ChemComm

View Online
Dynamic Article Links

Cite this: *Chem. Commun.*, 2011, **47**, 1698–1700

www.rsc.org/chemcomm

COMMUNICATION

Palladium-mediated site-selective Suzuki–Miyaura protein modification at genetically encoded aryl halides†

Christopher D. Spicer and Benjamin G. Davis*

Received 15th November 2010, Accepted 8th December 2010
DOI: 10.1039/c0cc04970k

Site-specific genetic incorporation of unnatural *p*-halophenylalanine amino acid residues as 'tags' coupled with Pd(0)-mediated Suzuki–Miyuara 'modification' has been enabled by discovery of an effective small molecule palladium scavenger.

The modification of proteins is a key process in living cells, greatly increasing diversity in protein functionality and structure. Yet our ability to model and mimic such modifications is limited by the relatively small number of bio-orthogonal and site-selective synthetic methodologies that are available.[1,2]

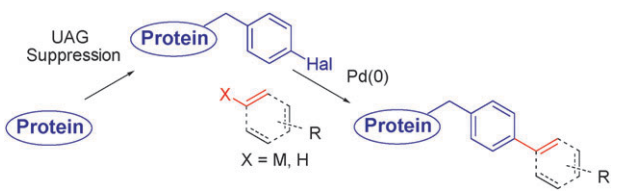
Traditionally, protein modification has been carried out at nucleophilic cysteine[3] or lysine residues.[4] However, in recent years a number of reactions have been developed, that allow site-selective modification at unconventional residues. Amongst the most attractive of these modifications are transition metal-catalysed reactions.[5] The reactive groups associated with such reactions are not commonly found in proteins and, through manipulation of metal and ligands, high conversions under mild conditions and with high functional group tolerance can be achieved.

We have previously demonstrated the utility of transition metals in protein modification. For instance, Ru-catalysis in protein olefin metathesis was achieved at privileged allyl chalcogenide containing amino acid residues.[6–8] In addition, a convenient, water-soluble palladium catalyst was recently developed for Suzuki–Miyaura cross-couplings at a cysteine-linked aryl halide. This catalyst's use in small molecule synthesis was also demonstrated.[9] Although related cross-couplings have previously been reported on proteins containing genetically incorporated coupling partners, they have suffered from low conversions (2% for Heck and 25% for Sonogashira couplings),[10,11] or the requirement of high temperatures.[12] Our initial report on Suzuki–Miyaura coupling was the first instance where cross-coupling had proceeded to full conversion on a protein substrate. To demonstrate the suitability of cross-coupling on proteins, the aryl halide was installed chemically. For more general application of such couplings it is desirable, however, to employ genetically encoded coupling partners. In doing so, this highly selective carbon–carbon bond formation could have potential *in vivo* applications. However, efficient cross-couplings at genetically encoded aryl halides remain an unsolved problem despite being first envisaged over a decade ago.[13]

A number of developments in the field of synthetic biology have allowed the genetic incorporation of such haloaryl amino acids through use of auxotrophic strains, and the expansion of the genetic code through amber stop codon suppression, and codon expansion.[4,14–17] In particular, the incorporation of *p*-iodophenylalanine (*p*IPhe) has been well documented in a range of proteins, with high fidelity and protein yields approaching those of the wild-type.[18–20]

We report here the incorporation of *p*IPhe into the model maltose binding protein (MBP) using UAG stop-codon suppression, and the use of this protein to explore Suzuki–Miyaura couplings on a protein surface as part of a general tag-and-modify strategy (Scheme 1).[21] We also demonstrate the use of a scavenger, 3-mercaptopropionic acid (3-MPrAc), to remove palladium from samples allowing ready and accurate monitoring and optimization of these protein reactions. We finally demonstrate the first Suzuki–Miyaura coupling at a genetically incorporated aryl halide, resulting in high conversion under ambient conditions.

In this study, His-tagged MBP was used as a model protein, due to its high stability and solubility.[22,23] MBP is a widely used protein, especially as a fusion construct. Importantly, both MBP and His_6 are sequences that are widely used in protein constructs. Using the pDB·His·MBP vector as a starting point, an amber stop codon was installed at sites corresponding to residue E13 by site-directed mutagenesis to generate the mutant plasmid pMBP·E13•. This site was chosen as the crystal structure of MBP indicated that this residue was



Scheme 1 Tag-and-modify strategy utilizing genetically incorporated *p*-halo-Phe as the 'tag' and Suzuki–Miyaura cross coupling as the 'modify' steps, respectively.

Chemistry Research Laboratory, University of Oxford, 12 Mansfield Road, Oxford, OX1 3TA, UK. E-mail: Ben.Davis@chem.ox.ac.uk; Fax: +44 (0)1865 275674; Tel: +44 (0)1865 275652

† Electronic supplementary information (ESI) available: Full experimental details including raw LC-MS data and protein expression and purification. See DOI: 10.1039/c0cc04970k

solvent accessible.[23] This plasmid was then co-transformed into *E. coli* with a pEVOL plasmid.[20] This second plasmid codes for a mutant ^{tyr}tRNA/tRNA synthetase pair that when expressed allows for suppression of the UAG amber codon at site E13• with a tRNA charged with *p*IPhe such that this 'stop' codon is instead translated as *p*IPhe.[20] The resulting protein product MBP–E13• (**1**) was expressed in good yields (14 mg L^{-1}) and high fidelity (Fig. S1, ESI†; in the absence of *p*IPhe no protein expression was observed).

With this site-selectively tagged protein in hand we examined its reactivity with Pd under various conditions. However, for certain reactions dramatic losses of signal intensity were observed, such that precise reaction monitoring by mass spectrometry was rendered difficult. To determine the potential sources of this signal loss, either loss of protein during reaction (through degradation or precipitation) or simply signal suppression (through reduced ionization or the nonspecific co-ordination of multiple metal cations to the protein surface generating many adduct isoforms, each at small quantities that are undetectable) we first investigated the product of reactions with Pd. SDS-PAGE showed no significant loss of protein at 500 equiv. of palladium (see ESI†). Similar observations were made with other proteins equine myoglobin (Fig. 1a–d) and the model protein Np276 (Fig. 1e), from *Nostoc punctiforme*. In the latter, the direct observation of small number adduct isoforms (corresponding to mono- and di-co-ordination) at low Pd concentrations strongly supported the second (metal suppression) hypothesis.

This problem of non-specific metal binding to proteins has been previously discussed by Francis and Antos.[5] It is important to note that this is a problem that is not specific to palladium catalysis, but is potentially relevant to all work utilising transition metals, or other Lewis acidic species, and may hinder future work in metal-catalyzed protein modification. We therefore looked to identify scavengers that might co-ordinate the palladium more strongly than the protein, and therefore remove it from the protein surface. In addition to allowing reaction analysis, such scavengers could in future be used for reaction purification *via* dialysis or size-exclusion chromatography. This scavenging therefore makes such reactions applicable to therapeutic proteins. Scavenging experiments were initially undertaken on the readily available protein equine myoglobin which is commonly used as a reference in LC-MS analysis.

EDTA, a common hexadentate chelating ligand, was the first potential scavenger to be identified.[24] DTT was also chosen, as it contains two nucleophilic thiol residues and is also compatible with proteins as shown by its use as a reducing agent. Additionally, cysteine was seen as a potential scavenger and has been used previously to greatly reduce palladium contaminants in small molecule synthesis.[25] However, none of these ligands showed any positive effects on the protein MS trace over a range of concentrations. Amongst the other ligands considered, 3-mercaptopropionic acid (3-MPrAc) has previously been used conjugated to magnetic nanoparticles to co-ordinate palladium.[26] We speculated that this material could form a 6-membered chelate to a metal centre. Gratifyingly, it was found that when this ligand was used as a scavenger, a strong re-emergence of the protein ion series was observed (Fig. 1a–d). A scavenger concentration of 4.4 μmol mL^{-1} (3 equiv. w.r.t. Pd), added as a dilute solution (5 μL mL^{-1}) in water to ensure accurate titration, was found to be optimal for

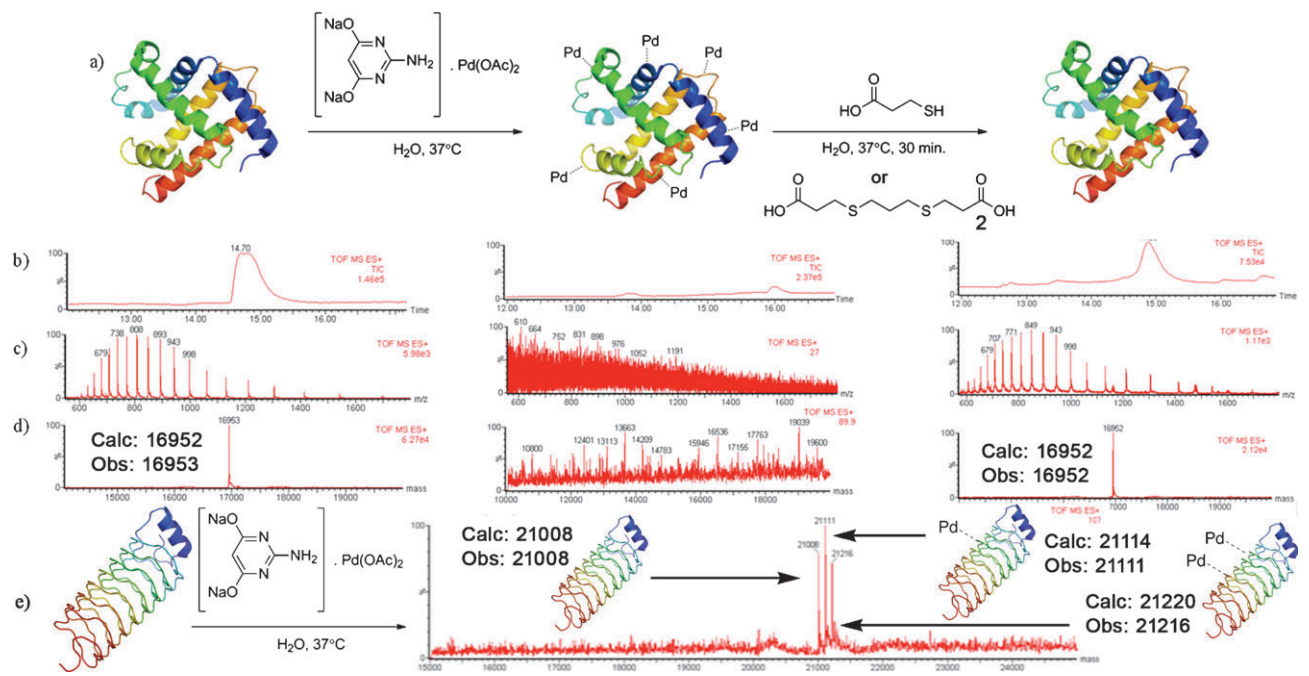
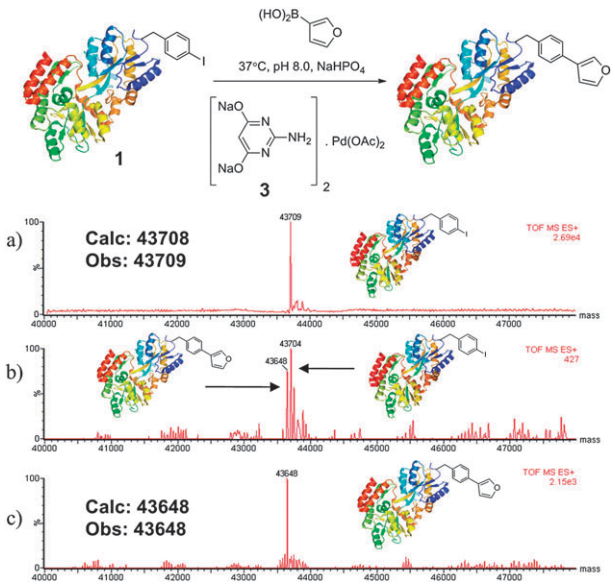


Fig. 1 (a) Addition of 500 equiv. palladium catalyst **3** to equine myoglobin, followed by addition of the scavenger 3-MPrAc, (b) total ion chromatogram of pure protein (left), protein in the presence of 500 equiv. Pd (middle), and after titration with 3-MPrAc (right), (c) corresponding ion series at time of protein elution, (d) corresponding deconvoluted mass chromatogram, (e) deconvoluted mass spectrum of Np276 in the presence of Pd showing discrete mono- and di-coordination of Pd.



Scheme 2 Coupling of furan-3-boronic acid to MBP–E13•. (a) $t = 0$, (b) $t = 1$ h, (c) $t = 2$ h.

myoglobin (1 mg mL^{-1}) (see ESI†). A diad of 3-MPrAc with a propane linker, **2**, was also synthesised, and showed similar scavenging ability at double the number of equivalents.

Now in a position to more readily investigate Suzuki–Miyaura couplings directly on proteins, this system was applied to MBP. Initial experiments showed optimal conditions to be highly dependent on both palladium and boronic acid concentrations. We chose to use furan-3-boronic acid as our coupling partner, as we had previously observed a high coupling rate and efficiency in model reactions (Scheme 2).[9] Gratifyingly, at 50 equiv. of Pd, and 680 equiv. of boronic acid, 50% conversion was observed at 37 °C after 1 hour (Scheme 2b). After 2 hours, complete conversion to the cross-coupled product was observed (Scheme 2c). In the absence of either palladium or boronic acid no reaction was observed, consistent with the proposed coupling mechanism. In addition, in the absence of 3-MPrAc no discernable protein peak could be observed by LC-MS, indicating the critical use of the scavenger during reaction analysis (see ESI†).

In conclusion, we have developed a method for scavenging palladium from protein surfaces using the ligand 3-MPrAc. This scavenging was essential for LC-MS analysis of the Suzuki–Miyaura cross-coupling. We then used this technique to undertake and optimize the first reported full-conversion Suzuki–Miyaura coupling on a protein at a genetically incorporated amino acid. This reaction did not require elevated temperatures or mixed solvents, and proceeded to completion in 2 hours. This site-specific reaction is an addition to the 'tool-kit' for protein modification, and also allows the formation of structures previously difficult to produce by conventional methods. For example, the furano-phenylalanine created here on MBP has not been previously incorporated by any other means, and exhibits a unique spectral signature which may prove useful for protein tracking. The work presented here generalizes Suzuki–Miyaura couplings on proteins and makes it more applicable as a methodology for chemical post-translational modification. In particular, the use of boronic acids as coupling partners opens up the possibility of installing hydrophobic groups onto a protein surface, without the need for organic solvents or denaturing conditions.[5,9]

We are now working towards unlocking the potential of Suzuki–Miyaura couplings, and investigating their scope for modification of proteins. We are particularly interested in the installation of hydrophobic groups, and the formation of bi-aryls as a method for fluorogenic protein labelling. We thank Prof. P.G. Schultz for donation of pEVOL (*p*IPhe) plasmid, Justin M. Chalker, Drs Huiwang Ai, Mark Batchelor, John Porter and Rikki Alexander for helpful discussions, and UCB and BBSRC for funding. BGD is a Royal Society Wolfson Research Merit Award recipient and is supported by an EPSRC LSI Platform grant.

Notes and references

1 E. M. Sletten and C. R. Bertozzi, *Angew. Chem., Int. Ed.*, 2009, **48**, 6974.
2 C. P. R. Hackenberger and D. Schwarzer, *Angew. Chem., Int. Ed.*, 2008, **47**, 10030–10074.
3 J. M. Chalker, G. J. L. Bernardes, Y. A. Lin and B. G. Davis, *Chem.–Asian J.*, 2009, **4**, 630–640.
4 A. J. de Graaf, M. Kooijman, W. E. Hennink and E. Mastrobattista, *Bioconjugate Chem.*, 2009, **20**, 1281–1295.
5 J. M. Antos and M. B. Francis, *Curr. Opin. Chem. Biol.*, 2006, **10**, 253–262.
6 J. M. Chalker, Y. A. Lin, O. Boutureira and B. G. Davis, *Chem. Commun.*, 2009, 3714–3716.
7 Y. A. Lin, J. M. Chalker, N. Floyd, G. J. L. Bernardes and B. G. Davis, *J. Am. Chem. Soc.*, 2008, **130**, 9642–9643.
8 Y. A. Lin, J. M. Chalker and B. G. Davis, *J. Am. Chem. Soc.*, 2010, **132**, 16805–16811.
9 J. M. Chalker, C. S. C. Wood and B. G. Davis, *J. Am. Chem. Soc.*, 2009, **131**, 16346–16347.
10 K. Kodama, S. Fukuzawa, H. Nakayama, T. Kigawa, K. Sakamoto, T. Yabuki, N. Matsuda, M. Shirouzu, K. Takio, K. Tachibana and S. Yokoyama, *ChemBioChem*, 2006, **7**, 1577–1581.
11 K. Kodama, S. Fukuzawa, H. Nakayama, K. Sakamoto, T. Kigawa, T. Yabuki, N. Matsuda, M. Shirouzu, K. Takio, S. Yokoyama and K. Tachibana, *ChemBioChem*, 2007, **8**, 232–238.
12 E. Brustad, M. L. Bushey, J. W. Lee, D. Groff, W. Liu and P. G. Schultz, *Angew. Chem., Int. Ed.*, 2008, **47**, 8220–8223.
13 N. Sharma, R. Furter, P. Kast and D. A. Tirrell, *FEBS Lett.*, 2000, **467**, 37–40.
14 A. J. Baldwin, J. A. J. Arpino, W. R. Edwards, E. M. Tippmann and D. D. Jones, *Mol. BioSyst.*, 2009, **5**, 764–766.
15 T. L. Hendrickson, V. d. Crécy-Lagard and P. Schimmel, *Annu. Rev. Biochem.*, 2004, **73**, 147–176.
16 K. Kirshenbaum, I. S. Carrico and D. A. Tirrell, *ChemBioChem*, 2002, **3**, 235–237.
17 D. R. Liu, T. J. Magliery, M. Pastrnak and P. G. Schultz, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 10092–10097.
18 J. Guo, E. M. Charles III, H. S. Lee, D. Groff and P. G. Schultz, *Angew. Chem., Int. Ed.*, 2009, **48**, 9148–9151.
19 J. Xie, L. Wang, N. Wu, A. Brock, G. Spraggon and P. G. Schultz, *Nat. Biotechnol.*, 2004, **22**, 1297–1301.
20 T. S. Young, I. Ahmad, J. A. Yin and P. G. Schultz, *J. Mol. Biol.*, 2010, **395**, 361–374.
21 B. G. Davis, *Pure Appl. Chem.*, 2009, **81**, 285–298.
22 J. D. Fox, R. B. Kapust and D. S. Waugh, *Protein Sci.*, 2001, **10**, 622–630.
23 M. B. Zwick, L. L. C. Bonnycastle, K. A. Noren, S. Venturini, E. Leong, C. F. Barbas, C. J. Noren and J. K. Scott, *Anal. Biochem.*, 1998, **264**, 87–97.
24 Y. Urawa, M. Miyazawa, N. Ozeki and K. Ogura, *Org. Process Res. Dev.*, 2003, **7**, 191–195.
25 M. Prashad, Y. Liu and O. Repič, *Adv. Synth. Catal.*, 2003, **345**, 533–536.
26 L. M. Rossi, L. L. R. Vono, F. P. Silva, P. K. Kiyohara, E. L. Duarte and J. R. Matos, *Appl. Catal., A*, 2007, **330**, 139–144.

