

CrossMark  
click for updatesCite this: *Chem. Sci.*, 2015, 6, 3712

# Functional modulation and directed assembly of an enzyme through designed non-natural post-translation modification†

Andrew M. Hartley,<sup>‡a</sup> Athraa J. Zaki,<sup>‡b</sup> Adam R. McGarrity,<sup>‡a</sup> Cecile Robert-Ansart,<sup>a</sup> Andriy V. Moskalenko,<sup>b</sup> Gareth F. Jones,<sup>c</sup> Monica F. Craciun,<sup>c</sup> Saverio Russo,<sup>c</sup> Martin Elliott,<sup>b</sup> J. Emyr Macdonald<sup>b</sup> and D. Dafydd Jones<sup>\*a</sup>

Post-translational modification (PTM) modulates and supplements protein functionality. In nature this high precision event requires specific motifs and/or associated modification machinery. To overcome the inherent complexity that hinders PTM's wider use, we have utilized a non-native biocompatible Click chemistry approach to site-specifically modify TEM  $\beta$ -lactamase that adds new functionality. *In silico* modelling was used to design TEM  $\beta$ -lactamase variants with the non-natural amino acid *p*-azido-L-phenylalanine (azF) placed at functionally strategic positions permitting residue-specific modification with alkyne adducts by exploiting strain-promoted azide-alkyne cycloaddition. Three designs were implemented so that the modification would: (i) inhibit TEM activity (Y105azF); (ii) restore activity compromised by the initial mutation (P174azF); (iii) facilitate assembly on pristine graphene (W165azF). A dibenzylcyclooctyne (DBCO) with amine functionality was enough to modulate enzymatic activity. Modification of TEM<sup>W165azF</sup> with a DBCO-pyrene adduct had little effect on activity despite the modification site being close to a key catalytic residue but allowed directed assembly of the enzyme on graphene, potentially facilitating the construction of protein-gated carbon transistor systems.

Received 16th December 2014  
Accepted 31st March 2015

DOI: 10.1039/c4sc03900a

[www.rsc.org/chemicalscience](http://www.rsc.org/chemicalscience)

## Introduction

Post-translational modification (PTM) is central to biology by expanding and modulating the function of a large number of proteins.<sup>1</sup> Many of these events essentially permanently covalently modify a protein, ranging from attachment of small moieties (*e.g.* methylation,<sup>2</sup> cofactors<sup>3</sup>) to larger events such as proteolysis and glycosylation. Each of these factors can impact significantly on protein structure and function thus influencing and even enabling inherent protein activity. The presence of sequence and/or structural motifs in combination with of a wide variety of subsidiary machinery, mostly enzymes, is required to achieve exquisite specificity both in terms of the target protein and the spatial position in the modified target. The complexity of these systems can be a significant hindrance with respect to their transfer to new proteins where such PTMs are not inherent. Furthermore, PTM is largely restricted to a select chemical set preexisting in nature. Precise modification

with non-natural adducts may be a more appropriate and useful means to expand and modulate protein function,<sup>4,5</sup> including for use in non-biological contexts.

Covalent modification with non-natural adducts is traditionally achieved using chemistry inherent to the natural amino acid repertoire, mostly amine, carboxyl and thiol groups. The main problem is lack of specificity as such chemistry is normally distributed across the surface of a protein and is ubiquitous in the proteome. A more general and powerful approach is the introduction of new chemically reactive handles not present in the native 20 amino acid set through the use of an expanded genetic code<sup>6</sup> (see ref. 7–10 for recent reviews). The unique reactivity of a chosen non-natural amino acid (nAA) together with the ability to select both the target protein and the residue within the target means that specificity comparable to natural PTM events can be achieved.<sup>10,11</sup> The use of highly specialized caged nAAs has been very effective in controlling activity<sup>12–15</sup> but is restricted to certain protein types and chemistry, and can add significant bulk to the amino acid side chain during the production and folding of the nascent polypeptide and its subsequent folded form.

The incorporation of phenyl azide chemistry into proteins through the use of *p*-azidophenylalanine (azF)<sup>17</sup> is an attractive alternative. As well as being only one atom bigger than the natural amino acid tyrosine and having been successfully incorporated into a wide variety of proteins,<sup>7–9</sup> the phenyl azide

<sup>a</sup>School of Biosciences, Cardiff University, Cardiff, UK. E-mail: jonesdd@cardiff.ac.uk<sup>b</sup>School of Physics and Astronomy, Cardiff University, Cardiff, Wales, UK<sup>c</sup>Centre for Graphene Science, University of Exeter, Exeter, Devon, UK

† Electronic supplementary information (ESI) available: Detailed experimental methods, supplementary Fig. 1 to 11 and supplementary Tables 1 to 3. See DOI: 10.1039/c4sc03900a

‡ AMH, AJZ and ARM contributed equally to this work.



moiety opens up different routes to non-natural PTM (nnPTM): photochemical transformations and Click chemistry adduct addition.<sup>9</sup> The use of phenyl azide photochemistry to control protein activity has recently been demonstrated.<sup>18–22</sup> Azide–alkyne cycloaddition is fast becoming a useful approach for orthogonal biomolecule conjugation but has largely been used in a passive way, for example, to label proteins.<sup>7,23–25</sup> Given the versatility in terms of the array of adducts available coupled with the inherent bioorthogonality and biocompatibility,<sup>26</sup> it is surprising Click chemistry has not been used more extensively as a general direct modulator of protein activity. Additionally, useful adducts can be placed at strategic positions to expand and facilitate protein function in a manner akin to co-factors. This includes attachment of entities to facilitate interfacing and assembly with secondary systems or materials.<sup>27–31</sup> As in many natural biomolecular assemblies, a defined and optimal protein–material interface is critical for maximal communication between the individual elements. Interfacing proteins with carbon  $sp^2$  materials such as graphene is gaining significant interest as it forms the basis for constructing hybrid bio-transistors in which events at even the single protein molecule level can be used to gate conductance through graphene.<sup>27,32</sup>

Here, we show that the activity and assembly onto pristine graphene of the antibiotic resistance protein TEM  $\beta$ -lactamase<sup>33,34</sup> can be directly controlled by Click chemistry. By genetically encoding phenyl azide chemistry at designed positions in TEM and using Cu-free biocompatible strain promoted azide–alkyne cycloaddition (SPAAC; Fig. 1)<sup>35,36</sup> different adducts can be attached (1, 2 and 3; Fig. 1) that either reduce or restore activity on modification, and define assembly of TEM on graphene.

## Results and discussion

The three main adducts with a core dibenzylcyclooctyne (DBCO) reactive handle (Fig. 1) chosen have very distinct properties: 1 is an amine derivative of DBCO that has hydrogen donor and acceptor groups opening up the potential to form H-bonds between residues not normally close enough to each other in the protein structure; 2 is a large, planar and hydrophobic rhodamine dye (Texas Red) that has proved useful in labelling proteins for fluorescent imaging<sup>23</sup> but could also act as an effective bulk spatial and steric blocking element if required; 3 is a pyrene derivative that will aid protein interfacing with extended carbon  $sp^2$  materials (*e.g.* graphene) through  $\pi$ – $\pi$  stacking.<sup>29,37</sup> Both 1 and 2 are “off-the-shelf” products while 3 can be generated by a simple succinimidyl ester reaction. Thus, all 3 adducts can easily be accessed by the wider bioscience community without any synthetic chemistry knowledge. The azF dependent production of active enzyme is shown in ESI Fig. 1.†

### Modulating enzyme activity

The *in silico* design process was performed using a combination of ROSETTA<sup>38</sup> and molecular dynamics<sup>39</sup> as outlined in the Supporting Methods.† Y105, a partially surface exposed residue ( $\sim 110 \text{ \AA}^2 \approx 50\%$  relative surface exposure), was chosen as a site for negative modulation through nnPTM due to its location close to the catalytically important SDN loop and its role in forming a partial lid over substrate binding cleft (Fig. 1; ESI Fig. 2†). Y105 is thought to be especially important in defining the size of the substrate binding cleft (therefore substrate specificity)<sup>40</sup> and local dynamics of residues critical for activity.<sup>41</sup> *In silico* modelling predicted exchanging the hydroxyl group for an azide should not have a major effect on enzyme structure around the locality of residue 105 (Fig. 2a). In fact ampicillin hydrolysis was slightly enhanced (higher  $k_{\text{cat}}$ ) for TEM<sup>Y105azF</sup> but had a slightly lower affinity (higher  $K_{\text{M}}$ ) making overall catalytic efficiency similar to wt TEM (Fig. 3 and ESI Table 1†). TEM<sup>Y105azF</sup> was receptive to modification by SPAAC as indicated by the estimated labelling efficiency of  $\sim 80\%$  with 2.

As predicted, TEM<sup>Y105azF</sup> was largely inhibited on modification with either 1 or 2 (Fig. 3) despite neither having any significant inherent inhibitory effect on wild-type TEM (ESI Table 1†). Both catalysis and substrate binding were disrupted as evident by the decrease in  $k_{\text{cat}}$  and increase in  $K_{\text{M}}$  (ESI Table 1†) for ampicillin resulting in an overall decrease in catalytic efficiency of  $\sim 70\%$  and  $85\%$  when modified with 1 or 2 respectively compared to unmodified protein. This is similar to the most deactivating site-directed mutants of Y105.<sup>40</sup>

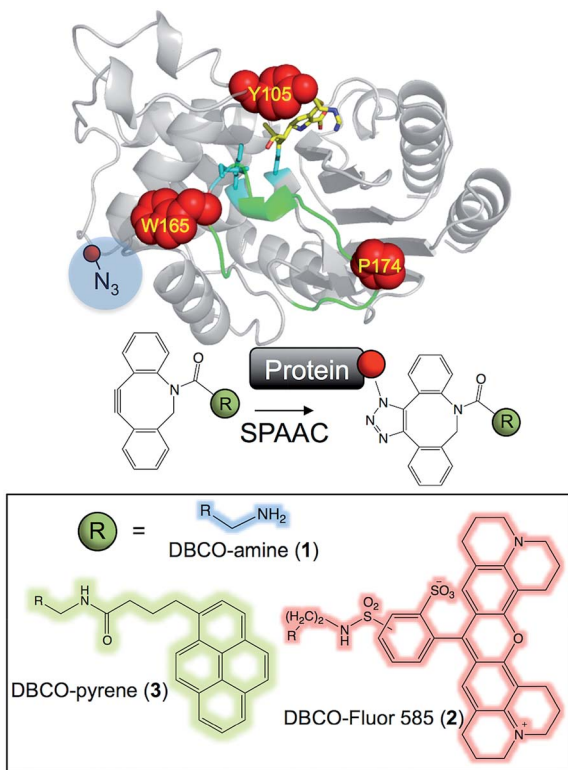


Fig. 1 Modification of TEM  $\beta$ -lactamase. The structure of TEM  $\beta$ -lactamase (top left) with the inhibitor imipenem (yellow sticks to emphasise the substrate binding pocket) highlights the residues targeted for replacement with azF (red spheres) together with the key catalytic residues (cyan sticks) and the  $\Omega$ -loop (green). The protein structures were generated using PyMol.<sup>16</sup> A representation of the SPAAC reaction using the activated alkyne present in the core DBCO moiety is shown (top right). The R groups attached to the DBCO used in this study are shown (bottom half).



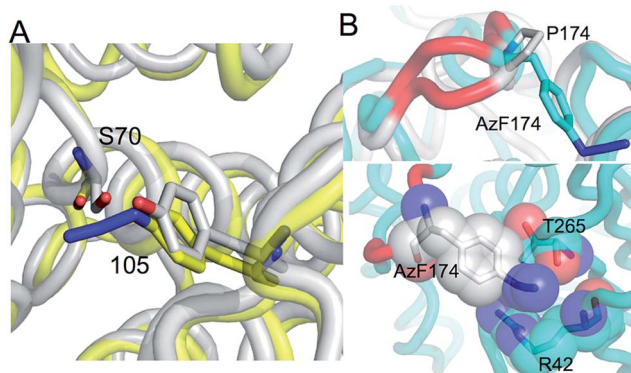


Fig. 2 Molecular design models of TEM azF variants. (A) Model of TEM<sup>Y105azF</sup> (yellow) aligned with wt TEM  $\beta$ -lactamase (grey; PDB 1btl), with the active site S70 shown as stick representation. (B) Alignment of the TEM<sup>P174azF</sup> (cyan) with wt TEM (grey; PDB 1btl). The region highlighted in red covers residues 173–175. (C) Interaction of AzF174 with neighbouring residues.

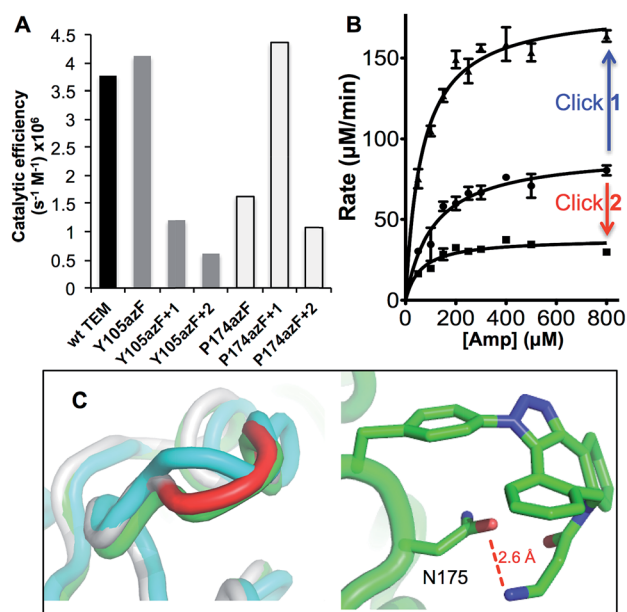


Fig. 3 Enzyme kinetics of ampicillin hydrolysis by TEM  $\beta$ -lactamase variants pre- and post-Click modification. (A) Catalytic efficiency of each variant pre- and post-Click modification with 1 (+1) or 2 (+2) derived from  $K_M$  and  $k_{cat}$  (ESI Table 1<sup>†</sup>). (B) The effect of modifying TEM<sup>P174azF</sup> with 1 (Click 1) or 2 (Click 2) on Michaelis–Menten kinetics. (C) Modelled changes to backbone structure of TEM<sup>P174azF</sup> (cyan) on modification with 1 (green) compared to wt TEM (grey) (top panel) and the configuration and local interactions (hydrogen bond in red) of 1 (bottom panel).

Modelling of the covalent complex (*vide infra*) between TEM<sup>Y105azF</sup> and 1 suggested that modification blocks access at one end of the substrate binding site, with the amine of 1 forming a hydrogen bond with the backbone carbonyl group of G238 (ESI Fig. 3<sup>†</sup>). The enzyme kinetics suggests that simple steric blocking of substrate binding is unlikely to be the sole mechanism of action as catalysis and substrate binding are equally affected, similar to that observed for classical mixed

inhibition models. The relatively small difference in catalytic efficiency of TEM<sup>Y105azF</sup> modified with either 1 or 2 indicates that the most significant effect is occurring close to the linkage site, as suggested by the covalent complex model (ESI Fig. 3<sup>†</sup>). However, the bulkier group of 2 is exerting a slightly greater effect with regards to  $k_{cat}$  rather than  $K_M$  (Fig. 3 and ESI Table 1<sup>†</sup>) suggesting that long range interactions made by the rhodamine dye moiety may be forcing the structure around residue 105 to adapt less catalytically proficient form. Currently, the separation of modified from unmodified protein has proved unsuccessful, so the observed activity may be contributed by unmodified protein only (estimated to be  $\sim 20\%$ ; *vide supra*). Therefore, the inhibition levels observed here represent the lowest currently achievable and we cannot rule out modified species are essentially inactive.

P174 is another largely surface exposed residue ( $\sim 80 \text{ \AA}^2 \approx 60\%$  relative surface accessibility) that is relatively distant from the active site (Fig. 1). P174 however contributes towards one of the turns that comprise the  $\Omega$  loop (Fig. 1 and ESI Fig. 4<sup>†</sup>), a region critical for substrate binding, substrate specificity and catalysis.<sup>42,43</sup> It was hypothesized in the design process that mutating P174 to a non-cyclic amino acid would result in local conformational changes around the turn region. Molecular modelling supported this hypothesis with the structure of the turn slightly shifted compared to wt TEM (Fig. 2b and ESI Fig. 5<sup>†</sup>). The side chain of azF174 lies approximately perpendicular to the native proline, with the azido moiety fitting into a shallow pocket and making a local polar interaction network with R42 and T265 (Fig. 2c). These new interactions may act as the driver of local conformational changes associated with the P174azF mutation, which in turn alters the local bonding network that subtly shifts the conformation of both the active site and substrate binding residues, including S70 (ESI Fig. 5<sup>†</sup>). Based on the TEM<sup>P174azF</sup> model it was further postulated that on modification with 1, the formation of the triazole link would break the interaction network with R42 and T265 so reconstituting the original loop structure and activating the protein. The amine group also has the potential to form new long range hydrogen bonds in spatially local regions (Fig. 3 and ESI Fig. 4<sup>†</sup>).

As suggested by the model, replacement of P174 with azF results in a significant change in activity. The overall catalytic efficiency of TEM<sup>P174azF</sup> was 60% lower compared to wt TEM (Fig. 3). The major effect was on  $k_{cat}$  which was 3 fold lower compared to wt TEM (ESI Table 1<sup>†</sup>), supporting the evidence from the model that the likely effect of azF incorporation at residue 174 is to disrupt interactions of residues associated with catalysis. This was offset by a slightly increased ampicillin affinity.

TEM<sup>P174azF</sup> was accessible to Click modification, but efficiency was lower (estimated  $\sim 33\%$  using the absorbance properties of 2) compared to TEM<sup>Y105azF</sup>. It is unknown how local protein microenvironment dictates SPAAC efficiency but dynamics, relative exposure to aqueous solvent and the character of shallow ‘‘pockets’’ have been suggested as possible determinants.<sup>23</sup> The molecular model suggests that interaction of the azide moiety with other residues, which may play a role



through, for example, altering the relative populations of resonance structures sampled by the azide moiety ( $R-N=N^+=N^-/R-N^--N^+\equiv N$ ), and/or its accessibility and available orientations to the incoming DBCO. The effect of the SPAAC nnPTM on the ability of TEM<sup>P174azF</sup> to hydrolyse ampicillin varied depending on the DBCO adduct. Modification with **2** exerted a similar but somewhat smaller inhibitory effect ( $\sim 40\%$  drop) as observed for TEM<sup>T105azF</sup> (Fig. 3 and ESI Table 1†). Modification with **1** resulted in a significant increase in overall activity restoring apparent catalytic efficiency essentially to wild-type levels (Fig. 3). This should be considered as a lower estimate of activation as  $\sim 70\%$  of the protein may be unmodified (*vide supra*). With both adducts, the most significant contribution was the change in  $k_{cat}$  of TEM<sup>P174azF</sup> compared to unmodified protein (ESI Table 1†); modification with **1** increased  $k_{cat}$  by 210% while addition of **2** reduced  $k_{cat}$  by almost half.

To understand how **1** exerts its beneficial effect on TEM<sup>P174azF</sup>, the nnPTM product was modelled. Using the TEM<sup>P174azF</sup> model as a starting point, the triazole linkage between azF and DBCO of **1** was generated *in silico* by parameterizing the azF-DBCO complex to calculate optimized geometries and the electrostatic potentials. The model was subjected to molecular dynamics for a total of 5 ns. The model of nnPTM product, termed TEM<sup>P174azF+1</sup> suggested that the general structure was closer to wt TEM than the original TEM<sup>P174azF</sup> model (Fig. 3c and ESI Fig. 5†). The  $\Omega$ -loop structure around residue 174 and the local bonding networks for TEM<sup>P174azF+1</sup> were largely comparable with wt TEM. The orientation of residue 174 differs significantly to accommodate **1** (ESI Fig. 5†). The amine group of **1** bends back and appears to come within hydrogen bonding distance ( $\sim 2.6$  Å) of the side-chain carbonyl group of the adjacent residue N175 (Fig. 3c). The loss of the azide group on formation of the triazole removes the interactions with R43 and T238. However, the reason for **1** acting as an activator and **2** as an inhibitor may be down to the reduced bulk and ability of **1** to form local H-bonds.

### Protein-graphene interfacing

In some instances, it would be attractive to select a residue intimately associated with active site regions so that secondary non-related events can be coupled to activity without significant effect on function. Linking proteins to carbon  $sp^2$  materials such as graphene is especially attractive; local charge and electrostatic changes in the protein can be used to “gate” the electronic properties of the  $sp^2$  material<sup>27,44</sup> so generating sensing systems with ultimate single molecule resolution. The pyrene moiety of **3** (Fig. 1) is especially attractive as an interfacing agent as it allows defined coupling of a protein to the  $sp^2$  material through  $\pi$  stacking.<sup>37</sup> This has been demonstrated previously using optimally placed cysteine residues as the reactive handle for attachment to pyrene-coated single walled carbon nanotubes.<sup>27</sup> However, this required removal by mutagenesis of native cysteine residues. Here, we demonstrate that defined enzyme interfacing on a clean graphene face is feasible though the use of designed azF placement.

W165 is another partially surface exposed residue ( $\sim 100$  Å<sup>2</sup>  $\approx 50\%$  relative surface accessibility) that fulfils the required

criteria (Fig. 1). It resides in the  $\Omega$ -loop and splits two functionally important residues; E166 is a conserved catalytic residue involved in proton shuttling<sup>34</sup> and R164 forms buried ionic interactions critical for stabilizing the  $\Omega$  loop and thus the positioning of E166.<sup>45</sup> The approach used to model TEM<sup>P174azF</sup> modified with **1** was applied to the W165azF mutation to ascertain the potential position of the adduct group and its effect on structure (Fig. 4a). Modelling suggested that the overall structure of the protein would be largely unperturbed and the adduct would point outwards into the solvent. Mutating W165 to azF did slightly reduce overall catalytic efficiency towards ampicillin (1.5 fold) but crucially, modification with **3** was feasible and had little effect on overall activity (ESI Fig. 6†).

By coupling **3** to TEM<sup>W165azF</sup> prior to surface assembly we can directly interface the protein with graphene in a defined manner without surface coating with pyrene, as has been previously used.<sup>27,46</sup> Intermittent-contact mode AFM (tapping mode) imaging indicated that TEM<sup>W165azF</sup> modified with pyrene binds stably to graphene surfaces (Fig. 4b and c; ESI Fig. 7 & Movie 1†). The protein molecules are not significantly disturbed by multiple scans. The average apparent heights were 3 nm, which is close to the predicted height of the protein bound in the designed orientation ( $\sim 3$  nm; ESI Fig. 7†). The average apparent lateral dimension was larger than predicted (10 nm *versus* 5 nm), as a result of tip convolution effects. In comparison, wt TEM did not stably bind to the surface, with tip contamination commonly observed after multiple scans (ESI Fig. 8†).

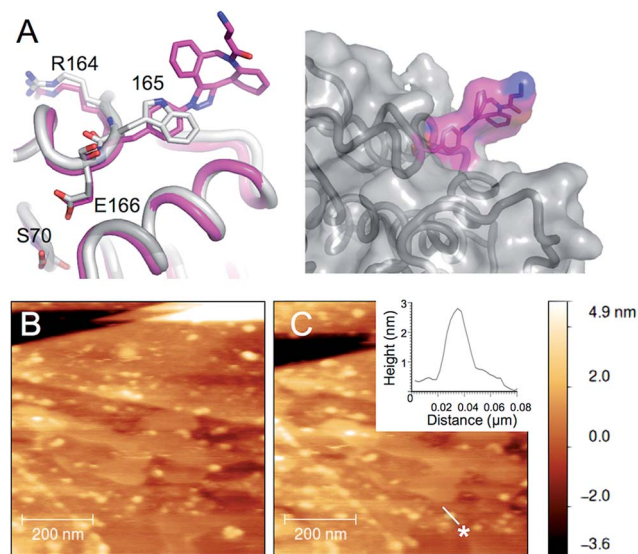


Fig. 4 AFM imaging of pyrene-modified TEM<sup>W165azF</sup> on graphene. (A) Model of TEM<sup>W165azF</sup> (magenta) modified with **1** aligned with the crystal structure of wt TEM (grey) with the residues either side of residue 165 and the catalytic S70 highlighted. The surface representation of TEM<sup>W165azF</sup> modified with **1** (left panel) illustrates that the Click adduct points out and away from the protein. Repeated AFM imaging of the surface after the (B) sixth and (C) tenth scans shows that the proteins bind stably to the surface. The height analysis of a single TEM<sup>W165azF-3</sup> molecule, which is selected in C (highlighted as \*) is shown as an inset to C. ESI Movie 1† provides a full trace for the imaged area above over 10 scans.



## Conclusions

SPAAC is a powerful approach for precise and designed protein PTM that supplement protein activity using a non-intrusive, non-caged reaction handle. We have demonstrated here that SPAAC can be used beyond simple passive labelling of proteins and can actively modulate activity through different mechanisms without the need to synthesize specialized molecules. It also facilitates the interfacing of proteins in a designed and defined manner with useful active materials such as graphene. This negates the need to remove inherent chemistry through mutagenesis allowing the focus to be on placing the reactive handle for optimal coupling between the protein and material. *In silico* modelling with both nAA and adducts can greatly aid the design process by which useful variants are generated and their affects rationalised, which will in turn lead to more accurate designs.

## Experimental

### *In silico* protein modelling

A detailed description of the *in silico* modelling of TEM containing the azF mutation and modification with the DBCO-amine adduct (**1**) are provided in the ESI.† Briefly, geometry optimised structure files, force field parameters for azF and azF-DBCO amine adduct complexes were generated. The published crystal structure of TEM  $\beta$ -lactamase to 1.8 Å resolution (PDB code 1BTL<sup>45</sup>) was used as a starting point for Monte Carlo simulations within the ROSETTA software package.<sup>38</sup> The lowest energy model was used as the starting point for molecular dynamics using GROMACS.<sup>39</sup> The modelling was run on the Raven cluster as part of the Advanced Research Computing @ Cardiff facility.

### Enzyme activity assays

The wt and mutant TEM variants were generated as outlined in the ESI.† The proteins were expressed in *E. coli* using a bespoke plasmid based on pBAD called pBADKAN, and purified as described in the ESI.† The TEM-dependent kinetics of ampicillin hydrolysis ( $\epsilon_{235} = 1500 \text{ M}^{-1} \text{ cm}^{-1}$ ) were determined spectrophotometrically using 1 cm path length QS quartz cuvette (Hellma). Hydrolysis assays were carried out in a 1 mL reaction volume. Purified enzyme was diluted to a final concentration of 250 ng  $\mu\text{L}^{-1}$  in 50 mM sodium phosphate buffer, pH 8 at room temperature. Reactions were started by addition of ampicillin and hydrolysis was measured by the decrease in absorbance at 235 nm. Ampicillin concentrations ranged from 50  $\mu\text{M}$  to 800  $\mu\text{M}$ . Kinetic parameters were calculated using initial rate of hydrolysis at each substrate concentration and then fitting to the Michaelis–Menten equation using GraphPad Prism. TEM  $\beta$ -lactamase activity using the colorimetric substrate nitrocefin was performed essentially as described previously.<sup>47,48</sup>

### Click modification

SPAAC reactions were performed on pure protein using a five-fold molar excess of dibenzylcyclooctyne (DBCO) reagent to protein. DBCO-amine (**1**) and DBCO-Fluor 585 (**2**) were obtained from Click chemistry tools, while DBCO-pyrene was synthesized from DBCO-amine and 1-pyrenebutanoic acid, succinimidyl ester *via* nucleophilic substitution. Reactions were left overnight at room temperature in PBS. SPAAC reactions with **2** were analysed by SDS-PAGE and subsequent imaging of the fluorescent dye on a transilluminator. Labelling efficiencies were calculated as described in the ESI.†

### Protein deposition on graphene and AFM imaging

Protein samples were deposited onto graphene on a copper foil substrate. Monolayer graphene was deposited on the copper substrate by chemical vapor deposition (CVD) using methane as carbon source. A detailed description of the growth procedure is provided in the ESI.† The foil pieces ( $\sim 4 \text{ mm}^2$ ) were immersed in PBS containing 1 nM protein and incubated at room temperature for 10 min. Following incubation, the foil pieces were rinsed with high purity deionised water and dried with  $\text{N}_2$  gas. AFM imaging was performed in using a Veeco Nanoscope IIa (Bruker) in tapping mode. The same area was scanned up to 10 times with a total of 6 different regions imaged to check the stability of the protein molecules on the surface of graphene. A video of the surface scans is provided as ESI.† Although the images show some drift, it is clear that the proteins are stably bound to the surface.

## Acknowledgements

DDJ and ARM would like to thank the Advanced Research Computing @ Cardiff facility, especially Thomas Green for help with access and usage of the Raven cluster. DDJ, JEM and ME would like to thank the BBSRC (BB/H003746/1 and BB/M000249/1), EPSRC (EP/J015318/1) and Cardiff SynBio Initiative/SynBioCite for supporting this work. SR and MFC acknowledge financial support from EPSRC (EP/J000396/1, EP/K017160/1, EP/K010050/1, EP/G036101/1, EP/M001024/1, EPM002438/1). AMH was supported by a BBSRC studentship, AJZ by a Ministry of Higher Education in Kurdistan Region and Kurdistan Regional Government studentship and ARM by a Cardiff School of Biosciences personal studentship.

## Notes and references

- 1 C. Walsh, *Posttranslational modification of proteins: expanding nature's inventory*, Roberts and Co. Publishers, Englewood, Colo., 2006.
- 2 S. I. Grewal and J. C. Rice, *Curr. Opin. Cell Biol.*, 2004, **16**, 230–238.
- 3 R. N. Perham, *Annu. Rev. Biochem.*, 2000, **69**, 961–1004.
- 4 K. Lang and J. W. Chin, *ACS Chem. Biol.*, 2014, **9**, 16–20.
- 5 C. D. Spicer and B. G. Davis, *Nat. Commun.*, 2014, **5**, 4740.
- 6 L. Wang, A. Brock, B. Herberich and P. Schultz, *Science*, 2001, **292**, 498–500.



- 7 K. Lang and J. W. Chin, *Chem. Rev.*, 2014, **114**, 4764–4806.
- 8 C. C. Liu and P. G. Schultz, *Annu. Rev. Biochem.*, 2010, **79**, 413–444.
- 9 S. Reddington, P. Watson, P. Rizkallah, E. Tippmann and D. D. Jones, *Biochem. Soc. Trans.*, 2013, **41**, 1177–1182.
- 10 W. H. Zhang, G. Otting and C. J. Jackson, *Curr. Opin. Struct. Biol.*, 2013, **23**, 581–587.
- 11 C. H. Kim, J. Y. Axup and P. G. Schultz, *Curr. Opin. Chem. Biol.*, 2013, **17**, 412–419.
- 12 A. Gautier, A. Deiters and J. W. Chin, *J. Am. Chem. Soc.*, 2011, **133**, 2124–2127.
- 13 A. Gautier, D. P. Nguyen, H. Lusic, W. An, A. Deiters and J. W. Chin, *J. Am. Chem. Soc.*, 2010, **132**, 4086–4088.
- 14 J. Li, S. Jia and P. R. Chen, *Nat. Chem. Biol.*, 2014, **10**, 1003–1005.
- 15 N. Wu, A. Deiters, T. A. Cropp, D. King and P. G. Schultz, *J. Am. Chem. Soc.*, 2004, **126**, 14306–14307.
- 16 The PyMOL Molecular Graphics System, Schrödinger, LLC.
- 17 J. Chin, S. Santoro, A. Martin, D. King, L. Wang and P. Schultz, *J. Am. Chem. Soc.*, 2002, **124**, 9026–9027.
- 18 J. L. Morris, S. C. Reddington, D. M. Murphy, D. D. Jones, J. A. Platts and E. M. Tippmann, *Org. Lett.*, 2013, **15**, 728–731.
- 19 S. C. Reddington, P. J. Rizkallah, P. D. Watson, R. Pearson, E. M. Tippmann and D. D. Jones, *Angew. Chem.*, 2013, **52**, 5974–5977.
- 20 S. C. Reddington, A. J. Baldwin, R. Thompson, A. Brancale, E. M. Tippmann and D. D. Jones, *Chem. Sci.*, 2015, **6**, 1159–1166.
- 21 S. Zhu, M. Riou, C. A. Yao, S. Carvalho, P. C. Rodriguez, O. Bensaude, P. Paoletti and S. Ye, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 6081–6086.
- 22 A. Grunbeck, T. Huber, P. Sachdev and T. P. Sakmar, *Biochemistry*, 2011, **50**, 3411–3413.
- 23 S. C. Reddington, E. M. Tippmann and D. D. Jones, *Chem. Commun.*, 2012, **48**, 8419–8421.
- 24 J. Y. Axup, K. M. Bajjuri, M. Ritland, B. M. Hutchins, C. H. Kim, S. A. Kazane, R. Halder, J. S. Forsyth, A. F. Santidrian, K. Stafin, Y. Lu, H. Tran, A. J. Seller, S. L. Biroc, A. Szydluk, J. K. Pinkstaff, F. Tian, S. C. Sinha, B. Felding-Habermann, V. V. Smider and P. G. Schultz, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 16101–16106.
- 25 C. S. McKay and M. G. Finn, *Chem. Biol.*, 2014, **21**, 1075–1101.
- 26 M. D. Best, *Biochemistry*, 2009, **48**, 6571–6584.
- 27 Y. Choi, I. S. Moody, P. C. Sims, S. R. Hunt, B. L. Corso, I. Perez, G. A. Weiss and P. G. Collins, *Science*, 2012, **335**, 319–324.
- 28 J. Han and C. Gao, *Nano-Micro Lett.*, 2010, **2**, 213–226.
- 29 J. A. Mann, T. Alava, H. G. Craighead and W. R. Dichtel, *Angew. Chem.*, 2013, **52**, 3177–3180.
- 30 E. A. Della Pia, M. Elliott, D. D. Jones and J. E. Macdonald, *ACS Nano*, 2012, **6**, 355–361.
- 31 E. A. Della Pia, J. E. Macdonald, M. Elliott and D. D. Jones, *Small*, 2012, **8**, 2341–2344.
- 32 Q. H. Wang, Z. Jin, K. K. Kim, A. J. Hilmer, G. L. Paulus, C. J. Shih, M. H. Ham, J. D. Sanchez-Yamagishi, K. Watanabe, T. Taniguchi, J. Kong, P. Jarillo-Herrero and M. S. Strano, *Nat. Chem.*, 2012, **4**, 724–732.
- 33 J. Petrosino, C. Cantu 3rd and T. Palzkill, *Trends Microbiol.*, 1998, **6**, 323–327.
- 34 A. Matagne, J. Lamotte-Brasseur and J. M. Frere, *Biochem. J.*, 1998, **330**(2), 581–598.
- 35 J. C. Jewett, E. M. Sletten and C. R. Bertozzi, *J. Am. Chem. Soc.*, 2010, **132**, 3688–3690.
- 36 J. M. Baskin, J. A. Prescher, S. T. Laughlin, N. J. Agard, P. V. Chang, I. A. Miller, A. Lo, J. A. Codelli and C. R. Bertozzi, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 16793–16797.
- 37 R. J. Chen, Y. Zhang, D. Wang and H. Dai, *J. Am. Chem. Soc.*, 2001, **123**, 3838–3839.
- 38 K. T. Simons, R. Bonneau, I. Ruczinski and D. Baker, *Proteins*, 1999, (3), 171–176.
- 39 S. Pronk, S. Pall, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess and E. Lindahl, *Bioinformatics*, 2013, **29**, 845–854.
- 40 N. Doucet, P. Y. De Wals and J. N. Pelletier, *J. Biol. Chem.*, 2004, **279**, 46295–46303.
- 41 N. Doucet, P. Y. Savard, J. N. Pelletier and S. M. Gagne, *J. Biol. Chem.*, 2007, **282**, 21448–21459.
- 42 W. Huang, J. Petrosino, M. Hirsch, P. S. Shenkin and T. Palzkill, *J. Mol. Biol.*, 1996, **258**, 688–703.
- 43 J. F. Petrosino and T. Palzkill, *J. Bacteriol.*, 1996, **178**, 1821–1828.
- 44 K. Besteman, J. O. Lee, F. G. M. Wiertz, H. A. Heering and C. Dekker, *Nano Lett.*, 2003, **3**, 727–730.
- 45 C. Jelsch, L. Mourey, J. M. Masson and J. P. Samama, *Proteins*, 1993, **16**, 364–383.
- 46 P. C. Sims, I. S. Moody, Y. Choi, C. Dong, M. Iftikhar, B. L. Corso, O. T. Gul, P. G. Collins and G. A. Weiss, *J. Am. Chem. Soc.*, 2013, **135**, 7861–7868.
- 47 D. D. Jones, *Nucleic Acids Res.*, 2005, **33**, e80.
- 48 A. M. Simm, A. J. Baldwin, K. Busse and D. D. Jones, *FEBS Lett.*, 2007, **581**, 3904–3908.

