A *de novo* peroxidase is also a promiscuous yet stereoselective carbene transferase

Stenner, Richard[1,2]; Steventon, Jack W.[1,3]; Seddon, Annela[2,4]; Anderson, J. L. Ross[1,3*].

[1]School of Biochemistry, University of Bristol, University Walk, Bristol, BS8 1TD, UK.
[2]Bristol Centre for Functional Nanomaterials, HH Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol, BS8 1TL, UK.
[3]BrisSynBio Synthetic Biology Research Centre, Life Sciences Building, University of Bristol, Tyndall Avenue, Bristol BS8 1TQ, UK.
[4]School of Physics, HH Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol, BS8 1TL, UK.

email: ross.anderson@bristol.ac.uk

By constructing an *in vivo* assembled, catalytically proficient peroxidase, C45, we have recently demonstrated the catalytic potential of simple, *de novo*-designed heme proteins. Here we show that C45's enzymatic activity extends to the efficient and stereoselective intermolecular transfer of carbenes to olefins, heterocycles, aldehydes and amines. Not only is this the first report of carbene transferase activity in a completely *de novo* protein, but also of enzyme-catalyzed ring expansion of aromatic heterocycles *via* carbene transfer by any enzyme.

Despite the significant advances in protein design, there still remain few examples of *de novo* enzymes that both approach the catalytic efficiencies of their natural counterparts and are of potential use in an industrial or biological context[1-7]. This reflects the inherent complexities experienced in the biomolecular design process, where approaches are principally focussed on either atomistically-precise redesign of natural proteins to stabilise reaction transition states[1-3] or imprinting the intrinsic chemical reactivity of cofactors or metal ions on simple, generic protein scaffolds[4-7]; both can be significantly enhanced by implementing powerful, yet randomised directed evolution strategies to hone and optimise incipient function[8,9]. While the latter approach often results in *de novo* proteins that lack a singular structure[4,10,11], the incorporation of functionally versatile cofactors, such as heme, is proven to facilitate the design and construction of *de novo* proteins and enzymes that recapitulate the function of natural heme-containing proteins in stable, simple and highly mutable tetrahelical chassis (termed maquettes)[12-14]. Since the maquettes are designed from first principles, they lack any natural evolutionary history and the associated functional interdependency that is associated with natural protein scaffolds can be largely circumvented[12].

We have recently reported the design and construction of a hyperthermostable maquette, C45, that is wholly assembled *in vivo*, hijacking the natural *E. coli* cytochrome *c* maturation system to covalently append heme onto the protein backbone (Fig. 1a)[4]. The covalently-linked heme C of C45 is axially ligated by a histidine side chain at the proximal site and it is likely that a water molecule occupies the distal site, analogous to the ligation state of natural heme-containing peroxidases[15] and metmyoglobin[16]. Not only does C45 retain the reversible oxygen binding capability of its ancestral maquettes[4,12,13], but it functions as a promiscuous and catalytically proficient peroxidase, catalyzing the oxidation of small molecules, redox proteins and the oxidative dehalogenation of halogenated phenols with kinetic parameters that match and even surpass those of natural peroxidases[4].

It has been recently demonstrated that several natural heme-containing proteins and enzymes (e.g. cytochromes P450[17], globins[18], cytochrome *c*[19]) are capable of accessing many chemistries intrinsic to the heme cofactor, not all of which are essential to life-supporting biological roles. These reported activities are mostly dependent on accessing hypothetical heme-based carbene and nitrene intermediates[20] analogous to the oxene intermediates, compounds I and II, observed in the catalytic cycles of heme-containing peroxidases and oxygenases[15]. Several groups have now reported several examples of natural and engineered hemoproteins that catalyze cyclo-propanations[17,21,22], C-H insertions[23], carbonyl olefinations[24,25], N-H insertions[26], aziridinations[27] and carbon-silicon bond formations[19], most of which have been exposed to several rounds of directed evolution to improve stereoselectivity and product yield. Excluding the aziridation reactions, an electrophilic metallocarbenoid intermediate (Fig. 1b) is hypothesised to be responsible for carbene transfer to a suitable nucleophile (e.g. olefin)[20,28]. We therefore reasoned that if the metallocarbenoid intermediate could be detected spectroscopically then C45 may function as a promiscuous carbene transferase, catalyzing a range of important and challenging organic transformations.

To spectroscopically isolate the metallocarbenoid intermediate, we rapidly mixed ferrous C45 with the carbene precursor, ethyl diazoacetate (EDA) at 5 ^{o}C in a stopped-flow spectrophotometer. Concomitant with the disappearance of the ferrous C45 spectrum was the appearance of a new spectroscopically distinct species over 60 seconds, with a red-shifted Soret peak at 426 nm and broad Q-bands centred at 543 and 584 nm (Fig. 1c). This new spectrum is consistent with those observed in experimentally-produced carbene:iron porphyrin complexes[29] and, most recently, an evolved cytochrome *c*[30]. The cold conditions on rapid mixing proved essential in observing the spectroscopic intermediate, as room temperature experiments resulted in spectra indicative of carbene-induced heme degradation, consistent with a mechanism proposed by Arnold[31].

We then investigated the ability of C45 to act as an active carbene transferase in the cyclopropanation of styrene, commonly used as a acceptor for heme protein-derived metallocarbenoids[17]. Following reactions initiated at 5 ^{o}C under anaerobic conditions, we analysed the cyclopropanation activity of C45 and hemin when mixed with EDA and styrene by chiral HPLC and LC-MS. Hemin exhibited little cyclopropanation activity, and under the reaction conditions employed here, only unreacted EDA was observed. However, analysis of the C45-catalyzed reaction revealed the presence of two peaks in the chiral HPLC at 7.23 and 9.00 mins, corresponding to the (*R,R*) and (*S,S*) enantiomers of the Et-CPC product (Supplementary Fig. 2). Following base-hydrolysis of the Et-CPC to the corresponding acid (CPC) and further LC-MS and chiral HPLC analysis against commercial standards (Supplementary Fig. 2), we determined that the C45-catalyzed cyclopropanation of styrene was highly diastereoselective (>95 % *de*), occurs at relatively high yield (69.7 %, TTN = 697) and exhibits significant enantioselectivity, with an enantiomeric excess (*ee*) of 77 % in favour of the (*R,R*) enantiomer. C45 will accept derivatised diazoacetates and *para*-substituted styrenes as substrates for cyclopropanation, producing products with varying stereoselectivities and yields (Fig. 2, Supplementary Figs. 3 & 4). We postulate that the observed stereoselectivity exhibited by C45 likely results from the intrinsic asymmetry of the protein scaffold, and offers a means by which we can improve subsequent C45-derived maquettes. In fact, small-scale screening of a C45 library generated by error-prone pcr identified a double mutant of C45 (JR1: C45 G67S/A93S) with significantly improved yield (>95 % *de;* yield = 79.8 %; TTN = 798) and a remarkably higher enantioselectivity

towards the (*R,R*) product (*ee* = 92 %) (Supplementary Fig. 2C). As is the case for many enzymes improved by random laboratory evolution methodologies, it is not immediately clear why these changes result in the observed enhancements[32], but it demonstrates well the mutability and catalytic potential of the scaffold and will be explored in more depth in future work.

With stereoselective carbene transferase activity firmly established for the cyclopropanation of styrene, we explored the substrate promiscuity of C45 towards similar reactions resulting in N-H insertions and carbonyl olefinations. Following reaction conditions identical to those described above for cyclopropanation and using HPLC and LC-MS to identify insertion products, we observed the C45-catalyzed insertion of EDA-derived carbene into the N-H bonds of *para*-chloroaniline and piperidine (Fig. 2, Supplementary Figs. 5 - 9), with single and double insertion products resulting from carbene transfer to the primary amine of *para*-chloroaniline. For carbonyl olefination, the addition of triphenylphosphine to the reaction mixture is necessary to facilitate the formation of an ylide that, through the generation of an oxaphosphetane intermediate, spontaneously collapses to the product[25]. We investigated the olefination of four potential aldehyde substrates: benzaldehyde, *p*-nitrobenzaldehyde, *p*-cyanobenzaldehyde and 3,4,5-trimethoxybenzaldehyde. After overnight reactions were performed, the products were again analysed by HPLC and LC-MS, and though yields and TTNs were typically low, they were consistent with values reported for several myoglobin variants (Fig. 2, Supplementary Figs. 10 - 12)[25]. Interestingly, C45 did not exhibit discernible diastereoselectivity for carbonyl olefination, possibly indicating that the ylide is released from the protein prior to rearrangement and final product formation. Nevertheless, the ability to catalyze these reactions further illustrates the utility of C45 as a general carbene transferase enzyme.

Ring expansion reactions are exceptionally useful in organic synthesis as they provide a reliable and facile method for acquiring large, expanded ring systems[33,34]. Though currently underused, the homologous ring expansion of nitrogen-containing heteroaromatics could be of considerable use in the synthesis of pharmaceuticals and natural products. While both the Buchner ring expansion[35] and Ciamician-Dennstedt reaction[36] proceed *via* a cyclopropane-containing bicyclic system that is subsequently ring-opened, the ring opening during the latter reaction is facilitated by the expulsion of a halogen leaving group. We therefore hypothesised that with a halogen-substituted carbene precursor, C45 would catalyze the cyclopropanation of a nitrogen-containing heteroaromatic to produce a cyclopropane-containing bicyclic system susceptible to spontaneous rearrangement (Fig. 3a). Using the same reaction conditions as those described for C45-catalyzed styrene cyclopropanation, we investigated whether C45 could catalyze the ring expansion of pyrrole with ethyl-2-bromo-2-diazoacetate[37] as the carbene precursor. After 2 hours, the reaction was analysed using HPLC and LC-MS, confirming the production of the ring-expanded product, ethyl nicotinate at 57 % yield (TTN = 570) (Fig. 3b, Supplementary Fig. 13). Subsequent base-catalyzed hydrolysis of the product yields the nicotinamide and NAD(P)H precursor niacin, raising the possibility of engineering a life-sustaining, artificial biochemical pathway from pyrrole to nicotinamide reliant on the *in vivo* activity of C45 or a related *de novo* designed enzyme. To the best of our knowledge, this is the first example of any enzyme - natural, engineered or *de novo* - that is capable of catalyzing homologous ring expansion reactions *via* a carbene transfer mechanism.

To further explore the possibility of employing C45 in a new, essential and life-sustaining pathway from pyrrole to the pyridine nucleotides, we examined the ring-expansion of

pyrrole to ethyl nicotinate in living *E. coli* cells. Using a well-established procedure for carbene transferase activity under such conditions[38], we tested the ability of *E. coli* bearing the C45-expression plasmid and pEC86 (harbouring the *c*-type cytochrome maturation apparatus)[39] to perform the ring expansion reaction of pyrrole with ethyl-2-bromo-2-diazoacetate under anaerobic conditions (Fig. 3b). Since the maturation apparatus is constitutively expressed and results in the production of several heme-containing membrane proteins (CcmC, CcmE & CcmF)[40,41], it was vital to establish whether any intrinsic ring expansion activity was detectable in the presence of these proteins. Whole cells that had been grown for 3 and 6 hours in the absence of the inducer, IPTG, had barely detectable ring-expansion activity (Fig. 3b). In contrast, cells that were grown for 3 hours, induced with IPTG and grown for a further 3 hours displayed significant ethyl nicotinate formation, indicating that under these conditions and *in vivo*, C45 exhibits the catalytic ring-expansion activity. It also indicates that both pyrrole and ethyl-2-bromo-2-diazoacetate are able to cross the outer membrane and access the periplasmically-located C45. Interestingly, despite the production of ethyl nicotinate by the cells, we did not observe an increase in the quantity of niacin in the extract, indicating that *E. coli* possibly lacks - or does not produce an appreciable quantity of - an endogenous periplasmic esterase capable of efficiently hydrolysing the product. We therefore tested a recombinant *Bacillus subtilis* esterase[42] expressed in, and purified from, *E. coli* for ethyl nicotinate hydrolysis activity. Incubation of this esterase with ethyl nicotinate under near-physiological conditions results in the production of niacin (Supplementary Fig. 14), thus highlighting another functional part in a potential novel biosynthetic pathway from pyrrole to the pyridine nucleotides through niacin.

To this end, we speculate that it is possible to make C45 and the *Bacillus* esterase essential to NAD+ biosynthesis. *E. coli* synthesises NAD+ through two pathways that both involve the production of NaMN (Supplementary Fig. 15), then NaAD prior to NAD+ formation[43-46]: the *de novo* pathway, using *L*-aspartate as a starting material to produce quinolate which is a NaMN precursor[44]; the pyridine ring salvage pathway where exogenous niacin or nicotinamide are utilised instead as NaMN precursors, representing the favoured pathway when niacin and nicotinamide are abundant in the environment[46]. Knocking out a key enzyme in the aspartate pathway - e.g. nicotinate-nucleotide diphosphorylase (NadC) - will provide an auxotrophic strain of *E. coli* that, when expressing both C45 and a suitable esterase in nicotinamide and niacin-lacking media, should be capable of converting pyrrole and ethyl-2-bromo-2-diazoacetate to niacin and ultimately NAD+, recovering the deleterious phenotype. Currently, production of such an *E. coli* strain is outwith the scope of this study, but this strategy highlights the mechanism by which we can tractably and rationally engineer the metabolism of a bacterium to rely on a *de novo* enzyme to sustain an essential pathway.

With this work, we have showcased the exceptionally diverse functionality available in this simple, *de novo* designed heme protein. The formation of the metallocarbenoid intermediate at the C45 heme facilitates not only the cyclopropanation of styrene and its derivatives, but also extends to the insertion of carbenes into N-H bonds, the olefination of carbenes and the first description of enzyme-catalyzed ring expansion of a nitrogen heterocycle. This demonstrates that abiological function is intrinsic to these *de novo* designed maquettes, and that the higher complexity of natural protein frameworks, such as the globins and cytochromes P450, is not necessary to support reactivity of this type. Given the significant enhancement of yield and stereoselectivity exhibited by JR1, we have also shown that with relatively minor changes to C45, reactivity can be profoundly affected. This indicates that further yield enhancements and alternative

stereoselectivities are undoubtedly achievable through directed evolution, consistent with the studies reported for the maquette's natural counterparts[17,19,23]. Furthermore, there are many beneficial features of C45 that compare favourably against the work described by others with natural enzymes: C45 is expressed at high levels in *E. coli*, and is functionally assembled *in vivo*[4], eliminating the requirement for the *in vitro* incorporation of abiological metalloporphyrins[23] and facilitating *in vivo* function[38]; C45 is thermally stable and displays excellent tolerance in organic solvents[4], thereby facilitating its use as a homogenous catalyst in aqueous:organic mixtures. In conclusion, this work demonstrates the biocatalytic potential and reactive promiscuity that maquettes possess, and future work will be concerned with improving the substrate scope and catalytic performance of these novel, *de novo* enzymes. It also reinforces our previous work[4], demonstrating that maquettes and related *de novo* proteins are much more than just lab curiosities or ornaments, and are instead powerful, green catalysts that can play a valuable role in facilitating challenging organic transformations.

Methods

General & Molecular biology, protein expression and purification

All chemicals were purchased from either Sigma or Fisher Scientific. Protein expression and purification in *E. coli* T7 Express (NEB) from the pSHT vector was carried out as described previously[4]. Error prone PCR (epPCR) libraries of C45, with a mutation rate of 2-3 amino acid mutations, were produced using the GeneMorph II EZClone Domain Mutagenesis Kit. The mutation rate was determined by randomly selecting and sequencing 10 colonies after transforming the PCR products into *E. coli* T7 Express cells. This mutation rate was achieved with 18.7 ng of target DNA (C45 coding sequence). JR1 was one of three randomly selected variants from the initial epPCR C45 library. Each variant was expressed and purified as perviously described for C45, and the purified proteins were used in the cyclopropanation assays.

Synthesis of ethyl 2-bromo-2-diazoacetate

Ethyl 2-bromo-2-diazoacetate was synthesised as previously reported[37]. Briefly, *N*-bromosuccinimide (260 mmol) was added to a solution of EDA (200 mmol) and 1,8-Diazabicyclo(5.4.0)undec-7-ene (280 mmol) in CH_2Cl_2 (5.0 ml) at 0 °C, and the reaction mixture was stirred for 10 mins. The crude reaction mixture was washed with cold $Na_2S_2O_3$ (aqueous, 3 x 5 ml) and quickly filtered through a silica column (cold CH_2Cl_2). Cold CH_2Cl_2 was added to bring the final volume up to 2.5 ml (~400 mM).

Stopped flow spectrophotometry

Stopped-Flow kinetics were conducted using a SX20 stopped-flow spectrophotometer (Applied Photophysics)housed in an anaerobic glovebox under N_2 ([O_2] < 5 ppm; Belle Technology). A solution containing a known concentration of C45 or JR1 (20 µM, 100 mM KCl, 20 mM CHES, pH 8.6) was placed in one syringe and a 20 mM solution of EDA (carbene precursor) in ethanol was placed in the second syringe. 50 µl from each syringe was simultaneously injected into a mixing chamber and the progression of the reaction was monitored spectroscopically at 5 ^{o}C and 25 ^{o}C examining the metallocarbenoid formation and porphyrin degradation pathway respectively.

Carbene Transfer Chemistry

Unless stated otherwise, all assays were conducted under scrubbed nitrogen inside an anaerobic glovebox ([O_2] < 5 ppm; Belle Technology). The assays were conducted inside 1.5 ml screw top vials sealed with a silicone-septum containing cap. All assays were conducted in CHES buffer (100 mM KCl, 20 mM CHES, pH 8.6) and the final reaction volumes were 400 µl unless otherwise stated.

Cyclopropanation assays

To 370 µl of a 10 µM C45 solution was added 10 µl of $Na_2S_2O_4$ (400 mM stock; de-ionised water) and the mixture was left to stir for 1 minute, ensuring complete reduction

of C45 from Fe^{3+} to Fe^{2+}. 10 µl of the selected styrene (1.2 M stock in EtOH) was added and the reaction left to mix for 30 seconds. A separate vial containing the selected diazo compound (400 mM stock in EtOH) was deoxygenated alongside the reaction vessel. The vials were sealed, transported out of the glovebox and cooled to 0 ^{o}C in an ice bath. Once cooled, 10 µl of the diazo compound was added *via* gastight syringe into the reaction vials to initiate the reaction. Final reaction concentrations were 10 µM enzyme, 10 mM sodium dithionite, 10 mM diazo compound, and 30 mM styrene. Once mixed, the reactions were stirred on a roller at room temperature. After 2 hours, the reaction was quenched by the addition of 20 µl of 3 M HCl. The vials were subsequently unscrewed and 1 ml of ethyl acetate was added to the vial. The solution was transferred to a 1.5 ml microcentrifuge tube, vortexed and centrifuged for 1 minute at 13,500 rpm. The upper organic layer was extracted, dried over $MgSO_4$ if necessary, and subsequently used for analysis. Where base hydrolysis of the resultant ester was necessary,400 µl of 3 M NaOH was introduced to the organic layer and the mixture was left to stir, at room temperature for 30 minutes. Products were analysed by chiral-HPLC and LC-MS as described below. The product yields, enantiomeric excesses and total turnover numbers (TTN; concentration of product formed/concentration of enzyme) where calculated *via* external calibration with commercial ethyl 2-phenylcyclopropane-1-carboxylate and 2-phenylcyclopropane-1-carboxylic acid.

Carbonyl olefination assays

Carbonyl olefination assays were conducted under the same reaction conditions as the cyclopropanation assays, with the following alterations: alkene substrates were substituted for selected benzaldehyde substrates; 10 mM of PPh_3 was included as an additional reagent in the reaction (10 µl of a 400 mM stock in acetone); the reactions were carried out overnight; after quenching, the products were extracted with 1 ml of dichloromethane; the products were analysed by C18 reverse-phase HPLC and LC-MS as described below. The product yields, cis/trans ratios and total turnover numbers where calculated *via* an external calibration with commercial ethyl cinnamate and cinnamic acid.

N-H insertion assays

N-H insertion assays were conducted under the same reaction conditions as the cyclopropanation assays, with the following alterations: the alkene starting materials were substituted for the selected amines; the reactions were carried out overnight; after quenching, the products were extracted with 1 ml of *n*-hexane; the products were subsequently analysed by C18 reverse-phase HPLC and LC-MS as described below. The product yield and total turnover numbers where calculated *via* an external calibration with commercial *n*-phenyglycine ethyl ester and *n*-ethylpiperidine acetate.

Ring expansion assays

Ring expansion assays were conducted under the same reaction conditions as the cyclopropanation assays described above, but with the following alterations: styrene was substituted for pyrrole; ethyl 2-bromo-2-diazoacetate was used as the carbene precursor; the products were analysed by C18 reverse-phase HPLC and LC-MS as described below. The product yields and total turnover numbers where calculated *via* an external calibration with commercial ethyl nicotinate and niacin.

Product Characterisation by HPLC

All the reactions performed using C45 and JR1 were quantified by High Performance Liquid Chromatography (HPLC). Two separate columns were employed for quantification: a reverse phase C18 column and a chiral column.

A C18 HPLC reverse phase column (Phenomenex, 150 x 15 mm, 5 µm) was used to analytically quantify ring expansion, carbonyl olefination and N-H insertion assays. A gradient (55:45 % H_2O:CH_3CN to 10:90 % H_2O:CH_3CN; 0.1 % v/v Formic acid; 2 ml.min^{-1}) was employed as the mobile phase for the ring expansion assays. An isocratic

polar ionic mobile phase (100 % CH_3CN: 0.1 % v/v TFA: 0.1 % v/v: Et_3N; 0.1 ml.min^{-1} for cyclopropanation, 1 ml.min^{-1} for carbonyl olefination and 0.2 ml.min^{-1} for N-H insertions) was employed for analysing the carbonyl olefination and N-H insertion assays. Injection volumes of the sample mixture were 20 µl and the elution was monitored spectroscopically at 220, 254 and 280 nm. The C18 column allowed the retention times of the starting materials and the reaction products to be determined, and to quantify the diastereoselectivity of a given reaction.

If appreciable product formation could be detected after the initial HPLC experiment, the product mixture was analysed again over a 20 minute gradient (30:70% H_2O:CH_3CN to 10:90% H_2O:CH_3CN; 0.1 % v/v Formic acid; 0.5 ml.min^{-1}) mobile phase. This protocol was subsequently adopted in the Liquid chromatography-Mass spectrometry (LC-MS) experiments and the retention times of the product using this protocol facilitated product hunting and data analysis.

A Chirobiotic V chiral column (Astec, 25 cm x 2.1 mm, 5 µm, with a vancomycin stationary phase) was employed to analytically separate mixtures of possible product enantiomers resulting from the cyclopropanation assays. A polar ionic mobile phase (100 % CH_3CN: 0.1 % v/v TFA: 0.1 % v/v: Et_3N; 0.1 ml.min^{-1}) was employed. Injection volumes of the sample mixture were 1-2 µl and the elution was monitored spectroscopically at 220 nm, 245 nm, 254 nm and 280 nm.

Liquid chromatography-Mass spectrometry (LC-MS) was employed to quantify product formation in each assay. A C18 reverse column (Phenomenex, 150 x 10, 5 µm), identical to the column employed in the HPLC experiments, was used and flow rate of 0.5 ml.min^{-1} over a 20 minute gradient (30:70 % H_2O:CH_3CN to 10:90 % H_2O:CH_3CN; 0.1% v/v Formic acid) was adopted as the mobile phase. Injection volumes of 20 µl were employed and the chromatogram was screened across a wavelength range of 240-300 nm. After eluting from the column, the mixture entered an isocratic solvent chamber where a 1:100 dilution was performed prior to injecting the sample into a positive electron-spray-ionisation (ESI) mass spectrometer (Waters Xevo G2-XS QTof). The mass spectrum contained peaks screened across a m/z range of 70-250. Retention times were initially determined using HPLC to facilitate product hunting and data analysis.

External calibrations were conducted using HPLC with commercial standards corresponding to the expected products of the assays. The conditions and mobile phases employed in the product analysis for each assay was employed for the calibrations. Each substrate was prepared at multiple concentrations (500 µM, 1 mM, 2 mM, 5 mM, 7.5 mM, 10 mM, and 20 mM) and the peak height response in the chromatogram was recorded. A plot of [substrate] vs peak height engendered a straight line which could be used to determine the product yields for each assay.

Whole cell C45-catalysed ring expansion experiments

Overnight starter cultures were prepared by adding 100 µl carbenicillin (50 mg/ml) and 100 µl chloramphenicol (50 mg/ml, C45 only) to 100 ml of LB before inoculating the media with a C45-expressing *E. coli* glycerol stock. Starter cultures were then incubated overnight at 37 ^{o}C and 180 rpm. 50 ml of the overnight starter culture was then used to inoculate 1 L of LB containing the same concentrations of antibiotics (see above). The 1 L cultures were grown in a shaking incubator at 37 ^{o}C and 180 rpm until an OD_{600} between 0.6-0.8 was obtained (usually after 3 hours), at which point 1 ml was extracted from both vessels, placed inside separate 1.5 ml screw top vials and placed on ice. 1 ml of IPTG solution (1 M stock, 1 mM final concentration at induction) was added to specific cultures to induce protein expression; induced and non-induced cultures were left in the incubator for an addition three hours (37 ^{o}C and 180 rpm). After three hours, 1 ml was extracted from the vessels and transferred to separate 1.5 ml screw top vials. The samples were then degassed inside an anaerobic glove-box (Belle Technology) for

30 minutes before 10 µl of pyrrole (3 M stock in EtOH) was added and the vials sealed. The vials were cooled on ice before 10 µl of ethyl 2-bromo-2-diazoacetate (1 M stock in DCM) was added *via* a needle, under nitrogen; the final concentration of the reagents were 10 mM ethyl 2-bromo-2-diazoacetate and 30 mM pyrrole. After the reactions had been left stirring for 2 hours, the samples were quenched with 3 M HCl (30 µl) and extracted with 1 ml ethyl acetate. Analysis was conducted as reported as above in the "Ring expansion assays" section.

Hydrolysis of ethyl nicotinate by a *Bacillus subtilis* esterase

To a 100 ml round-bottom flask equipped with a small magnetic stirrer bar was added 19.8 ml of CHES buffer (pH 8.6) and 200 µl of commercial ethyl nicotinate (5 M stock in DMSO). The final concentration of ethyl nicotinate was 50 mM. A 100 µl aliquot of the solution was extracted to an Eppendorf tube for HPLC analysis. 2 mg of esterase from *Bacillus subtilis* (SigmaAldrich) was added to the mixture, then the round-bottom flask was closed with a stopper and the solution was left to stir for 1 hour. After 1 hour, a 1 ml aliquot of the solution was extracted into a 1.5 ml Eppendorf tube, and 100 µl of 3 M trichloroacetate was added to precipitate the protein. The mixture was vortexed, centrifuged (30 seconds, 13000 rpm) and the resulting solution analysed directly *via* HPLC. The HPLC protocol for the analysis was identical to the procedure outlined in the ''ring expansion assays'' section above.

Acknowledgements

This work was supported at the University of Bristol by the BBSRC (grant no: BBI014063/1, BB/R016445/1 & BB/M025624/1), the Bristol Centre for Functional Nanomaterials (EPSRC Doctoral Training Centre Grant EP/G036780/1) through a studentship for R.S. and by the SynBioCDT (EPSRC and BBSRC Centre for Doctoral training in Synthetic Biology EP/L016494/1) through a studentship for J.S. We also wish to thank Dr. Peter Wilson for his assistance in collecting LC-MS data.

Author Contributions

R.S., J.S. and J.L.R.A. designed the experiments. R.S. and J.S. performed the experiments. All authors analysed the data and contributed to the manuscript preparation. R.S. and J.L.R.A. wrote the manuscript.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. Röthlisberger, D. *et al.* Kemp elimination catalysts by computational enzyme design. *Nature* **453**, 190–195 (2008).
2. Zymer, C., Zschoche, R., Hilvert, D. Optimization of Enzyme Mechanism along the Evolutionary Trajectory of a Computationally Designed (Retro-)Aldolase. *J. Am. Chem. Soc.* **139**, 12541-12549 (2017).
3. Preiswerk, N., *et al.* Impact of scaffold rigidity on the design and evolution of an artificial Diels-Alderase. *Proc. Natl. Acad. Sci. USA* **111**, 8013-8018 (2014).
4. Watkins, D.W., *et al.* Construction and in vivo assembly of a catalytically proficient and hyperthermostable de novo enzyme. *Nature Communications* **8**, 358 (2017).
5. Donnelly, A.E., Murphy, G.S., Digianantonio, K.M., Hecht, M.H. A de novo enzyme catalyzes a life-sustaining reaction in Escherichia coli. *Nat. Chem. Biol.* **14**, 253-255 (2017).

6. Snyder, R.A., Butch, S.E., Reig, A.J., Degrado, W.F., Solomon, E.I. Molecular-Level Insight into the Differential Oxidase and Oxygenase Reactivities of de Novo Due Ferri Proteins. *J. Am. Chem. Soc.* **137**, 9302-9314 (2015).
7. Zastrow, M. L., Peacock, A. F. A., Stuckey, J. A., Pecoraro, V. L. Hydrolytic catalysis and structural stabilisation in a designed metalloprotein. *Nat. Chem.* **4**, 118-123 (2012).
8. Blomberg, R. *et al.* Precision is essential for efficient catalysis in an evolved Kemp eliminase. *Nature* **503**, 418-421 (2013).
9. Patel, S.C. & Hecht, M.H. Directed evolution of the peroxidase activity of a de novo-designed protein. *Protein Eng. Des. Sel.* **25**, 445-452 (2012).
10. Nanda, V. K., Koder, R. L. Designing artificial enzymes by intuition and computation, *Nat. Chem.* **2**, 15-24 (2010).
11. Murphy, G. S., Griesman, J. B., Hecht, M. H. *De novo* proteins with life-sustaining functions are structurally dynamic. *J. Mol. Biol.* **428**, 399-411 (2016).
12. Farid, T. A. *et al.* Elementary tetrahelical protein design for diverse oxidoreductase functions. *Nat. Chem. Biol.* **9**, 826-833 (2013).
13. Koder, R. L. *et al.* Design and engineering of an O_2 transport protein. *Nature* **458**, 305-309 (2009).
14. Anderson, J. L. R. *et al.* Constructing a man-made *c*-type cytochrome maquette *in vivo*: electron transfer, oxygen transport and conversion to a photoactive light harvesting maquette. *Chem. Sci.* **5**, 507-514 (2014).
15. Poulos, T. Heme enzyme structure and function. *Chem. Rev.* **114**, 3919-3962 (2014).
16. Antonini, E. & Brunori, M. *Hemoglobin and myoglobin in their reactions with ligands* 47-48 (North-Holland Pub. Co., 1971).
17. Coelho, P.S., Brustad, E.M., Kannan, A., Arnold, F.H. Olefin cyclopropanation via carbene transfer catalyzed by engineered cytochrome P450 enzymes. *Science* **339**, 485-487 (2013).
18. Sreenilayam, G. & Fasan, R. Myoglobin-catalyzed intermolecular carbene N-H insertion with arylamine substrates. *Chem. Commun.* **51**, 1532-1534 (2015).
19. Kan, S.B., Lewis, R.D., Chen, K., Arnold, F.H. Directed evolution of cytochrome c for carbon-silicon bond formation: Bringing silicon to life. *Science* **354**, 1048-1051 (2016).
20. Brandenberg, O.F., Fasan, R., Arnold, F.H. Exploiting and engineering hemoproteins for abiological carbene and nitrene transfer reactions. *Curr. Opin. Biotechnol.* **47**, 102-111 (2017).
21. Oohora, K., *et al.* Catalytic cyclopropanation by myoglobin reconstituted with iron porphycene: acceleration of catalysis due to rapid formation of the carbene species. *J. Am. Chem. Soc.* **139**, 17265-17268.
22. Villarino, L., *et al.* An artificial heme enzyme for cyclopropanation reactions. *Angew. Chem. Intl. Ed. Engl.* **57**, 7785-7789.
23. Key, H.M., Dydio, P., Clark, D.S., Hartwig, J.F. Abiological catalysis by artificial haem proteins containing noble metals in place of iron. *Nature* **534**, 534-537 (2016).
24. Weissenborn, M.J., *et al.* Enzyme-catalyzed carbonyl olefination by the *E. coli* protein YfeX in the absence of phosphines. *ChemCatChem* **8**, 1636-1640 (2016).
25. Tyagi, V. & Fasan, R. Myoglobin-Catalyzed Olefination of Aldehydes. *Angew. Chem. Intl. Ed. Engl.* **55**, 2512-2516 (2016).
26. Wang, Z.J., Peck, N.E., Renata, H., Arnold, F.H. Cytochrome P450-Catalyzed Insertion of Carbenoids into N-H Bonds. *Chem. Sci.* **5**, 598-601 (2014).
27. Farwell, C.C., Zhang, R.K., McIntosh, J.A., Hyster, T.K., Arnold, F.H. Enantioselective Enzyme-Catalyzed Aziridination Enabled by Active-Site Evolution of a Cytochrome P450. *ACS Cent. Sci.* **1**, 89-93 (2015).

28. Wei, Y., Tinoco, A., Steck, V., Fasan, R., Zhang, Y. Cyclopropanations via Heme Carbenes: Basic Mechanism and Effects of Carbene Substituent, Protein Axial Ligand, and Porphyrin Substitution. *J. Am. Chem. Soc.* **140**, 1649-1662 (2018).
29. Liu, Y., *et al*. Electronic Configuration and Ligand Nature of Five-Coordinate Iron Porphyrin Carbene Complexes: An Experimental Study. *J. Am. Chem. Soc.* **139**, 5023-5026 (2017).
30. Lewis, D.L., *et al.* Catalytic iron-carbene intermediate revealed in a cytochrome *c* carbene transferase. *Proc. Natl. Acad. Sci. USA* **115**, 7308-7313 (2018).
31. Renata, H., *et al*. Identification of Mechanism-Based Inactivation in P450-Catalyzed Cyclopropanation Facilitates Engineering of Improved Enzymes. *J. Am. Chem. Soc.* **138**, 12527-12533 (2016).
32. Arnold, F.H. Directed Evolution: Bringing New Chemistry to Life. *Angew. Chem. Intl. Ed.* **57**, 4143-4148 (2018).
33. Hall, J.E., Matlock, J.V., Ward, J.W., Gray, K.V., Clayden, J. Medium-Ring Nitrogen Heterocycles through Migratory Ring Expansion of Metalated Ureas. *Angew. Chem. Intl. Ed.* **55**, 11153-11157 (2016).
34. Crombie, A.L., Kane, J.L., Shea, K.M., Danheiser, R.L. Ring expansion-annulation strategy for the synthesis of substituted azulenes and oligoazulenes. 2. Synthesis of azulenyl halides, sulfonates, and azulenylmetal compounds and their application in transition-metal-mediated coupling reactions. *J. Org. Chem.* **69**, 8652-8667 (2004).
35. Büchner, E. & Curtis, T. Synthesis of beta-keto esters from aldehydes and diazoacetic acid. *Ber. Dtsch. Chem. Ges.* **18**, 2371-2377 (1885).
36. Ciamician, G.L. & Dennstedt, M. *Ber.* Über die Einwirkung des Chloroforms auf die Kaliumverbindung Pyrrols. **14**, 1153 (1881).
37. Bonge, H.T., Pintea, B., Hansen, T. Highly efficient formation of halodiazoacetates and their use in stereoselective synthesis of halocyclopropanes. *Org. Biomol. Chem.* **6**, 3670-3672 (2008).
38. Coelho, P.S., *et al.* A serine-substituted P450 catalyzes highly efficient carbene transfer to olefins in vivo. *Nat. Chem. Biol.* **9**, 483-487 (2013).
39. Arslan, E. *et al.* Overproduction of the *Bradyrhizobium japonicum c*-type cytochrome subunits of the cbb3 oxidase in *Escherichia coli*. *Biochem. Biophys. Res. Commun.* **251**, 744-747 (1998).
40. San Francisco, B., Sutherland, M.C., Kranz, R.G. The CcmFH complex is the system I holocytochrome *c* synthetase: engineering cytochrome *c* maturation independent of CcmABCDE. *Mol. Microbiol.* **91**, 996-1008 (2014).
41. Mavridou, D.A.I., *et al*. Probing heme delivery processes in cytochrome *c* biogenesis system I. *Biochemistry*, **52**, 7262-7270 (2013).
42. Henke, E., Bornscheuer, U.T., Schmid, R., Pleiss, J. A molecular mechanism of enantiorecognition of tertiary alcohols by carboxylesterases. *ChemBioChem,* **4**, 485-493 (2003).
43. McLaren, J., Ngo, D.T.C., Olivera, B.M. Pyridine nucleotide metabolism in *Escherichia coli*. III. Biosynthesis from alternative precursors *in vivo*. *J. Biol. Chem.* **248**, 5144-5159 (1973).
44. Kurnasov, O. *et al*. NAD biosynthesis: Identification of the tryptophan to quinolinate pathway in bacteria. *Chemistry & Biology* **10**, 1195-1204 (2003).
45. Wang, X., *et al*. Engineering *Escherichia coli* nicotinic acid mononucleotide adenylyltransferase for fully active amidated NAD biosynthesis. *Appl. Environ. Microbiol.* **83**, e00692-17 (2017).
46. Bouvet, O., Bourdelier, E., Glodt, J., Clermont, O., Denamur, E. Diversity of the auxotrophic requirements in natural isolates of *Escherichia coli*. *Microbiology* **163**, 891-899 (2017).

Figures

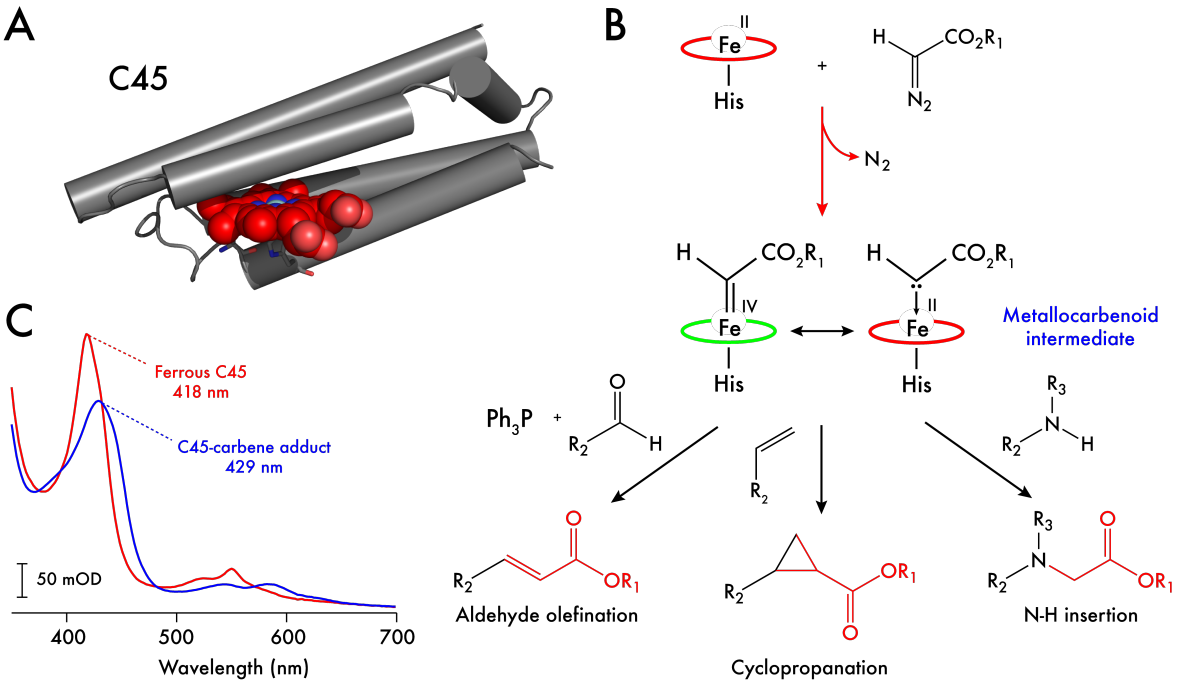

Figure 1. Metallocarbenoid formation and reactivity within a *de novo* designed *c*-type cytochrome maquette, C45. **A.** Single snapshot from a 1 µs Molecular Dynamics simulation of the C45 maquette[4]. **B.** Formation and potential reactivity of a heme-based metallocarbenoid intermediate, illustrating aldehyde olefination, olefin cyclopropanation and amine N-H insertion reactions. **C.** UV/visible spectra of ferrous C45 (red) and ethyldiazoacetate-treated C45 obtained by rapid mixing experiments in a stopped-flow spectrophotometer. The putative metallocarbenoid species was generated by mixing 20 mM ethyldiazoacetate with 20 µM ferrous C45 in 100 mM KCl, 20 mM CHES, pH 8.6.

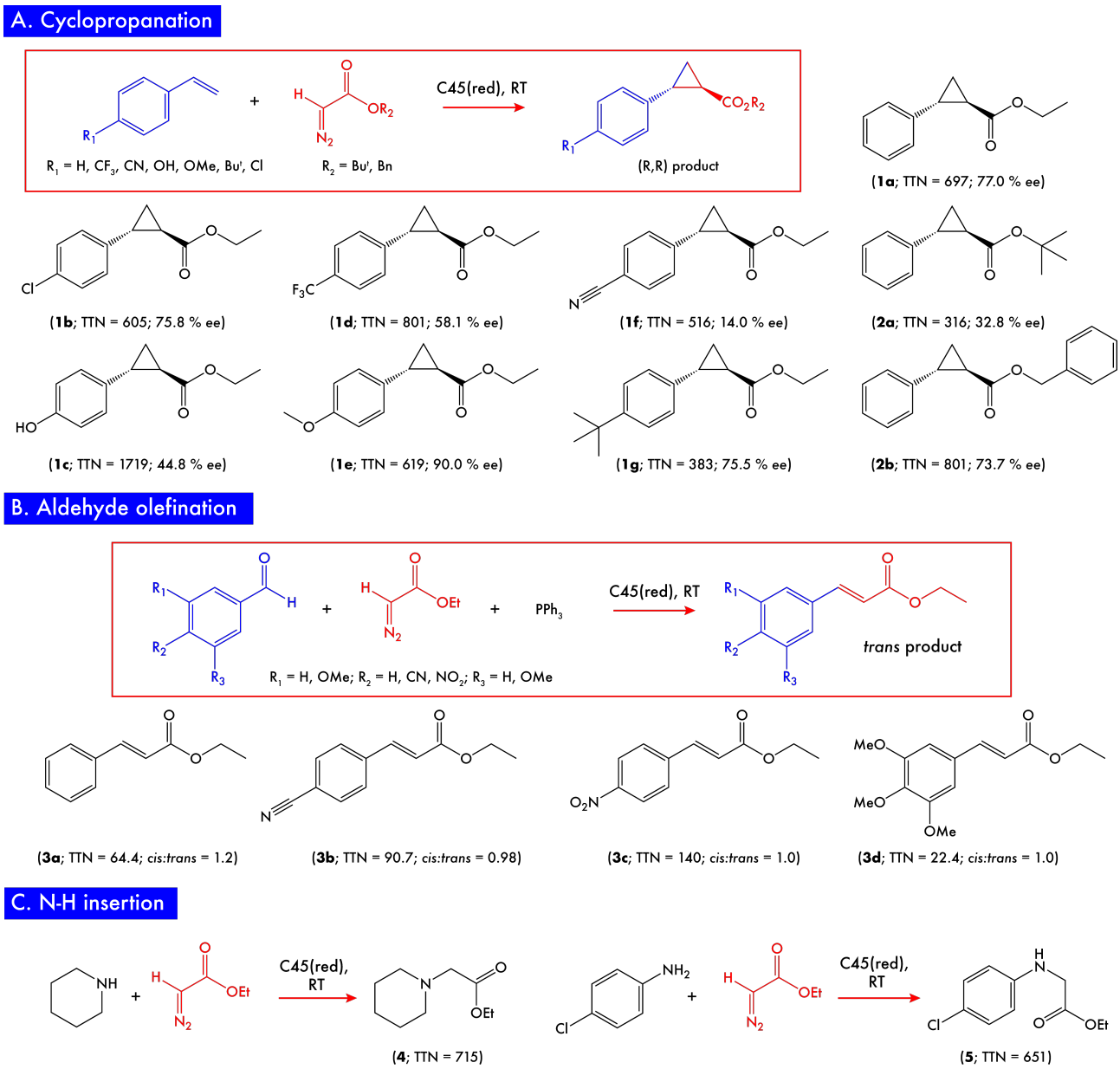

Figure 2. Carbene transferase activity of C45. **A.** Cyclopropanation of substituted styrenes catalyzed by C45. Total turnover numbers (TTN) and enantiomeric excesses (% *ee*) for each combination of ferrous C45 with para-substituted styrenes and functionalised diazoacetates. Only the (R,R) cyclopropanated product is displayed in the reaction scheme, representing the dominant product in all cases. All reactions were carried out with 0.01 % catalyst loading (10 µM C45) at the following concentrations of reagents: 10 mM sodium dithionite, 10 mM diazo compound, and 30 mM substituted styrene in 100 mM KCl, 20 mM CHES, pH 8.6. **B.** Olefination of substituted benzaldehydes catalyzed by C45. Only the *trans* product is displayed in the reaction scheme, though almost equal quantities of the *cis* product is also produced in the C45-catalyzed reactions. All reactions were carried out with 0.01 % catalyst loading (10 µM C45) at the following concentrations of reagents: 10 mM sodium dithionite, 10 mM PPh_3, 10 mM ethyldiazoacetate, and 30 mM substituted benzaldehyde in 100 mM KCl, 20 mM CHES, pH 8.6. **C.** N-H insertion of primary and secondary amines catalyzed by C45. Only the monofunctionalised product of the *p*-chloroaniline insertion reaction shown, and the corresponding TTN is calculated based on the yield of both mono- and di-substituted products. All reactions were carried out with 0.01 % catalyst loading (10 µM C45) at the following concentrations of reagents: 10 mM sodium dithionite, 10 mM ethyldiazoacetate, and 30 mM pyrrole in 100 mM KCl, 20 mM CHES, pH 8.6.

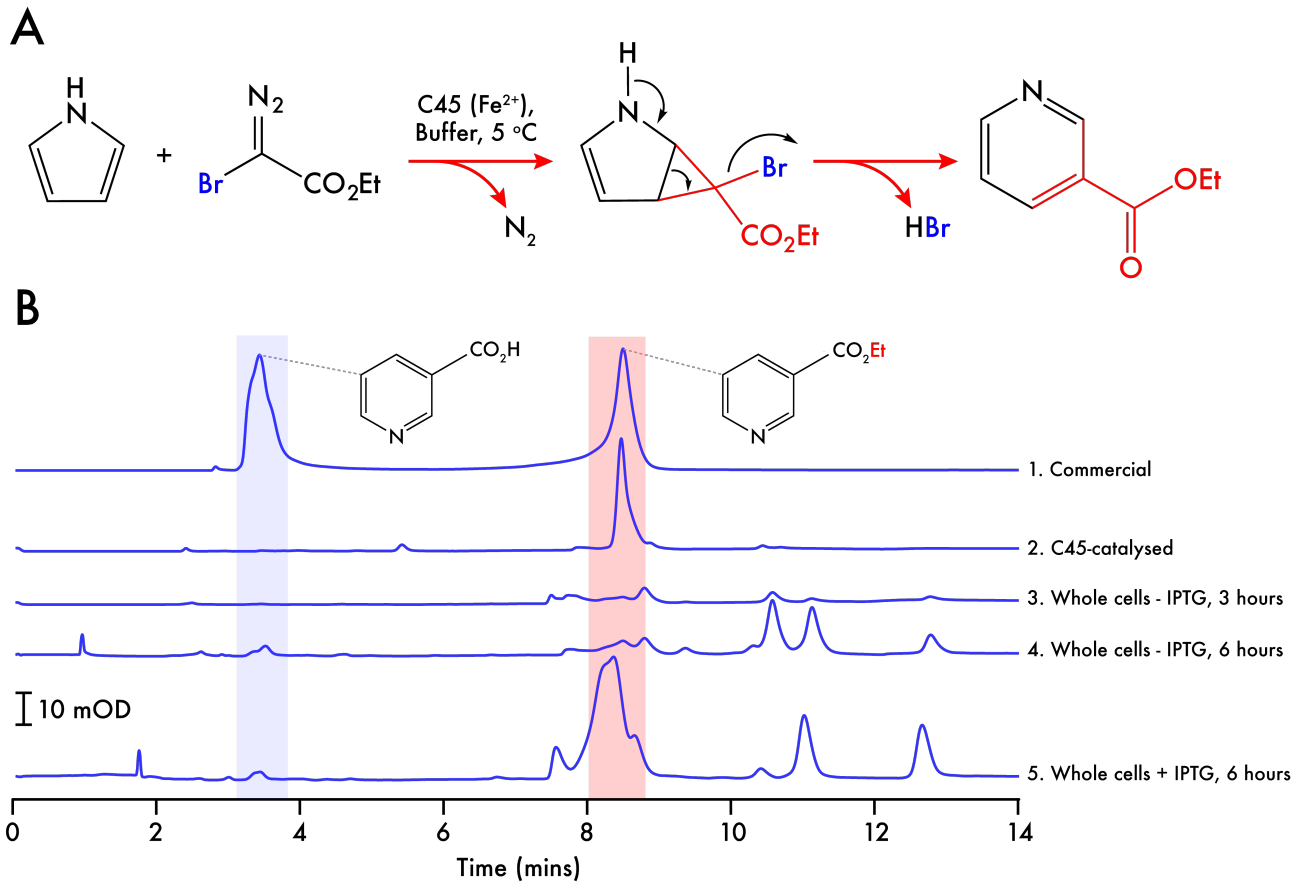

Figure 3. Heteroaromatic ring expansion catalyzed by C45. **A.** Reaction scheme for the ring expansion strategy using ethyl 2-bromo-2-diazoacetate, pyrrole and ferrous C45. Following carbene transfer to the pyrrole, spontaneous rearrangement of the bicyclic ring system leads to elimination of HBr and formation of a 6-membered pyridine ring. **B.** C18 reversed phase HPLC traces of the C45-catalyzed ring expansion of pyrrole to ethyl nicotinate. Traces 1 & 2 show the C45-catalysed ring expansion compared to a partially hydrolysed commercial standard of ethyl nicotinate. The ring expansion was carried out with 0.01 % catalyst loading (10 µM C45) at the following concentrations of reagents: 10 mM sodium dithionite, 10 mM ethyl 2-bromo-2-diazoacetate, and 30 mM pyrrole in 100 mM KCl, 20 mM CHES, pH 8.6. Traces 3,4 & 5 show the results of incubating whole cells containing the C45 expression vector and pEC86 harbouring the *E. coli* cytochrome *c* maturation apparatus. Traces 3 & 4 represent reactions between whole cells, pyrrole and ethyl 2-bromo-2-diazoacetate at 3 and 6 hours after inoculation and in the absence of the inducer, IPTG. Trace 5 represents the reaction with C45-expressing whole cells, pyrrole and ethyl 2-bromo-2-diazoacetate. In this case, the cells were grown for 3 hours, induced with 1 mM IPTG and C45 was expressed for a further 3 hours prior to use in the whole cell transformation. Reaction conditions are fully described in the materials and methods section.