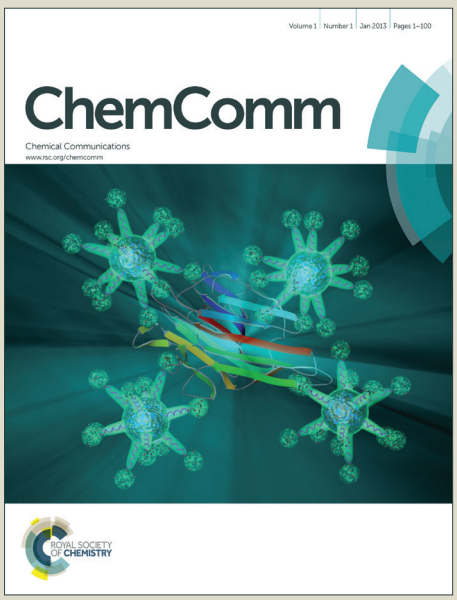
ChemComm

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/chemcomm

Journal Name ChemComm Dynamic Article Links ▶

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

Protein Stapling via Azide-Alkyne Ligation†

Diya M. Abdeljabbar,[a] Frank J. Piscotta,[a] Siyan Zhang[a] and A. James Link[a,b]*

Received (in XXX, XXX) Xth XXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX
DOI: 10.1039/b000000x

Here we demonstrate a methodology, termed protein stapling, for the introduction of covalent constraints into recombinant proteins. Using the azide-alkyne click reaction as the stapling chemistry, we have improved the thermostability of a model leucine zipper protein. Additionally, stapling the core of the small, globular protein G resulted in improved binding to its target, immunoglobulin G.

Constraining a polypeptide to a restricted set of conformations can endow the molecule with several desirable properties including increased stability in the face of thermal or chemical denaturation, resistance to proteases, and tighter interactions with binding partners. Nature uses an array of different posttranslational modifications to constrain short polypeptides into macrocycles[1] or more exotic shapes.[2] In the laboratory, short peptides, especially helical peptides, have been constrained and fortified by sidechain-sidechain ligation chemistries, a technique now known as peptide stapling.[3-5] Recombinant proteins have also been "stapled" using either the incorporation of two noncanonical amino acids (ncAAs)[6] or an electrophilic ncAA that reacts with a canonical amino acid sidechain.[7] The conformational constraints placed within the proteins using these techniques are not particularly rigid because long-chain amino acids with many freely rotatable bonds were used in the staple. Here we have used residue-specific incorporation techniques[8, 9] to simultaneously introduce two unnatural amino acids, azidohomoalanine (AHA) and *p*-ethynylphenylalanine (PEP) into proteins (Figure 1). This pair of amino acids can be specifically crosslinked via azide-alkyne ligation chemistry,[10, 11] and the resulting linkage (Figure 1) is expected to be quite rigid. AHA can be introduced into proteins as a surrogate for methionine (Met)[12] and is a substrate for the native translational machinery in *E. coli*. The replacement of phenylalanine (Phe) by PEP in recombinant proteins has also been described in *E. coli* and requires a phenylalanyl-tRNA synthetase (PheRS) active-site variant.[13] Here we demonstrate the protein stapling methodology on two proteins: an artificial leucine zipper A1[14] and the IgG-binding domain of protein G.[15] In both cases protein properties were improved, demonstrating for the first time the effects of conformational constraints introduced by protein stapling.

We generated a variant of the artificial leucine zipper A1[14] with the sequence $MAGSH_6GS$ appended to its N-terminus. We hypothesized that this simple protein could be stabilized against thermal denaturation if a rigid azide-alkyne staple were introduced to it. It has been previously demonstrated that an

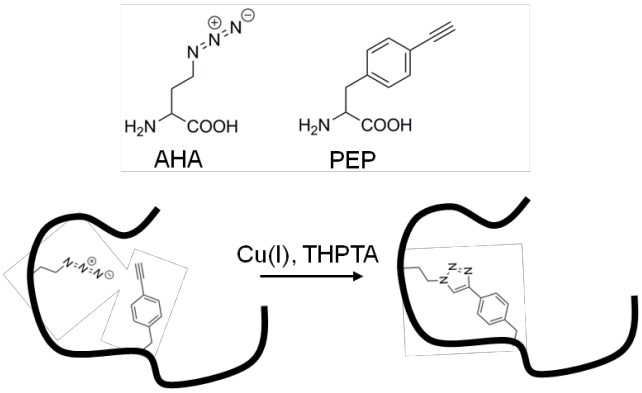


Fig. 1. Protein stapling with azide- and alkyne-bearing amino acids. Top: structures of azidohomoalanine (AHA) and *p*-ethynylphenylalanine (PEP). Bottom: Schematic of the protein stapling reaction between AHA and PEP. The resulting linkage is more rigid than a disulfide bond.

alanine in the second position of a protein allows for efficient removal of the N-terminal amino acid by methionine aminopeptidase (MetAP) in *E. coli*, even when the N-terminal amino acid is AHA.[16] In the case of the A1 protein, this allows for "pseudo site-specific" insertion of a single AHA residue within the protein. Two A1 variants were constructed, one with a central staple and one with a C-terminal staple. The central staple construct includes a single internal ATG codon encoding for Met at position 31 of the protein. A Phe codon was introduced at position 34 of the protein, placing the Met and Phe residues on the same face of the helix comprising the leucine zipper. The C-terminal staple construct has Met at position 52 and Phe at position 55. In both proteins, a second Phe residue is present in the N-terminal portion of the protein (see Figure S1 for complete sequences). A1 protein harboring AHA and PEP was expressed in an *E. coli* BL21 derivative rendered auxotrophic for both Phe and Met via knockouts of the *pheA* and *metE* genes, respectively. Both proteins were successfully produced in M9 minimal media lacking Met and Phe, but supplemented with AHA and PEP (Figure S2). Yields of the purified proteins were reasonable with 5.8 mg/L of the central staple construct and 7.8 mg/L for the C-terminal staple construct. The proteins were purified, dialyzed, lyophilized, and resuspended in dilute sodium phosphate buffer in preparation for CuAAC. Analysis of a tryptic digest of the doubly-substituted proteins by MALDI-MS revealed essentially complete replacement of Met with AHA and near-complete replacement of Phe with PEP (Figure S3), in accordance with previous studies.[12, 13]

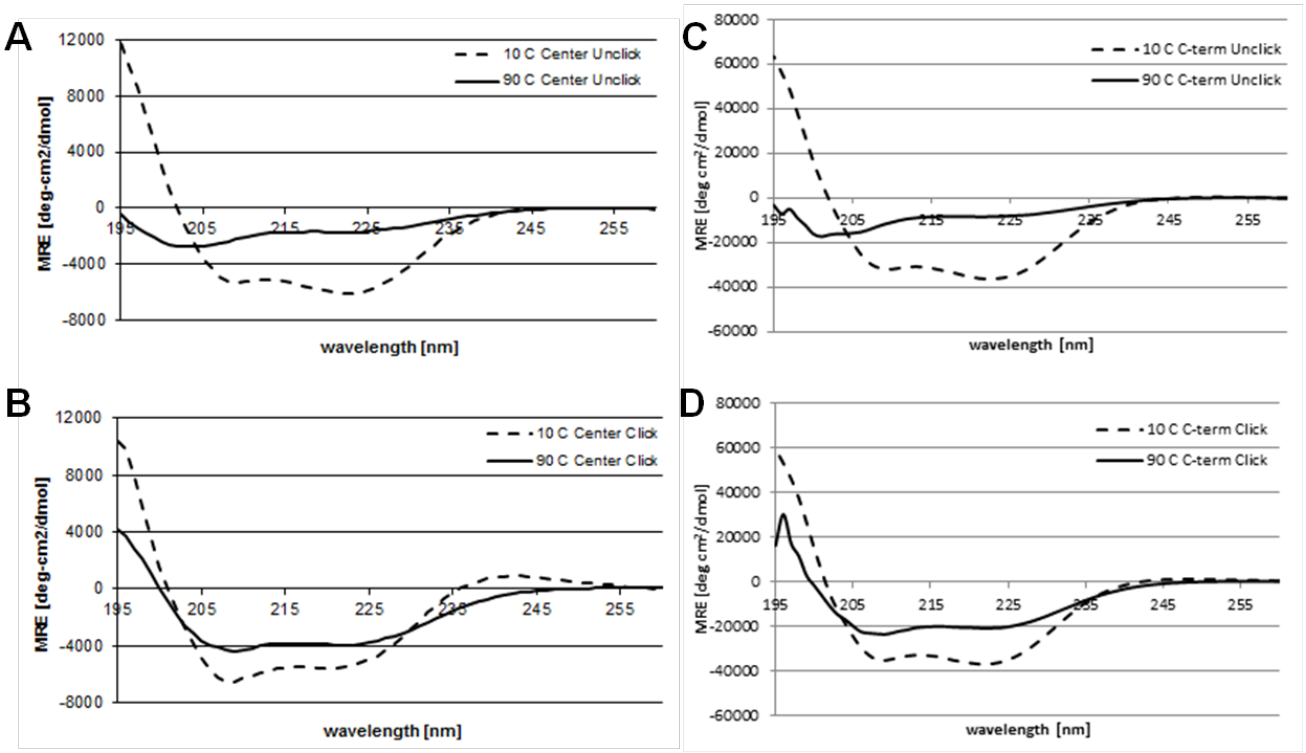


Fig. 2. Circular dichroism analysis of unstapled and stapled leucine zipper A1 at 10 °C and 90 °C. CD spectra of A: unclicked center staple construct, B: clicked center staple construct, C: unclicked C-terminal staple construct, D: clicked C-terminal staple construct. In both the center staple and C-terminal staple constructs, the introduction of the staple results in retention of helicity at 90 °C.

The use of *tris*(triazole) ligands is known to accelerate CuAAC reactions, but the most commonly used ligand, TBTA,[17] is poorly soluble in water and may cause protein precipitation.[18] A more water-soluble, biocompatible ligand, THPTA, has been described,[19] and we utilized this ligand for our protein stapling studies. The requisite Cu(I) catalyst was generated *in situ* by reduction of $CuSO_4$ by ascorbic acid. Reactions containing protein, THPTA, the copper source, and ascorbic acid were incubated at 4 °C for 6-8 h. The click reagents were removed using a centrifugal filter device, and the putatively stapled proteins were examined using circular dichroism (CD) spectroscopy at a range of temperatures (Figure S4). Addition of a single staple to the 67 aa A1 protein, either in the center of the helical region or at its C-terminus, led to a substantial increase in thermostability. (Figure 2) While the "unclicked" forms of the proteins were completely unfolded at 90 °C, the clicked proteins retained about 50% of the helical content of the proteins at 10 °C. These results indicate that a single staple between AHA and PEP can result in stabilization of a helix, and that either a centrally-located staple or a C-terminal staple can provide the stabilization.

We next turned our attention to the stapling of a small, globular protein, the IgG-binding domain of protein G.[15] Our hypothesis was that stapling protein G into a productive confirmation could enhance its affinity toward IgG by decreasing the entropic penalty incurred during the binding of these two proteins. The protein G construct we used here is 72 aa including an N-terminal sequence containing a His tag and a tryptophan for concentration determination purposes (see Figure S5 for entire sequence). There is a phenylalanine residue on the single helix of protein G (position 46 in our construct) which points into the core of the protein. We reasoned that placing a Met residue nearby would allow stapling of the core. Position Leu-21 is in close proximity (4.3 Å from γ-carbon to γ-carbon) to Phe-46 (Figure 3), so we mutated this position to Met. The protein was expressed,

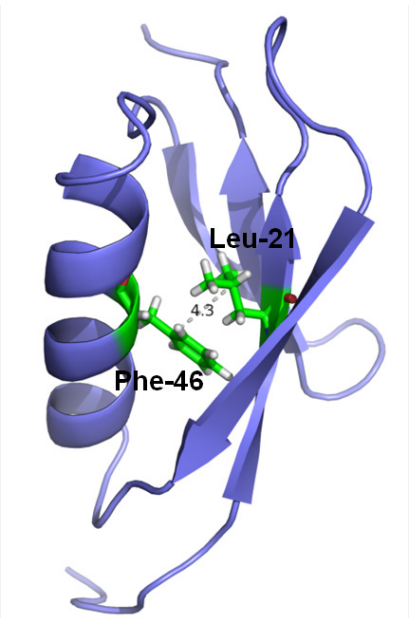


Fig. 3. Structure of the IgG-binding domain of protein G (PDB 2IGG) showing the location of Leu-21 and Phe-46 in the core of the protein. The γ-carbon to γ-carbon distance between these residues is 4.3 Å. Leu-21 was replaced with Met in order to incorporate AHA while Phe-46 was substituted with PEP to generate a protein that can be stapled.

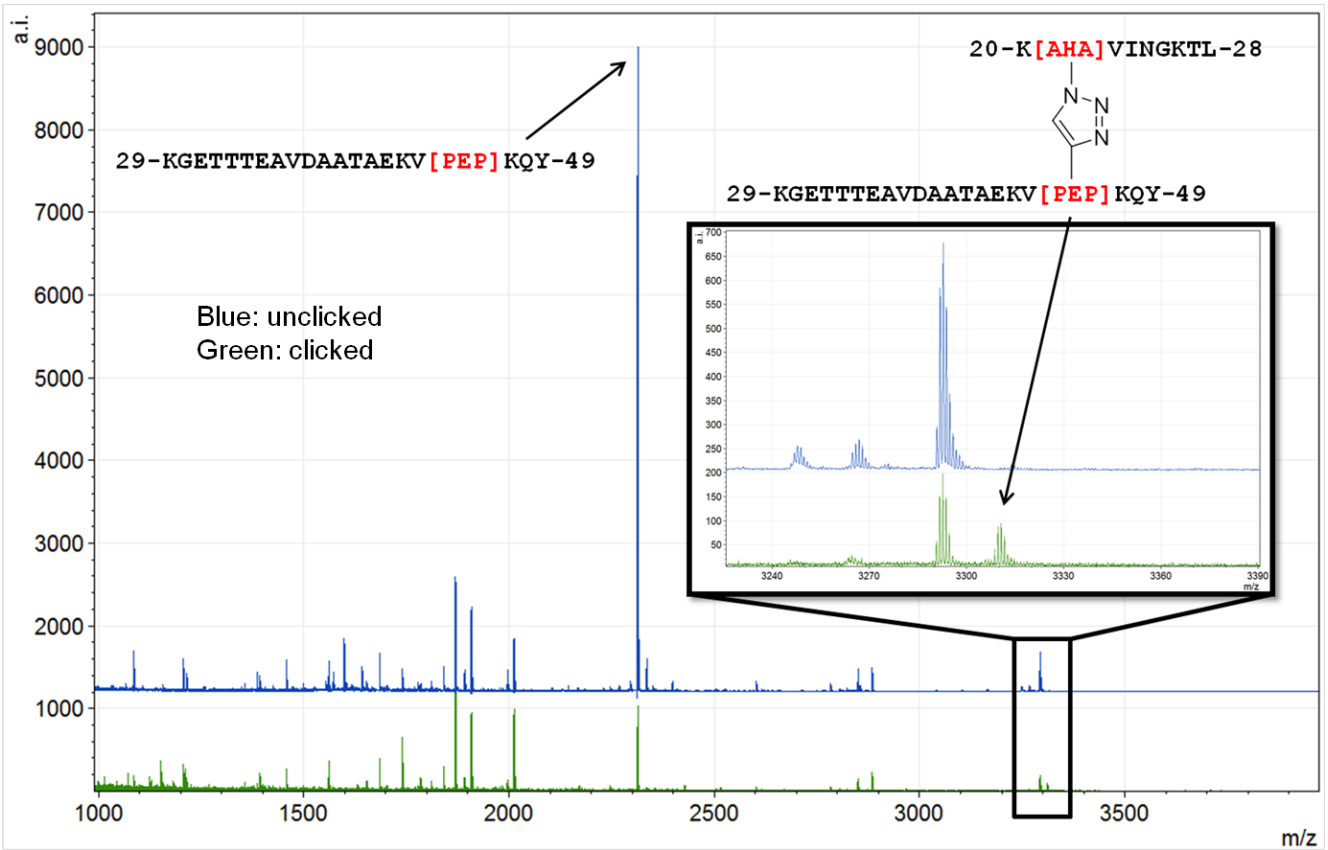


Fig. 4. Mass spectrometric analysis of chymotryptic digest of AHA- and PEP-substituted protein G. The blue spectrum is from a digest of unclicked protein. The prominent peak at 2312.1 amu includes the PEP residue at position 46. The green spectrum is from a digest of the clicked protein. The intensity of the 2312.1 amu peak is reduced, and a new peak at 3310.6 amu appears (inset). This new peak corresponds to the adduct between the AHA- and PEP-containing peptides.

purified (yield 2.5 mg/L), and subjected to the click reaction as described above for A1.

Since CuAAC is completely atom efficient, a mass change cannot be used as a diagnostic for a successful reaction. Instead, the attachment of the chymotryptic fragment containing AHA to the chymotryptic fragment containing PEP was used to determine whether stapling had occurred. For the L21M variant of protein G, the fragment including the AHA in position 21 of protein G has a mass of 999.6 amu while the fragment including PEP at position 46 has a mass of 2312.1 amu. While the fragment containing AHA was not observed, the PEP-containing fragment was prominent in the mass spectrum (Figure 4). After stapling, however, the 2312.1 amu peak was diminished by ~85%, and a new peak with mass 3310.6 amu, corresponding to an adduct between the peptides, appeared (Figure 4). This indicates that the substituted protein is folded in a near native state prior to stapling, and that the stapling reagents can access the core of this small protein.

We next characterized the binding affinity of the unstapled and stapled protein G variants using biolayer interferometry, a technique which provides data on the kinetics of protein-protein interactions. Biosensors coated in protein A were first incubated with total IgG. The association and dissociation of both unstapled and stapled protein G to IgG was then measured at three different concentrations of protein G, 600 nM, 800 nM, and 1000 nM. The data were analysed using both individual sensorgrams (Figure S6) and global analysis over all concentrations tested (Table 1). The clicked protein exhibited a ~4-fold decrease in the equilibrium dissociation constant, K_d, from 598 nM in the unstapled protein to 167 nM in the stapled version. This result indicates that the conformational restriction of a protein via azide-alkyne stapling can lead to an improvement in the binding of the protein to its partners.

Here we have demonstrated azide-alkyne click reactions on proteins containing two unnatural amino acids, AHA and PEP. The judicious placement of this pair of amino acids within a protein followed by a click reaction can lead to significantly

Table 1. Binding of stapled and unstapled protein G to IgG

Sample	Protein G Conc. [nM]	K_d [M]	k_a [1/M*s]	k_d [1/s]
		Local analysis		
Click	600	1.59×10^{-7}	$2.47\pm0.09\times10^{4}$	$3.93\pm0.07\times10^{-3}$
No click	600	7.71×10^{-7}	$2.45\pm0.16\times10^{4}$	$18.9\pm0.30\times10^{-3}$
Click	800	2.00×10^{-7}	$2.29\pm0.08\times10^{4}$	$4.59\pm0.10\times10^{-3}$
No click	800	10.4×10^{-7}	$1.86\pm0.11\times10^{4}$	$19.3\pm0.30\times10^{-3}$
Click	1000	3.26×10^{-7}	$2.02\pm0.08\times10^{4}$	$6.59\pm0.14\times10^{-3}$
No click	1000	7.65×10^{-7}	$2.35\pm0.10\times10^{4}$	$18.0\pm0.32\times10^{-3}$
		Global analysis		
Click		1.67×10^{-7}	$2.64\pm0.08\times10^{4}$	$4.42\pm0.10\times10^{-3}$
No click		5.98×10^{-7}	$2.85\pm0.15\times10^{4}$	$17.0\pm0.37\times10^{-3}$

improved protein properties. The insertion of a single click staple into the artificial leucine zipper A1 led to it retaining helical character even at 90 °C while a single click staple in the core of protein G led to a 5-fold improvement in its equilibrium binding to IgG. In contrast to previous protein stapling reports that used larger flexible amino acids to generate the staple,[6, 7] the linkage between AHA and PEP is even more rigid than nature's most common method for introducing covalent constraints, the disulphide bond. We suggest that this rigidity is likely the reason why such large gains in protein properties were observed in the studies presented here. Given that structural plasticity of proteins can be a major variable leading to their aggregation,[20] the protein stapling methodology presented here may be useful in stabilizing therapeutic proteins and endowing them with improved shelf-life.

This research was supported by the National Science Foundation (Award CBET-0952875)

Notes and references

[a]*Department of Chemical and Biological Engineering, Princeton University, Princeton, NJ 08544*
[b]*Department of Molecular Biology, Princeton University, Princeton, NJ 08544*
** To whom correspondence should be addressed: ajlink@princeton.edu*
† Electronic Supplementary Information (ESI) available: experimental methods and supporting figures. See DOI: 10.1039/b000000x/

1. D. J. Craik and U. Malik, *Curr. Opin. Chem. Biol.*, 2013, **17**, 546-554.
2. P. G. Arnison, M. J. Bibb, G. Bierbaum, A. A. Bowers, T. S. Bugni, G. Bulaj, J. A. Camarero, D. J. Campopiano, G. L. Challis, J. Clardy, P. D. Cotter, D. J. Craik, M. Dawson, E. Dittmann, S. Donadio, P. C. Dorrestein, K.-D. Entian, M. A. Fischbach, J. S. Garavelli, U. Goransson, C. W. Gruber, D. H. Haft, T. K. Hemscheidt, C. Hertweck, C. Hill, A. R. Horswill, M. Jaspars, W. L. Kelly, J. P. Klinman, O. P. Kuipers, A. J. Link, W. Liu, M. A. Marahiel, D. A. Mitchell, G. N. Moll, B. S. Moore, R. Muller, S. K. Nair, I. F. Nes, G. E. Norris, B. M. Olivera, H. Onaka, M. L. Patchett, J. Piel, M. J. T. Reaney, S. Rebuffat, R. P. Ross, H.-G. Sahl, E. W. Schmidt, M. E. Selsted, K. Severinov, B. Shen, K. Sivonen, L. Smith, T. Stein, R. D. Sussmuth, J. R. Tagg, G.-L. Tang, A. W. Truman, J. C. Vederas, C. T. Walsh, J. D. Walton, S. C. Wenzel, J. M. Willey and W. A. van der Donk, *Nat. Prod. Rep.*, 2013, **30**, 108-160.
3. H. E. Blackwell and R. H. Grubbs, *Angew. Chem.-Int. Edit.*, 1998, **37**, 3281-3284.
4. L. D. Walensky, A. L. Kung, I. Escher, T. J. Malia, S. Barbuto, R. D. Wright, G. Wagner, G. L. Verdine and S. J. Korsmeyer, *Science*, 2004, **305**, 1466-1470.
5. R. N. Chapman, G. Dimartino and P. S. Arora, *J. Am. Chem. Soc.*, 2004, **126**, 12252-12253.
6. H. Neumann, K. H. Wang, L. Davis, M. Garcia-Alai and J. W. Chin, *Nature*, 2010, **464**, 441-444.
7. X.-H. Chen, Z. Xiang, Y. S. Hu, V. K. Lacey, H. Cang and L. Wang, *ACS Chemical Biology*, 2014.
8. A. J. Link and D. A. Tirrell, *Methods*, 2005, **36**, 291-298.
9. N. Budisa, *Angew. Chem.-Int. Edit.*, 2004, **43**, 6426-6463.
10. V. V. Rostovtsev, L. G. Green, V. V. Fokin and K. B. Sharpless, *Angew. Chem.-Int. Edit.*, 2002, **41**, 2596-2599.
11. C. W. Tornoe, C. Christensen and M. Meldal, *J. Org. Chem.*, 2002, **67**, 3057-3064.
12. K. L. Kiick, E. Saxon, D. A. Tirrell and C. R. Bertozzi, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 19-24.
13. K. Kirshenbaum, I. S. Carrico and D. A. Tirrell, *Chembiochem*, 2002, **3**, 235-237.
14. W. A. Petka, J. L. Harden, K. P. McGrath, D. Wirtz and D. A. Tirrell, *Science*, 1998, **281**, 389-392.
15. L. Y. Lian, J. P. Derrick, M. J. Sutcliffe, J. C. Yang and G. C. K. Roberts, *J. Mol. Biol.*, 1992, **228**, 1219-1234.
16. A. Wang, N. W. Nairn, R. S. Johnson, D. A. Tirrell and K. Grabstein, *Chembiochem*, 2008, **9**, 324-330.
17. T. R. Chan, R. Hilgraf, K. B. Sharpless and V. V. Fokin, *Organic Letters*, 2004, **6**, 2853-2855.
18. A. E. Speers, G. C. Adam and B. F. Cravatt, *J. Am. Chem. Soc.*, 2003, **125**, 4686-4687.
19. V. Hong, S. I. Presolski, C. Ma and M. G. Finn, *Angew. Chem.-Int. Edit.*, 2009, **48**, 9879-9883.
20. E. Sahin, J. L. Jordan, M. L. Spatara, A. Naranjo, J. A. Costanzo, W. F. Weiss, A. S. Robinson, E. J. Fernandez and C. J. Roberts, *Biochemistry*, 2011, **50**, 628-639.