Site-specific N-terminal labelling of proteins *in vitro* and *in vivo* using *N*-myristoyl transferase and bioorthogonal ligation chemistry†

William P. Heal,[a] Sasala R. Wickramasinghe,[a] Paul W. Bowyer,‡[a] Anthony A. Holder,[b] Deborah F. Smith,[c] Robin J. Leatherbarrow*[a] and Edward W. Tate*[a]

Received (in Cambridge, UK) 18th October 2007, Accepted 5th November 2007
First published as an Advance Article on the web 13th November 2007
DOI: 10.1039/b716115h

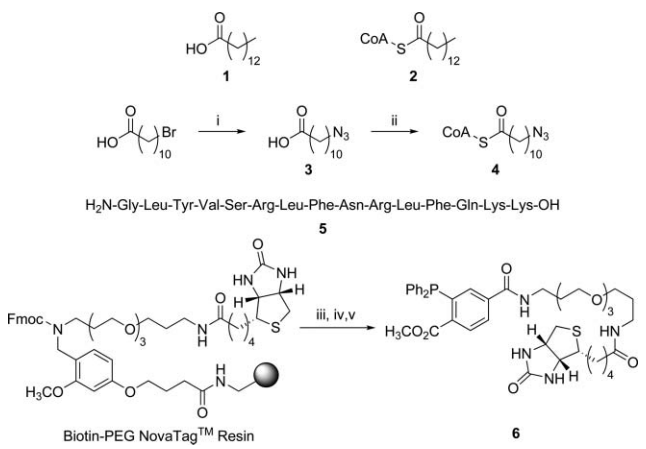
***N*-Myristoyl transferase-mediated modification with azide-bearing substrates is introduced as a highly selective and practical method for *in vitro* and *in vivo* N-terminal labelling of a recombinant protein using bioorthogonal ligation chemistry.**

Site-specific labelling of proteins with chemical and fluorescent probes is a prerequisite for many established and emerging techniques in chemical and cell biology.[1] Protein labelling technology has been revolutionised by the discovery of highly selective, water-compatible reactions[2] that permit bioorthogonal ligation of a wide variety of synthetic probes at one or more sites defined by a small chemical tag such as an azide, alkyne or aldehyde. Whilst bioorthogonal ligation chemistry has proven to be quite robust,[3] there remains a need for new methods that allow the introduction of tags into a protein in a site-specific manner. Recent innovations in this area, including expanded codon usage for incorporation of tagged amino acids[4] and site-directed post-translational modification,[1,5] can require relatively complex biochemical and/or synthetic protocols that can limit their widespread use for labelling applications. Here, we describe a practical transferase-based technique that permits site-specific *in vitro* and *in vivo* generation of N-terminal azide-tagged recombinant protein, and subsequent labelling by bioorthogonal ligation chemistry.

N-Myristoylation at an N-terminal glycine is a constitutive co-translational modification that occurs in all eukaryotic cells, and contributes to regulation of signalling and trafficking by modulating protein–membrane and protein–protein associations.[6] Myristoyl-CoA:protein *N*-myristoyl transferase (NMT), the enzyme that catalyses myristoylation, has been characterised in a wide range of organisms, from yeast to humans.[7] Comprehensive studies by Gordon *et al.* found that the specificity of *Saccharomyces cerevisiae* NMT (ScNMT) towards fatty acyl-CoA analogues is limited to compounds that closely mimic myristoyl-CoA, particularly with respect to chain length (~14 heteroatoms) and flexibility.[8] Nevertheless, early work suggested that ScNMT tolerates the introduction of some functionality in the acyl chain of the myristoyl-CoA substrate[8,9] and recent reports have described the use of modified myristate for non-radioactive labelling in mammalian cell lines.[20] Encouraged by this work, we examined the possibility of using azide-modified myristate analogues as a means for introducing a bioorthogonal ligation tag at the N-terminus of proteins bearing an N-terminal glycine. A well-characterised and widely-used NMT cloned from *Candida albicans* (CaNMT)[10] was selected for the present work: it exhibits excellent stability, activity and solubility, and may be over-expressed and purified without the need for an affinity tag.[11]

To test this approach we synthesised ω-azido undecanoyl-CoA **4** and a model peptide substrate **5**, corresponding to the N-terminal region of *Plasmodium falciparum* ADP ribosylation factor 1 (PfARF1),[11,12] carrying the canonical N-terminal myristoylation target motif Gly-Xaa-Xaa-Xaa-Ser (Scheme 1). Incubation of equimolar quantities of **4** and **5** with a catalytic quantity of CaNMT (0.25 mol%) in a suitable buffer for 18 h was sufficient for complete transfer of the ω-azido undecanoyl group from **4** to the target peptide as determined by HPLC (Fig. 1), and at a rate comparable to myristoyl-CoA **2** (data not shown).

Following these promising initial results, the tagging technique was applied to the labelling of a recombinant protein. An ω-azido



Scheme 1 Myristic acid (**1**), myristoyl-CoA (**2**) and the azido-tagged analogues of myristic acid (**3**) and CoA ester (**4**), PfARF1 peptide substrate (**5**) and capture reagent (**6**) synthesised and/or used in this study. *Reagents and conditions*: (i) NaN_3, DMSO; (ii) (a) CDI, THF, (b) CoASH, THF–$NaHCO_3$ (0.5 M); (iii) piperidine, DMF; (iv) 3-(diphenylphosphino)-4-(methoxycarbonyl)benzoic acid, HATU, DIPEA; (v) 95% $TFA_{(aq)}$. See ESI for full experimental details.

[a]*Department of Chemistry, South Kensington Campus, Imperial College, London, UK SW7 2AZ. E-mail: e.tate@ic.ac.uk; r.leatherbarrow@ic.ac.uk; Fax: +44 (0)20 75941139; Tel: +44 (0)20 75943752*
[b]*Division of Parasitology, MRC National Institute for Medical Research, Mill Hill, London, UK NW7 1AA*
[c]*Immunology and Infection Unit, Department of Biology, University of York, UK YO10 5YW*

† Electronic supplementary information (ESI) available: Representative experimental procedures and characterisation data for compounds. See DOI: 10.1039/b716115h

‡ *Present address*: Department of Pathology, Stanford University School of Medicine, 300 Pasteur Drive, Stanford, CA 94305-5324, USA.

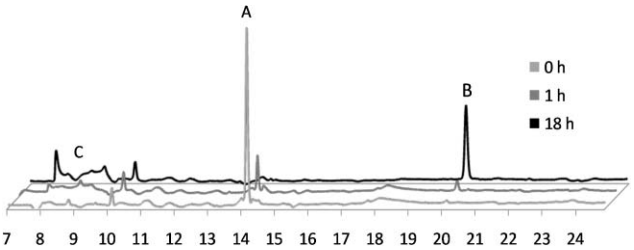


Fig. 1 HPLC traces showing transfer of azido-myristate **3** (*via* CoA ester **4**) to peptide **5** *in vitro* after reaction times of 0 h, 1 h and 18 h. (A) Peptide **5**; (B) azido-myristoylated **5**; (C) CoASH evolved during the reaction.

undecanoyl group was successfully transferred under similar conditions from **4** to full-length recombinant PfARF1[11,12] as shown by gel electrophoresis (Fig. 2). A capture reagent **6** bearing a biotin label, PEG linker and a phosphine capture group was conveniently synthesised using solid-phase chemistry on Biotin-PEG NovaTag™ resin (Scheme 1). Staudinger–Bertozzi ligation[3a] with **6** effectively transferred a biotin label to the tagged protein, as visualised with NeutrAvidin™–horseradish peroxidase (HRP) conjugate following gel electrophoresis and blotting to nitrocellulose membrane (Fig. 2B). Pull-down of biotinylated protein from the reaction mixture using Neutravidin™–agarose beads enabled the overall efficacy to be measured directly by virtue of a C-terminal His_6 affinity tag present in the PfARF1 construct; quantification by immunoblotting using anti-His–HRP conjugate demonstrated a transfer efficiency >99% (Fig. 2; lanes 3 *vs.* 5). There is no observable gel-shift of this protein upon myristoylation and capture, which is indicative of the minor influence of this N-terminal label on overall hydrophobicity.[5b]

By taking advantage of *E. coli* engineered to co-express CaNMT and a substrate protein,[13] this highly effective tagging technique may be performed in a single step *in vivo*, concurrent with protein over-expression. CaNMT and PfARF1 protein expression constructs bearing complementary resistance genes were co-transformed into *E. coli* BL21, and bacterial growth and subsequent expression of both proteins were induced under conditions of double antibiotic selection. Simultaneous feeding with **3** results in uptake, transformation into the active CoA analogue by endogenous *E. coli* acyl-CoA synthetase[8,9,13,14] and subsequent incorporation into PfARF1 by CaNMT. Bacteria were lysed, and the remarkable efficiency and specificity of this 'one-pot' *in vivo* process were demonstrated by labelling with reagent **6** and immunoblotting as described for *in vitro* experiments (Fig. 3).†§ Pull-down with NeutrAvidin™-agarose beads demonstrated a >99% efficiency of the overall process, relative to expressed PfARF1 (Fig. 3; lanes 1 *vs.* 2). The strict requirement for an N-terminal glycine,[10] and thus the presence of a single tag, was demonstrated by the absence of labelling in bacteria co-expressing CaNMT and PfARF1[G2A], a PfARF1 mutant bearing an N-terminal alanine (Fig. 3B, lanes 3 and 4).[11] It also appears that endogenous activity of *E. coli* methionine aminopeptidase (EcMetAP)[15] is sufficient for removal of the N-terminal initiator methionine in PfARF1 to reveal the requisite N-terminal glycine residue in quantitative yield. The utility of this *in vivo* protocol is further enhanced by the need for only the free acid: in contrast to acyl-CoA analogue **4**, ω-azido undecanoic acid **3** is readily synthesised in high yield in a single step from inexpensive 11-bromoundecanoic acid (Scheme 1), and may be used in feeding experiments without chromatographic purification.

Taken together, these results demonstrate that CaNMT is an effective tool for *in vitro* and *in vivo* transfer of an azide-modified acid to the N-terminus of a polypeptide derived from a species entirely unrelated to *C. albicans*. Previous studies have shown that the protein substrate specificity of NMT *in vitro* or in bacterial co-expression systems is dependent on the presence of a short N-terminal canonical myristoylation motif of general sequence GXXX[T|F|S],[16] whereas the *in vivo* specificity in eukaryotic cells is complicated by the presence of NMT isoforms and varying levels of expression and localisation of both enzyme and substrates.[10] In common with recombinant N-terminal labelling (*e.g.* His_6 tag), the addition of a myristoylation motif to proteins in which the N-terminal region is not involved in folding or other interactions does not usually impede its normal function,[17] and azido-myristoylation[18] with *N*-myristoyl transferase presents a potentially useful alternative to native chemical ligation[19] for N-terminal

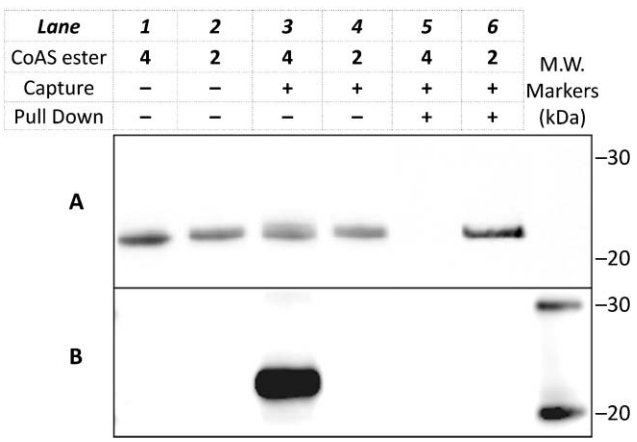


Fig. 2 *In vitro* transfer of **4** to PfARF1. **A**: Anti-His–HRP blot. **B**: NeutrAvidin™–HRP blot showing transfer of biotin after capture with reagent **6** (lane 3) and pull-down with NeutrAvidin™–agarose (lane 5). Control experiments using myristoyl-CoA **2** are shown for comparison.

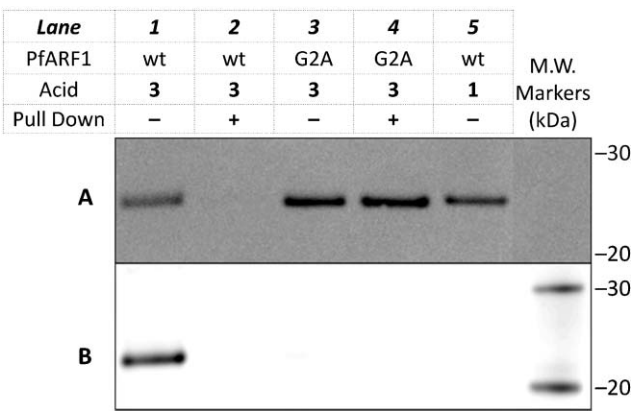


Fig. 3 *In vivo* transfer to PfARF1 and capture from cell lysates. **A**: Anti-His–HRP blot. **B**: NeutrAvidin™–HRP blot of pull-down showing transfer of biotin only to wild-type PfARF1 co-expressed with CaNMT in the presence of analogue **3**. wt = wild type PfARF1, G2A = PfARF1[G2A].

labelling of recombinant proteins under non-denaturing conditions.

The widespread adoption of a novel protein labelling technique is dependent on the simplicity and efficacy of the protocol.[5a] We show here that the requisite tagged substrates and labelling reagents are readily synthesised in a few high-yielding steps and that transfer can be performed concurrent with recombinant expression with very high efficiency, a combination that will allow rapid adoption for N-terminal labelling of proteins. Enzymatic tagging *via N*-myristoylation is also orthogonal to each of the alternative site-specific enzymatic tagging methods reported to date,[1,5] suggesting that the methods might be combined for sequential or simultaneous site-specific labelling with different probes. Furthermore, our work provides strong support for recent results that indicate azido-myristoylation may be used to tag proteins metabolically for proteomic applications.[20] Work is ongoing in our labs to exploit this approach for general protein labelling and targeted chemical proteomics in a variety of cell lines.

Notes and references

§ Lysates were pre-depleted of all endogenous biotinylated proteins by pull-down using NeutrAvidin™–agarose before treatment with **6** (see ESI†).

1 (*a*) T. L. Foley and M. D. Burkart, *Curr. Opin. Chem. Biol.*, 2007, **11**, 12–19; (*b*) K. M. Marks and G. P. Nolan, *Nat. Methods*, 2006, **3**, 591–596; (*c*) L. P. Tan and S. Q. Yao, *Protein Pept. Lett.*, 2005, **12**, 769–775; (*d*) E. W. Tate, *Signal Transduction*, 2006, **6**, 144–159.

2 J. E. Moses and A. D. Moorhouse, *Chem. Soc. Rev.*, 2007, **36**, 1249–1262.

3 (*a*) N. J. Agard, J. M. Baskin, J. A. Prescher, A. Lo and C. R. Bertozzi, *ACS Chem. Biol.*, 2006, **1**, 644–648; (*b*) J. A. Prescher, D. H. Dube and C. R. Bertozzi, *Nature*, 2004, **430**, 873–877.

4 J. Xie and P. G. Schultz, *Nat. Rev. Mol. Cell Biol.*, 2006, **7**, 775–782.

5 (*a*) I. S. Carrico, B. L. Carlson and C. R. Bertozzi, *Nat. Chem. Biol.*, 2007, **3**, 321–322; (*b*) U. T. Nguyen, J. Cramer, J. Gomis, R. Reents, M. Gutierrez-Rodriguez, R. S. Goody, K. Alexandrov and H. Waldmann, *ChemBioChem*, 2007, **8**, 408–423.

6 T. A. Farazi, G. Waksman and J. I. Gordon, *J. Biol. Chem.*, 2001, **276**, 39501–39504.

7 (*a*) D. K. Giang and B. F. Cravatt, *J. Biol. Chem.*, 1998, **273**, 6595–6598; (*b*) D. A. Towler, S. P. Adams, S. R. Eubanks, D. S. Towery, E. Jackson-Machelski, L. Glaser and J. I. Gordon, *Proc. Natl. Acad. Sci. U. S. A.*, 1987, **84**, 2708–2712; (*c*) R. S. Gunaratne, M. Sajid, I. T. Ling, R. Tripathi, J. A. Pachebat and A. A. Holder, *Biochem. J.*, 2000, **348 Pt 2**, 459–463; (*d*) H. P. Price, M. R. Menon, C. Panethymitaki, D. Goulding, P. G. McKean and D. F. Smith, *J. Biol. Chem.*, 2003, **278**, 7206–7214; (*e*) C. Panethymitaki, P. W. Bowyer, H. P. Price, R. J. Leatherbarrow, K. A. Brown and D. F. Smith, *Biochem. J.*, 2006, **396**, 277–285.

8 (*a*) N. S. Kishore, T. B. Lu, L. J. Knoll, A. Katoh, D. A. Rudnick, P. P. Mehta, B. Devadas, M. Huhn, J. L. Atwood, S. P. Adams, G. W. Gokel and J. I. Gordon, *J. Biol. Chem.*, 1991, **266**, 8835–8855; (*b*) R. O. Heuckeroth, E. Jackson-Machelski, S. P. Adams, N. S. Kishore, M. Huhn, A. Katoh, T. Lu, G. W. Gokel and J. I. Gordon, *J. Lipid Res.*, 1990, **31**, 1121–1129.

9 R. O. Heuckeroth, D. A. Towler, S. P. Adams, L. Glaser and J. I. Gordon, *J. Biol. Chem.*, 1988, **263**, 2127–2133.

10 R. C. Wiegand, C. Carr, J. C. Minnerly, A. M. Pauley, C. P. Carron, C. A. Langner, R. J. Duronio and J. I. Gordon, *J. Biol. Chem.*, 1992, **267**, 8591–8598.

11 (*a*) P. W. Bowyer, R. S. Gunaratne, M. Grainger, C. Withers-Martinez, S. R. Wickramsinghe, E. W. Tate, R. J. Leatherbarrow, K. A. Brown, A. A. Holder and D. F. Smith, *Biochem. J.*, 2007, DOI: 10.1042/BJ20070692; (*b*) W. H. L. Stafford, R. W. Stockley, S. B. Ludbrook and A. A. Holder, *Eur. J. Biochem.*, 1996, **242**, 104–113.

12 (*a*) R. S. Gunaratne, M. Sajid, I. T. Ling, R. Tripathi, J. A. Pachebat and A. A. Holder, *Biochem. J.*, 2000, **348 Pt 2**, 459–463; (*b*) R. R. Rees-Channer, S. R. Martin, J. L. Green, P. W. Bowyer, M. Grainger, J. E. Molloy and A. A. Holder, *Mol. Biochem. Parasitol.*, 2006, **149**, 113–116; (*c*) E. W. Tate, P. W. Bowyer, K. A. Brown, D. F. Smith, A. A. Holder and R. J. Leatherbarrow, *Signal Transduction*, 2006, **6**, 160–166.

13 (*a*) R. J. Duronio, E. Jackson-Machelski, R. O. Heuckeroth, P. O. Olins, C. S. Devine, W. Yonemoto, L. W. Slice, S. S. Taylor and J. I. Gordon, *Proc. Natl. Acad. Sci. U. S. A.*, 1990, **87**, 1506–1510; (*b*) L. J. Knoll, D. R. Johnson, M. L. Bryant and J. I. Gordon, *Methods Enzymol.*, 1995, **250**, 405–435.

14 (*a*) R. S. Bhatnagar and J. I. Gordon, *Methods Enzymol.*, 1995, **250**, 467–486; (*b*) J. H. Yoo, O. H. Cheng and G. E. Gerber, *Biochem. J.*, 2001, **360**, 699–706.

15 H. A. Van Valkenburgh and R. A. Kahn, *Methods Enzymol.*, 2002, **344**, 186–193.

16 (*a*) M. D. Resh, *Biochim. Biophys. Acta*, 1999, **1451**, 1–16; (*b*) J. A. Boutin, *Cell. Signalling*, 1997, **9**, 15–35; (*c*) D. R. Johnson, R. S. Bhatnagar, L. J. Knoll and J. I. Gordon, *Annu. Rev. Biochem.*, 1994, **63**, 869–914; (*d*) D. A. Towler, S. R. Eubanks, D. S. Towery, S. P. Adams and L. Glaser, *J. Biol. Chem.*, 1987, **262**, 1030–1036.

17 K. J. French, Y. Zhuang, R. S. Schrecengost, J. E. Copper, Z. Xia and C. D. Smith, *J. Pharmacol. Exp. Ther.*, 2004, **309**, 340–347.

18 Preliminary studies in our lab suggest that ω-alkynyl myristic acid analogues are equally applicable as NMT substrates for protein labelling applications; this work will be reported elsewhere.

19 (*a*) I. E. Gentle, D. P. De Souza and M. Baca, *Bioconjugate Chem.*, 2004, **15**, 658–663; (*b*) R. David, M. P. Richter and A. G. Beck-Sickinger, *FEBS J.*, 2004, **271**, 663–677.

20 (*a*) H. C. Hang, E. J. Geutjes, G. Grotenbreg, A. M. Pollington, M. J. Bijlmakers and H. L. Ploegh, *J. Am. Chem. Soc.*, 2007, **129**, 2744–2745; (*b*) D. D. O. Martin, G. L. Vilas, J. A. Prescher, G. Rajaiah, J. R. Falck, C. R. Bertozzi and L. G. Berthiaume, *FASEB J.*, 2007, DOI: 10.1096/fj.07-9198com.

