

ChemComm

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

COMMUNICATION

Sulfonyl Azide–Mediated Norbornene Aziridination for Orthogonal Peptide and Protein Labeling

Cite this: DOI: 10.1039/x0xx00000x

Michael J. Gattner,^a Michael Ehrlich^a and Milan Vrabel^{a,b}

Received 00th January 2014,

Accepted 00th January 2014

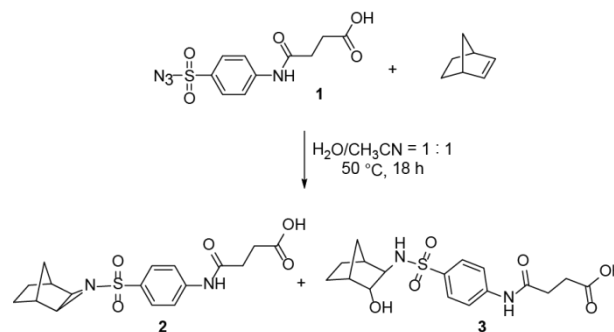
DOI: 10.1039/x0xx00000x

www.rsc.org/

We describe a new bioconjugation reaction based on the aziridination of norbornenes using electron-deficient sulfonyl azides. The reaction enables to attach various useful tags to peptides and proteins under mild conditions.

Bioconjugation reactions substantially extend our ability to chemically manipulate proteins. Numerous chemical strategies for the attachment of synthetic molecules to proteins have been developed.¹ Early approaches focused on native functional groups present on endogenous amino acids.² The main drawback of this approach is the lack of specificity since multiple copies of each amino acid are present in the primary protein structure. Unique recognition elements can be introduced into the protein structure to improve the selectivity of the ligation. One possibility is to add a specific amino acid sequence to the target protein that can be recognized by either an appropriate protein modifying enzyme³ or by specific chemical probes.⁴ Another approach uses the biosynthetic machinery of the cell for incorporation of unnatural amino acids containing artificial functional groups.⁵ We and others have used this approach for the incorporation of unnatural amino acids into proteins and have shown that the introduced functionality can be efficiently tagged by orthogonal chemical reactions.⁶ The right choice of the reacting functional groups and the proper ligation technique plays here the crucial role.⁷ Among other suitable functional groups that enable efficient protein labeling various derivatives of cyclooctyne, cyclooctene and cyclopropene gained special attention in the field.⁸ An extraordinary fast kinetics were observed especially in combination with tetrazines in inverse electron-demand Diels-Alder reactions or in dipolar cycloadditions with nitrile imines.^{8a, 8b, 8c} Unfortunately, the high reactivity of such systems is often accompanied by reduced stability and an increased tendency toward side reactions.⁹ Moreover, synthetic access to these reagents often represents a considerable challenge. It is therefore desirable to develop methodology that not only enables robust protein labeling but also utilizes easily available starting materials.

Based on our experience using norbornenes as highly reactive compounds in combination with nitrile oxides, nitrile imines and tetrazines¹⁰ we searched for reagents that do react with norbornenes, but also meet the criteria for better synthetic accessibility. Here we show that aziridination of norbornenes using electron-deficient sulfonyl azides represents an excellent balance between reactivity, stability and availability of the starting materials. We demonstrate that this reaction can be used for efficient labeling of norbornene containing peptides and proteins.



Scheme 1. Model reaction of norbornene with sulfonyl azide **1**. Yields: 78% of **2** and 18% of **3**.

Inspired by the pioneering studies on norbornene aziridination by Franz et al.,¹¹ we decided to investigate whether this reaction can proceed in an aqueous environment and can be used for biomolecule labeling. Sulfonyl azides have been previously successfully applied as reagents for detecting thiocarboxylates in bacterial proteome¹² and as ligation agents for thioacid-containing peptides and proteins.¹³ These studies clearly indicate that sulfonyl azides are compatible with natural systems. To investigate the reactivity of sulfonyl azides with norbornene under aqueous conditions we first performed a model reaction. We reacted sulfonyl azide **1** with norbornene in water/acetonitrile (1:1) at moderate temperature overnight (Scheme 1 and Scheme S1). After purification using semipreparative HPLC we obtained the desired aziridine **2** in 78% yield. We also isolated

the corresponding sulfonamide **3** (18%) as a by-product formed by nucleophilic attack of a water molecule on the aziridine and subsequent ring opening. 2D-NMR analysis of **2** showed that only the *exo*-aziridine product was formed in the reaction (see Supporting Information). The reaction mechanism leading to the formation of the desired aziridine can involve the initial formation of a triazole intermediate followed by the extrusion of nitrogen. Alternatively, the aziridine formation parallels to that of epoxidation involving a concerted addition of the azide to the double bond with concomitant loss of nitrogen.^{11a} No direct experimental evidence that would support or disprove any of these mechanisms was observed during our experiments. A more detailed study is required to address this issue.

To gain more insight into the reaction kinetics we measured the bimolecular rate constant of the reaction in water/acetonitrile (9:1) to simulate more relevant conditions required for its use on biomolecules. To ensure sufficient water solubility of the norbornene substrate we used *endo*-5-norbornene-2-methanol in these experiments. The measurements were performed under pseudo first order conditions using an excess of norbornene and were performed in triplicate (for details see Supporting Information). The determined second order rate constant was $k = 1.7 \times 10^{-3} \pm 0.21 \times 10^{-3} \text{ M}^{-1} \text{ s}^{-1}$. This value is comparable to rate constants of the strain-promoted azide-alkyne cycloaddition reaction of the first generation cyclooctyne derivatives that are commonly used for biomolecule labeling applications.¹⁴ These promising results prompted us to further examine the reaction as potentially new methodology for modification of biomolecules.

To investigate the stability of sulfonyl azides and products formed in their reaction with norbornenes, we first incubated **1** and the isolated products **2** and **3** in 50 mM MES buffer at pH 5.5 (MES = 2-(*N*-morpholino)ethanesulfonic acid) and in 50 mM TRIS buffer at pH 8.5 (TRIS = 2-amino-2-hydroxymethyl-propane-1,3-diol). HPLC analysis of the mixtures did not show any reaction or decomposition even after prolonged incubation time (Figure S6). Only slow hydrolysis of aziridine **2** to the corresponding sulfonamide **3** was detected. The observed slow hydrolysis of the originally formed aziridine in water is in fact not an obstacle for biomolecule labeling since the desired tag will stay attached to its target.

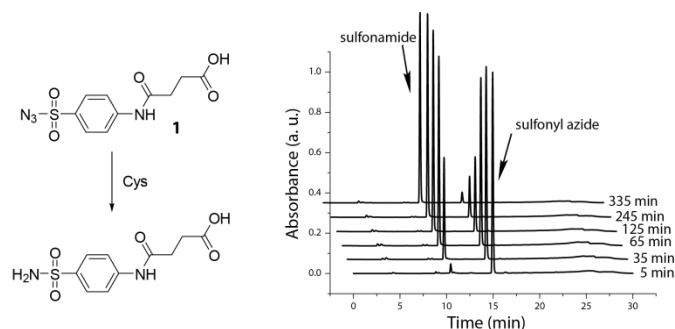


Fig. 1 Reduction of the sulfonyl azide **1** in the presence of cysteine. The new peak (10.3 min) was identified as the corresponding sulfonamide by LC-MS (calc. for $\text{C}_{10}\text{H}_{11}\text{N}_2\text{O}_5\text{S}$ $[\text{M}-\text{H}]^-$: 271.0389, found: 271.0393).

Although sulfonyl azides have been previously used in biological context,¹²⁻¹³ to exclude possible side reactions with endogenous amino acids we performed additional experiments. The electron deficient sulfonyl azides could react with nucleophilic groups on proteins (e.g. cysteines or lysines). Our stability study showed that sulfonyl azides were not affected by amino groups since incubation of compound **1** in TRIS buffer that contains a primary amino group did not show any reaction even at pH 8.5. When we incubated **1** (0.5 mM) with cysteine (50 mM or 2.5 mM) the corresponding sulfonamide was formed as a result of azide reduction (Figure 1, S6 and S7). No nucleophilic substitution reaction was detected. This reaction again will not interfere with protein labeling itself. However, considering the relatively high concentrations of cysteine and glutathione in cells, this side reactivity may hinder the use of this chemistry for *in vivo* applications. To examine a possible reactivity of the aziridine product with other nucleophiles such as amines or thiols we incubated **2** in 50 mM TRIS or in 50 mM cysteine solution. We again observed only slow hydrolysis to the corresponding sulfonamide **3**. These experiments demonstrate that the aziridine moiety is stable toward these nucleophiles. Also, sulfonyl azides can react with nucleophilic double bonds including indole and *N*-methyl indole.¹⁵ To investigate whether this reaction takes place on tryptophan residues (Trp) we incubated compound **1** (0.5 mM) with Trp (0.5 mM). We did not observe any detectable reaction after two days (Figure S8). These additional experiments indicate that sulfonyl azides can be used for peptide and protein labeling in the presence of these native functional groups.

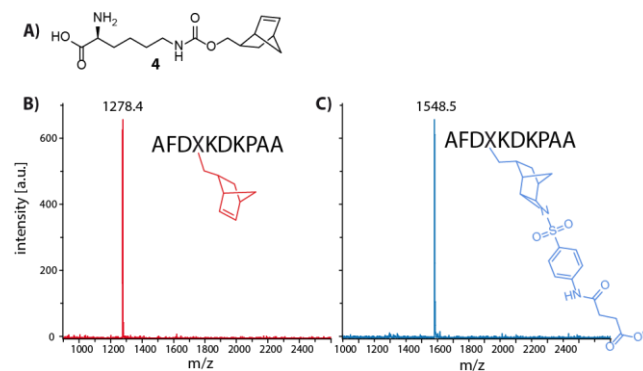


Fig. 2 Aziridination of norbornene-containing peptide AFDXKDKPAA (X = norbornene amino acid **4**) using sulfonyl azide **1**. A) Structure of norbornene containing amino acid **4**. B) MALDI-TOF spectrum of the starting norbornene-containing peptide (calc. mass: $[\text{M}-\text{H}]^-$: 1279.7 Da). C) MALDI-TOF spectrum showing the formation of the aziridinated peptide (calc. mass: $[\text{M}-\text{H}]^-$: 1549.7 Da). $\Delta\text{M}_{\text{calc.}} = 270.0 \text{ Da}$; $\Delta\text{M}_{\text{found}} = 270.1 \text{ Da}$; Conditions: 50 μM peptide, 2.5 mM sulfonyl azide **1**, $\text{H}_2\text{O}:\text{CH}_3\text{CN} = 9:1$, 37 $^\circ\text{C}$, 35 h.

To evaluate the reaction on biomolecules we next synthesized a norbornene-containing peptide: AFDXKDKPAA, where X = *endo* norbornene-containing amino acid **4**. The peptide was incubated at 50 μM final concentration with sulfonyl azide **1** (2.5 mM) in water/acetonitrile (9:1) and the reaction was followed by MALDI-TOF spectrometry. The analysis of the reaction mixture showed complete conversion of the starting peptide to the aziridinated

product within 35 hours (Figure 2 and Figure S3). The hydrolysis of the aziridine to the corresponding sulfonamide was in this case observed only after prolonged incubation time (72 hours, see Figure S4).

Encouraged by these results we next moved to proteins. Using the pyrrolysine amber suppression system we introduced the *endo* norbornene amino acid **4** into *E. coli* thioredoxin (Trx) and human carbonic anhydrase (HCA) as model proteins.^{10c, 16} The norbornene functionality was incorporated at positions Asn65 of Trx (Trx N65X) and His36 of HCA (HCA H36X). (for expression and purification details see the Supporting Information). We next synthesized sulfonyl azide derivatives **5** and **6** bearing a biotin tag or a dansyl fluorophore to examine the norbornene aziridination on proteins (Figure 3). We first incubated a 40 μ M solution of Trx N65X with 50 equivalents of **5** overnight at 37 $^{\circ}$ C in 50 mM Tris (pH 7.5). The successful, almost quantitative labeling of Trx N65X with **5** was indicated by a gel shift of the protein band after SDS-PAGE and further confirmed by intact mass spectrometry (Figure 3A, estimated yield 95%).

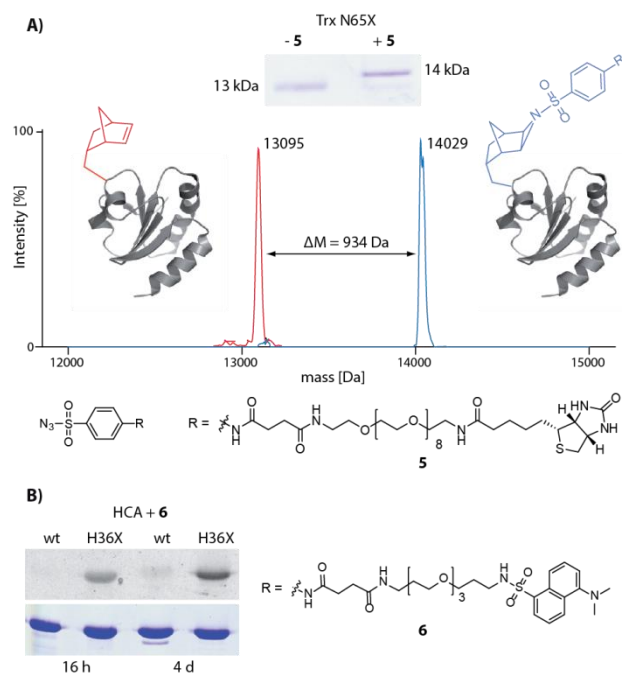


Fig. 3 Aziridination of norbornene-containing proteins. A) Labeling of thioredoxin N65X (X = norbornene amino acid **4**) using biotin sulfonyl azide **5**. Coomassie stained gel and overlaid deconvoluted ESI-MS spectra of the norbornene-containing protein before (calc. mass: 13097 Da, obs. mass: 13095 Da) and after reaction with sulfonyl azide **5** (calc. mass: 14032 Da, obs. mass: 14029 Da, calc. Δ M = 935 Da, obs. Δ M = 934 Da, estimated yield 95%). Depiction of Trx was generated using PDB 2TRX. B) Fluorescent labeling of HCA H36X (X = norbornene amino acid **4**) using dansyl sulfonyl azide **6**. SDS-PAGE analysis of the reaction showing specific labeling of the norbornene-containing HCA mutant after 16 h and 4 d. The gel was analyzed by fluorescence detection (upper part) before it was coomassie stained (lower part).

In addition, after tryptic digestion and MS analysis we found the expected mass of the peptide LNIDHXPGTAPK containing

the desired biotin modification at position X ($[M+2H]^{2+}_{\text{calc.}} = 1188.0978$, $[M+2H]^{2+}_{\text{obs.}} = 1188.0919$ Δ M = 4 ppm). To further evaluate the chemistry on proteins we used the wild type HCA (HCA wt) and norbornene containing HCA H36X respectively. The proteins (40 μ M final concentration) were incubated with the fluorescent sulfonyl azide **6** (2 mM, 50 equiv.) in 50 mM MES buffer, pH 5.5 at 37 $^{\circ}$ C overnight (Figure 3B). SDS-PAGE and subsequent fluorescence detection showed selective labeling of the norbornene containing HCA H36X. Only after prolonged incubation time (four days) we observed a weak fluorescent signal in the reaction of the HCA wt protein, which could be due to non-specific labeling. The origin of the non-specific reaction was not clear since our investigations regarding possible side reactions with endogenous amino acids did not show any reaction. Moreover, the sulfonyl azides were stable under the labeling conditions used (50 mM MES, pH 5.5). Similar non-specific reaction was previously described in the literature, where fluorescent sulfonyl azides were used for visualization of thiocarboxylates in bacterial proteome.¹² However, also in this case the side reactivity could not be appropriately explained. Nevertheless, the specific reaction with norbornene is much faster and therefore the side reaction can be eliminated by simply removing the excess reagent after the reaction. To verify the presence of the desired fluorescent dansyl tag we digested the protein and analyzed the peptide mixture using HPLC-MS. These data demonstrated that the modification was present in the correct position within the protein structure (peptide QSPVDIDTXTAK, X = norbornene amino acid **4** aziridinated by **6**: $[M+2H]^{2+}_{\text{calc.}} = 1079.5082$, $[M+2H]^{2+}_{\text{obs.}} = 1079.5039$, Δ M = 4 ppm).

In summary, we show that norbornene aziridination using electron-deficient sulfonyl azides can be used for orthogonal peptide and protein labeling. The reaction proceeds efficiently under mild conditions, does not require any catalysis and is orthogonal to functional groups of native proteins. Since norbornenes, sulfonyl azides and derivatives thereof are easily accessible compounds the presented technology constitutes an attractive alternative to currently used bioconjugation techniques especially for *in vitro* applications. Further optimization and evaluation of this chemistry toward its use for *in vivo* peptide and protein labeling is ongoing.

Notes and references

^a Department of Chemistry, Ludwig-Maximilians-University, Butenandtstrasse 5-13, 81377 Munich (Germany)

^b Current address: Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo nám. 2, 166 10, Prague (Czech Republic), E-mail: vrabel@uochb.cas.cz; Tel.: +420-220183317

† Electronic Supplementary Information (ESI) available: Synthesis, experimental details, NMR, mass spectrometry and additional analysis data. See DOI: 10.1039/c000000x/

1. (a) M. F. Debets, J. C. M. van Hest and F. P. J. T. Rutjes, *Org. Biomol. Chem.*, 2013, **11**, 6439-6455; (b) C. P. Ramil and Q. Lin, *Chem.*

- Commun.*, 2013, **49**, 11007-11022; (c) E. M. Sletten and C. R. Bertozzi, *Angew. Chem. Int. Ed.*, 2009, **48**, 6974-6998.
- (a) E. Basle, N. Joubert and M. Pucheault, *Chem. Biol.*, 2010, **17**, 213-227; (b) G. E. Means and R. E. Feeney, *Bioconjug. Chem.*, 1990, **1**, 2-12.
 - (a) M. Rashidian, J. K. Dozier and M. D. Distefano, *Bioconjug. Chem.*, 2013, **24**, 1277-1294; (b) J. M. Chalker, G. J. Bernardes and B. G. Davis, *Acc. Chem. Res.*, 2011, **44**, 730-741.
 - (a) T. L. Halo, J. Appelbaum, E. M. Hobert, D. M. Balkin and A. Schepartz, *J. Am. Chem. Soc.*, 2009, **131**, 438-439; (b) J. Zhang, R. E. Campbell, A. Y. Ting and R. Y. Tsien, *Nat. Rev. Mol. Cell Biol.*, 2002, **3**, 906-918.
 - (a) J. W. Chin, *Annu. Rev. Biochem.*, 2014, **83**, 379-408; (b) K. Lang and J. W. Chin, *Chem. Rev.*, 2014, **114**, 4764-4806; (c) C. C. Liu and P. G. Schultz, *Annu. Rev. Biochem.*, 2010, **79**, 413-444; (d) W. Wan, J. M. Tharp and W. R. Liu, *Biochim. Biophys. Acta*, 2014, **1844**, 1059-1070; (e) T. S. Young and P. G. Schultz, *J. Biol. Chem.*, 2010, **285**, 11039-11044; (f) Q. Wang, A. R. Parrish and L. Wang, *Chem. Biol.*, 2009, **16**, 323-336.
 - (a) T. Fekner, X. Li, M. M. Lee and M. K. Chan, *Angew. Chem. Int. Ed.*, 2009, **48**, 1633-1635; (b) Z. Hao, Y. Song, S. Lin, M. Yang, Y. Liang, J. Wang and P. R. Chen, *Chem. Commun.*, 2011, **47**, 4502-4504; (c) D. P. Nguyen, H. Lusic, H. Neumann, P. B. Kapadnis, A. Deiters and J. W. Chin, *J. Am. Chem. Soc.*, 2009, **131**, 8720-8721; (d) L. Wang, Z. Zhang, A. Brock and P. G. Schultz, *Proc. Natl. Acad. Sci.*, 2003, **100**, 56-61; (e) T. Yanagisawa, R. Ishii, R. Fukunaga, T. Kobayashi, K. Sakamoto and S. Yokoyama, *Chem. Biol.*, 2008, **15**, 1187-1197; (f) E. Kaya, K. Gutsmedl, M. Vrabel, M. Muller, P. Thumbs and T. Carell, *ChemBioChem*, 2009, **10**, 2858-2861; (g) N. Li, R. K. V. Lim, S. Edwardraja and Q. Lin, *J. Am. Chem. Soc.*, 2011, **133**, 15316-15319; (h) J. Li, S. Lin, J. Wang, S. Jia, M. Yang, Z. Hao, X. Zhang and P. R. Chen, *J. Am. Chem. Soc.*, 2013, **135**, 7330-7338; (i) A. Dumas, C. D. Spicer, Z. Gao, T. Takehana, Y. A. Lin, T. Yasukohchi and B. G. Davis, *Angew. Chem. Int. Ed.*, 2013, **52**, 3916-3921.
 - (a) C. P. Hackenberger and D. Schwarzer, *Angew. Chem. Int. Ed.*, 2008, **47**, 10030-10074; (b) N. Stephanopoulos and M. B. Francis, *Nat. Chem. Biol.*, 2011, **7**, 876-884.
 - (a) A. Borrmann, S. Milles, T. Plass, J. Dommerholt, J. M. Verkade, M. Wiessler, C. Schultz, J. C. M. van Hest, F. L. van Delft and E. A. Lemke, *ChemBioChem*, 2012, **13**, 2094-2099; (b) K. Lang, L. Davis, S. Wallace, M. Mahesh, D. J. Cox, M. L. Blackman, J. M. Fox and J. W. Chin, *J. Am. Chem. Soc.*, 2012, **134**, 10317-10320; (c) T. Plass, S. Milles, C. Koehler, C. Schultz and E. A. Lemke, *Angew. Chem. Int. Ed.*, 2011, **50**, 3878-3881; (d) T. Plass, S. Milles, C. Koehler, J. Szymanski, R. Mueller, M. Wiessler, C. Schultz and E. A. Lemke, *Angew. Chem. Int. Ed.*, 2012, **51**, 4166-4170; (e) Z. Yu and Q. Lin, *J. Am. Chem. Soc.*, 2014, **136**, 4153-4156; (f) Z. Yu, Y. Pan, Z. Wang, J. Wang and Q. Lin, *Angew. Chem. Int. Ed.*, 2012, **51**, 10600-10604.
 - R. van Geel, G. J. M. Pruijn, F. L. van Delft and W. C. Boelens, *Bioconjugate Chem.*, 2012, **23**, 392-398.
 - (a) K. Gutsmedl, C. T. Wirges, V. Ehmke and T. Carell, *Org. Lett.*, 2009, **11**, 2405-2408; (b) M. Vrabel, P. Kolle, K. M. Brunner, M. J. Gattner, V. Lopez-Carrillo, R. de Vivie-Riedle and T. Carell, *Chem. Eur. J.*, 2013, **19**, 13309-13312; (c) E. Kaya, M. Vrabel, C. Deiml, S. Prill, V. S. Fluxa and T. Carell, *Angew. Chem. Int. Ed.*, 2012, **51**, 4466-4469.
 - (a) J. E. Franz, C. Osuch and M. W. Dietrich, *J. Org. Chem.*, 1964, **29**, 2922-2927; (b) J. E. Franz and C. Osuch, *Tetrahedron Lett.*, 1963, **13**, 837-840.
 - K. Krishnamoorthy and T. P. Begley, *J. Am. Chem. Soc.*, 2010, **132**, 11608-11612.
 - (a) R. Merckx, A. J. Brouwer, D. T. S. Rijkers and R. M. J. Liskamp, *Org. Lett.*, 2005, **7**, 1125-1128; (b) X. Zhang, F. Li, X.-W. Lu and C.-F. Liu, *Bioconjugate Chem.*, 2009, **20**, 197-200.
 - (a) N. J. Agard, J. A. Prescher and C. R. Bertozzi, *J. Am. Chem. Soc.*, 2004, **126**, 15046-15047; (b) J. M. Baskin, J. A. Prescher, S. T. Laughlin, N. J. Agard, P. V. Chang, I. A. Miller, A. Lo, J. A. Codelli and C. R. Bertozzi, *Proc. Natl. Acad. Sci.*, 2007, **104**, 16793-16797.
 - (a) A. S. Bailey and J. J. Merer, *J. Chem. Soc. Perkin 1*, 1966, **15**, 1345-1348; (b) R. E. Harmon, G. Wellman and S. K. Gupta, *J. Heterocyclic Chem.*, 1972, **9**, 1191-1192.
 - S. Schneider, M. J. Gattner, M. Vrabel, V. Flügel, V. López-Carrillo, S. Prill and T. Carell, *ChemBioChem*, 2013, **14**, 2114-2118.