Organic & Biomolecular Chemistry

RSC Publishing

PAPER

View Article Online
View Journal | View Issue

Site-selective azide incorporation into endogenous RNase A *via* a "chemistry" approach†

Cite this: *Org. Biomol. Chem.*, 2013, **11**, 353

Xi Chen,[a,b] Lars Henschke,[a] Qianzhen Wu,[a] Kasturi Muthoosamy,[b] Boris Neumann[c] and Tanja Weil*[a,b]

Downloaded by University of Science and Technology of China on 23 December 2012
Published on 07 November 2012 on http://pubs.rsc.org | doi:10.1039/C2OB26561C

Received 7th August 2012,
Accepted 30th October 2012

DOI: 10.1039/c2ob26561c

www.rsc.org/obc

Site-selective labeling of endogenous proteins represents a major challenge in chemical biology, mainly due to the absence of unique reactive groups that can be addressed selectively. Recently, we have shown that surface-exposed lysine residues of two endogenous proteins and a peptide exhibit subtle changes in their individual reactivities. This feature allows the modification of a single residue in a highly site-selective fashion if kinetically controlled labeling conditions are applied. In order to broaden the scope of the "kinetically-controlled protein labeling" (KPL) approach and highlight additional applications, the water-soluble bioorthogonal reagent, biotin–TEO–azido–NHS (**11**), is developed which enables the site-selective introduction of an azido group onto endogenous proteins/peptides. This bioconjugation reagent features a biotin tag for affinity purification, an azido group for bioorthogonal labeling, a TEO (tetra-ethylene oxide) linker acting as a spacer and to impart water solubility and an *N*-hydroxysuccinimidyl (NHS) ester group for reacting with the exposed lysine residue. As a proof of concept, the native protein ribonuclease A (RNase A) bearing ten available lysine residues at the surface is furnished with a single azido group at Lys 1 in a highly site-selective fashion yielding azido–(K1)RNase A. The K1 site-selectivity is demonstrated by the combined application and interpretation of high resolution MALDI-ToF mass spectroscopy, tandem mass spectroscopy and extracted ion chromatography (XIC). Finally, the water soluble azide-reactive phosphine probe, rho–TEO–phosphine (**21**) (rho: rhodamine), has been designed and applied to attach a chromophore to azido–(K1)RNase A *via* Staudinger ligation at physiological pH indicating that the introduced azido group is accessible and could be addressed by other established azide-reactive bioorthogonal reaction schemes.

Introduction

Over the past few years, the site-selective modification of proteins has received increasing interest in chemical biology.[1–4] Many important *in vivo* and *ex vivo* events are triggered by modification processes that alter the structure and function of proteins.[5] These processes are commonly denoted as post-translational modifications (PTMs) and they play a crucial role in many cellular events such as the *in vivo* phosphorylation, which regulates endocytosis and downstream protein–protein interactions.[6] *In vitro*, the site-specific chemical modification of proteins represents a key tool for numerous biochemical studies[7,8] and has earned also high commercial interest for the development of next generation biotherapeutics with improved pharmacokinetic properties for disease treatment.[9,10]

Even though PTMs play a crucial role in many areas and they can be readily accomplished in living organisms,[11] the site-selective chemical modification of endogenous proteins still represents a major challenge. Since many copies of each amino acid are present in a typical protein, it is often considered impossible to produce a homogeneously functionalized variant.[12] In contrast, recombinant proteins can be furnished with reactive functionalities *via* point mutations such as cysteine residues that allow the site-specific attachment of further functionalities.[13] In addition, non-canonical amino acids (NCAA) bearing reactive functionalities could be introduced by more elaborate biological techniques such as residue-specific NCAA incorporation[14,15] or nonsense codon suppression[16,17] that facilitate bioorthogonal post-modifications such as 1,3 Huisgen cycloaddition reactions ("click

[a]Institute of Organic Chemistry III/Macromolecular Chemistry and Biological Chemistry, University of Ulm, Albert-Einstein 11, 89069 Ulm, Germany. E-mail: tanja.weil@uni-ulm.de; Fax: (+)49-73150-22883; Tel: (+)49-73150-22880
[b]Department of Chemistry, National University of Singapore, 3 Science Drive 3, 117543, Singapore
[c]Proteome Factory AG, Magnusstr. 11, Haus 2, 1. OG, 12489 Berlin, Germany

†Electronic supplementary information (ESI) available: The auto-syringe pump setup in KPL, RNase A functional assay, HRMS (ESI) spectra of K1 modified *K*ETAAAKF (**1–8**), interpretation of the false positive signals in Fig. 4a, the argon column chromatography setup, details of organic synthesis as well as the full NMR spectra of all novel compounds. See DOI: 10.1039/c2ob26561c

This journal is © The Royal Society of Chemistry 2013

View Article Online

Downloaded by University of Science and Technology of China on 23 December 2012
Published on 07 November 2012 on http://pubs.rsc.org | doi:10.1039/C2OB26561C

reactions")[18–20] or Sonogashira cross-coupling reactions[21,22] on proteins. Other strategies involve enzymatic posttranslational modifications of proteins.[23,24] These biological approaches are nowadays widely used but they can be technically more demanding and they often require several optimization cycles until the protein of choice is expressed and purified in acceptable quantities.

Labeling of native proteins *via* chemical approaches usually lacks site-selectivity since addressing a single functional group is challenging and heterogeneous protein mixtures are often obtained. A few attempts have been proven successful such as targeting accessible disulfide bonds[25] or the N-terminus of a protein.[26–29] However, possible limitations of these approaches include the necessity to apply non-physiological reaction conditions, lack of accessible disulfides for intercalation or N-termini not suitable for further chemical modifications. In addition, heterogeneous mixtures containing both native and modified proteins are often obtained, and the purification of the modified protein can be challenging. Recently, affinity-guided concepts based on the specificity of ligand–receptor interactions have received considerable attention for the site-specific modification of endogenous proteins.[30–33]

We have recently reported the concept that surface-exposed lysine residues of native proteins and peptides can reveal subtle differences of their individual reactivity, which allows introducing reactive groups in a highly site-selective fashion. This "*kinetically-controlled protein labeling*" (KPL) approach has been applied successfully to furnish the model proteins RNase A, lysozyme C and the peptide hormone somatostatin-14 with a single biotin ligand in a highly site-selective fashion.[34] In addition, the protein RNase A was also functionalized with an ethynyl group.[34] In order to further extend this bioconjugation scheme to the attachment of even more attractive and versatile groups, we have developed herein the novel azido-containing biotin-tagged bioorthogonal reagent biotin–TEO–azido–NHS (**11**) for the site-selective introduction of a bioorthogonal azido group onto proteins and peptides. Biotin-tagged affinity conjugation agents have received considerable attention in chemical biology, particularly when an affinity separation is required.[35–37] Besides, the azido substituent reacts with a high chemoselectivity in the presence of proteinogenic groups *via e.g.* copper-free click labeling[38] and Staudinger ligation labeling;[39,40] both approaches have even been performed in living organisms. Therefore, the azide-mediated "tag-and-modify"[41] concept that this functionality offers allows chemical modifications under rather mild conditions and is particularly suited for sensitive proteins. In addition, the azido group also represents an interesting IR probe for elucidating protein folding, protein dynamics and electrostatics[42] which further stresses the significance of decorating proteins with azides.

In this contribution, the protein RNase A is furnished with an azido group site-selectively at Lys 1 and further modified *via* Staudinger ligation following an entirely chemistry-based scheme for RNase functionalization. The novel azide-containing reagent biotin–TEO–azido–NHS (**11**) has been developed that reacts with native RNase A under KPL conditions with a high labeling yield. The water soluble azido-reactive phosphine probe, rho–TEO–phosphine (**21**), is prepared demonstrating the potential to site-selectively introduce additional functional groups into azido–(K1)RNase A *via* Staudinger ligation at physiological pH in PBS buffer. In addition, we have shown previously that KPL allows modifying lysozyme and somatostatin as well again stressing the potential of our approach for the introduction of azido groups into proteins other than RNase A. The approach reported herein is fast, convenient and minimizes the risk for protein misfolding. It complements the portfolio of recombinant techniques or chemoenzymatic approaches and opens up various opportunities for precisely functionalizing endogenous proteins and peptides with azido groups.

Experimental section

Methods and materials

Ribonuclease A was obtained from *MP Biomedicals, LLC* (Cat. No. 101076). Pierce Monomeric Avidin Agarose for affinity chromatography was from *Thermo Scientific* (Cat. No. 20267). MALDI-ToF-MS spectra of protein samples were recorded on a Bruker Daltonics mass spectrometer using sinapinic acid as the matrix. The analysis of the protein modification site by *nano*LC-MS^2(ESI) was accomplished at *Proteome Factory AG* (Germany). NuPAGE® Novex 4–12% Bis-Tris gel and its corresponding NuPAGE MES SDS running buffer for gel electrophoresis were purchased from *Invitrogen*. Phosphate buffered saline, ultra-pure grade, was purchased from 1st BASE (Cat. No. 2040). A Pierce BCA assay kit for protein quantifications was obtained from *Novagen* (Cat. No. 71285-3). Baker's yeast ribonucleic acid (*S. cerensiae*) was purchased from *Sigma-Aldrich* (Cat. No. R6750). SYBR® Safe DNA gel stain (10 000× concentrated) in DMSO was obtained from *Invitrogen Molecular Probes* (USA). Ultra-pure grade Tris (Cat. No. 1400) and ultra-pure grade EDTA (Cat. No. 1050) were purchased from 1st BASE. 384-Well black microtiter plates for RNase A functional assay and sterile 96-well plates were received from *Greiner Bio-One Ltd* (UK) and *Nunclon Surface* (Denmark), respectively.

Protein modifications

(A) "Kinetically-controlled protein labeling" for the preparation of azido–(K1)RNase A (13). A solution of the *in situ* prepared biotin–TEO–azido–NHS (**11**) (0.1 M in DMF) (58.4 μL, 5.84 μmol, 2 equiv.) was diluted in 2 mL of cold DMF (anhydrous). The resultant DMF solution was injected dropwise into the stirring RNase A (**12**) solution (40 mg, 2.92 μmol, 1.0 equiv.) in PBS buffer (0.1 M, pH 7.2) (40 mL) *via* an auto-syringe pump during 200 min at +4 °C in a cold room (Fig. S1†). After the injection was completed, the resultant protein solution was further stirred at +4 °C overnight (*step I: kinetic conjugation*). Ethanolamine solution (7.84 mg, 128 μmol, 43.8 equiv.) was added and the solution was stirred at RT for an additional 9 h in order to quench any unreacted NHS ester (**12**) (*step II: quenching step*). The protein solution

 This journal is © The Royal Society of Chemistry 2013

Downloaded by University of Science and Technology of China on 23 December 2012
Published on 07 November 2012 on http://pubs.rsc.org | doi:10.1039/C2OB26561C

was freeze dried, re-constituted in 4 mL of MQ-H_2O and dialyzed against MQ-H_2O (MWCO 3.5 kDa) overnight in order to remove biotin-containing small molecules that could perturb the subsequent affinity purification (*step III: 1st biotin-removal step*). The obtained protein solution (~8 mL) exclusively contained azido–(K1)RNase A (**13**) and excess of unreacted native RNase A (**12**). Due to the loading capacity limit, 0.5 mL of this solution was loaded into a Pierce Monomeric Avidin Resin Column (2 mL of resin supplied, Thermo Scientific, Prod. #20267) and the column was eluted by PBS buffer (pH 7.2), yielding the recovered RNase A solution and quantified by BCA assay (2.28 mg) (*step IV: recovery step*). Afterwards, the column was eluted by a 2 mM (+)-biotin solution in PBS buffer (pH 7.2), giving mono-modified RNase A (**13**) (*step V: affinity elution*). The resultant azido–(K1)RNase A (**13**) solution contained undesired (+)-biotin molecules and needed to be removed *via* a second dialysis (MWCO 3.5 kDa) (*step VI: 2nd biotin-removal step*). The obtained protein solution was azido–(K1)RNase A (**13**) which was quantified by BCA assay (195 μg) and then lyophilized to yield the targeted azido–(K1) RNase A (**13**). On account of recovered RNase A, the labeling yield was calculated to be 84%. **MALDI-ToF-MS:** m/z 14 336 $[M + H]^+$.

(B) Staudinger ligation labeling of azido–(K1)RNase A (**13**) using rho–TEO–phosphine (**11**). Rho–TEO–phosphine (**21**) (1.2 mg, 992 nmol) wetted by DMSO was added into azido–(K1)RNase A (**13**) (100 μg, 3.5 nmol) solution in PBS buffer (200 μL, 1×, pH 7.2) under Ar in a small glass vial. The resultant violet solution was shaken under Ar at 37 °C for 24 h. Gel-electrophoresis reveals a highly fluorescent band at ~15 kDa suggesting the rhodamine was covalently attached onto azido–(K1)RNase A *via* Staudinger ligation. Pure azido–(K1)RNase A (**13**) and the control reaction using native RNase A (**12**) instead of azido–(K1)RNase A (**13**) only display a non-fluorescent gel band at ~14 kDa.

Results and discussion

Design of the key building blocks for KPL of endogenous RNase A

RNase A is selected as a model protein since it is stable, well-characterized and of therapeutic relevance as one of its close variants, onconase, is currently under clinical investigations as a promising cancer therapeutic.[43,44] Endogenous RNase A contains as many as ten lysine residues on its surface, and according to solvent accessibility[45] calculations, Lys 1 displays the highest solvent accessibility.[34] However, except for Lys 41, all other lysine residues of RNase A are located on the protein surface and should still be readily accessible (Fig. 1a). KPL could be applied if one Lys residue has a higher reactivity compared to other Lys residues based on, *e.g.*, a better solvent accessibility or activating or deactivating interactions of neighboring amino acids.[34] In this case, by adding a low concentration of the bioconjugation reagent to a large excess of native protein over an extended period of time, the most reactive amino acid can be labeled selectively. Herein, the novel multifunctional reagent (**11**) has been designed that offers the site-selective incorporation of an azido group *via* KPL (Fig. 1b). In addition, the water-soluble azide-reactive probe, rho–TEO–phosphine (**21**), facilitates efficient post-modifications of azido-containing proteins in aqueous solution at physiological pH *via* Staudinger ligation to prove the accessibility of the newly introduced azido group (Fig. 1c). The two steps allow an endogenous protein to be converted to its site-selectively modified variants (Fig. 1d) where all related bioconjugation and

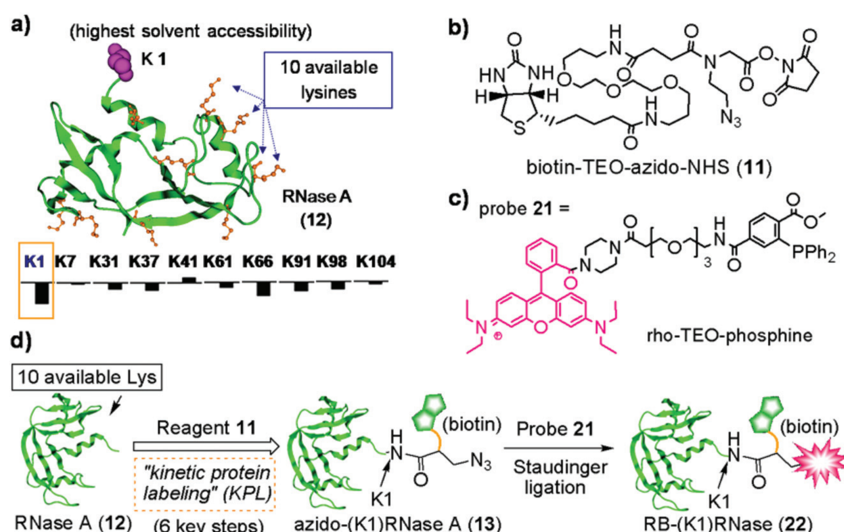


Fig. 1 The crystal structure of RNase A (PDB file: 1RCA.pdb) with nine out of ten lysine residues highlighted as a yellow ball-and-stick model (above); calculated solvent accessibility of each lysine residue where lysine 1 (K1) (violet sphere) gives the highest value (below); (b) the chemical structure of reagent **11**—biotin–TEO–azido–NHS; (c) the chemical structure of the phosphine probe—rho–TEO–phosphine (**21**); (d) the "kinetic protein labeling" approach for site-selectively introducing an azido handle onto endogenous RNase A (**12**) using reagent **11** and further Staudinger ligation labeling of azido-functionalized RNase A by phosphine probe **21**.

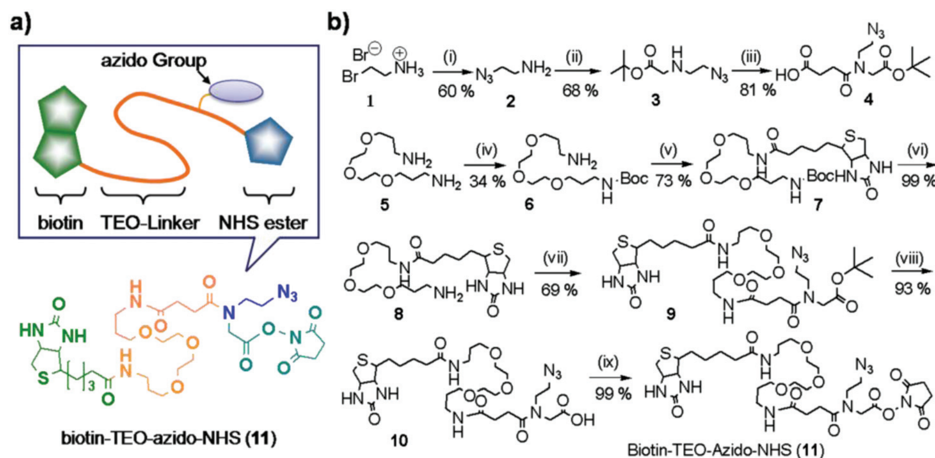


Fig. 2 (a) Design of biotin–TEO–azido–NHS (**11**) featuring a biotin tag for affinity purification, a TEO-spacer for imparting water solubility, an azido group for bioorthogonal labeling and an amine-reactive NHS ester for targeting lysine. (b) The total synthetic scheme of biotin–TEO–azido–NHS (**11**): (i) NaN_3, H_2O, 75 °C, 20 h, then NaOH (aq), 60%. (ii) $BrCH_2COO\textit{t}Bu$, K_2CO_3, DMF, RT, 48 h, 68%. (iii) Succinic anhydride, DMF, RT, overnight, 81%. (iv) $(Boc)_2O$, dioxane, RT, overnight, 34%. (v) (+)-Biotin, EDC·HCl, DMF, RT, 73%. (vi) TFA, DCM, RT, 30 min then NH_3 (aq), 99%. (vii) **4**, EDC·HCl, DMF, RT, 48 h, 69%. (viii) TFA, DCM, RT, 6 h, 93%. (ix) NHS, EDC·HCl, DMF, RT, 24 h, 99%.

Downloaded by University of Science and Technology of China on 23 December 2012
Published on 07 November 2012 on http://pubs.rsc.org | doi:10.1039/C2OB26561C

purification steps proceed under very mild conditions (pH 7.2) and without the necessity to apply recombinant techniques.

Synthesis and characterization of biotin–TEO–azido–NHS (11)

The multifunctional, bioorthogonal linker biotin–TEO–azido–NHS (**11**) (TEO: tetraethylene oxide; NHS: *N*-hydroxysuccinimidyl ester) has been designed to meet several unique characteristics such as a biotin tag for affinity purification of proteins, a TEO spacer for imparting water solubility and an amine-reactive NHS ester group for targeting lysine residues as well as the azido group as a reactive handle (Fig. 2a). The total synthesis scheme is depicted in Fig. 2b. Briefly, 2-azidoethylamine (**2**)[46] is prepared *via* nucleophilic attack of 2-bromoethylamine (**1**) with sodium azide under heating. Then, 1 equiv. of *tert*-butyl-2-bromoacetate is treated with 2 equiv. of 2-azidoethylamine (**2**) under basic conditions to afford *tert*-butyl 2-azidoethylaminoacetate (**3**) which is subsequently acylated by succinic anhydride to yield **4**. In parallel, 4,7,10-trioxa-1,13-tridecanediamine (**5**) is protected to yield *N*-Boc-4,7,10-trioxa-1,13-tridecanediamine (**6**) and then coupled with (+)-biotin *via* EDC (1-ethyl-3-(3-dimethylaminopropyl)carbodiimide) condensation to obtain biotin–TEO–NH(Boc) (**7**). Afterwards, **7** is deprotected by trifluoroacetic acid (TFA) to give biotin–TEO–NH_2 (**8**), which is then coupled with **4** in the presence of EDC to yield **9**. Subsequently, **9** is deprotected by TFA to afford biotin–TEO–azido–COOH (**10**). Finally, **10** is coupled with *N*-hydroxysuccinimide to give the desired bioorthogonal reagent **11** in its activated NHS ester form. Biotin–TEO–azido–NHS (**11**) is synthesized in nine steps with an overall yield of 30% starting from 2-bromoethylamine hydrobromide salt (**1**).

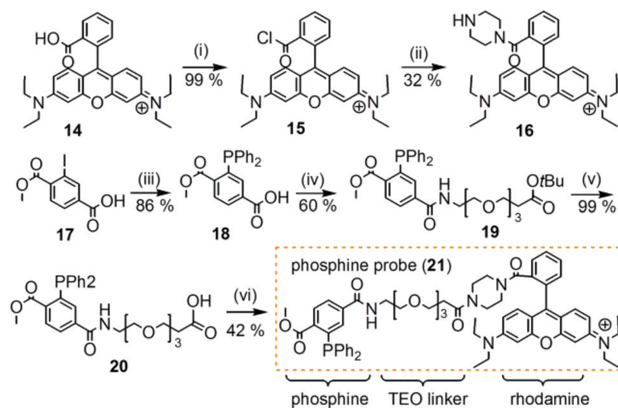
Since the final bioconjugation reagent biotin–TEO–azido–NHS (**11**) represents an active species and displays only limited stability, the free acid form, biotin–TEO–azido–COOH (**10**), has been characterized in detail by HRMS, LC-MS and NMR spectroscopy. Since biotin–TEO–azido–COOH features a tertiary amide bond, both of its ^1H-NMR and carbon NMR spectra show two sets of signals due to the presence of rotamers. Therefore, 2D-NMR including ^1H–^1H COSY and HMQC as well as DEPT 135 spectra have been recorded to fully assign all the NMR signals (see ESI†).

Synthesis of the water soluble azide-reactive probe—rho–TEO–phosphine (21)

Subsequently, the water soluble phosphine probe—rho–TEO–phosphine (**21**) has been designed featuring an azido-reactive phosphine group, a water soluble TEO spacer and a rhodamine B chromophore (Scheme 1). Previously, similar rhodamine–phosphine conjugates have been reported and applied, *e.g.* as bioorthogonal chemical reporters.[47] The rho–TEO–phosphine presented herein features a highly hydrophilic tetraethylene oxide (TEO) linker for water solubility. Synthesis of **21** is accomplished in six steps. First, rhodamine B piperazine amide[48] is prepared from rhodamine B (**14**) using a new procedure *via* a rhodamine B acyl chloride intermediate (**15**). Then, 1-methyl-2-iodoterephthalate (**17**) is coupled to diphenylphosphine under the catalysis of palladium acetate to yield 2-diphenylphosphino-1-methylterephthalate (**18**).[49] Thereafter, this phosphine intermediate is connected to the water soluble linker, *tert*-butyl 12-amino-4,7,10-trioxadodecanate, using HBTU (*O*-benzotriazol-1-yl-*N*,*N*,*N*′,*N*′-tetramethyl-uronium hexafluorophosphate) as the coupling reagent. After deprotection of the *tert*-butyl ester group, the resultant free acid (**20**) is attached to rhodamine B piperazine amide (**16**) *via* HBTU coupling yielding the rho–TEO–phosphine (**21**) azide probe in an overall yield of 21% starting from 2-iodo-1-methylterephthalate (**17**).

 This journal is © The Royal Society of Chemistry 2013

Downloaded by University of Science and Technology of China on 23 December 2012
Published on 07 November 2012 on http://pubs.rsc.org | doi:10.1039/C2OB26561C



Scheme 1 Synthesis of the water soluble rho–TEO–phosphine probe (**21**): (i) $SOCl_2$, RT, overnight. (ii) Piperazine, DCM, RT, overnight, 31% from **14**. (iii) $HPPh_2$, $Pd(OAc)_2$, K_2CO_3, MeCN, 85 °C, 24 h, 86%. (iv) *tert*-Butyl 12-amino-4,7,10-trioxadodecanate, HBTU, DIEA, DMF, RT, 48 h, 60%. (v) TFA, DCM, RT, overnight, 99%. (vi) HBTU, rhodamine B piperazine amide (**16**), DIEA, DMF, RT, 48 h, 42%.

Preparation of azido–(K1)RNase A *via* an optimized "kinetic protein labeling" (KPL) procedure

Preparation of azido–(K1)RNase A *via* an optimized KPL scheme. The general procedures of the KPL (Scheme S1) are given in the ESI,† featuring six essential steps. In comparison to our previous KPL study,[34] an optimized labeling procedure has been applied herein for the preparation of site-selectively mono-functionalized azido–(K1)RNase A. First, an increased kinetic bioconjugation time (from 100 min to 200 min) and an overnight incubation after labeling (instead of 1 h of incubation) were applied in order to increase the labeling efficiency of **11**. As a result, only 2 equiv. of the bioconjugation reagent compared to 8 equiv. used in the previous protein ethynylation were applied that allowed saving the precious reagent **11**. Additionally, an auto-syringe pump set-up (Fig. S1†) as opposed to tedious manual addition used in a previous biotinylation protocol[34] has been applied which guaranteed higher reproducibility. Dialysis was chosen instead of gel filtration which better suits a larger scale reaction. Furthermore, in *step I*, the bioconjugation reaction is performed at a reduced temperature (+4 °C) in a cold room and not at RT to further increase the site-selectivity.

In short, 2 equiv. of biotin–TEO–azido–NHS (**11**) in DMF solution (1/20 v/v relative protein solution volume) are injected dropwise into the RNase A solution at +4 °C in PBS buffer (pH 7.2) during 200 min *via* an auto-syringe pump yielding mono-modified RNase A co-existing with native RNase A which was used in excess (Fig. 3a). Herein, minor amounts of the organic solvent DMF are applied in order to reduce the autohydrolysis of **11** in aqueous solution. Even though 2 equiv. of **11** are applied, it is still kept in deficiency since **11** reveals autohydrolysis in water and is partially deactivated before reacting with RNase A (**12**). According to the MALDI-ToF-MS spectrum, only mono-modified RNase A (**13**) is formed in the presence of native RNase A, which has been used in excess (Fig. 3b, middle). It is noteworthy that no multi-modified side-products have been detected. Ethanolamine is added to the reaction mixture to quench any remaining unreacted NHS ester and the solution is purified on the monomeric avidin resin. Azido–(K1)RNase A (**13**) which carries a biotin motif is reversibly bound into the resin while excess of unreacted RNase A is recovered by eluting the resin column with PBS buffer (pH 7.2). Thereafter, azido–(K1)RNase A (**13**) is eluted from the resin with PBS buffer (pH 7.2) containing 2 mM of (+)-biotin. Finally, the free biotin is removed by gel filtration, yielding azido–(K1)RNase A (**13**) which is quantified by the BCA assay. **13** exhibits a single peak with m/z 14 336 $[M + H]^+$ matching the calcd M.W. of 14 335 g mol^{-1} (Fig. 3b, right). Considering the amount of recovered RNase A, the labeling yield is about 84% and minor product loss might be due to the purification steps including dialysis.

It is noteworthy that the total amount of the bioconjugation reagent relative to the amount of protein required for this reaction step does not represent a fixed number but merely depends on the properties of the reagent itself (*e.g.* its water-solubility, reactivity, steric hindrance of the reactive group) as well as the protein of choice. Therefore, it is advisable that a kinetic labeling study is conducted in order to identify the optimal reagent/protein ratio. This kinetic labeling study of a new reagent/protein pair could be easily investigated *via* MALDI-ToF-MS spectrometry. In the case that a less water soluble or less reactive labeling reagent is used, the required amount of the bioconjugation reagent could also be less than *e.g.* one equiv. However, if the reagent is very water-soluble but of limited stability in water such as (**11**), which is deactivated during the labeling process, the required quantity to achieve an optimal labeling result exceeds one equiv. In this study, the application of 0.5 equiv. of biotin–TEO–azido–NHS (**11**) results in a very low labeling ratio and very little amounts of azido–(K1)RNase A (**13**) are formed (data not shown). Therefore, 2 equiv. of **11** have been used thus yielding an improved labeling efficiency (Fig. 3b, middle).

RNase A functional assay. The enzymatic activities of both azido–(K1)RNase A and recovered RNase A relative to native RNase A are determined by an RNase A functional assay[34] in order to understand the effects of the azide-modification as well as the optimized KPL procedures on the activity of this enzyme. According to the experimental results, the retained enzymatic activity of the K1 functionalized RNase A variant, azido–(K1)RNase A (**13**), is about 29% of native RNase A (Fig. S3†) which is lower than the retained activity of biotin–LC–(K1)RNase A (88%)[34] which carries a smaller biotin substituent. This activity loss often occurs after chemical modification and in the case of RNase A, it might be due to the fact that the modified Lys 1 residue is located in close vicinity to the active site of RNase A and a more bulky group (*e.g.* biotin–TEO–azido moiety) might further reduce the enzymatic activity. However, the introduction of functional groups close to the active center could be very attractive for achieving responsive proteins whose activity could be switched upon an external stimulus *e.g.* by attaching responsive polymers.[50] It is

Downloaded by University of Science and Technology of China on 23 December 2012
Published on 07 November 2012 on http://pubs.rsc.org | doi:10.1039/C2OB26561C

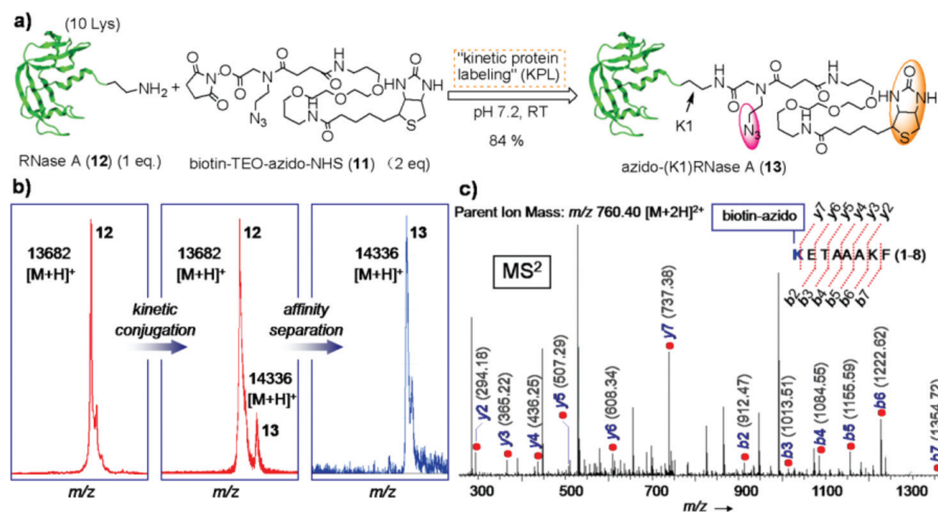


Fig. 3 (a) "Kinetic protein labeling" scheme of the site-selective introduction of an azido group at K1 of RNase A (**12**) applying biotin–TEO–azido–NHS (**11**); (b) MALDI-ToF-MS spectra of native RNase A (**12**) (left), KPL reaction mixture (middle) and purified mono-modified azido–(K1)RNase A (**13**); (c) full interpretation of the MS^2 spectrum of modified *K*ETAAAKF (**1–8**), suggesting that the modification site occurred at K1 rather than K7.

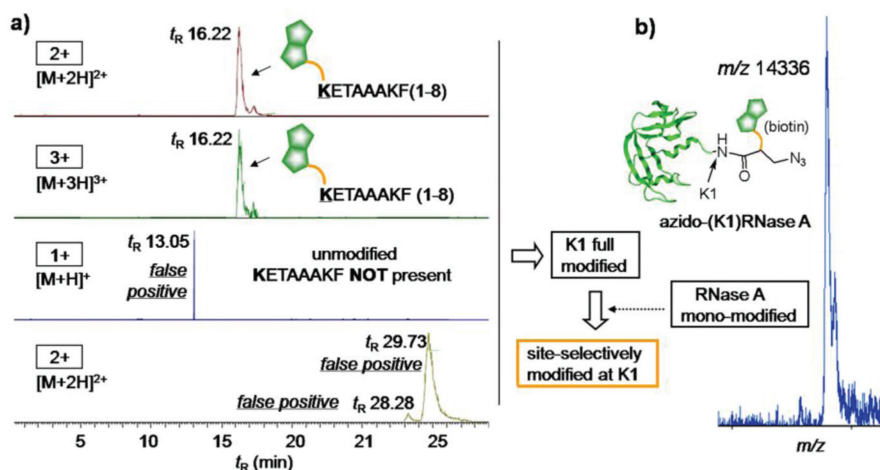


Fig. 4 (a) The XIC chromatographs of modified *K*ETAAAKF (**1–8**) (top, 2+ and 3+) as well as unmodified KETAAAKF (**1–8**) (bottom, 1+ and 2+) proving the full modification at K1; (b) the MALDI-ToF-MS spectrum of azido–(K1)RNase A revealing mono-modification of RNase A. Both data demonstrate the site-selective modification at K1.

noteworthy that the recovered RNase A displays the same enzymatic activity of native RNase A suggesting that the optimized KPL procedure and the purification protocol have no negative impact on the activity of RNase A (Fig. S3†).

Identification of the location of the azido-group at the surface of azido–(K1)RNase A (**13**). Identification of the modification site is achieved by using *nano*LCMS^2.[34] Briefly, this protein sample is subjected to gel-electrophoresis, in-gel digested by chymotrypsin and the resultant peptide fragments are analyzed by *nano*LC-MS^2. The peptide fragment, KETAAAKF (**1–8**), has been identified to contain the biotin–TEO–azido moiety (calcd M.W. of modified ***K***ETAAAKF: 1518.790 g mol^{-1}) due to the presence of the parent ions that match *m*/*z* 760.403 $[M + 2H]^{2+}$ and *m*/*z* 507.271 $[M + 3H]^{3+}$ (Fig. S4†). Since this peptide fragment contains two lysine residues, *i.e.* Lys 1 and Lys 7, the MS^2 spectrum of this peptide fragment has been recorded in order to differentiate the two sites (Fig. 3c). All the *b* and *y* ions of this peptide fragment have been elucidated and are marked in the MS^2 spectrum, proving that the modification occurred at K1 rather than K7.

Labeling site-selectivity at K1. Some of the peptide fragments containing free thiol groups cannot be determined since they are generally lost after *Michael addition* with the polyacrylamide gel during in-gel digestion. In order to circumvent this limitation, the extracted ion chromatographs (XIC) of both modified ***K***ETAAAKF (**1–8**) (Fig. 4a top, $[M + 2H]^{2+}$ and $[M + 3H]^{3+}$) and unmodified **KETAAAKF** (**1–8**) (Fig. 4a bottom, $[M + H]^{+}$ and $[M + 2H]^{2+}$) fragments have been recorded. While the modified ***K***ETAAAKF (**1–8**) shows intensive peaks in its XIC chromatograph at t_R 16.22 min, unmodified **KETAAAKF** (**1–8**) peaks have not been identified. The peak at t_R 13.05 min in Fig. 4a, and the peaks at t_R 28.28 min and t_R 29.73 min in

 This journal is © The Royal Society of Chemistry 2013

Published on 07 November 2012 on http://pubs.rsc.org | doi:10.1039/C2OB26561C
Downloaded by University of Science and Technology of China on 23 December 2012

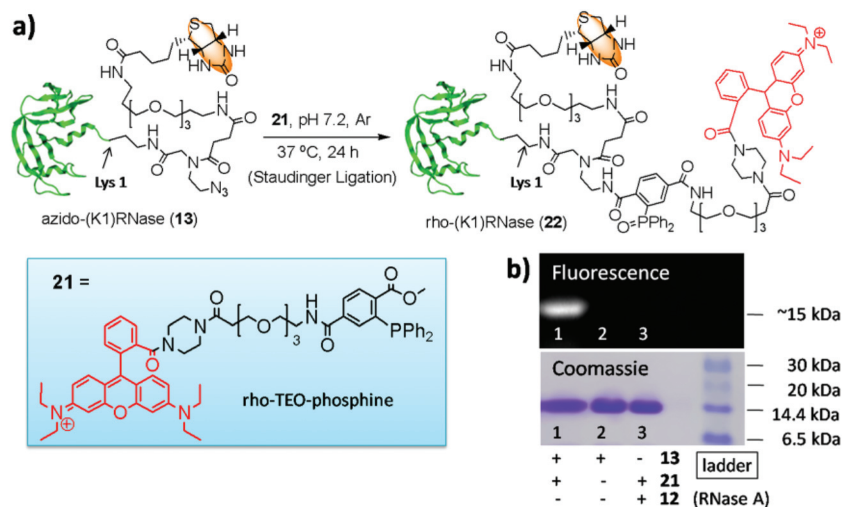


Fig. 5 (a) Staudinger ligation of azido–(K1)RNase A (**13**) with the rho–TEO–phosphine probe (**21**); (b) gel-electrophoresis image of the Staudinger ligation labeling reaction (left), pure azido–(K1)RNase A (middle) and the control reactions using native RNase A (**12**) instead of azido–(K1)RNase A (**13**) to react with rho–TEO–phosphine (**21**) (right) (above: fluorescence; below: Coomassie blue stain; Pre-Stained Protein Marker II (AppliChem) was used as the gel ladder).

Fig. 4a (bottom) represent false positive signals due to their unmatched charges (see ESI, Fig. S4†). This information collected from XIC proved the full modification of the K1 residue of RNase A. As RNase A is mono-modified (Fig. 4b), *the combined information demonstrates the K1 site-selectivity of the "kinetic protein labeling" approach yielding azido–(K1)RNase A* (**13**). However, this study does not exclude the possibility of other lysine residues than K1 that might have been modified to a very low extent which is below the detection limit of MALDI-ToF-MS spectrometry.

Staudinger ligation of azido–(K1)RNase A (13) by rho–TEO–phosphine (21)

In the last step, the accessibility of the azido group on azido–(K1)RNase A (**13**) for the Staudinger ligation modification as well as the azide-reactivity of the rho–TEO–phosphine (**21**) is investigated. For this purpose, the azido–(K1)-RNase A (**13**) is labelled with the probe (**21**) under an Ar atmosphere at 37 °C for 24 hours (Fig. 5a). The resultant protein solution is analysed *via* gel electrophoresis under UV irradiation to test if the rhodamine chromophore has been covalently attached to azido–(K1) RNase A (Fig. 5b). According to the gel image, the protein product displays a gel band with intense emission (Fig. 5b, lane 1). On the other hand, azido–(K1)RNase A (**13**) (Fig. 5b, lane 2) and native RNase A (**12**) as the control which carries no azido group (Fig. 5b, lane 3) only reveal non-fluorescent gel bands. This study demonstrates the availability of the introduced azido group for subsequent Staudinger ligation in aqueous media at physiological pH as well as the reactivity of the phosphine probe, rho–TEO–phosphine (**21**), toward azides.

Discussion of the advantages and limitations of the KPL approach

KPL represents a straightforward and efficient approach to site-selectively modify endogenous proteins and therefore it complements recombinant techniques. Compared to other endogenous protein labeling methods, KPL reveals both advantages and limitations. On the one hand, the KPL labeling scheme is straightforward and a modified protein variant can be readily achieved directly from the endogenous protein. As a result, high quantities of modified proteins might become accessible in particular if removal of the native protein is not crucial for the downstream reactions (*e.g.* in the preparation of protein biohybrids and high performance protein chips *etc.* where only modified protein can react). Besides, compared to other techniques, *e.g.* N-terminal transamination which employs highly reactive pyridoxal-5-phosphate (PLP),[27] disulfide intercalation[25] which necessitates a reducing reagent to reduce the disulfide bond(s) or others, both the labeling and purification steps of the KPL can be performed under mild, physiological pH conditions.

On the other hand, the synthesis of the bioconjugation reagent can be tedious and often requires multistep reactions. In addition, two to four days are required to accomplish all essential synthesis and purification steps in KPL and upscaling in a batch process has not yet been demonstrated. Additionally, the KPL labeling only occurs at the most reactive lysine residues which might not be the desired site. However, this challenge might be addressed when the origin of the site-selectivity is better understood and more reactive lysine residues can be introduced at the desired position by point mutations.

Conclusions

In summary, the synthesis of the novel bioorthogonal linker, biotin–TEO–azido–NHS (**11**), has been developed allowing the site-selective introduction of a single azido group on proteins *via* the KPL approach. The scheme of the KPL approach contains in total six key steps after which azido–(K1)RNase A (**13**)

Downloaded by University of Science and Technology of China on 23 December 2012
Published on 07 November 2012 on http://pubs.rsc.org | doi:10.1039/C2OB26561C

is achieved from native RNase A (**12**). Azido–(K1)RNase A (**13**) contains a single azido group as shown by the MALDI-ToF-MS spectrum and the modification site is located at K1 as demonstrated by *nano*LC-MS^2 analysis which is consistent with previous biotinylation results.[34] K1 site-selectivity has been further verified from the XIC chromatographs of both modified and unmodified KETAAAKF (**1–8**) peptide fragments suggesting that all detected K1 fragments have been modified. In order to assess the reactivity of the introduced azido group, the water soluble azide-reactive chromophore, rho–TEO–phosphine (**21**), has been prepared facilitating the successful labeling of azido–(K1)RNase A (**13**) yielding a fluorescent rho–(K1) RNase A (**22**) variant.

The work described herein represents an extension of our previous attempt to introduce a bioorthogonal ethynyl group for click labeling *via* the KPL approach.[34] The present study aims at introducing the more attractive bioorthogonal azido group to further broaden the general applicability of KPL. The azido group has proven to be an ideal bioorthogonal anchor since aside from the copper(I)-catalyzed click reaction that both azido and ethynyl can participate in,[20] azido groups also allow other attractive posttranslational modifications such as copper-free click reactions[38] or Staudinger ligations that proceed under mild physiological conditions which can be performed even in living animals. The introduction of the azido group by this new bioconjugation reagent, biotin–TEO–azido–NHS (**11**), has been accomplished following an optimized procedure which allowed increasing the labeling efficacy.

Modified azido-proteins represent attractive macromolecular building blocks that offer the great opportunity to attach a variety of functional groups *e.g. via* Staudinger ligation. This opens up exciting prospects for achieving protein libraries in a straightforward fashion to allow a deeper investigation of the impact of unnatural substituents such as polymer chains, lipids or peptides on cell uptake, cell trafficking and cytotoxicity. Such investigations could have a high impact on the rational design of improved protein or protein–polymer therapeutics. In this context, it would be attractive to further improve the therapeutic potential of clinically relevant RNase A variants by introducing the appropriate post-modifications by high content protein synthesis. The novel bioorthogonal linker presented herein could be applied in principle for the site-selective introduction of an azido group onto other proteins and peptides as well and we have already demonstrated the modification of the protein lysozyme and the peptide hormone somatostatin following this strategy. These findings underline that the KPL approach is not only applicable if the protein of interest carries a lysine residue at the N-terminus since the most accessible lysine residue will be modified. In addition, the KPL approach could probably be extended to recombinant proteins where an accessible lysine residue is introduced *via* a point mutation. In this way, it might be also suitable for dual labeling strategies in combination with cysteine point mutations.

Acknowledgements

We acknowledge funding from the German Research Foundation (DFG) *via* the SFB 625 as well as grant P3246029 and the Volkswagenstifung (Aktenzeichen 86 366) for their financial support of this work.

Notes and references

1 T. L. Foley and M. D. Burkart, *Curr. Opin. Chem. Biol.*, 2007, **11**, 12–19.
2 M. B. Francis and I. S. Carrico, *Curr. Opin. Chem. Biol.*, 2010, **14**, 771–773.
3 L. Yi, H. Sun, A. Itzen, G. Triola, H. Waldmann, R. S. Goody and Y. W. Wu, *Angew. Chem., Int. Ed. Engl.*, 2011, **50**, 8287.
4 M. Hejjaoui, S. Butterfield, B. Fauvet, F. Vercruysse, J. Cui, I. Dikiy, M. Prudent, D. Olschewski, Y. Zhang, D. Eliezer and H. A. Lashuel, *J. Am. Chem. Soc.*, 2012, **134**, 5192.
5 E. Basle, N. Joubert and M. Pucheault, *Chem. Biol.*, 2010, **17**, 213–227.
6 H. B. Moeller, J. Praetorius, M. R. Rutzler and R. A. Fenton, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 424–429.
7 X. Michalet, S. Weiss and M. Jager, *Chem. Rev.*, 2006, **106**, 1785–1813.
8 J. Y. Yang and W. Y. Yang, *J. Am. Chem. Soc.*, 2009, **131**, 11644–11645.
9 J. M. Harris and R. B. Chess, *Nat. Rev. Drug Discovery*, 2003, **2**, 214–221.
10 S. J. Bell, C. M. Fam, E. A. Chlipala, S. J. Carlson, J. I. Lee, M. S. Rosendahl, D. H. Doherty and G. N. Cox, *Bioconjugate Chem.*, 2008, **19**, 299–305.
11 G. Walsh, *Drug Discovery Today*, 2010, **15**, 773–780.
12 S. D. Tilley and M. B. Francis, *J. Am. Chem. Soc.*, 2006, **128**, 1080–1081.
13 Y. Kim, S. O. Ho, N. R. Gassman, Y. Korlann, E. V. Landorf, F. R. Collart and S. Weiss, *Bioconjugate Chem.*, 2008, **19**, 786–791.
14 A. J. Link, M. L. Mock and D. A. Tirrell, *Curr. Opin. Biotechnol.*, 2003, **14**, 603–609.
15 K. L. Kiick, E. Saxon, D. A. Tirrell and C. R. Bertozzi, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 19–24.
16 L. Wang, A. Brock, B. Herberich and P. G. Schultz, *Science*, 2001, **292**, 498–500.
17 H. S. Park, M. J. Hohn, T. Umehara, L. T. Guo, E. M. Osborne, J. Benner, C. J. Noren, J. Rinehart and D. Soll, *Science*, 2011, **333**, 1151–1154.
18 A. Dondoni, *Angew. Chem. Int., Ed. Engl.*, 2008, **47**, 8995–8997.
19 J. E. Moses and A. D. Moorhouse, *Chem. Soc. Rev.*, 2007, **36**, 1249–1262.
20 H. C. Kolb, M. G. Finn and K. B. Sharpless, *Angew. Chem., Int. Ed. Engl.*, 2001, **40**, 2004–2021.
21 H. Pfeiffer, A. Rojas, J. Niesel and U. Schatzschneider, *Dalton Trans.*, 2009, 4292–4298.

This journal is © The Royal Society of Chemistry 2013

Downloaded by University of Science and Technology of China on 23 December 2012
Published on 07 November 2012 on http://pubs.rsc.org | doi:10.1039/C2OB26561C

22 N. Li, R. K. Lim, S. Edwardraja and Q. Lin, *J. Am. Chem. Soc.*, 2011, **133**, 15316–15319.

23 J. Z. Yao, C. Uttamapinant, A. Poloukhtine, J. M. Baskin, J. A. Codelli, E. M. Sletten, C. R. Bertozzi, V. V. Popik and A. Y. Ting, *J. Am. Chem. Soc.*, 2012, **134**, 3720–3728.

24 D. Rabuka, *Curr. Opin. Chem. Biol.*, 2010, **14**, 790–796.

25 S. Shaunak, A. Godwin, J. W. Choi, S. Balan, E. Pedone, D. Vijayarangam, S. Heidelberger, I. Teo, M. Zloh and S. Brocchini, *Nat. Chem. Biol.*, 2006, **2**, 312–313.

26 J. Alam, T. H. Keller and T. P. Loh, *J. Am. Chem. Soc.*, 2010, **132**, 9546–9548.

27 J. M. Gilmore, R. A. Scheck, A. P. Esser-Kahn, N. S. Joshi and M. B. Francis, *Angew. Chem., Int. Ed. Engl.*, 2006, **45**, 5307–5311.

28 G. B. Postnikova, *Biokhimiya*, 1994, **59**, 1445–1457.

29 K. F. Geoghegan and J. G. Stroh, *Bioconjugate Chem.*, 1992, **3**, 138–146.

30 S. Tsukiji, M. Miyagawa, Y. Takaoka, T. Tamura and I. Hamachi, *Nat. Chem. Biol.*, 2009, **5**, 341–343.

31 M. Kunishima, S. Nakanishi, J. Nishida, H. Tanaka, D. Morisaki, K. Hioki and H. Nomoto, *Chem. Commun. (Camb)*, 2009, 5597–5599.

32 G. Chen, A. Heim, D. Riether, D. Yee, Y. Milgrom, M. A. Gawinowicz and D. Sames, *J. Am. Chem. Soc.*, 2003, **125**, 8130–8133.

33 Z. Chen, B. V. Popp, C. L. Bovet and Z. T. Ball, *ACS Chem. Biol*, 2011, **6**, 920–925.

34 X. Chen, K. Muthoosamy, A. Pfisterer, B. Neumann and T. Weil, *Bioconjugate Chem.*, 2012, **23**, 9.

35 J. Szychowski, A. Mahdavi, J. J. Hodas, J. D. Bagert, J. T. Ngo, P. Landgraf, D. C. Dieterich, E. M. Schuman and D. A. Tirrell, *J. Am. Chem. Soc.*, 2010, **132**, 18351–18360.

36 A. A. Poloukhtine, N. E. Mbua, M. A. Wolfert, G. J. Boons and V. V. Popik, *J. Am. Chem. Soc.*, 2009, **131**, 15769–15776.

37 L. Chaisemartin, P. Chinestra, G. Favre, C. Blonski and J. C. Faye, *Bioconjugate Chem.*, 2009, **20**, 847–855.

38 P. V. Chang, J. A. Prescher, E. M. Sletten, J. M. Baskin, I. A. Miller, N. J. Agard, A. Lo and C. R. Bertozzi, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 1821–1826.

39 S. S. van Berkel, M. B. van Eldijk and J. C. van Hest, *Angew. Chem. Int., Ed. Engl.*, 2011, **50**, 8806–8827.

40 C. I. Schilling, N. Jung, M. Biskup, U. Schepers and S. Brase, *Chem. Soc. Rev.*, 2011, **40**, 4840–4871.

41 J. M. Chalker, G. J. Bernardes and B. G. Davis, *Acc. Chem. Res.*, 2011, **44**, 730–741.

42 H. Taskent-Sezgin, J. Chung, P. S. Banerjee, S. Nagarajan, R. B. Dyer, I. Carrico and D. P. Raleigh, *Angew. Chem., Int. Ed. Engl.*, 2010, **49**, 7473–7475.

43 W. Ardelt, K. Shogen and Z. Darzynkiewicz, *Curr. Pharm. Biotechnol.*, 2008, **9**, 215–225.

44 M. Lillich, X. Chen, T. Weil, H. Barth and J. Fahrer, *Bioconjugate Chem.*, 2012, **23**, 1426–1436.

45 M. J. Thompson and R. A. Goldstein, *Proteins*, 1996, **25**, 38–47.

46 M. K. Sinha, O. Reany, M. Yefet, M. Botoshansky and E. Keinan, *Chemistry*, 2012, **18**, 5589–5605.

47 P. V. Chang, J. A. Prescher, M. J. Hangauer and C. R. Bertozzi, *J. Am. Chem. Soc.*, 2007, **129**, 8400–8401.

48 T. Nguyen and M. B. Francis, *Org. Lett.*, 2003, **5**, 3245–3248.

49 S. T. Laughlin, N. J. Agard, J. M. Baskin, I. S. Carrico, P. V. Chang, A. S. Ganguli, M. J. Hangauer, A. Lo, J. A. Prescher and C. R. Bertozzi, *Methods Enzymol.*, 2006, **415**, 230–250.

50 P. M. Mendes, *Chem. Soc. Rev.*, 2008, **37**, 2512–2529.