

Organic & Biomolecular Chemistry

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

Photo-induced Conjugation of Tetrazoles to Modified and Native Proteins

Cite this: DOI:

10.1039/x0xx00000x

Received 00th January 2012,
Accepted 00th January 2012

DOI: 10.1039/x0xx00000x

www.rsc.org/

Winna Siti,^a Amit Kumar Khan^a, Hans-Peter M. de Hoog,^b Bo Liedberg^a and Madhavan Nallani^a

Bio-orthogonal chemistry has been used widely for conjugation of polymer molecules to proteins. Here, we demonstrate conjugation of polyethylene glycol (PEG) to bovine beta-lactoglobulin (BLG) by the photo-induced cyclo-addition of tetrazole-appended PEG and allyl-modified BLG. During the course of conjugation, a significant side-reaction was found to occur for the conjugation of PEG-tetrazole to native BLG. Further exploration of the underlying chemistry reveals that the presence of a tryptophan residue is sufficient for conjugation of tetrazole-modified molecules.

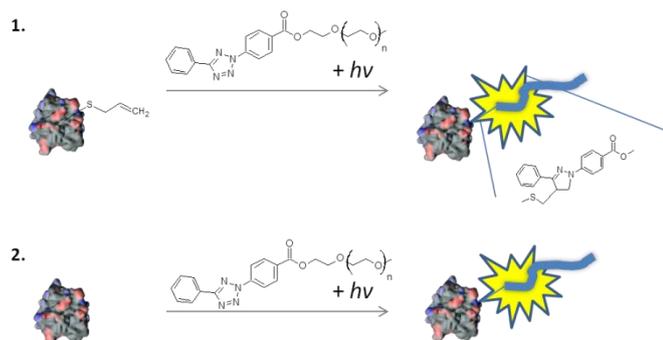
The conjugation of polymers to biomolecules or aggregates thereof forms an important means to enhance their stability or stimulate their controlled self-assembly, constituting an exciting area of research that aims to develop sophisticated drug carriers. An important aspect of these endeavours is the chemistry by which polymers are linked to the biomolecule.¹⁻⁵ Here, the application of click chemistry, which in its ideal form encompasses the seamless, high-yield, and bio-orthogonal conjugation of the two macromolecules, has been especially popular. The most widely adopted reactions used in such conjugations are those based on the “classical” click reaction involving (strained) alkynes and azides. However, variations on this theme have constantly been introduced in the past years, most notably involving conjugations based on inverse electron-demand Diels Alder reactions such as that of tetrazines and strained alkenes as well as the reaction of thiols with terminal alkenes, which may or may not be light activated (so-called thiol-ene chemistry).⁶⁻¹⁰

Undoubtedly, azide remains a popular choice as it can be relatively easily encoded into biomolecules, either post-translationally or, by genetic encoding.

Another approach, which relies on a relatively simple conjugation moiety that is amenable to genetic encoding, is based on the photo-conjugation of tetrazole-appended molecules to alkene moieties – where an alkene can be introduced into proteins by non-canonical amino acids such as homoallylglycine and *O*-allyl-tyrosine.¹¹⁻¹⁶ The reaction has the additional benefit of giving rise to a fluorescent linker moiety (a pyrazoline) which makes *in-vivo* reactivity easy to monitor. Although having received some initial attention from different groups, it has lagged behind in types of applications reported, perhaps because it has been shown to be quenched by high concentrations of amines and carboxylic acids, which evidently are ubiquitous in biological systems, as has recently been investigated into some detail by Zhao *et al.*¹⁷ Still, its high rate, photo-inducibility and the fluorescence of the resulting pyrazoline pose some genuine advantages for niche areas where conjugation is difficult to induce and monitor otherwise such as under *in-vivo* conditions.

Recently, we have demonstrated the application of this reaction in the straightforward covalent attachment of enzymes to polymersomes.¹⁸ Specifically, we have shown that by modifying the biomolecule with the simple bisaryl tetrazole 4-(5-phenyltetrazol-2-yl)benzoic acid^{12,19} (hereafter referred to as tetrazole (**tz**)) the resulting macromolecule could readily be conjugated to methacrylate moieties on the surface of polymersomes. Reaction progress and purification were greatly facilitated by the presence of the

fluorescent moiety on the eventual conjugate. In this study, we report our findings obtained during the course of a follow-up study with the greater aim of constructing polymer protein conjugates, where we initially were mainly interested in the photo-induced conjugation of simple polyethylene glycol (PEG) chains to a protein of interest. The original experimental design, however, gave rise to a significant side reaction, which we found interesting to explore in greater detail. Hence, this paper focuses on the exploration of this side reaction, the chemistry behind it and its general applicability in the conjugation of polymers to proteins.



Scheme 1 Preparation of BLG-PEG conjugates by the photo-induced cycloaddition of tetrazole-modified PEG. Apart from the procedure outlined (1), which involves reaction of allyl-modified BLG, a significant side-reaction occurs when performing the photo-induced cyclo-addition in presence of the native protein (i.e., without allyl modification) (2).

Results

Our original design involved the use of bovine beta-lactoglobulin (BLG) as the protein of choice, as the protein itself shows interesting self-assembly behaviour and is readily available in virtually pure format.²⁰ Furthermore, it carries two cysteines, of which only one is readily accessible,²⁰⁻²² allowing the site-specific conjugation of a single polymer chain. To make the BLG amenable to photo-click conjugation, cysteine was derivatized to an alkene by reacting it with allyl-chloride as reported by Chalker *et al.*²³ Tetrazole-appended PEG (PEG-tz) was prepared following procedures reported in literature (for detailed preparation see SI).^{12,18-20,23}

Since – upon photo-irradiation – the pyrazoline exhibits strong fluorescence emission between 470–540 nm,¹² the conjugation of PEG-tz to allyl-modified BLG (BLG-allyl) was analyzed by fluorescence spectroscopy (Fig 1a, for UV spectrum see Fig S1). As a control experiment, BLG without allyl group was subjected to photo-irradiation at 302 nm for 15 min in presence of PEG-tz, followed by analysis with fluorescence spectroscopy. Interestingly, the results showed that photo-irradiation of allyl-BLG with PEG-tz as well as the native BLG with PEG-tz reproducibly generated fluorescent adducts. This was confirmed by gel electrophoresis, showing a clear band at 22 kDa and a minor band at 25 kDa, the latter band possibly indicating formation of doubly conjugated BLG (Fig 1b). Judging from the shape of the fluorescence spectrum, the BLG-allyl:PEG-tz conjugate constituted a different species than that of the “direct” conjugate between PEG-tz and native BLG. The unimodal peak of the former conjugate is reminiscent of that observed earlier for pyrazoline formation and therefore most likely shows the successful addition of PEG-tz to BLG-allyl. The nature of the “direct” conjugate was, however, unclear. Because the reaction occurs to a significant extent under presumed bio-orthogonal conditions, we considered it worthwhile to study its occurrence in more detail to shed light on the protein site involved.

First, the efficiency of the conjugation was studied by reacting increasing equivalents of PEG-tz with BLG (0.5–5 eqs of PEG-tz vs BLG), irradiating the relevant mixtures for 15 min, and evaluation of the extent of conjugation by SDS-PAGE (Fig S2). The analysis suggested an increase in conjugate formation with increasing concentration of PEG-tz, with 5 eqs of PEG-tz showing over 50 % conversion, which is, given the short reaction time and the macromolecular nature of the comparatively high. The fluorescence spectra of the PEG-tz:BLG conjugates showed an interesting evolution with increasing amounts of PEG-tz when compared to that of PEG-tz:BLG-allyl (Fig 1c–e). For PEG-tz:BLG conjugate at low concentration (0.5 eq), a peak was observed at 415 nm, which was not observed for allyl-modified BLG. This peak increased only slightly in intensity at higher PEG-tz concentrations, while the peak at 506 nm showed a significant and steady increase. The observation may be attributed to reaction of PEG-tz at multiple sites of BLG, dependent on the concentration of the protein. Indeed, MALDI-ToF mass spectrometry showed the occurrence of peaks at m/z 21.9 kDa and 25.1 kDa, indicative of singly conjugated BLG as well as doubly conjugated BLG (Fig S3). Overall, the experiments demonstrate that, at least for BLG, PEG-tz, which is essentially a fully water-soluble tetrazole derivative, directly reacts to a considerable extent with a site on the protein, without the presence of an externally introduced dipolarophile.

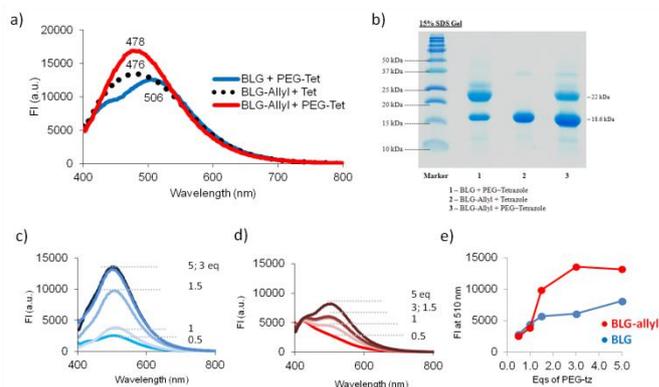


Figure 1 Characterization of the products formed by photo-induced cyclo-addition of 5 eqs of PEG-tz with BLG-allyl and native BLG by fluorescence spectroscopy (a) and SDS-PAGE (b), indicative of conjugate formation. The fluorescence spectra after reaction of BLG-allyl and native BLG with increasing equivalents of PEG-tz (0.5 – 5 eqs) show formation of chemically distinct conjugates (c–e).

Because of the simplicity of the observed side reaction, we then investigated the applicability of the reaction to other proteins / enzymes: horseradish peroxidase (HRP), *Candida antarctica* lipase (CalB), and lysozyme, being enzymes commonly used in bionanotechnology. These were reconstituted in PBS and, without any modification, incubated with PEG-tz at various concentrations under irradiation for 15 min at 302 nm. Interestingly, all enzymes showed formation of a conjugate with PEG-tz, as was evident from fluorescence, SDS-PAGE and MALDI-ToF analysis of the reaction products (see further). This then indicated that the side-reaction observed is common and occurs readily with a diverse subset of enzymes.

More in detail, lysozyme exhibited similar behaviour as that observed for BLG with respect to its fluorescence spectrum, with peak emission at 500 nm and an increasing extent of conjugate formation with increasing equivalents of PEG-tz (Fig 2). SDS-PAGE showed the clear formation of the PEG-tz conjugate, which was corroborated with MALDI-ToF, showing a peak at 17.8 kDa, albeit

that the extent of conjugate formation as observed by MALDI-ToF did not fully correlate with that observed by SDS-PAGE. As this discrepancy was also apparent for the BLG-PEG conjugates, we attribute this to a reduced propensity for desorption / ionization of the PEG-conjugate as compared to the free protein. The observed rapid conjugation of unmodified lysozyme to PEG-tz is especially interesting since its conjugation with PEG-tz has been studied earlier, showing no apparent conjugation of unmodified lysozyme.^{12,24}

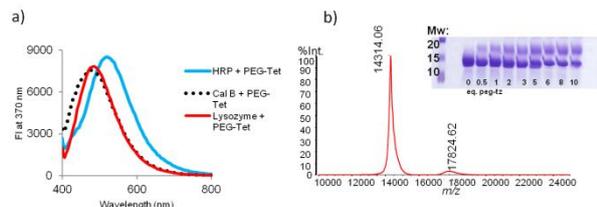


Figure 2 Photo-induced cyclo-addition of PEG-tz to unmodified HRP, CalB, and lysozyme (a). For lysozyme, conjugate formation was confirmed by MALDI-ToF and SDS-PAGE (b).

With respect to CalB, fluorescence as a function of PEG-tz concentration after irradiation exhibited somewhat dissimilar behaviour (Fig S4a). At low concentrations of PEG-tz an intense peak was observed at 450 nm for CalB, in analogy with BLG, yet this peak initially decreased and gradually shifted to a bimodal peak with a maximum of 515 nm at higher PEG-tz / enzyme ratios. SDS-page and MALDI-ToF clearly indicated the reaction of a single site on the protein, which is in contrast to BLG that showed reaction at multiple sites. Finally, HRP showed a steadily increasing monomodal emission peak as a function of concentration of PEG-tz centred at 525 nm (Fig S4b). HRP, furthermore, displayed relatively high fluorescence intensity as compared to the other proteins and therefore appears to exhibit a different reactivity than BLG, CalB and lysozyme. We tentatively suggest this observation to stem from the presence of the heme prosthetic group. Indeed, the heme contains a chemically modifiable double bond, which we suggest to be the site of reaction.²⁵ Considering the extent of conjugation, a predominantly singly conjugated species was observed with MALDI-ToF in addition to a small peak at 51 kDa, indicating the attachment of two PEG chains. SDS-PAGE was not entirely conclusive in this particular case because of the relatively broad bands of the native HRP, probably stemming from some apo-HRP in the sample (as also apparent from MALDI-ToF of HRP itself).²⁵

In summary, the heterogeneous fluorescence behaviour of the conjugates pointed to the contribution of a fluorescent amino acid, most likely tryptophan, to the reaction. Its role was confirmed by the following lines of evidence. First, the relatively small peptide JR2EC (see Fig S5 for sequence),²⁶ did not show conjugation as judged from the absence of a fluorescent product at all PEG-tz / peptide ratios (Fig 3a). Strong fluorescence was observed for JR2EC, however, upon modifying the sole cysteine with allyl chloride. The amino acid sequence revealed the absence of tryptophan residues in the polypeptide, pointing to a role of the amino acid in the conjugation reaction. We then probed the selectivity of the reaction by assessing the reactivity of tetrazole only (without PEG) with two additional custom-designed 12 amino-acid peptides, one containing 3 histidine residues and one containing the 3 histidines plus a single tryptophan at the C-terminal (Fig 3b). Here we chose histidine since we suspected it could display reactivity to tetrazole as well. Analyzing the products before and after irradiation by MALDI-ToF, showed a mass increase of 238.9 for the tryptophan-containing peptide, indicative of the addition of tetrazole. The peptide containing only His and no tryptophan was found to be unreactive.

Finally, reaction of PEG-tz with tryptophan led to considerable fluorescence as well, while reaction with the control amino acid (histidine) did not show noticeable fluorescence (Fig 3c). Although we could not extract detailed structural information by 1H-NMR we acquired MALDI-ToF spectra before and after irradiation. The spectra showed a clear shift in *m/z* values between PEG-tz before and after irradiation, amounting to a mass difference of 176.9. Assuming loss of dinitrogen during irradiation of PEG-tz, this is close to the theoretically expected value of 176.21 (+ 28.01 = 204.23), suggesting addition of tryptophan to the nitrilimine. Here, it should be noted that the MALDI-ToF does not allow differentiation between quenching of the nitrilimine by the amine or cycloaddition. Still, the occurrence of the formation of a fluorescent moiety indicates formation of pyrazoline.

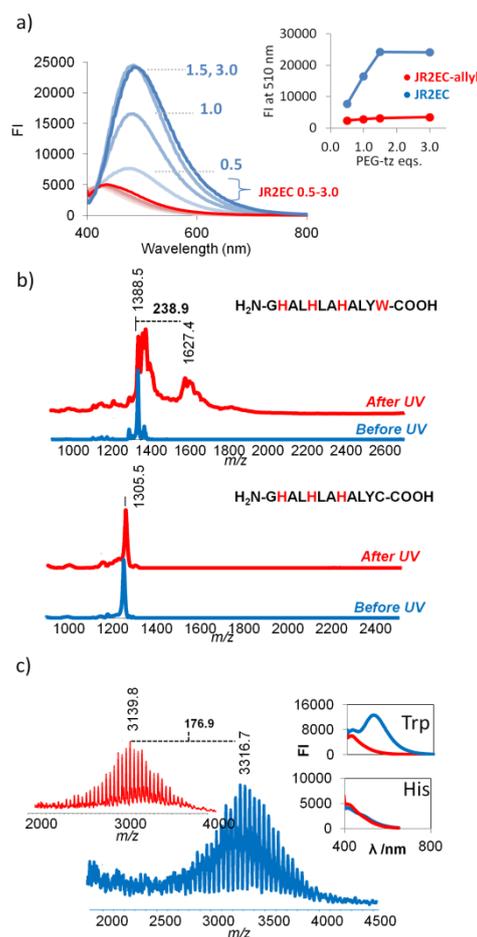


Figure 3 a) Fluorescence spectra showing formation of fluorescent conjugates with increasing eqs of PEG-tz upon irradiation with UV for unmodified JR2EC (red), and allyl-modified JR2EC (blue). b) Selectivity of the reaction of tetrazole upon incubation, under UV irradiation ($\lambda = 302$ nm), with the two custom peptides, as probed by MALDI-ToF. For the tryptophan-functional peptide (top), which carries a tryptophan at the C-terminal, a shift is observed of 238.9 mass units, corresponding to addition of photolysed tetrazole (266.25 – 28.01 = 238.24). For the control peptide (bottom), which lacks the tryptophan and instead carries a cysteine, no reaction is observed. c) MALDI-ToF of PEG-tz before (red) and after photo-irradiation in presence of excess tryptophan (blue), indicating addition of the amino acid to the polymer. The fluorescence spectrum shows the reaction mixture before (red trace) and after (blue) irradiation of PEG-tz with tryptophan or histidine.

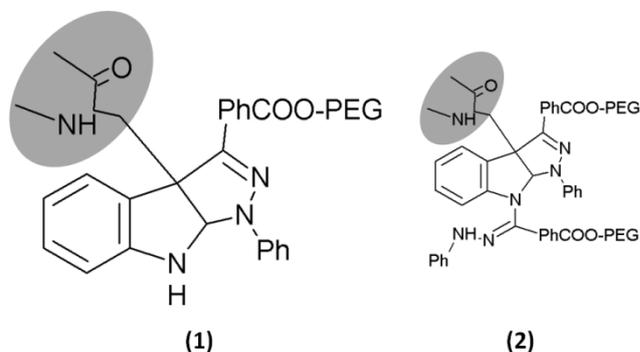
Discussion

Overall, the results presented above strongly suggest the involvement of tryptophan in the direct photo-induced conjugation of tetrazoles to proteins as observed here for PEG-modified 4-(5-phenyltetrazol-2-yl)benzoic acid where it should be noted that all enzymes studied have one or more tryptophan residues present (Fig S6). The reaction observed seems to be different than that reported by Zhao *et al.*, who described quenching of the nitrilimine by the amino acid amine moiety, a reaction that was observed irrespective of the nature of the amino acid. In our case, the reaction appears to be specific to tryptophan and therefore does not seem to be affected by quenching by amines, which are ubiquitously present. We suggest that this may result from the reaction stoichiometry (an excess of nucleophile in the case of Zhao *et al.* versus equimolar or excess amounts of tetrazole in the present study) and the use of aqueous buffer, which is known to accelerate cycloaddition,¹⁷ as a result of the use of fully water soluble PEG-tz. Very likely, the protein environment may play a role as well. BLG, for instance, contains a tryptophan residue in its binding site, and it may be envisioned that the PEG-tz interacts here, as the tetrazole itself carries quite some hydrophobic character. The close proximity may explain the relatively high yield and speed of the reaction, with > 50 % of the protein having reacted after 15 minutes irradiation. Further evidence for an effect of the protein structure on reactivity is provided by the observations on lysozyme and CalB, both of which contain multiple tryptophan residues, with lysozyme having three solvent-exposed residues. In both cases, predominantly singly conjugated proteins were observed that indicate the presence of a single tryptophan with optimal reactivity.

Given the compelling evidence for the direct reaction of tryptophans with photo-activated nitrilimines, it appears somewhat remarkable that this reaction has not been observed earlier by others or even ourselves.^{12, 13, 15-16, 18} In our case, this might result from the experimental design, which involved modifying the protein with tetrazole, instead of a double bond, as shown here. Considering other reports, the presence of an alkene-modified amino acid of high reactivity may lead to the alkene being preferentially modified, precluding reaction of the tetrazole with tryptophan. The relative abundance of papers that report incorporation of such alkene-appended amino-acids suggests that this indeed is the case. For reactions conducted in-vivo this may only be part of the explanation, as a considerable "pool" of tryptophan may be present. Here, the concentration of the alkene may have been essential and / or the proximity of the reactant to the probe. Furthermore, it should be noted that the reaction documented here involves excitation at $\lambda = 302$ nm so that photo-activation of tryptophan may play an active role; nitrilimines generated from tetrazoles excitable at higher wavelengths may be less, or not at all, susceptible to reaction with tryptophan.

Investigation of less recent literature (i.e., before the advent of bio-orthogonal chemistry) reveals that direct addition of nitrilimines to tryptophan is not unprecedented. Indeed, in 1973, Ruccia *et al.* reported the reaction of nitrilimines with indole - being the functional side-group of tryptophan.²⁷ For the most closely related indole, cycloadduct (1) was observed, albeit it in a very low yield. This then is the most likely product of our current investigation. Apart from this cycloadduct, a bisadduct (2) was observed for indole carrying a single free nitrogen at the 3-position, which may explain the formation of di-substituted protein in case of HRP (with only one tryptophan) and tri-substituted BLG (carrying only two tryptophans). Definite proof of the occurrence of such di-substituted protein

adducts requires a more detailed investigation, however, and is outside the scope of the current manuscript.



Scheme 2 Suggested structure (1) of the possible conjugate formed by photo-induced cyclo-addition of PEG-tz to tryptophan residues in proteins. Structure (2) shows a possible bisadduct of the PEG-tz. The amide back bone of the protein is shown in grey.

Conclusion

In brief, during our exploration of the photo-induced conjugation of tetrazole-modified PEG (PEG-tz) to alkene-modified proteins, we observed significant fluorescent product formation for the photo-addition of PEG-tz to unmodified protein as well. Exploration of this reaction using BLG showed the side reaction to be nearly as efficient and high-yielding as the original cyclo-addition of photolyzed tetrazoles to alkenes. Exploration of its applicability showed that it reacted with all proteins (in our case three different enzymes) except for a small peptide that lacked tryptophan. Further investigation revealed that PEG-tz rapidly reacts with tryptophan to form a fluorescent adduct. Except from the potential synthetic utility of this reaction, the results suggest that, at least for tetrazoles that are photolyzed at low UV wavelengths, care should be taken when interpreting conjugation solely on fluorescent product formation in conditions where tryptophan can be ubiquitous, as is the case for in-vivo conditions.

Notes and references

^a Centre for Biomimetic Sensor Science, School of Materials Science and Engineering, Nanyang Technological University, 50 Nanyang Drive 637553, Singapore. E-mail: MNallani@ntu.edu.sg

^b ACM Biolabs, Nanyang Technological University, 50 Nanyang Drive 637553, Singapore.

† Electronic Supplementary Information (ESI) available: experimental procedures, fluorescence spectra, MALDI-ToF spectra and SDS-PAGE gels, as well as structures of the proteins employed. See DOI: 10.1039/c000000x/

- 1 R. Duncan, *Nat. Rev. Drug Discovery*, 2003, **2**, 347-360.
- 2 R. M. Broyer, G. N. Grover, H. D. Maynard, *Chem. Commun.*, 2011, **47**, 2212-2226.
- 3 P. Wonganan, M. A. Croyle, *Viruses*, 2010, **2**, 468-502.
- 4 A. Kolate, D. Baradia, S. Patil, I. Vhora, G. Kore, A. Misra, *J. Control. Rel.*, 2014, **192**, 67-81.
- 5 N. Nischan, C.P.R. Hackenberger, C.P.R., *J. Org. Chem.*, 2014, **79**, 10727-10733.
- 6 J. F. Lutz, *Angew. Chem. Int. Ed.*, 2007, **46**, 1018-1025.
- 7 M. D. Best, *Biochemistry*, 2009, **48**, 6571-6584.

Journal Name

- 8 R. K. V. Lim, Q. Lin, *Chem. Commun.*, 2010, **46**, 1589-1600.
- 9 C. P. Ramil, Q. Lin, *Chem. Commun.*, 2013, **49**, 11007-11022.
- 10 A. C. Fahrenbach, J. F. Stoddard, *Chem. Asian J.*, 2011, **6**, 2660-2669.
- 11 W. Song, Y. Wang, J. Qu, Q. Lin, *J. Am. Chem. Soc.*, 2008, **130**, 9654-9655.
- 12 W. Song, Y. Wang, J. Qu, M. M. Madden, Q. Lin, *Angew. Chem. Int. Ed.*, 2008, **47**, 2832-2835.
- 13 Y. -J. Lee, B. Wu, J. E. Raymond, Y. Zeng, X. Fang, K. L. Wooley, W. R. Liu, *ACS Chem. Biol.*, 2013, **8**, 1664-1670.
- 14 F. Li, H. Zhang, Y. Sun, Y. Pan, J. Zhuo, J. Wang, *Angew. Chem. Int. Ed.*, 2013, **52**, 9700-9704.
- 15 Z. Yu, Y. Pan, Z. Wang, J. Wang, Q. Lin, *Angew. Chem. Int. Ed.*, 2012, **51**, 10600-10604.
- 16 P. An, Z. Yu, Q. Lin, *Chem. Commun.*, 2013, **49**, 9920-9922.
- 17 Y. Zhang, W. Liu, Z. K. Zhao, *Molecules*, 2014, **19**, 306-315.
- 18 H. -P. M. de Hoog, M. Nallani, B. Liedberg, *Polym. Chem.*, 2012, **3**, 302-306.
- 19 S. Ito, Y. Tanaka, A. Kakehi, K. Kondo, K., *Bull. Chem. Soc. Jpn.*, 1976, **49**, 1920-1923.
- 20 N. Kitabatake, R. Wada, Y. Fujita, *J. Agric. Food Chem.*, 2001, **49**, 4011-4018.
- 21 K. A. Piez, E. W. Davie, J. E. Folk, J. A. Gladner, *J. Biol. Chem.*, 1961, **236**, 2912-2916.
- 22 K. Sakai, K. Sakurai, M. Sakai, M. Hoshino, Y. Goto, *Protein Science*, 2000, **9**, 1719-1729.
- 23 J. M. Chalker, Y. A. Lin, O. Boutureira, B. G. Davis, *Chem. Commun.*, 2009, 3714-3716.
- 24 Z. Yu, R. K. V. Lim, Q. Lin, *Chem. Eur. J.*, 2010, **16**, 13325-13329.
- 25 L. Huang, G. Wojciechowski, P. R. O. de Montellano, *J. Biol. Chem.*, 2006, **281**, 18983-18988.
- 26 D. Aili, M. Mager, D. Roche, M.M. Stevens, *Nano Lett.*, 2011, **11**, 1401-1405.
- 27 M. Ruccia, N. Vivona, G. Cusmano, M. L. Marino, F. Piozzi, *Tetrahedron*, 1973, **29**, 3159-3164.