
9. Lang K, Chin JW. *Chem Rev*. 2014;114:4764–4806.
10. Chen X, Wu Y-W. *Org Biomol Chem*. 2016;14:5417–5439.
11. Koniev O, Wagner A. *Chem Soc Rev*. 2015;44:5495–5551.
12. Gunnoo SB, Madder A. *ChemBioChem*. 2016;17:529–553.
13. Kapanidis AN, Weiss S. *J Chem Phys*. 2002;117:10953–10964.
14. Toda N, Asano S, Barbas CF. *Angew Chem Int Ed*. 2013;52:12592–12596.
15. Spicer CD, Davis BG. *Nat Commun*. 2014;5. 4740–4740.
16. Boutureira O, Bernardes GJL. *Chem Rev*. 2015;115:2174–2195.
17. Bernardim B, Cal PMSD, Matos MJ, et al. *Nat Commun*. 2016;7:13128.
18. Kim Y, Ho SO, Gassman NR, et al. *Bioconjugate Chem*. 2008;19:786–791.
19. Takaoka Y, Ojida A, Hamachi I. *Angew Chem Int Ed*. 2013;52:4088–4106.
20. Chalker JM, Bernardes GJL, Davis BG. *Acc Chem Res*. 2011;44:730–741.
21. Smith MEB, Schumacher FF, Ryan CP, et al. *J Am Chem Soc*. 2010;132:1960–1965.
22. Rabuka D, Rush JS, DeHart GW, Wu P, Bertozzi CR. *Nat Protoc*. 2012;7:1052–1067.
23. Gavrilyuk J, Ban H, Nagano M, Hakamata W, Barbas CF. *Bioconjugate Chem*. 2012;23:2321–2328.
24. Chalker JM, Gunnoo SB, Boutureira O, et al. *Chem Sci*. 2011;2:1666–1676.
25. Rowan FC, Richards M, Bibby RA, Thompson A, Bayliss R, Blagg J. *ACS Chem Biol*. 2013;8:2184–2191.
26. Timms N, Windle CL, Polyakova A, et al. *ChemBioChem*. 2013;14:474–481.
27. Freedy AM, Matos MJ, Boutureira O, et al. *J Am Chem Soc*. 2017;139:18365–18375.
28. Wright TH, Bower BJ, Chalker JM, et al. *Science*. 2016;354.
29. Smith ME, Schumacher FF, Ryan CP, et al. *J Am Chem Soc*. 2010;132:1960–1965.
30. Chudasama V, Smith ME, Schumacher FF, et al. *Chem Commun (Camb)*. 2011;47:8781–8783.
31. Morgan RE, Chudasama V, Moody P, Smith ME, Caddick S. *Org Biomol Chem*. 2015;13:4165–4168.
32. Boutureira O, Martínez-Sáez N, Brindle KM, Neves AA, Corzana F, Bernardes GJL. *Chem Eur J*. 2017;23:6483–6489.
33. Sun S, Companon I, Martinez-Saez N, et al. *ChemBioChem*. 2018;19:48–52.
34. Alam IS, Neves AA, Witney TH, Boren J, Brindle KM. *Bioconjugate Chem*. 2010;21:884–891.
35. Liu J, Chen Q, Rozovsky S. *J Am Chem Soc*. 2017;139:3430–3437.
36. Massa S, Xavier C, De Vos J, et al. *Bioconjugate Chem*. 2014;25:979–988.
37. Komander D, Rape M. *Annu Rev Biochem*. 2012;81:203–229.
38. Hamamoto T, Sisido M, Ohtsuki T, Taki M. *Chem Commun*. 2011;47. 9116-9116.
39. Xiang Z, Ren H, Hu YS, et al. *Nat Methods*. 2013;10:885–888.
40. Chen XH, Xiang Z, Hu YS, Lacey VK, Cang H, Wang L. *ACS Chem Biol*. 2014;9:1956–1961.
41. Wu G, Fang Y, Yang S, Lupton JR, Turner ND. *Environ Health*. 2004;134:489–492.