Integrated catalysis opens new arylation pathways via regiodivergent enzymatic C–H activation

Jonathan Latham1,2, Jean-Marc Henry1,2, Humera H. Sharif1,2, Binuraj R.K. Menon1,2, Sarah A. Shepherd1,2, Michael F. Greaney1 & Jason Micklefield1,2

Despite major recent advances in C–H activation, discrimination between two similar, unactivated C–H positions is beyond the scope of current chemocatalytic methods. Here we demonstrate that integration of regioselective halogenase enzymes with Pd-catalysed cross-coupling chemistry, in one-pot reactions, successfully addresses this problem for the indole heterocycle. The resultant ‘chemobio-transformation’ delivers a range of functionally diverse arylated products that are impossible to access using separate enzymatic or chemocatalytic C–H activation, under mild, aqueous conditions. This use of different biocatalysts to select different C–H positions contrasts with the prevailing substrate-control approach to the area, and presents opportunities for new pathways in C–H activation chemistry. The issues of enzyme and transition metal compatibility are overcome through membrane compartmentalization, with the optimized process requiring no intermediate work-up or purification steps.
Selective activation of C–H positions has transformed synthesis in recent years. Using transition metal catalysis, a plethora of direct, fundamental bond formations are now achievable under mild, catalytic conditions, changing the way chemists make molecules. The field is currently defined by substrate control, with a relatively small number of C–H motifs undergoing predictable and reliable metatation by a transition metal catalyst; for example, electron-rich heteroarenes, electron-poor fluorarenes, activation ortho (and more recently, meta), to a directing group via cyclometalation all being dominant in poor fluoroarenes, activation ortho (and more recently, meta), to a directing group via cyclometalation all being dominant examples (Fig. 1a)2–9. Effective catalyst control, where choice of metal catalyst; for example, electron-rich heteroarenes, electron-undergoing predictable and reliable metalation by a transition is very rare and remains an outstanding challenge in the field10–13. Progress in this area will profoundly enhance the scope of C–H activation, as the majority of C–H bonds do not fit the current substrate-control criteria for chemocatalytic activation chemistry. We were interested in addressing this problem using a combination of bio- and chemo-catalysis. The complex coordination sphere of an enzyme active site enables exquisite selectivities that cannot be achieved by simple metal coordination complexes. Correspondingly, the manifold power of noble metal catalysis in C–C bond formation presents synthesis pathways that are not possible through biocatalysis. Their integration offers a complementarity that could create new C–H activation strategies, whereby C–H positions beyond the scope of chemocatalysis alone can be targeted for C–C, C–N, C–Hal and other fundamental bond-forming steps. In addition, the merging of bio- and chemo-catalysis into single process, one-pot operations creates a step-change in efficiency through enhanced space-time yields, better energy consumption and elimination of auxiliary chemicals14,15.

Despite the potential advantages of combining biocatalysts with chemocatalysts, their successful integration has been largely restricted to organic solvents15–17, which are incompatible with many enzymes and cofactors—therefore limiting the number of biocatalysts that can be utilized. In contrast, integration of chemocatalysts with enzymes in their preferred aqueous media is relatively rare14. Indeed, the inherent differences in operating conditions of enzymes and chemocatalysts present serious challenges, often requiring compartmentalization or immobilization of at least one of the catalysts18–23.

Our plans for a catalyst-controlled, integrated C–H activation system involve the merging of flavin-dependent halogenase (Fl-Hal) enzymes with palladium-catalysed cross-couplings. The pivotal role played by aryl halide functionality in metal-catalysed transformations suggested to us that a regioselective, biohalogenation would have great versatility, with the incipient Ar-X group being adaptable to a multitude of Pd-catalysed chemistries. We chose to integrate Fl-Hals and Suzuki–Miyaura coupling (SMC) with boronic acids in the first instance, given the critical role of SMC in producing biaryl products for pharmaceutical, agrochemical and materials science applications. The most widely studied Fl-Hals are tryptophan halogenases, which are responsible for chlorinating different positions of the indole ring of tryptophan en route to various bacterial natural...
products. Several fungal Fl-Hals have also been reported that chlorinate phenolic natural products. Together this suite of enzymes have the potential to regioselectively halogenate a range of aromatic scaffolds. With appropriate cofactor-recycling systems (Fig. 1b) Fl-Hals can afford halogenation reactions using benign inorganic halide (typically MgCl₂ or NaBr), oxygen (from air), and glucose or isopropanol as the only stoichiometric reagents. Studies on broadening the substrate scope of the tryptophan halogenases RebH, PrnA and PyrH, as well as improving their productivity and scaleability, have meant that Fl-Hals are becoming increasingly viable as biocatalysts for aromatic halogenation. In addition, Fl-Hals have been used to halogenate natural products and synthetic substrates, which have been isolated or extracted and then derivatized in separate cross-coupling reactions under standard synthetic operating procedures.

Biohalogenation is also attractive, given that traditional aromatic halogenation chemistry utilizes deleterious reagents, lacks regiocontrol and can afford mixtures, including halogenated by-products that can be toxic and persistent in the environment.

We anticipated that the unique regiocontrol inherent to Fl-Hal enzymatic halogenation could set-up an overall catalyst-controlled C–H arylation system, if a working SMC could be successfully integrated under aqueous conditions. The indole heterocycle has been central to the development of contemporary C–H activation chemistry, driving the discovery of new chemocatalytic methods for functionalizing the C2 and C3 positions (the ‘innate’ positions for functionalization with an electrophile). Direct C–C bond formation at the arene C6–H positions (the ‘innate’ positions for functionalization with an electrophile) has been challenging, with the exception of the C6–H position in anthranilamide. Here, PdII is known to react with glucose under certain conditions, and the use of PdII catalysts has been shown to produce the desired halogenation products in good yields. In this study, we show that PdII catalysts are also effective at functionalizing the C6–H position in anthranilamide and its derivatives, allowing for the synthesis of a range of C6-halogenated indoles in good yields.

Results

Biocatalytic halogenation. To effect one-pot Fl-Hal-SMC reactions (Fig. 1c), we required enzymes and substrates that afford high yields of structurally diverse brominated products. First, we exploited our previous findings that the tryptophan 5-halogenase PyrH can accept anthranilamide (1) to develop conditions for effective enzymatic synthesis of 5-bromoanthranilamide (2) (Fig. 2). Given the exquisite natural regio-complementarity of the flavin-dependent tryptophan halogenases we then set out to identify potential indole substrates for biocatalyst-controlled bromination and subsequent arylation. We found that tryptophol (3) could be efficiently brominated by RebH (a tryptophan 7-halogenase) to afford 7-bromotryptophol (4), after optimizing the conditions (Supplementary Fig. 1) that had been previously reported for chlorination of this substrate. Initial attempts to brominate the 5-position of tryptophol (3) using PyrH in place of RebH gave very low conversion. The larger substrate 3-indolepropionic acid (5), however, was transformed by PyrH to give 5-bromo-3-indolepropionate (6) as a single regiosomer in good yield. In addition, the 6-bromo-3-indolepropionate (7) could be obtained in good regioisomeric excess and yield by using the tryptophan 6-halogenase SttH. Collectively, these three halogenase enzymes established a clean, regiodivergent synthesis of C5, C6 and C7 bromoindoles.

Finally, to extend the substrate diversity further we produced the radicicol halogenase RadH enzyme from the fungus Chaetomium chiversii. RadH is similar in sequence to Rdc2—a related fungal halogenase that had been shown previously to chlorinate 6-hydroxy-isoquinoline (8) and other related substrates. Accordingly conditions were developed whereby RadH could be used to prepare 5-bromo-6-hydroxyisoquinoline (9) in good yields.

Cross-coupling optimization. We next sought conditions whereby the SMC could be integrated with the biocatalysts, using the coupling of 5-bromoanthranilamide (2) with phenyl boronic acid as a test reaction (Fig. 3). On the basis of previous studies of SMCs in aqueous media with an open atmosphere, a number of pyrimidine and guanidine ligands were selected for screening. Of these, L2 afforded highest conversion to 5-phenylanthranilamide (10) at 2.0 mM substrate concentration, 2.5 mol% Pd loading, in potassium phosphate buffer in the absence of any cofactors or enzymes. In addition, a number of water-soluble phosphine ligands were also screened. Of these, tppts and sulfonated SPhos gave full conversion to the desired product. These two ligands were therefore carried forward in a base and palladium source screen (Supplementary Fig. 2). The reaction was found to be tolerant of many bases, including DIPEA, whilst the lack of any base afforded no product. In addition, conducting the reaction using L5 and L6 under air as opposed to inert atmosphere afforded no conversion, whilst the same reaction using L2 under open air afforded good conversion.

The tolerance of the selected Pd catalysts towards the biocatalytic halogenation conditions was then probed. As NAD⁺/NADH and FAD/FADH₂ are rich in heteroatoms and redox active they could possibly interfere with Pd-catalysis either through redox chemistry or coordination to metal species. In addition, it has been previously reported that soft donor atoms (for example, sulfur) on the surface of proteins can coordinate to transition metals potentially inhibiting their catalytic activity. Pd₄ is also known to react with glucose under certain conditions (utilized to recycle NADH). The test reaction (Fig. 3) was therefore run in the presence of each of these components to
determine their effect on SMC conversion. This revealed a somewhat deleterious effect from the presence of the cofactors and a major reduction in cross-coupling yields in the presence of proteins (Supplementary Fig. 3). Increasing the loading of palladium catalyst was generally advantageous in terms of conversion in the presence of NADH, FAD and glucose, but did not enhance cross-coupling when either the reductase or dehydrogenase proteins were present (Supplementary Fig. 4).

Halogenase SMC integration. Given that the presence of proteins caused the greatest reduction in cross-coupling yields, we reasoned that the enzymes would need to be compartmentalized or removed before cross-coupling, at least in the first instance. Pleasingly, we found that passing the PyrH biotransformation of anthranilamide (1) through a 10-kDa molecular-weight cut-off (MWCO) cellulose membrane to remove protein before cross-coupling was effective in this regard, affording very good conversion to arylated product (10). Following further optimization around this process (Supplementary Fig. 5), we could achieve an overall isolated yield of 79% of arylated product (10) using tppts ligand L6, Pd(OAc)2 (50 mol%), and an excess of phenyl boronic acid (Fig. 4). Moreover, we showed that the enzymes that had been removed using the membrane filter could be recycled affording intermediate bromide (2) in 62% and 42% in the second and third cycles, respectively. Subsequent arylation of the combined biotransformation filtrates afforded arylated product (10) in 54% overall yield; a ca. twofold increase in the amount of material that would be obtained from the same amount of enzyme in a single arylation reaction. After determining the scope of this method with respect to boronic acids (Supplementary Fig. 6) and finding the SMC compatible with the conditions required for RadH-catalysed bromination of 8 (Supplementary Fig. 7), we then extended this membrane filtration method (method A) to the selective 5-arylation of 6-hydroxyisoquinoline (8), using RadH to produce 15–17 in good yields.

With a workable integrated arylation process in hand, we looked at ways to improve our method, principally with a view to increasing scale and reducing the high Pd content in the SMC reaction. The requirement for high catalyst loadings appeared to arise from a combination of dilute reaction volumes and inhibitory effect of cofactors present in the reaction medium. Indeed, other attempts to conduct palladium-catalysed chemistry in biological media and in the presence of proteins have required very high loading or suffered from poor yields47,48. Enhancing enzyme efficiency would mitigate these two issues, by affording higher concentrations of aryl bromide for cross-coupling. The group of Sewald have recently reported improvements in RebH productivity through the use of crosslinked enzyme aggregates (CLEAs)37. The individual subunits of RebH are crosslinked with each other, as well as the partner reductase and dehydrogenase proteins, via surface lysines using glutaraldehyde forming a RebH-reductase-dehydrogenase CLEA that is more efficient and tractable to handle37. Accordingly, we successfully prepared these

| Ligand | Temperature (°C) | Conversion |
|--------|-----------------|------------|
| L1     | 80              | 32         |
| L2     | 80              | 86<sup>a</sup> |
| L3     | 80              | 56         |
| L4     | 80              | 32         |
| L5     | 50              | 100        |
| L6     | 50              | 100<sup>b</sup> |
| L7     | 50              | n.c.       |
| L8     | 50              | n.c.       |

Figure 3 | Ligand screening for the SMC of 2 with phenyl boronic acid. Pd(OAc)₂ was used as the palladium source in each case in a 2:1 ligand:Pd ratio. Conversions determined by analytical high-performance liquid chromatography (HPLC) or liquid chromatography mass spectrometry (LC-MS). n.c., no conversion. <sup>a</sup>81% isolated yield; <sup>b</sup>80% isolated yield. Further optimization results are available in the Supplementary Material (Supplementary Fig. 2).
materials as active heterogeneous biocatalysts containing either PyrH, SttH or RebH. Although active CLEAs of the phenolic halogenase RadH were also obtained, they did not prove to be significantly more efficient than purified enzyme.

Initial screening of the PyrH-CLEA with anthranilamide (I) demonstrated improved efficiency, affording up to 3.0 mM aryl bromide. This enabled one-pot halogenase-SMC reactions to proceed to reasonable conversion (ca. 50%) without the need to remove or compartmentalize the CLEA from the Pd catalyst and other reagents, presumably as many heteroatoms of the protein had been effectively protected during the crosslinking process (Supplementary Fig. 8). We also found that the arylation of tryptophol (3) could be carried out in the same one-pot manner, with the RebH-CLEA and the Pd catalyst both present throughout (Supplementary Fig. 8). However, highest conversions were obtained when the CLEA was removed, by centrifugation or filtration, before cross-coupling of the supernatant from biocatalytic halogenation (Supplementary Figs 8 and 9). The scope of this method (method B) was found to be broad with respect to boronic acids (Supplementary Fig. 10), allowing the efficient arylation of anthranilamide (1) with both electron-rich, electron-poor and heterocyclic boronic acids using PyrH-CLEA with reduced palladium loading (10 mol%), affording the 5-arylated compounds in very good yield (10–14, Fig. 4).

Tryptophol (3) was also efficiently coupled using electron-rich, electron-poor, heterocyclic and alkenyl boronic acids to afford 7-substituted products 18 to 22 in good yield with RebH-CLEA at lower Pd loading. 3-Indole propionate (5) was likewise effectively arylated to biaryl products 23 and 24, with PyrH- and SttH-CLEAs, respectively. Along with the tryptophol examples

Figure 4 | Isolated yield of the regioselective and regiodivergent arylation. *Obtained using purified enzymes and a MWCO membrane (method A). Obtained using a CLEA of the appropriate Fl-Hal and 10 mol% Pd-catalyst loading (method B). Obtained from three successive cycles of halogenation using the same Fl-Hal. Biocat., biocatalysis; chemocat., chemocatalysis.
(18–22), this demonstrates the efficient and highly selective manipulation of each of the 5-, 6- and 7- C–H positions of the indole nucleus, under catalyst control, which is unattainable using existing synthetic methodology. In addition, the heterogeneous nature of the CLEA biocatalyst allowed it to be recycled over repeated cycles of bromination, with little loss in efficiency (Supplementary Fig. 11). In the case of the 6-arylation of 3-indolepropionic acid (5), this allowed an overall yield of 75% of 24 over three cycles with the SttH-CLEA. The increased stability and ease of preparation of the heterogeneous biocatalyst allowed this methodology to be scaled up to a 1.0-mmol scale to afford reasonable yields of 10 (>100 mg, 52%).

**Polydimethylsiloxane compartmentalization.** To further enhance the scope and efficiency of the halogenase-SMC process, we looked for alternative ways of separating the proteins from the palladium catalyst. While membrane filtration worked well, it did introduce an additional operation in the middle of the process and necessitated a deoxygenation step before SMC. To better streamline the process and develop a more efficient one-pot procedure, we examined polydimethylsiloxane (PDMS) thimbles as a compartmentalization strategy. This approach had been reported for a number of reactions that require the separation of incompatible reagents19,49, including the combination of a Pd-catalyzed oxidation with a biocatalytic reduction19.

Typically, one of the reactions is assembled inside a small thimble made of PDMS, which is then placed inside the larger bulk reaction vessel. As PDMS is a hydrophobic polymer, non-polar compounds (that is, substrates and intermediates in the cascade) should dissolve sufficiently well to pass through the PDMS walls, whilst charged reagents (Pd catalysts, enzymes and cofactors) would be unable to diffuse through PDMS and might therefore be compartmentalized. Initial tests showed that 5-bromoanthranilamide (2) would flux through PDMS well while glucose, NADH, FAD and the L2.Pd(OAc)2 complex were all unable to effectively penetrate PDMS (Supplementary Fig. 12). SMC conditions using L2.Pd(OAc)2 were therefore chosen to test the PDMS compartmentalization approach, as this catalyst functions well under open air and because alternative phosphine ligands have been shown to promote flux through PDMS.

Accordingly conditions were optimized, which allowed the PyrH-halogenation of anthranilamide (2) using pure protein to be carried out with the cross-coupling reagents within the PDMS thimble (Fig. 5), resulting in a ‘one-pot’ reaction to give 5-phenylanthranilamide (10) in 74% yield. The immediate advantages of this approach (method C) are that protein removal and deoxygenation of the reaction buffer are not required. Moreover a lower catalyst loading (10 mol%) is achievable under this regime, compared with the removal of pure protein using filtration (50 mol%), due to the compartmentalization of cofactors and enzymes from the Pd catalyst. By using a PyrH-CLEA to afford higher concentration of aryl bromide in conjunction with PDMS compartmentalization (method D), the Pd loading was further reduced to 2 mol% to afford 10 in good yield (59%). This combination also allowed the coupling of deactivated, electron-rich boronic acids such as 3-furyl boronic acid to give 25 in good yield (64%), a reaction that was unsuccessful using a cellulose membrane (MWCO filter). Finally, this method was also extended to include coupling of the 7-position of tryptophol (3), using RebH-CLEA, to give 7-pent-1-enyl tryptophol (18) and 4′-tertbutyl-7-phenyl tryptophol (26) in good yields.

**Discussion**

In summary, we have combined four regioselective flavin-dependent halogenase enzymes with a palladium-catalysed Suzuki–Miyaura cross-coupling to afford the regio-controlled arylation of a number of aromatic scaffolds (benzamides, isoquinolines and indoles). In each case, the regioselective C–H arylation transformations were previously inaccessible using stand-alone chemocatalysis or biocatalysis methods. Specifically, we report the selective C–H arylation of the 5-, 6- and 7- positions of the indole nucleus under catalyst control.

To overcome issues of catalyst incompatibility, we initially developed a filtration method using a 10-kDa MWCO filter (method A). This allowed efficient biocatalyst recycling for a number of cycles, affording approximately double the overall productivity that the same batch would in a single cycle. Biocatalyst efficiency was further improved by using CLEAs, which afforded higher aryl bromide concentration—enabling lower palladium loading to be employed (method B). We also
developed a PDMS thimble system to compartmentalize the chemo- and bio-catalysts. This method afforded similar yields of regioselectively arylated product as the MWCO filter system, but without the need for protein removal or solution deoxygenation before the cross-coupling chemistry (method C). Finally, the implementation of CLEAs with PDMS compartmentalization led to a further reduction in the required Pd loading down to 2 mol% (method D). Combined with directed evolution and other research aimed at improving the productivity and scalability of heteroarylation reactions involving heteroarenes containing one heteroatom. Adv. Synth. Catal. 356, 171–117 (2014).

1. Yu, J.-Q. & Shi, Z.-C. GH Activation Vol. 292 (Springer, 2010).

2. Côté, E. et al. C–F and C–H bond activation of fluorobenzones and fluoropyridines at transition metal centers: how fluorine tips the scales. Acc. Chem. Res. 44, 333–348 (2011).

3. Rossi, R., Bellina, F., Llesi, M. & Manzini, C. Cross-coupling of heteroarenes by C–H Functionalization: recent progress towards direct arylation and
32. Zhou, H., Qiao, K., Gao, Z., Vederas, J. C. & Tang, Y. Insights into radicicol biosynthesis via heterologous synthesis of intermediates and analogs. *J. Biol. Chem.* **285**, 41412–41421 (2010).
33. Shepherd, S. A. *et al.* Extending the biocatalytic scope of regiocomplementary flavin-dependent halogenase enzymes. *Chem. Sci.* **6**, 3454–3460 (2015).
34. Payne, J. T., Andorfer, M. C. & Lewis, J. C. Regioselective arene halogenation using the FAD-dependent halogenase RebH. *Angew. Chem. Int. Ed.* **52**, 5271–5274 (2013).
35. Hölzer, M., Burd, W., Reißig, H.-U. & Péé, K.-H. V. Substrate specificity and regioselectivity of tryptophan 7-halogenase from *Pseudomonas fluorescens* BL915. *Adv. Synth. Catal.* **343**, 591–595 (2001).
36. Frese, M., Guzowska, P. H., Voß, H. & Sewald, N. Regioselective enzymatic halogenation of substituted tryptophan derivatives using the FAD-dependent halogenase RebH. *ChemCatChem.* **6**, 1270–1276 (2014).
37. Frese, M. & Sewald, N. Enzymatic halogenation of tryptophan on a gram scale. *Angew. Chem. Int. Ed.* **54**, 298–301 (2015).
38. Glenn, W. S., Nims, E. & O’Connor, S. E. Reengineering a tryptophan halogenase to preferentially chlorinate a direct alkaloid precursor. *J. Am. Chem. Soc.* **133**, 19346–19349 (2011).
39. Roy, A. D., Gruschow, S., Cairns, N. & Goss, R. J. M. Gene expression enabling synthetic diversification of natural products: chemogenetic generation of pacidamycin analogs. *J. Am. Chem. Soc.* **132**, 12243–12245 (2010).
40. Durak, L. J., Payne, J. T. & Lewis, J. C. Late-stage diversification of biologically active molecules via chemoenzymatic C–H activation. *ACS Catal.* **6**, 1451–1454 (2016).
41. Sandtorp, A. H. Transition metal-catalyzed C–H activation of indoles. *Adv. Synth. Catal.* **357**, 2403–2435 (2015).
42. Itahara, T., Ikeda, M. & Sakakibara, T. Alkenylation of 1-acylindoles with olefins bearing electron-withdrawing substituents and palladium acetate. *J. Chem. Soc. Perkin. Trans. 1*, 1361–1363 (1983).
43. Grimmer, N. P., Gauntlett, C., Godfrey, C. R. A. & Gaunt, M. J. Palladium-catalyzed intermolecular alkenylation of indoles by solvent-controlled regioselective C–H functionalization. *Angew. Chem. Int. Ed.* **44**, 3125–3129 (2005).
44. Stuart, D. R. & Fagnou, K. The catalytic cross-coupling of unactivated arenes. *Science* **316**, 1172–1175 (2007).
45. Chalker, J. M., Wood, C. S. C. & Davis, B. G. A convenient catalyst for aqueous and protein Suzuki–Miyaura cross-coupling. *J. Am. Chem. Soc.* **131**, 16346–16347 (2009).
46. Gao, Z., Gouverneur, V. & Davis, R. G. Enhanced aqueous Suzuki–Miyaura coupling allows site-specific polypeptide 18F-labeling. *J. Am. Chem. Soc.* **135**, 13612–13615 (2013).
47. Poor, C. B., Andorfer, M. C. & Lewis, J. C. Improving the stability and catalyst lifetime of the halogenase RebH by directed evolution. *ChemBioChem.* **15**, 1286–1289 (2014).

**Acknowledgements**

We acknowledge BBSRC (grant BB/K00199X/1), EPSRC, CoEBio3 and GlaxoSmithKline for support. We also thank Dr Anna-Winona Struck for helpful conversations.

**Author contributions**

J.M. and M.F.G. conceived and supervised the project; J.L., J.-M.H. and H.H.S. performed integrated arylation reactions; B.R.K.M. and S.A.S. developed expression and initial screening conditions for the SMC; C.J.L. prepared and carried out integrated reactions with PDMS thimbles; J.L. and M.F.G. conceived and supervised the project; J.L., J.-M.H. and H.H.S. optimized the integrated process whilst J.L. and H.H.S. wrote the manuscript.

**Additional information**

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Latham, J. et al. Integrated catalysis opens new arylation pathways via regiodivergent enzymatic C–H activation. *Nat. Commun.* 7:11873 doi: 10.1038/ncomms11873 (2016).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material.

To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/