Author:
Stenner, Richard A

Title:
A Promiscuous and Catalytically Proficient De Novo Carbene Transferase

A PROMISCUOUS AND CATALYTICALLY PROFICIENT *DE NOVO* CARBENE TRANSFERASE

BY RICHARD STENNER

School of Biochemistry

Richard Stenner

School of Biochemistry

rs15090@bristol.ac.uk; 07721727592

A dissertation submitted to the University of Bristol in accordance with the requirements for the award of degree of Doctor of Philosophy in the Faculty of Biomedical Sciences, November 2019

Word count: 109,016

ABSTRACT

Biocatalysis is the use of biological materials, such as whole cells or enzymes, to perform chemical transformations. One of the major obstacles limiting the widespread adoption of biocatalytic methodologies in organic synthesis is the restricted number of chemical transformations enzymes are capable of performing. To address this, the field of rational design utilises the techniques of molecule biology, such as mutagenesis, molecular cloning, and recombinant protein technology, to engineer novel enzymes with abiological activity. Rational design can be split into two complementary fields: protein redesign and *de novo* design. Protein redesign involves identifying a suitable, naturally occurring protein scaffold to function as a molecular blueprint into which desired functionality can be engineered. *De novo* design is a bottom-up approach, in which novel proteins are designed from scratch using first principles.

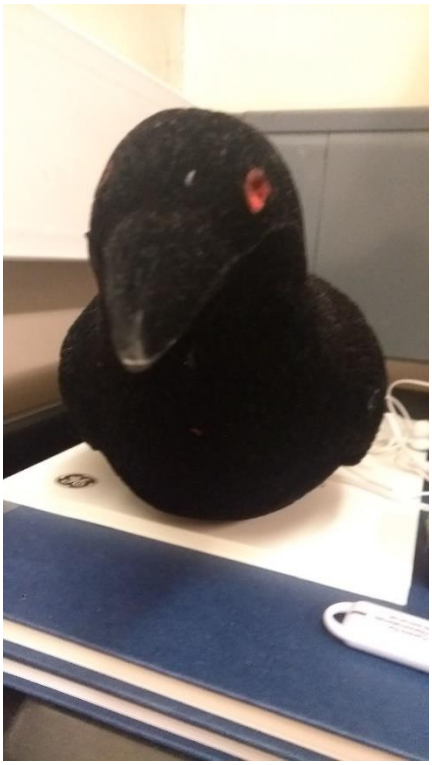
The recent success in engineering carbene transferases using artificial cytochrome P450 and globin scaffolds has expanded the repertoire of biocatalysis into carbene transfer chemistry. The reactive intermediate in all the reported artificial carbene transferases is a heme-localised electrophilic metallocarbenoid intermediate. C45 is a tetra-α-helical c-type heme-containing *de novo* protein which functions as a proficient oxidoreductase. The first question raised is whether C45 is capable of forming a metallocarbenoid intermediate, and, if it can, can C45 catalyst carbene transfer reactions with an assortment of suitable substrates, such as olefins (cyclopropanation), X-H α-bonds (X-H insertions), aldehydes (carbonyl olefinations), and nitrogen-containing heteroaromatics (homologous ring expansions)?

C45 was screened for metallocarbenoid formation activity using spectroscopic techniques and mass spectrometry, and its reactivity was quantified using a variety of small molecule analytical techniques. In addition, the inherent simplicity and evolvability of C45 rendered the protein scaffold a perfect blueprint for direct evolution, and novel stereocomplementary tetra-α-helical c-type heme-containing maquettes were developed, biophysically characterised, and enzymatically evaluated. Another significant obstacle hindering the widespread adopting of biocatalytic methodologies in organic synthesis is the limited number of operating conditions hospitable to enzymes. In particular, the presence of organic solvents is often completely detrimental to enzymatic function. The stability of C45 encapsulated inside an organic-solvent resistance hydrogel allowed for the development of a novel heterogeneous biocatalyst that remains catalytically active towards peroxidase and carbene transferase chemistry when suspended in a range of neat organics solvents.

Overall, a *de novo* c-type heme-containing protein proficient at i) forming a number of stable, spectroscopically distinctive bioorganometallic intermediates, ii) catalysing a wide range of abiological chemical reactions *via* a metallocarbenoid intermediate, and iii) performing biocatalytic conversions under conditions typically inhospitable to solvents (i.e. neat organic solvents) is reported.

ACKNOWLEDGEMENTS

I would like to thank Ross Anderson and Annela Seddon for their supervision and support throughout my PhD. I would also like to thank all the members of C101, past and present, who made the lab a wonderful working environment. In particular I would like to thank all the members of the Anderson group: Angelique, Katie, Jonnie, Kris, Bettina, Jack, George, Claire, Paulina (<3), Ben, Dan, and Adrian. I would also like to thank Peter Wilson. Finally, I would like to thank the crow (pictured) for its inalienable support and companionship throughout my PhD.



AUTHOR'S DECLERATION

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programs and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the word is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: …… **DATE:** ……

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

LIST OF ABBREVIATIONS

Abbreviation	Meaning
Å	Ångström (10^{-10} m)
2AP	2-aminophenol
3AP	3-aminophenol
4AP	4-aminophenol
AAP	Aqueous Alginate Phase
ABTS	2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid)
ACN	Acetonitrile
AE-C45	Alginate-encapsulated C45
AE-HRP	Alginate-encapsulated horseradish peroxidase
APX	Ascorbate peroxidase
BAI	Bulk Alginate Interface
bCTM(s)	Mixed b/c-type maquette(s)
BnDA	Benzyl diazoacetate
BOP	Bulk organic phase
Br-EDA	Ethyl 2-bromo-2-diazoacetate
Cb_5	Cytochrome b_5
Cc/Cyto c	Cytochrome *c*
CcP	Cytochrome *c* peroxidase
cCTM(s)	di-c-type maquette(s)
CD	Circular Dichroism Spectroscopy

CpI	Compound I
CpII	Compound II
CP	Coproporphyrin I
CPR	Cytochrome P450 reductase
CSS	Closed-Shell Singlet
CST	Closed-Shell Triplet
CTM(s)	c-type maquette(s)
DBU	1,8-Diazabicyclo[5.4.0]undec-7-ene
DCM	Dichloromethane
ddNTP	Deoxynucleotide triphosphates
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleosidetriphosphates
DE	Directed Evolution
DFT	Density Functional Theory
DMPS	Dimethylphenylsilane
dr	Diastereoselectivity
ε	Molar Extinction Coefficient (M^{-1} cm^{-1})
EDA	Ethyl diazoacetate
ee	Enantioselectivity
EDG	Electron Donating Group
EPR	Electron Paramagnetic Resonance Spectroscopy
erPCR	Error Prone Polymerase Chain Reaction

ESI-MS	Electron Spray Ionisation-Mass Spectrometry
ET	Electron Transfer
EtOH	Ethanol
EtOAc	Ethyl acetate
EWG	Electron Withdrawing Group
FAD	Flavin adenine dinucleotide
Fd	Ferredoxin
FMN	Flavin mononucleotide
Fmoc	Fluorenylmethyloxycarbonyl protecting group
Fmoc-Cl	Fluorenylmethyloxycarbonyl chloride
FR	Ferredoxin reductase
HNMD-NMR	Heteronuclear Multidimensional NMR Spectroscopy
HOMO	Highest Occupied Molecular Orbital
HPLC	High Performance Liquid Chromatography
HRP	Horseradish peroxidase
HSQC NMR	Heteronuclear Single-Quantum Correlation NMR Spectroscopy
IPA	Isopropyl alcohol
IPC(s)	Iron Porphyrin Carbene Complex(es)
LC-MS	Liquid Chromatography-Mass Spectrometry
LOMO	Lowest Unoccupied Molecular Orbital
malE	Maltose/maltodextrin-binding periplasmic protein gene
MALDI(-TOF)	Matrix Assisted Laser Desorption/Ionization (-time of flight)

Mb	Myoglobin
MD	Molecular Dynamics
Me-EDA	Ethyl 2-diazopropanoate
MeOH	Methanol
MP-11	Microperoxidase-11
MPIX	Mesoporphyrin IX
MS	Mass Spectrometry
NAD^+	Nicotinamide adenine dinucleotide (oxidised)
NadC	Nicotinate-nucleotide diphosphorylase
NADH	Nicotinamide adenine dinucleotide (reduced)
NaDN	Nicotinic acid dinucleotide
$NAD(P)^+$	Nicotinamide adenine dinucleotide phosphate (oxidised)
NAD(P)H	Nicotinamide adenine dinucleotide phosphate (reduced)
NaMN	Nicotinic acid mononucleotide
NCA(s)	Nitrogen-containing heteroaromatic(s)
NCH(s)	Nitrogen-containing heterocyclic(s)
NMR	Nuclear Magnetic Resonance Spectroscopy
NO(E)SY	Nuclear Overhauser Effect NMR Spectroscopy
OSS	Open-Shell Singlet
P450	Cytochrome P450
PCR	Polymerase Chain Reaction
Pether	Petroleum ether

PPIX	Protoporphyrin IX
QM/MM	Quantum Mechanics/Molecular Mechanics
RNA	Ribonucleic acid
SRSs	Substrate Recognition Sites
^{t}BuDA	*tert*-butyl diazoacetate
TFE	2,2,2-trifluoroethanol
THF	Tetrahydrofuran
T_m	Denaturation Midpoint Temperature
TPP	Tetraphenylporphyrin
TOF	Total Turnover Frequency/Time-of-flight
t_R	Retention Time (min)
TTN	Total Turnover Number
UV-VIS	Ultraviolet-Visible Spectroscopy
WT	Wild Type

CHAPTER 1: INTRODUCTION TO CARBENE CHEMISTRY

1.1 INTRODUCTION TO CARBENE CHEMISTRY

A carbene is a divalent carbon atom, with six valence electrons and two unpaired electrons, and can be represented by the general formula :CR_1R_2, where R_1 and R_2 represent the substituents.[1] The existence of carbenes was first theorized, in 1855, by Geuther and Hermann who identified dichlorocarbene as a possible reaction intermediate in the hydrolysis of chloroform.[2,3] Dichloromethane was mentioned again, by Nef in 1897, as a possible mechanistic intermediate in the Reimer-Tiemann reaction.[3,4] In 1912, Staudinger isolated cyclopropane-containing products when investigating reactions between alkenes and diazomethane, and methylene, :CH_2, a carbene generated by the elimination of molecular nitrogen from diazomethane, was postulated to be the reactive species.[3,5] Interest in carbenes was rejuvenated in 1953 when Doering and Knox demonstrated their synthetic versatility after photochemically decomposing diazomethane as a key step in the synthesis of substituted tropolones.[6] In 1954, Hofmann and Doering further demonstrated the utility of carbenes by synthesising an assortment of cyclopropane-containing products from alkenes using dichlorocarbene (CCl_2)[7] in addition to providing evidence for the existence of a dibromocarbene species in the reaction between cyclohexene and bromoform under basic conditions.[3,7] In 1959, DeMore attempted to isolate methylene spectroscopically *via* photolysis of diazomethane in a nitrogen matrix at 20 K; although the authors reported IR and UV spectra consistent with methylene, it was stated the species itself could not be identified and could only be inferred from the assortment of reaction products obtained from the experiment.[8] In 1964, Fischer isolated and reported a tungsten carbene complex, although the existence of transition metal carbenes had been considered as early as 1925.[9,10]

In 1992, Goddard provided a theoretical framework to predict the ground-state configuration for an assortment of carbenes.[11] The electronic configuration of a carbene, which regulates its stability and reactivity, is a function of the geometry of the X-C-X bond, and the chemical properties of the substituents, X, bonded to the divalent carbon.[11–16] The geometry of the carbene determines the nature of the frontier orbitals. Linear carbenes have two degenerate p orbitals (p_x and p_y) and are characterised by sp hybridisation.[1] Linear carbenes are relatively rare, and the vast majority exist in a bent confirmation[1] where the deviation away from linearity lifts the degeneracy of the system. The p_y orbital remains generally unaffected by the deviation from linearity and is denoted as the p_π orbital. The p_x orbital, however, acquires a small degree of s-like character and is denoted the σ-orbital.[1] Bent carbenes are characterised by sp^2 hybridisation and two non-degenerate orbitals, p_π and σ.[1] The magnitude of the energy gap between the p_π and σ orbitals is a function of the electronic properties

of the substituents neighbouring the carbene carbon. The electronic configuration of the carbene, which has pronounced effects on its properties and reactivity, is determined by the energy gap between the p_π and σ orbitals. [1,3,11] If the energy difference is sufficiently large it becomes energetically favourable to pair the two electrons together in the lower energy σ orbital, resulting in an electronic configuration of σ^2 and a spin multiplicity of one; these systems are called *singlet* carbenes (another possible singlet configuration is p_π^2) (**figure 1**).[1,3,8-13] Alternatively, if the pairing of the two electrons comes at an energetic penalty greater than the difference between the p_π and σ orbitals, one electron occupies the vacant p_π and the other occupies the vacant σ orbital; this results in an electronic configuration of $\sigma^1p_\pi^1$ and a spin multiplicity of three, and hence these systems are called *triplet* carbenes (an alternative configuration $\sigma^1p_\pi^1$ with differing spin states is also possible, resulting in a singlet carbene) (**figure 1**).[1,3,8-13] How the multiplicity manifests as differences in the reactivity of singlet and triplet carbenes will be discussed in detail shortly.

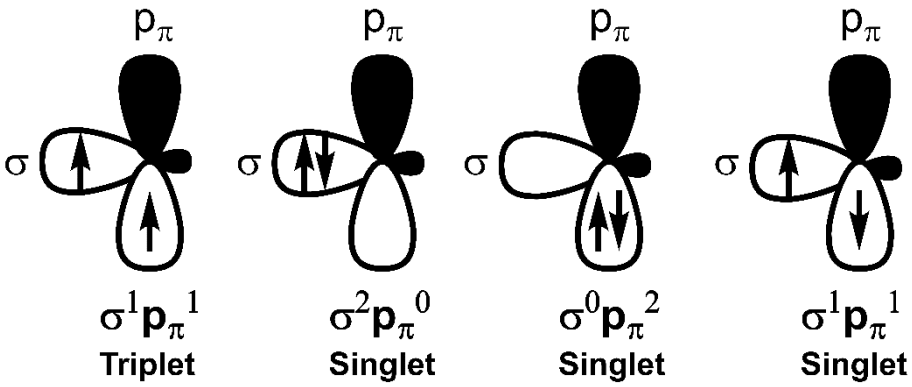


Figure 1: The electronic configuration of singlet and triplet carbenes (adapted from de Frémont *et al*).[3]

1.1.1 Electronic effects of the –R substituents

The nature of the substituents neighbouring the divalent carbon influences the multiplicity of the carbene *via* induction and mesomeric effects.[1,13] For induction, electron-withdrawing groups stabilise the singlet state by withdrawing electron density away from the σ orbital, increasing the s-like character of the σ orbital and lowering its energy while leaving the p_π unaffected. The increase in the σ-p_π energy gap therefore favours a singlet configuration. Electron-donating groups have the inverse effect, by donating electron density into the σ orbital the σ-p_π energy gap is decreased and hence the system favours a triplet configuration.[1,3] This principle is conveniently demonstrated by the differences between CH_2 (triplet) and CF_2 (singlet) carbenes (**figure 2**).[11,14]

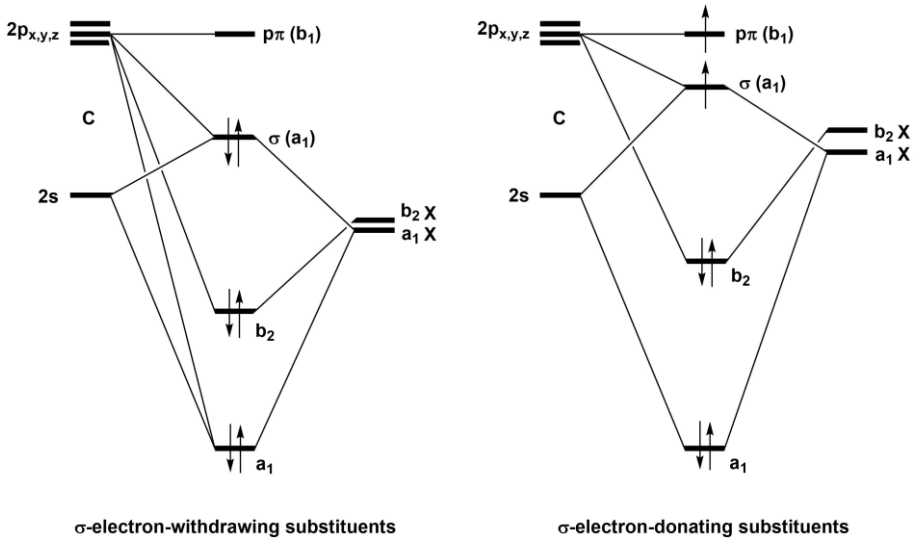


Figure 2: Inductive effects of EWG and EDG substituents on the singlet-triplet energy gap in free carbenes (adapted from Bourissou *et al*).[1]

Although inductive effects play a critical role in determining the ground state multiplicity of a carbene, mesomeric effects are often more significant.[1,12] There are three systems worth considering: CX_2, CY_2 and CXY carbenes were X = π-electron-donating substituents and Y = π-electron-withdrawing substituents (**figure 3**). In CX_2 carbenes, a lone pair of electrons on the substituent interact with the p_π orbital while leaving the σ largely unchanged. The increased electron density in the p_π orbital results in an elevated σ-p_π energy gap and hence CX_2 carbenes are normally bent singlet carbenes. The donation of electron density into the p_π orbital also results in the C-X bonds acquiring a degree of π-character and can cause a build-up of negative charge on the central carbon. CX_2 carbenes are common when X = NR_2, OR, F, Cl, Br, I, PR_2 or $SR_{(1\text{-}3)}$.[1,3,12,14,15] In CY_2 carbenes the dominant interaction is between vacant orbitals on the substituent and the carbene p_y orbital. The p_x orbital is unaffected and the removal of electron density from the p_y orbital lifts the degeneracy of the system resulting in a p_y-p_x energy gap. The p_y-p_x energy gap means CY_2 carbenes are usually linear species with a p_y^2 singlet configuration. Analogous to CX_2 systems, the interaction between the vacant orbitals of the substituent and the carbene p_y orbital imparts a degree of π-character into the C-Y bonds, but a positive charge develops on the carbene carbon on account of the electron-withdrawing nature of the substituents. CY_2 carbenes are common when Y = COR, CN, CF_3, BR_2 and SiR_3.[1,11] Lastly, in the case of mixed carbenes, CXY, the X substituent interacts with the p_y orbital while the Y substituent interacts with the p_x orbital. The net result is a destabilisation of the p_y orbital and a stabilisation of the p_x orbital, which results in a singlet configuration being favoured with π-like character associated with both the C-X and C-Y bond, with negative charge localised on the Y substituent and positive charge localised on the X substituent (**figure 3)**.[1,3,8-13]

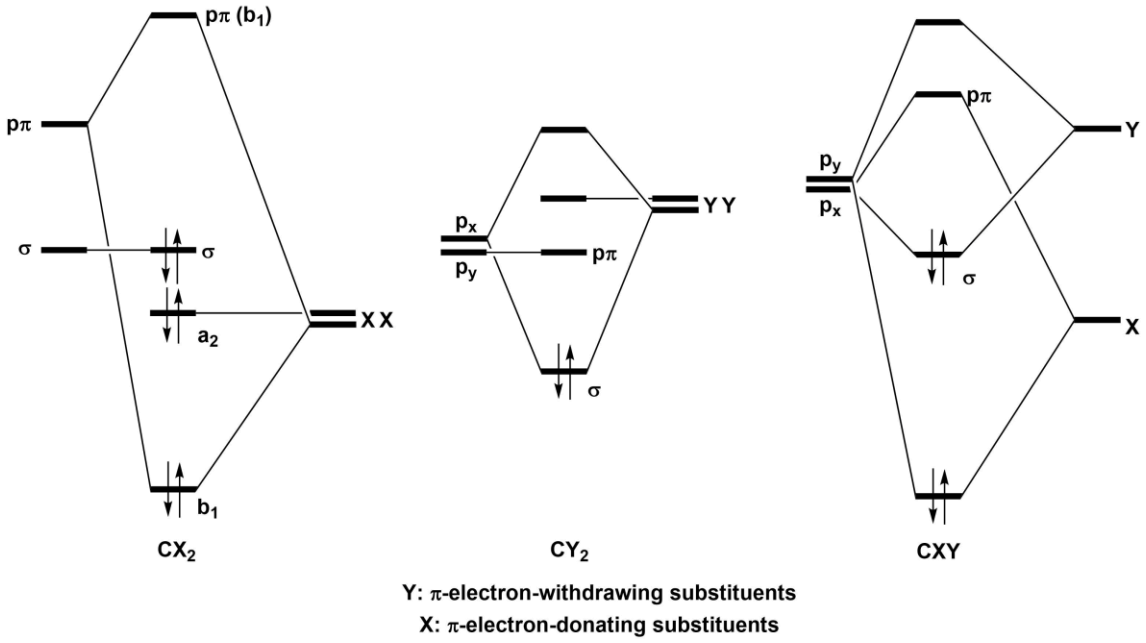


Figure 3: Mesomeric effects of EDG (X) and EWG (Y) substituents on the singlet-triplet energy gap in free carbenes (adapted from Bourissou *et al*).[1]

1.2 TRANSITION METAL CARBENE COMPLEXES

The majority of free carbenes are too reactive to be conveniently isolated (unless sterically bulky substituents are employed to kinetically stabilise the carbene). A common method of stabilising (and hence isolating) a carbene species is *via* complexation with a transition metal. The d-orbitals of a transition metal can provide stabilising interactions with the carbene's p_π and σ orbitals to form a stable transition metal carbene complex with the carbene functioning as a ligand.[1,3,18–22] The additional ligands in the coordination sphere of the metal determine the nature of the metal-carbene bond and reactivity of the carbene ligand. The two major types of transition metal carbene complexes are the Fischer complexes and the Schrock complexes, which are distinguished by the oxidation state of the metal, oxidation state of the carbene ligand, and the properties of the metal-carbene bond.[3,21,23,24] The distinctions between the two types of carbene complexes will be outlined below, followed by a brief account of ''intermediate'' transition metal carbene complexes known as radical carbenes. *N*-heterocyclic carbenes have been omitted from the discussion on account of irrelevancy.

1.2.1 Fischer Carbenes

Fischer carbene complexes are transition metal-carbene complexes formed by low oxidation state middle-to-late transition metals (i.e. Fe(0), Mn(0), Ni(0)). The other ligands coordinated to the transition metal are often strong-field π-acceptor ligands (i.e. CO, NO, CN, PR_3). A distinctive characteristic of Fischer carbenes is one of the carbene's substituents is a π-donating ligand (i.e. NR_2,

$^{-}$NR, $^{-}$OR). The π-donating interaction from the substituent lone pair into the vacant p-orbital of the carbene provides additional stabilisation to the Fischer complex.[3,21,23,24] A few examples of Fischer complexes include $(CO)_5Cr=C(NR_2)Me$, $(CO)_5Cr=C(OR)Ph$, $(CO)_5W=C(OR)Me$, and $(CO)_4(PPh_3)Mo=C(OEt)Ph$ (**figure 4**).

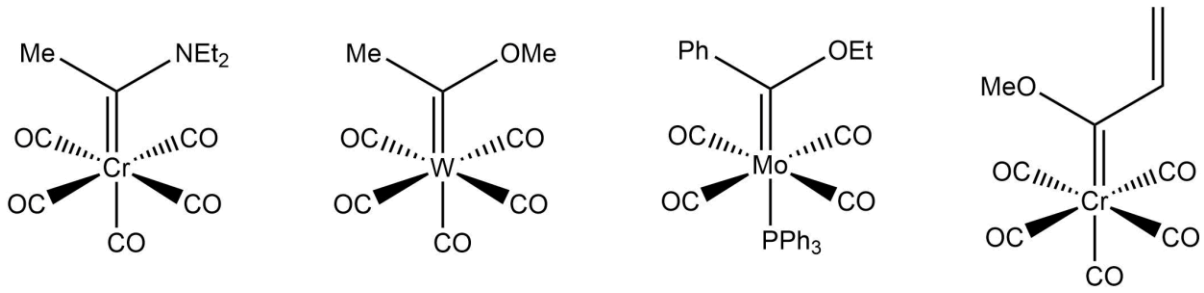


Figure 4: The molecular structures of some typical Fischer carbene complexes.

The π-donating interaction between the carbene and a heteroatom stabilises the carbene, with the increased electron density in the p-orbital elevating the singlet-triplet barrier in favour of a singlet complex. Fischer carbene complexes often adopt a $\sigma^2p_\pi^0$ singlet electronic configuration, and this is reflected in the nature of the M=C bond. The M=C bond in Fischer complexes are characterised by a strong σ-donation from the lone pair situated in the filled carbene σ-orbital into a vacant metal d-orbital and a reciprocal π-back bonding from a filled d-orbital into the vacant p-orbital on the carbene. The M=C bond in Fischer complexes is shorter than a typical M-C bonds but longer than a typical M=C bond. In addition, the M=C bond length is inversely related to the strength of C-X interaction between the carbene carbon and the alpha-heteroatom (X), resulting in an elongation (and hence weakening) of the M=C bond as the C-X bond strength increases. The overall net result of the metal-carbene bond is a flow of electron density onto the metal from the carbene carbon, resulting in an electron-rich, nucleophilic metal centre and an electron deficient, electrophilic carbene centre. The electrophilic nature of the carbene imparts reactivity analogous to a ketone carbonyl, and therefore Fischer carbenes can function as intermediates and/or catalysts in nucleophilic substitutions, aldol reactions, Diels-Alder reactions, and Michael additions. The most noteworthy reaction involving Fischer carbenes is the Wulff-Dötz benzannulation, where aromatic/vinylic alkoxy pentacarbonyl chromium carbene complexes (**figure 4, far right**) are employed in the synthesis of $Cr(CO)_3$-coordinated substituted phenols from a terminal alkyne and CO.[25,26]

1.2.2 Schrock Carbenes

Schrock complexes are transition metal-carbene complexes formed by high oxidation state early transition metals (i.e. Ti(V), Zr(V), Ta(V)). The other ligands coordinated to the transition metal are

typically π-donor ligands (i.e. NR_2, ^-NR, ^-OR,).[3,21,24] The substituents on the carbene are hydrogen and/or alkyl groups, and, therefore, unlike in Fischer complexes, Schrock carbenes lack the stabilising π-donating interaction provided by a neighbouring heteroatom. The lack of electron donation into the vacant p-orbital means the σ-p_π energy gap is smaller relative to Fischer carbenes, and therefore Schrock carbenes often adopt $\sigma^1p_\pi^1$ electronic configurations which is reflected in the nature of the M=C bond. The M=C bond in Schrock complexes is characterised by a σ-bond resulting from the coupling between an unpaired electron situated in the carbene σ-orbital and an unpaired electron situated in the metal d-orbital. This interaction is complemented by a reciprocal π-bond formed by the coupling between an unpaired electron in the d-orbital of the metal with an unpaired electron in the p-orbital on the carbene. A few examples of Shrock complexes include $Cl_2CpTa=C(H)^tBu$, $Cp_2Ti=CMe_2$, $Cp_2HTa=CH_2$, and $Cl_2(P(Cy)_3)_2Ru=C(H)Ph$. The M=C bond in Schrock complexes is shorter and stronger than the M=C bond in Fischer complexes on account of the triplet coupling between the carbene and the metal. The overall net result of the metal-carbene bond is a flow of electron density from the metal to the carbene carbon, resulting in an electron-deficient metal centre and an electron-rich carbene centre. The polarisation of the M=C bond towards the carbon renders the carbene a nucleophilic species, and Schrock complexes have been employed as synthetic reagents in carbonyl olefinations, alkene metathesis reactions, ring opening metathesis polymerisations (ROMP) and ring closing metathesis (RCM) reactions. One Schrock carbene of interest is the dimetal-containing organometallic compound Tebbe's reagent (**figure 5**).[27] Tebbe's reagent is a titanium tetrahedral complex coordinated to two cyclopentadienyl $(C_5H_5)^-$ ligands and bridged with a dimethyl aluminium fragment *via* a chlorine ligand and a methylene (carbene) ligand, to yield a complex with a molecular formula of $(C_5H_5)_2TiCH_2ClAl(CH_3)_2$. Tebbe's reagent is employed as a catalyst in carbonyl methylenation reactions of aldehydes, esters, lactones, and amides.[27,28] A chlorinated analogue of Tebbe's reagent has been successfully employed to convert carbonyls into cyclopropanes, and the reaction mechanism is hypothesised to proceed *via* a $TiCl_2(=CH_2)$ carbene intermediate.[28]

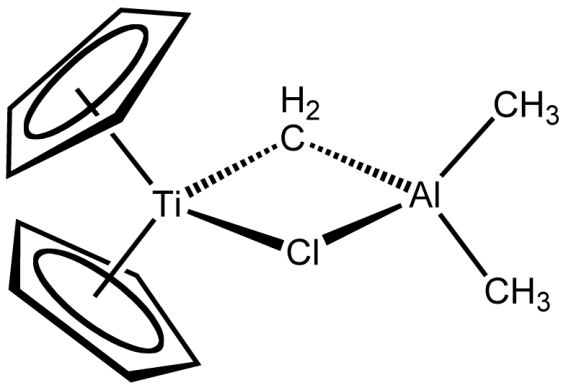


Figure 5: Tebbe's reagent, an example of a Shrock carbene.

1.2.3 Radical Carbenes

Radical carbenes are a unique class of organometallic carbenes formed when a carbene ligand is coordinated to an open-shell metal. Radical carbenes can be prepared by the one-electron reduction of a Fischer complex or by the direct coordination of a carbene to an open-shell metal.[29] In radical carbenes, the carbene carbon have radical-like properties and exhibit features intermediate between Shrock and Fischer carbenes. Radical-carbene complexes are common when a carbene is coordinated to a low-spin group 9 transition metal. In particular radical-carbene complexes with cobalt (Co^{3+})[29] are very stable and generate long-lived intermediates.[26–30] Radical carbene complexes with other metals, such as ferrous iron (Fe^{2+}), have also been reported,[34] although these complexes are shorter-lived and less stable relative to the analogous Co^{3+} complexes. The mechanism behind radical carbene formation is analogous to Fischer carbene formation, with the formation of a σ-bond between the carbene carbon and the metal accompanied by π back-bonding from the metal to the carbene π* orbital.[29,35] The nature of the M=C bond imparts significant electron density to the carbene carbon, predominantly localised in the p_z. The complex is further stabilised by a back bonding interaction from the carbene p_z orbital into an adjacent substituent's π*-orbital; this interaction is critical for stabilising the overall complex.[29,35] Tetrapyrrole ligands, especially porphyrin-derived macrocycles, are renowned for their ability to stabilise a carbene radical complex by functioning as an ''electron buffer''. The porphyrin macrocycle donates electron density into the π d-orbitals of the metal, stabilising the metal and localising the radical-like electron density on the carbene carbon.[29,35]

1.3 METALLOCARBENOIDS

Carbenoids is the name assigned to a class of compounds that possess a complete valency but exhibit properties similar to free carbenes.[1,18,20] The most common precursors for carbenoids are diazo compounds (C-N_2).[36,37] The reactivity of diazo compounds is determined by the nature of the substituents neighbouring the carbenoid carbon. In the presence of a suitable metal (Cu, Rh, Pt), a diazo compound will decompose to release a molecule of N_2 and form a metal-carbon double bond.[37] The bonding to the metal stabilises the carbenoid and allows the reactivity to be controlled. The bonding consists of a strong σ-bond between the metal and carbon formed by the donation of a lone pair of electrons localised on the carbon to the metal.[38–41] The metal donates electron density from a d-orbital into the carbon p-orbital, affording a stable bond with a degree of π-character. The metallocarbenoid is an electrophilic species as a consequence of the strong σ-bond, with the electrophilicity controlled by the neighbouring substituents of the carbenoid carbon and the strength of the π-donation from the metal. [38–41] Metallocarbenoids have reactivity similar to singlet carbenes, and the nature of the C=M bond allows the selectivity of metallocarbenoid reactions to be controlled.

A notable metallocarbenoid is the organozinc carbenoid $IZnCH_2I$ (generated from CH_2I_2) which is employed in the stereospecific cyclopropanation of alkenes in a reaction called the Simmons-Smith reaction (discussed below in section 2.2.4).[42] Many transition metal complexes of carbenes can be classified as metallocarbenoids.

1.4 REACTIONS OF FREE CARBENES AND METALLOCARBENOIDS

1.4.1 [2+1] cycloaddition reactions

The most important reactions involving carbenes and metallocarbenoids are the [2+1] cycloadditions.[37,42,43] [2+1] cycloadditions are cheletropic, with both new σ-bonds being formed on the carbene(oid) carbon to give a three-carbon ring. The substrates are most commonly alkenes which yield cyclopropanes products (cyclopropanation), but alkynes can also be employed to yield cyclopropene products (cyclopropenation).[37,42,43] Although the majority of cyclopropanation reactions involve a transition-metal stabilised carbenoid, free carbenes can also be employed as reagents in cycloaddition reactions, although the variable electronic configurations of free carbenes has important implications on the reaction stereoselectivity.

1.4.1.1 Carbene multiplicity and cyclopropanation stereochemistry

Singlet carbenes, characterised by a σ^2 electronic configuration, behave as electrophiles, and react with alkenes in a concerted single step cheletropic reaction which proceeds *via* a four-electron-three-centred transition state. The bonding in the transition-state involves the donation of the σ^2 electrons into the vacant π* orbital of the alkene with a weak back-donation from the alkene p-orbitals into the vacant carbene p_z orbital. The diastereoselectivity of the alkene substrate is preserved in the cyclopropane-product because both new σ-bonds form simultaneously; singlet carbene [2+1] cycloaddition reactions are therefore stereospecific (**figure 6**).

Triplet carbenes, characterised by a $\sigma^1p_\pi^1$ electronic configuration, behave as diradicals, and react with alkenes in a stepwise radical addition reaction which proceeds *via* a diradical intermediate. The first step involves a reaction between the carbene σ electron and the alkene π-bond to produce a diradical species with the two lone electrons separated by a tetrahedral bridge. The second step involves coupling of the two lone electrons to form a new σ-bond and generate the final cyclopropane product. The diastereoselectivity of the alkene substrate is not preserved in the cyclopropane-product because the two new σ-bonds form separately; triplet carbene [2+1] cycloaddition reactions are therefore not stereospecific but can stereoselective. This is a consequence of the diradical intermediate, which

allows for free rotation around the newly formed tetrahedral carbon and the two neighbouring carbon radicals (**figure 6**).

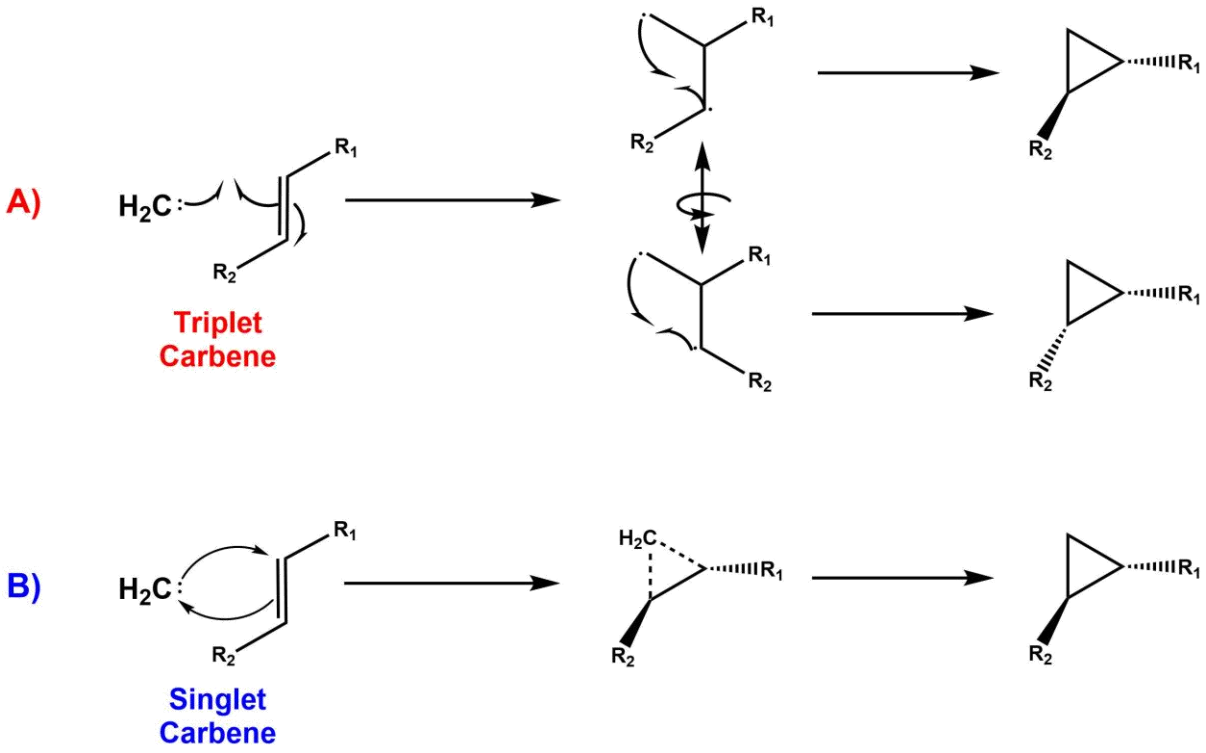


Figure 6: The selectivity of triplet (A) vs singlet (B) free carbene cyclopropanation. Triplet carbenes react with olefins in a step-wise mechanism, proceeding *via* a diradical intermediate which affords free rotation around the newly formed sp^3-sp^3 C-C bond; the reaction is not stereospecific. Singlet carbenes react with olefins in a concerted manner and are stereospecific.

1.4.1.2 Cyclopropanation using free carbenes

The [2+1] cycloaddition of a carbene across an alkene C=C double bond can be readily achieved by generating a reactive carbene intermediate *in situ via* thermolysis, photolysis, dihalogen reduction, or base hydrolysis.[3] A major limitation with free carbene reagents is the lack of chemo-, regio- and stereo-control because of the highly reactive nature of free carbenes and the singlet/triplet multiplicity. Nonetheless, several important cyclopropanation reactions proceed *via* the generation and subsequent consumption of a reactive carbene in the presence of olefins. Treating chloroform with alkoxides (i.e. $(^tBu)_3ONa$, EtONa, ^tBuOK, NaOH etc) generates the common carbene dichlorocarbene (Cl_2C:).[7] Cl_2C**:** is a bent singlet carbene which reacts with alkenes to yield geminal dichlorocyclopropanes. Analogous geminal dibromocyclopropanes can be readily obtained by treating $CHBr_3$ with an alkoxide to generate dibromocarbene (Br_2C**:**) in the presence of a suitable olefin. The cyclopropantion of alkenes with dihalocarbenes is a stereospecific process because the singlet multiplicity of dihalocarbenes means the reaction proceeds *via* a concerted, cheletropic mechanism.

The resulting geminal dihalocyclopropanes can be subsequently treated with an organolithium base in a reductive dehalogenation reaction to yield a cyclopropane-carbene intermediate. The cyclopropane-carbene intermediate is incredibly unstable, and spontaneously rearranges to yield an allene product. This reaction is called the Skattebøl reaction (**figure 7**).[44] An analogous reaction, the Doering–LaFlamme allene synthesis, also involves the *in-situ* generation of a dihalocarbene which reacts (with an olefin) to form a geminal dihalocyclopropane. The dihalocyclopropane is subsequently reduced in the presence of magnesium or sodium metal to generate the cyclopropane-carbene intermediate which spontaneously rearranges to the final allene product (**figure 7**).[45] Another modification of the Skattebøl/Doering–LaFlamme allene synthesis involves the reductive dehalogenation of a 1,1-dihalo-2-vinylcyclopropane to yield a 2-vinylcyclopropane carbene, which spontaneously rearranges to yield a cyclopentadiene product; this modified reaction is known as the vinylcyclopropane rearrangement.

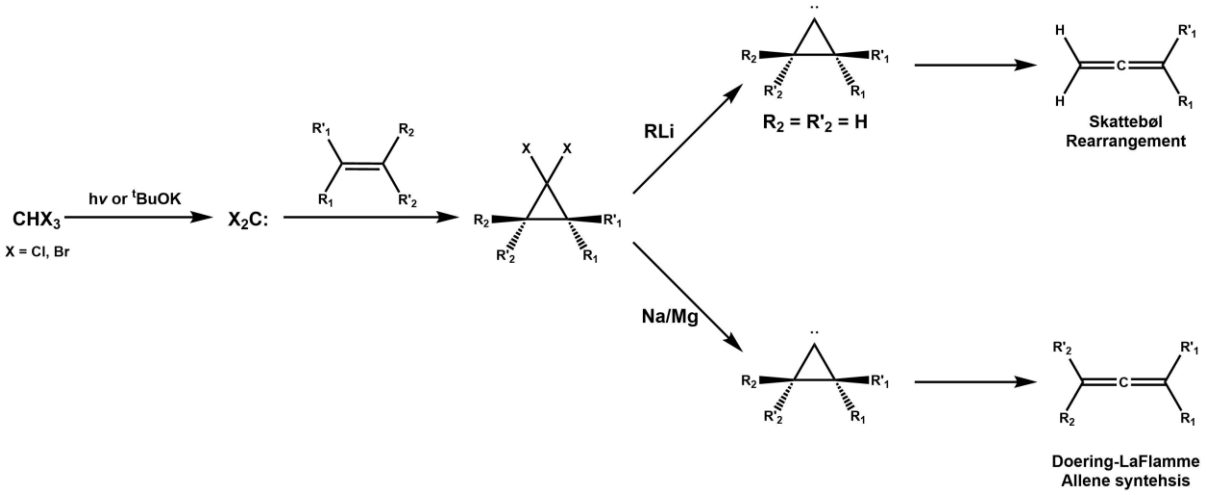


Figure 7: Allene synthesis *via* dihalocyclopropanes: Skattebøl rearrangement (top) and Doering-LaFlamme synthesis (bottom). In both reactions, the first step involves the generation of a dihalocyclopropane moiety, which is reduced in a second step to yield a reactive cyclopropane carbene. The highly reactiove cyclopropane carbon is susceptible to spontaneously rearrangement to yield an allene (C=C=C proudct).

Fluorenylidene is a carbene derived from the photolysis of 9-diazofluorene, with the carbene carbon localised on the bridging methylene connecting the two aryl groups (**figure 8**). The most distinctive property of fluorenylidene is the remarkably low energy gap between the singlet and triplet states (approximately 4.5 KJ mol^{-1}).[46,47] Flash photolysis of 9-diazofluorene results in a high population of singlet fluorenylidene molecules which gradually relax down to triplet fluorenylidene *via* intersystem crossing. The rate of intersystem crossing is dependent on temperature and the physical properties of the solvent, which translates to the observed stereochemistry of cyclopropanation reactions between

fluorenylidene and a suitable olefin. If the rate of intersystem crossing greatly exceeds the rate of singlet fluorenylidene cyclopropanation, the system equilibrates to the triplet state and the cyclopropanation proceeds in a non-stereospecific, stepwise fashion resulting in a mixture of stereoisomers. However, the presence of halogenated solvents was found to decelerate the rate of intersystem crossing. In particular, hexafluorobenzene (C_6F_6) was reported to significantly reduce the rate of intersystem crossing and the reported rate of reaction with a triplet-quenching molecule such as MeOH. The authors conjectured that a halogen lone pair from the solvent interacts with the vacant fluorenylidene orbital and stabilises the singlet state, and subsequent cyclopropanation with fluorenylidene in the presence of a suitable olefin and C_6F_6 resulted in a stereospecific reaction, indicative of a singlet fluorenylidene carbene species.[46–48]

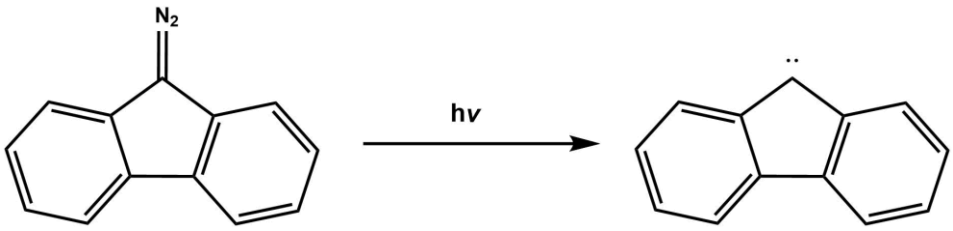


Figure 8: Fluorenylidene synthesis *via* flash photolysis of 9-diazofluorene.

1.4.1.3 Metallocarbenoid-catalysed cyclopropanation

Although free carbenes can be employed as reagents in cyclopropanation, problems associated with their inherent instability, selectivity, and side-reactions (i.e. carbene dimerization) limit their scope and applications. Transition-metal stabilised carbenoids, synthesised from diazo compounds, present a safer and more reliable methodology for intra- and intermolecular cyclopropanation.[42] The stereoselectivity of transition metal catalysed carbene transfer reactions can also be finely tuned by coordinating select chiral ligands to the transition metal.[49–53] Metallocarbenoid catalysed cyclopropanations are well established and have been readily employed as key steps in the enantioselective synthesis of important molecules, such as the pyrethroid precursor chrysanthemic acid.[54]

A very common reaction is the Simmons-Smith cyclopropanation (**figure 9**),[42,55] a cheletropic [2+1] cycloaddition between an alkene and an organozinc metallocarbenoid species: iodomethylzinc iodine ($IZnCH_2I$). Iodomethylzinc iodine is the active reagent in the Simmons-Smith reaction, and is generated *in-situ* by reacting diiodomethane (CH_2I_2) with a zinc-copper alloy, Zn(Cu). In the presence of a suitable olefin, iodomethylzinc iodine transfers a methylene group to the alkene C=C bond *via* a concerted cheletropic mechanism to afford a cyclopropane product.[42,55] The reaction is stereospecific, with the

stereochemistry of the alkene substrate being preserved in the cyclopropane product. The reaction is subject to steric effects, with the methylene adding to the least-hindered face of the alkene.[42,55] The reactivity of the Simmons-Smith reaction can be significantly enhanced by the Furukawa modification, which involves substituting the Zn(Cu) alloy for diethylzinc, Et_2Zn.[56] The Furukawa modification, like the original Simmons-Smith reaction, is also stereospecific and the methylene groups adds across the least hindered face of the alkene. A notable exception occurs for alkenes situated in close proximity to hydroxyl groups; the hydroxyl group coordinates to the zinc and positions the metallocarbenoid in the same plane of the hydroxyl prior to the methylene transfer step, meaning the methylene transfer step proceeds *via* a syn-addition relative to the hydroxy group irrespective of any steric considerations.[42,57]

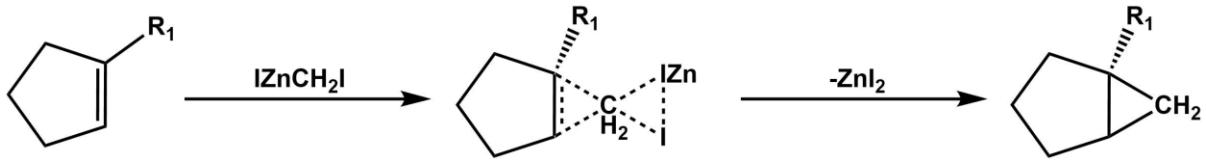


Figure 9: The Simmons-Smith cyclopropanation. The reaction proceeds *via* an organozinc metallocarbenoid intermediate. The reaction is stereospecific.

The Simmons-Smith reaction is one of the most popular reactions for synthesising cyclopropanes, but limitations concerning the expense of the reagents, unwanted side-reactions occurring because of ZnI_2, and the requirement for a stoichiometric quantity of ICH_2ZnI means alternative methodologies have been explored. An alternative cyclopropanation methodology is the transition metal catalysed decomposition of a diazo compound to form a metallocarbenoid complex followed by a subsequent carbene transfer step to a suitable substrate.[42,49–53] The mechanism involves the initial attack of the diazo compound on the metal centre to yield a tetrahedral zwitterion alkyl-metal complex. Rapid expulsion of a N_2 molecule thermodynamically drives the formation of the π-bond between the metal and carbon to yield the metallocarbenoid complex. The metallocarbenoid complex is electrophilic, and reacts with an alkene *via* a concerted, cheletropic mechanism to afford a cyclopropane product.[42,58]

1.4.2 X-H insertion reactions

An insertion reaction occurs when a reagent inserts into a pre-existing σ-bond on another substrate, resulting in the formation of a new σ-bond. Insertion reactions are often limited because of the strength of C-X σ-bonds, but the extraordinarily high reactivity of carbenes renders them useful reagents for insertions into X-C σ-bonds (where X = C, N, S, Si, H).[37,59,60] In the absence of a stabilising metal, carbene insertion reactions tend to be non-regioselective, but the selectivity can be improved *via* coordination of the carbene to a transition metal.

There are two plausible mechanisms for carbene X-H insertion reactions,[39,61] although at present it is unclear which mechanism is accurate. The first is a concerted formation of a tricyclic C-X-H intermediate state which is characterised by the concurrent formation of the C-X σ-bond, breaking of the X-H σ-bond and transfer of the H to form a X-H σ-bond. The second involves the nucleophilic addition of the X-H substrate to the electrophilic metallocarbenoid followed by the subsequent dissociation of the resultant ylide. A 1,2-hydrogen shift from the X-group to the carbene carbon completes the reaction. The two plausible mechanistic pathways are illustrated in **figure 10** for the reaction between aniline and methylene.

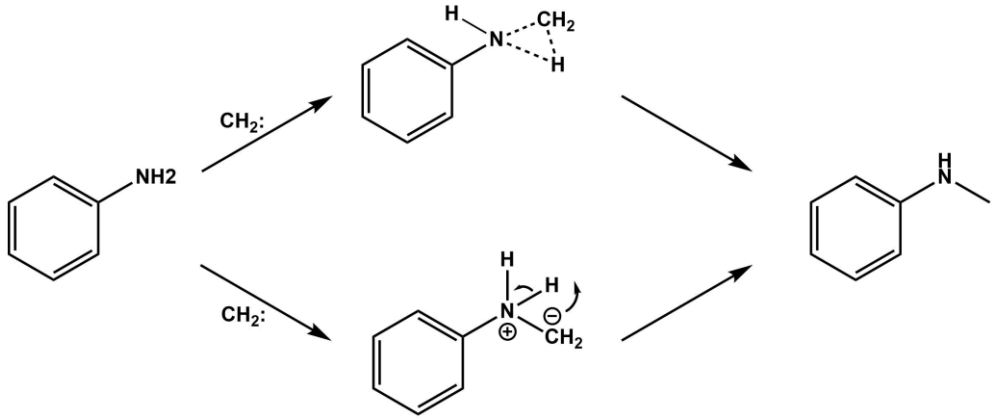


Figure 10: The two possible mechanisms for the carbene N-H insertion reaction between aniline and methylene. The first pathway involves the formation of three-membered cyclic intermediate; the second pathway involves the initial formation of an ylide species followed by the 1,2-hydrogen shift.

1.4.2.1 C-H insertion reactions

The synthesis of C-C σ-bonds is one of the most important reactions in organic chemistry because it affords the possibility of synthesising large and complex molecules from smaller starting materials.[62,63] The problem with such reactions concerns the strong bong energy of C-H bonds, which needs to be compensated for.[62–64] The common methodologies involve an initial C-H activation/substitution step to lower the energetic barrier to C-C σ-bond formation, but this is unfavourable because it adds additional steps to a synthesis.[62,63] Carbenes can form new C-C σ-bonds without the need for C-H activation, which renders them useful reagents.[37,59,64] In the absence of a catalyst, carbene C-H insertion reactions proceed without selectivity and yield a mixture of different possible insertion products. The employment of a suitable transition metal catalyst can circumvent the selectivity issues.

Dirhodium polycarboxylate catalysts are the most commonly employed class of catalysts for this type of chemistry, with $Rh(II)_2(AcO)_4$, which exhibits high promiscuity towards regioselective intermolecular and regiospecific C-H insertion of carbenes into aliphatic and aromatic C-H bonds, being the most

utilised.[65–67] An assortment of dirhodium complexes with bulky chiral ligands have been reported for asymmetric intermolecular and intramolecular C-H carbene insertion reactions.[66]

1.4.2.2 N-H insertion reactions

In addition to C-H insertion reactions, carbene intermediates have been utilised in the synthesis of C-X σ-bonds (X = O, N, S, Si).[37] Amine insertions, the insertion of a carbene into a N-H σ-bond to form a novel C-N σ-bond,[65] is a valuable synthetic reaction because it affords the ability to synthesise complex important molecules under mild reaction conditions. The N-H insertion reaction is catalysed by a variety of transition metals, including Cu, Pd, Ru, which affords a degree of chemo-, regio- and stereoselectivity to the reactions.[65] For example, a Cu(I) derived catalyst afforded the chemoselectivity insertion of reactive carbenes into N-H in the presence of O-H groups. In another study, an iridium-derived catalyst demonstrated a remarkable catalytic turnover towards the mono-insertion of a carbene for aromatic-NH_2 functional group containing molecules in water.[68]

The importance of N-H insertion reactions in forming novel C-N σ-bonds is of special interest in amino acid chemistry and peptide synthesis.[69] The current synthetic methods employed in peptide synthesis involves multiple protection/deprotection steps, adding additional reactions and purifications to the overall synthesis which can negatively affect reaction yields. The chemo-, regio- and stereoselectivity of transition metal catalysed carbene insertion reactions offers a new and easier methodology in the controlled synthesis of peptides and non-natural amino acid derivatives.

1.4.3 Homologation reactions

The simplest free carbene is the compound methylene (H_2C**:**) generated *in situ* by the photolysis of diazomethane (CH_2N_2).[6,37] A colourless gas detectable at low temperatures and pressures, methylene is a highly reactive species only isolatable in solution an adduct or coordinated to a suitable metal. In organic chemistry, methylene is synthesised and immediately consumed *in situ* as a reactive intermediate. A good example of methylene in synthesis is the Arndt–Eistert reaction,[59,70] the one carbon homologation reaction of carboxylic acids to yield the homologue carboxylic acid. In the reaction, the nucleophilic methylene, generated by photolysis of diazomethane, attacks an acyl chloride to yield a tetrahedral intermediate. The carbonyl group reforms, expelling HCl, and yields a α-diazo carbonyl; the subsequent rearrangement of the α-diazo carbonyl (*via* a Wolff rearrangement) expels N_2 and results in a ketene species, which is subsequently hydrolysed to yield the homologation product.[59,70] Diazomethane is also a crucial reagent in the synthesis of substituted tropolones, in which the photolysis of diazomethane generates the reactive methylene species which subsequently reacts

with substituted benzene to yield a substituted 7-membered cycloheptatriene ring, which is further oxidised to the tropolone product.[6]

1.4.4 Carbene dimerization reactions

The reactions discussed above involve a carbene or carbenoid reacting with a non-carbene(oid) substrate in a biomolecular reaction. Another organic reaction observed is a carbene dimerization, a biomolecular reaction where a carbene(oid) reacts with another carbene(oid) to yield a homodimer (when the two carbene(oids) are identical) or a heterodimer (when the two carbene(oids) are not identical) alkene product. Although carbene dimerization reactions are usually undesirable side-reaction (carbene dimerization reactions are particularly problematic when α-carbonyl diazo compounds are employed as carbene precursors), carbene dimerizations have attracted some attention in the synthesis of symmetrical alkene products. A good example is the dimerization of ethyl diazoacetate (EDA) to the symmetrical cis-alkene ethyl maleate which is catalysed by the organoruthenium complex $RuCl(PPh_3)_2(C_5H_5)$.[71]

1.4.4.1 The Wanzlick Equilibrium

The most famous and well documented carbene dimerization reaction is the Wanzlick dimerization,[72–77] a reaction characterised by a chemical equilibrium between a stable carbene and its dimer product (**figure 11**). In the 1960s, Wanzlick and Schikora studied the vacuum pyrolysis of 2-trichloromethyl dihydroimidazole substrates and hypothesised the elevated temperatures encouraged the thermal decomposition of 2-trichloromethyl dihydroimidazole to afford dihydroimidazol-2-ylidene substrates *via* the elimination of chloroform (**figure 11**).[78,79] It was proposed that dihydroimidazol-2-ylidene derivatives would exist in an unfavourable equilibrium with the corresponding dimer.[72] Investigations using bis-(1,3,-diphenyl-2-imidazolidinylidene), generated from the elimination of chloroform, resulted in a colourless crystal which exhibited a strong band in the Raman spectrum at 1640 cm^{-1} (C=C bond).[72] The thermal decomposition of bis-(1,3,-diphenyl-2-imidazolidinylidene) in the presence of tetracyanoethylene resulted in the isolation of a cyclopropane product. The authors demonstrated the dimerization product was chemically inactive towards a selection of electrophiles, and postulated that bis-(1,3,-diphenyl-2-imidazolidinylidene), which reacted with the selected electrophiles, was the reactive nucleophilic species.[72]

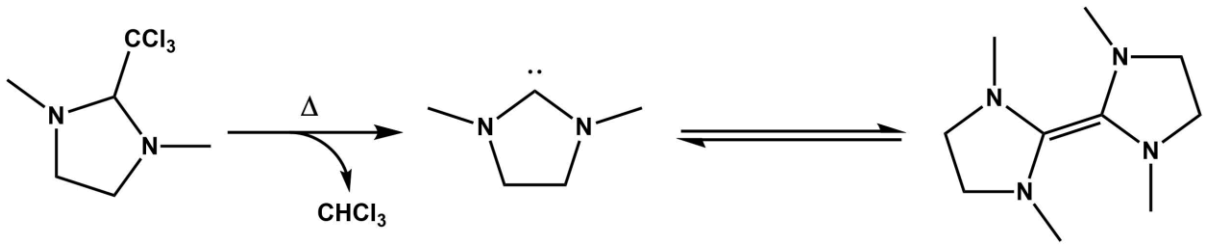


Figure 11: An illustration of the Wanzlick equilibrium. Flash photolysis yields a carbene intermediate which exists in equilibrium with its dimer.

The carbene-dimer equilibrium conjecture presented by Wanzlick and Schikora was challenged by a series of experiments conducted in the mid-60s. In 1964, Lemal and co-workers investigated the dissociation of two tetraaminoethylenes substrates, the *N*-Ph (1) and the *N*-*p*-tolyl derivative (2).[80] It was claimed that if Wanzlick's hypothesis was correct a cross over-species between the two tetraaminoethylenes derivatives would be expected. After refluxing equimolar quantities of substrates (1) and (2), under anaerobic conditions, the product mixture was collected, purified, and analysed by NMR spectroscopy. The authors reported that ''*no absorption attributable to the crossed ethylene was present*'' in the NMR spectrum and concluded that the carbene-dimer equilibrium proposed by Wanzlick was incorrect; a new reaction mechanism for the reactivity observed by Wanzlick was subsequently proposed.[80] In 1965, the Wanzlick equilibrium was also criticised by Winberg and co-workers, after reporting that no cross-over products could be detected by gas chromatography after refluxing equimolar volumes of tetrakis(dialkylamino)ethylene derivatives, corroborating the arguments of Lemal and co-workers.[81]

The Wanzlick carbene-dimer equilibrium was revisited in 1999 by Denk and co-workers who reported the synthesis and isolation of metal carbene complexes from electron-rich tetraaminoethylenes precursors.[75] Denk and co-workers reported a statistical mixture of A=A, A=B and B=B products in the ration 1:2:1 in the free metathesis reaction of all tetraaminoethylenes explored; the only mixture were no mixed product (A=B) could be detected was the mixture employed by Lemal in 1964. The authors claimed the experiments supported the existence of the Wanzlick equilibrium but commented that the results did not irrefutably prove it; the authors highlighted that a [2+2]-cycloaddition/[2+2] cycloreversion could account for the reported statistical mixture of alkene products.[75] In 2000, Hahn and co-workers provided further evidence for the existence of the Wanzlick equilibrium after detecting a free *N*-heterocyclic carbene, *via* NMR spectroscopy, in solution (prepared from its corresponding pure entetramine) after 24 hours.[76] The simultaneous detection of the free carbene and the dimer also provided convincing evidence for the dissociation mechanism proposed by Wanzlick and Denk respectively, as opposed to the metathesis route. This conclusion was further

supported by Lemal after observing the dissociation of methyl-/ethyl-functionalised entetramine *via* NMR spectroscopy.[73,82]

1.5 CONCLUSION

Carbenes are highly reactive and electron deficient species that can be employed as useful intermediates in a number of important reactions, including [2+1] cycloadditions, X-H insertions, homologation reactions and dimerization reactions. The stability and reactivity of a carbene species is a function of the electronic and spatial properties of the substituents bonded to the carbene carbon. A few classes of carbene species can be established from the disinvite differences in the electronic nature and reactivity. For example, singlet carbenes react in a concerted fashion whereas triplet carbenes react in a two-step stepwise fashion. The differences in reactivity can also afford control over the selectivity in syntheses employing carbenes. A common methodology for stabilising a carbene intermediate is by decomposing a carbene intermediate, such as a diazo compound, in the presence of a suitable transition metal to form a metallocarbenoid complex. The nature of the substituents coordinated to the transition metal can be finely tuned to generate three classes of metallocarbenoid complexes: Fischer, Shrock and radical complexes. The different natures of the complexes result in different reactivity with a variety of substituents. Once a metallocarbenoid complex has been prepared, the complex is then free to react with a variety of functional groups to construct larger complex organic molecules, typically with a reasonable degree of chemo-, regio- and stereo-selectivity.

CHAPTER 2: A HISTORY OF CARBENE TRANSFERASES: FROM RATIONALLY DESIGNED ARTIFICIAL METALLOENZYMES TO *DE NOVO* MAQUETTES.

2.1 A BRIEF HISTORY OF THE MAJOR DEVELOPMENTS IN MOLECULAR BIOLOGY

Biocatalysis is the use of biological material, such as an enzyme or whole cell, to catalyse a chemical transformation.[83–85] The ability to isolate, clone and express a specific protein has rejuvenated interest in biocatalysis. The non-toxicity, biodegradability, remarkable selectivity, and specificity of enzymes underlines them as promising alternatives to the hazardous organometallic catalysts ubiquitously used in synthesis. However, three historical obstacles continue to hinder the widespread adoption of enzymes: i) the difficulty in isolating a sufficient quantity of a desired enzyme, ii) the limited number of chemical transformations enzymes can perform, and iii) the limited operating conditions of enzymes. The revolutionary developments in molecular biology and biotechnology have allowed these limitations to be addressed. Three of the most important developments are i) DNA sequencing technology, ii) the polymerase chain reaction (PCR), and iii) DNA/protein recombinant technology. Each of these developments shall be briefly discussed below.

A major breakthrough came from the development of DNA sequencing technology. The first method introduced for DNA sequencing, by Wu, involved the binding of location-specific primers to a DNA sequence followed by an extension reaction catalysed by DNA polymerase.[86] Wu and colleagues used this methodology to sequence the lysozyme gene of bacteriophage T4.[86] In 1977, Sanger adopted the principle of the primer-extension strategy to develop faster DNA sequencing techniques.[87] In Sanger sequencing, the four necessary canonical bases deoxynucleosidetriphosphates (dNTPs), a DNA primer, a single-stranded DNA template, and DNA polymerase are present, in addition to one of four modified dNTPs called deoxynucleotide triphosphates (ddNTPs).[87] The importance of the ddNTPs in the Sanger sequencing is a due to the absence of a 3'-hydroxyl group which is required for the formation of the phosphodiester bond between two nucleotides which form the backbone of a DNA strand. The random incorporation of ddNTPs into a sequence of DNA inhibits further elongation of the strand and terminates the reaction. The random nature of ddNTP addition means after the reaction the mixture will contain an ensemble of different-length single-stranded DNA strands, each containing a terminal ddNTP. The ddNTPs can be radioactively or fluorescently tagged for detection and, thus, repeating the experiment in the presence of all four ddNTPs separately allows the sequence of a DNA strand to be accurately determined by corresponding the positions of the terminal ddNTPs to the original sequence.[87] The first complete genome of an organism, of the bacterium *Haemophilus influenzae*, was published in 1995.[88]

Another crucial development was the polymerase chain reaction (PCR). In PCR, an infinitesimal quantity of a specific DNA sequence is amplified by repeated cycles of thermally denaturation of the double-stranded DNA, the annealing of specifically designed primers to the single strands of DNA , and the subsequent replication of the gene by DNA polymerases. The key component behind the development of PCR was the isolation of *taq* DNA polymerase from the thermophilic bacterium *Thermus aquaticus*.[89] The protein *taq* DNA polymerase is thermally-stable (up to 80 ^{O}C) and can tolerate the higher temperatures required for denaturing double-stranded DNA without any loss of functionality or specificity. Through repeated cycles of denaturation, annealing and replication, PCR allows a sufficient quantity of a desired gene to be acquired.

The techniques of molecular cloning and the introduction of recombinant DNA and protein technology was the final development.[90] Molecular cloning introduced a facile way of acquiring large quantities of a desired biomolecule, such as a protein. Molecular cloning involves the introduction of laboratory prepared recombinant DNA into a host organism and directing its expression. A large variety of host organisms and cloning vectors have been developed, but the majority of applications involve *E. coli* as the host organism and a plasmid expression vector as the cloning vector.[90] A plasmid is a small, circular fragment of extrachromosomal DNA, commonly found in bacteria cells, that is capable of autonomous replication within its host. Plasmids can be introduced into competent cells *via* transformation and can be passed on to future generations during cell division. These biochemical properties highlighted plasmids as perfect tools for molecular cloning. A plasmid cloning vector must contain several important genes in order to function adequately: i) an origin of replication, ii) a minimum of one selective marker, and iii) a minimum of one unique restriction endonuclease recognition site. The origin of replication allows for the plasmid to be expressed inside the host organisms. The selective marker is a gene that encodes for antibiotic resistance towards a specific antibiotic (e.g. ampicillin, tetracycline, etc) and ensures only cells possessing the plasmid will survive in the presence of the specific antibiotic. The restriction endonuclease recognition sites allow for endonucleases to cut the plasmid DNA and so a gene encoding the desired gene product (i.e. protein) can be inserted into the plasmid prior to transformation. The inserted DNA is appended to the plasmid *via* DNA ligases. In enzymology, the recombinant DNA contains a gene encoding for a desired enzyme, which can be over-expressed once transformed into the host organisms.[90]

The three techniques above allow sufficient quantities of a desired protein to be collected and also provide an easy methodology for introducing modifications into a sequence of DNA, which is translated into modifications to a gene target's (protein) structure and function. This has afforded the ability to start addressing the aforementioned limitations of biocatalysis by redesigning enzymes

towards novel functionality and chemistry. The design of novel proteins is sometimes called *rational design*.[83,91–93] Rational design is divided into two complementary approaches: *protein redesign* and *de novo* design.[94] Protein redesign is a top-down approach in which a natural protein is selected as the blueprint from which new proteins can be engineered. *De novo* design is a bottom-up approach in which non-naturally occurring proteins are designed from scratch.[95,96] The application of rational design has afforded valuable insights into complex phenomena such as protein folding, protein structure, protein dynamics, and mechanistic enzymology. In addition, rational design has resulted in an explosion of new functional enzymes with abiological activity. A testament to the success of *rational protein design* is the recent work achieved in oxidoreductase (re)design.[94,97–99] The oxidoreductase class is an enormous group of ubiquitously distributed enzymes that are responsible for catalysing redox biochemistry.[100] The biological importance of oxidoreductases makes them attractive targets for protein (re)design. One major class of enzymes in the oxidoreductase class are the hemoproteins, a group of metalloenzymes with a range of biological roles including small molecule transportation, electron transfer chemistry, biosynthesis, and xenobiotic detoxification.[101] The hemoproteins have been subjected to protein redesign in attempts to unlock abiological chemistry intrinsic to the heme cofactor while *de novo* design has focused on accomplishing the same objective by manufacturing simplified heme-containing protein scaffolds. Before a discussion on the rational design of hemoproteins is presented, it is first necessary to provide an overview of the structure, biological function, and biochemistry of naturally occurring hemoproteins.

2.2. AN INTRODUCTION TO HEMOPROTEINS

Hemoproteins are a large and diverse class of metalloproteins characterised by the integration of a heme prosthetic group/cofactor into the protein structure.[100–103] Abundant in nature, hemoproteins are involved in an assortment of vital biological processes, including small molecule binding and transportation, oxygen reduction, peroxidase/monooxygenase chemistry, and electron transport processes.[100,101] The incorporation of a heme group allows the hemoproteins to perform biological functions that would be otherwise inaccessible *via* the twenty canonical amino acids.

2.2.1 The heme group – the molecular structure and electronic spectral classification

The heme group is a family of tetrapyrrole macrocycles derived from the simplest tetrapyrrole, porphine. The porphine scaffold consists of four pyrrole heterocycles connected *via* methylene bridges (-CH_2=) to afford a planar macrocyclic molecule. The carbon backbone is divided into three positions: i) the meso-carbons forming the methylene bridges, ii) the α (connected to the meso-carbons), and iii) the β positions of the pyrrole units, respectively. The porphyrin macrocycle contains

a total of 26 π-electrons, with 18 π-electrons forming a continuous, conjugated system connecting all four pyrrole subunits, rending porphyrin an aromatic compound. In solution, porphin can exist in three protonated states, depending on the pH of the system, on account of the four pyrrolic nitrogen: $PorH_4^{2+} \leftrightarrow PorH_2 \leftrightarrow Por^{2-}$. In the fully deprotonated state, porphin can function as a tetradentate coordination ligand and can form complexes with a variety of different transition metals.

The heme groups are iron-containing porphine derivatives and are distinguished by the various functionalisations at the β-position of the pyrrole subunits (**figure 12**). The four planar pyrrolic nitrogens of the porphine ring coordinate to the iron and the axial sites of the iron are occupied by one or two coordinating amino acids to generate pentacoordinate or hexacoordinate complexes respectively.[100,101] The iron has two accessible oxidation states, the ferric (Fe^{3+}) and ferrous (Fe^{2+}) states, with the redox potential of the Fe(III)/Fe(II) transition often being finely tuned by the primary and secondary coordination spheres of proteins and the functionalisation of the tetrapyrrole ring. In addition, a change in oxidation state of the heme iron results in a conformational change to the heme and to the surrounding protein on account of the axial ligand and other interaction residues provided from the protein backbone.[100,101]

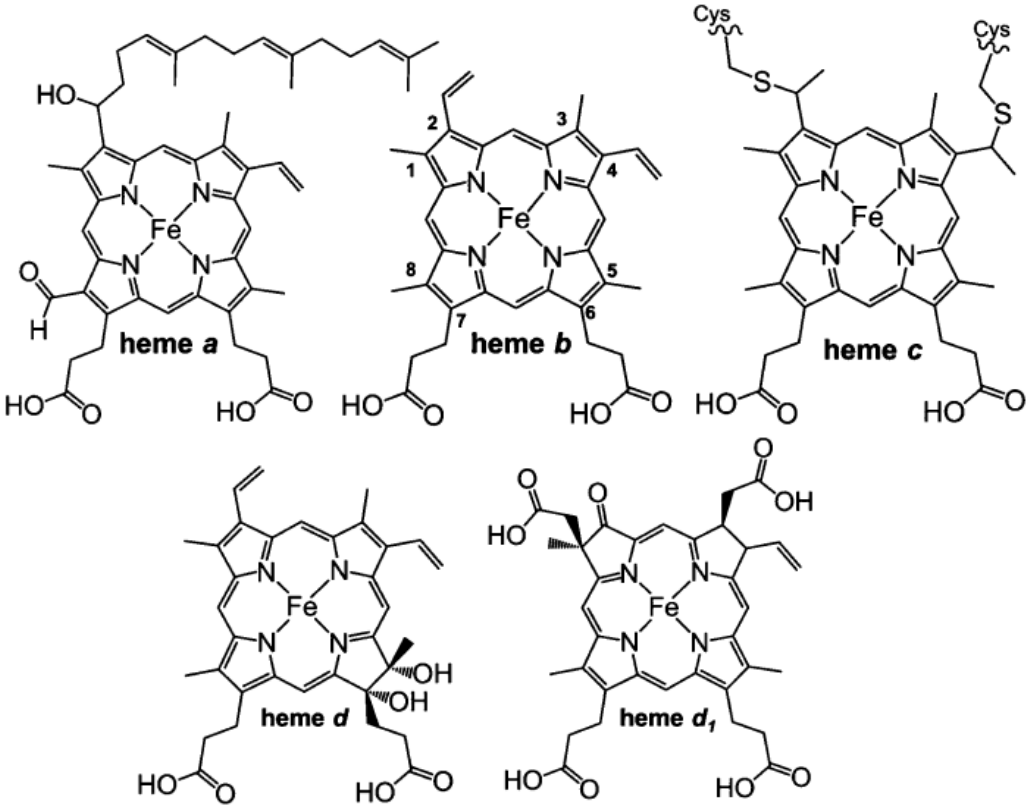


Figure 12: The molecular structure of the most common and biologically relevant heme cofactors (figure taken from Poulus *et al*).[101]

The heme cofactors can be divided into distinct groups based on the substituents at the β-position of the pyrrolic subunits constituting the porphin ring (**figure 12**). The most important heme types are heme A, heme B, heme C and heme O. Heme A and heme O are characterised by an isoprenoid hydroxyethylfarnesyl substituent at the 2nd position of the porphyrin ring; heme A also possess a formyl group at the 8th position of the porphyrin ring, which distinguishes it from heme O, which contains a methyl side chain at the 8th position. Heme A plays an important role in the conversion of energy from oxygen reduction in mammalian cells, often acting as terminal oxidases, and heme O performs a similar role in *E. coli*, on account of the much higher redox potentials of heme A and heme O relative to heme B and heme C.

Heme B and heme C are the most common and biologically relevant of the heme cofactors. The structure of heme B and heme C is illustrated in (**figure 12**). Heme B is the most abundant of the heme cofactors and is important in biological roles such as small-molecule binding and electron transfer processes; for example, heme B is the cofactor found in the oxygen-binding hemoproteins myoglobin and haemoglobin (discussed below) and the ubiquitous oxidoreductase hemoprotein cytochrome b_5.[100,101] Heme B is bound to the protein by one or two coordinating amino acids, typically histidine, cysteine, and methionine, but is not covalently attached to the protein scaffold itself. Heme C differs from heme B in that the two vinyl groups (attached to positions 2 and 4 (**figure 12**)) are modified to covalent thioether bonds, formed by the reaction with the thiol groups of two cysteine residues provided from the protein backbone. The tethering of heme C to the protein backbone *via* the thioether linkages is suspected to provide a greater degree of functional flexibility, allowing the Fe(III)/Fe(II) redox potential to be finely tuned across a far greater range than is possible with heme B (-400 to +450 mV (vs SHE) for heme C vs -400 to 180 mV (vs SHE) for heme B). Like heme B, the axial positions of heme C are often occupied by one or two coordinating amino acids, typically histidine, cysteine and/or methionine.[100]

One defining characteristic of the heme cofactors are the distinctive, and diagnostic, UV-VIS spectra of each heme originating from the allowed π-π* electronic transitions of the porphyrin ring.[100,104] The π-π* electronic transition gives an intense peak, called the Soret peak, in the 400-430 nm range and is accompanied by two weaker peaks, called the α and β peaks (collectively called the Q-band region), in the 500-600 nm range (**figure 13**); the Q-band originates from vibrational coupling to the electronic transitions (addressed in more detail in *chapter 3 section 3.4*). The positions, and intensities, of the Soret peak and the Q-band are highly sensitive to the oxidation state of iron, the coordination state of iron, the chemophysical properties of the coordinating ligands, and the substituents on the porphyrin ring. Incidentally, the heme cofactors can also be classified according to their distinctive

electronic absorption profiles when coordinated to a pyridine ligand in its ferrous state (called the pyridine hemochrome); the letters assigned to each heme are associated with the distinctive electronic spectrum for the pyridine-Fe(II)-heme complex. The heme groups can be further divided based on the distinctive UV-VIS spectra arising from variable coordinate states and axial ligands; for example, mono- or bis-histidine ligation or type I/type II cysteine ligation produce distinctive electronic absorption spectra.[100,101,104]

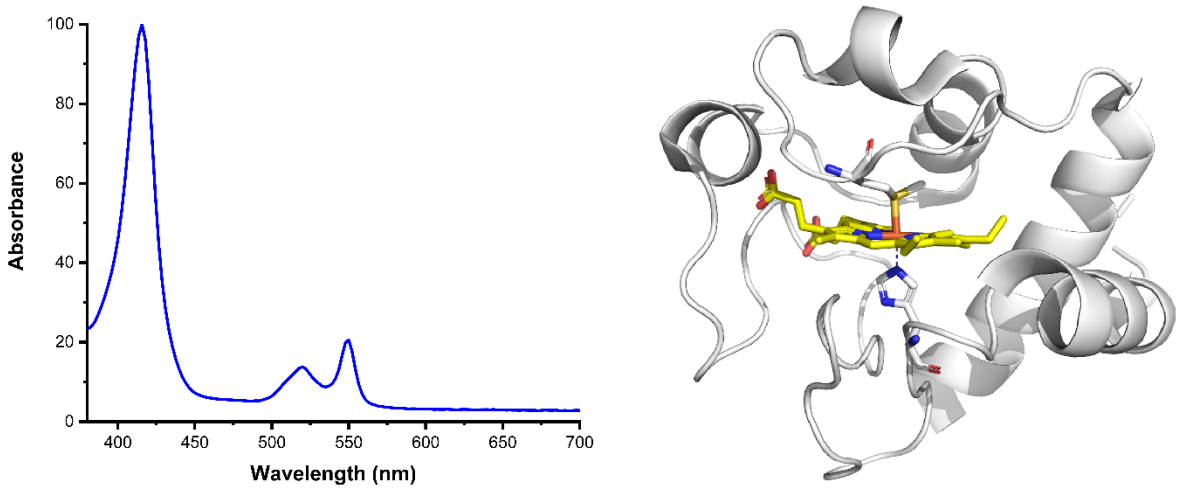


Figure 13: The UV-VIS spectrum of cytochrome *c*, **(left)** the UV-VIS spectrum of ferrous cytochrome *c* displaying the characteristic Soret peak (415 nm) and the Q-band (500-551 nm), and **(right)** the structure of horse heart cytochrome *c* showing the histidine/methionine bis-ligation to a c-type heme (heme is shown in yellow, *PDB: 1HRC*).

The unique and diagnostic electronic spectra of heme cofactors has proved indispensable for studying hemoprotein function, allowing for facile detection of electron transfer reactions ($Fe^{3+} \leftrightarrow Fe^{2+}$ transitions), small molecule binding (penta- ↔ hexa- coordinate and/or $Fe^{3+} \leftrightarrow Fe^{2}$ transitions) and the formation of reaction intermediates. In all situations, the function exhibited by the hemoprotein is a consequence of the intrinsic nature of the heme-cofactor itself, with the protein scaffold usually establishing key interactions with the heme-cofactor to finely tune parameters such as the oxidation state, standard electrode potential and/or reactivity of the iron.[101] These interactions between the protein and the heme-cofactor are often critical for the specific function of a specific hemoprotein. This versatility of the heme cofactor accounts for the ubiquitous distribution of, and the vast array of biological functions performed by, hemoproteins across all kingdoms of life.

2.2.2 Hemoprotein function – dioxygen binding, peroxidation and monooxygenation enzymology

2.2.2.1 Oxygen-binding and transporting hemoproteins

The electron-transport chain is a complex sequence of electron transfer reactions, where electrons descend an energy pathway from donor to acceptors.[105] The flow of electrons is coupled to the flow of protons across the cellular membrane which generates an electrochemical gradient. The electrochemical gradient drives the synthesis of ATP.[105] ATP functions as an energy reserve in the cell, preserving energy in highly strained triphosphate bonds which can later be released to provide the necessary energy for important cellular processes, such as biosynthesis.[105] The electron-transport chain requires a ''terminating acceptor'', the final step in the sequence of electron transport steps. In aerobic respiration the final step is the reduction of molecular oxygen to water – hence, molecular oxygen serves as the ''terminating acceptor'' in aerobic respiration.[105,106] Aerobic respiration is an immensely important biological process because it is the pathway responsible for the majority of ATP synthesised (in vertebrates).[107] It is therefore imperative that molecular oxygen is efficiency transported throughout the organism in order for the electron-transport chain to function properly. The electron transport chain itself involves several key hemoproteins, but the oxygen-binding and transportation process itself (in vertebrates) is the responsibility of two extremely important hemoproteins, myoglobin, and haemoglobin.

2.2.2.1.1 Myoglobin

Myoglobin, the first protein for which the three-dimensional tertiary structure was solved using X-ray crystallography, is a small hemoprotein. The structure of myoglobin consists of eight α-helices which adopt a tertiary structure called the 'globin fold' – a fold common to globins (**figure 14**). Myoglobin binds a single pentacoordinate, monohistidine-ligated b-type heme. The distal coordination site of the heme remains vacant and functions as the site of oxygen binding. Myoglobin can bind many small molecules other than oxygen (CO, CN^-, N_3^-), but its primary biological role is oxygen binding. The relevant physiological forms of myoglobin and deoxymyoglobin (Mb) and oxymyoglobin (MbO_2).[101,107–109]

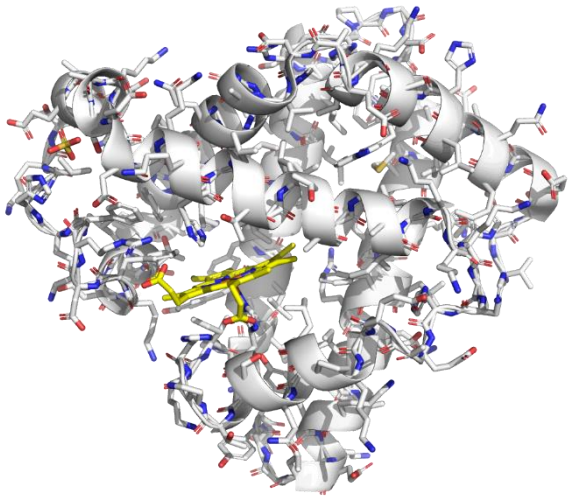


Figure 14: WT sperm whale deoxymyoglobin (heme group is shown in yellow; *PDB: 1VXA*).

The heme in deoxymyoglobin is a pentacoordinate mono-histidine ligated complex. The EPR spectra of ferrous deoxymyoglobin indicates the d^6 iron has a high-spin electronic structure with an electronic configuration of $(dxz)^2(dyz)^1(dxy)^1(dz^2)^1(dx^2\text{-}y^2)^1$ (**figure 15**).[110] The electron density situated in the $dx^2\text{-}y^2$ orbital means ferrous iron has an ionic radius too large to be accommodated inside the porphyrin cavity, and so ferrous iron deviates away from the xy-plane of the porphyrin ring towards the coordinating histidine situated along the z-axis. However, the high spin state of ferrous iron is imperative for deoxymyoglobin to bind molecular oxygen, as the ground state of molecular oxygen is a high spin triplet species (**figure 15**) meaning interactions between ground state molecular oxygen and a low spin complexes is strictly spin forbidden.[111] The binding of oxygen to deoxymyoglobin produces oxymyoglobin, a hexacoordinate, diamagnetic low-spin d^6 ferrous iron complex with a ground state electronic configuration of $(dxz)^2(dyz)^2(dxy)^2(dz^2)^2(dx^2\text{-}y^2)^0$ (**figure 15**).[110] One key structural difference between deoxy- and oxymyoglobin is the decreased ionic radius of iron in oxymyoglobin renders the iron small enough to be comfortably accommodated in the centre of the porphyrin ring. This structural change is a consequence of spin-pairing between the d-electrons of iron and the π^* electrons of oxygen which results in a destabilisation, and hence removal of electron density, from the $dx^2\text{-}y^2$ orbital (**figure 15**).[110] The active site of myoglobin also contains a histidine residue, uncoordinated to the iron, at the distal face of the heme. The histidine residue provides a stabilising interaction with molecular oxygen in the oxymyoglobin state, and is a key residue behind the preferential binding of molecular oxygen over the binding of competitive small molecules such as CO. It has been shown that mutating this key histidine residue generates a mutant myoglobin capable of peroxidase activity.[112,113] Myoglobin's biological role is to serve as an oxygen storage reserve and is

situated chiefly in muscle tissue. The transportation (and delivery) of oxygen around (and to) the body, *via* the bloodstream, is the function of another hemoprotein called haemoglobin.

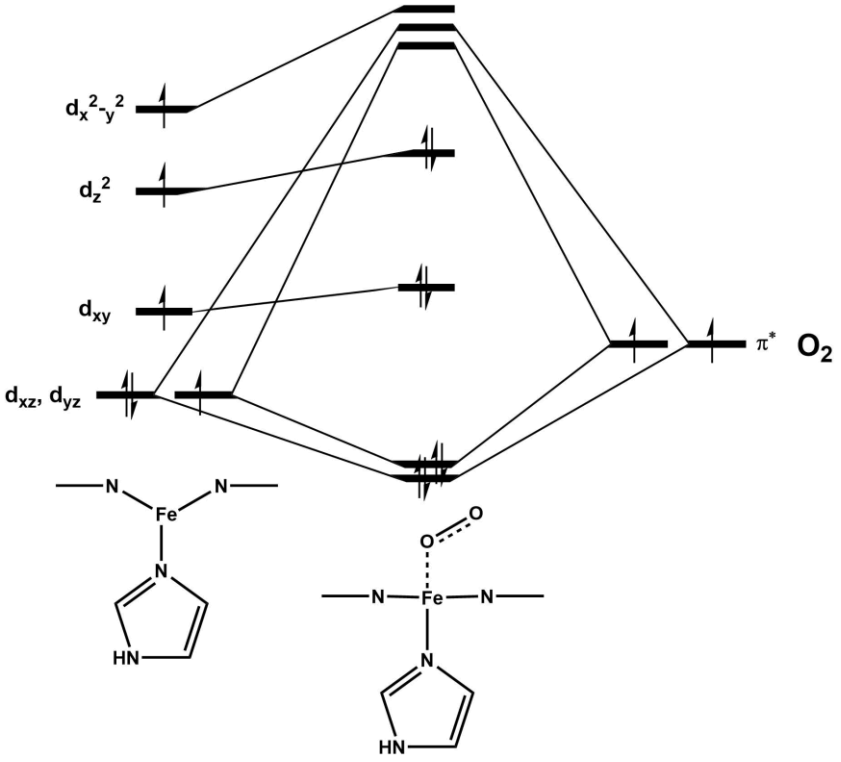


Figure 15: The electronic configuration of deoxy- and oxy- myoglobin.[110]

2.2.2.1.2 Haemoglobin

Haemoglobin is an oxygen-binding hemoprotein responsible for the transportation of oxygen around an organism.[101,107,114] Haemoglobin consists of two nearly identical all-α-helical subunits, the α and β units, which non-covalently assemble to yield a tetrameric quarternary $\alpha_2\beta_2$ structure in an approximate tetrahedral arrangement. The subunits α and β adopt the globin fold motif, and create a tight binding pocket for heme-binding similarly to myoglobin. The four subunits are assembled through a combination of non-covalent interactions, predominantly salt bridge formation and hydrogen bonding. Each subunit can bind a single heme, meaning haemoglobin is a tetrameric tetraheme-containing protein (**figure 16**). Each subunit provides a histidine ligand which coordinates to the iron in the proximal position to give a pentacoordinate heme with a vacancy in the axial position for oxygen binding. The iron can exist in two oxidation states, ferric (Fe^{3+}) and ferrous (Fe^{2+}); however, ferric-haemoglobin is incapable of binding molecular oxygen, and so ferric-haemoglobin must be reduced to ferrous-haemoglobin prior to oxygen binding. Ferrous-deoxyhaemoglobin is a high-spin paramagnetic complex, whereas oxyhaemoglobin is a ferric, low-spin diamagnetic complex. Likewise,

the structural changes in the heme group after molecular oxygen binds to deoxyhaemoglobin parallels the changes observed in myoglobin.[101,107,110,91]

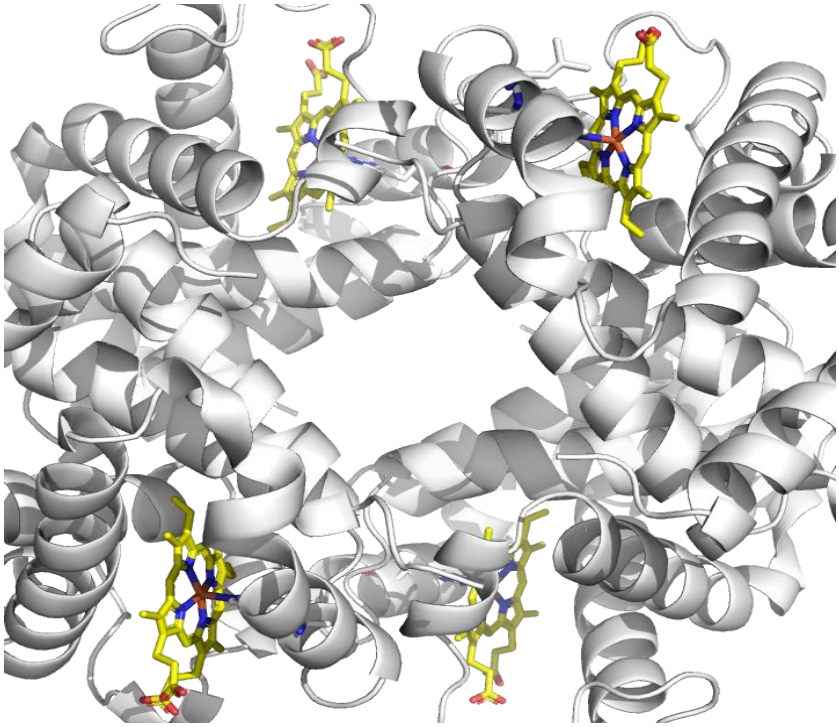


Figure 16: WT human deoxyhaemoglobin (the four heme groups are shown in yellow; *PDB: 1A3N*)

The presence of four hemes in the final structure of haemoglobin results in a more complex oxygen binding process than for myoglobin. This is best illustrated by comparing oxygen-binding curves for haemoglobin and myoglobin.[108,115] Myoglobin exhibits the distinctive hyperbolic curve for small molecule binding, showing initial binding of molecular oxygen followed by saturation once all deoxymyoglobin has been exhausted. Haemoglobin deviates away from the hyperbolic curve and exhibits a sigmoidal relation with respect to oxygen binding.[108,115,116] The plot indicates that, for haemoglobin, the binding of a second molecule of oxygen is facilitated by the first oxygen binding event, which in turn facilitates the binding of a third oxygen molecule and so on.[114,115,117–119] The phenomenon of the first ligand binding event enhancing successive ligand binding is called *cooperativity* and can be quantified by the Hill equation (which was first postulated to account for the O_2-sigmodial binding cure exhibited by haemoglobin).[117,118,120,121] The structural rationale behind cooperative binding suggests that the four subunits of haemoglobin exist in two interconverting states: the *'T'-state* (tense-state) and the *'R'-state* (relaxed state).[115,116,118,119,122,123] A subunit of deoxyhaemoglobin exists predominantly in the T-state which possesses a lower binding affinity for molecular oxygen relative to the R-state. The T-state and R-state are in equilibrium, and the binding of molecular oxygen stabilises the R-state of the oxygenated subunit, which, in the quarternary state,

induces a confirmation transition within the subunit. The conformational change is intramolecularly transmitted through the remaining deoxygenated subunits to shift the equilibrium in favour of the R-state which facilitates the binding of a second molecule of molecular oxygen because of the higher binding affinity in the R-state. The process repeats until saturation.[115,118,119,122,123] The two models postulated to structurally account for cooperative binding are the *concerted*[124] and *sequential*[125] models. In the concerted model, all four subunits of haemoglobin interconvert between the R and T states simultaneously, with each molecular oxygen binding event stabilising the R-state until saturation.[118,121,124] In the sequential model, the four subunits are in the R-T equilibrium independently, and the binding of molecular oxygen stabilises the R-state of an individual subunit only; however, the stabilisation of the R-state in a single subunit translates through the remaining domains to shift the equilibrium in the remaining deoxygenated subunits towards the R-state which facilitates additional oxygen binding events.[121,125]

Haemoglobin transports oxygen through the bloodstream before transferring oxygen to myoglobin, which possess a higher affinity of oxygen relative to haemoglobin.[107] The remarkable system established by these two hemoproteins beautifully highlights how the protein structure can modify and finely tune the functionality of a cofactor. Hemoproteins are not limited to small-molecular binding, and other hemoproteins exhibit catalytic activity with the protein scaffolding serving to finely tune the heme cofactor towards a particular type of chemistry.

2.2.2.2 Catalase and peroxidase hemoproteins

The peroxidases are a large group of heme-containing enzymes that catalyse the peroxide-dependent oxidation of an assortment of substrates.[101,126,127] The common feature of all heme-peroxidases is the use of hydrogen peroxide or organic hydroperoxides to catalyse a range of oxidation reactions *via* a reactive intermediate called compound I, a heme-localised cationic-radical oxene complex (**figure 17**). The formation of compound I allows peroxidase enzymes to perform important biological roles, including reactive-oxygen-species scavenging, providing protection from oxidative stress, and biopolymer degradation. Unlike oxygen-binding hemoproteins, the binding of hydrogen peroxide to heme is not spin-forbidden[111], and an iron-localised peroxide complex forms rapidly without the need for prior reduction from the ferric to the ferrous state.[101]

The catalytically cycle of a heme peroxidase is illustrated in (**figure 17**).[101-104] The initial step involves the formation of a ferrous-heme-peroxide complex. The second step involves the heterolytic cleavage of hydrogen peroxide to eliminate water. In this step, the ferric iron is oxidised by oxygen to generate a ferryl species, with another electron being supplied *via* the oxidation of the porphyrin ring or a

suitable amino acid side chain, such as tryptophan. After the heterolytic elimination of water, the resultant complex is a highly reactive ferryl cationic radical intermediate called *compound I*.[101] In the third step, the substrate is oxidised by compound I, resulting in the reduction of the cationic radical to a ferryl complex called *compound II.*[101] In the final step, a second substrate is oxidised which completes the catalytic cycle by regenerating the initial ferric-heme. A distal histidine in the active site of many peroxidases facilitates the heterolytic cleavage of peroxide by facilitating the initial deprotonation of the oxene oxygen and the transfer of the proton to the departing hydroxyl group. The distal histidine functions as an acid-base catalysis during the catalytic cycle.[101]

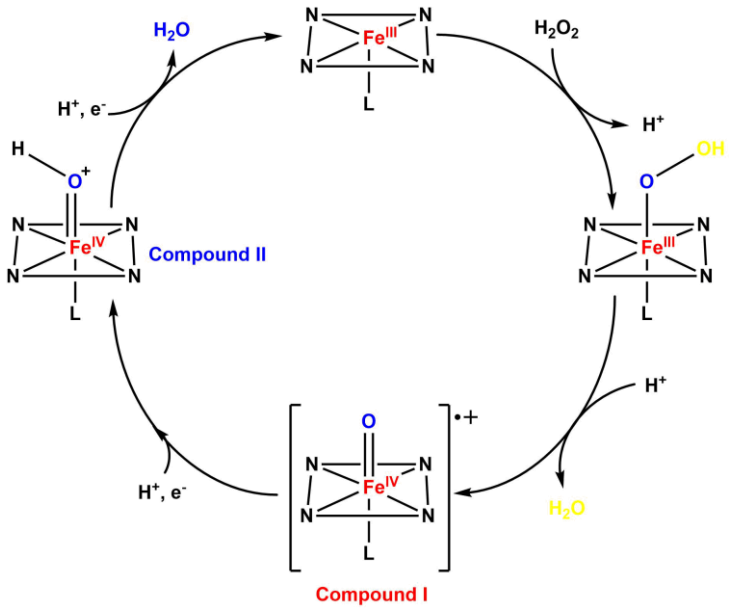


Figure 17: A general reaction cycle for heme peroxidase chemistry. The reaction of ferric heme with hydrogen peroxide results in the heme-peroxide complex. A proton transfer and loss of water results in the formation of the highly reactive cationic-radical species compound I. A proton-coupled reduction of compound I yields compound II, and a second proton-coupled reduction produces water and regenerates the starting ferric enzyme.

The heme peroxidases are split into three distinctive class, intercellular (class I), fungal extracellular (class II) and plant extracellular (class III).[101,129,130] The class I peroxidases are intercellular enzymes which function as biological scavengers for reactive oxygen species. Although the class I peroxidases share a similar biological function, the enzymology of each specific heme peroxidase can vary greatly in terms of the enzyme's active site, heme axial ligand, quarternary structure, necessary redox partners, substrates employed in the oxidative mechanism and the localisation of the cationic radical in compound I.[101,129,130] These variances are amply highlight by considering the distinctive differences between catalase and two common class I heme peroxidases, cytochrome *c* peroxidase, and ascorbate peroxidase.

2.2.2.2.1 Catalase

The first enzyme to break down peroxides was reported in 1818 to account for an unidentified biological substance that catalysed the disproportionation of cellular hydrogen peroxide to molecule oxygen and water.[131] The unknown biological substance was later deciphered to be a tetraheme containing metalloenzyme. The metalloprotein was named *catalase* and it was purported to have a tetrameric quarternary structure formed by the assembly of four single heme-containing domains (**figure 18**).[131–134] The four hemes are penta-coordinate, with a tyrosine residue coordinated at the axial position. The main biological function of catalase is to protect cells from oxidative damage caused by reactive oxygen species, and as such catalase has one of the largest turnovers reported for an enzyme. Although technically not a peroxidase, as catalase catalyses the disproportionation of peroxide instead of generating an activate substrate, catalase and peroxidases provide similar biological roles, and both proceed *via* the active cationic radical intermediate compound I.

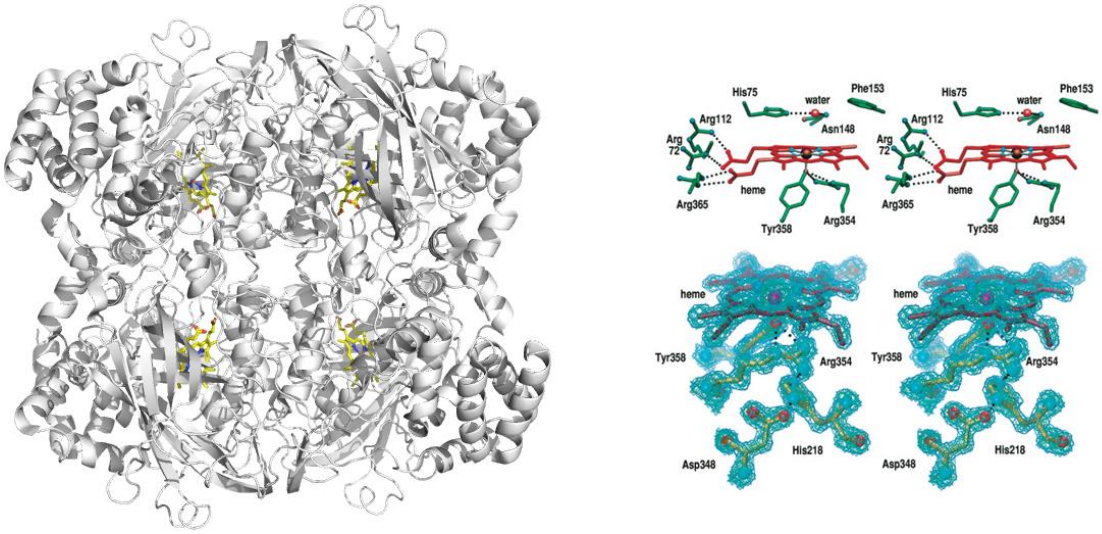


Figure 18: (**left**) Human erythrocyte catalase (hemes shown in yellow; *PDB: 1QQW*), and (**right**) active site of human catalase (figure taken from Putnam *et al*).[133]

The catalytic cycle for catalase epitomises the aforementioned peroxidase catalytic cycle (despite being distinguished from peroxidases). The active site of wild-type catalase is illustrated in **figure 18** with the key amino acids involved in the peroxidase cycle highlighted.[133,135] The initial diffusion of hydrogen peroxide into the active site is facilitated by an asparagine (Asp148) which favourably interacts with hydrogen peroxide.[106] The histidine (His75) functions as an acid-base catalyst and facilitates the heterolytic cleave of the O-O peroxide σ-bond; the whole process is mediated by a water molecule.[101,106] The tyrosine axial ligand (Tyr358) promotes the oxidation of Fe(III) to Fe(IV) and is susceptible to oxidation by the oxene during the formation of compound I.[106] Although the cationic

radical of compound I is suspected to be localised on the porphyrin ring there is evidence to indicate the formation of a tyrosyl radical on the axial tyrosine.[101,136–138] A second hydrogen peroxide molecule is deprotonated by compound I, forming compound II, and is concurrently deprotonated by His75 to release molecular oxygen. The final stage involves the transfer of the proton from His75 to compound II, liberating a water molecule and completing the catalytic cycle.[106]

2.2.2.2.2 Cytochrome *c* peroxidase

Cytochrome *c* peroxidase (CcP) is a small water-soluble hemoprotein first reported in 1939.[139] The structure of yeast CcP is highly α-helical, with a single b-type heme bound within a hydrophobic pocket in the active site.[101,140] The heme is bound to CcP *via* monohistidine ligation (in contrast to mono-tyrosine ligation in catalase), resulting in a pentacoordinate, high-spin, ferric complex resting state. The distal side of the heme is vacant (often coordinated by a liable water molecule) and therefore available for catalysis (**figure 19**).

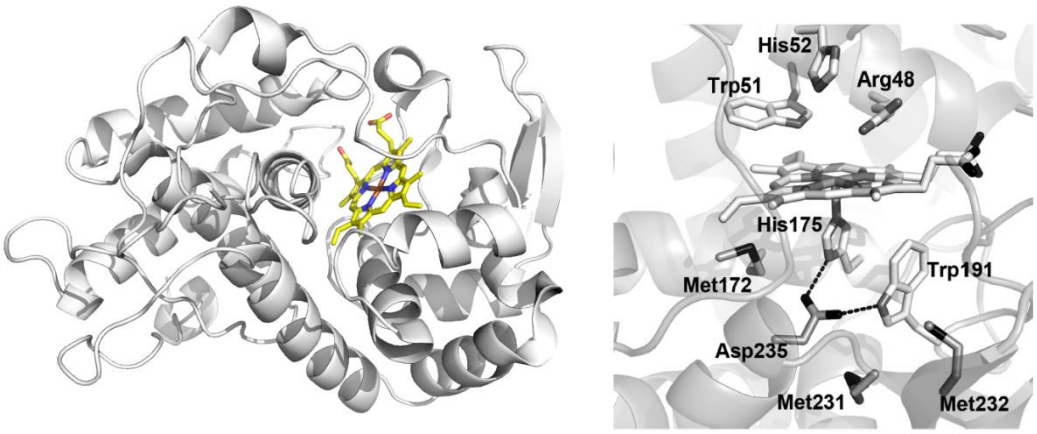


Figure 19: (**left**) Crystal structure of yeast cytochrome *c* peroxidase (CcP) (heme is shown in yellow; *PBD: 2CYP*), and (**right**) the active site of yeast cytochrome *c* peroxidase (figure taken from Poulos)[101]

CcP, similar to catalase, catalytically degrades hydrogen peroxide (or alkyl hydroperoxides) to two equivalents of water; however, unlike catalase, CcP requires an additional enzyme, the heme-containing metalloenzyme cytochrome *c*, in order to adequately function.[141] Cytochrome *c* (Cc) is a small c-type hemoprotein structurally consisting of 5 α-helices and a β-sheet.[142] The axial ligands for Cc are provided by a histidine and methionine residue, resulting in a hexacoordinate, low spin, paramagnetic ferric complex.[101,141–143] Cc functions as an electron-transfer protein, readily converting between ferric and ferrous oxidation states. The main biological role of Cc is to function as a redox partner for other metalloenzymes, such as CcP. The formation of a Cc-CcP complex (**figure 20**)

precedes peroxidase activity, and the catalytic cycle includes the breakdown of hydrogen peroxide and the concurrent oxidation of two ferrous cytochrome *c* enzymes.[101,141]

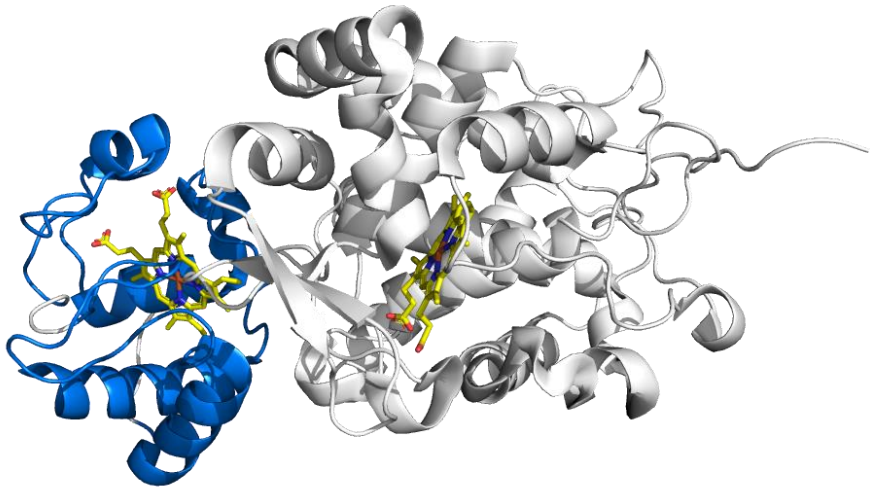


Figure 20: The crystal structure of the 1:1 complex formed between cytochrome *c* peroxidase (CcP, white) and yeast cytochrome *c* (Cc, iso-1, blue) (hemes are shown in yellow; *PDB: 2PCB*)

The first step of the Cc-CcP catalytic cycle (**figure 21**) is the formation of compound I from the heterolytic cleavage of hydrogen peroxide, accompanied by binding of one equivalent of ferrous cytochrome *c*.[101,141] An electron transfer from Cc to CcP reduces compound I to compound II, which is followed by the disassociation of ferric-cytochrome *c* and binding of a second equivalent of ferrous cytochrome *c*. Another electron transfer step occurs, followed by disassociation of the second ferric cytochrome *c* to yield the starting ferric CcP. The active site of CcP resembles that of catalase, with an acid-base histidine residue (His-52) and a stabilising arginine residue (Arg-48); however, the presence of a tryptophan residue (Trp-191) further distinguishes CcP from catalase. The Trp-191 residue is critical for the two electron transfer steps.[101,141] In the formation of compound I (first electron transfer step) tryptophan provides the necessary reducing electron resulting in a Fe(IV) $Trp^{\bullet+}$ intermediate; therefore, unlike catalase, the formation of compound I results in a low-spin ferryl complex with the cationic radical localised on the tryptophan (unlike in catalase, where the cationic radical is suspected to be a porphyrin/tyrosyl species). Following the reduction of compound I to compound II, *via* one reducing equivalent of Cc, an equilibrium exists between the Fe(IV)-Trp and Fe(III)-$Trp^{\bullet+}$ species, and, in addition to providing extra stabilisation to the intermediate, Fe(III)-$Trp^{\bullet+}$ behaves as an electron acceptor for the second electron transfer step provided by ferrous cytochrome *c*.[101,141]

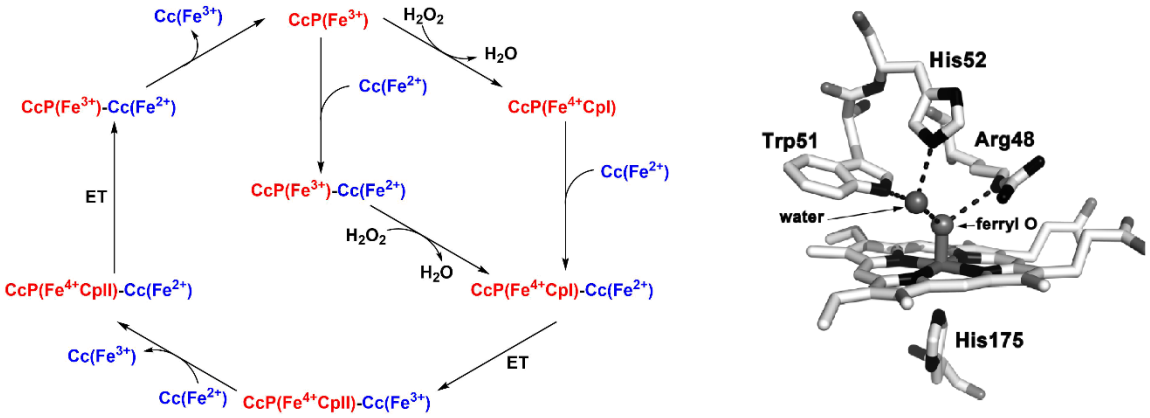


Figure 21: (**left**) The cytochrome *c* peroxidase catalytic cycle (CcP = cytochrome *c* peroxidase, Cc = cytochrome *c*, CpI = compound I, CpII = compound II, ET = electron transfer; figure adapted from Volkov *et al*),[141] and (**right**) structure of compound I in yeast cytochrome *c* peroxidase (figure taken from Poulos).[101]

A small distinct group of cytochrome *c* peroxidases are the diheme-containing cytochrome peroxidases, which containing two distinctive c-type heme containing domains.[144–147] The two hemes have drastically different redox potentials, with a high-potential heme acting as the reducing equivalent for a low-potential heme.[144,146,147] The high-potential heme is first reduced by redox partners, and transports electrons to the low-potential heme which functions as the site of peroxide reduction in a fashion analogous to catalase.

2.2.2.2.3 Ascorbate peroxidase

A final peroxidase worth mentioning is ascorbate peroxidase (APX). APX was first identified in 1979[148] as the metalloenzyme responsible for the peroxide-dependent oxidation of ascorbate in plants.[149] APX is notable for its high-sequence identity to CcP.[129,130,150] The heme-group in APX is a pentacoordinate, b-type heme with the proximal site coordinated to a histidine and a tryptophan residue (Trp-179) situated in the active site.[101,151–153] Unlike CcP, however, APX possess a monovalent cation in close proximity to the active site tryptophan (**figure 22**).[101]

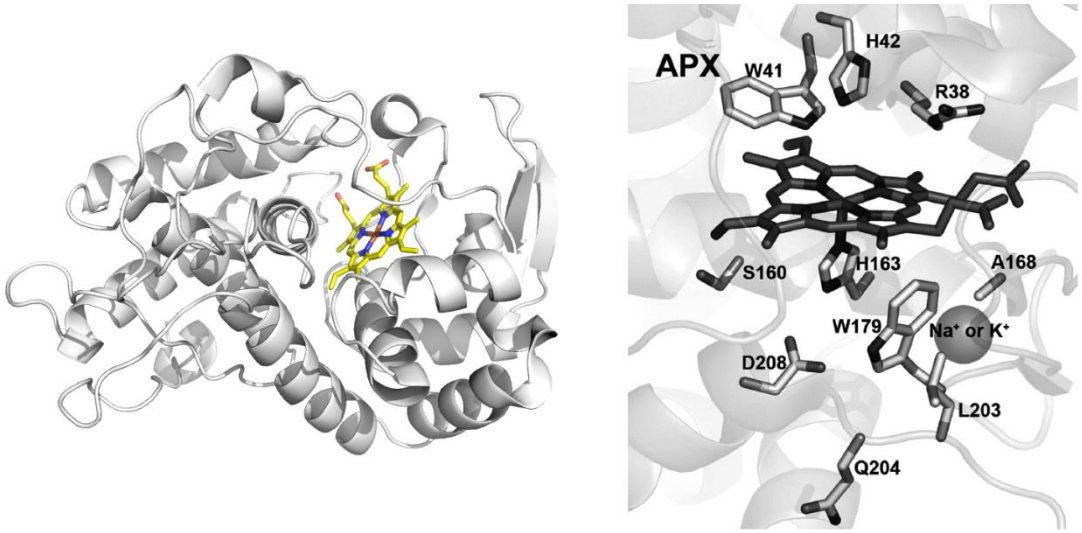


Figure 22: (**left**) Crystal structure of soybean ascorbate peroxidase (heme is shown in yellow; *PDB: 1OAG*), and (**right**) the active site of ascorbate peroxidase, showing the binding site of the monovalent cation (figure taken from Poulos).[101]

Although APX exhibits a general peroxidase catalytic cycle, the presence of a monovalent cation disfavours the formation of a tryptophan radical and so the structure of compound I is accounted for as a Fe(IV)-por$^{\bullet+}$ species, with the cationic radical localised on the porphyrin ring of the heme group.[101,153] The catalytic cycle of APX is distinguished from the cycle observed for catalase and CcP because the reduction of compound I and compound II is achieved by a single substrate molecule, ascorbate, in two sequential electron transfer steps (**figure 23**).[101,151,152] The reaction cycle involves the formation of compound I from hydrogen peroxide, followed by the subsequent oxidation of ascorbate to yield compound II and the anionic radical monodehydroascorbate. A subsequent electron transfer step reduces compound II, completing the catalytic cycle and yielding the oxidative product dehydroascorbate and the original ferric-APX complex.[101,151,152]

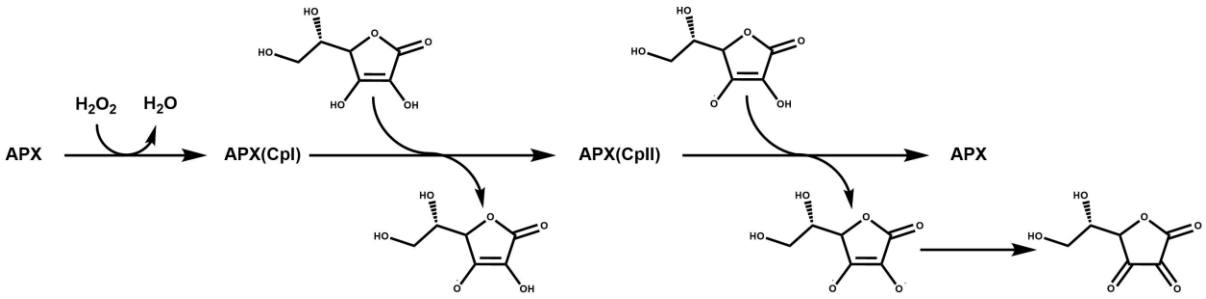


Figure 23: The overall reaction mechanism for ascorbate peroxidase (CpI = compound I, CpII = compound II).

2.2.2.3 Monooxygenase hemoproteins

The class I heme peroxidases are distinctive in their ability to utilise hydrogen peroxide to oxidase substrates. The monooxygenases are another group of heme-containing metalloenzymes that reduce molecular oxygen in the oxidation of substrates.[101,154–157] The typical catalytic cycle (**figure 24**) involves the reduction of molecular oxygen to water and the concurrent oxidation of an organic substrate by the insertion of a single oxygen to form a hydroxyl or epoxide group.[101] One of the most important groups of monooxygenases are the cytochrome P450s, a superfamily of thiolate-ligated heme-containing monooxygenases involved in an assortment of important biological processes including steroidogenesis, aerobic respiration, xenobiotic detoxification and secondary metabolite biosynthesis.[101,104,157] The 'P450' classification arises from the characteristic maximum absorbance peak at 450 nm observed in the electronic spectra for ferrous-P450-CO complexes.[101,104]

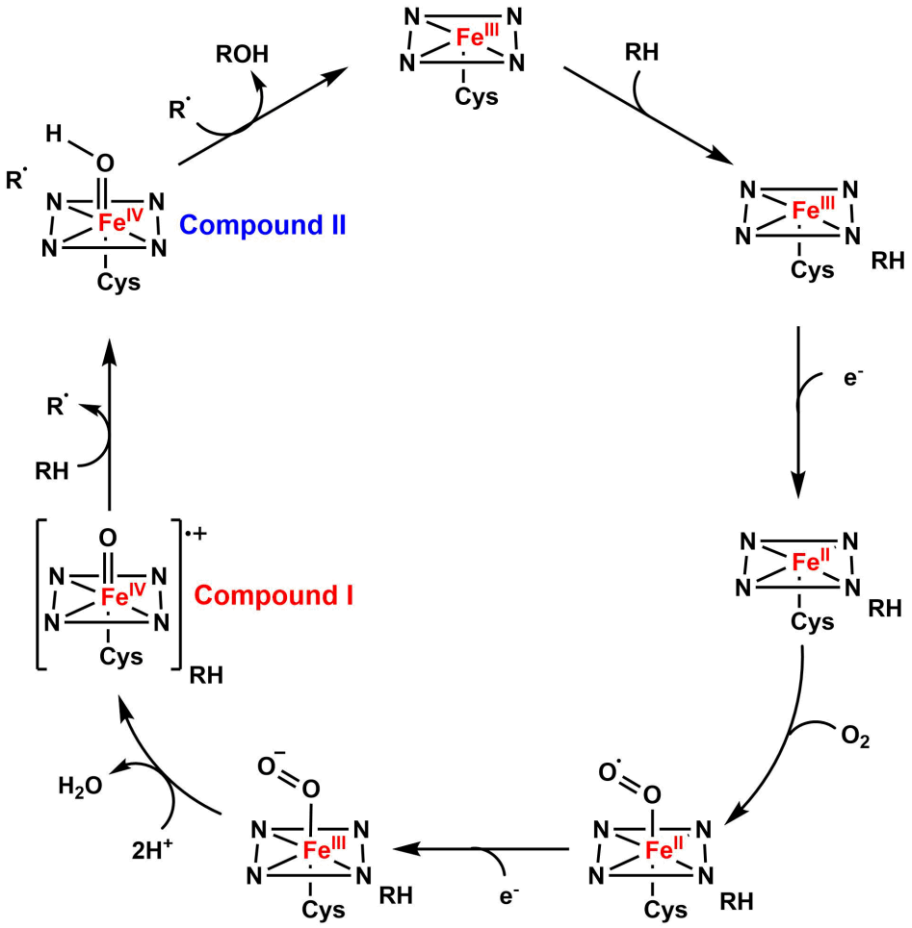


Figure 24: A general catalytic cycle for heme monooxygenase chemistry.

2.2.2.3.1 Monooxygenase catalytic cycle

In addition to the initial reduction, the formation of additional intermediates in the catalytic cycle prior to the formation of compound I further distinguishes monooxygenases from peroxidases.[101,158,159] The general catalytic cycle of a cytochrome P450 is illustrated in (**figure 24**). The resting state of cytochrome P450s is typically a low-spin thiolate-ligated ferric complex with a Fe(II)/Fe(III) reduction potential lower than the reduction potential of NAD(P)H/NADH. As a consequence, NAD(P)H/NADH cannot reduce a P450 in its resting state. The binding of a substrate to the complex displaces a liable water ligand and converts the complex into a high-spin state, which also elevates the Fe(II)/Fe(III) reduction potential above the lower limit for NAD(P)H/NADH reduction, rendering the P450 susceptible to reduction by NAD(P)H/NADH. After reduction, molecular oxygen binds to P450 to generate a ferric-$O_2^{\bullet}$ superoxide complex. A second electron transfer step results in the ferric-OO^- peroxy complex which, upon protonation, heterolytically cleaves to release a molecule of water and form the ferryl Fe(IV)=$O^{\bullet+}$ cationic radical intermediate compound I.[101,158,159] The thiol ligand coordinated to the heme iron in the proximal position (provided by a backbone cysteine residue) is thought to facilitate the heterolytic cleavage of the peroxy intermediate through increased electron donation from the sulphur onto iron, which increases the electron density on iron and enhances backbonding into the O-O $p\pi^*$ orbital, weakening the O-O bond and therefore facilitating its heterolytic cleavage.[160–162] The substrate is subsequently deprotonated by compound I to yield compound II which functions as an 'activated hydroxyl group' and is readily transferred to the deprotonated/oxidised substrate to complete the catalytic cycle, yielding the monooxygenated product and the original low-spin hexacoordinate ferric-P450.[101,158,159]

Although the reactivity of cytochrome P450 monooxygenases proceeds *via* the same highly reactive intermediate as the peroxidases (compound I), there are key differences in the overall catalytic cycle. The first major difference is that, unlike peroxidases which are active in the ferric form, the P450s required a reduction to the ferrous state prior to binding and activating molecular oxygen. To achieve reduction, the P450s are often associated with a redox-partner which transports electrons from common biological reductants, such as NADPH, to the heme iron of the P450.[101,163,164] The need for reduction is a consequence of the spin-forbidden interaction between high-spin molecular oxygen and the low-spin ferric resting state of cytochrome P450s; a one-electron reduction of the low-spin ferric heme results in a high-spin ferrous heme complex which can freely interact with molecular oxygen without violating spin-interaction laws.

2.2.2.3.2 A case study: the heme-domain and chemistry of cytochrome $P450_{BM3}$

Cytochrome $P450_{BM3}$ is one the best studied of the cytochrome P450 superfamily. $P450_{BM3}$ was first isolated from the soil bacterium *Bacillus mageterium* by Miura and Fulco in 1974.[165] $P450_{BM3}$ functions as a cell-free NAD(P)H and O_2 dependent monooxygenase and catalyses the regio- and stereoselective hydroxylation of numerous long-chain fatty acids, alcohols, and amides, typically hydroxylating at the ω1-3 positions.[165,166] It was also demonstrated that $P450_{BM3}$ can catalyse the regio- and stereoselective epoxidation of unsaturated substrates.[167] $P450_{BM3}$ reached great popularity as a model P450 on account of its self-sufficiency: its electron-donating redox partner, a FAD/FMN reductase domain (cytochrome P450 reductase (CPR)), is covalently fused to the C-terminal end of the heme-containing domain, circumventing the need for co-expressing and binding any necessary redox partners.[101] In light of its popularity, $P450_{BM3}$ has also received a great deal of attention from the rational design community as a model hemoprotein for introducing novel functionality. It is therefore prudent to give a comprehensive description of the structural, functional, and catalytic properties of $P450_{BM3}$ before introducing its applications in protein design.

The structure of the heme-domain of $P450_{BM3}$ has been determined crystallographically (**figure 25**) and its biophysical properties have been exceptionally well characterised. The heme-domain of $P450_{BM3}$ consists of an α and a β domain.[168,169] The α-domain accounts for 70% of the overall structure and is mostly α-helical,[168] with the 12 α-helices labelled A-L.[168,170–172] The main structural feature of the α-domain is a right-handed four-α-helix bundle formed through the parallel arrangement of helices D, I and L which are anti-parallel to helix E.[145] The β-domain constitutes 22% of the overall structure and consists of two β-sheets blanketed by three α-helices and two 3_{10} helices.[145] The active site of the heme is accessible through a channel formed mostly from nonaromatic hydrophobic residues in the β domain.[168–171] The heme cofactor is situated between helices I and L of the α-domain which shields the heme from the solvent.[168–171] $P450_{BM3}$ is classified as a low-spin, pentacoordinate, type I thiolate-ligated hemoprotein,[169] and a liable water molecule occupies the *trans*-axial position relative to the cysteine thiolate.[169] A phenylalanine protrudes into the active site and is orientated perpendicular to the heme,[168,171] and four solvent exposed hydrophobic residues constitute the solvent docking site.[145] It was demonstrated that $P450_{BM3}$ possesses one of the highest known catalytic efficiencies for P450s, turning over arachidonate with an activity of 15,000-17,000 min^{-1}.[169,173] The structural study of the $P450_{BM3}$ heme domain complexed with the fatty acid substrate palmitoleic acid revealed significant conformational differences between the free-form and the complexed domain.[171] Analysis of the substrate-free and substrate-bound structures revealed the most important conformational changes occurred in the F, G and H helices.[171] The F/G loops form one side of the

hydrophobic substrate access channel, and upon substrate binding a conformational change in the F/G loops results in a narrowing of the access tunnel.[171] The narrowing of the access channel allows for new hydrophobic and hydrogen-bonding interactions to form between various amino acid residues, which augments the rigidification of the F/G loop region.[171]

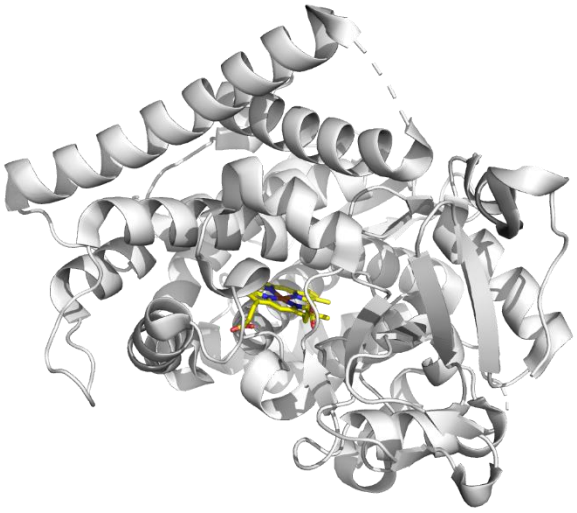


Figure 25: The atomic structure of the P450-BM3 heme domain (heme shown in yellow; *PBD: 2IJ2*).

P450-BM3 catalyses the monooxygenation of saturated and mono-unsaturated substrates, including fatty acids, alcohols, and amides.[154,166,167,174] The reactions are typically regio- and stereoselective, occurring at the ω-1, ω-2, and ω-3 positions.[166] The mechanism of monooxygenation epitomises the classical oxygenase catalytic cycle.[101] The reactivity and selectivity of P450-BM3 catalysed monooxygenation is governed by an intricate network of interactions between key amino acids and the substrate in the active site. In the resting state, the ferric iron is a low-spin, six-coordinate complex (a liable water molecule occupies the proximal site) and therefore the resting state is spin-forbidden from interacting with triplet molecular oxygen.[101,175] The first step of the catalytic cycle involves the binding of the substrate, which displaces the liable water molecule and yields a ferric high-spin iron; the interaction between a high spin heme and triplet oxygen is not spin forbidden, and so the binding of the substrate converts the catalytically inactive resting state to the catalytically active state.[101,175] The binding of two substrates, sodium laurate and 12-bromo-laurate, to the ferric heme domain was studied using paramagnetic relaxation techniques and revealed the substrates bind in an extended conformation, with the reactive ω-1, ω-2, and ω-3 positions situated too far from the iron for hydroxylation to proceed.[175] After binding of a substrate, ferric iron is reduced to ferrous iron and the reduction is accompanied by a 6 Å displacement of the substrate towards the catalytic iron.[175] It was reported the ω-1 and ω-2 positions are situated 3 Å from iron in the ferrous state, as opposed to 9 Å in the ferric state.[175] **Figure 26** shows the overlapping structure of the substrate-free and substrate-

bound P450-BM3 heme domain when ferrous P450-BM3 is bound to the inhibiting substrate omeprazole (**figure 26**) and illustrates the significant conformational changes accompanying a substrate binding event and subsequent reduction.

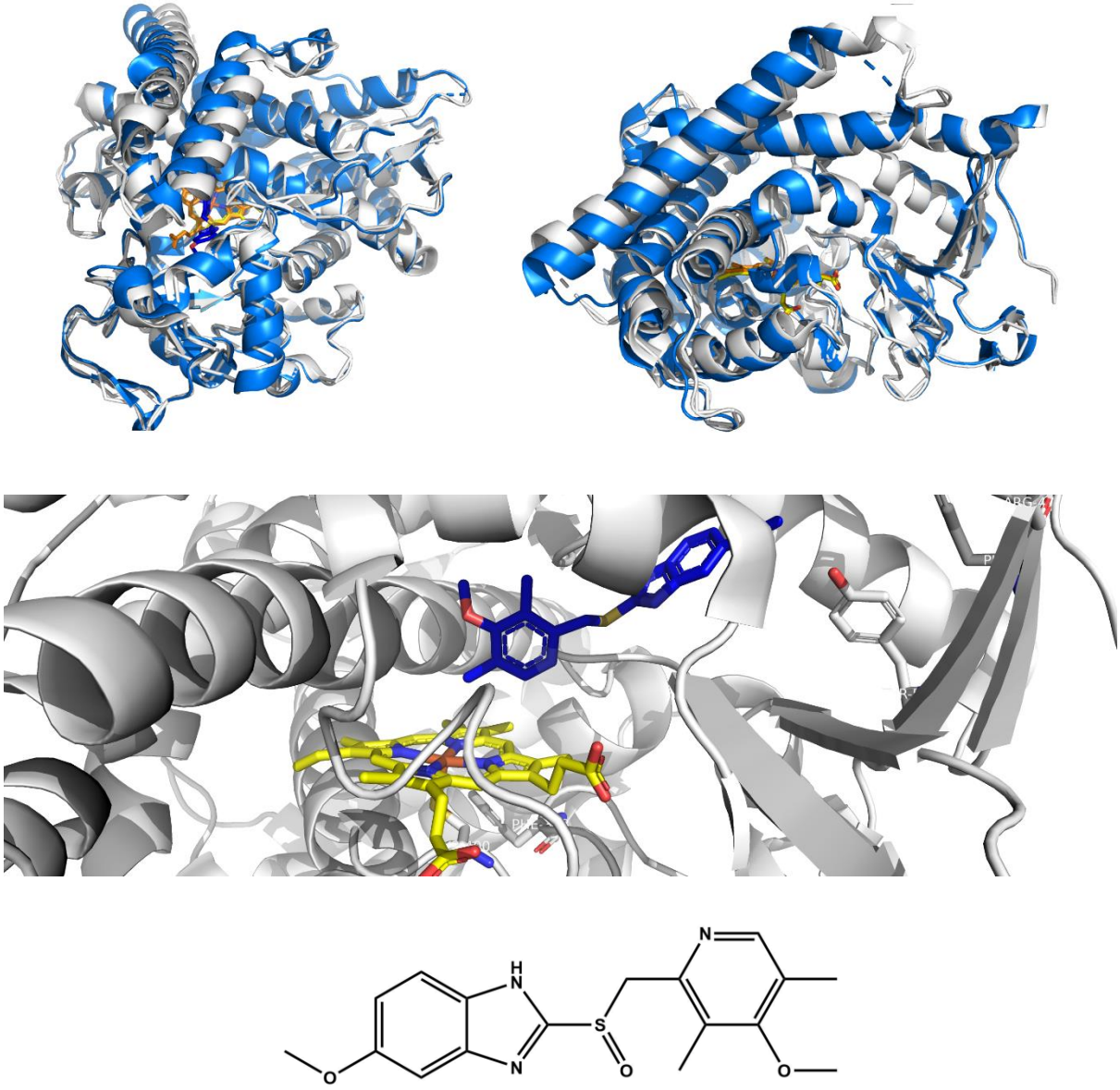


Figure 26: (top) The overlapping crystal structure of the substrate-free (blue, heme is shown in orange; *PDB: 2IJ2*) and substrate-bound P450-BM3 (white, heme is shown in yellow, omeprazole is shown in purple; *PDB: 4KEW*) heme domain when ferrous P450-BM3 is bound to the inhibiting substrate omeprazole. **(middle)** The active site crystal structure of the omeprazole-bound P450-BM3 heme domain (white, heme is shown in yellow, omeprazole is shown in purple; *PDB: 4KEW*), and **(bottom)** the molecular structure of omeprazole.

It appears binding of the substrate raises the redox potential of the iron sufficiently to allow for reduction by NAD(P)H.[101,175] The reduction results in a significant conformational change which orientates the ω-1, ω-2 and ω-3 positions into proximity of the catalytically active iron.[175] Four active

site amino acids, Phe-42, Arg-47, Tyr-51, and Phe-87, have been identified as critical residues involved in substrate binding in the active site (**figure 27**).[176] Arg-47 is proposed to establish favourable interactions with the carboxylate terminal of fatty acids through the formation of an electrostatic guanidinium-carboxylate pair.[176] A kinetic study of P450-BM3 mutants containing Arg-47 mutations revealed higher K_M and reduced k_{cat} values for laurate and arachidonate, indicating the guanidinium-carboxylate interaction is key in the initial substrate binding step and in stabilising the transition state.[176] A BM3 mutant containing a Y51F mutation exhibited increased K_M values, but the k_{cat} was practically unaffected, indicating Tyr-51 plays a key role in the initial substrate binding step, probably through establishing a hydrogen-bond with the carboxylate group of the fatty acid.[176] The residue Phe-42 resides close to the entrance of the active site on the protein surface. After binding of the substrate, Phe-42 is re-orientated and provides a ''cap'' over the binding site which is postulated to protect the hydrophobic substrate and increase the hydrophobicity by excluding the diffusion of solvent molecules into the active site. It is also postulated that Phe-42 constitutes a fraction of the hydrophobic docking site of the substrate, which is supported by observations that a P450 variant containing a F42A mutant exhibited a 7-fold increase in K_M.[176] The re-orientation of P450-BM3 upon substrate binding displaces the residue Phe-87 from a plane perpendicular to the heme cofactor to a conformation parallel with the bound substrate. Phe-87 is orientated in close proximity to the heme cofactor and has been demonstrated to govern the regiospecificity to monooxygenation reactions.[176]

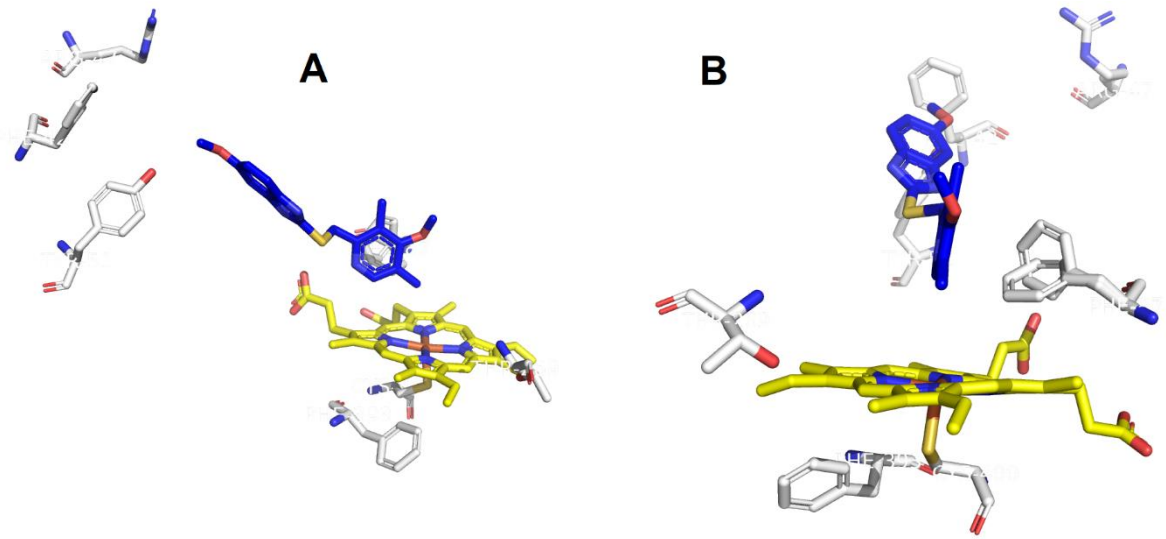


Figure 27: The active site crystal structure of the omeprazole-bound P450-BM3 heme domain (white, heme is shown in yellow, omeprazole is shown in purple; *PDB: 4KEW*) displaying key amino acids involved in substrate binding (note the two distinctive rotamers of Phe87) from (**A**) a side-on and (**B**) a longitudinal perspective relative to the plane of omeprazole.

The highly conserved residue Phe-393 in P450-BM3 resides close to the Cys400 thiolate ligand in the active site and has been demonstrated to regulate the redox potential of P450-BM3.[177–179] In regulating the redox potential of the heme, Phe-393 plays a critical role in the FMN-to-heme electron transfer step, the binding of molecular oxygen, and its subsequent reduction.[177] The kinetic analysis for the turnover of laurate and arachidonate by a series of BM3 variants containing mutations targeting Phe-393 displayed the trend WT > F393Y > F393H > F393A.[178] Interestingly, the kinetic analysis of the FMN-to-heme transfer revealed the exact opposite trend.[178] The trend supports the notion that an increase in free and substrate-bound redox potential favours the stabilisation of the ferrous complex relative to WT. Increasing the redox potential facilitates electron transfer but also hinders the rate of oxygen reduction, resulting in a decreased turnover rate. In a follow-up study, the spectroscopic and structural comparisons between WT-BM3 and the F393H mutant revealed a number of factors governing the redox potential.[179] No reported differences in the oxidation products produced by each variant were identified, indicating Phe-393, although critical for catalysis, exerts little influence over the selectivity of the monooxygenation reaction.[179] This was supported as the crystal structures of both WT-BM3 and the F393H mutant exhibited minimal conformational deviations. The authors could not conclude how the phenylalanine to histidine mutation resulted in such a drastic change in redox potential but conjectured that the histidine exists in a protonated state which alters the dielectric constant of the microenvironment surrounding the heme and the cysteine thiolate.[179] In an additional study, Raman spectroscopy was employed to study the Fe-S stretching frequency ($\nu_{Fe\text{-}S}$) in a series of Phe-393 mutants. The authors surprisingly reported that the Fe-S bond is largely unaffected by a Phe-393 mutation.[177] The authors discovered the extent of conformational changes to the heme cofactor upon reduction correlated reasonably well with the steric bulk of the Phe393 mutants and the heme redox potential.[177]

A final residue important in governing the catalytic properties of P450-BM3 is Thr-268.[180] The Thr-268 residue is conserved in many P450s and is located in the distal side of the heme cofactor on the I-helix. Thr-268 has been accredited with playing critical roles in protonation, oxygen activation, and substrate binding.[180] In P450-CAM a similar residue, Thr-252, has been identified as critical for dioxygen activation.[181] The introduction of a T252A mutation engendered a P450-CAM mutant with little hydroxylation activity towards P450-CAM's natural substrate, camphor. It was reported that the T252A mutation increased solvent accessibility into the active site which is suspected to destabilise dioxygen binding, promoting hydrogen peroxide and water production instead of hydroxylation.[181] Thr-268 in P450-BM3 performs the same role as Thr-252 in P450-CAM, behaving as a proton-shuffling residue, controlling solvent diffusion and activating dioxygen.[180] It was demonstrated that Thr-268 is important in substrate recognition, as was illustrated when the metabolism of several peroxyacids by

P450-BM3 was revealed to be governed through a substrate recognition pathway that is mediated by residues Thr-268 and Phe-87.[182]

An understanding of the key roles of individual amino acids, and the intricate network of molecular interactions between the protein and the substrate, is paramount for understanding enzymatic function. However, the problem of clarifying the interdependency of single amino acids, on account of the evolutionary history of the enzyme, drastically obscures the precise interconnectivity between the biophysical, kinetic, thermodynamic, and structural properties of an enzyme. This principle was illustrated, for example, by the difficulty associated with unravelling the relationship between the highly conserved Phe-393 residue and the redox potential in P450-BM3. Nonetheless, a structural understanding of the key components governing an enzyme's function provides a platform from which novel functionality can be engineered through the introduction of intelligent modifications.

2.3 REDESIGN OF NATURALLY OCCURRING HEMOPROTEINS TOWARDS CARBENE TRANSFERASE ACTIVITY

2.3.1 Rational design of enzymes towards carbene transferase chemistry *via* directed evolution

In protein redesign, a naturally occurring protein provides a blueprint for engineering a novel protein with desired functionality, usually by varying rounds of mutagenesis.[183,184] The introduction of mutations into a gene, *via* mutagenesis, can be done randomly or selectively, depending on the nature of the desired gene product and the information known about the blueprint protein's structure. In *random mutagenesis* (**figure 28**) the mutations are completely random and are introduced to a gene *via* mutagenic chemicals, exposure to UV/X-ray radiation, or by error-prone PCE in which the conditions employed reduce the fidelity of the DNA-polymerase.[185] The variants generated from random mutagenesis are then cloned into an expression vector, expressed and characterised. The advantage of random mutagenesis is that it requires little-to-no structural knowledge of the blueprint protein. If structural knowledge of the protein is known, selective mutagenesis can be performed to i) target key residues responsible for the undesired functionality, and/or ii) introduce key mutations at certain residues to enhance desired functionality. A technique called *site-directed mutagenesis* can be employed to introduce specific changes into a DNA sequence, allowing control over where and which changes are introduced (**figure 28**). Site-direct mutagenesis is often used to investigate how a key residue, usually in the active site, participates in the functionality of the protein.[185] Another technique, called *saturation mutagenesis*, also targets a specific residue in the DNA sequence, but outputs a library of variants containing all the possible amino-acid substitutions at the identified position.[186]

Saturation mutagenesis can be used to identify optimal residues, at key positions, for desired functionality.

In protein engineering, a technique called *direct evolution* (DE) is used to generate large numbers of mutant libraries (**figure 28**).[83,93,187,188] DE is a process that mimics natural evolution, with the selective pressure being defined by the desired functionality. DE involves iterative rounds of mutagenesis to create mutant libraries that are screened to identify variants that exhibit the desired functionality. The best variant(s) are then amplified to serve as a/the template(s) in subsequent rounds of mutagenesis.[83,93,187,188] The isolation of iron-porphyrin carbene complexes,[189] and the analogous nature between porphyrin metallocarbenoids and the intermediate compound I in peroxidases/monooxygenases, provoked the question of whether hemoproteins that are proficient towards the formation of compound I could be selectively fine-tuned towards metallocarbenoid formation instead *via* DE.[190] If at all possible, another question concerning whether such an intermediate would be catalytically active towards carbene transfer chemistry, as has been reported for iron-porphyrin carbenes in the literature, would be raised.[190] The first experiments involved subjecting the hemoprotein, $P450_{BM3}$, to DE in attempts to generate the first heme-containing carbene transferase, with the selective pressure being activity towards a cyclopropanation reaction.[190]

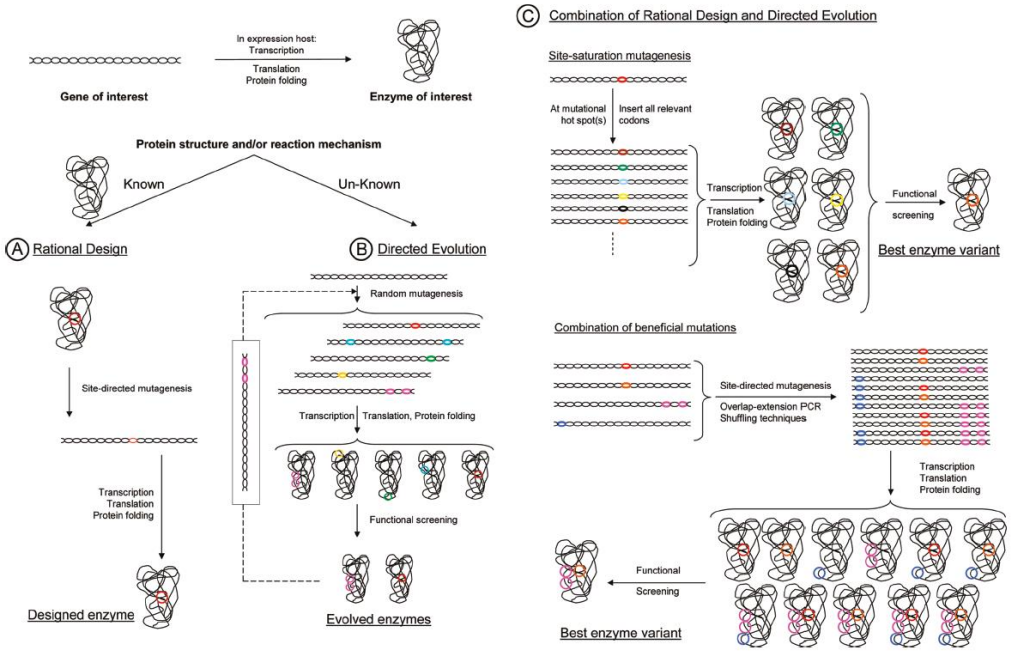


Figure 28: Schematic illustration of the different approaches to enzyme design (figure taken from Strohmeier *et al*).[93]

2.3.2 Cyclopropanation

The most common reactions involving carbene reagents are [2+1] cycloadditions to unsaturated C-C bonds.[42] [2+1] cycloadditions are cheletropic, with two new σ-bonds being formed on the carbene carbon, across a C-C π-bond, to yield a three-carbon ring. The substrates are most commonly alkenes which yield cyclopropane products, and hence [2+1] cycloadditions involving alkenes are commonly called cyclopropanations.[42] As discussed in *section 1.4.1* of chapter 1 the vast majority of cyclopropanation reactions involving a carbene intermediate require a stabilising agent, such as an organometallic complex, to circumvent the lack of chemo-, regio- and stereoselectivity exhibited by free carbene reagents.

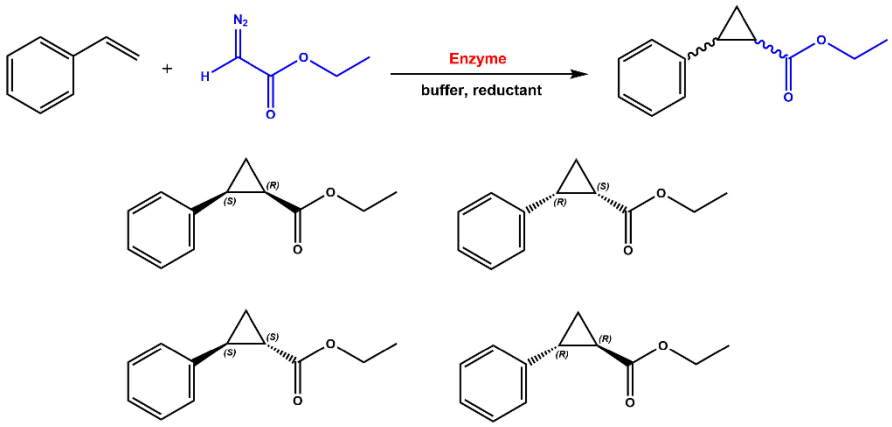


Figure 29: The cyclopropanation reaction between ethyl diazoacetate (blue) and styrene used for detecting the carbene transferase activity and of hemoproteins. The four possible stereoisomers of the reaction are also presented.

In January 2013, Coelho and Arnold reported the activity of a selection of mutant cytochrome P450s engineered (*via* DE) towards the cyclopropanation reaction between styrene and ethyl diazoacetate (EDA) (**figure 29**).[190] Ninety-two variants of $P450_{BM3}$ were screened for cyclopropanation activity, with the ten best variants being taken forward for subsequent rounds of evolution. Interestingly, a single mutation (T268A) generated a highly active and stereoselective mutant, with product yields and cis:trans ratios of 65% and 1:99 (the (*S,S*)-product was favoured).[190] Another mutant, BM3-CIS, 13 mutations away from WT-$P450_{BM3}$, demonstrated reasonable activity towards cyclopropantion (with product yields of 40%), but with inverted stereoselectivity, exhibiting a cis:trans ratio of 71:29. The introduction of an additional mutation to generate the variant BM3-CIS-(T438S) elevated the product yield to 59% and the cis:trans ratio to 92:8. Interestingly, the variant BM3-CIS-(Q263A) afforded a cis:trans ratio of 9:91, demonstrating succinctly the sensitive relationship between the enzyme's sequence and the resultant stereoselectivity. All variants were tolerant to electron-withdrawing and

electron-donating olefins, as well as α-methyl styrene. Furthermore, in the presence of CO all the variants were inhibited and were characterised by the presence of an intense absorption in the UV-VIS spectra (λ_{MAX}) at 450 nm, a diagnostic feature of cytochrome P450 Fe^{II}-CO complexes. The importance of this observation is it demonstrated the discerned carbene transferase activity was ceased once the vacant site of the heme was occupied, signalling that the iron-porphyrin cofactor is critical for carbene transferase activity.[190]

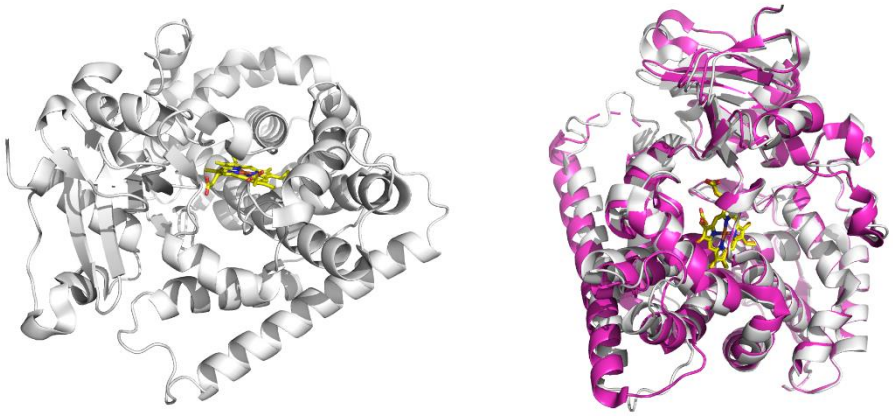


Figure 30: (**left**) The crystal structure of the P450-BM3 carbene transferase mutant P411-CIS (white, heme is show in yellow, *PDB: 4H23*), and (**right**) the overall of the crystal structures for P411-CIS (white, heme is show in yellow, *PDB: 4H23*) and P450-BM3 (blue, heme is show in orange, *PDB: 2IJ2*)

The generation of P450s active towards cyclopropanation (after a single mutation) demonstrated the powerful methodological approaches available for unlocking the intrinsic, but unexplored in nature, chemistry of heme cofactors. Another report by Coelho *et al*, published in August 2013,[191] investigated the nature of the axial ligand coordinating to the heme cofactor, and demonstrated its importance in governing heme functionality. A cysteine to serine mutation (C400S) generated a library of cytochrome mutants characterised by a blue shifted Fe^{II}-CO peak in the electronic spectra from 450 to 411 nm; the C400S mutant were therefore named the cytochrome P411s.[191] The P411s possessed an increased resting state Fe^{II}/Fe^{III} reduction potential and a distinctive lack of monooxygenation activity. Introducing the C400S mutant to BM3-CIS resulted in the *cis*-selective P411-CIS (**figure 30**), a variant proficient in catalysing the cyclopropanation reaction between styrene and EDA with reported product yields and enantioselectivities of 74% and 99% respectively.[191] In October 2014, Heel *et al* introduced the Cys→Ser mutation to four other hemoproteins, CYP119-(C316S), CYP153A6-(C363S), TxtE-(C356S) and $P450_{CAM}$-(C357S), and created a new library of four serine-ligated but structurally distinctive hemoproteins.[192] The authors reported the Cys→Ser mutation resulted in elevated cyclopropanation activity for all four mutants under anaerobic and whole *E. coli* cell conditions (except TxtE-(C356S) under anaerobic conditions). In addition, a stereochemical outcome for each mutant was

different, reflecting the author's hypothesis that the different active site volumes and topologies for each parent mutant would influence the stereochemical distribution of the resultant products.[192]

A critical step in P450 monooxygenation chemistry is the cessation of the molecular oxygen bond, which is facilitated by the weakening of the O=O bond through the donation of electron density from the iron into the π^*_p orbitals of molecular oxygen.[104,161] The observed diminishment in monooxygenation chemistry accompanying the Cys→Ser mutation was therefore attributed to the differing electronic properties of a thiol group relative to a hydroxyl group. The SH group of cysteine is more polarisable and therefore a stronger electron-donating group, compared to the OH group of serine.[104,161] The SH group provides more electron density to the iron, providing more electron density to be donated to dioxygen's π^*_p orbitals.[104,161] The OH group of serine, being a weaker electron-donating group, provides less electron density to iron and consequently hinders the cessation of the O=O bond. As monooxygenation chemistry competes with carbene transfer chemistry, it was further theorized that mutating the serine in the P411 mutant library to an even less electron-donating group would further diminish monooxygenation activity while concurrently enhancing carbene transfer chemistry. In 2014, Wang *et al* investigated the influence of multiple axial ligands on the functionality and catalytic activity of $P450_{BM3}$.[193] The amino acids screened were alanine, histidine, methionine, tyrosine and serine, in addition to another mutation, T268A (a mutation previously found to be highly beneficial for cyclopropanation), to generate five mutants called T268A-axX. After examining all mutants for cyclopropanation activity, the histidine-ligated $P450_{BM3}$-(T268A-axH) mutant was reported to be the best variant, catalysing the reaction between *N,N*-diethyl-2-phenylacrylamide and EDA, in whole *E. coli* cells, with a product yield, enantioselectivity and cis:trans ratio of 81%, 42% and, 94:6 respectively.[193] Additional rounds of directed evolution resulted in the identification of two other variants that showed improved enantioselectivity without loss of reactivity. The most influential mutations were subsequently combined to produce a final variant, T268A-axH-L437W-V78M-L181V (named "BM3-Hstar").[193]

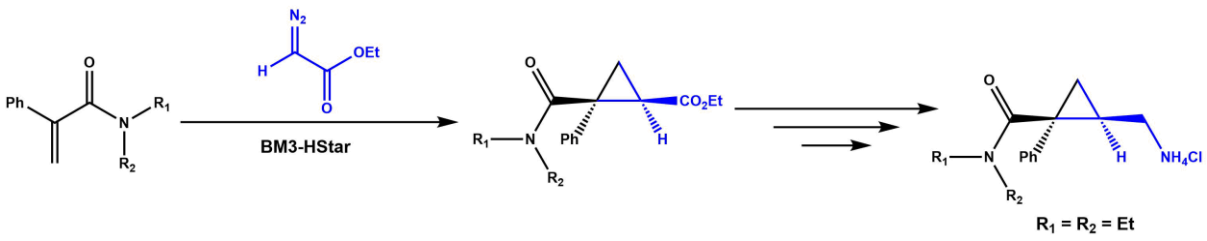


Figure 31: The cyclopropanation of acrylamides by the P450-BM3 mutant HStar; the reaction was used in the synthesis of the pharmaceutically important antidepressant levomilnacipran.[194]

BM3-Hstar exhibited substantially enhanced carbene transferase activity, with reaction yields of the desired product exceeding 92%, alongside reported values of 92% and 2:98 for the reaction enantioselectivity and diastereoselectivity respectively.[193] In a later study, published in June 2014 by Renata *et al*, BM3-Hstar was demonstrated to be an extremely proficient carbene transferase, catalysing the cyclopropanation of numerous acrylamide substrates (**figure 31**), with product yields, enantioselectivities and diastereoselectivities of 75-90%, 60-99%, and 97:3 respectively.[194] The substrate scope of BM3-Hstar was also expanded to acrylamides functionalised at either the aryl or amine moiety of the molecule and was later employed in the gram-scale synthesis of levomilnacipran (**figure 31**).[194]

The successful engineering of carbene transferases by utilising cytochrome P450 scaffolds raised the question of whether other hemoproteins could be utilized, and in December 2014 Bordeaux *et al* reported a library of myoglobin variants engineered towards carbene transferase activity.[195] One variant, Mb(H64V,V68A) (**figure 32**), was reported to catalyse the cyclopropanation of aryl-olefins with yields, enantioselectivities and diastereoselectivities of up 99% respectively for the trans-(*S,S*) product. Mb(H64V,V68A) also exhibited good substrate scope, displaying catalytic proficiency towards several *ortho-*, *meta-*, and *para-* functionalised styrene substrates. Interestingly, a Hammett plot of *para*-functionalised styrenes revealed a positive relationship between the electron-donating ability of the para-substituent and the rate of the cyclopropanation reaction (ρ_+ = -0.34 ± 0.07).[195] This observation is consistent with the suggestion that cyclopropanation reactions proceed through an electrophilic, iron localised metallocarbenoid intermediate.

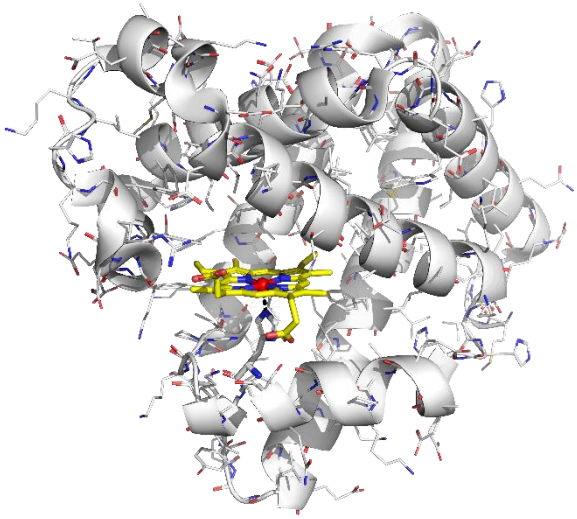


Figure 32: The crystal structure of the sperm-whale myoglobin mutant Mb(H64V,V68A) (white, the heme is shown in yellow; *PDB: 6M8F*).

In April 2017, Tinoco *et al* expanded the substrate scope of Mb(H64V,V68A) to include new carbene precursors.[196] The carbene precursor 2-diazo-1,1,1-trifluoroethane (CF_3CHN_2) was prepared *in situ* from $CF_3CH_2NH_2$.HCl and $NaNO_2$ and was utilised in the synthesis of trifluoromethyl-substituted cyclopropanes (**figure 33**). The reaction was catalysed by isolated Mb(H64V,V68A), with product yields, enantioselectivities and diastereoselectivies of 22%, >99.9% and 98.5% respectively. When the reaction was conducted in whole *E. coli* cells overexpressing Mb(H64V,V68A), product yields, enantioselectivities and diastereoselectivies between 43-82%, 97-99.9% and >99.9% were reported for an assortment of functionalised styrene substrates.[196]

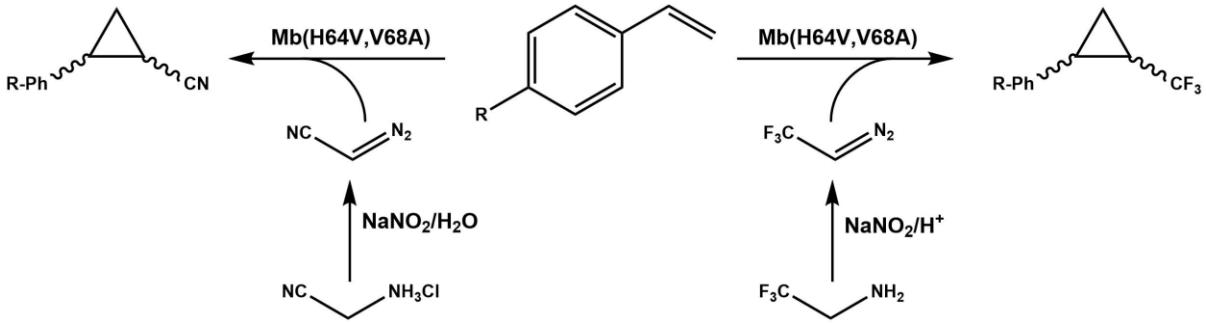


Figure 33: The Mb(H64,V68A)-catalysed cyclopropanation of styrene by *in situ* generated aliphatic diazo compounds (**left**) diazoacetonitrile[197] and (**right**) 2-diazo-1,1,1-trifluoroethane.[196]

Later that year, in October 2018, the substrate scope of Mb(H64V,V68A) was further expanded, by Fasan *et al*, to include the synthesis of nitrile-functionalised cyclopropanes using *in situ* generated diazoacetonitrle as the carbene precursor (**figure 33**).[197] A range of nitrile-functionalised cyclopropane products were obtained using *E. coli* overexpressing Mb(H64V,V68A), with product yields ranging from 50-99%, and enantioselectivities and diastereoselectivities being consistently >99.9% for the trans-(*S,S*)-product. The nitrile-functionalised cyclopropanes are of particular interest as synthetic targets because of the reaction scope of nitrile groups.[197]

In May 2018, Villarino *et al* developed an artificial hemoprotein using the lactococcal multidrug resistance regulator (LmrR) protein as the metalloenzyme scaffold.[198] Incorporation of the heme B cofactor (followed by UV-VIS spectroscopy) afforded a modestly proficient metalloenzyme capable of catalysing cyclopropanation reactions (products yields of up to 25%), but with low enantioselectivity (17% for the trans-(*R,R*) product). Introducing a single mutation, M8A, afforded the variant LmrR-(M8A), which, despite only showing a moderate increase in product yield relative to WT-LmrR (45% vs 25% respectively), enhanced the enantioselectivity relative to WT-LmrR (51% vs 17% for the trans-(*R,R*) product respectively).[198]

In November 2016, Bajaj *et al* explored the relationship between the stereoselectivity of cyclopropanation reactions and an enzyme's structure by engineering a library of myoglobin variants with stereoselectivities complementary to Mb(H64V, V68A).[199] The authors had previously recognized five key amino acids in the distal pocket of myoglobin that influenced the activity and selectivity of myoglobin in carbene transfer reactions, namely Leu29, Phe43, His64, Val68, and Ile107.[199] The mutation of a distal histidine residue to a valine (H64V) was demonstrated to significantly enhance the cyclopropanation activity of myoglobin, which the authors attributed to an increased accessibility of the heme pocket to the reactants.[199] The authors generated ten notable myoglobin mutants, including two trans-(*S,S*) selective variants and eight trans-(*R,R*)-selective variants. The substrate scope was subsequently screened for the (*R,R*)-selective myoglobin variants, against *ortho*-, *meta*-, and *para*- functionalised styrenes, with reported product yields, enantioselectivities and diastereoselectivies of >95% (excluding *p*-CF_3; 45%), 45-99% and 92-99% respectively.[199]

In February 2018, Knight *et al* further explored the relationship between the stereoselectivity of biocatalytic cyclopropanation reactions by screening a diverse range of native and engineered hemoproteins.[200] The authors investigated the cyclopropanation of unactivated terminal alkene substrates catalysed by variants isolated from distinctive organisms. Two naturally occurring enzymes, WT-*Aeropyrum pernix* protoglobin (ApePgb) and WT-*Rhodothermus marinus* nitric oxide dioxygenase (RmaNOD), were identified as suitable candidates for engineering as both WT-enzymes exhibited low but measurable cyclopropanation activity towards 1-octene, catalysing the reaction with TTNs of 18 and 27 respectively.[200] Notably, WT-ApePgb and WT-RmaNOD displayed complementary diastereoselectivity, preferentially producing cis-(*1R,2S*) and trans-(*1S,2S*)-products respectively (**figure 34**). After subjecting RmaNOD and ApePgb to directed evolution, the two variants RmaNOD-(Q52V) and ApePgb-(W59A,Y60G,F145W) showed drastically improved activity towards the cyclopropanation of 1-octene. Two addition mutants, P411-UA-(V87C) and P411-UA-(V87F) (derived from a previous variant P411-CIS-(L437F,T438Q,L75Y,L181I), called P411-UA) were identified as proficient catalysts towards the cyclopropanation of 1-octene, affording the separate cis-(*1S,2R*) and trans-(*1R,2R*) stereoisomers (**figure 34**). Overall, the authors reported four mutant hemoproteins proficient towards the cyclopropanation of 1-octene with complementary stereoselectivities (**figure 34**). The enantioselectivities and diastereoselectivities of the four mutants were: 97% and 95:5 for cis-(*1S,2R*), P411-UA-(V87C); 99% and 89:11 for cis-(*1R,2S*), ApePgb-(W59A,Y60G,F145W); 96% and 4:96 for trans-(*1R,2R*), P411-UA-(V87F); and >99.9% and 1:99 for cis-(*1S,2S*), RmaNOD-(Q52V).[200]

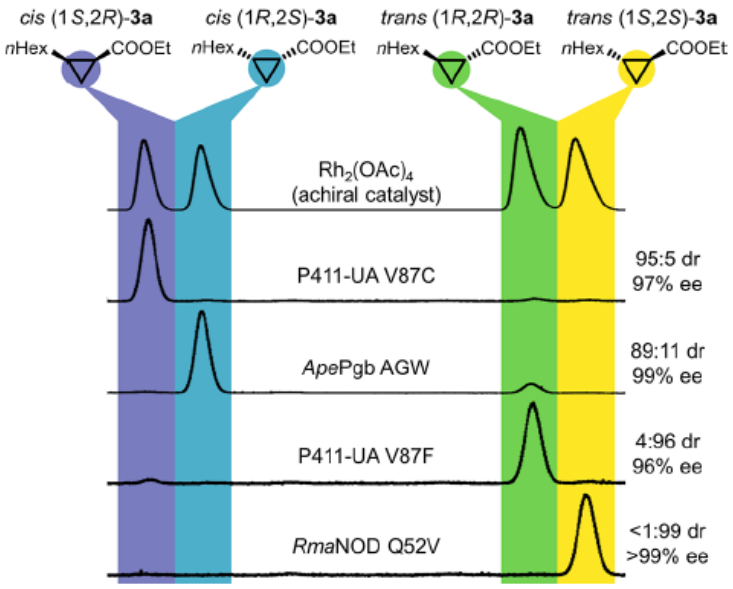


Figure 34: Stereoselective whole cell catalysed cyclopropanation of the aliphatic alkene 1-octene and EDA to obtain each of four stereoisomers of the cyclopropane product, two serine-ligated cytochrome P450-BM3 variants (P411-UA-V87C and P411-UA-V87F), *Aeropyrum pernix* protoglobin-(W59A Y60G F145W) (ApePgb AGW), and *Rhodothermus marinus* nitric oxide dioxygenase-(Q52V) (RmaNOD Q52V); reactions were contrasted against a racemic mixture synthesised using a rhodium acetate dimer ($Rh_2(OAc)_4$) catalyst (figure taken from Knight *et al*).[200]

In addition to the axial ligand and the first-coordination sphere, the identity of the metal and its coordinating ligand(s) also govern the functionality and catalytic activity of a metalloenzyme. As a variety of non-iron metalloporphyrin catalysts are well-established abiological catalysts, a considerable amount of research has been devoted to incorporating non-natural and/or semi-synthetic cofactors into metalloenzymes in efforts to enhance reactivity and selectivity.[201–208] In June 2016, Key and Hartwig[209] generated an eight variant library of apo-mOCR-Myo-H93X, with each variant characterised by a distinctive axial ligand (H93X) coordinating to the heme metal. Each variant was reconstituted with nine different porphyrin cofactors which differed by the coordinated metal, including Fe, Co, Cu, Mn, Rh, Ir(Cl), Ir(Me), Ru, and Ag.[209] The authors rapidly screened the 72 artificial metalloenzymes and reported activity profiles distinct from wild-type myoglobin. The variants were screened for cyclopropanation activity, employing β-methylstyrene and EDA as substrates. The authors reported that the variants reconstituted with Fe-PPIX exhibited negligible activity, in contrast to the Rh-, Ru-, and Ir-PPIX reconstituted variants. The variants reconstituted with Ir(Me)-PPIX were the most active, exhibiting TONs of >20 with enantioselectivity and diastereoselectivity ratios, for the trans-isomer, of 70:30 and 33:1 respectively.[209] The authors then expanded the substrate scope of the Ir(Me)-PPIX-reconstituted variants by investigating the cyclopropanation of 1-octene. Although unselective C-H insertion products were detected, the variants successfully catalysed the

cyclopropanation reaction between EDA and 1-octene with TONs of >40 and enantioselectivity and diastereoselectivity ratios of 91:9 and 40:1, for the trans-isomer, respectively.[209]

In April 2017, Key *et al* continued investigated how incorporating an iridium containing analogue of heme B, Ir(Me)-protoporphyrin IX, into a hemoprotein scaffold influenced the catalytic activity.[210] The authors derived a library of iridium-containing P450 variants form the parent WT-CYP119 hemoprotein and screened for cyclopropanation activity. A double mutant, CYP119(+), was reported to yield the cis-cyclopropane product with an enantioselectivity of 73%, while a triple mutant, CYP119(-), afforded the trans-cyclopropane product with an enantioselectivity of 74%.[210] The authors utilised CYP119(+) and CYP119(-) to generate another library of 100+ variants and reported one variant, CYP119(+)-(L155W), that catalysed the reaction between to give the (*1S,2R*)-cis cyclopropane product with a yield, enantioselectivity, and diastereoselectivity of 80%, 98%, and 90:1 respectively. The same reaction catalysed by of another variant, CYP119(-)-(V254L) gave a product yield, enantioselectivity and diastereoselectivity of 74%, 98%, and 73:1%, respectively. Perhaps more importantly, the Ir(Me)-PPIX containing CYP119 variants catalysed the cyclopropanation of a large variety of alkenes, displaying catalytic activity towards terminal and internal, activated and unactivated, electron-rich and electron-deficient, conjugated and nonconjugated alkenes as well as 1,2-disubstituted styrenes.[210]

Also, in April 2017, Wolf and co-workers reported several sperm whale myoglobin variants reconstituted with ruthenium mesoporphyrin IX (RuMPPIX) (mesoporphyrin differs from protoporphyrin IX by the complete saturation of the vinyl groups).[211] Reconstituted WT-Mb and the three variants containing the H64X mutation, with X equalling aspartic acid (D), aniline (A) and valine (V) were demonstrated to be very weakly active towards the cyclopropanation of styrene and EDA with reaction yields from 0.6 - 3.9 % compared to 26% for free RuMPPIX respectively (the Mb catalysts were, however, moderately proficient at N-H insertion [see below]).[211] The reaction between *para*-methoxystyrene and EDA, catalysed by RuH64AMb, was moderately proficient, however, with a reported reaction yield of 36%. The authors did report an ESI-MS for the RuMb prior to and after the addition of EDA, with dominant peaks present at 17995.4 and 18081.4 m/z respectively corresponding to a mass difference of 86 m/z, the anticipated mass change expected to be observed upon the addition of EDA; however, the authors were not capable of discerning whether EDA was localised on the metal or whether the carbene had inserted into the porphyrin/protein.[211]

At the same time, Sreenilayam *et al* investigated the influence of the metal on the catalytic activity of the previously reported Mb(H64V,V68A) mutant.[212] Apo-Mb(H64V,V68A) was reconstituted with a protoporphyrin IX (PPIX) cofactor coordinated to either manganese, cobalt, ruthenium, or iridium,

generating a series of five reconstituted holoenzymes (all four metal substitutions were referenced against Fe(PPIX)). In addition, three M(PPIX)-Mb(H64V,V68A,H93F) (where M = Fe, Mn, Ir) variants were also prepared and characterised. The authors reported that the variant Fe(PPIX)-Mb(H64V,V68A) remained the best biocatalyst for the cyclopropanation reaction between EDA and styrene, with neither the metal substitution nor the introduction of the additional H93F mutation producing any appreciable improvements in product yield or stereoselectivity.[212] The authors did, however, discover some interesting functionality with other variants towards X-H insertion chemistry (discussed below). In September 2017, the same authors reported a variant of Mb(H64V,V68A) reconstituted with iron-chlorin e6 (Fe(Ce6) (**figure 35**)) was highly proficient towards carbene transferase activity, catalysing the reaction between styrene and EDA with a reported product yield, enantioselectivity and diastereoselectivity of >99% for all values respectively.[213] Importantly, although Fe(Ce6)-Mb(H64V,V68A) was less thermally stable than its Fe(PPIX) analogue, unlike the vast majority of metalloenzyme variants proficient towards cyclopropanation chemistry the Fe(Ce6)-Mb(H64V,V68A) variant is highly proficient under aerobic conditions and is capable of catalysing the cyclopropanation reaction of a diverse collection of functionalised styrenes with reported product yields, enantioselectivities and diastereoselectivities of 70->99%, 92-98% and 93-99% for all values respectively.[213]

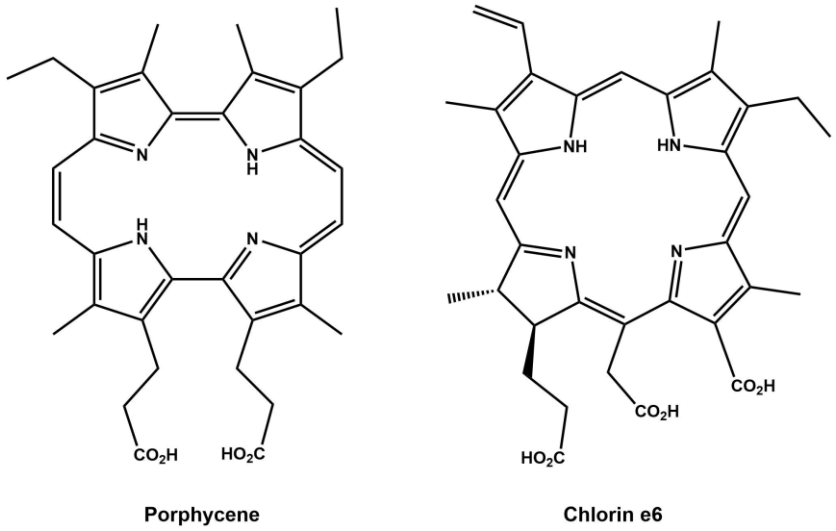


Figure 35: The chemical structure of the tetrapyrrole derivatives porphycene (left) and chlorin e6 (right).

The above work highlighted that for inorganic cofactors both the metal and the coordinating ligand(s) are important in catalysis. This principle was wonderfully demonstrated by Oohora *et al*, in November 2017, who reported that WT-myoglobin reconstituted with iron porphycene (rMb (**figure 35**)) afforded a metalloenzyme 615-fold faster than WT-myoglobin.[214] The authors reported that rMb was a

proficient cyclopropanation biocatalyst, affording the trans-(*R,R*)-product at 99% with no modifications introduced into the WT scaffold. The authors also reported kinetic parameters for the reaction and concluded that rMb presented a 26-fold enhanced catalytic efficiency (k_{cat}/K_M) relative to WT-Mb.[214]

2.3.3 X-H insertions

2.3.3.1 X-H insertions where X = N, S, Si, B

As discussed in section chapter 1, an X-H (X = C, N, B, Si) insertion reaction is the insertion of a chemical species (R) into a pre-existing σ-bond (X-H), resulting in the formation of a new σ-bond (X-R). Carbenes are good reagents for insertion reactions (and are particularly useful for C-H bond), but require a stabilising agent, such as an organometallic complex, to overcome problems associated with the lack of chemo-, regio- and stereo- selectivity. Having engineered myoglobin and P450s variants towards cyclopropanation activity, the question of whether reactivity and specificity towards X-H σ-bonds instead of olefin π-bonds was raised.

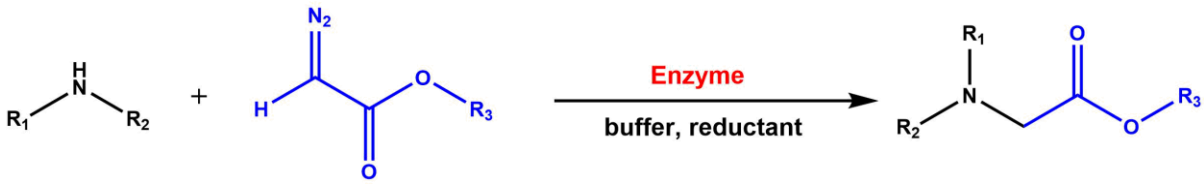


Figure 36: The general reaction scheme for an enzymatic catalysed carbene insertion reaction between an N-H σ-bond and a diazo compound (blue).

The first carbenes insertion reactions reported concerned insertions into N-H σ-bonds (**figure 36**). N-H insertion reactions are important in organic synthesis because the reactions can divide complex synthetic pathways into multiple pathways and allow for a non-linear synthetic approach for complex molecules. In early 2014, Wang *et al* reported a cytochrome P450 variant proficient towards N-H insertion chemistry.[215] The variant H2-5-F10 (15 mutations away from WT-$P450_{BM3}$) was thermally stable and catalysed the N-H insertion reaction between aniline and EDA under anaerobic conditions and when expressed in whole *E. coli* cells (47% and 38%). H2-5-F10 is tolerant to primary and secondary anilines, and anilines with either *ortho*- or *para*- substitutions.[215] H2-5-F10 is also chemo- and regioselective towards N-H insertion: in the presence of aniline and styrene N-H insertion/cyclopropanation yields were 63% to 9% respectively, and when *para*-vinylaniline was employed as the substrate a 32% yield for the N-H insertion product was accompanied by only trace amounts of the cyclopropanation product.[215] The authors reported that the introduction of a C400S mutation decreased activity towards N-H insertion, contrasting the reported observation that a C400S

mutation enhances cyclopropanation activity. In addition, only single-insertion products were detected in reactions catalysed by H2-5-F10, circumventing a common problem in amine chemistry (single:double insertion ration is 3:1 when the same reaction is catalysed by hemin).[215]

In early 2015, the Fasan group engineered myoglobin variants towards N-H insertion.[17] The same variant reported for cyclopropanation, Mb(H64V,V68A), demonstrated a 61% reaction yield for the reaction between aniline and EDA. For comparison, WT-Mb exhibited a product yield of 21% for the same reaction. Mb(H64V,V68A) also exhibited good substrate scope, reacting with a variety of *ortho*- , *meta*-, and *para*- functionalised anilines with yields and TTNs between 67-99% and 540-7000 respectively.[17] Another variant, Mb(L29A, H64V), tolerated the N-H insertion using both EDA and ^{t}BuDA, with a TTN of 4000 and for the reaction between EDA and *N*-methylaniline and 8000 for the reaction between ^{t}BuDA and aniline. As mentioned above, Wolf and co-workers reported several sperm whale myoglobin variants reconstituted with ruthenium mesoporphyrin IX (RuMPPIX). Although the reconstituted WT-Mb and the three variants containing the H64X mutation (X = D, A and V) were very weakly active towards cyclopropanation (especially when compared to free RuMPPIX), the RuMb variants demonstrated moderate catalytic proficiency towards the N-H insertion of aniline, with reaction yields between 32-52% observed.[211] The variant RuH64AMb was a good improvement relative to WT-Mb and RuMb (52% vs 21 and 36% respectively), but only provided a marginal improvement over free RuMPPIX (52% vs 48% respectively). The variant RuH64AMb is also capable of catalysing several *para*-functionalised anilines, but with decreased reactivity (product yields ranged from 2.9 – 45 %).[211]

Having demonstrated that hemoproteins can be engineered towards N-H insertion chemistry, the question of what other X-H functional groups could be targeted as potential substrates was raised. In February 2015, Tyagi *et al* reported the first example of a biocatalytic methodology for the synthesis of thioethers *via* an intermolecular carbene insertion into a thiol S-H bond (**figure 37**).[216] Employing the sperm whale myoglobin scaffold previously used for engineering, a library of mutants were generated *via* site-directed mutagenesis and screened for reactivity using thiophenol and EDA as substrates. The previously reported variant, Mb(H64V,V68A),[17,195] catalysed the reaction but with modest activity compared to the cyclopropanation and N-H insertion activity previously reported. Screening the initial library, two single point mutations, L29A and F43W, resulted in an elevated S-H insertion activity relative to WT-Mb. Subsequent mutations resulted in a second generation of variants, with one mutant, Mb(L29A,H64V), eclipsing all the other variants with regards to S-H insertion activity (TTN = 2750).[216] Mb(L29A,H64V) was screened for its substrate scope, and reaction yields ranging from 60% to >99% were reported for an assortment of functionalised thiophenols.

Three additional α-diazo precursors, *tert*-butyl diazoacetate, cyclohexyl diazoacetate, and benzyl diazoacetate, were screened using Mb(L29A,H64V) and product yields of 95, 95 and 75% were reported respectively (**figure 37**).[216] The authors considered the effect of the two mutations on the observed catalytic activity of Mb(L29A,H64V). The mutation L29A is postulated to facilitate the approach of the thiol nucleophile to the iron-bound metallocarbenoid, and the mutation H64V is postulated to enhance the catalytic activity by increasing the accessibility of the heme active site to both the thiol and α-diazo ester substrate. These propositions are supported by the observation that introducing both mutations to WT-Mb enhanced its catalytic activity towards S-H insertion. The same mutations also activate WT-Mb towards N-H insertion reactions, suggesting N-H and S-H insertion reactions proceed *via* a similar reaction mechanism.[216]

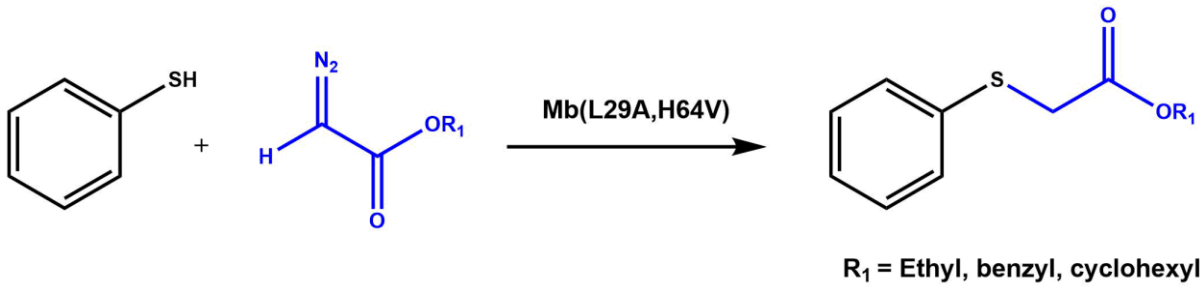


Figure 37: The general reaction scheme for the thiol insertion reaction catalysed by Mb(L29A,H64V).[216]

The successful development of biocatalytic methodologies to synthesize novel N-C and S-C σ-bonds raised the exciting question of whether more exotic X-H groups could be targeted. This question was answered in November 2016 by Kan *et al*, who reported an engineered enzyme capable of converting Si-H σ-bonds to Si-C σ-bonds, *via* a carbene insertion mechanism.[217] Initially, free hemin and a selection of cytochrome *c* variants were screened for Si-H insertion activity using phenyldimethylsilane and ethyl 2-diazopropanoate as substrates (**figure 38**). A wild-type cytochrome *c* isolated from *Rhodothermus marinus* (*Rma* cyto *c*) exhibited low activity but remarkable enantioselectivity (97%). For this reason, *Rma* cyto *c* was subjected to directed evolution to enhance its catalytic activity towards Si-H insertion.[217] A first-generation *Rma* cyto *c* mutant, containing a M100D mutation, displayed a 12-fold improvement in activity compared to the WT-enzyme. The *Rma*(M100D) variant was then selected as the enzyme scaffold for subsequent directed evolution.[217] Two amino acids in close proximity to the iron heme centre, Val-75 and Met-103, were identified as possible candidates for mutation, and sequential site-saturation mutagenesis at these positions resulted in the triple mutant *Rma*-(V75T,M100D,M103E) (called *Rma*-TDE (**figure 38**)). *Rma*-TDE showed drastically improved catalytic activity towards Si-H carbene insertion while retaining the high enantioselectivity observed for the WT-enzyme (TTN > 1500; and >99.9% enantioselectivity).[217] The two mutations, V75T

and M103E, were proved to activate WT-*Rma* cyt *c* towards Si-H insertion individually, but provided a synergistic effect when combined in the triple-mutant. *Rma*-TDE was then investigated for Si-H substrate scope, and was demonstrated to turnover a diverse variety of aryl-silane substrates when *E. coli* cells over-expressing *Rma*-TDE were employed as catalyts. Although the TTNs for some aryl-silane substrates were relatively low the enantioselectivity reported for each reaction was consistently >95%.[217]

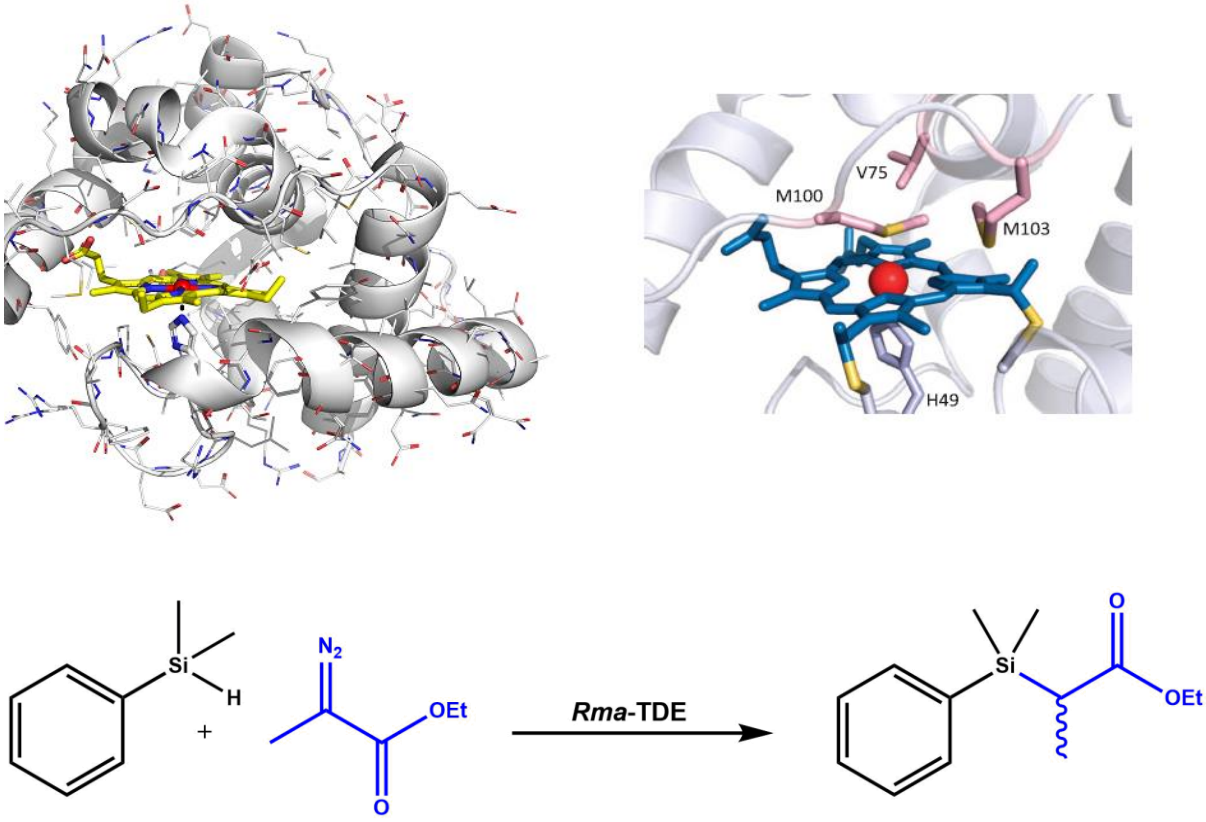


Figure 38: (**top left**) The crystal structure of the *Rhodothermus marinus* cytochrome *c* variant *Rma*-TDE (white, heme is shown in yellow; *PDB: 6CUK*), (**top right**) structure of the active site of *Rma*-TDE (picture taken from Kan *et al*) and (**bottom**) the general reaction scheme for the *Rma*-TDE catalysed Si-H insertion reaction.[217]

An interesting question concerning chemoselectivity arises when a molecule containing two differing X-H groups in its structure is employed as a substrate. As carbene transferases have previously shown proficiency towards N-H (and S-H) insertions, the chemoselectivity of *Rma* cyto *c* mutants was investigated using 4-(dimethylsilyl)aniline, a Si-H and N-H containing compound (**figure 39**).[217] WT-*Rma* exhibited no chemoselectivity towards N/Si-H insertion, affording the Si-C and N-C containing products in equal quantities.[217] The introduction of the M100D mutation to WT-*Rma*, yielding *Rma*(M100D), introduced a minor preference for Si-H insertion, affording the Si-C and N-C containing products in a ratio of 2:1. Introduction of the V75T mutation shifted the Si-C/N-C ratio to 7:1, while

the final triple-mutant variant, *Rma*-TDE, exhibited a high chemoselectivity in favour of Si-H insertion, with a reported Si-C/N-C ratio of 29:1, a >30 fold increase relative to WT-*Rma*.[217] The authors later demonstrated, through MD and QM/MM simulations, that the three mutations (V75T, M100D, M103E) affect the flexibility of the T98-E103 loop on the pro-*R*-face of the protein.[218,219] The surrounding protein residues on the pro-*S*-face are rigidified by an extensive network of hydrophobic contacts which blocks the channel between the substrate and heme active site. Silane binding is more favourable in the *Rma*-TDE triple mutant compared to WT-*Rma* because of the increased flexibility at the front of the T98-E103 loop.[218,219] The authors performed DFT calculations on the geometries of the insertion transition states in WT-*Rma* and *Rma*-TDE for both N-H and Si-H insertions, and discovered that the N-H and Si-H insertions proceed *via* geometrically distinct transition states on account of the differences associated with accommodating the substrate's approach towards the carbene.[218,219] It was noted that N-H carbene insertion has a lower energy barrier than Si–H carbene insertion when catalysed by free hemin, and therefore the authors concluded the observed chemoselectivity for Si-H carbene insertion was a result of the conformational changes in the front T98-E103 loop, which lowered the energy barriers of silane binding and the Si-H insertion transition state relative to WT-*Rma* cyto *c*.[218,219] The authors provided evidence for this conclusion by engineering another variant that is chemoselective towards N-H insertion.[218] The group rationalised that Si-H insertion could be disfavoured by carefully introducing key mutations to alter the conformation of the front loop and ultimately the binding of the silane. A novel mutant *Rma*(V75T,M100D,N80F,M99P,E103I) (abbreviated as *Rma*-TDFPI) was engineered and catalysted the N-H insertion reaction between 4-(dimethylsilane)-aniline and ethyl 2-diazopropanoate chemoselectively with product yields of 90% (**figure 39**).[218] The IPC of *Rma*-TDFPI was analysed using MD simulations and revealed a significant change in the front loop conformation in *Rma*-TDFPI compared to *Rma*-TDE. It was also emphasised that the mutations M99P and E103I work synergistically to disfavor the Si-H insertion reaction transition state by changing the conformational of the front loop and blocking the preferred binding conformation necessary for the silane to approach the IPC.[218]

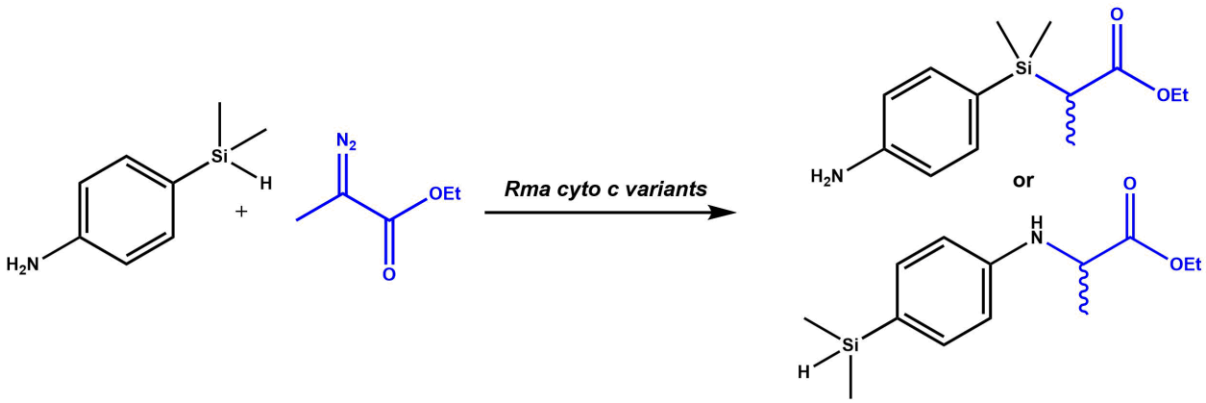


Figure 39: The possible mono-insertion products of the X-H carbene insertion when employing 4-(dimethylsilyl)aniline as the substrate and a *Rma* cyto *c* variant as the catalyst.

In December 2017, the Arnold group expanded the scope of biocatalytic X-H insertion methodologies to include B-H carbene insertion reactions (**figure 40**).[220] Employing the *Rma*-cytochrome *c* scaffold acquired from *Rhodothermus marinus*, and subjecting the wild type to directed evolution, the group engineered an enzyme proficient at catalysing B-C σ-bond formation when overexpressed in *E. coli* cells. Initial screenings using either free hemin or WT-*Rma* (named BOR^{WT}) incubated with an *N*-heterocyclic carbene borane (NHC-borane) and ethyl 2-diazopropanoate (Me-EDA) resulted in low-activities (TTNs = 80 and 120 respectively), but reasonable enantioselectivity for BOR^{WT} ([*R*]/[*S*] = 6).[220]

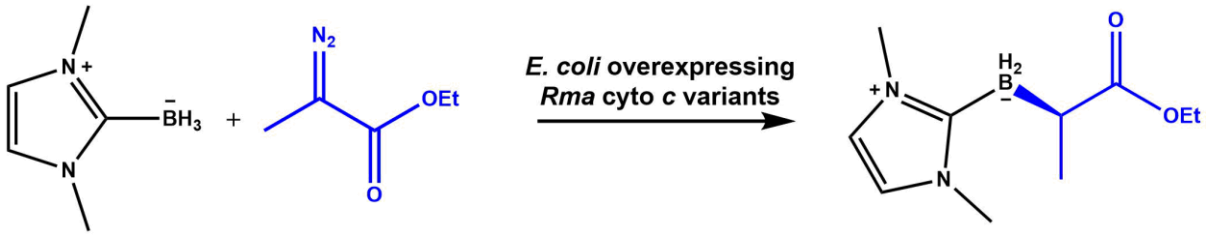


Figure 40: The *in vivo* borylation reaction between an *N*-heterocyclic carbene borane and ethyl-2-diazopropanoate catalysed by *E. coli* overexpressing *Rma* cyto *c* variants.

The introduction of the previously reported activating mutation, M100D, engendered the first generation BOR(M100D) variants that exhibited a drastic improvement in catalytic activity. The M100D mutation is known to increase the flexibility of the pro-*R*-face loop of BOR^{WT} (while simultaneously blocking the pro-*S*-face) and was demonstrated to be an important mutation in stabilising silane substrate.[218] The introduction of two additional mutations, V75R and M103T, resulted in the triple-mutant BOR^{R1} which catalysed the B-H insertion reaction with a reported TTN of 2,500 and a significant improvement in enantioselectivity ([*R*]/[*S*] = 39) (**figure 41**).[220] The increase in catalytic activity and *R*-stereoselectivity parallels the observations reported for *Rma*-TDE and Si-H insertion chemistry, indicating the conformational flexibility of the E98-M103 loop and the pro-*R*-face

of the heme active site plays an important role in binding boranes and stabilising the B-H carbene insertion transition state.[217,218,220] The substrate scope and whole cell compatibility of BOR^{R1} overexpressed in *E. coli* cells was explored, with TTNs and (*R*)-enantioselectivities ranging from 890-2760 and 92->99 across 16 different borane substrates.[220] Further mutagenesis studies engendered another variant, BOR^{WT}(V75R,M100D,M103D) (called BOR^{P2}) which catalysed the B-H insertion of borane-NHC and benzyl 2-diazopropanoate with a TTN and (*R*)-enantioselectivity of 4,200 and 92% respectively. Interestingly, *E.coli* cells harbouring BOR^{P2} also catalysed the B-H insertion of an NHC-borane substrate when (1-diazo-2,2,2-trifluoroethyl)benzene was employed as the carbene precursor substrate, opening up the potential towards highly coveted organofluoride synthesis (**figure 41**).[220]

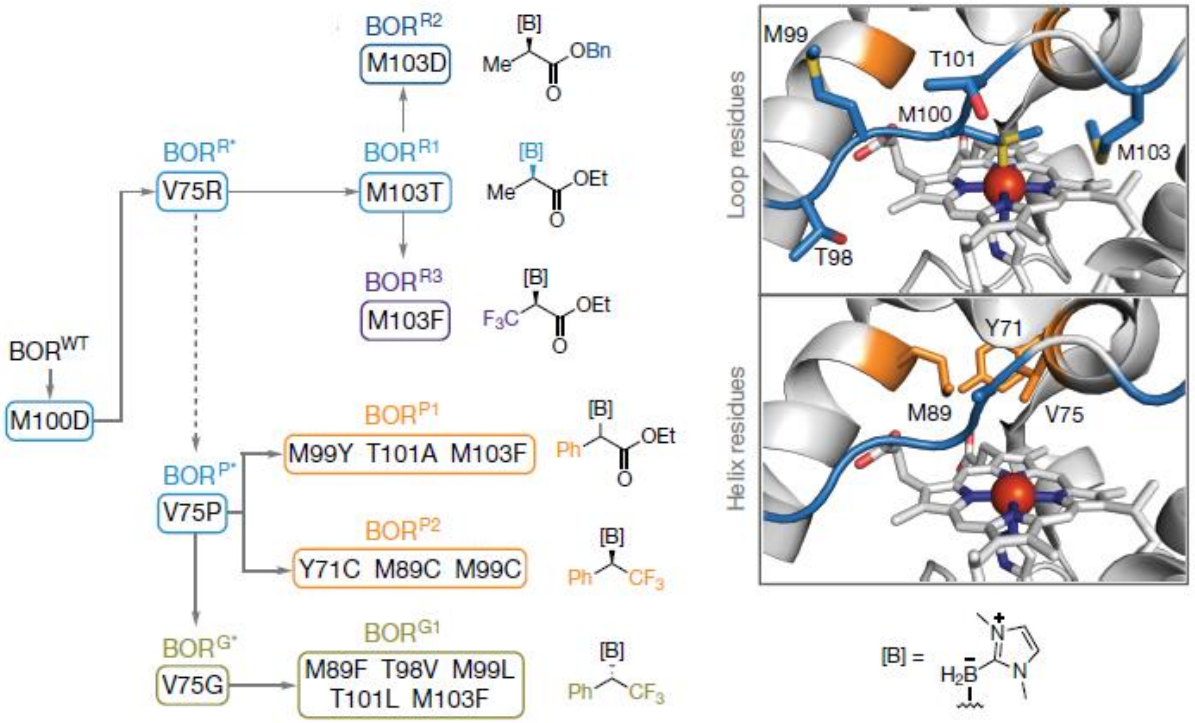


Figure 41: The expansion of the substrate scope of *in vivo* borylation reactions catalysed by various *Rma* cyto *c* variants accessed *via* direct evolution and the amino acid residues targeted during directed evolution (figure adapted from Kan *et al*).[220]

In 2019, Huang *et al* reported a novel variant, BOR-CF_3, five mutations away from WT-*Rma* cyto *c* (Y44I,V75S,M99A M100L,M103D) which catalysed the borylation reaction between (3-diazo-4,4,4-trifluorobutyl)benzene and an NHC-borane substrate (**figure 42**), affording the chiral α-trifluoromethylated organoborane product with a TTN of 2460 and an enantioselectivity of >95% (*R*-stereoisomer). The substrate scope was enlarged to include 12 α-CF_3-diazo precursors, producing TTNs and (*R*)-enantioselectivities ranging from 730-2900 and >96% respectively.[221] As with *Rma*-TDE,[218] the authors computationally demonstrated that the catalytic activity of BOR-CF_3 is attributable to the

dynamic flexibility of the protein caused by the introduction of key mutations in the surface loop.[221] This postulation was later corroborated by the computational studies of Alonso-Cotchico and Roelfes.[222]

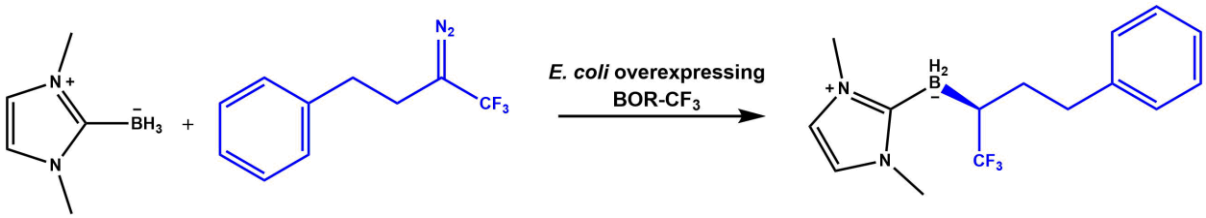


Figure 42: The general *in vivo* enantioselective borylation reaction between (3-diazo-4,4,4-trifluorobutyl)benzene and an NHC-borane substrate catalysed by *E. coli* overexpressing the *Rma* cyto *c* variant BOR-CF_3.

2.3.3.2 C-H insertions

One of the most challenging reactions in organic chemistry is the conversion of C-H σ-bonds to C-C σ-bonds. The functionalisation of C-H σ-bonds to C-X σ-bonds usually involve an intermediate step, called C-H activation, to active the C-H bond towards reactivity. The activation step usually involves a transition metal and the formation of a M-C organometallic intermediate. The direct conversion of C-H σ-bonds to C-C σ-bonds is one of the most highly coveted transformations, as direct functionalisation circumvents the C-H activation step. In 1981, Teyssié reported the rhodium-catalysed carbene insertion into the C-H σ-bonds of several paraffins.[67] The reaction mechanism involves the metal-catalysed decomposition of a diazo compound to yield a metallocarbenoid intermediate which then reacts with a C-H σ-bond in a concerted manner, proceeding *via* a three-membered cyclic transition state.[223] In 1993, Doyle demonstrated that dirhodium(II) carboxylates and carboxamides are extremely proficient catalysts for the intramolecular C-H insertion reaction of diazoacetoacetates.[223] Since the seminal paper in 1981, several copper, dirhodium and metalloporphyrin complexes have been established as proficient catalysts for numerous intra- and intermolecular C-H carbene insertion reactions.[64]

Having demonstrated that engineered carbene transferases are proficient catalysts for X-H carbene insertion reactions (X = N, S, B, and Si), the question of whether a carbene transferase capable of direct carbene insertion into C-H bonds was raised. It was mentioned above that in June 2016 Key *et al*[209] generated a library of apo- Myo-H93X variants each characterised by a distinctive mutation to the axial ligand (H93X) coordinating to the heme iron. Each variant was reconstituted with nine different porphyrin cofactors which differed by the coordinated metal, and included Fe, Co, Cu, Mn, Rh, Ir(Cl), Ir(Me), Ru, and Ag.[209] The authors rapidly screened the 72 artificial metalloenzymes and reported

cyclopropanation activity profiles that are distinct from wild-type myoglobin, as mentioned above. In addition to cyclopropanation activity, the authors screened the variants for intramolecular C-H insertion activity, employing methyl 2-diazo-2-(2-methoxyphenyl)acetate as both the substrate and the diazo compound (**figure 43**). Several variants reconstituted with Ir(Me)-PPIX exhibited remarkable C-H insertion activity (TONs of >60). The authors generated a novel library of Ir(Me)-mOCR-Myo mutants and screened across 7 substrates to access the activity and selectivities towards intramolecular C-H carbene insertion. It can be seen from **figure 43** that both the activity and selectivity of C-H carbene insertion towards distinctive substrates can be finely tuned by introducing key mutations into Ir(Me)-mOCR-Myo.[209]

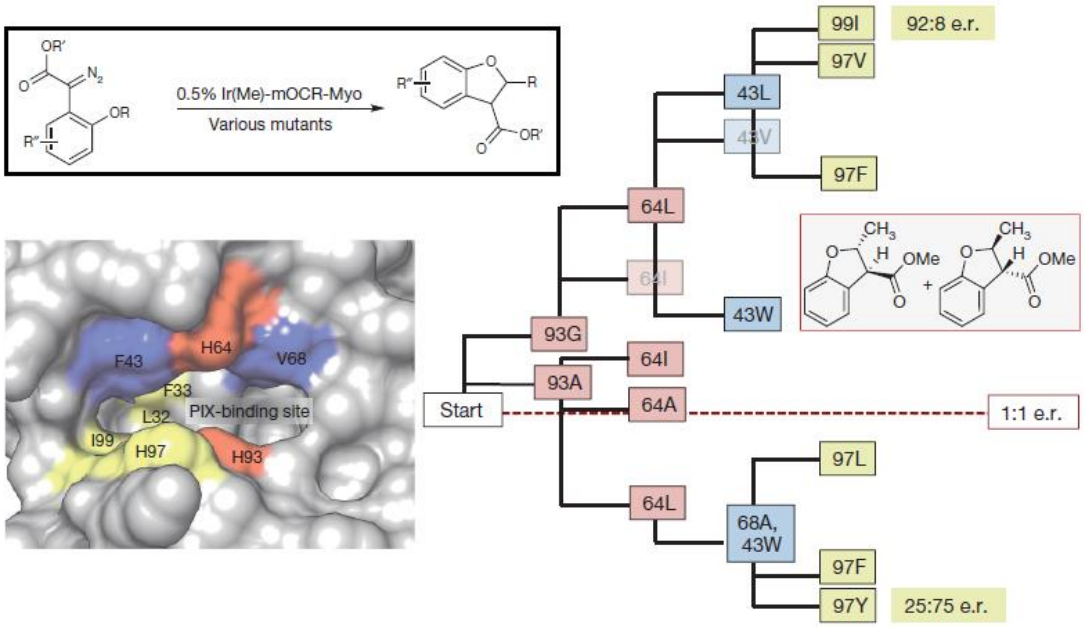


Figure 43: The directed evolution strategy used by Hartwig to obtain Ir(Me)-mOCR-Myo mutants capable of producing either enantiomer of the intramolecular C–H carbene insertion reactions of varied substrates. Inner sphere (red), middle sphere (blue), and outer sphere (yellow) residues are highlighted in the depiction of the active site (bottom left) and in the evolutionary tree (figure adapted from Key and Dyido *et al*).[209]

In October of the same year, Dyido *et al* re-examined the intramolecular C-H carbene insertion reaction used previously but using iridium-reconstituted-CYP119 variants to catalyse the reaction.[224] A variant called Ir(Me)-CYP-Max was reported to catalyse the intramolecular C-H insertion of methyl 2-diazo-2-(2-methoxyphenyl)acetate with a reaction yield and enantioselectivity of >95% respectively.[224] The authors reported remarkable kinetic values for Ir(Me)-CYP-Max, with K_M, k_{cat} and k_{cat}/K_M values of 170 mM, 45.8 min^{-1} and 269 $min^{-1}\,mM^{-1}$ for the intramolecular C-H carbene insertion reaction respectively. The authors also reported other Ir(Me)-CYP119 variants proficient towards intramolecular carbene insertion into activated primary sp^3 C-H bonds (TON = 120-600), activated

secondary sp^3 C-H bonds (TON = 402-414), unactivated primary sp^3 C-H bonds (TON = 31), sterically hindered sp^3 C-H bonds (TON 107-582), and intermolecular C-H carbene insertions into activated primary sp^3 C-H binds (55% yield).[224]

In 2018, Vargas *et al* reported the myoglobin-catalysed carbene insertion into sp^2 C-H bonds in indole.[225] The myoglobin variant, Mb(H64V,V68A), previously reported as a proficient cyclopropanation catalyst was employed to catalyse the C-H carbene insertion reaction between unprotected indole and EDA (**figure 44**). The reaction was conducted using free enzyme and in whole cells, and reaction yields of 85 and >99% were reported respectively.[225] The remarkably high yields for the whole cell catalysed biotransformations encouraged the authors to explore the substrate promiscuity, and reported Mb(H64V,V68A) could catalyse the C-H carbene insertion of a variety of functionalized indoles and reported yields of >99% all substrates. Interestingly, another variant, Mb(L29F,H64V), exhibited a preference for bulkier indole substrates relative to Mb(H64V,V68A), catalysing the C-H insertion of a variety of bulky indoles with reaction yields of 40-80%[225] The authors rationalised the enhanced activity of Mb(L29F,H64V) (relative to Mb(H64V,V68A)) by suggesting the phenylalanine, situated close to the heme cofactor, can establish favourable π-π interactions with the indole ring, unlike the H64V mutation which sterically clashes with bulkier substrates. Furthermore, the authors postulated a possible reaction mechanism for indole C-H insertion (**figure 44**). The C3 atom of indole behaves as a nucleophile and attacks the electrophilic carbon in the metallocarbenoid intermediate. The resultant complex possesses a formal negative charge on the carbene carbon (distributed with iron) and a formal positive charge on the nitrogen in the indole fragment. At this point, either a [1,2]-proton shift from the C3 carbon of indole to the α-carbonyl ester or dissociation of the ylide complex and protonation of the fragment by the solvent (accompanying an elimination of the C3 hydrogen) are both plausible mechanisms. To confirm the mechanism, the authors repeated the assays with 1-methyl-3-deutero-indole. The product obtained did not retain deuterium which indicates the reaction proceeds *via* the solvent-mediated protonation pathway. If this is the case, the mechanism cannot be classified as a C-H insertion reaction; instead, the mechanism would be better described as resembling a regioselective electrophilic aromatic substitution reaction (**figure 44**).[225] The reaction is also chemoselective, as no N-H insertion products were detected in any of the assays.

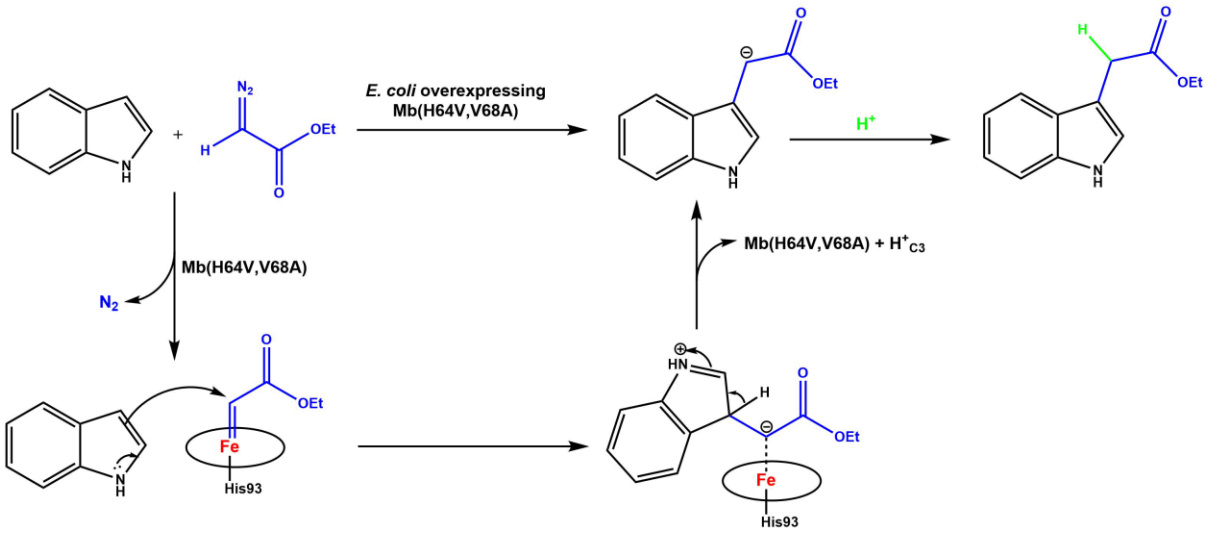


Figure 44: The general reaction and mechanism of the Mb(H64,V68A) catalysed regioselective and chemoselective electrophilic aromatic substitution reaction between EDA and indole.

In 2019, Zhang *et al* reported the first sp^3 C-H functionalisation reaction catalysed by a heme-containing protein.[226] A selection of P450, globin and cytochrome *c* variants were screened for C-H functionalisation activity to identify the blueprint protein for DE. The reaction between 1-methoxy-4-(methoxymethyl)benzene and EDA was employed as the selective pressure. One variant, P411-P-4(A82L) (18 mutations from WT-$P450_{BM3}$), gave the product with a TTN of 13, and was subsequently subjected to DE.[226] After multiple rounds of mutagenesis, one novel variant, P411-CHF, exhibited a 140-fold improvement in activity relative to P411-P-4(A82L). Furthermore, P411-CHF catalysed the reaction with an enantioselectivity of 97%. P411-CHF was active towards an assortment of substrates, exhibiting TTNs and enantioselectivities of 50-2150 and 69-99%.[226] P411-CHF was also remarkably chemoselective, catalysing sp^3 C-H functionalisation over cyclopropanation in a ratio of 25:1 when 1-allyl-4-methoxybenzene was employed as the substrate (**figure 45**).[226] *E. coli* overexpressing P411-CHF was also proficient at chemo-, regio-, and stereoselectively catalysing sp^3 C-H functionalisations (when the C-H group was alpha to an sp^2 carbon or nitrogen), successfully converting several products with TTNs ranging between 950-3750.[226]

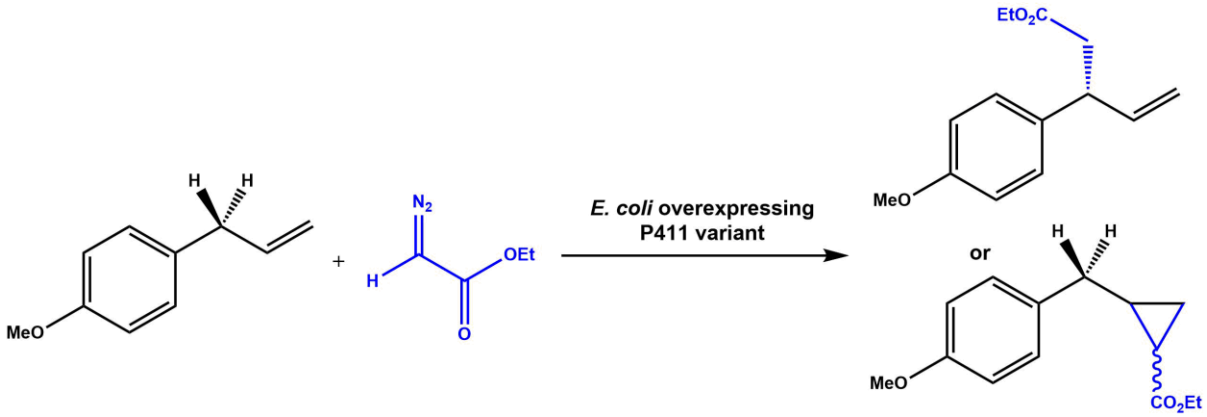


Figure 45: The possible reaction products from the P411-variant catalysed reaction between EDA and 1-allyl-4-methoxybenzene. The reaction catalysed by the variant P411-CHF exhibited a 25:1 ratio for the formation of the sp^3 C-H insertion product.[226]

In May 2019, Brandenberg *et al* re-explored the regioselective sp^2 C-H functionalisation of the C3 C-H bond in indole substrates.[227] Several P411 variants were initially screened in whole-cell biotransformation assays, with the variant P411-HF emerging as the best enzyme, catalysing the reaction between 1-methylindole and EDA with a TTN of up to 5350 (83% yield).[227] The substrate scope of P411-HF extended to 10 functionalised 1-methylindole substrates, affording TTNs and yields of 700-3380 and 17-74% respectively.[227] The authors also demonstrated P411-HF could be engineered towards either chemoselective C-H functionalisation or cyclopropanation by the introduction of key mutations. It was also shown that key mutations afforded control over the regioselectivity of a reaction, by engineering two distinctive P411-HF variants that catalysed the C-H functionalisation of methylpyrrole at the C2 and C3 positions, with ratios of 9:91 and 98:2, respectively,.[227]

In June of the same year, Zhang *et al* reported another P411 variant, P411-PFA, that was proficient at the fluoroalkylation of α-amino sp^3 C-H bonds.[228] P411-PFA catalysed the C-H functionalisation reaction between phenylpyrrolidine and 2,2,2-trifluoro-1-diazoethane with a TTN and enantioselectivity (for the *(R)*-product) of 4000 and 98% respectively.[228] The substrate scope of P411-PFA extended to 10 phenyl-functionalised phenylpyrrolidine substrates, affording TTNs and enantioselectivities of 120-2100 and 84->99%. The authors also engineered another variant, P411-PFA-(S) (8 mutations away from P411-PFA), which catalysed the C-H trifluoroalkylation of phenylpyrrolidine in favour of the (*S*)-product (94% enantioselectivity), demonstrating the evolvability of the P411 scaffold towards enantiodivergent biocatalysis in the synthesis of chiral organofluoride molecules (**figure 46**).[228]

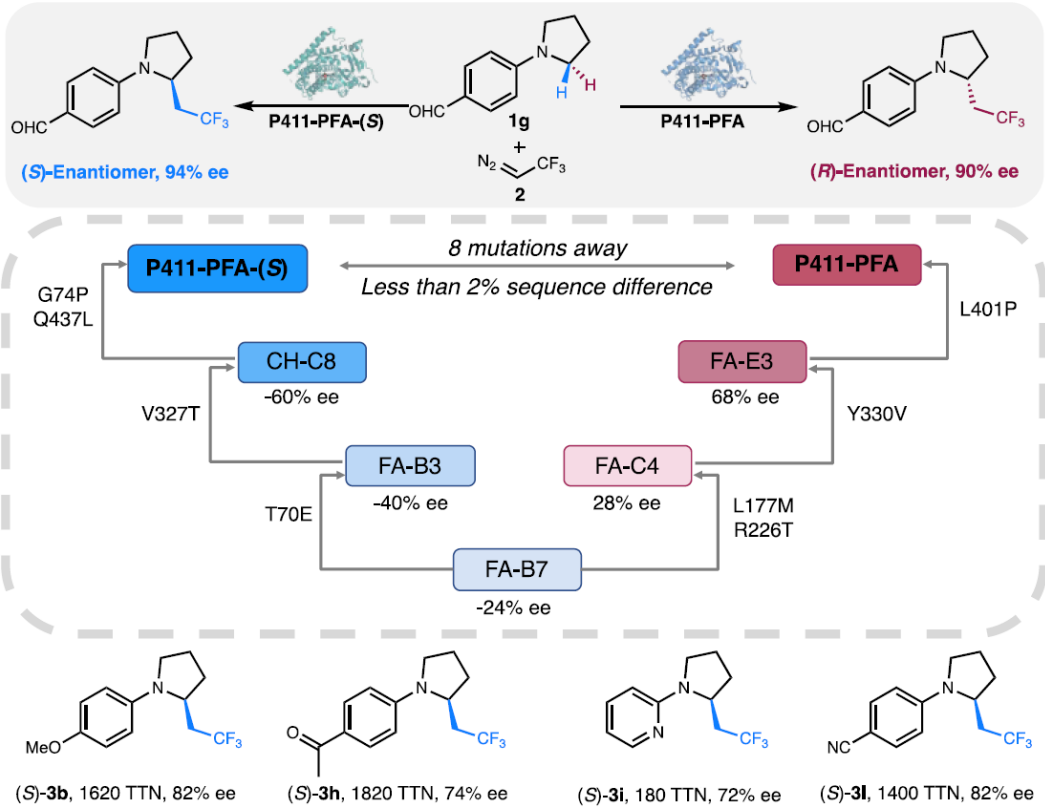


Figure 46: Enantiodivergent C-H sp^3 trifluoroalkylation using the P411 variant P411-PFA (figure taken from Zhang *et al*).[228]

2.3.4 Carbonyl Olefination

In 2015, the Fasan group demonstrated a few key mutations to sperm-whale myoglobin afforded a carbene transferase proficient towards cyclopropanation.[195] In 2016, the same group postulated that an electrophilic metallocarbenoid intermediate could react with tertiary phosphine nucleophiles to engender a phosphonium ylide.[229] The resultant intermediate, a heme localised phosphonium ylide, could further react with electrophilic carbonyl groups in a fashion analogous to the Wittig reaction, but with the protein scaffolding providing an asymmetric environment to govern the stereoselectivity. After screening several conditions and mutants, a myoglobin variant, Mb(F43V,V68F), proved the most efficient, with a reported TON and E/Z diastereoselectivity of 34 and 64 respectively in the synthesis of ethyl cinnamate from benzaldehyde and EDA in the presence of triphenylphosphine.[229] Repeating the assays with triphenyl arsine resulted in a TON and d_E of 92 and >99.9% respectively. The same mutant in later screens (with lower catalytic loading) exhibited good substrate specificity for carbonyls when $AsPh_3$ and EDA/cyclohexyl-diazoacetate were employed, with reported TONs and cis:trans diastereoselectivities ranging from 735-3,400 and 91-99.9% respectively.[229]

In the same year, Weissenborn *et al* published a report on the carbonyl olefination reaction catalysed by the hemoprotein YfeX, isolated from *E. coli*.[230] YfeX was reported to catalyse the carbonyl olefination reaction between EDA and benzaldehyde under aerobic conditions and in the absence of phosphines. The group started by screening for carbonyl olefination activity in *E. coli* lysate, and identified four active hemoproteins, flavohemoprotein (Hmp), bacterioferritin (Bfr), hydroperoxidase II (KatE), and YfeX, *via* fast-protein-liquid-chromatography (FPLC). All four proteins, once identified, were overexpressed, purified and screened for carbonyl olefination activity; of the four candidates, YfeX catalysed the reaction with a TTN of 2.9 and a trans:cis selectivity of 60%.[230] The introduction of PPh_3 to the reaction mixture elevated the TTN to 16, and the introduction of various mutations into YfeX engendered a variant, YfeX(S234A), with an increased TTN of 17.3.[230] The authors presented two differing mechanisms to account for the reactivity observed in the presence and absence of phosphine additives (**figure 47**).

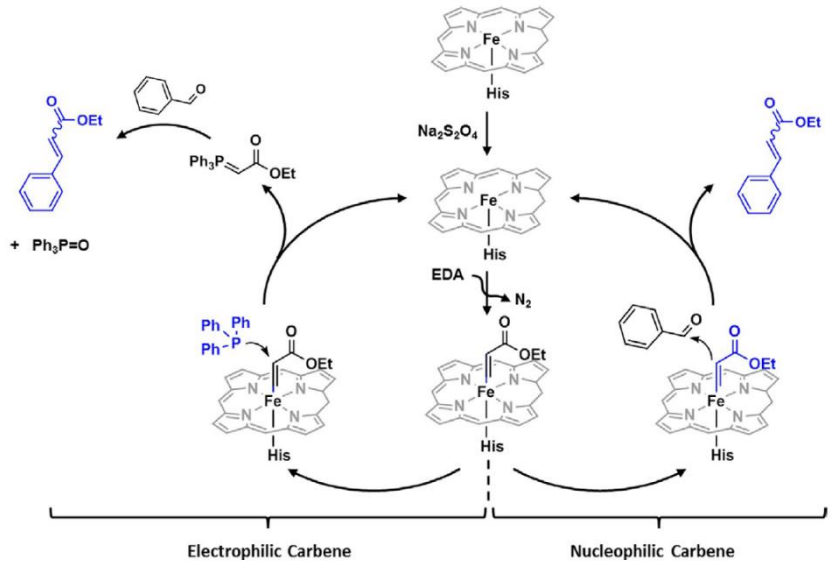


Figure 47: Proposed mechanistic differences for the hemoprotein catalysed carbonyl olefination of aldehydes in the presence (electrophilic carbene) and absence (nucleophilic carbene) of phosphine reagents (figure taken from Weissenborn *et al*).[230]

2.4 INTRODUCTION TO *DE NOVO* PROTEINS

The ability to introduce novel functionality into naturally occurring protein scaffolds ignited interest in engineering new enzymes towards abiological chemistry. However, one problem with natural proteins is their long, complex evolutionary history.[95] The interdependency of amino acids, a consequence of natural evolution, means simple mutations can unpredictably impair the function of a protein, sometimes in an untraceable way. Unravelling the intricate relationship between amino acids and protein function, which have been fostered over two billion years of evolution, is an

insurmountable task. This problem with natural protein scaffolds demanded a novel approach to protein design: *de novo* design.[95,96] The basic premise underpinning *de novo* protein design is simplistic, small protein scaffolds, inspired by nature, can be created to mimic the functionality of natural proteins without harbouring any evolutionary baggage. In addition, the chemistry exhibited by many enzymes is a consequence of the intrinsic chemistry of an incorporated cofactor, such as a heme group. *De novo* design asserts the functionality of hemoproteins, for example, can be replicated by incorporating heme cofactors into simple, engineerable protein scaffolds and utilising the protein scaffold to fine-tune the functionality by introducing conformational restraints, facilitating substrate/cofactor binding, and/or tuning redox potentials.[231–233]

There are three prominent approaches to designing *de novo* proteins: i) the maquette approach, ii) computational design, and iii) combinatorial assembly.[233] In practice a combinational of al the approaches is used. The three methodologies are founded on basic principles. The maquette approach is a non-computational method which employs an evolutionary naïve protein scaffold as a starting template.[98,231,233] The protein scaffold is designed to fold into a particular stable confirmation and provides a blueprint for functional design (**figure 48**). The intrinsic stability and mutability of maquettes facilitates enzyme design, with simple point mutations in the protein scaffold often proving sufficient for acquiring the desired functionality.[98,231,233] Furthermore, multiple cofactors, such as hemes,[234–238] flavins,[97,239] small metals ions (such as Zn^{2+}, Co^{2+}),[94,231] iron-sulphur clusters[240–242] and other tetrapyrrole cofactors[97,241,243,244] can be readily incorporated into a maquette backbone by introducing a known sequence motif (i.e. heme c-type C-X-X-C-H motif) or binding pocket without disrupting the protein scaffolding or suppressing the desired functionality (**figure 48**). The computational design of *de novo* proteins utilises computational methods (such as Rosetta software for protein design, coupled with QM/MM and MD to analyse Rosetta outputs) to design suitable protein scaffolds optimised towards a particular function, protein fold, or theozyme.[95,245–250] Iterative simulations can be performed to determine the optimal geometric conditions for cofactor binding and small metal binding, and for tuning redox potentials by optimising the amino acids in the active site and/or secondary sphere. Computational design is frequently utilised alongside the maquette approach to assist in designing new catalysts, for example in the development of maquettes with oxidoreductase activity.[232,251,252] The combinatorial approach employs a small, monomeric protein scaffold to be redesigned, and generates large libraries of mutants which are subject to high-throughput screening methods to assess function; iterative rounds of diversification can be performed to further enhance a protein scaffold towards a desired function.

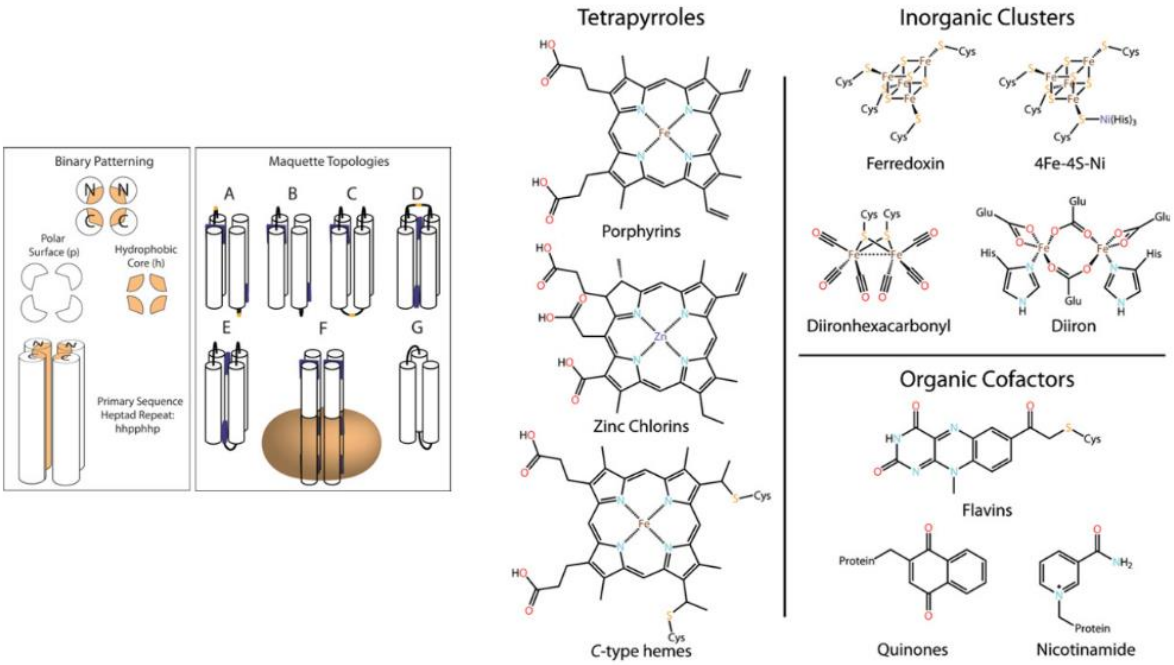


Figure 48: (**left**) the pattern, sequence and common topologies of maquette protein scaffolds, and (**right**) common redox cofactors that can be readily incorporated into maquette scaffolds (figure taken from Lichtenstein *et al*).[253]

The oxidoreductase class of proteins encompasses an enormous set of functionally diverse and biologically important enzyme families.[99,143,254] Oxidoreductase enzymes are involved in redox chemistry and find critical roles in several biological processes including aerobic/anaerobic respiration, secondary metabolite biosynthesis, xenobiotic detoxification and energy/light harvesting.[99,143,254] The design of small, simplistic enzymes capable of replicating oxidoreductase activity is a very attractive endeavour, owing to the immense biological importance of oxidoreductases, and has been the subject of a wealth of research in the protein design community.

2.4.1 Historical development of α-helical b-type heme-containing maquettes

The redox chemistry performed by oxidoreductases is achieved through the employment of various redox active cofactors. [99,143,254] One of the most important classes are the heme cofactors, which facilitate redox chemistry *via* Fe^{II}/Fe^{III} redox transitions.[143] For example, the hemoproteins (heme-containing proteins) catalyse electron transfer reactions in a variety of biological processes, including anaerobic/aerobic respiration, sulphate, nitrate and carbon fixation, biosynthesis and small-molecular transportation, all processes governed by the Fe^{II}/Fe^{III} redox potential.[143] The ubiquitous distribution of hemoproteins across all biological kingdoms makes them attractive targets for protein design.

The transmembrane respiratory complex cytochrome bc_1 contains a four α-helical subunit capable of binding two bis-histidine coordinated b-type heme cofactors (**figure 49**).[255–257] The structure of the respective active sites neighbouring each heme finely tunes the Fe^{II}/Fe^{III} redox potential of each heme, resulting in a low- and high-potential heme which facilitates electron transfer.[255–257] The sequence of the tetra-α-helix bundle was recognized to be comprised of a certain pattern of hydrophobic and hydrophilic amino acid residues situated at particular positions to facilitate protein folding, stability and heme binding. The four helices are connected by three disordered helix-loop-helix domains to yield the monomeric tetra-α-helix di-heme containing subunit of the enzyme complex cytochrome bc_1.[255–257] The chemical principles underlining the tetra-α-helix bundle served as inspiration for the design of simplified protein scaffolds for heme-containing oxidoreductase *de novo* design.

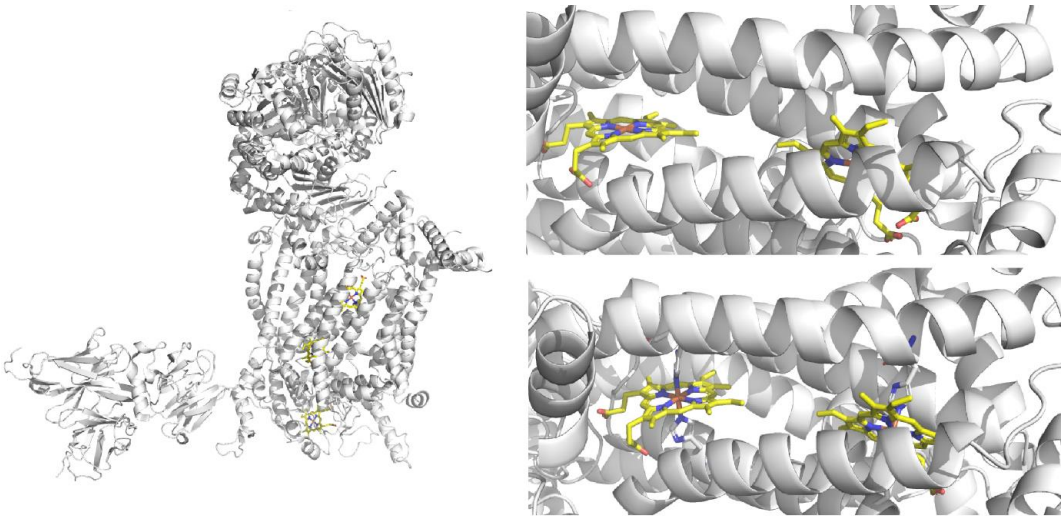


Figure 49: (**left**) The crystal structure of the yeast cytochrome bc_1 complex (white, hemes are shown in yellow; *PDB: 1KB9)*, and (**top left and top right**) close of views of the di-heme tetra-α-helix bundle in the core of yeast cytochrome bc_1, the bottom panel shows the bis-histidine ligation of each heme cofactor (white, hemes are shown in yellow; *PDB: 1KB9*).

In 1987, Degrado and co-workers reported the design of monomeric amphiphilic 16-residue synthetic peptides, α_1A and α_1B, designed to tetramerise into 4-helix bundles (**figure 50, A**).[258] Each peptide was confirmed to form stable helices by circular dichroism (in the presence and absence of a denaturant). The peptide α_1B was subjected to further design and two peptides, α_1B-Pro-α_1B and α_1B-Pro-Arg-Arg-α_1B, were reported, with the -Pro- and -Pro-Arg-Arg- links incorporated into the scaffold to covalently link two α_1B peptides. The α_1B-Pro-α_1B peptide was reported to trimerise in solution to give a hexa-α-helix protein, while the α_1B-Pro-Arg-Arg-α_1B was reported to form stable helical dimers (**figure 50, B**). The difference between the two is attributed to the inclusion of the two Arg residues in the connecting loop, which reportedly destabilises the trimeric structure of α_1B-Pro-α_1B. This work demonstrated the facile nature at which tetra-α-helix bundles can be designed from simple chemical principles.[258]

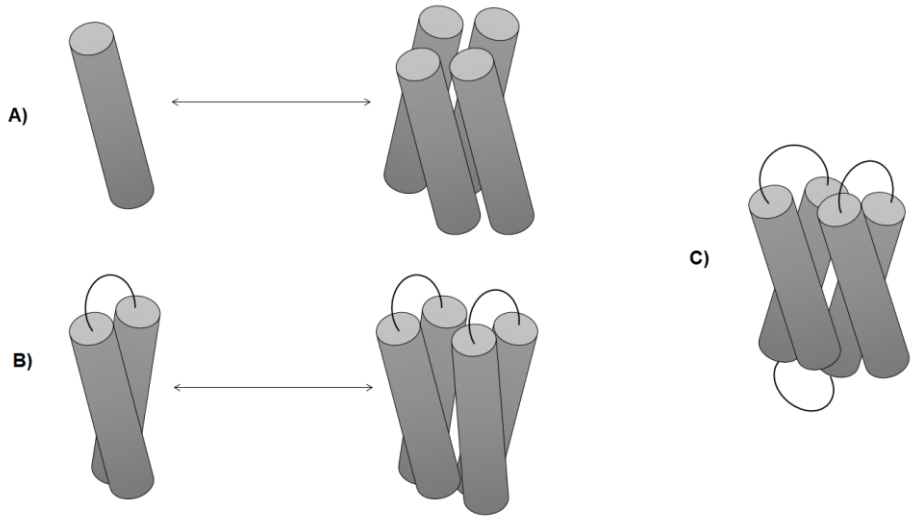


Figure 50: The design of tetra-α-helical bundles, from **A)** the self-assembly of monomeric peptides into a tetramer, **B)** the self-assembly of a helix-loop-helix monomeric unit into a dimer, and **C)** a monomeric four-helix bundle formed by connecting loops

In 1988, Degrado and co-workers extended their designs to include monomeric tetra-α-helix bundles (**figure 50, C**).[259] A series of proteins, named α_1, α_2, and α_4, were designed to emphasize how simple point mutations in the helical backbone can finely tune the dynamics, folding and stability of simple tetra-α-helix maquettes. The general 'helix' sequence, employed for all three maquettes, was intentionally selected to favour 4-helix bundle formation. Hydrophobic leucine residues were introduced at certain positions to enhance the hydrophobic interactions in the core of the maquette. Likewise, glutamic acid and lysine residues were introduced at certain positions to facilitate the formation of ion pairs between the helices and provide additional stabilisation to the final 4-helix structure.[259] The α_1 protein consisted of a single helix monomeric unit which tetramerised in solution; the α_2 protein consisted of a helix-loop-helix monomeric structure which favourably dimerised into a tetra-α-helix protein.[259] The α_4 protein was a monomeric, single polypeptide chain that folded into a tetra-α-helix confirmation, could be expressed in *E.coli*, and possessed considerable stability towards denaturants (such as GuHCl) relative to small hemoproteins ($\Delta G_{\alpha 4}$ = -22.5 kcal mol^{-1} vs -7.6 kcal mol^{-1} for myoglobin).[259]

In 1994, DeGrado and Dutton reported a library of synthetic 62-residue di-α-helical peptides that, through self-assembly into tetra-helical bundles, were capable of accommodating two bis-histidine ligated heme B cofactors (**figure 51, C**).[260] A series of four dimeric maquettes, collectively named IIa_2, IIb_2, IIc_2, and IId_2 respectively, were designed based on two differing di-α-helical monomers (the two helices were linked by an S-S bond) to yield four distinctive $(\alpha\text{-SS-}\alpha)_2$ maquettes. The defining

structural difference between the four maquettes was the number and position of histidine residues (H10 and H24) in each helix respectively.[260] The IIa_2 and IId_2 maquettes both contained a H10-H10 and H24-H24 histidine pair, rendering them tetraheme-containing dimeric bundles, while the IIb_2 (H10-H10) and IIc_2 (H24-H24) maquettes yielded dimers possessing only two bis-histidine ligated hemes (at H10-H10 and H24-H24) respectively (**figure 51, A/B**).[260] Importantly, all four maquettes possessed biophysical characteristics similar to natural heme-containing proteins, displayed sub-micromolar heme dissociation constants and exhibiting characteristic spectroscopic profiles for bis-histidine ligataion.[260] These maquettes are collectively named the H10H24 maquettes, and beautifully illustrate how simple, mutable protein scaffolds can be readily engineered to accommodate biologically important cofactors. In 1996, Dutton and co-workers incorporated a $[4Fe\text{-}4S]^{2+/1+}$ cluster into the loop domain of H10H24 maquettes, further illustrating the facile nature at with tetra-α-helix bundles can be engineered towards desired functionality.[240]

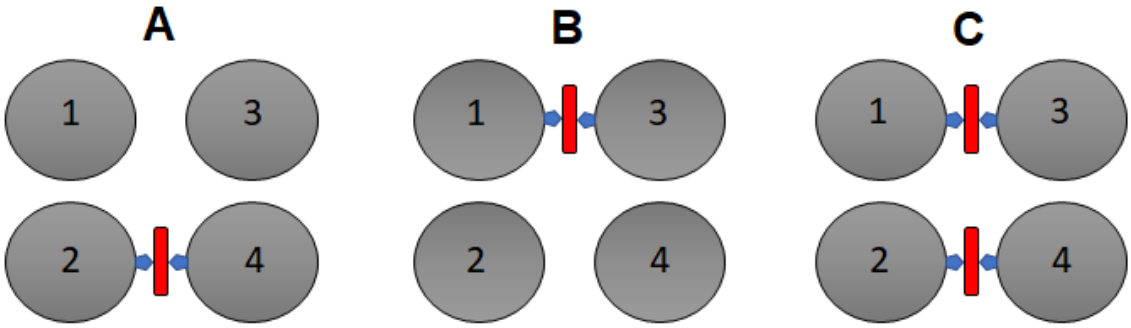


Figure 51: The incorporation of bis-histidine ligated b-type heme cofactors into tetra-α-helical maquette. **A and B)** incorporation of a single b-type heme, and **C)** incorporation of two b-type hemes.

In 1996, DeGrado *et al* reported a di-α-helical maquette, α_2, which upon dimerization assembled into a parallel configuration with two-fold symmetry to engender the tetra-α-helix maquette, $[\alpha_2]_2$.[261] All four α-helices of the bundle had a common sequence, containing two histidine residues to function as heme binding sites. Coproporphyrin I (CP) was selected to represent chlorophyll. A parent chain of α was derivatised with CP to yield CP-α, which heterodimerised with another α chain to yield the CP-α_2 which, like the parent chain, self-assembles to give the dimeric maquette $[CP\text{-}\alpha_2]_2$. The CP monomer-to-dimer equilibrium ($K_D < 10^{-8}$) indicated that $[CP\text{-}\alpha_2]_2$ stabilises and promotes CP dimerization; the authors reported up to four CP cofactors could be incorporated into the maquette without any detectable disruption to the CP monomer-to-dimer spectra.[261] The postulation was further support after examining the monomer-to-dimer equilibrium in the presence of 2,2,2-trifluoroethanol (TFE), a solvent known to disrupt the tertiary and quaternary structure of a protein without disrupting the helicity. The addition of TFE to $[CP\text{-}\alpha_2]_2$ yielded UV-VIS spectra indicative of a dimer-to-monomer

transition, caused by the dissociation of the $[CP\text{-}\alpha_2]_2$ into monomeric CP-α_2 molecules. The ability of $[CP\text{-}\alpha_2]_2$ to stabilise the dimerization of CP by >1000-fold lead the authors to postulate that $[CP\text{-}\alpha_2]_2$ could function as a primitive, simplified structure in future efforts towards the design of photoactive redox centres.[261]

In 1997, Dutton *et al* explored the biophysics of the H10H24 maquette by substituting the hydrophobic leucine residues for β-branched and/or aromatic amino acids.[262] Initial studies introduced isoleucine mutations at L6 and phenylalanine mutations at L13 to yield three mutant maquettes, H10H24-(L6I), H10H24-(L13F), and H10H24-(L6I, L13F).[262] The CD spectrum for each mutant had 222nm/208nm ratios >1, the hallmark for α-helicity. The authors also reported NMR data for each mutant, with each exhibiting enhanced chemical shift dispersions relative to the parent H10H24 maquette.[262] The enhanced chemical shift dispersion is a result of a more native-like, stable tertiary structure for the mutant maquettes provided by the increased side-chain packing specificity in the maquette core. Furthermore, each mutant maquette exhibited heme-binding dissociation constants in the 10-20 nM range for the binding of the initial pair of hemes, followed by dissociation constants in the 5-15 μM range for the second pair of hemes. The EPR spectra indicating the resting state of each maquette is a low-spin ferric complex.[262] The importance of this work is it demonstrates how introducing key substitutions into the maquette scaffold (leucine for isoleucine/phenylaniline) can result in the emergence of native-like structures for the simple maquette proteins.

In the following year, Dutton and co-workers further examined the nature of the double variant H10H24-(L6I, L13F) in contrast to parent maquette H10H24.[263] It was reported that apo-H10H24 has a defined secondary structure, but exists as an ensemble of slowly interconverting confirmations in solution. The authors utilised ^{15}N-HSQC and NOSY NMR to investigate the structure of apo-H10H24-(L6I, L13F), which was characterised by a single set of narrow NMR resonances and a chemical shift dispersion characteristic of a native protein. It was reported that apo-H10H24-(L6I, L13F) exists, in solution, as a four-α-helix bundle comprised of two asymmetric $(\alpha\text{-SS-}\alpha)_2$ helices with a well-defined tertiary structure. The authors indicated the maquette apo-H10H24-(L6I, L13F) would be employed in future investigations to better understand the dynamics of protein stability and folding.[263]

In 1999, two papers using NMR spectroscopy to study the solution structure and dynamics of *de novo* helical-bundles were reported: a *de novo* three-helix bundle by Walsh, Degrado and co-workers[264] and a *de no*vo four-helix bundle by Skalicky *et al*.[265] In the Degrado report, the NMR solution structure of a non-native three-α-helix bundle, α_3D, was resolved. Although α_3D is not based on any native protein, its biophysical characterisation revealed its thermodynamic and spectroscopic properties paralleled native proteins.[264] The solution structure of α_3D was solved by heteronuclear multidimensional NMR

(HNMD-NMR) spectroscopy which revealed the helices adopt a counterclockwise topology, in agreement with computational predictions. The structure acquired from HNMD-NMR spectroscopy also confirmed the presence of a well-defined and tight hydrophobic core.[264] In the Dutton report, the NMR solution structure of a non-native, two-fold symmetrical four-α-helix bundle, $(\alpha'\text{-SS-}\alpha')_2$, was resolved.[265] It was observed that α'-SS-α' (**figure 52**) dimerises and adopts a parallel configuration which is stabilised by the formation of a well-ordered hydrophobic core. One particular feature of $(\alpha'\text{-SS-}\alpha')_2$ is its lack of any heme-binding pocket because of the tight hydrophobic packing in its core. Despite this, $(\alpha'\text{-SS-}\alpha')_2$ is still capable of binding four heme cofactors with K_d values of 10-20 nM (for the first two hemes) and 5-15 μM (for the last two hemes). The authors note the binding of four hemes causes the reorganisation in the hydrophobic core, resulting in a disorder mutli-heme containing protein with a well-defined helical structure.[265]

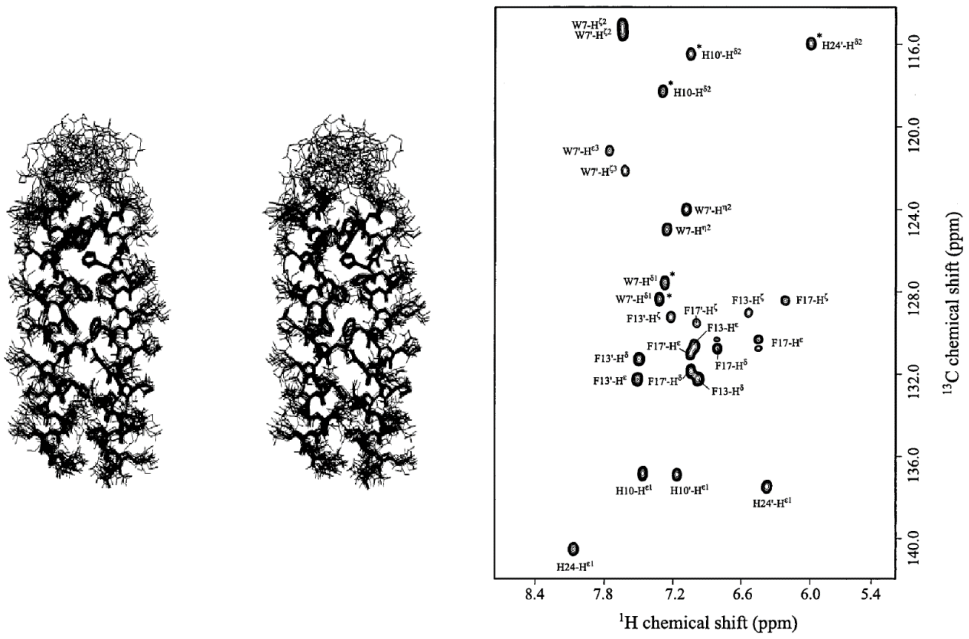


Figure 52: (left) the three-dimensional superposition of 25 (α'-SS-α') structures and **(right)** the 1H/13C HSQC NMR spectrum for α'-SS-α' (figures taken from Dutton *et al*).[265]

The establishment of well-defined α-helical scaffolds provided a platform to begin investigating how oxidoreductase functionality could be engineered into maquettes. In 2000, Shifman *et al* explored how varying the structure of a co-factor, changing the pH and introducing key mutations in the heme-binding site affected the heme-redox potential when the cofactor was bound to a tetra-α-helical maquette.[244] A combination of all possible approaches allowed the authors to obtain a library of bis-histidine ligated heme-containing maquettes possessing redox potentials across a 435 mV range.[244] The work demonstrated that using simple chemical principles the redox potential of heme-containing

maquettes can be finely tailored towards a particular application, highlighting the fact that *de novo* maquettes are more than mere laboratory curiosities.

In 2003, to corroborate the NMR structures of α-helical maquettes reported by the Degrado[264] and Dutton[265] labs in 1999, Dutton reported the X-ray crystal structure of the maquette $[I_6F_{13}M_{31}]_2$ (called L31M), a third-generation H10H24 maquette (H10H24-(L6I, L13F, L31M)).[266] The X-ray structure, resolved to 2.8 Å, of L31M revealed the maquette is a tetra-α-helical bundle, formed by the dimerization of two (α-SS-α) monomers, and adopts an anti-topology confirmation.[266] In 2004, Dutton and co-workers expanded on the L31M maquette.[267] It was reported that both $[I_6F_{13}]$ (the predecessor to L31M) and L31M were conformationally specific in the apo-state, but binding of a ferric heme engendered molten globular maquettes with poorly defined tertiary structures.[267] The L31M maquette scaffold was subsequently subjected to seven additional mutations to yield a novel, fourth generational maquette named HP-1.[267] HP-1 was reported to be a molten globule in the apo-state, in direct contrast to $[I_6F_{13}]$ and L31M maquettes.[267] HP-1 was, however, reported to bind ferric heme tightly, with reported dissociation constants in the nM range (< 20 nM). Halo-HP-1 exhibited UV-VIS absorption spectra consistent with bis-histidine ligation for both the ferric and ferrous states, and the EPR spectra indicated the ground electronic state of di-heme HP-1 to be two identical low-spin bis-histidine ligated ferric hemes.[267] Unlike apo-HP1, halo-HP-1 possess a defined tertiary structure in solution (determined by NMR spectroscopy), in direct contrast to $[I_6F_{13}]$ and L31M.[266,267] Interestingly, HP-1 binds zinc-protoporphyrin IX with high affinity, in a similar fashion to iron-protoporphyrin IX; however, the resulting complex is a pentacoordinate ligated complex and the NMR chemical shift dispersions indicate Zn-PPIX-HP-1 has no definitive tertiary structure in. This observation indicated that bis-histidine ligation is necessary for HP-1 to adopt a stable, defined tertiary structure.[267]

The maquette approach began using simple repeating monomeric structures that self-assemble into small tetrameric bundles in solution. Over the course of two decades, the maquette approach matured and ultimately afforded small, stable well-defined tetra-α-helical heme-containing protein scaffolds. The next question concerned whether tetra-α-helical heme-containing maquettes could mimic hemoprotein function. In 2009, Dutton and co-workers reported a tetra-α-helical heme-maquette capable of binding oxygen with affinities similar to natural histidine-ligated globins (**figure 53**).[268] The evolutional history of the oxygen-binding maquette starts with a simplified scaffold consisting of three amino acids (glutamate, lysine, and leucine) arranged in an almost repeating heptad structure. The monomers generated were observed to spontaneously assemble into four-helix bundles in solution, with the polar residues (glutamate and lysine) exposed to the solvent and the non-polar residues (leucine) forming a hydrophobic interior.[268] From this simple scaffold a novel bis-

histidine ligated di-heme maquette was engineered. The maquette consisted of two di-helical containing identical sequences fused together by a single disulphide linked bridge to yield a monomeric di-heme-containing tetra-α-helical maquette (**figure 53**).[268] The maquette bound heme B tightly (K_d < 1 nM) and exhibited UV-VIS spectra consistent with bis-histidine coordinated ferric and ferrous heme. The ferrous maquette displayed UV-VIS spectra consistent with the formation of an oxy-ferrous complex, with spectra remarkably similar to that of neuroglobin (**figure 53**).[268] It was reported the oxy-ferrous was stable for approximately 10 seconds before superoxide is released. The authors reported an apparent K_{dO2}/K_{dCO} of 0.1, a ratio significantly larger than that observed for other distal histidine oxygen binding hemoproteins.[268] Although the rate of formation of the oxy-complex, (k_{onO2} ; 0.017 μM^{-1} s^{-1}) was lower than for natural hemoproteins, the rate of oxygen dissociation (k_{offO2}) was drastically lower. The stability of the oxy-complex is attributed to a stabilising interactions between a histidine residue and the oxygen ligand.[268] The importance of the work is two-fold. Firstly, it demonstrates that *de novo* protein scaffolds are functional proteins capable of replicating naturally occurring enzymes. Secondly. It highlighted the underlining principle of *de novo* design that the complex structures found in nature (such as the globin fold in oxygen-binding proteins) are unnecessary to achieve functionality.

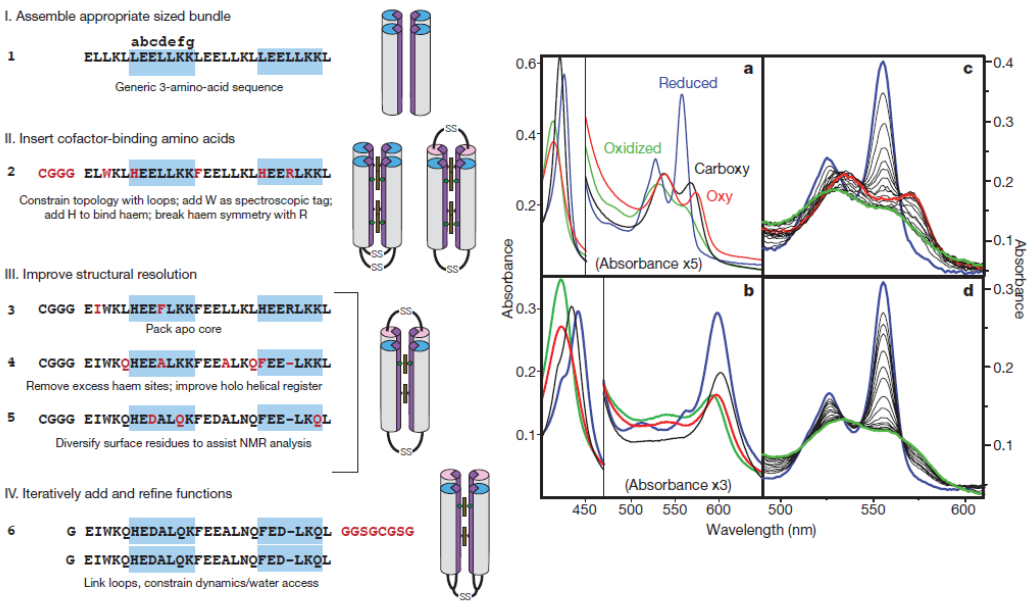


Figure 53: The design of an artificial oxygen binding protein **(left)** the design pathway from a simple self-assembling monomeric peptides into a 4-helix bundle (**1**) to the di-b-type heme containing monomeric 4-helix bundle (**6**) and **(right)** the UV-VIS spectra for the ferric, ferrous, carboxy-ferrous and oxy-ferrous of the maquette **6** with either heme B (**a**) or heme A (**b**). The spectra **c** and **d** show the stopped-flow spectral changes after rapid mixing of ferrous heme B containing **6** with oxygen (figure adapted from Koder *et al*).[268]

A seminal paper in *de novo* protein design came in 2013, when Dutton and co-workers reported a series of non-symmetrical monomeric tetra-α-helical maquettes.[269] The homodimeric maquette structure was redesigned to engender a series of monomeric maquettes structurally consisting of four helices connected by three loops. The elementary scaffold design was engineered to afford a minimalist interior that allowed for the functionalisation with a range of light- and redox- active cofactors, including hemes, zinc porphyrins, zinc chlorins, and flavins.[269] The non-symmetrical sequence of the principle maquette scaffold allowed more sophisticated functionalisations to be accessed.[269] A monoheme-containing maquette was reportedly efficient at oxygen-binding, producing spectra consistent with the formation of a ferrous-oxy complex. The absorbance in the spectra doubled when the same maquette scaffold containing two hemes was mixed with an oxygen saturated buffer.[269] The same diheme-containing reduced maquette was reported to catalyst an inter-protein reduction, transferring an electron to cytochrome *c* with a remarkable bimolecular rate constant of 2 x 10^7 M^{-1} s^{-1}. The mutation of three glutamates to alanines in each helix afforded a new di-heme containing maquette that demonstrated proficiency at superoxide formation, with a half-life of approximately 0.15 seconds observed for the formation of the oxy-ferrous complex and the subsequent generation of superoxide. Finally, functionalising mono-heme scaffolds with either zinc porphyrin, zinc chlorin or flavin engendered a series of maquettes with photoinducible intraprotein electron transfer capabilities.[269]

2.4.2 The emergence of the tetra-α-helical c-type heme-containing maquettes

There are two primary objectives of *de* novo protein design. The first is the ability to engineer simple, protein scaffolds capable of mimicking the biological function. The second is the ability to incorporate *de novo* proteins into cellular environments for *in vivo* function. The development of the heme-containing maquettes utilised non-covalently bound heme cofactors (heme A and heme B), with the heme cofactor being introduced to the protein scaffold post-purification. In nature, many hemoproteins are functionalised with c-type heme cofactors, in which the heme is post-translationally appended to the protein *via* the formation of covalent thioether bonds with cysteine residues.[100] The covalent attachment of a c-type heme to apo-c-type proteins is catalysed by a large membrane-localised protein complex called the cytochrome *c* maturation (*ccm*) apparatus.[270,271] The attachment of c-type hemes to apo-c-type proteins requires three things: i) the translocation of the biosynthesised heme cofactor from the cytoplasm to the periplasm, ii) the translocation of the polypeptide sequence from the cytoplasm to the periplasm, and iii) the reduction of any disulphide bridges to thiols to allow covalent attachment.[270] In bacterial c-type hemoproteins, the *ccm* apparatus perform all three crucial roles (**figure 54**).[270] In histidine ligated c-type hemoproteins, a common sequence motif $C_\alpha X_1 X_2 C_\beta H$ is

considered essential for translocation across the lipid bilayer and post-translational modification.[271] In particular, the histidine residue which provides the axial ligand post-functionalisation is considered absolutely vital for the functioning of the *ccm* apparatus, as mutations of the histidine residue in c-type hemoproteins result in only small quantities of the apoprotein being formed.[271]

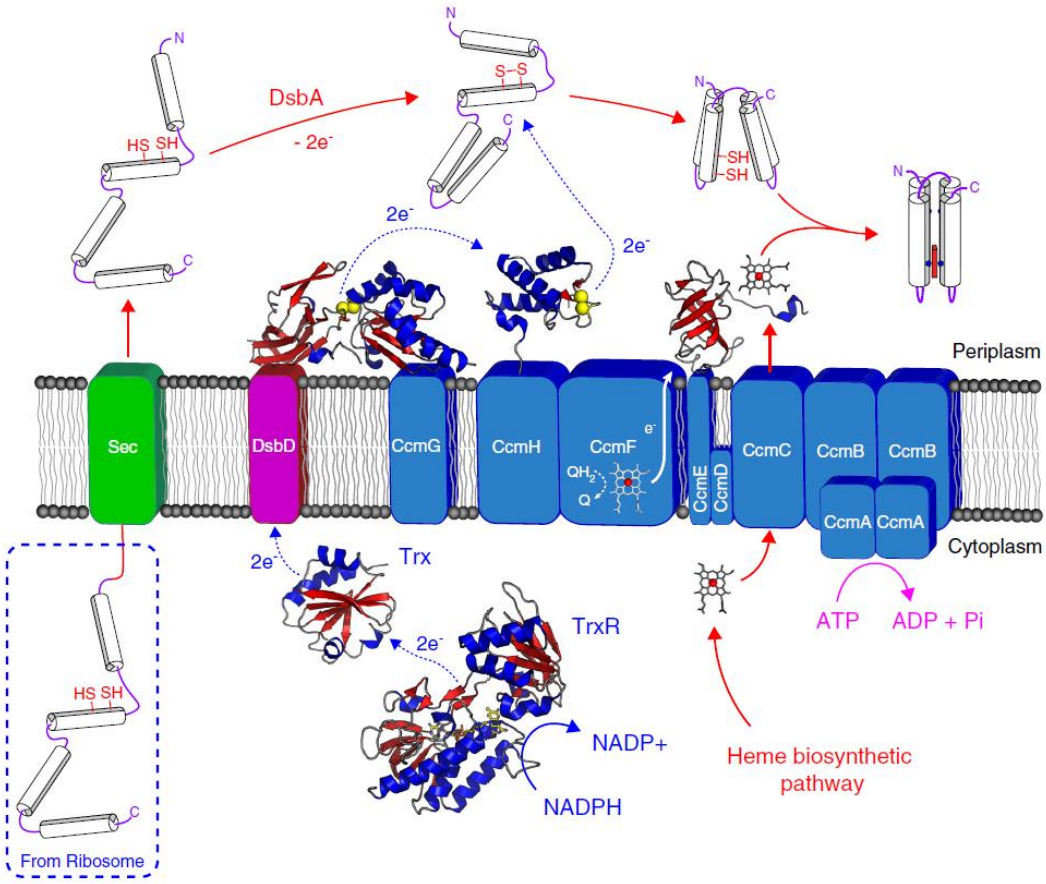


Figure 54: The cytochrome *c* maturation (*ccm*) pathway (taken from Watkins *et al*)[272]

The success of engineering b-type maquettes raised the question of whether the *ccm* apparatus could be utilised *in vivo* to post-translationally modify an apo-maquette (containing the crucial $C_{\alpha}X_1X_2C_{\beta}H$ sequence motif) into a c-type maquette. The hijacking of the bacterial *ccm* machinery had already been demonstrated in 2004 when Braun and Thöny-Meyer utilised the *ccm* machinery of *E. coli* to produce a library of c-type functionalised artificial microperoxidases *in vivo*.[273] In 2014, Anderson and co-workers reported a fully functional, monomeric tetra-α-helical c-type maquette (CTM), possessing the necessary consensus sequence, that was successfully constructed *in vivo* when co-expressed with a pEC86 expression vector harbouring the *ccm* apparatus.[251] After 15 hours of co-expression with the pEC86 vector the authors reported up to 90% c-type heme incorporation into the maquette scaffold. The incorporation of a c-type heme was verified by MALDI-TOF mass spectrometry and UV/VIS spectroscopy.[211] The Q-band of the ferrous c-type maquette observed in the pyridine hemochrome

was identical to the Q-band exhibited by bovine cytochrome *c*, and treating the c-type maquette with acidified 2-butanone resulting in no heme partitioning into the organic phase.[211]

Anderson and co-workers had demonstrated that *de novo* maquettes can be engineered and expressed, without disruption of functionality, in cellular environments. In 2015, Watkins *et al* extended their work by exploring the engineering tolerance of the *ccm* machinery and the c-type maquette prototype.[234] The authors generated a diverse library of c-type maquettes (CTMs) including C1, C2, C3, and C4 respectively, distinguished by the helix the c-type heme is covalently appended to (**figure 55**).[234] In addition, a suite of diheme containing maquettes, functionalised with ether two c-type hemes (cCTMs) or a c-type and b-type heme (bCTMs), named C1C2, C1C4, C2C4, and C3C4 respectively, were also created.[234]

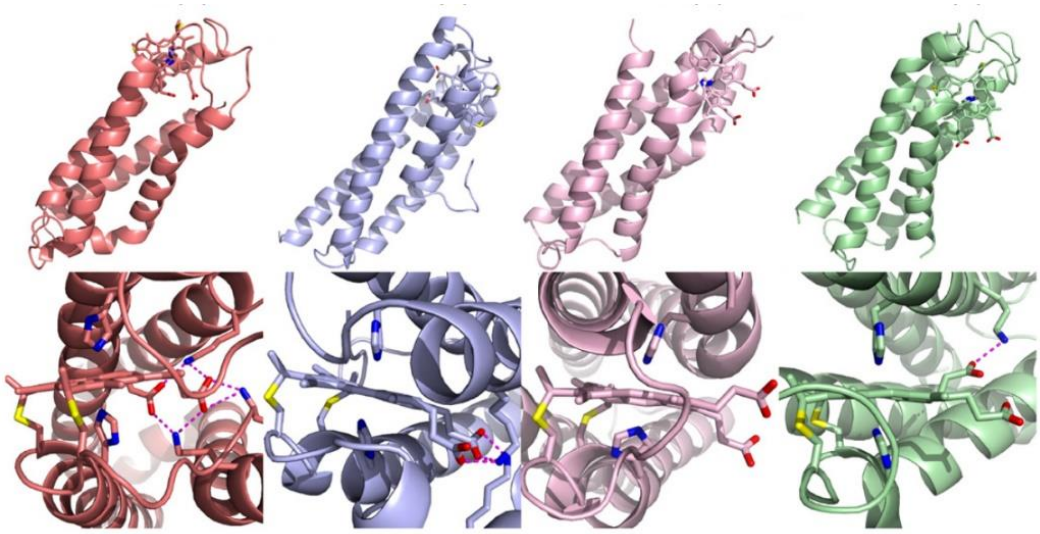


Figure 55: 1 μs molecular dynamic simulations of CTMs. (**top row**) side views of C1 (red), C2 (blue), C3 (pink), and C4 (green), and (**bottom row**) zoomed in view of the heme c active site in C1 (red), C2 (blue), C3 (pink), and C4 (green). Figures taken from Watkins *et al*.[234]

The CTM and C1C4 and CC2C3 cCTM variants exhibited UV-VIS spectra consistent with ferric and ferrous c-type hemoproteins.[234] The redox properties of the mono-heme CTMs, bCTMs and C1C4 was also investigated. The mono-heme CTMs exhibited midpoint potentials across a range of 39 mV, from -181 mV for C3 to -220 mV (vs NHE) for C4. The bCTMs had two distinctive redox events separated by a midpoint potential difference of 40 mV, 49 mV and 62 mV (vs NHE) for BC1, BC3 and BC4, respectively. In addition, the cCTM C1C4 maquette also recorded two separate redox events, at -154 mV and -207 mV (vs NHE).[234] The thermal stability of the CTMs, bCTMs and C1C4 was also investigated by circular dichroism. The CTMs exhibited thermal melting temperatures (T_m) across a $24^{O}C$ range; incorporation of a b-type heme into the CTMs elevated the T_m of all the maquettes into the 60-$70^{O}C$

range.[234] For the CTMs, the authors report the differences in the orientation of heme at the binding site plays a critical role in the observed T_M. The heme propionates are either positioned towards the solvent (C3 and C4) or towards the hydrophobic core (C1 and C2); the displacement of the charged propionates from the hydrophobic core engenders a more stable conformation, which is reflected in the respective T_m values (C3 > C4 > C1 > C2). It was also suggested, *via* MD simulations, that the charged propionates of the C3 and C4 heme formed stabilizing salt bridges with lysine residues from the hydrophilic helix regions.[234] In the bCTM maquettes, the orientation of the b-type heme is not restricted by covalent attachment to the protein scaffold, and therefore adopts an orientation that positions the propionates towards bulk solvent thus facilitating the formation of electrostatic interactions with surface lysines, accounting for the elevated T_m values for bCTMs vs CTMs.[234] Finally, the oxygen binding and autoxidation of the CTM variants was examined. The oxygen binding affinity for all CTMs was determined under anaerobic conditions by mixing the maquette with oxygenated buffer and monitoring the binding event spectroscopically. The observed first-order rates for dioxygen binding (k_{obs}; s^{-1}) indicated the CTM variants retained the ability to bind oxygen despite changes in the heme C binding site.[234] The formation of an oxy-ferrous complex has also previously been demonstrated to be succeeded by an autoxidation event, resulting in the generation of superoxide. The observed rates of autooxidation (k_{obs}; s^{-1}) for C1-4 were 0.25, 0.27, 0.46, and 0.27 respectively.[234] The work demonstrated, again, the flexibility and plasticity of the maquette approach, and demonstrated that the biophysical properties of CTMs could be fine-tuned towards a desired oxidoreductase function.

In 2017, Watkins *et al* extended the maquette approach towards the design of an *in vivo* constructed CTM with proficient peroxidase activity.[235] The maquette was designed from the C4 CTM by introducing three key histidine to phenylalanine mutations: mutating the two histidines that coordinate to the iron and form the heme b binding site (on helices 1 and 3) and the histidine that provides the sixth ligand to the iron at the c-type active site (on helix 2) (**figure 56**).[235] The result was a pentacoordinate CTM called C45 (C = c-type, 4 = helix the axial histidine residues on, 5 = the coordinate number of the iron (**figure 56**)).

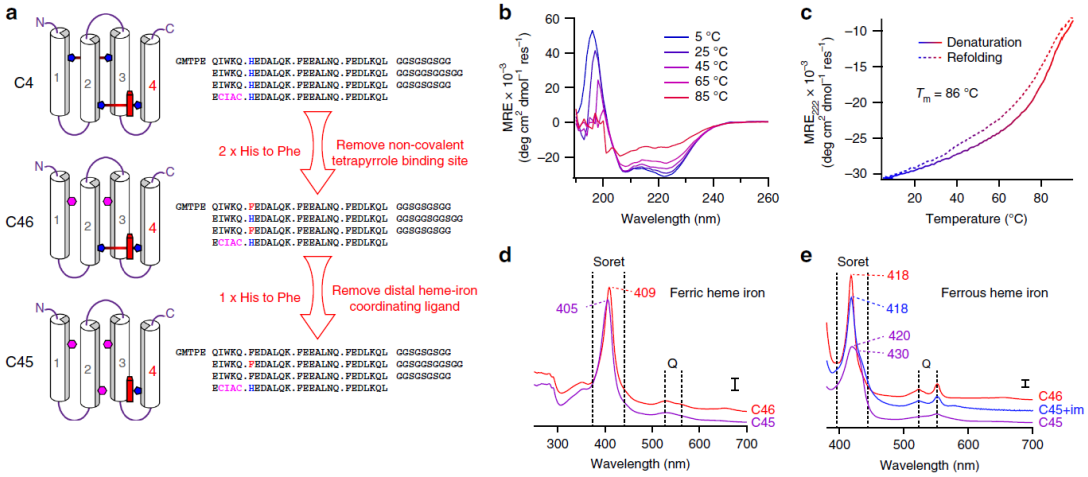


Figure 56: (**a**) The engineering of the monohistidine ligated CTM C45 from C4 *via* the intermediary bis-histidine CTM C46, (**b**) CD spectra of C45 across an 80^{O}C temperature range, (**c**) denaturation and refolding of C45 measured at 222 nm, (**d**) the UV-VIS spectra of ferric C46 (red) and C45 (blue), (**e**) UV-VIS spectra of ferrous C46 (red), C45 (blue), and C45 in excess imidazole (magenta). Figure taken from Watkins *et al*[235]

The UV-VIS spectra of ferric and ferrous C45 resembled the electronic spectra recorded for the other monohistidine-ligated cytochrome *c* variants. The electronic spectra of ferric C45 is pH-dependent, indicting the vacant sixth coordinate site in C45 is occupied by a titratable water molecule (in the resting state).[235] The presence of a labile water molecule is also indicated from the pH-dependence of the heme redox potential, which varies from -208 mV (pH = 10) to -174 mV (pH = 8.6), indicative of a water/hydroxide equilibrium.[235] The UV-VIS spectra of C45 in excess imidazole reproduces the spectra of C4 and C46 (a bis-histidine-ligated CTM resulting from two histidine to phenylalanine mutations on helices 1 and 3 respectively). The three histidine to phenylalanine mutations afford remarkable thermostability, with a T_m of 86^{O}C reported for C45. C45 was also proficient at binding O_2, exhibiting a 10-fold increase in binding affinity relative to C4. The presence of a liable water axially coordinated to the histidine residue opens up the possibility that C45 could be employed for catalysis.[235]

The enzymology of peroxidases was introduced in *section 2.2.2.2*. The catalytic cycle of peroxidases, irrespective of structural differences, are characterised by the formation of an active intermediate called compound I. A common assay for measuring peroxidase activity is mixing an enzyme and hydrogen peroxide together in the presence of 2,2'-azino-bis(3-ethylbenzothiazoline 6-sulfonic acid (ABTS).[274] In the presence of an active peroxidase, ABTS is converted into a green cationic radical, allowing for peroxidase activity to be detected spectroscopically (**figure 57**). Mixing ABTS with ferric C45 and H_2O_2 resulted in the rapid formation of a green solution, diagnostic of peroxidase activity.[235]

The steady-state kinetics of C45-catalysed ABTS turnover, as a function of peroxide concentration, were determined and displayed classic Michaelis-Menten behaviour (**figure 57**).

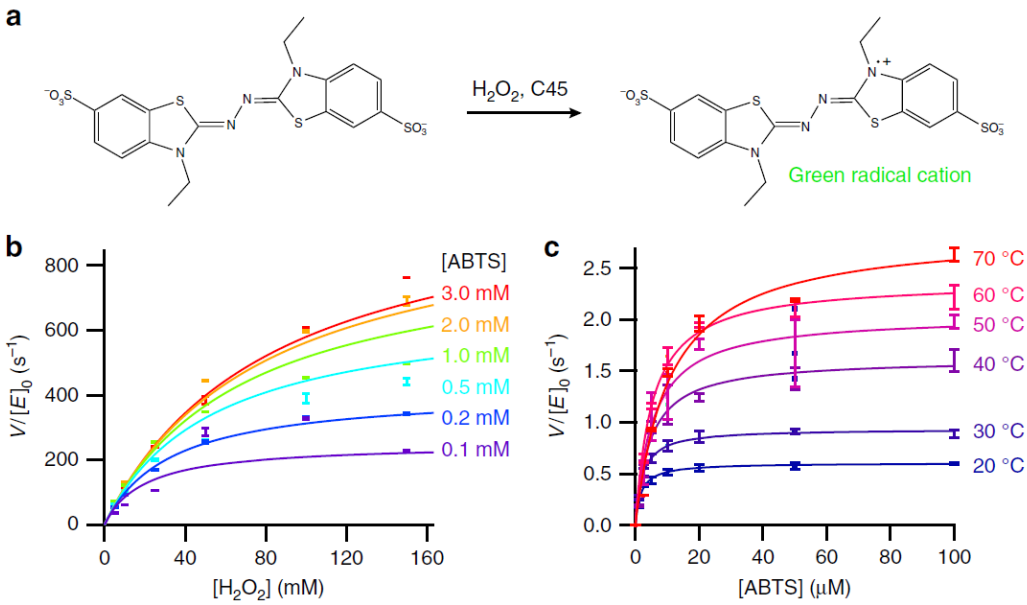


Figure 57: The steady-state of ABTS and hydrogen peroxide turnover by C45. **(a)** The reaction between peroxide-activated C45 and ABTS generates a green cationic radical, **(b)** Steady-state kinetics plot for ABTS turnover by C45 at varying [ABTS] and [H_2O_2], **(c)** The thermal stability of C45 enables classical Michaelis-Menten kinetics to be determined up to 70^{O}C. Figure taken from Watkins *et al*[235]

The kinetic behaviour was fitted to a ping-pong mechanism, characteristic of peroxidases, and allowed the kinetic parameters for i) the activation of peroxide by C45, and ii) the electron transfer between peroxide-activated C45 and ABTS to be quantified. The Michaelis-Menten parameters reported for the electron transfer step afforded a catalytic efficiency of k_{cat}/K_M = 3.2 x 10^6 M^{-1} s^{-1}, which is close to the catalytic efficiency reported for the naturally occurring peroxidase horseradish peroxidase (HRP; k_{cat}/K_M = 5.13 x 10^6 M^{-1} s^{-1}).[235] In contrast to the electron transfer step, C45 is less adept at the peroxide activation step (k_{cat}/K_M = 1.3 x 10^4 M^{-1} s^{-1} for C45; k_{cat}/K_M = 4.6 x 10^6 M^{-1} s^{-1} for HRP), which the authors attribute to a lack of peroxidase-activating residues in the active site of C45. C45 was also demonstrated to function as an efficient artificial cytochrome *c* peroxidase (k_{cat}/K_M = 4.3 x 10^5 M^{-1} s^{-1} for C45; k_{cat}/K_M = 1.62 x 10^7 M^{-1} s^{-1} for yeast cytochrome *c* peroxidase).[235] C45 exhibits significant substrate promiscuity, catalysing the oxidation of guaiacol, *p*-anisidine, *o*-phenylenediamine, 5-aminosalicylic acid, luminol, reactive blue 4, reactive black 5, and the oxidative dehalogenation of halogenated phenols such as 2,4,6-trifluorophenol, 2,4,6-trichlorophenol, and 2,4,6-tribromophenol, and 4-bromophenol.[235] The reactive intermediate in all the aforementioned reactions was investigated *via* UV-VIS and EPR spectroscopy. The UV-VIS spectrum of C45 in the presence of H_2O_2/peroxyacids produced a spectroscopic intermediate with a profile similar to the spectrum

reported for compound I in cytochrome *c* peroxidase; in addition, the EPR spectrum indicated the formation of a tryptophan radical species (the residue W43 residues close to the active site heme) also resembling the EPR spectra recorded for compound I in cytochrome *c* peroxidase.[235]

2.5 CONCLUSION: IS C45 A POSSIBLE CARBENE TRANSFERASE?

Hemoproteins are an extensive class of metalloenzymes characterised by the incorporation of a redox-active iron-metallated tetrapyrrole cofactor (heme) into the protein backbone. The incorporation of the heme cofactor allows hemoproteins to perform biological functions that would otherwise be inaccessible *via* the twenty canonical amino acids. These functions include electron transport, small ligand sensing, small molecule transportation, peroxidase chemistry, and monooxygenase chemistry. The catalytic cycle exhibited by peroxidase and monooxygenase enzymes all proceed *via* a universal, heme-localised cationic radical reactive intermediate called compound I. The intrinsic chemistry of the heme cofactor is, however, a lot richer than is otherwise observed in nature. The developments in molecular biology have afforded novel methodologies for the re-engineering of proteins towards abiological functionality, with the introduction of key mutations and modifications allowing for novel, abiological chemistry to be unlocked. It was demonstrated that cytochrome P450 monooxygenases and globins were identified as suitable candidates for designing artificial hemoproteins proficient towards carbene transferase chemistry on account of the analogous nature of compound I and isolated iron-tetrapyrrole metallocarbenoids. Through iterative rounds of directed evolution naturally occurring hemoproteins could be engineered towards metallocarbenoid formation and subsequently demonstrated proficient chemo-, regio-, and stereoselective carbene transferase activity towards a host of suitable substrates, including olefins (cyclopropanation), thiols, amines, phosphines, borides, C-H bonds (X-H insertion) and carbonyls (carbonyl olefinations).

C45 (**figure 58**) is a thermostable pentacoordinate, c-type tetra-α-helical maquette capable of being assembled *in vivo* and functions as a promiscuous and proficient *de* novo peroxidase. The UV-VIS and EPR spectrum acquired for C45 in the presence of peroxides indicates its catalytic activity is a consequence of its ability to form a heme-localised reactive intermediate called compound I, with the EPR spectra suggesting the cationic radical is localised on an active site tryptophan residue, analogously to compound I in cytochrome *c* peroxidase. The similarities between the carbene transferases engineered from naturally occurring hemoprotein scaffolds (P450s and globins) and C45 engendered two simple questions: i) can a heme-localised metallocarbenoid intermediate by isolated and characterised with C45, and ii) if it can, is it possible to use C45 as a *de novo* carbene transferase *in vitro* and/or *in vivo*? These two questions will be addressed in detail in the following chapters.

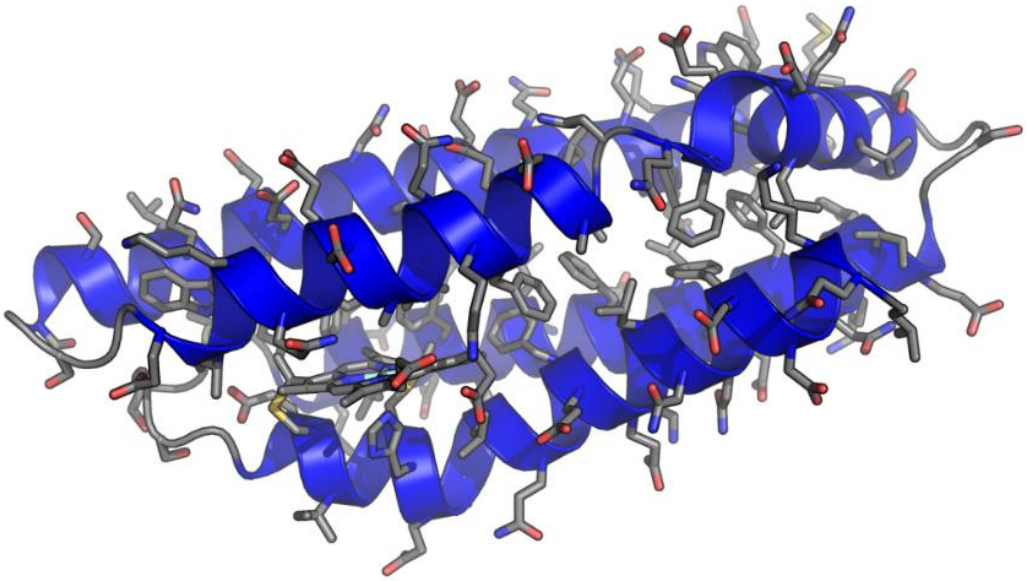


Figure 58: An MD simulation of the solution structure of the CTM C45.

CHAPTER 3: SPECTROSCOPIC CHARACTERISATION AND MASS SPECTROMETRY OF BIOORGANOMETALLIC METALLOCARBENOID COMPLEXES IN HEMOPROTEINS AND *DE NOVO* MAQUETTES

3.1 A BRIEF HISTORY OF IRON PORPHYRIN CARBENES (IPCS)

An iron-bound tetrapyrrole metallocarbenoid (IPC) was first hypothesised in 1974 by Mansuy and Ulrich in attempts to explain the metabolism of halothane in the body.[275] It was postulated that halothane is broken down by a microsomal cytochrome P450 and its anabolism proceeds *via* a metallocarbenoid intermediate. The UV-VIS spectrum obtained for P450 in the presence of halothane produces a distinctive spectrum characterised by a shift in the Soret peak.[275] It was shown the UV-VIS spectrum for P450 in the presence of halothane could be reproduced when a known carbene precursor, trifluorodiazoethane, was added to a solution of P450 under reducing conditions. Furthermore, trifluoroacetic acid was obtained under aerobic conditions, which hinted at the formation, and subsequent hydrolysis, of trifluorodiazoethane *via* a carbene intermediate.[275]

In 1977 it was demonstrated that cytochrome P450s (isolated from liver) could reduce tri- and tetra-halogenated methanes (i.e. CBr_4, CCl_4, CCl_3F, CCl_3Br, CCl_3CN, CHI_3, $CHBr_3$, and $CHCl_3$).[276] Under anaerobic conditions NADPH was oxidised and the production of CO was detected *via* IR-spectroscopy; the production of CO supports the formation of a metallocarbenoid intermediate (**figure 59**), as it had previously been isolated as a breakdown product of Cr(III)-halogenated-carbene complexes.[276] A crystal structure for the carbene complex, Fe(TPP)(CCl_2)(H_2O), was reported in 1978 (**figure 59**), and confirmed the hypothesis that a metallocarbenoid intermediate was responsible for the reduction of halothane and polyhalogenated methanes.[189] In addition, the IR-spectrum of the complex exhibited a band at 872 cm^{-1} which was assigned to a C-Cl stretching vibration (the band is absent in Fe(TPP)$(py)_2$ and Fe(TPP)(py)(CO) complexes) and shifts to 841 cm^{-1} in the IR-spectrum for Fe(TPP)($^{13}CCl_2$).[189] Interestingly, the crystal structure revealed the carbene bound to the metal in the terminal mode, as opposed to the possible bridging mode (a result of insertion into a Fe-N bond). It was later demonstrated, in 1978, that cytochrome P450s metabolise DDT *via* a vinylidene complex,[277] and in 1979 a metallocarbenoid intermediate was attributed to the potency of the 1,3-benzodioxole series of insecticide synergists, suggesting metallocarbenoid formation irreversible inhibits P450 and thus prevents the metabolic breakdown of the active insecticide.[278]

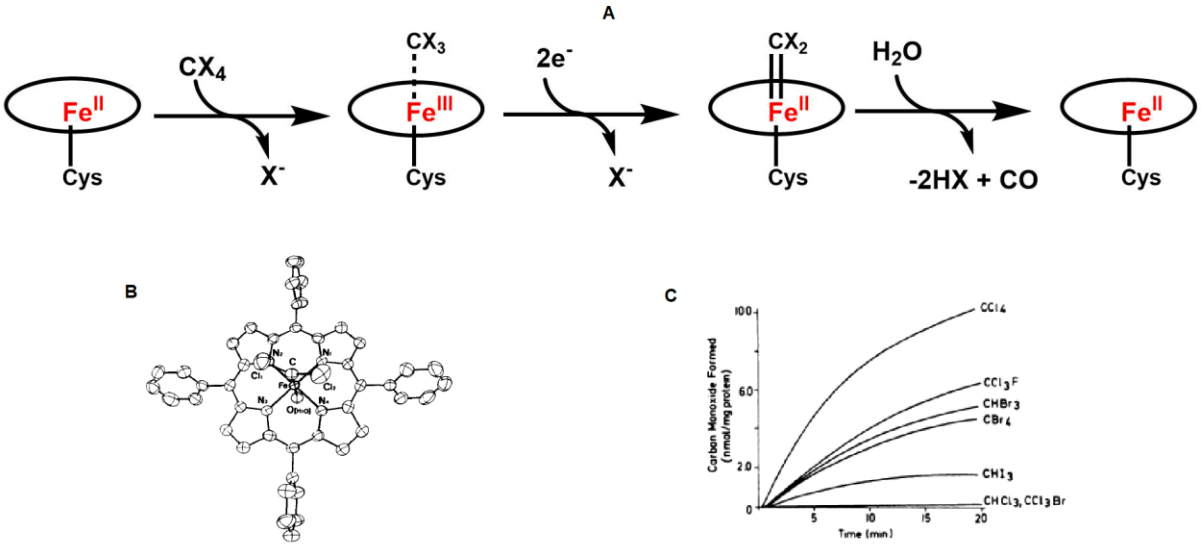


Figure 59: A) The general reaction scheme for the cytochrome P450 catalysed dehalogenation of polyhalogenated methanes, **B)** the crystal structure of the carbene complex Fe(TPP)(CCl_2)(H_2O) (figure taken from Mansuy *et al*),[189] and **C)** the rate of carbon monoxide production following the addition of various polyhalogenated methanes to reduced rate liver cytochrome P450 (figure taken from Wolf *et al*).[276]

Iron tetraporphyrins are well established catalysts in hydroxylation and epoxidation chemistry.[279–282] In nature, the cytochromes P450 are a superfamily of hemoproteins that perform monooxygenation reactions using molecular oxygen or hydrogen peroxide.[101,283,284] P450s are found ubiquitously in biology, with roles ranging from the synthesis of biomolecules (i.e. steroidogenesis) to the metabolism of xenobiotics. All P450s capable of performing monooxygenation reactions proceed *via* a ubiquitous intermediate called compound I – a radical cation iron(IV)-oxene species.[101,126,283–285] In 1995 Robbins Wolf and Woo employed ferrous tetraphenylporphyrin (Fe^{II}(TPP)) as a catalyst in the stereoselective cyclopropanation of alkenes.[286] The active intermediate was postulated to be an iron-localised metallocarbenoid, a postulation that was supported by secondary kinetic isotope effects using styrene-d_8.[286] Interestingly, an iron-localised-metallocarbenoid intermediate is analogous to compound I, the active intermediate in monooxygenases.[101,285] As carbene transfer chemistry and monooxygenation chemistry are both catalysed by an iron-tetrapyrrole catalyst, and the mechanism proceeds *via* analogous intermediates, the question of whether hemoproteins could catalyse carbene transfer reactions *via* a metallocarbenoid intermediate, instead of or in addition to monooxygenation reactions, emerged.

3.2 ARTIFICIAL CARBENE TRANSFERASES

The successful employment of iron tetrapyrrole complexes as catalysts in abiological reactions, such as carbene transfer chemistry, prompted the question of whether hemoproteins could be altered towards abiological chemistry. The rationale was that key mutations to an enzyme's sequence, for example *via* iterative rounds of directed evolution, could supress its natural functionality while concurrently unlocking novel functionality undocumented in nature.[91,94,287] Through directed evolution, engineered hemoproteins (starting from myoglobin and cytochrome P450 blueprints) have been successfully utilised as biocatalysts in cyclopropanation,[192,195,199,209,210] intra- and intermolecular X-H insertion (where X = C, N, S, Si, B)[17,210,215–217,220,221,288] and carbonyl olefination reactions.[229,230]

The mechanism of biocatalytic carbene transfer reactions is hypothesised to proceed *via* a heme-localised metallocarbenoid intermediate.[187,289] Computational analysis has indicated a metallocarbenoid complex in a tetrapyrrole-ligated iron would afford an electrophilic carbene as electron density is donated from the carbene orbitals into the vacant orbitals of iron (discussed in detail below).[38,39,290,291] These predictions are consistent with two pieces of empirical evidence. Firstly, a mutation of the axial ligand that coordinates to the heme iron, and trans to the reactive intermediate (oxene, carbene etc), to a less-efficient electron donor (i.e. cystine to serine) results in diminished monooxygenation activity and improved carbene transfer activity.[192,193,292,293] This is rationalised on account of electronics. In the case of monooxygenation, as less electron-donation from the occupied axial ligand orbitals to the unoccupied iron orbitals results in less electron density being available in the iron orbitals for donating into the π^* orbital of O_2, a crucial step in the heterolytic cleavage of O_2 leading to the formation of compound I. In the case of the metallocarbenoid intermediate, the decreased electron donation from the occupied orbitals of the axial ligand to the unoccupied orbitals of iron results in a greater degree electron donation from the occupied carbene orbitals to the empty iron orbitals, meaning electron density is shifted away from the carbene carbon and results in increased electrophilicity. Secondly, the improved reaction efficiency observed in cyclopropanation/X-H insertion reactions catalysed by engineered hemoproteins, where an olefin/X-H group (NH, SH etc) serves as the nucleophile in the reaction. This is corroborated by the Hammett plots and Hammett constants reported in the literature for various functionalised styrenes: electron-donating groups in the *para*-position of the styrene aromatic ring increase the nucleophilicity of the olefin nucleophile and result in an increased reaction efficiency, whereas electron-withdrawing groups decrease the nucleophilicity and consequently decrease the reaction efficiency.[52]

In 2018 the Arnold group reported the first crystal structure of an iron-localised metallocarbenoid intermediate using the hemoprotein cytochrome *c* isolated from *Rhodothermus mainus (Rma)* (**figure**

60).[219] The importance of the crystal structure is two-fold. Firstly, the crystal structure provided the first direct evidence for the existence of an enzymatic metallocarbenoid intermediate, and confirmed that a bioorganometallic intermediate, previously undocumented in nature, was involved in the abiological chemistry exhibited by modified hemoproteins.[219] Secondly, the crystal structure allowed for the key interactions between amino acids in the protein's primary/secondary spheres and the metallocarbenoid intermediate to be elucidated, affording an insight into how further modifications of the protein structure could be introduced to stabilise the formation of the metallocarbenoid intermediate, improve the rate of carbene transfer kinetics, and increase the stereoselectivity of carbene transfer reactions.[218,219,294]

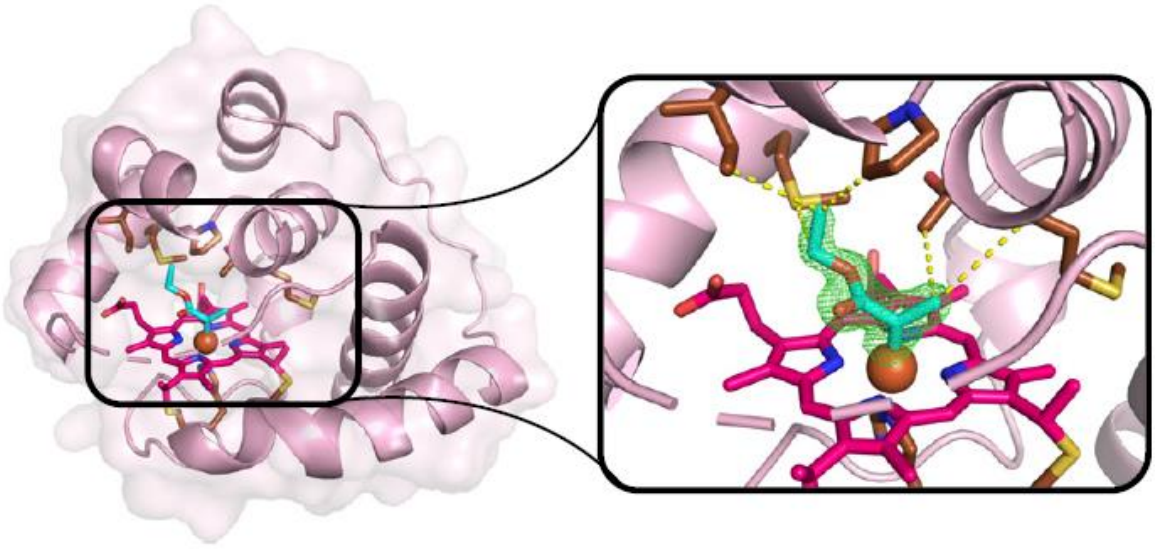


Figure 60: The crystal structure of the metallocarbenoid complex formed between ethyl diazoacetate (EDA) and the *Rhodothermus mainus* cytochrome *c* variant *Rma*-TDE (figure taken from Lewis *et al*).[219]

3.3 COMPUTATIONAL STUDIES OF IRON PORPHYRIN METALLOCARBENOID (IPCS): THEORETICAL EXAMINATION OF IPC FORMATION AND ELECTRONIC STRUCTURE

The isolation of metallocarbenoid intermediates in crystal structures of iron-porphyrins[189] and cytochrome *c*[219] validated the role of metallocarbenoid intermediates in carbene transferase activity previously reported in the literature by Arnold,[184,187] Fasan,[195,199,213,216,288,291] Brustad,[289,295–297] Coelho.[190,191,298] However, further questions concerning i) the reaction mechanism for the formation of IPCs from diazo compounds, ii) the electronics governing IPC formation, and iii) the electronic structure of IPCs needed to be addressed. To date, several computational studies have been reported and have provided valuable insights into the aforementioned questions and have made predictions that are in line with experimental observations and reports.[38–41,290,291]

3.3.1 Iron porphyrin carbenes (IPCs): formation pathway and complex electronics

The theoretical reaction mechanism for iron-porphyrin carbenoid formation from diazo compounds is illustrated in **figure 61**. The proposed pathway starts with a tetrahedral intermediate following the formation of a σ-bond between iron and the carbenoid carbon. A concerted elimination ensues where a π-bond between the iron and the carbenoid forms at the expense of N_2, which is eliminated to yield the iron-porphyrin carbenoid.[61,291] The loss of entropy following conversion of the carbenoid carbon from a sp^3 hybridisation state in the tetrahedral intermediate to a sp^2 hybridisation state in the IPC is compensated by the significant increase in entropy accompanying the expulsion of N_2, and likely acts as the thermodynamic driving force of the reaction.

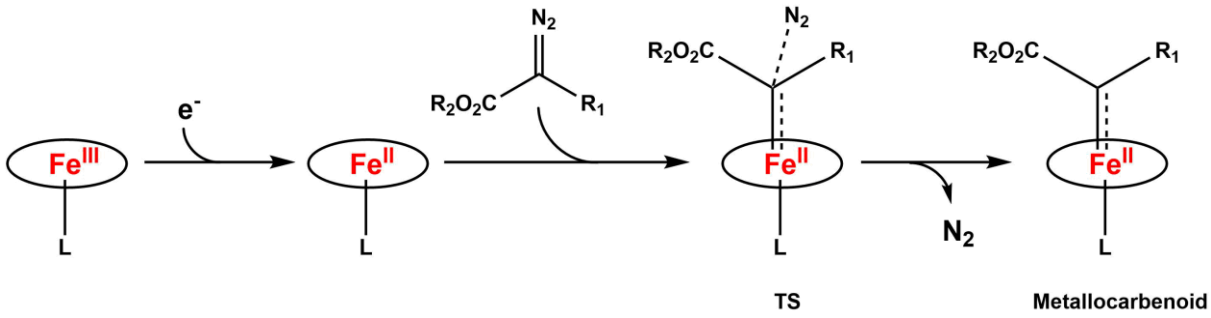


Figure 61: The formation of an iron porphyrin metallocarbenoid (IPC) from a diazo compound proceeding *via* a tetrahedral transition state (TS).[38,40,290]

3.3.1.1 Formation of IPCs via a tetrahedral intermediate

Zhang and co-workers studied the IPC formation pathway computationally and provided some key insights into the nature of the initial tetrahedral intermediate.[38,40,290] It was reported that the largest bond change accompanying tetrahedral formation is the C-N bond on the diazo compound, with an elongation of approximately 0.5 Ä.[40] The elongation of the C-N bond is accompanied by an increase in the bond angles around the carbon substituents as a consequence of a deviation away from the planar conformation of the diazo compound.[40]

It was noted that a large charge transfer from the N_2 substituent to the carbene carbon is a crucial electronic step governing the rate of formation of the tetrahedral intermediate. A subsequent charge donation from the carbene to the iron in the transition state results in an increase in negative charge on the iron which is accompanied by an increase in positive charge on the carbene carbon.[40] This is corroborated by considering the influence the electronic nature of the carbene substituents have on the rate of IPC formation: electron-donating groups (EDG) facilitate the formation of the IPC because the charge donation helps stabilise the development of a positive charge on the carbene carbon, EWG

result in an elevated reaction barrier to IPC formation by destabilising the carbene carbon in the transition state. It was concluded from the computational experiments that the carbene-to-iron charge transfer is the dominant step governing the formation of IPCs

3.3.1.2 Electronic structure of IPCs

An important question concerns the electronic structure of a ground-state iron-porphyrin-carbene complex. An understanding of the frontier orbitals, electronic configuration, nature of the iron-carbene bond, and oxidation state of the iron are important parameters that need to be properly understood to improve the stability and reactivity of IPCs. Zhang and co-workers initially addressed the nature of the iron-carbene bond and the Fe(II)/Fe(IV) problem using Mossbauer, NMR, and computational simulations.[38,290] Zhang and co-workers were particularly interested in addressing whether the bonding in IPCs is best described as a Fe(II)←[:C(X)(Y)] or Fe(IV)=[C(X)(Y)]$^{2-}$ resonance structure. The calculations performed using geometrically optimised structures of IPCs indicated the ground-state of IPCs are best described as diamagnetic low-spin complexes.[35] The Fe(II)←[:C(X)(Y)] resonance structure results from a dominant carbene-to-iron charge transfer which renders a positive charge on the carbene carbon and is consistent with the reported chemical shifts of IPCs in NMR spectroscopy. The Fe(IV)=[C(X)(Y)]$^{2-}$ resonance structure is characterised by π-back donation from the iron to the carbene which would result in the build-up of a negative charge on the carbene carbon; an electronic structure inconsistent with the NMR data and the observed electrophilicity of IPCs.[35]

The next question concerned the electronic configuration of the frontier molecular orbitals in IPCs. Shaik and co-workers expanded upon the work of Zhang by calculating the molecular orbital configurations for both penta- and hexa- coordinated IPCs for both Fe(II) and Fe(IV) complexes, and demonstrated that in the d^6 Fe(II) complexes the orbital configuration is $(dxy)^2(dxz)^2(dyz)^2(dz^2)^0(dx^2\text{-}y^2)^0$, consistent with the proposed low-spin ground state.[39] In d^4 Fe(IV) complexes the orbital configuration is $(dxy)^2(dxz)^2(dyz)^0(dz^2)^0(dx^2\text{-}y^2)^0$, also consistent with a low-spin ground state.[39] Shaik and co-workers also investigated the electron distribution in Fe(II) d^6 ground-state IPCs.[39] The three possible electron configurations that are possible in a Fe(II) d^6 IPC are closed-shell-singlet (CSS), open-shell singlet (OSS) and closed-shell-triplet (CST) (**figure 62, A and B**).[39]

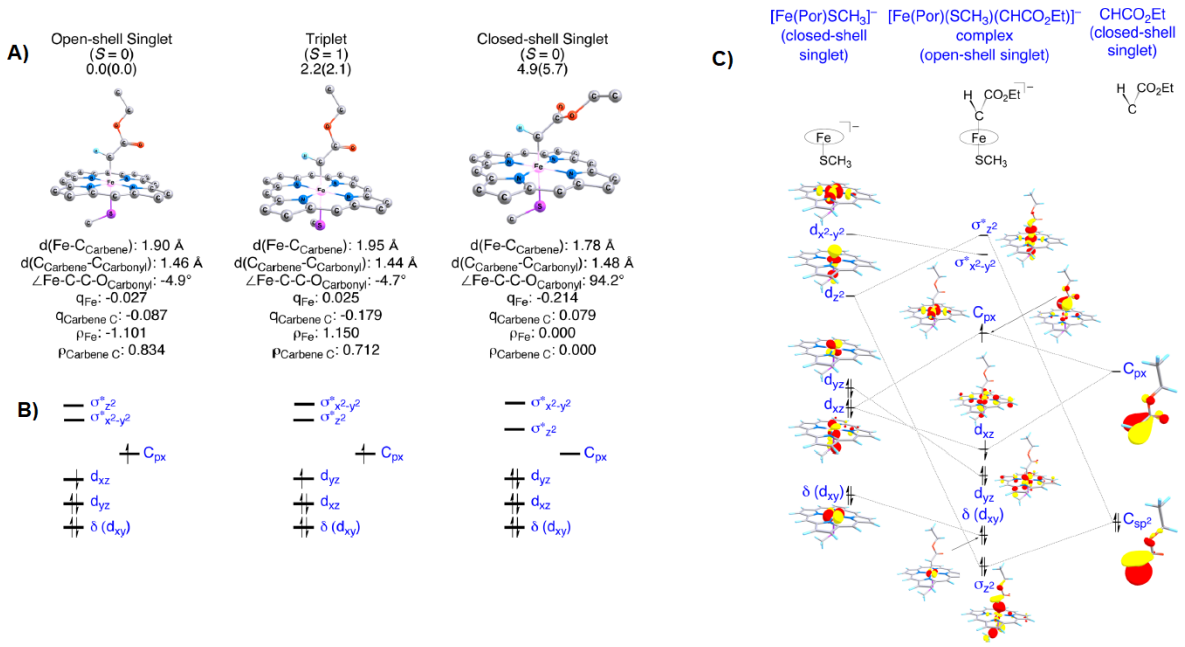


Figure 62: IPC orbitals (figure taken from Sharon *et al*)[39]

The model presented by Zhang operated under the assumption that a CSS was the energetically favourable configuration of IPCs, but the calculations performed by Shaik and co-workers rendered the OSS electronic configuration as the energetically favourable electron distribution (**figure 62, C**).[39] It was rationalised that an OSS electronic state occurs from the nature of the interacting orbitals on the iron and carbene in IPCs. A carbene sp^2 orbital with the correct symmetry overlaps with the dz^2 orbital of iron, donating its electron pair, to form the initial Fe-C σ-bond.[39] Additionally, the doubly occupied dxz orbital on iron, which possesses π-symmetry, weakly interacts with the vacant carbene p_x-orbital. This configuration can be best described as generating an iron-localised dxz orbital and a carbene localised px orbital, with one electron occupying each. This results in an electronic configuration of $(dxy)^2(dyz)^2(dxz)^1(Cpx)^1(dz^2)^0(dx^2\text{-}y^2)^0$, with the lone-electrons in dxz and Cpxz anti-ferromagnetically coupled, giving a diamagnetic open-shell low-spin configuration.[39] The question concerning ground-state IPCs has yet to be completely answered, but what can be deduced is that i) carbene-to-metal charge transfer is the dominant process governing IPC formation, and ii) IPCs are low-spin singlet, diamagnetic complexes possessing a Fe(II) metal.[38,290] The singlet nature of IPCs could also partially account for the observed stereoselectivity observed with iron-tetrapyrrole catalysts and the reported hemoprotein carbene transferases.

One last consideration investigated by Zhang and co-workers explored the energetic differences between terminal- and bridging- carbene complexes.[40] Bridging carbene complexes have been reported in Co(II) complexes, but the calculations reported by Zhang indicate the bridging mode in

IPCs is significantly higher than in Co(II) carbene complexes (13.17 kcal/mol higher).[40] This suggests the terminal-binding mode in IPCs is the dominant interaction, a conclusion consistent with the reported crystal structures of IPCs and the IPC in the cytochrome *c* crystal structure reported by Lewis *et al*.[219] It was further reported that the preference in terminal binding in IPCs is a consequence of the ordering of the d-orbitals in IPCs relative to Co(II)-carbene complexes.[40] In terminal binding IPCs, the dz^2 is unoccupied on account of the greater repulsion between the carbene and the dz^2 orbital, while the dyz orbital, less-aligned with the carbene, is occupied. In the bridging mode, the electron density on the carbene is shifted away from the dz^2 axis and is more aligned with the dyz orbital, resulting in greater repulsion between the carbene and the dyz orbital in the bridging mode. This results in the dz^2 and dyz orbitals being energetically reversed in the bridging mode relative to the terminal mode and as a consequence the dz^2 is occupied while the dyz orbital is unoccupied.[40]

3.4 UV-VIS SPECTROSCOPY OF HEME-CONTAINING PROTEINS – APPLICATIONS TO STUDY METALLOCARBENOID FORMATION IN NATURAL HEMOPROTEINS AND *DE NOVO* HEME-CONTAINING MAQUETTES

Although crystal structures provide the most reliable and direct evidence for enzymatic reaction intermediates, not all proteins are capable of being crystallised. The *de novo* tetra-α-helix c-type maquettes are a class of proteins that have proved notoriously difficult to crystallise. Therefore, an alternative methodology for studying the formation and characteristics of a metallocarbenoid intermediate needs to be employed.

Heme-containing proteins are characterised by distinctive UV-VIS spectroscopic profiles as a consequence of the π-conjugation of the porphyrin cofactor.[143,299,300] The UV-VIS profiles of porphyrins are a consequence of electronic transitions between two occupied (HOMO) ground state (S_0) π orbitals, a_{1u} and a_{2u}, and two unoccupied (LUMO) degenerate (S_1 and S_2) π* orbitals, e_{gx} and e_{gy} (**figure 63**).[104,143,300] Electronic transitions between these orbitals results in an intense Soret peak around 400-430 nm and a region of multiple, less intense peaks in the 500-650 nm range called the Q-band.[299,300] The Soret peak arises from an $S_0 \rightarrow S_2$ electronic transition from the heme π_{a1u} to π* orbitals, with the intensity arising from the transition being spin and symmetry allowed. The Q band arises from an $S_0 \rightarrow S_1$ electronic transition from the π to π* orbitals, with the splitting of the Q-band a consequence of vibrational coupling with the $S_0 \rightarrow S_1$ electronic transition, resulting in two possible transitions, $Q_{0,0}$(v=0) and $Q_{1,0}$,(v=1) from the ground electronic state, S_0 ($S_{0v0} \rightarrow S_1$ and $S_{0v1} \rightarrow S_1$).[299–301] (In free base porphyrins, the presence of two NH groups breaks the symmetry of the porphyrin macrocycle and further splits the two Q-band peaks resulting in four peaks ($Q_{X0,0}$, $Q_{Y0,0}$, $Q_{X1,0}$ and $Q_{Y1,0}$) in the Q-band region (**figure 63**)).[104,143,299–301]

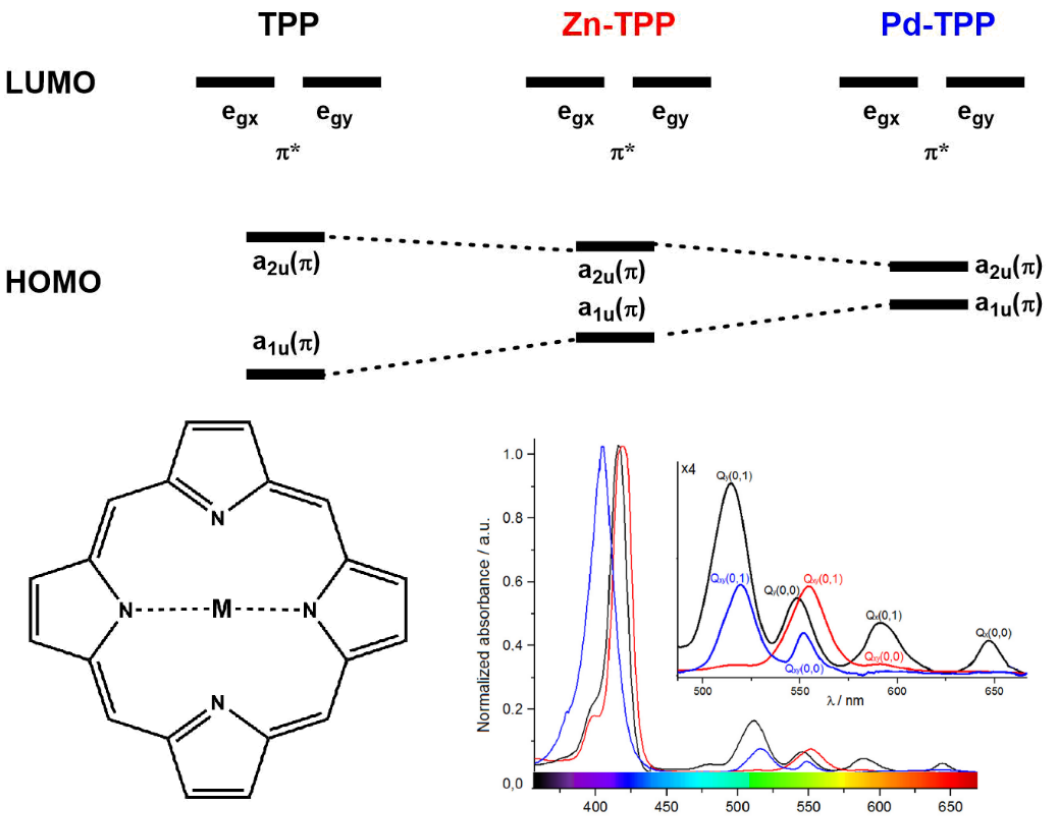


Figure 63: The UV-VIS spectra of free-base and metalled porphyrins. **(top)** the relative molecular orbitals of tetraphenylporphyrin (TPP, black) Zn-TPP (red) and Pd-TPP (blue), **(bottom left)** the general molecular structure of a metalloporphyrin, and **(bottom right)** the UV-VIS spectra for TPP (black), Zn-TPP (red) and Pd-TPP (blue) with the Q-bands magnified (figure taken from Dabrowski *et al*).[301]

The spectroscopic profiles of heme-containing proteins are sensitive to the oxidation state of the metal, the coordination state of the central metal, and the properties of the coordinating ligands (i.e. electronics).[104,143] Any changes in these parameters, caused by a chemical transformation, would alter the frontier orbitals of the system and would engender pronounced changes in a hemeprotein's electronic spectrum. The formation of a metallocarbenoid intermediate, and hence a shift from a pentacoordinate to a hexacoordinate complex, should, therefore, have a traceable effect on the absorbance profile of the UV-VIS spectrum.[104,143] In the presence of a suitable carbene precursor (i.e. diazo compounds) three scenarios are possible: (i) no chemical reaction between the protein and the carbene precursor, (ii) the formation of a ferrous hexacoordinate metallocarbenoid intermediate terminally bound at the iron, and (iii) an insertion reaction where a carbene inserts into a Fe-N bond of the porphyrin ring to yield a Fe-C-N bridged complex. Scenario (i) would result in no observable changes in the UV-VIS spectrum, whereas scenarios (ii) and (iii) would produce clear and distinctive changes. The bridged complex would result in modification of a porphyrin nitrogen and would result a detectable change in the Q-band profile; the formation of a terminally bound metallocarbenoid

intermediate would not disrupt the porphyrin, resulting in only noticeable shifts in the Soret peak and Q-band profile respectively. Therefore, UV-VIS spectroscopy is a suitable technique for investigating the formation, and possible subsequent decay, of a metallocarbenoid intermediate in the absence/presence of a reactive substrate (i.e. styrene). Initial investigations into metallocarbenoid formation (using ethyl diazoacetate (EDA) as the carbene precursor) employed several naturally occurring heme protein scaffolds, including myoglobin, zinc-chlorin-e6 reconstituted myoglobin, haemoglobin, horseradish peroxidase, cytochrome *c*, and the cytochrome *c* digestion product microperoxidase-11, before extending the methodology to examine metallocarbenoid formation in the *de novo* tetra-α-helix c-type maquettes C45 and JR1. The cytochrome *c* variant reported by Arnold *et al*[219] and a myoglobin variant ((Mb(H64V,V68A)) reported by Fasan *et al*[195] were also examined for comparative purposes.

3.5 RESULTS AND DISCUSSION

3.5.1 Stopped-Flow Kinetics and UV/VIS Spectroscopy

3.5.1.1 Spectroscopic characterisation of metallocarbenoid formation for naturally occurring hemoproteins

The formation of a metallocarbenoid intermediate was investigated by monitoring changes in the distinctive spectroscopic profile of heme-containing proteins. To avoid the rapid decay of a metallocarbenoid intermediate, single-mixing stopped-flow techniques were employed. A solution containing the ferrous enzyme of interest (15 µM protein, 10 mM $Na_2S_2O_4$, CHES buffer, pH 8.6) and a solution containing ethyl diazoacetate (EDA) (100 µM-10 mM, 20-100% EtOH) were isolated in separate compartments (syringes) before a small volume (2x20 µl) of each were injected concurrently into a mixing chamber. The mixing chamber was monitored by UV-VIS spectroscopy over several timescales to monitor any reactions occurring in the mixing chamber. Final concentrations were 7.5 µM enzyme, 50 µM-5 mM EDA, 5mM $Na_2S_2O_4$, and 10-50% EtOH. The spectra acquired for myoglobin, zinc-chlorin-e6 reconstituted myoglobin, haemoglobin, horseradish peroxidase, and cytochrome *c* are presented in **figure 64** and discussed below.

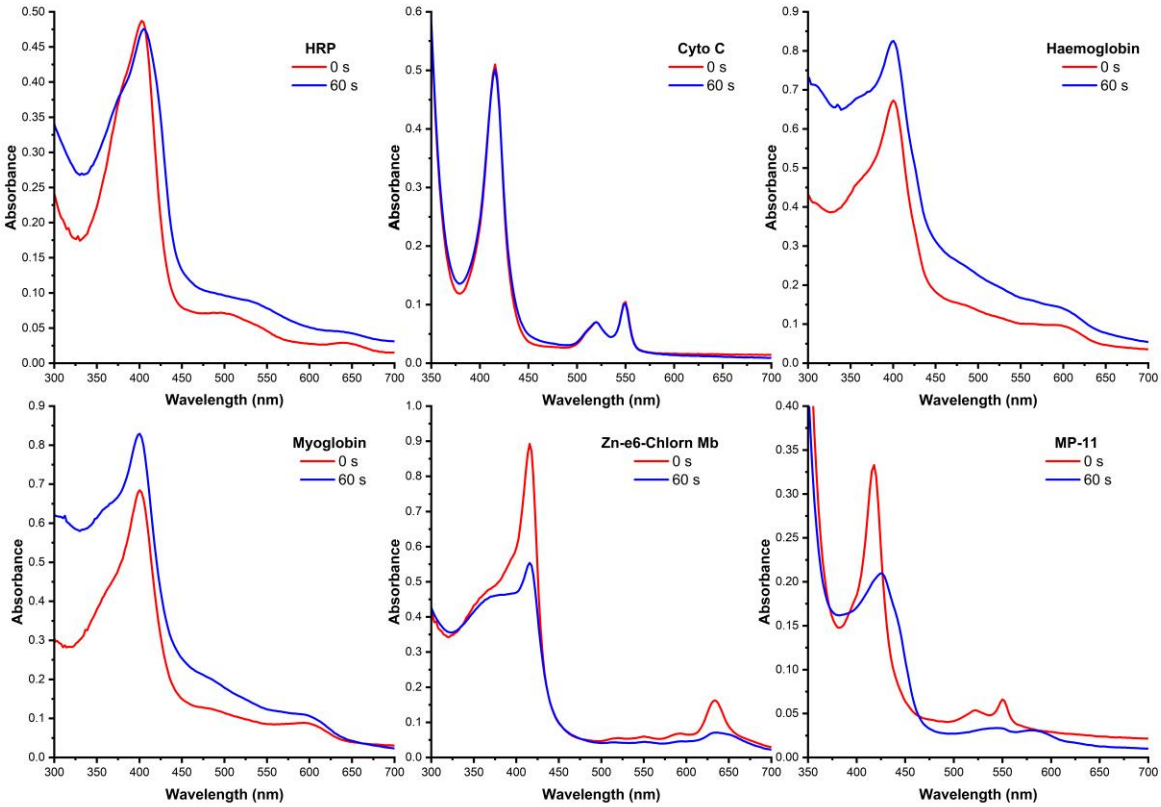


Figure 64: The stopped-flow UV-VIS spectra for metallocarbenoid formation in the presence of ethyl diazoacetate (EDA) at 0 seconds (red) and 60 seconds (blue) for the ferrous hemoproteins (**top left**) horseradish peroxidase, (**top middle)** cytochrome *c*, (**top right**) haemoglobin, (**bottom left**) myoglobin, (**bottom middle**) zinc chlorin-e6 reconstituted myoglobin, and (**bottom right**) microperoxidase-11. All spectra were recorded at 5^{O}C using stopped-flow spectroscopy; final concentrations were 1 mM EDA, 10 µM protein, 1 mM $Na_2S_2O_4$ and 20% (v/v) ethanol.

Horseradish peroxidase is a predominately α-helical protein that coordinates a b-type heme cofactor *via* a backbone histidine ligand to yield a pentacoordinate iron.[101] The UV-VIS spectrum (**figure 64, top left**) for ferrous horseradish peroxidase is dominated by the ferrous Soret peak and a small less-intense Q-band at 500-650 nm. After 60 seconds of mixing with EDA, the UV-VIS spectrum for ferrous HRP displayed a minor shift in the Soret peak and a decrease in intensity of the Q-band. The UV-VIS indicates no metallocarbenoid formation/Fe-N insertion has occurred and the changes are probably associated with degradation. Cytochrome *c* (equine heart) is a small, soluble c-type hemoprotein.[110] The heme is covalently bound to the protein backbone *via* the sequence motif C-X-X-C-H and two axial ligands are provided from a histidine and a methionine, resulting in a hexacoordinate iron.[110] The UV-VIS spectrum for cytochrome *c* (**figure 64, top middle**) is dominated by the ferrous Soret peak and is accompanied by two, sharp and well resolved peaks in the Q-bands at 525 nm and 551 nm respectively. After 60 seconds of mixing with EDA the UV-VIS spectrum displayed no major differences, indicating no metallocarbenoid formation/Fe-N insertion had occurred. This is unsurprising because

the heme is hexacoordinate, meaning there is no vacant site for the carbene ligand. Haemoglobin is an all α-helix protein structurally consisting of two α and two β sub-units that form a tetrameric quarternary structure $(\alpha\beta)_2$ in aqueous solution.[101] Each subunit possesses one b-type heme group tethered to the protein backbone *via* a coordinating histidine, meaning haemoglobin is a tetraheme containing protein.[101] The UV-VIS spectrum for haemoglobin (**figure 64, top right**) is dominated by the ferrous Soret peak and is accompanied by a poorly resolved Q-band in the range 500-600 nm. After 60 seconds of mixing with EDA, the UV-VIS spectrum for ferrous haemoglobin displayed only a decrease in absorbance, indicative of possible heme degradation. Myoglobin is a small, α-helical protein that possesses a pentacoordinate $Fe^{3+/2+}$ b-type heme cofactor.[101] A backbone histidine provides the coordinating ligand that binds the cofactor, leaving a vacant site at the *trans* position for catalysis/binding in the holoprotein (usually occupied by a liable water molecule in the resting state in solution). Myoglobin is structurally characterised by the globin fold, with the heme group sandwiched between two α-helices.[101] The ferrous UV-VIS spectrum for myoglobin (**figure 64, bottom left**) is dominated by the ferrous Soret peak and is accompanied by a less-intense, poorly resolved Q-band in the range 500-600 nm. The UV-VIS spectrum for ferrous-myoglobin parallels the spectra acquired for HRP and haemoglobin: after 60 seconds of mixing with EDA the Soret peak and Q-bands are retained, but there is a general decrease in absorbance, indicative of possible heme degradation instead of formation of a metallocarbenoid intermediate. Zinc-chlorin-e6 reconstituted myoglobin is the myoglobin scaffold with the b-type heme group substituted for Zn-chlorin-e6. Like myoglobin, a backbone histidine coordinates to the zinc to yield a pentacoordinate Zn^{2+} complex. Zinc reconstituted hemoproteins have similar UV-VIS spectra to iron-containing hemoproteins, but the Q-band is red-shifted to 600-700 nm. Unlike iron, zinc can only occur in a M^{2+} oxidation state, but computational analysis of heme-localised metallocarbenoids indicate the metal adopts a M^{2+} oxidation state rendering it possible for a zinc-metallocarbenoid intermediate to form; however, after 60 seconds of mixing with EDA, the UV-VIS spectrum (**figure 64, bottom middle**) was significantly disrupted, probably caused by cofactor degradation, indicating no metallocarbenoid formation/Zn-N insertion had occurred.

Microperoxidase oligopeptides are the proteolysis products of cytochrome *c* enzymes and consist of a c-type heme pentacoordinated to an oligopeptide *via* a backbone histidine ligand. For microperoxidase-11, the oligopeptide is 11 amino acids long and contains the conserved C-X-X-C-H c-type heme binding sequence motif.[302,303] As free heme has been demonstrated to form metallocarbenoid complexes in solution, it is reasonable to expect microperoxidase oligopeptides to display similar chemistry because the heme cofactor is considerably solvent exposed. The UV-VIS spectrum for microperoxidase-11 (**figure 64, bottom right**) is dominated by a Soret peak at 415 nm

with two peaks in the Q-bands at 525 nm and 551 nm respectively (like cytochrome *c*). After 60 seconds of mixing with EDA the UV-VIS spectrum displayed a distinctive red shift in the Soret peak, from 415 nm to 427 nm, which is accompanied by pronounced changes in the Q-band profile (**figure 64, bottom right**). These shifts can be accounted for by two possible explanations: (i) metallocarbenoid formation, or (ii) carbene insertion into a N-Fe bond, disrupting the iron coordination state. Carbene insertion into a N-Fe bond, to form a bridged Fe-C-N complex, would modify a porphyrin nitrogen sufficiently to alter the Q-band profile and produce a spectrum completely uncharacteristic of a six-coordinate heme-containing protein. The formation of a metallocarbenoid would be characterised by concurrent shifts in the Soret peak and Q-band region consistent with the diagnostic spectroscopic features of a hexacoordinate heme-containing protein. The UV-spectra recorded in the stopped-flow experiment for microperoxidase-11 indicates the species formed after rapid mixing is the terminally-bound metallocarbenoid complex, consistent with the crystal structure reported by Arnold.[219]

3.5.1.2 Spectroscopic characterisation of C45 metallocarbenoid formation

Having observed a spectrum consistent with terminal metallocarbenoid formation for the cytochrome *c* derived heme-containing oligopeptide microperoxidase-11, stopped-flow spectroscopy was extended to investigate metallocarbenoid formation in the *de novo* hemoprotein C45. Identical experimental conditions to those outlined above were employed for C45.

A time course stopped-flow UV-VIS spectrum acquired for C45 in the presence of EDA (at 40% EtOH) is presented in **figure 65**. The spectrum of C45 parallels the UV-VIS obtained for microperoxidase-11 (figure 10) under identical conditions. After 60 seconds, a significant red shift in the Soret peak, from 417 nm to 428 m, was observed and is accompanied by a concurrent red-shift and broadening of the Q-band profile. The formation of the metallocarbenoid is visible in the spectrum after 60 seconds. **Figure 66** displays a time-dependent profile of the spectroscopic changes at wavelengths 417 and 433 nm, corresponding to ferrous C45 and the C45-metallocarbenoid complex, respectively. The time-dependent plots show an initial increase in absorbance at 433 nm, accompanied by a paralleled decrease in absorbance at 417 nm, with both traces plateauing after approximately 12 seconds indicating saturation has been established. The spectroscopic data is consistent with the formation of a terminal metallocarbenoid intermediate; however, the rate of formation of the metallocarbenoid intermediate is initially slow. This is probably a consequence of the low temperature and ethanol concentration of the experimental. The spectroscopic data indicates metallocarbenoid intermediate, once formed, is stable and persistent (no appreciable decay over 1000 seconds) and does not react with any residues in the active site (**figure 66**).[294] The UV-VIS spectrum obtained for the formation of

the C45-metallocarbenoid intermediate was unaffected by changes in the ethanol concentration from 10% to 50%, with an observable increase in the initial rate of formation being the only appreciable difference. Tetra-α-helix maquettes, in particular C45, have previously demonstrated good tolerance to high ethanol concentrations[235] without any disruption to their structure or function, and therefore it is unsurprising that the profile of the obtained spectra is unaffected in 40-50% ethanol (**figure 65**).

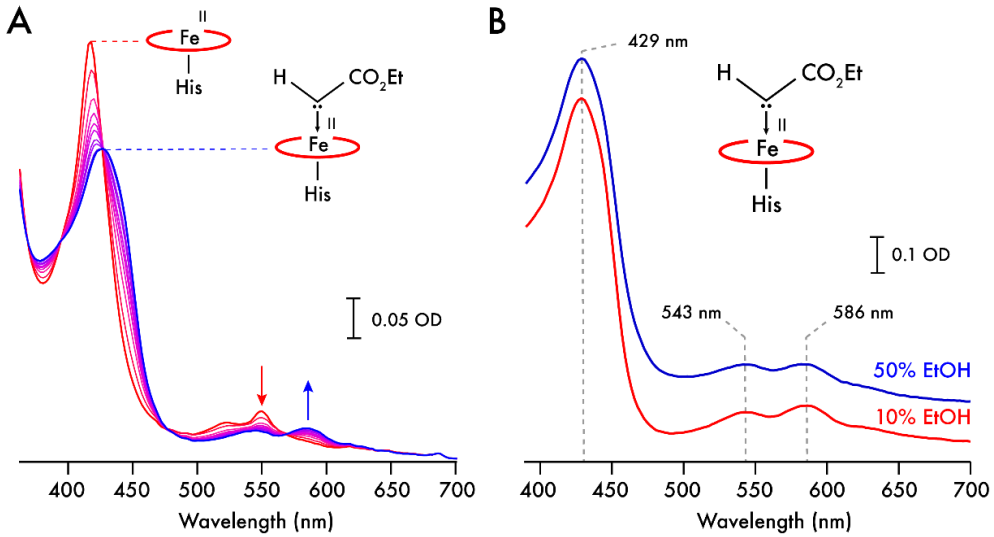


Figure 65: A) The UV-VIS spectra of C45 (7.5 µM, CHES buffer, pH 8.6, 40% (v/v) EtOH) after 0 (red) and 60 (blue) seconds following rapid mixing of ferrous C45 with 500 µM ethyl diazoacetate at 5^{O}C, and **B)** The UV-VIS spectra of C45 (7.5 µM, CHES buffer, pH 8.6i) n the presence of EDA (1 mM) after 60 seconds in 10% (red) and 50% (blue) EtOH.

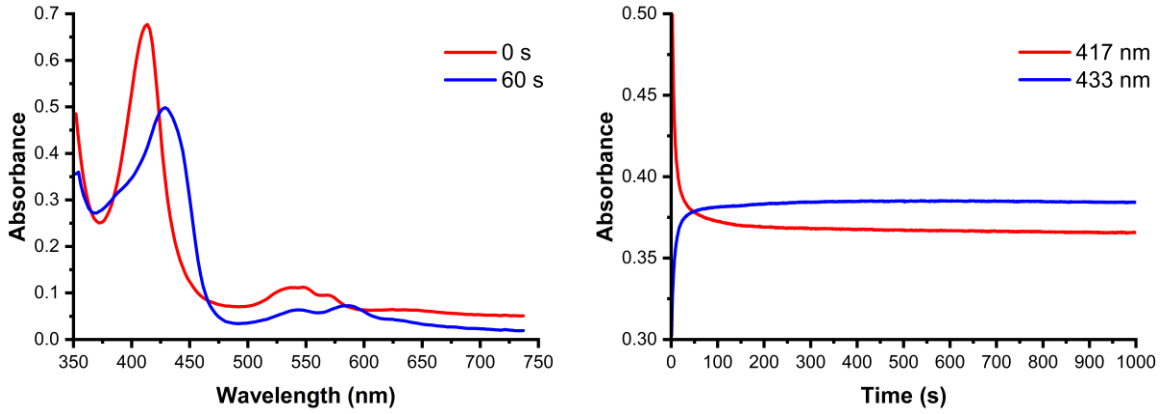


Figure 66: (**Left**) Metallocarbenoid formation and stability. The UV-VIS spectra of C45 (7.5 µM, CHES buffer, pH 8.6, 40% (v/v) EtOH) after 0 (red) and 60 (blue) seconds following rapid mixing of ferrous C45 with 500 µM ethyl diazoacetate (EDA) at 5^{o}C. (**Right**) Single wavelength traces represent the time course of ferrous C45 (417 nm, red, 7.5 µM protein, 10% EtOH) and metallocarbenoid:C45 adduct (433 nm, blue) following rapid mixing of ferrous C45 with 500 µM EDA at 5^{O}C. Once formed, the metallocarbenoid:C45 adduct persists for the duration of the experiment (1000 seconds).

Having demonstrated the formation of an iron-bound C45-metallocarbenoid intermediate using EDA as the carbene precursor, the versatility of the carbene precursor was next investigated using two different diazo compounds: *tert*-butyl diazoacetate (^{t}BuDA) and benzyl diazoacetate (BnDA). These two carbene precursors were selected because of the increased sterics of the ester substituents.

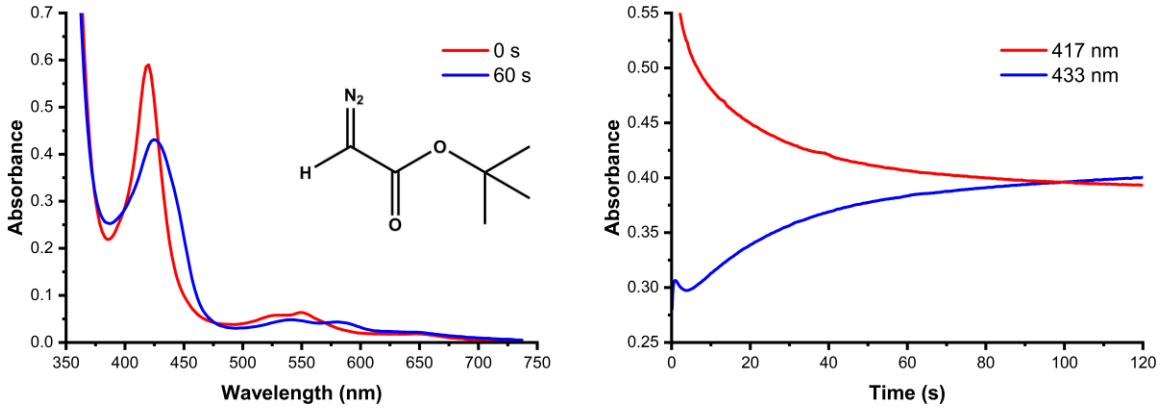


Figure 67: (**Left**) The UV-VIS spectra of C45 (7.5 µM, CHES buffer, pH 8.6, 40% (v/v) EtOH) after 0 (red) and 60 (blue) seconds following rapid mixing of ferrous C45 with 500 µM *tert*-butyl diazoacetate at 5^{o}C. (**Right**) Single wavelength traces represent the time course of ferrous C45 (417 nm, red, 7.5 µM protein, 10% EtOH) and metallocarbenoid:C45 adduct (433 nm, blue) following rapid mixing of ferrous C45 with 500 µM *tert*-butyl diazoacetate at 5^{O}C.

Figure 67 shows the UV-VIS spectrum acquired for ferrous C45 in the presence of ^{t}BuDA. Analogous to **figures 65-66** for EDA, **figure 67** exhibits a pronounced red-shift in the Soret peak from 417 nm (ferrous C45) to the metallocarbenoid intermediate (430 nm), which is accompanied by a concurrent red shift and subtle diminishment in the Q-band region. **Figure 67** displays a time-dependent trace of the reaction at 417 and 433 nm, and demonstrates that, analogous to the C45-EDA intermediate, the C45-^{t}BuDA intermediate is stable and persists after initial formation. In comparison to EDA the duration required to reach saturation, indicated by the plateauing of the 417 and 433 nm absorbance profiles respectively, is approximately 80 s, approximately seven-times greater than the duration required for EDA to reach saturation under identical conditions. Although the electron-donating properties of the ^{t}Bu substituent would stabilise the electrophilic metallocarbenoid intermediate, the increased steric bulk of the ^{t}Bu, relative to the small and linear Et substituent, would hinder the accessibility of ^{t}BuDA into the active site and would be subject to enhanced steric clashes once situated near the heme. The unfavourable steric clashes of ^{t}Bu therefore likely account for the observed seven-fold increase in the duration required to reach saturation.

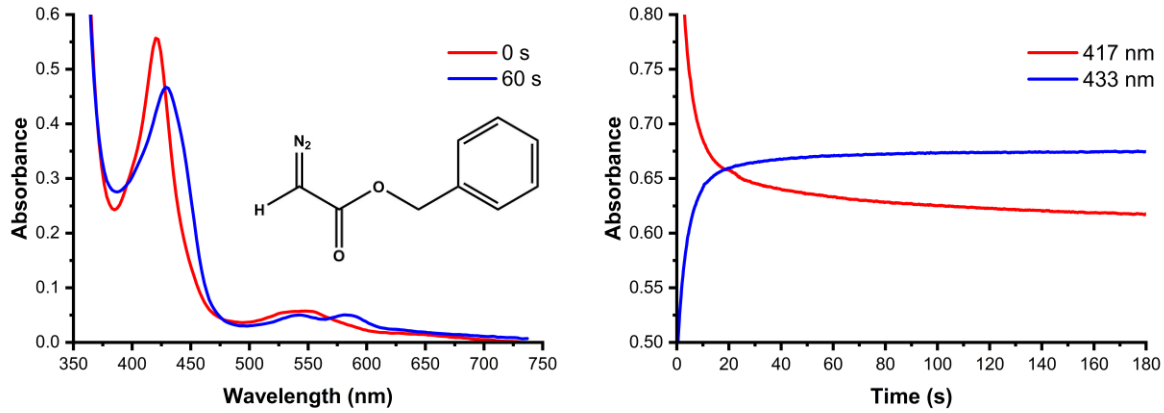


Figure 68: (**Left**) The UV-VIS spectra of C45 (7.5 µM, CHES buffer, pH 8.6, 40% (v/v) EtOH) after 0 (red) and 60 (blue) seconds following rapid mixing of ferrous C45 with 500 µM benzyl diazoacetate at 5°C. (**Right**) Single wavelength traces represent the time course of ferrous C45 (417 nm, red, 7.5 µM protein, 10% EtOH) and metallocarbenoid:C45 adduct (433 nm, blue) following rapid mixing of ferrous C45 with 500 µM benzyl diazoacetate at 5°C.

Figure 68 shows the UV-VIS spectrum acquired for ferrous C45 in the presence of BnDA. Analogous to **Figure 65-67** for EDA and ^{t}BuDA, **figure 68** exhibits a pronounced red-shift in the Soret peak from 417 nm (ferrous C45) to the metallocarbenoid intermediate (428 nm), which is accompanied by a concurrent red shift and subtle diminishment in the Q-band region. **Figure 68** displays the time-dependent trace of the reaction at 417 nm and 433 nm, respectively. Interestingly, the time-dependent trace for the formation of the C45-BnDA complex demonstrates that a similar duration, approximately 12 seconds, is required for saturation to occur, indicated by the plateauing of the 417 and 433 nm absorbance changes. The spectroscopic data for BnDA parallels the spectroscopic profile exhibited by EDA. The similarity between BnDA and EDA is difficult to completely account for by taking the electronics and sterics of the benzyl substituent into account. The increased sterics of the Bz substituent should parallel the decrease in rate formation exhibited by ^{t}Bu if sterics alone was responsible for governing the formation of a metallocarbenoid intermediate. However, the phenyl group of the benzyl substituent is an aromatic species and can potentially establish stabilising π-π interactions with other aromatic species in the active site, such as the porphyrin on the heme cofactor or backbone residues such as Phe46 or Trp43. Any stabilising π-π interactions would compensate for the increased and unfavourable steric clashes between Bz and residues in the active site, facilitating both the diffusion of BnDA into the active site and the formation of the metallocarbenoid intermediate.

3.5.1.3 Kinetic characterisation of C45 metallocarbenoid formation

In order to understand the nature of the metallocarbenoid complex the kinetic parameters for each diazo compound were determined. Repeating the stopped-flow experiments using varying concentrations of EDA, ^{t}BuDA and BnDA allowed the single turnover kinetic parameters, K_1 and k_{obs}, for the formation of the C45-metallocarbenoid intermediates to be determined (**table 1**). **Figure 69** exhibits the calculated curves acquired for each diazo compound, respectively. The analysis of the single turnover formation of a metallocarbenoid involves three distinct identities: i) free enzyme and substrate, ii) a tetrahedral transition state, and iii) the metallocarbenoid intermediate.[40] There are three underlining assumptions that have been made to assist in the analysis: i) the rate of decomposition of the tetrahedral transition state to the metallocarbenoid intermediate (k_2) is first order, ii) the reaction is irreversible (k_2 >>> k_{-2}) and iii) a steady-state equilibrium is established between the free enzyme and the tetrahedral transition state (k_1 = k_{-1}). The overall process, with associated rate constants, is illustrated in **figure 69 (bottom)**.

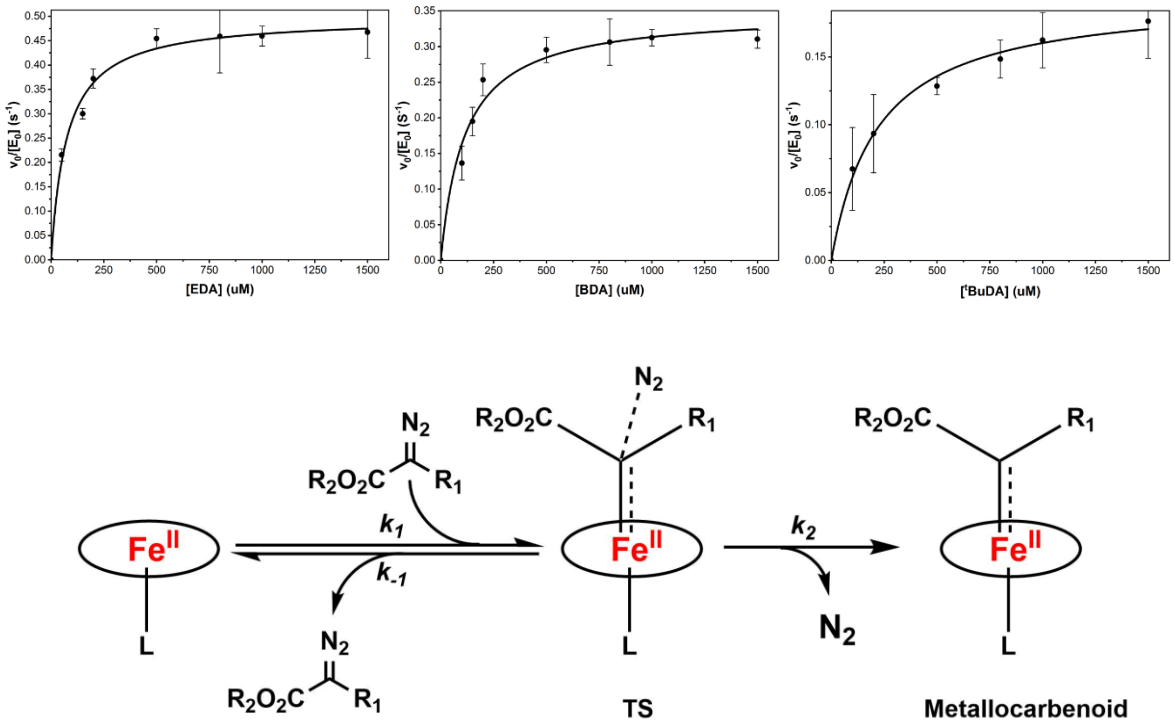


Figure 69: (**top**) The concentration-dependent single turnover kinetics of C45-metallocarbenoid formation (7.5 µM, CHES buffer, pH 8.6, 40% (v/v) EtOH) with the diazo precursors EDA (**left**), BnDA (**middle**) and ^{t}BuDA (**right**). Initial rates were determined from the rate of change in absorbance for the C45:metallocarbenoid adduct at 5°C. (**bottom**) the mechanism of metallocarbenoid formation showing the respective rate constants. The existence of the tetrahedral transition state has been postulated and computational studied by Zhang *et al*.[40]

Table 1: The single turnover kinetic parameters for C45-metallocarbenoid formation for diazo precursors EDA, BnDA, and ^{t}BuDA.

C45	EDA	BDA	^{t}BuDA
k_{obs} (s^{-1})	0.498 (± 0.013)	0.349 (± 0.015)	0.195 (± 0.007)
K_1 (μM)	74.36 (± 9.75)	113.16 (± 21.15)	219.43 (± 27.31)
k_{obs}/K_1 (s^{-1} M^{-1})	6.7 x 10^3	3.1 x 10^3	888.6
$ln(2)/k_{obs}$	1.39 s	1.98 s	3.55 s

The K_1 values represent the affinity of C45 for the binding of a carbene precursor, with lower K_1 values reflecting a higher affinity for binding. The K_1 values do not contain any information on the thermodynamic stability of a resultant metallocarbenoid intermediate, but it does provide insight into the relative energy barriers to formation from the tetrahedral transition state. The calculated K_1 values for EDA, ^{t}BuDA and BnDA are 74.36, 219.43 and 113.16 μM, respectively. The K_1 values are consistent with the computational reports on the relative activation energies for metallocarbenoid intermediate formation as a function of carbene substituents.[40] EDA contains a small, linear ethyl substituent which can be readily accommodated in the active site of C45 without disrupting the local organisation. BnDA, which possess a benzyl substituent, has a larger K_1 relative to EDA, but the phenyl group can establish favourable π-π interactions with the heme group and/or nearby aromatic residues in the active site which can provide some energetic compensation for any necessary reorganisation required in the active site to accommodate the bulkier ligand. ^{t}BuDA has a K_1 two times that of BnDA and three times that of EDA respectively, and this is accounted for by the steric bulk of the *tert*-butyl group which would need to be accommodated in the active site, but unlike BnDA cannot establish any compensatory interactions with local residues. The relationship between carbene substituent and formation energy is further corroborated by the k_{obs} values, 0.498, 0.349 and 0.195 s^{-1} for EDA, ^{t}BuDA and BnDA respectively, which represent the relative rates for the formation of the metallocarbenoid intermediate from the tetrahedral transition state. The ratio k_{obs}/K_1 for EDA, ^{t}BuDA and BnDA are reported in **table 1** and demonstrate the C45 preferentially binds to EDA with an efficiency twice that of BnDA and 7.5 times that of ^{t}BuDA. This is reflected in the time-course stopped-flow spectroscopic data reported above, which demonstrated C45 reaches saturation with EDA and BnDA relatively quickly (12 and 20 seconds respectively) but requires approximately 80 seconds to reach saturation with ^{t}BuDA.

Although the kinetic parameters for C45 against EDA, ^{t}BuDA and BnDA demonstrate a kinetic preference in the order EDA > BnDA > ^{t}BuDA, the parameters cannot provide any insight into the

thermodynamic stability of the resultant metallocarbenoid complexes. It was reported above that although bulkier ligands elevate the activation energy of formation the resultant complexes tend to be more thermodynamically stable.[40] If this observation is correct it should be reflected in the reactivity of the metallocarbenoid intermediates when presented with suitable substrates. The kinetic analysis of C45 metallocarbenoid formation with EDA, ^{t}BuDA and BnDA also suggests the steric bulk of the diazo substituent influences the reaction. This will be explored further in chapter 4.

3.5.1.4 Spectroscopic characterisation of metallocarbenoid formation in mutant cytochrome *c* and myoglobin carbene transferases

Several carbene transferases engineered through the directed evolution of naturally occurring hemoproteins, such as cytochrome P450s, cytochrome *c*, and myoglobin, have been reported in the literature. The most successful variants have demonstrated product yields, diastereoselectivies and enantioselectivities of >99.9%. In all the reported cases, the reactive intermediate is postulated to be a terminally bound, iron-localised electrophilic metallocarbenoid complex. In 2018 the Arnold lab published a crystal structure of a c-type hemoprotein cytochrome *c Rhodothermus mainus (Rma)* mutant with a bound IPC.[75] The enzyme was initially generated after three rounds of mutagenesis and was demonstrated to be highly efficient at catalysing carbene insertion reactions into Si-H σ-bonds. The three mutations discovered to enhance its reactivity where V75T, M100D, and M103E, and hence the mutant enzyme was named *Rma*-TDE. *Rma*-TDE was subsequently selected as a suitable candidate for capturing and characterising an IPC. In December of the same year, the Hilvert lab published a crystal structure of an IPC in a sperm-whale myoglobin mutant called Mb(H64V,V68A).[304] Mb(H64V,V68A), two mutations from WT-sperm whale myoglobin, was previously demonstrated to be an efficient and highly stereoselective carbene transferase in the cyclopropanation of vinylarenes.

The IPC crystal structures for *Rma*-TDE and Mb(H64V,V68A) provided key insights into the nature of IPCs, and into the phenomena governing metallocarbenoid formation and stability. Despite these insights a thorough spectroscopic investigation of the enzymatic IPCs has yet to be conducted. As the methodology for investigating metallocarbenoid formation had been established using C45 and microperoxidase-11, *Rma*-TDE and Mb(H64V,V68A) were selected for spectroscopic investigation and comparison to C45. Identical experimental conditions used to investigate metallocarbenoid formation with C45 were employed for each characterisation.

3.5.1.4.1 *Rma*-TDE

The UV-VIS spectra for the ferric and ferrous states of *Rma*-TDE are presented in **figure 70**. As is characteristic of hemoproteins, the UV-VIS spectrum for ferric *Rma*-TDE is dominated by an intense

Soret peak at 410 nm followed by a less-intense Q-band region split into two peaks, α and β, with α > β in relative intensity. The UV-VIS spectrum for ferrous *Rma*-TDE, acquired by the addition of sodium dithionite, is characterised by a pronounce red-shift in the Soret peak, from 410 nm to 419 nm, alongside a concurrent modification of the Q-band region; in particular, both the α and β peaks comprising the Q-band are slightly blue-shifted and are accompanied with a reversal in the relative intensities of each respective peak, such that α < β.

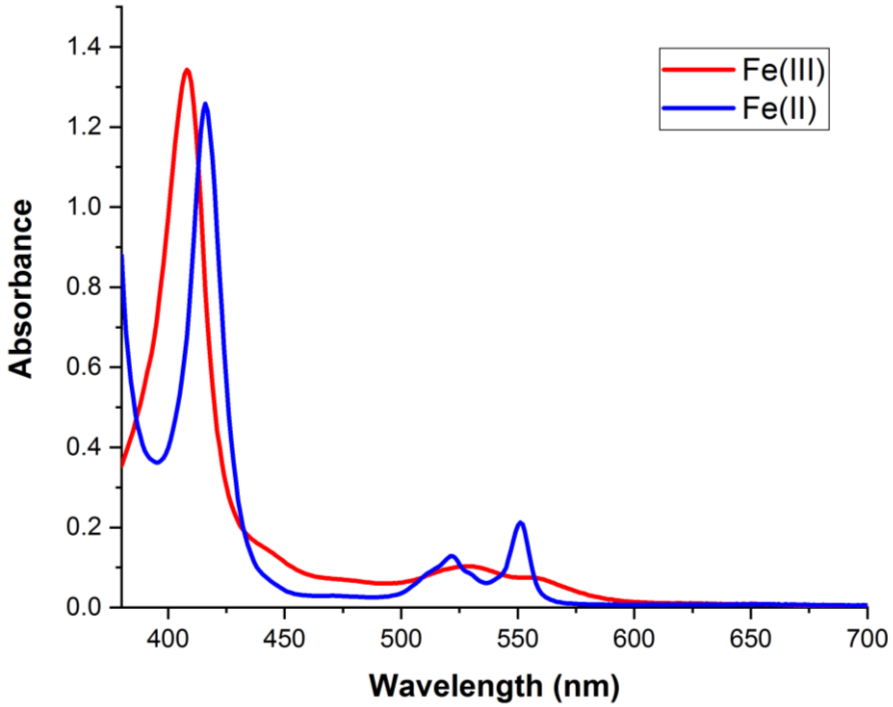


Figure 70: The UV-VIS spectrum for ferric (red) and ferrous (blue) *Rma*-TDE.

The acquired UV-VIS spectra for *Rma*-TDE is consistent with the UV-VIS spectra reported in the literature for cytochrome *c*.[143,219] *Rma*-TDE was subsequently investigated for metallocarbenoid activity using stopped-flow spectroscopy. The UV-VIS spectra acquired for ferrous *Rma*-TDE in the presence of EDA is exhibited in **figure 71**.

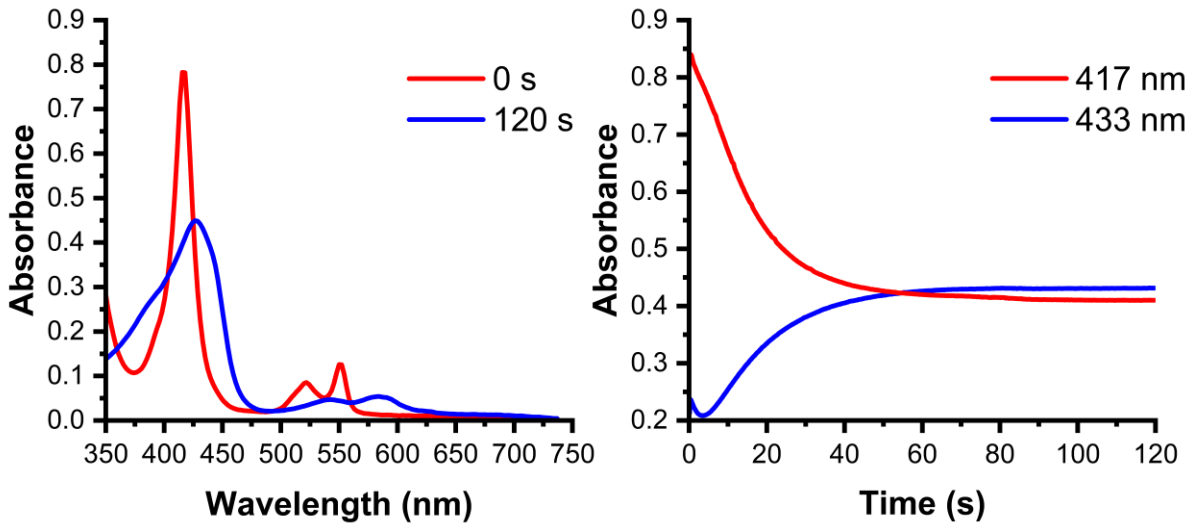


Figure 71: (**Left**) The UV-VIS spectra of *Rma*-TDE (7.5 µM, CHES buffer, pH 8.6, 40% (v/v) EtOH) after 0 (red) and 60 (blue) seconds following rapid mixing of ferrous C45 with 500 µM ethyl diazoacetate at 5°C. (**Right**) Single wavelength traces represent the time course of ferrous *Rma*-TDE (417 nm, red, 7.5 µM protein, 10% EtOH) and metallocarbenoid:*Rma*-TDE adduct (433 nm, blue) following rapid mixing of ferrous *Rma*-TDE with 500 µM ethyl diazoacetate at 5°C.

The UV-VIS spectrum is characterised by a significantly red-shifted Soret peak, from 418 to 430 nm, which is accompanied by a red-shifted and broadened Q-band. The time-dependent plots presented in **figure 71** show the time-varying absorbance changes at 417 and 433 nm, corresponding to the ferrous species and metallocarbenoid complex, respectively. The decrease in absorbance at 417 nm is accompanied by a paralleled increase in absorbance at 433 nm, with a 40 second duration required before saturation is established. The spectroscopic data acquired for *Rma*-TDE metallocarbenoid formation in the presence of EDA is analogous to the spectroscopic data acquired for C45 metallocarbenoid formation. The two data sets are characterised by a significant red-shift in the Soret peak, a red-shift and broadening of the Q-band profile, and the eventual attainment of saturation. In comparison to C45, *Rma*-TDE takes approximately 3.5 times longer to reach saturation. The ferrous *Rma*-TDE UV-VIS spectrum is identical to the spectrum reported by Lewis *et al*;[219] however, the spectrum acquired for the metallocarbenoid intermediate contrasts the spectrum reported by Lewis *et al*, which is characterised by a Soret peak at 416 nm and two Q-band peaks at 529 and 561 nm.[219] The spectrum reported here has a similar profile but with the Soret peak and Q-band peaks red-shifted. There are, however, several notable differences between the spectrum reported here and by Lewis *et al*:[219] the spectrum reported by Lewis *et al* was acquired using ethyl 2-diazopropanoate (Me-EDA), at room temperature under aerobic conditions,[219] while the spectrum reported here was acquired using EDA, at 5°C, under anaerobic conditions and 10% (v/v) EtOH. As the spectrum presented here was acquired under more rigorous conditions (stopped-flow spectrometry under

anaerobic conditions ([O_2] < 5 ppm) it is postulated that the species reasonable for the spectrum reported by Lewis *et al* is an alternative, unidentified species.

3.5.1.4.2 Mb(H64V,V68A)

The UV-VIS spectra for ferric and ferrous Mb(H64V,V68A) are presented in **figure 72**. The spectrum for ferric Mb(H64V,V68A) is dominated by an intense Soret peak at 410 nm, with a significantly less-intense and broad Q-band region between 500-600 nm. Addition of sodium dithionite yields the spectrum for ferrous Mb(H64V,V68A) which is characterised by a pronounced red-shift in the Soret peak, from 410 nm to 435 nm. The Q-band is characterised by a narrowing of the observed range, to 550-600 nm, but remains low in intensity and poorly resolved. Mb(H64V,V68A) was subsequently investigated for metallocarbenoid activity using stopped-flow spectroscopy.

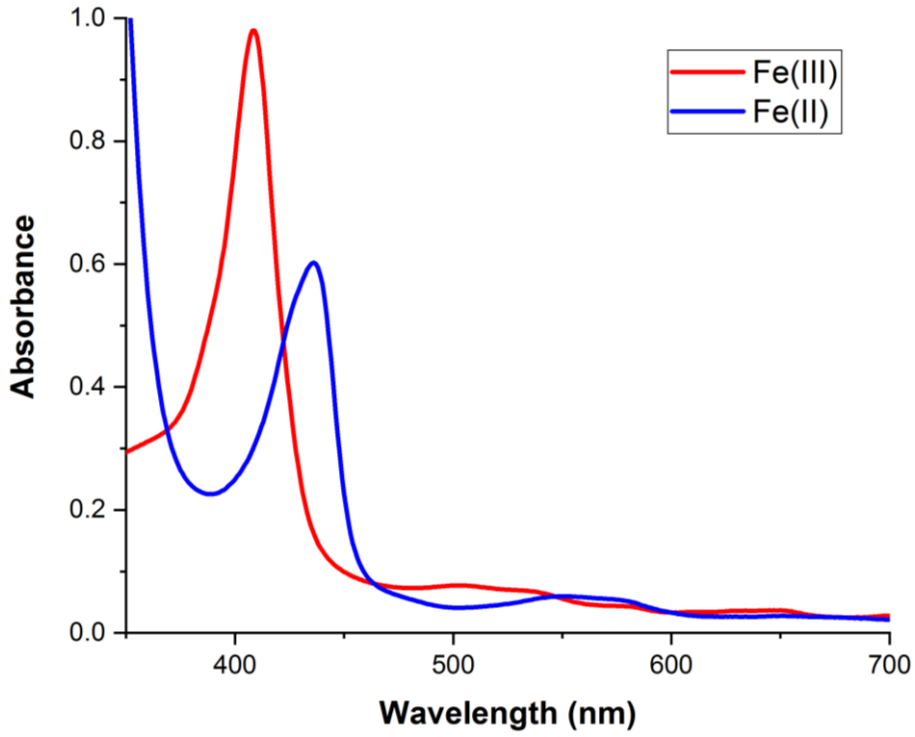


Figure 72: The UV-VIS spectrum for ferric (red) and ferrous (blue) Mb(H64V,V68A).

Figure 73 is the UV-VIS spectra acquired for ferrous Mb(H64V,V68A) in the presence of EDA. The spectra is characterised by a significantly blue-shifted Soret peak, from 435 to 424 nm, which is accompanied by the emergence of a blue-shifted but well-resolved Q-band profile, with α and β peaks at 525 and 555 nm respectively, with α < β in relative intensity. The time-dependent plots presented in **figure 73** show the changes in absorbance at 435 and 424 nm over time, corresponding to the ferrous species and metallocarbenoid complex, respectively. The decrease in absorbance at 435 nm is accompanied by a paralleled increase in absorbance at 424 nm, with a 120 second duration required

before saturation is established. The time-dependent spectra for Mb(H64V,V68A) in the presence of EDA is distinctive from the spectra acquired for C45 and *Rma*-TDE as the initial changes in absorbance are fast, with a drastic change occurring in the initial 20 seconds, followed by a considerable lag phase before saturation is established (80 seconds). The time-dependent spectra acquired for Mb(H64V,V68A) in the presence of EDA resembles the time-dependent spectra acquired for C45 in the presence of ^{t}BuDA, but is nonetheless consistent with the formation of a metallocarbenoid complex.

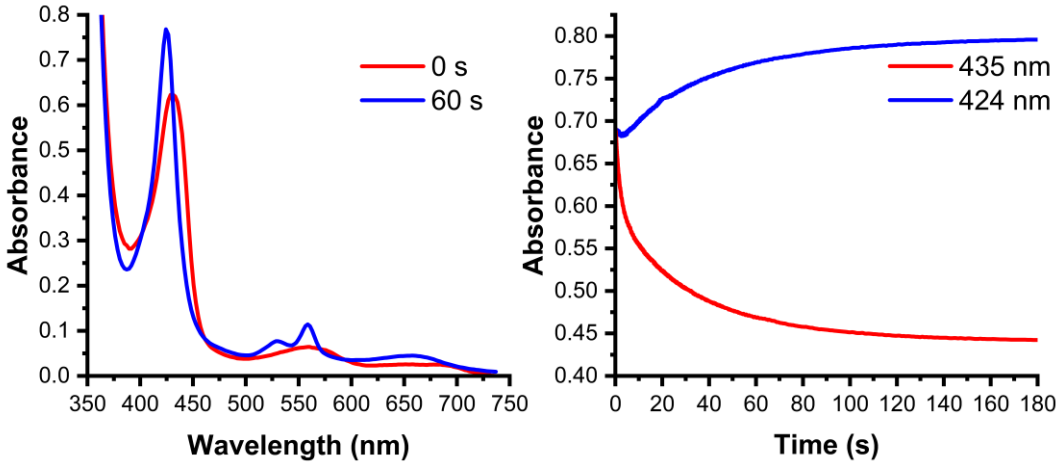


Figure 73: The UV-VIS spectra of Mb(H64V,V68A) (7.5 µM, 0.1 M KPi, pH 7, 40% (v/v) EtOH) after 0 (red) and 60 (blue) seconds following rapid mixing of ferrous Mb(H64V,V68A) with 500 µM ethyl diazoacetate at 5^{o}C. (**Right**) Single wavelength traces represent the time course of ferrous Mb(H64V,V68A) (435 nm, red, 7.5 µM protein, 10% EtOH) and metallocarbenoid:Mb(H64V,V68A) adduct (424 nm, blue) following rapid mixing of ferrous Mb(H64V,V68A) with 500 µM ethyl diazoacetate at 5^{O}C.

3.5.1.5 Kinetic characterisation of metallocarbenoid formation in *Rma*-TDE and Mb(H64V,V68A) carbene transferases

The initial experiments using *Rma*-TDE and Mb(H64V,V68A) exhibited UV-VIS spectra consistent with the formation of an iron-localised, terminally bound metallocarbenoid intermediate and are consistent with the crystal structures of these complexes. The next step was to employ the stopped-flow methodology used to determine the single turnover kinetic parameters, K_1 and $k_{obs,}$ for metallocarbenoid formation using *Rma*-TDE and Mb(H64V,V68A) respectively. The single turnover kinetic parameters for *Rma*-TDE and Mb(H64V,V68A) could then compared to the single turnover kinetic parameters obtained for C45 (**table 2**). As in the C45 experiments, varying concentrations of EDA were used and the formation of the metallocarbenoid intermediate was followed spectroscopically as a function of time.

Table 2: The kinetic parameters for the metallocarbenoid formation between EDA and the *de novo* heme-containing maquette C45, and the rationally designed heme-containing proteins Mb(H64V,V68A), and *Rma*-TDE.

EDA	C45	Mb(H64V,V68A)	*Rma*-TDE
$k\ (s^{-1})$	0.498 (± 0.013)	0.224 (± 0.013)	0.144 (± 0.010)
$K_1\ (\mu M)$	74.36 (± 9.75)	15.04 (± 4.68)	487.94 (± 49.22)
$k_{obs}/K_1\ (s^{-1}\ M^{-1})$	6.7×10^3	1.50×10^4	294.26
$\ln(2)/k_{obs}$	1.39 s	3.1 s	4.83 s

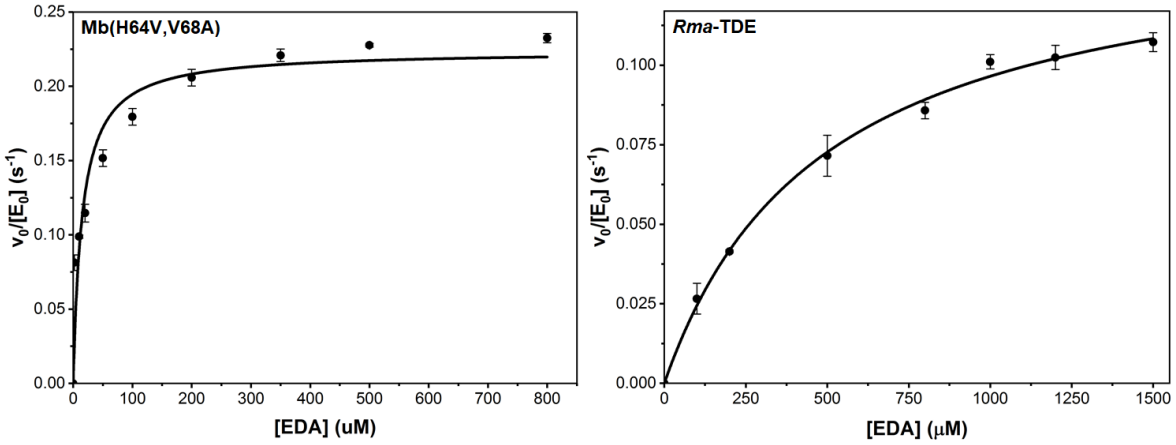


Figure 74: The concentration-dependent kinetics of metallocarbenoid formation of (**left**) Mb(H64V,V68A) (7.5 µM, 0.1M KPi buffer, pH 7.1, 40% (v/v) EtOH) and of (**right**) *Rma*-TDE (7.5 µM, CHES buffer, pH 8.6, 40% (v/v) EtOH) with varying concentration of EDA. Initial rates were determined from the rate of change in absorbance for the enzyme:metallocarbenoid adduct at 5°C

The calculated curves for *Rma*-TDE and Mb(H64V,V68A) are presented in **figure 74** for EDA. As with C45 the analysis of the formation of a metallocarbenoid involves three distinct identities: i) free enzyme and substrate, ii) a tetrahedral transition state, and iii) the metallocarbenoid intermediate. There are three underlining assumptions that have been made to assist in the analysis: i) the rate of decomposition of the tetrahedral transition state to the metallocarbenoid intermediate (k_2) is first-order, ii) the reaction is irreversible ($k_2 >>> k_{-2}$) and iii) a steady-state equilibrium is established between the free enzyme and the tetrahedral transition state. The K_1 values represent the affinity of an enzyme for the binding of a carbene precursor, with lower K_1 values reflecting a higher affinity for binding. The calculated K_1 values for *Rma*-TDE and Mb(H64V,V68A) are 487.94 and 15.05 µM, respectively. The calculated K_1 for EDA/C45 is 74.36 µM, rendering K_1 for Mb(H64V,V68A) 5 times smaller, but for *Rma*-TDE K_1 is approximately 6 times larger, relative to C45. In order to comprehend the difference in magnitudes for the three enzymes it is necessary to begin by analysing the structural

differences between the three enzymes, with particular attention dedicated to the active-site and first-coordination sphere.

The crystal structure of *Rma*-TDE (1.3 Å resolution, **figure 75**) reveals the introduction of the three mutations results in significant changes to the loop distal to the heme.[218,219] Three amino acids in the loop region of IPC-bound *Rma*-TDE, Asp-100, Thr-101, and Asp-102, were unresolved in the crystal structure (**figure 75**). The authors suggested the mutations introduced significant flexibility into the loop region, opening up the active site and increasing solvent accessibility.[218,219] It was also suggested the topological changes in the loop region preorganised *Rma*-TDE towards substrate binding and metallocarbenoid formation. In addition to increased loop flexibility and solvent accessibility, the crystal structure revealed five short-range non-polar interactions between amino acid residues and the iron-carbene. The five residues where identified as Thr-75, Met-76, Pro-79, Ile-83, and Met-89, with average interaction distances in the range 3.5-4.3 Å, suggesting a strongly hydrophobic active site despite the increased solvent accessibility (**figure 75**).[218,219]

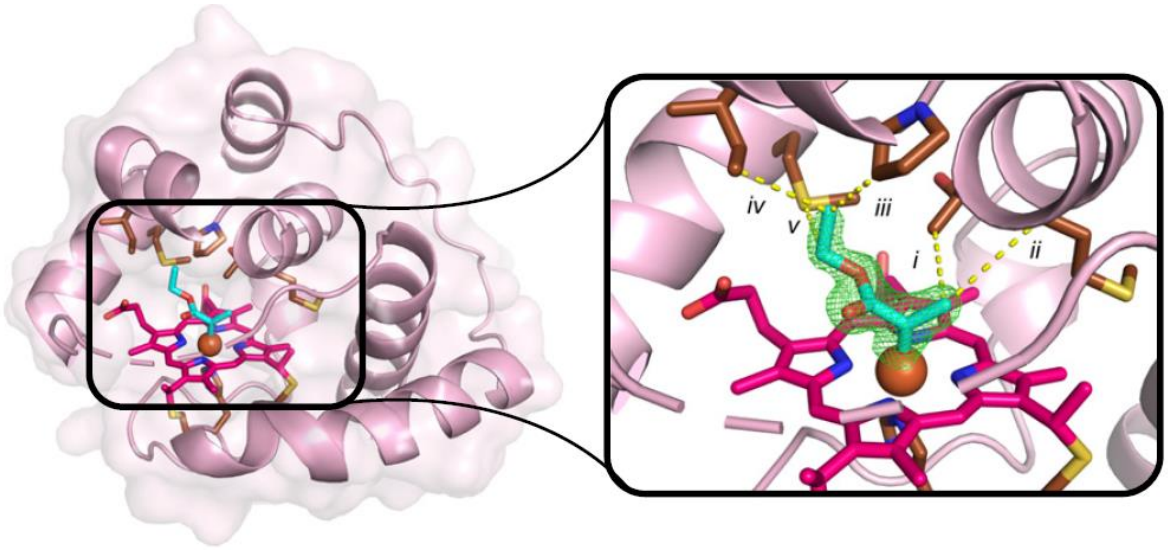


Figure 75: The crystal structure *Rma*-TDE bound to the carbene generated from ethyl 2-diazopropanoate (1.29 Å; PDB ID code 6CUN). The residues Asp-100, Thr-101, and Asp-102 are unresolved. Interactions between the carbene and amino acid residues are shown with yellow lines. (i) Thr-75, 3.7 Å. (ii) Met-76, 4.3 Å. (iii) Pro-79, 3.5 Å. (iv) Ile-83, 3.6 Å. (v) Met-89, 3.5 Å (figure taken from Lewis *et al*).[219]

The computational analysis of the electronic structures of IPCs in model complexes reported in the literature indicate the lowest energy state for an IPC is an open shell singlet (OSS) with an antiferromagnetic coupling between the iron and the carbene.[39] In contrast, the authors reported a water molecule in the active-site of *Rma*-TDE stabilises the carbene complex through hydrogen-bonding interactions. The authors reported the stabilisation provided by the water molecule is

significantly more substantial in a closed shell singlet (CSS) IPC, and suggested this stabilising interaction contributed to the formation of a single conformation.[218,219] The authors reported the observed electron density of the IPC crystal structure supports the proposition of a singlet carbene complex adopting a single conformation, corroborating the suggestion the dominant species in solution is a CSS carbene complex.[218,219]

The crystal structure of Mb(H64V,V68A) (1.1 Å resolution) revealed the introduction of the two mutations did not disrupt the three-dimensional fold of the protein relative to wild-type sperm whale myoglobin.[304] In addition, the residues participating in hydrogen-bonding with the heme cofactor, and the distances of such interactions, were also unaffected by the two mutations relative to the wild-type. The most drastic structural change occurring as a consequence of the introduction of the two mutations was an enlargement of the heme distal pocket volume from approximately 125 $Å^3$ in the wild-type to 243 $Å^3$ in Mb(H64V,V68A) respectively.[304] The authors noted the expansion was most pronounced on the solvent-exposed face of the heme cofactor, which should enhance substrate accessibility into the active site.[304]

The H64V mutation was reported to enhance the cyclopropanation activity relative to WT-myoglobin, and it is postulated this is accomplished *via* increased hydrophobicity of the active site.[304] In addition, analysis of the crystal structure allowed four key residues occupying the area above the heme cofactor to be identified – Val-64, Ala-68, Phe-43, and Ile-107.[304] Fasan and co-workers subsequently utilised the structural information obtained from crystallography to perform an extensive conformational analysis of the IPC. DFT calculations were performed to identify the energetically favourable conformations and to elucidate the stabilise factors.[305] Initially, the number of conformations was established to be 8, with four conformations corresponding to the orientation of the carbene plane along a Fe-N axis and four corresponding to an orientation in-between two Fe-N axes. The σ-bond connecting the ester moiety to the carbene carbon can freely rotate, and two particular conformations, labelled (+) and (-), were identified for this C-C σ-bond, resulting in a total of 16 possible conformations to be explored.[305] The conformations could be conveniently divided into four spatial groups of four conformations, with each spatial region defined as the inter-residue space between the four primary amino acids above the plane of the porphyrin; therefore, the four spatial regions were i) Ala-68-Ile-107, ii) Ile-107-Phe-43, iii) Phe-43-Val-64, and Val-64-Ala-68 (**figure 76**).[305]

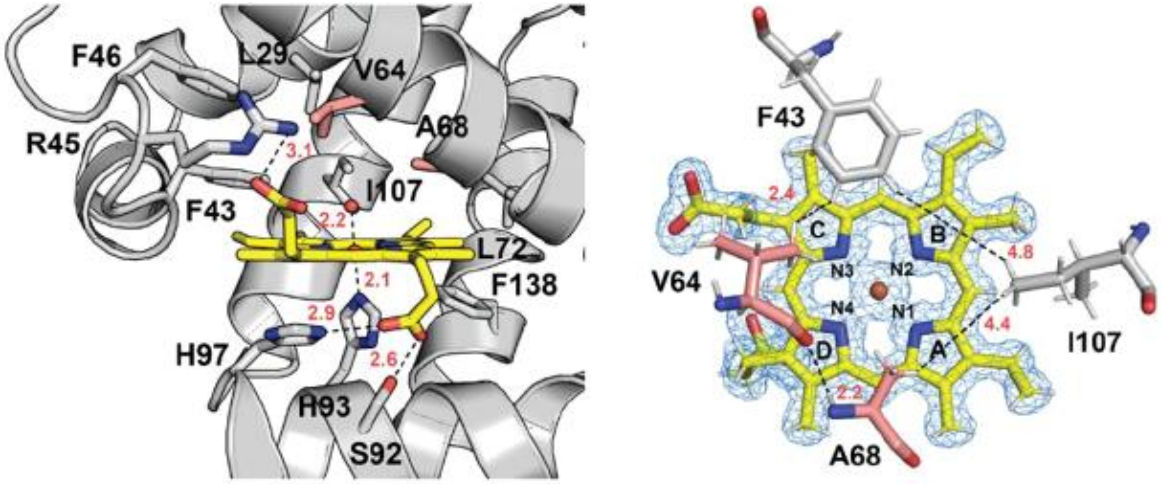


Figure 76: (**left**) The arrangement of amino acids surrounding the heme cofactor in Mb(H64V,V68A), and (**right**) view of the bound heme and nearby amino acid residues, along with the labels for pyrrole N atoms, porphyrin rings and the four spatial regions utilised in the DFT calculations (picture adapted from Tinoco *et al*).[305]

The DFT calculations indicated that conformer 2(+), a conformer in the Ala-64-Ile-107 spatial group, was the most energetically favourable by approximately 3-10 kcal mol^{-1} relative to the remaining fifteen conformers.[305] The results of the calculations indicated that steric effects play a dominant role in the conformational stability of the heme-bound carbene, as conformer 2(+) is situated in a region with the second biggest inter-residual volume. Further examination of the remaining fifteen conformers reveals that two additional conformers, 3(+) and 6(+) were only slightly higher in energy relative to 2(+) (3.2 and 3.4 kcal mol^{-1} respectively) despite 3(+) occupying the least hindered region, Ile-107-Phe-43. It was evident from the crystal structure, and corroborated by the DFT simulations, that the stability of a heme-bound carbene was mediated by van der Walls interactions between protein residues and the carbene-bound ester moiety. In 2(+), a total of four short interactions (~2.5 Å) between the oxygen atoms in the ester moiety (O_1 and O_2) and neighbouring C-H bonds were present, in comparison to only 2 short-distance bonds in the 3(+) and 6(+) conformations.[305]

The analysis of the IPCs in *Rma*-TDE and Mb(H64V,V68A) revealed two factors governing the Gibbs energy of formation and resultant conformational stability of a metallocarbenoid species: i) sterics, and ii) protein-carbene non-covalent interactions. A discussion on the importance of protein-carbene non-covalent interactions has already been presented above, in particular with regard to *Rma*-TDE and Mb(H64V,V68A). For C45, an account of important carbene-protein interactions is difficult to provide because of a lack of structural information; although it is likely that a tryptophan residue in close proximity to the active site may provide stabilising π-π interaction with the benzyl group in BnDA. In terms of sterics, the reported single turnover kinetic parameters **(table 1)** hint towards a possible

relationship between the rate of metallocarbenoid formation and the steric bulk of the diazo precursor. It is abundantly clear that more structural information is required before any atomistic account of metallocarbenoid formation can be presented. Analysing the formation of the metallocarbenoid:C45 adduct with a variety of other functionalised diazo precursors, Raman spectroscopy, NMR spectroscopy, and computational simulations (DFT), could collectively provide key insights into the protein-carbene interactions governing the binding of the diazo precursor, and stabilisation of the tetrahedral intermediate and the resultant metallocarbenoid complex.

3.5.2 Electron spray ionisation mass spectrometry of C45 and C45-metallocarbenoid complexes

The difficulty associated with crystallising tetra-α-helix maquettes rendered the crystallisation of a heme-localised metallocarbenoid intermediate, like the crystal structure reported by Arnold using cytochrome *c*, unfeasible. Electron-spray-ionisation mass spectrometry (ESI-MS), a soft ionisation technique, has been previously employed to characterise the binding of myoglobin and haemoglobin to small molecules such as O_2, CO and CN^-. It was therefore rationalised that ESI-MS could be a good technique for isolating a heme-bound maquette metallocarbenoid intermediate.

An initial ESI-MS experiment was conducted using only C45 (350 µM, 100 mM KCl, 20 mM CHES, pH 8.6) in the absence of any carbene precursor. Injections were inserted immediately into the ionisation chamber without an initial chromatographic purification step. The injections into the ionisation chamber were monitored at 280 and 407 nm (the Soret peak for ferric C45) respectively. The acquired mass spectrum exhibited a dominant peak at 1974.36 Da, with smaller significant peaks at 1755.18 Da and 2256 Da respectively (**figure 77**). The peak at 1974.36 Da was calculated to correspond to the M^{8+} species, with the peaks at 1755.18 Da and 2256 Da corresponding to the M^{9+} and M^{7+} species, respectively. The clear resolution of these peaks indicated that any change in mass of the species, corresponding to the formation of a metallocarbenoid complex, would be detectable in the ESI-MS spectrum.

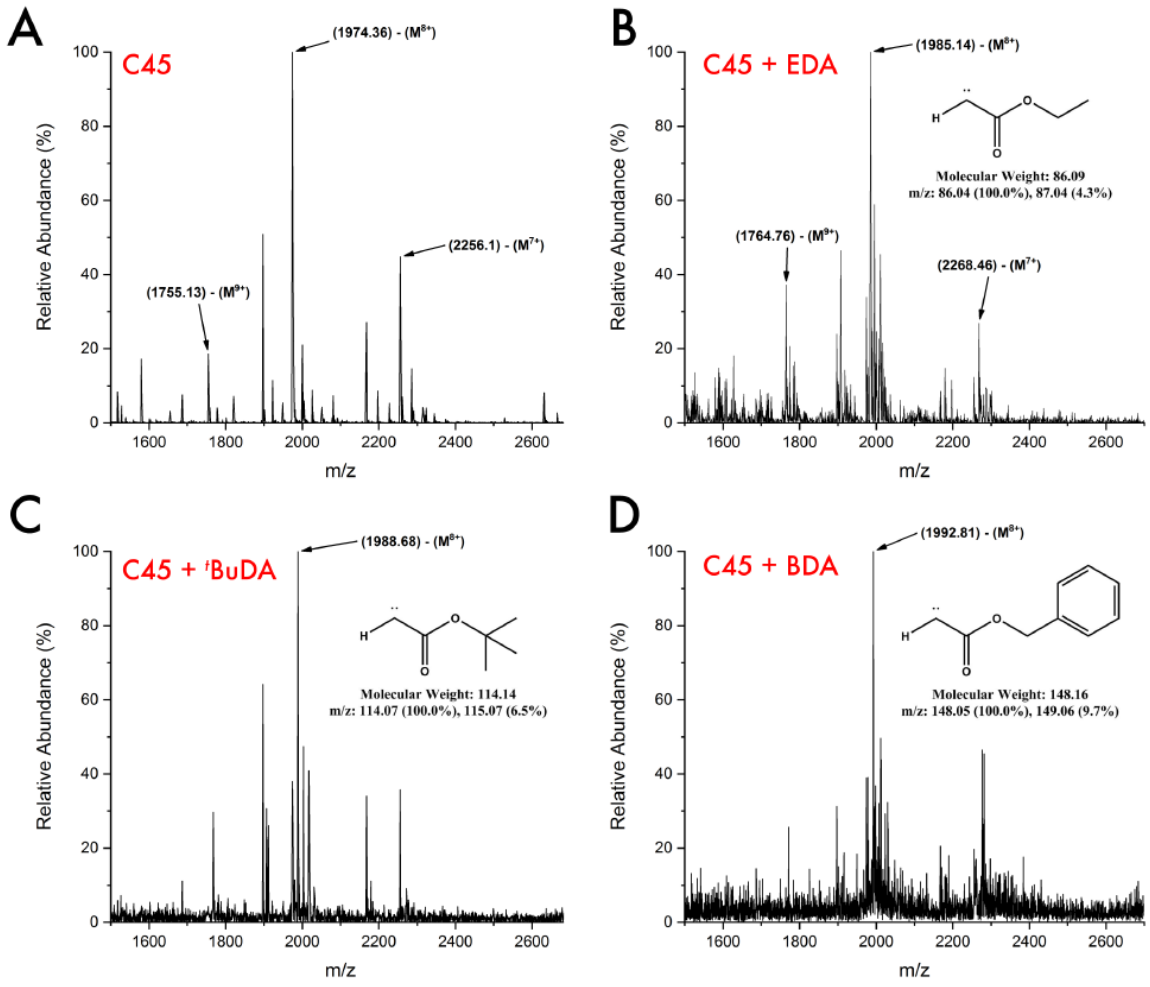


Figure 77: ESI mass spectrometry of C45 metallocarbenoid complexes. **A.** Mass spectrum of C45 (350 µM) in buffer. The three dominant peaks occur at m/z values of 1755.13, 1974.36, and 2256.1 Da, respectively. These fragments correspond to a charged species of (from left to right) 9^+, 8^+, and 7^+, as determined by the atomic mass of C45 (15789.17 Da). **B.** Mass spectrum for C45 after the addition of 50 µl of ethyl diazoacetate (EDA, 20 mM, 5% ethanol). The three dominant peaks in the spectrum have shifted, relative to the peaks exhibited in the C45 spectrum, to 1764.76, 1985.14, and 2268.46 Da for the $M^{9+,}$ M^{8+} and M^{7+} species, respectively. The spectrum indicates an average mass increase of 86.52 Da, which corresponds approximately with the predicted molecular weight of an ethyl diazoacetate carbene ($-N_2$) species (86.09 Da). **C.** Mass spectrum for C45 after the addition of 50 µl of *tert*-butyl diazoacetate (^{t}BuDA, 20 mM, 5 % ethanol). The dominant peak corresponding to the M^{8+} species has shifted to an m/z value 1988.68, indicating an increase in mass of 114.56 Da, which corresponds with the predicted molecular weight of a *tert*-butyl diazoacetate carbene ($-N_2$) species (114.14 Da). **D.** Mass spectrum for C45 after the addition of 50 µl of benzyl diazoacetate (BnDA, 20 mM, 5% ethanol). The dominant peak corresponding to the M^{8+} species has shifted to an m/z value 1992.81, indicating an increase in mass of 148.08 Da, which corresponds with the predicted molecular with of a benzyl diazoacetate carbene ($-N_2$) species (148.05 Da). All spectra were recorded in ES+ mode and monitored at 280 and 430 nm. The samples were injected directly into the ionisation chamber, with no LC separation step (0.1M ammonium acetate buffer, pH 7). Injection volumes were 20 µl.

A solution containing ferrous C45 (350 µM, 100 mM KCl, 20mM CHES, pH 8.6, 10 mM $Na_2S_2O_4$) was prepared, under nitrogen, and cooled to 5 ^{O}C. Ethyl diazoacetate (EDA, 400 mM stock in EtOH) was prepared under nitrogen, cooled, and 20 µl was transferred into the sealed-vial containing ferrous C45 using an air-tight syringe. The vial was left to mix for one minute, on ice, after which effervescence was visible (which was attributed to the expulsion of N_2 following the formation of the metallocarbenoid intermediate). After one minute, the vial was transported to the mass spectrometer loading bay and immediately injected into the ionisation chamber. The injected sample passed through a UV-VIS detector prior to entering the ionisation chamber, which was set at 428 nm (the Soret peak for the maquette-metallocarbenoid complex). The spectrum acquired exhibited the following shifted peaks, 1764.82 Da, 1985.1 Da, and 2268 Da, corresponding to the M^{7+}, M^{8+} and M^{9+} species respectively (the M^{8+} peak, at 1985.1 Da, was the dominant peak) (**figure 77**). The shift in each respective peak was calculated *via* subtraction and subsequent multiplication by the respective charge and resulted in an average m/z shift of 86 Da, which corresponds very well with the expected mass of an EDA carbene adduct after the loss of N_2. As the injection of the species into the ionisation chamber was monitored at 428 nm, the Soret peak for the maquette-metallocarbenoid intermediate, this rules out the possibility that EDA is reacting with an amino acid in the protein backbone or inserting into a Fe-N bound of the heme cofactor, as these chemical transformations would not result in detectable species at 428 nm, as had been previously established spectroscopically. The mass spectrum obtained therefore corroborates the spectroscopic results, indicating that a heme-localised metallocarbenoid complex is readily formed and persistent.

The spectroscopic profiles for metallocarbenoid formation employing ^{t}BuDA and BnDA as the carbene precursors paralleled the spectra obtained for EDA and literature reports for metallocarbenoid complexes. As the maquette-metallocarbenoid intermediate had been successfully detected for EDA using ESI-MS, the experiments were repeated, under identical conditions, for ^{t}BuDA and BnDA. The mass spectrum obtained for C45 in the presence of ^{t}BuDA resulted in a dominant peak at 1988.7 Da, corresponding to the M^{8+} species (**figure 77**). The mass shift calculated *via* subtraction and subsequent multiplication by the respective charge resulted and produced an m/z shift of 114.1 Da, corresponding to the mass difference expected for the formation of a ^{t}BuDA:M^{8+} species. Similarly, the mass spectrum for C45 in the presence of BnDA resulted in a dominant peak at 1992.8 Da, corresponding to the M^{8+} species and a m/z difference of 148.1 Da, the mass difference expected for the formation of a BnDA:M^{8+} species (**figure 78**).

3.6 CONCLUSION

The carbene transfer activity exhibited by numerous carbene transferases in the literature are postulated to proceed *via* a heme-localised electrophilic metallocarbenoid intermediate. Computational studies have revealed the most energetically favourable state of a heme-localised metallocarbenoid intermediate would be a closed-shell-singlet electronic configuration with a ferrous iron. C45 is a pentacoordinate, histidine ligated heme-containing four-helix bundle which exhibits oxygen binding, electron transfer and peroxidase functionality. In the literature, it was reported an axial histidine ligation enhances carbene transferase activity for select hemeproteins. Given the similarities between C45 and the reported engineered carbene transferases, the question of whether a metallocarbenoid intermediate could be detected using C45 was raised. A combination of stopped-flow spectroscopy, single turnover experiments, and ESI-MS conducted with three diazo precursors (EDA, ^{t}BuDA and BnDA) revealed C45 forms a stable, spectroscopically distinct intermediate with each precursor, with ESI-MS spectrum revealing the new intermediates each have an exact mass corresponding to the loss of molecular nitrogen. A relationship between the steric bulk of the diazo ester and the kinetic parameters has also been suggested. A comparison of the spectroscopic profiles and kinetics of the metallocarbenoid complex formed by EDA with two engineered carbene transferases, *Rma*-TDE and Mb(H64V,V68A), was reported. The metallocarbenoid intermediate of *Rma*-TDE and Mb(H64V,V68A) has been isolated and successfully crystallised, and the strong similarities between the spectroscopic profile of the EDA adducts of *Rma*-TDE and Mb(H64V,V68A) and the EDA, ^{t}BuDA and BnDA adducts of C45 strongly corroborates the existence of a terminal-bound heme-localised metallocarbenoid intermediate in C45.

CHAPTER 4: CYCLOPROPANATION CATALYSED BY A *DE NOVO* CARBENE TRANSFERASE

4.1 INTRODUCTION TO THE CYCLOPROPANYL GROUP

The cyclopropyl group is a three-membered carbon ring consisting of three sp^3 hybridised carbons arranged in a planar ring *via* three C-C σ-bonds.[306] The carbon-carbon bond angles are 60^O, a significant deviation away from the ideal 109.5^O adopted in linear sp^3 carbons (**figure 78, left**). The three sp^3 carbon atoms align in the same plane, forming a planar ring structure, which forces the hydrogen atoms into an eclipsed conformation (**figure 78**). As a consequence of the C-C and C-H bond angles a cyclopropyl fragment is a highly strained fragment, which is manifested in its interesting properties and unique reactivity.[42,306] This is reflected in the reported C-C and C-H bond dissociation energies for cyclopropane compared to propane, which are 238 and 456 KJ mol^{-1} for cyclopropane and 335 and 413 KJ mol^{-1} for propane respectively.[307–309] The high ring strain in cyclopropane is also reflected in the C-C bond dissociation energies for the homologous cycloalkane series, which follows the trend $C_3 < C_4 < C_5$, C_6, C_7, C_8.[307] Interestingly, the C-H bond in a cyclopropane is stronger than the C-H bond in propane, and this observation is accounted for by the hybridisation of the relative orbitals involved in cyclopropane bonding.[42,306]

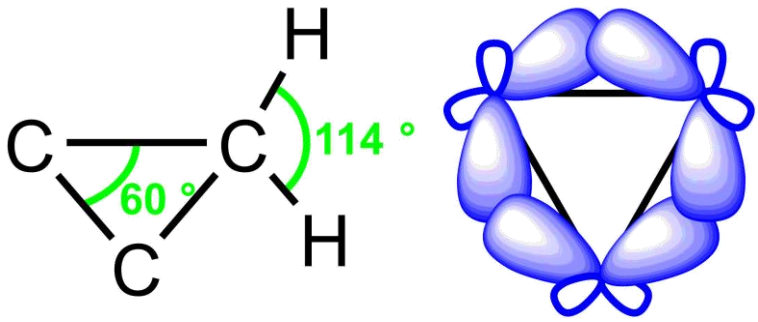


Figure 78: The cyclopropyl fragment. **(left)** the C-C-C and H-C-H bond angles, and **(right)** the orientation of the carbon sp^3 orbitals.

The C-C bonds in cyclopropanes are best described as ''bent'' or ''banana'' bonds (**figure 78, right)**. In bent bonds, the density distribution function describing the electron density of the C-C bond is deviated away from the planar cyclopropane ring. This arrangement increases the orbital angles and compensates for some of the ring strain.[42,306] In this arrangement, the hybridisation of the carbon is better described as an sp^5 hybridisation state. A consequence of sp^5 hybridisation is that the C-C bonds possess significant π-character and are shorter in length relative to alkanes.[42,306] Furthermore, the sp^5 hybridisation results in a significant s-character in the C-H bonds which renders the C-H bonds in cyclopropanes shorter and stronger than C-H bonds in ethane.[42,306] A consequence of the stronger C-H bond is that proton abstraction from cyclopropanes is more difficult than from a methyl group; this

property, in addition to the increased hydrophobicity, has seen the widespread incorporation of cyclopropyl-fragments into an assortment of pharmaceuticals.[42,306–161]

4.1.2 Synthesis of cyclopropane functional groups – a general overview

In general, the synthesis of cyclopropanes falls into six categories: i) carbene based reactions, ii) ylide reactions, iii) cyclisation reactions, iv) nucleophilic displacement reactions, vi) cyclisation reactions, and vi) miscellaneous reactions.[42,306] The carbene based syntheses can be further subdivided into i) free carbene reactions, ii) metallocarbenoid catalysed reactions, and iii) enzymatic cyclopropanation catalysed by a designed carbene transferase. The traditional synthetic approaches to the synthesis of cyclopropanes involving carbene reagents were briefly discussed in chapter 1 *section 1.4.1*. Chapter 2 *section 2.3.2* was dedicated to the recent advances in enzyme engineering and the successful development of carbene transferases, especially from P450s and myoglobin scaffolds, for cyclopropane synthesis. Having successfully demonstrated the proficiency of C45 towards the formation of a stable metallocarbenoid intermediate, the next question concerned whether C45 could be utilised as a carbene transferase, thus introduction a novel methodology to the synthesis of cyclopropanes from carbene reagents: enzymatic cyclopropanation catalysed by a *de novo* carbene transferase.

4.2 RESULTS AND DISCUSSION

4.2.1 Spectroscopic characterisation of C45 metallocarbenoid formation in the presence of receptive olefins

In chapter 3, the spectroscopic characterisation of C45 metallocarbenoid formation was reported with three distinct carbene precursors. The single turnover kinetic parameters for the formation of metallocarbenoids from EDA, BnDA and ^{t}BuDA for C45 were also determined. It was observed that the K_1 values for C45-metallocarbenoid formation followed the order EDA < BnDA < ^{t}BuDA. The three examined C45-metallocarbenoid intermediates were demonstrated to be stable complexes, with no observable degradation detectable in the UV-VIS spectra over a 1000 second period. This affords the possibility that C45 could be readily utilised as a carbene transfer catalyst *via* the reaction of the metallocarbenoid intermediate with a suitable substrate. Initial experiments were conducted using stopped-flow spectroscopy to examine the carbene transfer activity of the C45-EDA complex.

The experimental conditions employed to investigate carbene transfer activity paralleled the conditions initially used to investigate metallocarbenoid formation, with a single modification being the excess introduction of a suitable substrate, styrene. A solution containing ferrous C45 (15 µM

protein, stochiometric $Na_2S_2O_4$, CHES buffer, pH 8.6) and a solution containing ethyl diazoacetate (EDA) and styrene (100 µM -10 mM and 200 µM - 3mM respectively, 20%:80% CHES:EtOH) were prepared in separate compartments (syringes) before a small volume (2x20 µl) of each were injected concurrently into a mixing chamber. The mixing chamber was monitored by UV-VIS spectroscopy over several timescales to monitor any reactions occurring over the duration of the experiment. Final concentrations were 7.5 µM enzyme, stochiometric $Na_2S_2O_4$, 50 µM - 5 mM EDA, 100 µM - 1.5 mM styrene, and 40% EtOH.

Figure 79 displays the time-course spectra (obtained from the stopped-flow experiments) for C45 in the presence of EDA and styrene at wavelengths 417 and 433 nm. 417 nm corresponds to the ferrous C45 species and 433 nm corresponds to the C45-EDA metallocarbenoid complex, respectively. The data indicates that in the initial 20 seconds post-mixing there is a gradual decrease in the 417 nm absorbance and a concurrent increase in the 433 nm absorbance – corresponding to the formation of the C45-EDA metallocarbenoid complex from the ferrous species. It was demonstrated that once formed the C45-EDA metallocarbenoid complex is stable and persists over a duration of 1000 seconds without any appreciable decay (chapter 3, *section 3.5.1.4*; **figure 75**). In the presence of styrene, the spectra initially resemble the spectra acquired in the absence of styrene, corresponding to the formation of the metallocarbenoid intermediate. After approximately 30 seconds, however, the absorbance at 433 nm starts linearly decreasing and is accompanied by a concurrent linear increase in absorbance at 417 nm – indicating the decay of the metallocarbenoid species and the return to the starting ferrous complex. The differences between the time-dependent traces in the presence and absence of styrene at 417 and 433 nm respectively are presented in **figure 79**. **Figure 79** clearly demonstrates the plateauing of the absorbance changes at 417 and 433 nm in the absence of styrene, corresponding to saturation, is never accomplished in the presence of styrene. UV-VIS traces captured at numerous timepoints over 120 seconds is also presented in **figure 79.** In the initial stages of the reaction (1-20 seconds) the UV-VIS exhibits the red-shift corresponding to the formation of the metallocarbenoid intermediate, alongside the concurrent shift and broadening of the Q-band profile. The spectra acquired in the absence of styrene is characterised by the disappearance of the sharp Soret peak at 417 nm and the emergence of a new Soret peak at 428-433 nm, corresponding to the formation of the metallocarbenoid intermediate from the starting ferrous complex. In the presence of styrene, the emergence of the 433 nm peak appears as a shoulder alongside the 417 nm, resulting in a broad, split Soret peak, without either the 417 nm or 433 nm peak dominating the spectra. After 20 seconds, the maximum absorbance begins to shift towards the 417 nm peak and is accompanied by the gradual re-emergence of the ferrous Q-band profile.

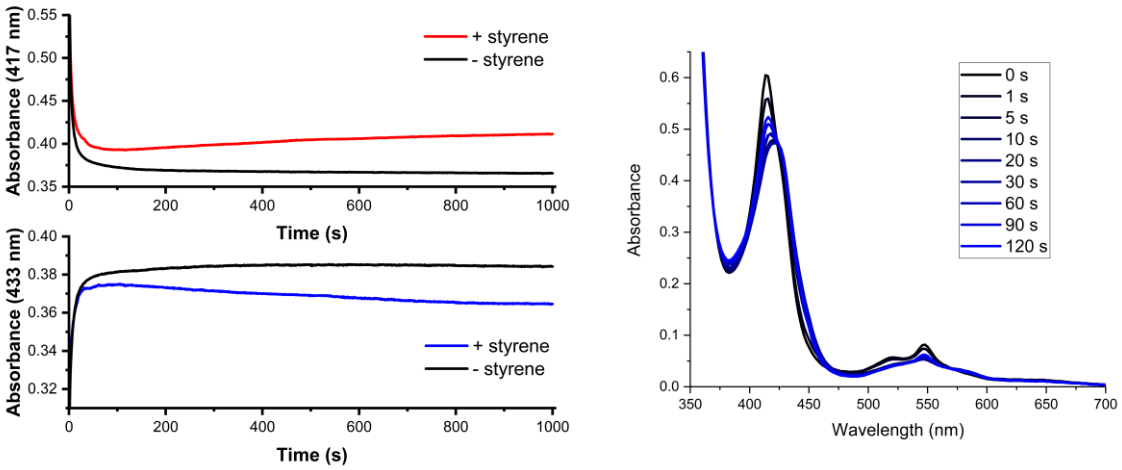


Figure 79: (**left**) Single wavelength traces representing the time course absorbance of ferrous C45 (417 nm, 7.5 µM protein, 10% EtOH) and the metallocarbenoid:C45 complex (433 nm) following rapid mixing of ferrous C45 with 500 µM ethyl diazoacetate at 5^{O}C in the absence (black traces) and presence (red trace for 417 nm; blue trace for 433 nm) of 3 mM styrene. (**right**) UV-VIS traces recorded at various time intervals following the rapid mixing of ferrous C45 with 500 µM ethyl diazoacetate at 5^{O}C in the presence of 3 mM styrene.

The spectroscopic data acquired from these initial experiments provides two observations: i) the presence of styrene results in the decay of a C45-metallocarbenoid complex, and ii) the decay of a metallocarbenoid complex results in the re-emergence of the initial C45 ferrous complex. It cannot, however, be concluded whether the metallocarbenoid is i) reacting with styrene to form a cyclopropane product, or ii) the presence of styrene destabilises a metallocarbenoid and causes its decomposition, possibly *via* displacement or dimerization. Additional experiments that address product identification and characterisation need to be conducted to assign a dynamic explanation to the above UV-VIS spectra.

4.2.3 C45 catalysed cyclopropanation assays

The spectroscopic information acquired for C45-metallocarbenoid formation in the presence of styrene indicates that some reactivity is taking place. In order to quantify the reactivity, it is necessary to identify the resultant products. In the literature, it has been well established that hemoproteins (such as P450, cytochrome *c*, and myoglobin) engineered towards metallocarbenoid formation are also capable of subsequently reacting with olefins to yield cyclopropane-containing products *via* a concerted carbene transfer reaction.[187] The typical substrates employed in the screening of carbene transferase activity are ethyl diazoacetate (EDA) and styrene (**figure 80**). The reaction between these two species results in a cyclopropane-containing product and is further distinguished by the generation of two stereogenic centres. The generation of two new stereogenic centres means that

four possible stereoisomers can be obtained from the reaction; therefore, in addition to carbene transfer activity, the assay can provide information on the stereoselectivity of the reaction.

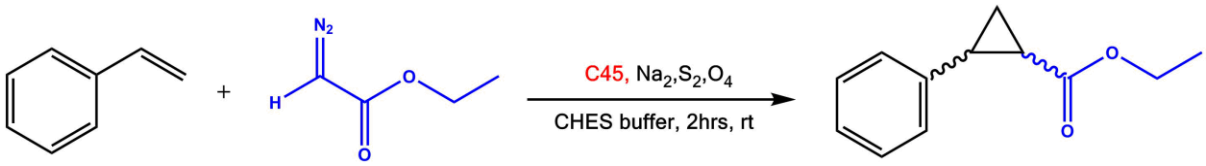


Figure 80: The general cyclopropanation reaction between styrene and EDA catalysed by C45.

A well-established protocol from the literature for screening for carbene transferase activity was employed to screen for C45 activity.[190] The protocol involves a small-scale (400 µl) reaction between EDA and styrene which is catalysed by the enzyme of interest (C45) under anaerobic conditions. The reaction is left to mix for 2 hours at room temperature before the reaction is quenched and the product is extracted into an organic solvent and analysed. The product mixture is analysed by HPLC, and a chiral column is employed as the stationary phase; chiral-HPLC affords the quantification of the stereoselectivity, in addition to product identification and reaction yields. LC-MS was subsequently used post-HPLC to corroborate product identification. Initial carbene transfer assays were conducted i) in the absence of any catalyst, ii) in the presence of free hemin, and iii) the heme-containing proteins myoglobin, haemoglobin, catalase, and horseradish peroxidase. The assays were performed in 1.5 mL seal-top reaction vials and the vials were purged with nitrogen prior to conducting the assays. The final reaction volumes were 400 µl, and consisted of 30 mM styrene, 10mM EDA, 10 mM sodium dithionite and 10 µM of the appropriate catalyst in CHES buffer (100 mM KCl, 20mM CHES, pH 8.6). The reactions were stirred for 2 hours, at room temperature, before being quenched with 3M HCl and the products extracted with ethyl acetate for analysis. The chromatograms for the six initial assays performed are presented in **figure 81**. It can be seen in **figure 81** that no product was identified in any of the performed assays.

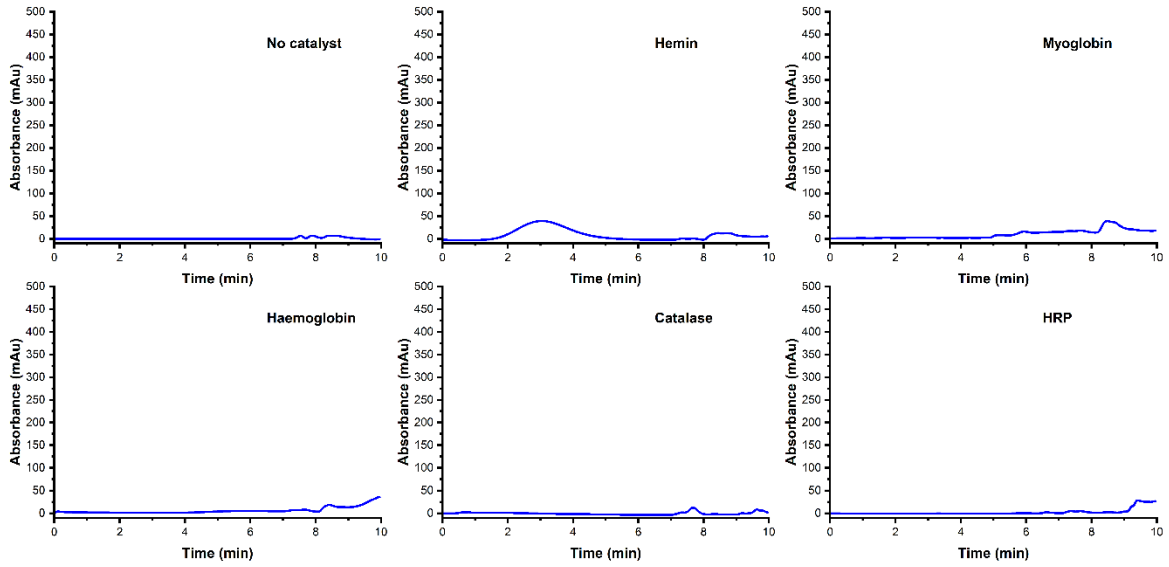


Figure 81: Chiral-HPLC chromatograms for the cyclopropanation assays between styrene (30 mM) and EDA (10 mM) (CHES buffer, pH 8.6, 254 nm) catalysed (10 µM, 0.1 % catalyst loading) by **(top left)** no catalyst, **(top middle)** hemin, **(top right)** myoglobin, **(bottom left)** haemoglobin, **(bottom middle)** catalase, and **(bottom right)** horseradish peroxidase (HRP). A polar organic mobile phase (100% MeCN: 0.1% (v/v) TFA:0.1% (v/v) Et_3N) was employed and injection volumes were 2 µl.

The expected cyclopropane product following the reaction between styrene and EDA is ethyl 2-phenylcyclopropane-1-carboxylate. A chromatogram of a commercial racemic sample of *trans*-ethyl 2-phenylcyclopropane-1-carboxylate was acquired; the chromatogram contained two peaks with retention time ranges 7.1-7.3 and 8.9-9.5 seconds, respectively. A commercial sample of enantiomerically pure *trans*-ethyl (*1R,2R*)-2-phenylcyclopropane-1-carboxylate was subsequently analysed, using chiral-HPLC and LC-MS, to identify which peak corresponded to which enantiomer, and to facilitate subsequent product analysis. The chiral-HPLC chromatogram of the commercial (*R,R*)- and (*S,S*)- products is presented in **figure 82 (top)**, and exhibits a symmetrical peak, characteristic of the classic Gaussian distribution, at a retention time of 7.1-7.3 seconds; by deduction, the 7.1-7.3 retention time corresponds to the (*R,R*)-enantiomer and the 8.9-9.5 retention time to the (*S,S*)-enantiomer. The enantiomerically-pure (*R,R*)-enantiomer was analysed *via* LC-MS, and the spectrum is displayed in **figure 83 (bottom)**. The spectrum is dominated by a peak at 145 m/z, corresponding to a stable carbonyl ion following the loss of EtOH from the parent $[M\text{-}H]^+$ *via* fragmentation. The spectrum also possess a $[M\text{-}H]^+$ parent ion at 191 m/z, corresponding to the exact mass of the protonated (*R,R*)-product.

The assay was then repeated, under identical conditions employed in the controls, in the presence of C45. The chromatogram acquired for the C45-cataysed assay is presented in **figure 82** against the chromatogram acquired for commercial enantiomerically-pure (*R,R*)-product. The chromatogram

possesses a dominant peak at 7.1 minutes shortly followed by a few minor peaks in the 8.5-9.1 minutes region. The comparison between the normalised average chromatograms for the C45-catalsyed carbene transferase assay, obtained from triplicate repeats, against the normalised chromatogram for the commercial (*R,R*)-sample reveals a very clear overlap between the two chromatograms, indicating the peak at 7.1-7.3 in the C45-catalysed carbene transfer assay can be assigned as the (*R,R*)-product.

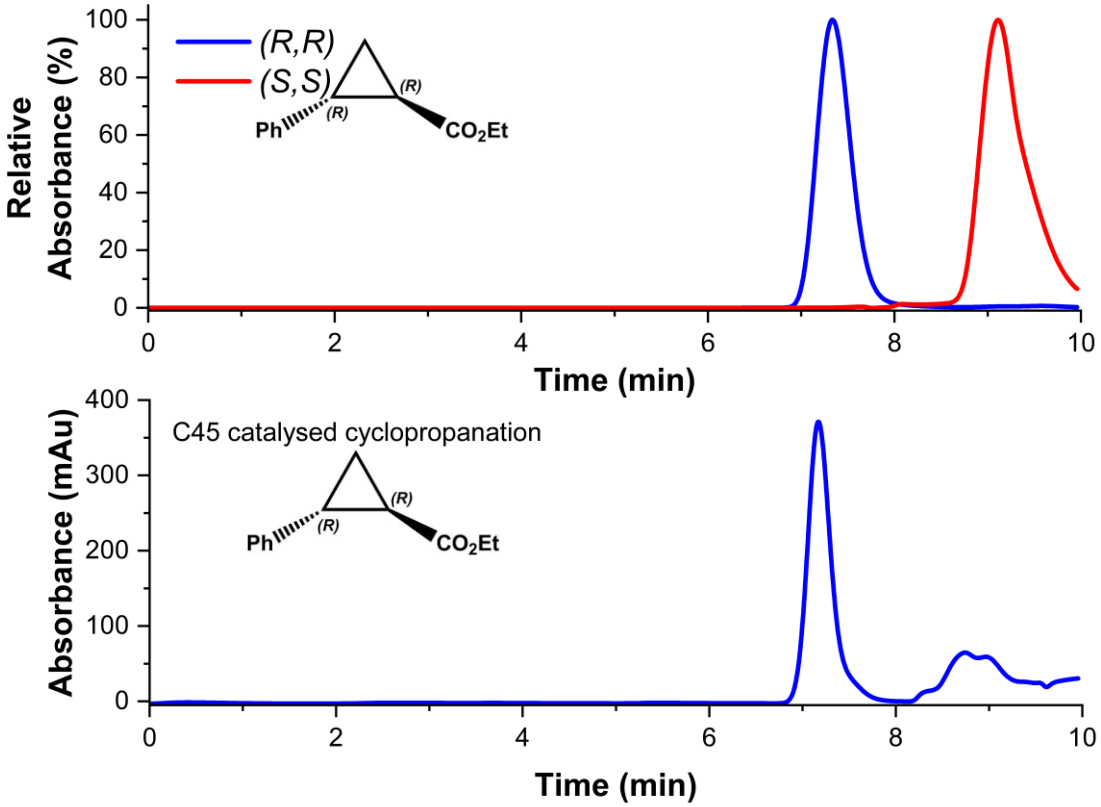


Figure 82: Chiral-HPLC chromatograms for the cyclopropanation assays. A polar organic mobile phase (100% MeCN: 0.1% (v/v) TFA:0.1% (v/v) Et_3N) was employed and injection volumes were 2 µl. (**top**) normalised chiral-HPLC chromatograms for commercial (*R,R*)-ethyl 2-phenylcyclopropane-1-carboxylate (254 nm, EtOH, blue) and (*S,S*)-ethyl 2-phenylcyclopropane-1-carboxylate (254 nm, EtOH, red). (**bottom**) Averaged chromatogram from the C45 (10 µM, 0.1 % catalyst loading) catalyzed cyclopropanation assay between styrene (30 mM) and EDA (10 mM) (CHES buffer, pH 8.6, 254 nm).

In order to confirm the identity of the 7.1-7.3 peak, the C45-catalysed assay mixture was analysed by LC-MS, and the acquired spectrum is displayed in **figure 83**. The mass spectrum possesses the dominant 145 m/z peak and the parent $[M\text{-}H]^+$ at 191 m/z, and excluding relevant intensities is practically identical to the spectrum obtained for the commercial sample.

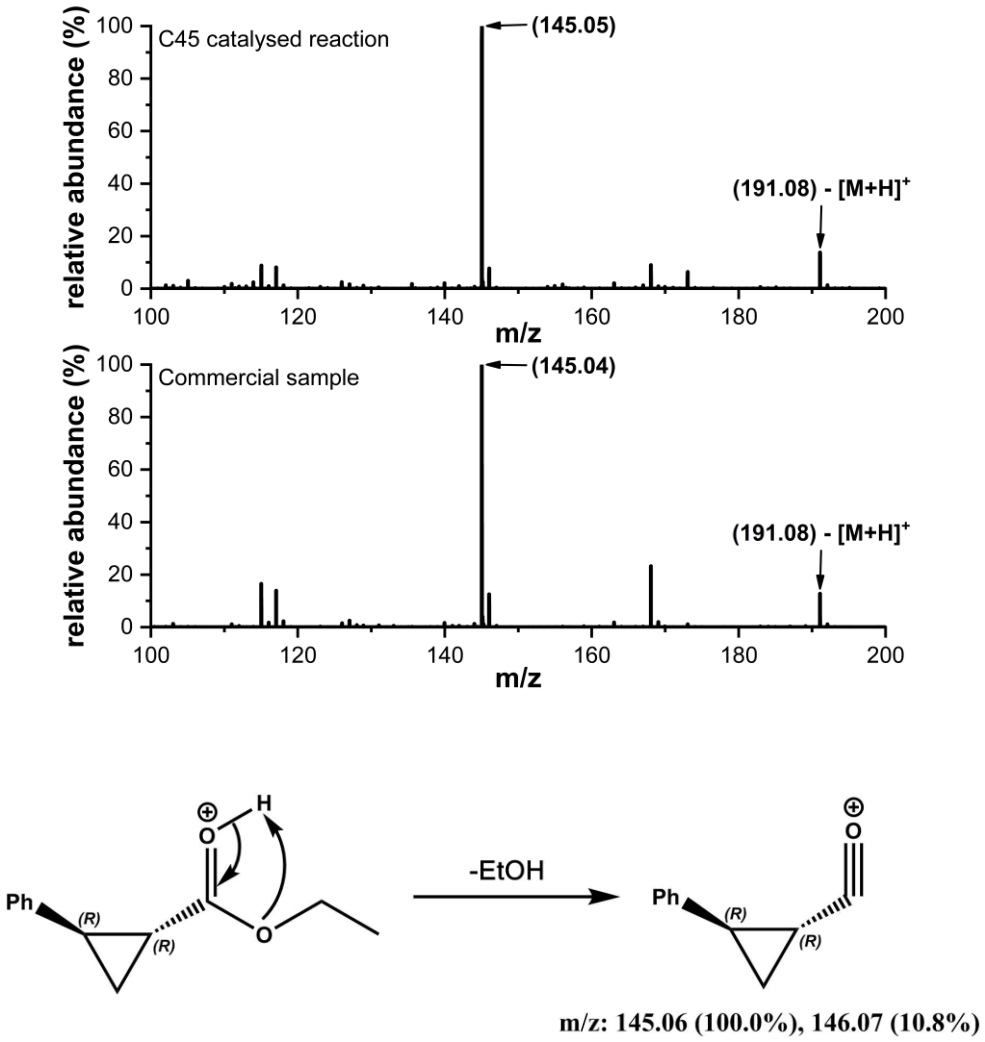


Figure 83: LC-MS spectra of C45 catalyzed cyclopropanation assay products. Commercial ethyl 2-phenylcyclopropane-1-carboxylate (in EtOH) exhibiting the dominant oxonium ion fragment at 145 m/z (**top MS**) and C45 catalyzed cyclopropanation assay between styrene (30 mM) and EDA (10 mM) (**bottom MS**). All spectra were recorded in ES+ mode and monitored at 254 and 280 nm. A C8 column was employed for the LC separation with a gradient mobile phase (95:5:0.1 % v/v water/MeCN/formate 10:90:0.1 % v/v water/MeCN/formate). Injection volumes were 20 µl. (**bottom**) Fragmentation pathway leading to the dominant oxonium ion at 145 m/z from the elimination of EtOH from the parent $[M+H]^+$ ion.

The spectroscopic data obtained for C45 metallocarbenoid formation in the presence of EDA and styrene indicated that the presence of styrene disrupts the formation of the metallocarbenoid, which was concluded from the simultaneous changes in the metallocarbenoid (433 nm) and ferrous species absorbance (417 nm) respectively. It could not be concluded, however, whether the disruption was a consequence of carbene transferase reactivity or destabilisation/degradation of the metallocarbenoid

complex. The chiral-HPLC and LC-MS data obtained from the C45-catalysed assay is consistent with the postulation that C45 is catalysing the carbene transfer reaction between styrene and EDA to yield a cyclopropane-product; therefore, the absorbance differences in the time-dependent spectra of C45/EDA in the absence and presence of styrene can be attributed to carbene transfer activity. The chromatograms obtained from the assays also allow the carbene transferase activity of C45 to be quantified (**figure 164**, *appendix 1*). An average product yield of 80% (*R,R*), TTN of 802, diastereoselectivity of >99.9% (trans) and enantioselectivity of 77% (*R,R*) for the C45-catalysed reaction between EDA and styrene was calculated.

In the literature, several carbene transferases, engineered from naturally occurring hemoproteins, have been reported.[190] **Figure 84** is a non-exhaustive list of data reported for various engineered hemoproteins catalysing the reaction between EDA and styrene presented as a scatter plot of TTN vs ee% (**table 17**, *appendix 2*).[190–192,195,198,199] It is immediately observable that the reported values for the product yields, enantioselectivities and TTNs in **figure 84** are relatively broad, with product yields ranging from 37% to >99.9%, enantioselectivities ranging from *racemic* to >99.9% and TTNs ranging from 186-3700.[190–192,195,198,199] The two best mutants are the myoglobin variants Mb(F43V,V68F) and Mb(H64V,V68A), each with reported values of >99.9% product yields, >99.9% enantioselectivity, and TTNs of 500.[195] In comparison to C45, the product yields and enantioselectivities are objectively better, but the dominant product of the reactions is the (*S,S*)-enantiomer, the opposite product to C45. Furthermore, C45 possess a higher TTN (802 vs 500), a consequence of the authors reporting catalytic loadings of 20 µM for each enzyme relative to 10 µM used in the C45 cyclopropanation assays. In terms of product yield, four additional mutants, Mb(F43V,V68A),[195] Mb(L29T,H64V),[199] Mb(H64V,V68S)[199] and T268-axH,[193] have a higher yield (on average) relative to C45, with 91%, 93%, 82% and 81% respectively relative to 80% (C45). However, three of the four mutants, Mb(F43V,V68A), Mb(L29T,H64V) and T268-axH, exhibit lower enantioselectivities, 67%, 59% and 42% respectively, relative to C45 (77%) as well as lower TTNs (455, 460, 710 vs 802).[195,199] Mb(H64V,V68S) (in addition to Mb(F43V,V68F) and Mb(H64V,V68A)) has a superior product yield and enantioselectivity, 82% and 99.9% respectively, relative to C45, but possess a TTN almost half that of C45 (410 vs 802).[195,199] Only two mutants, P411BM3 and P411BM3-CIS, possess higher TTNs relative to C45 (2900 and 3700 respectively vs 802), but present low product yields (50 and 55% respectively);[191] in addition, these assays were conducted using whole cells expressing the mutants and so cannot be directly compared (and are excluded from **figure 85**).[191] An assay conducted with purified P411BM3-CIS resulted in a 40% product yield and a 199 TTN, although a 94% enantioselectivity was reported.[195] The data demonstrates that functionality can drastically vary among engineered hemoproteins, with C45, initially designed as a *de novo* peroxidase, behaving remarkably similar to designed carbene

transferases. The data indicates the C45 maquette, in addition to displaying reasonable carbene transferase activity, is a suitable scaffold for further modification to manufacture more selective and efficient *de novo* carbene transferases. This will be explored in detail in chapter 7.

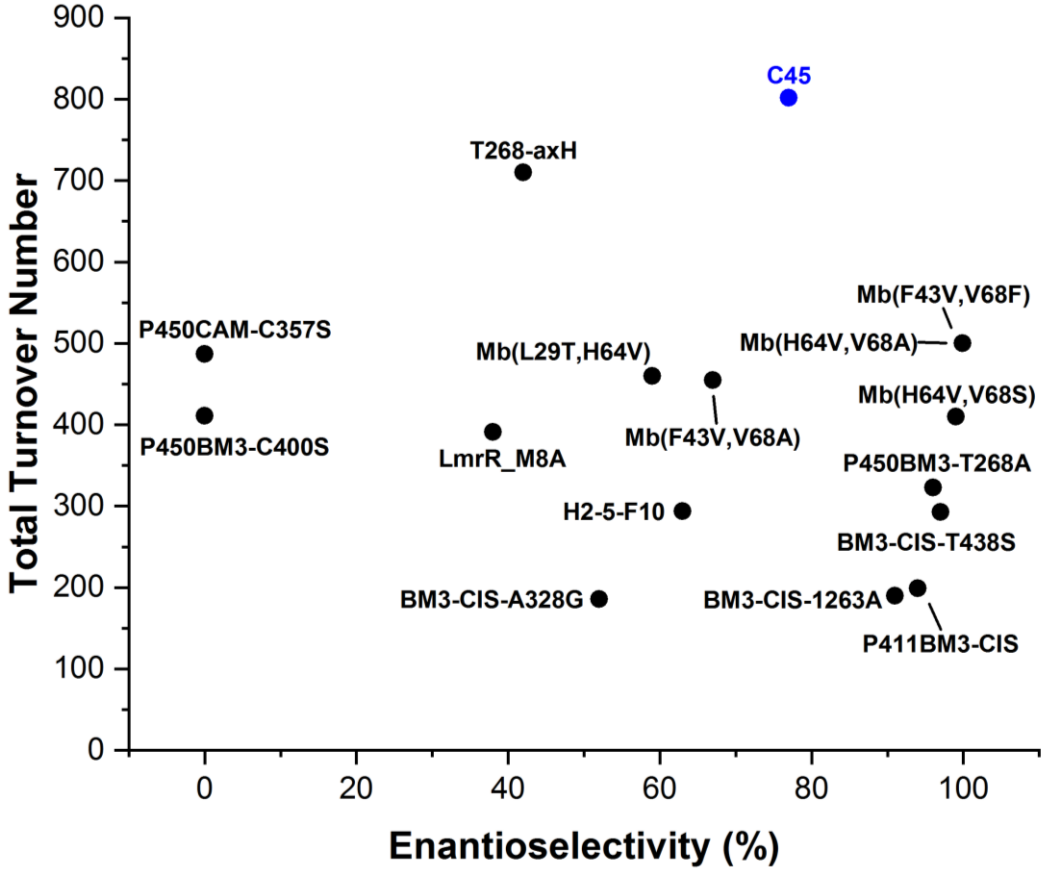


Figure 84: A comparison of TTN vs enantioselectivity for several rationally designed carbene transferases relative to C45 (blue).[190–192,195,198,199]

4.2.3.1 Substrate specificity – substituted diazo compounds and para-functionalised styrenes

The C45-catalysed cyclopropanation assay was repeated using ^{t}BuDA and BnDA as the diazo precursors (the chiral-HPLC traces for the reaction products are presented in **figure 85**). As in the assays with EDA, the chromatograms for ^{t}BuDA and BnDA are dominant by a major peak with a retention time of 7.0-7.3, the known retention time for the (*R,R*)-cyclopropane product. Although activity can be detected for both ^{t}BuDA and BnDA, the product yields and TTNs of the assays are considerably lower (^{t}BuDA = 18.45%, 189; BnDA = 19.6%, 196 respectively; **table 3**) while the enantioselectivities, although less affected, are also lower relative to EDA (^{t}BuDA 40.64%; BnDA 66.83%; **table 3**).

Table 3: The product yields, diastereoselectivities, enantioselectivities and TTN of the C45-catalysed cyclopropanation reaction between styrene and EDA, ^{t}BuDA and BnDA.

Substrates (vs styrene)	% Yield$_{(R,R)}$	% Yield$_{(S,S)}$	dr$_{(E)}$ %	ee$_{(R,R)}$ %	TTN$_{(R,R)}$
EDA	80.20 (± 8.01)	10.42 (± 1.8)	>99	77.00 (± 2.82)	802
^{t}BuDA	18.45 (± 8.12)	7.00 (± 1.54)	>99	40.64 (± 16.5)	189
BDA	19.45 (± 1.44)	4.00 (± 1.41)	>99	66.83 (±7.43)	196

The difference in activity towards ^{t}BuDA and BnDA relative to EDA can be partially accounted for by sterics, metallocarbenoid stability, and protein-carbene intermolecular interactions. The K_1 values for C45-metallocarbenoid formation for each diazo precursor was presented in chapter 3, *section 3.5.1.3*. The values were calculated to be 74.36 113.16 and 219.43 µM for EDA, BnDA and ^{t}BuDA respectively (**table 1**, chapter 3, *section 3.5.1.2*). The rates of formation for C45-metallocarbenoid formation were also calculated as 0.498, 0.349 and 0.195 s^{-1} for EDA, BnDA and ^{t}BuDA, respectively. The single turnover kinetic data indicates that C45 *preferentially* forms a metallocarbenoid with EDA, but the kinetic data provides no information on the thermodynamic stability of the resultant complex. Computational studies have provided insight into metallocarbenoid stability and highlighted an increased steric bulk on the carbene carbon's substituents decreases the rate of formation but also provides a stabilising effect on the resultant complex post-formation. The product yields acquired from the cyclopropanation assays support this claim, as the product yields negatively correlate with the K_1 values for metallocarbenoid formation. The increase in K_1 for metallocarbenoid formation is a consequence of the increased steric bulk of the carbene carbon's substituents, which hinders accessibility into the active site and requires a reorganisation of the active site residues to accommodate the bulky ligand; however, upon formation, the bulkier ligand *stabilises* the metallocarbenoid, which hinders reactivity with a suitable substrate. The decrease in enantioselectivity for ^{t}BuDA and BnDA is probably a consequence of reorganisation of the active site residues to accommodate the bulkier substituents. However, additional C45-cataysed cyclopropanation assays employing a range of different functionalised diazo precursors would need to be conducted, and subsequently analysed, to corroborate the above hypothesis.

The moderate enantioselectivity exhibited for BnDA could also be possibly accounted for by π-π interactions between the phenyl group of BnDA and nearby aromatic residues, an interaction that has been identified as important for governing the stereoselectivity in the literature. Any π-π interactions would facilitate the formation of the metallocarbenoid (K_1(BnDA) < K_1(^{t}BuDA)) but would also stabilise the resultant conformation with the strongest π-π interactions. As the strength of π-π interactions is a function of the relative angle between the two π systems, the conformer of the metallocarbenoid

that can provide the strongest interactions would be energetically preferred and be the foremost conformer adopted. In terms of carbene transfer reactivity, the existence of a dominant conformation(s) would account for the observed stereoselectivity of BnDA relative to ^{t}BuDA. At present, however, it cannot be concluded whether sterics, electronics or both govern the observed stereoselectivity of the C45-catalysed cyclopropanation reaction. A combination of experimental techniques and computational analysis (DFT, MD, QM/MM) will be employed to elucidate these questions in the future.

Having demonstrated that C45 can function as a stereoselective carbene transferase towards styrene the next line of investigation concerned substrate specificity. An assortment of *para*-functionalised styrenes were used to explore how the electronics/sterics of functionalised styrenes influenced the reaction yields and selectivity. The selected substrates are presented in **table 4**. EDA was employed as the diazo precursor substrate in all the assays. The reaction conditions employed for each substrate were identical to those employed in the aforementioned assays were styrene served as the substrate. All assays were subject to chiral-HPLC and LC-MS analysis to quantify product yields and stereoselectivities. All product yields were determined by reference to an external calibration, with each product synthesised by following a well-established protocol that uses $Rh_2(OAc)_4$ as the catalyst (see chapter 9, *section 9.10.2.1*).

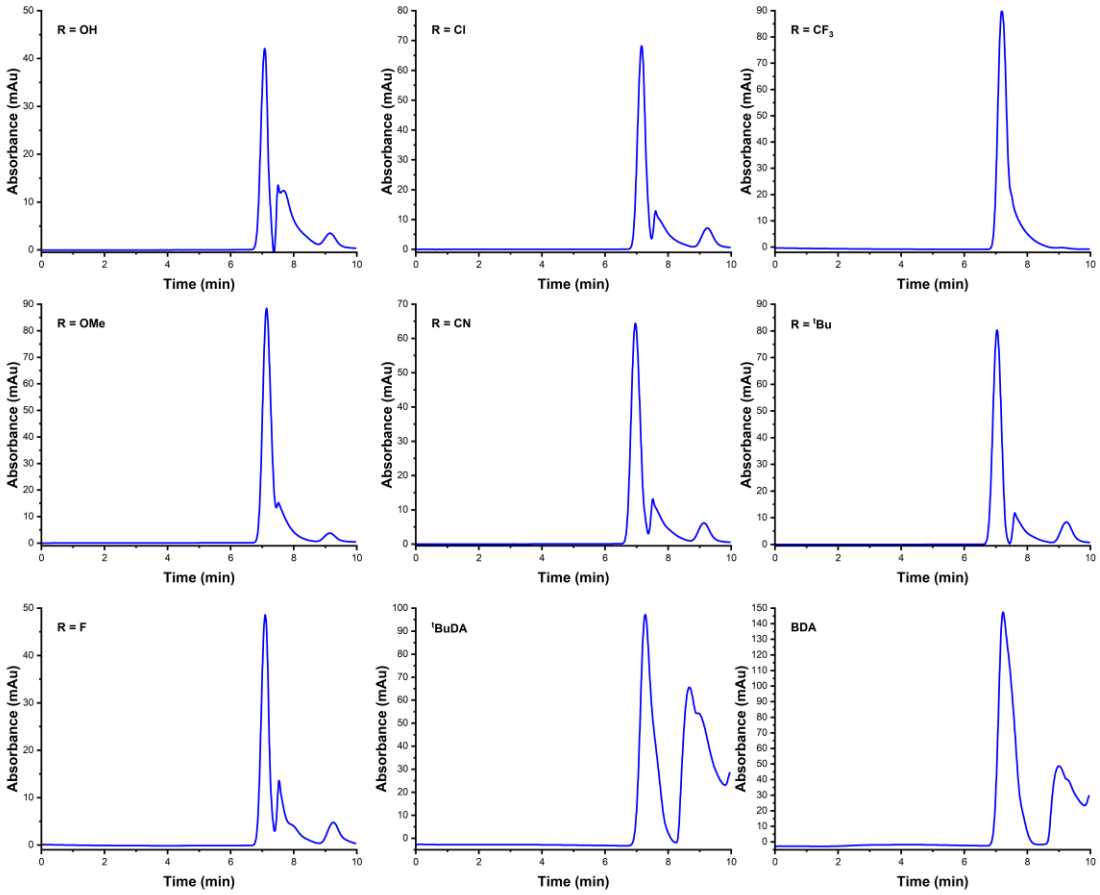


Figure 85: Chiral-HPLC chromatogram for the C45 (10 µM) catalyzed cyclopropanation assays. (**top left**) *p*-hydroxystyrene (30 mM) and EDA (10 mM); (**top middle**) *p*-chlorostyrene (30 mM) and EDA (10 mM); (**top right**) *p*-trifluoromethylstyrene (30mM) and EDA (10 mM); (**middle left**) *p*-methoxystyrene (30 mM) and EDA (10 mM); (**middle**) *p*-cyanostyrene (30 mM) and EDA (10 mM); (**middle right**) *p-tert*-butylstyrene (30 mM) and EDA (10 mM); (**bottom left**) *p*-fluorostyrene (30 mM) and EDA (10 mM); (**bottom middle**) styrene (30 mM) and *tert*-butyl diazoacetate (10 mM); (**bottom right**) styrene (30 mM) and benzyl diazoacetate (10 mM). All assays were performed in CHES buffer (pH 8.6) with 10 µM C45 (0.1 % catalyst loading). A polar organic mobile phase (100% MeCN: 0.1% (v/v) TFA:0.1% (v/v) Et_3N) was employed and injection volumes were 2 µl.

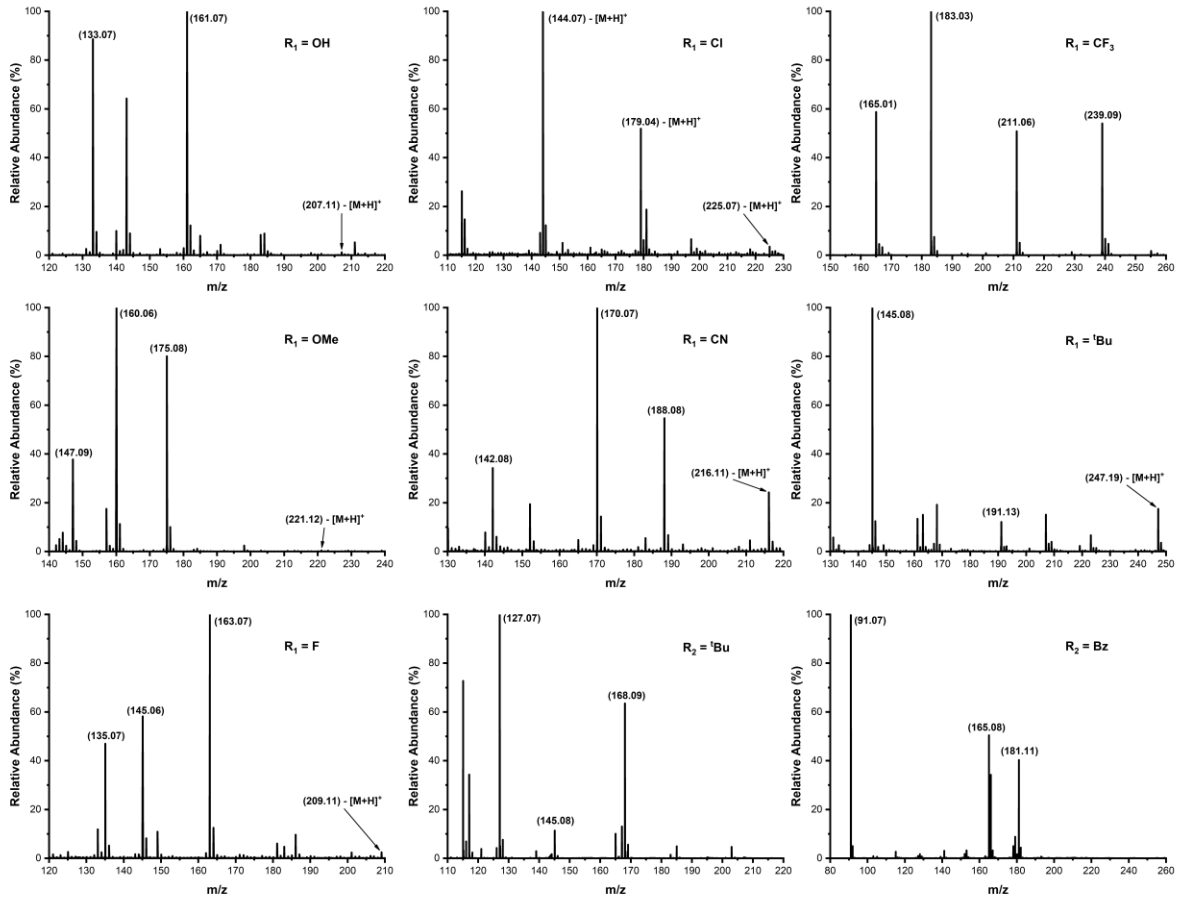


Figure 86: LC-MS spectra for the C45 (10 µM) catalyzed cyclopropanation assays. (**top left**) *p*-hydroxystyrene (30 mM) and EDA (10 mM); (**top middle**) *p*-chlorostyrene (30 mM) and EDA (10 mM); (**top right**) *p*-trifluoromethylstyrene (30mM) and EDA (10 mM); (**middle left**) *p*-methoxystyrene (30 mM) and EDA (10 mM); (**middle**) *p*-cyanostyrene (30 mM) and EDA (10 mM); (**middle right**) *p-tert*-butylstyrene (30 mM) and EDA (10 mM); (**bottom left**) *p*-fluorostyrene (30 mM) and EDA (10 mM); (**bottom middle**) styrene (30 mM) and *tert*-butyl diazoacetate (10 mM); (**bottom right**) styrene (30 mM) and benzyl diazoacetate (10 mM). All spectra were recorded in ES+ mode and monitored at 254 and 280 nm. A C8 column was employed for the LC separation with a gradient mobile phase (95:5:0.1 % v/v water/MeCN/formate 10:90:0.1 % v/v water/MeCN/formate). Injection volumes were 20 µl.

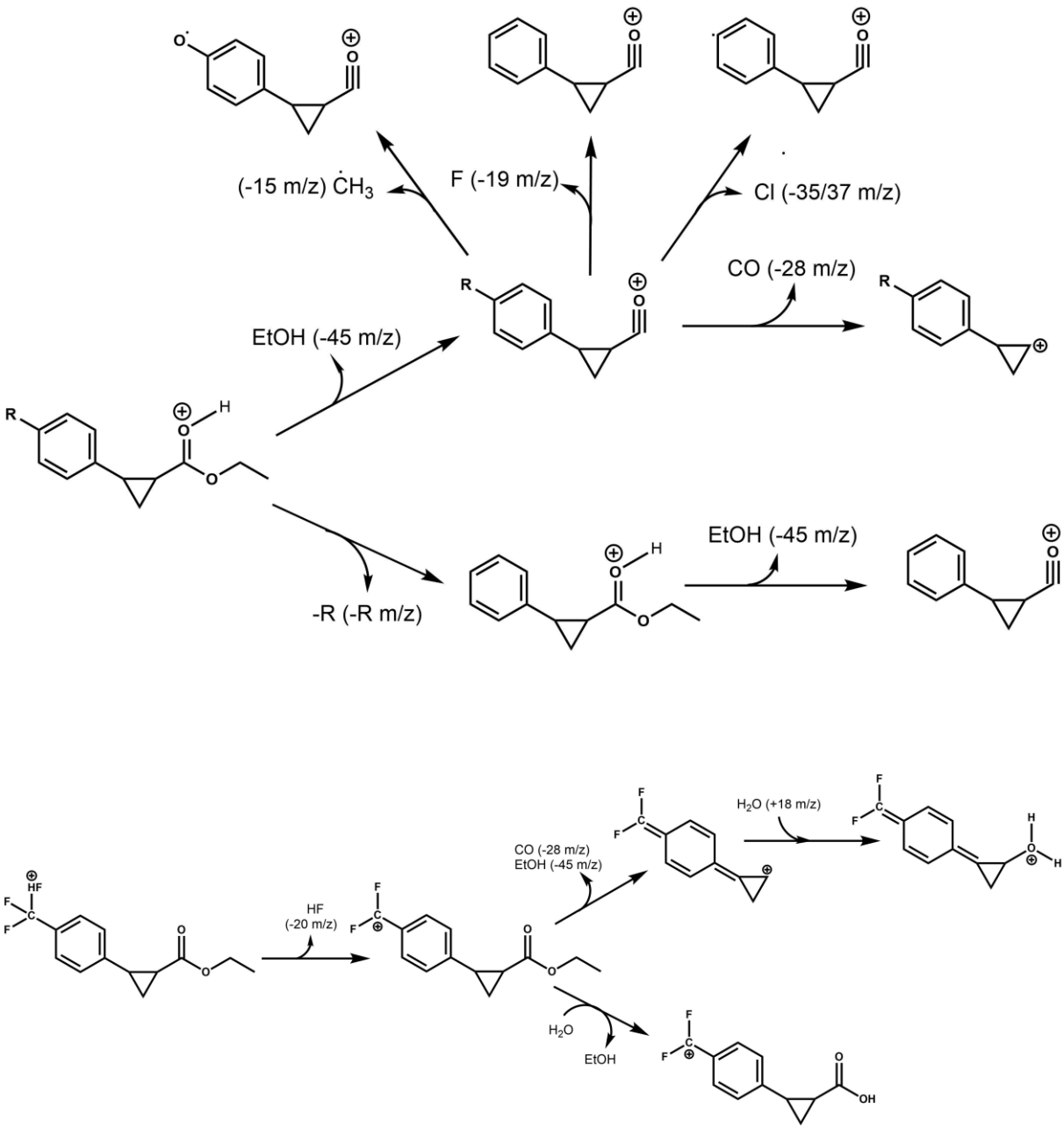


Figure 87: (**top**) Fragmentation pathways leading to the observed peaks in the mass spectra. (**bottom**) Fragmentation pathway for the product obtained using *p*-trifluoromethylstyrene as the starting substrate, resulting in the peaks observed in the product mass spectrum.

The chiral-HPLC traces acquired for the C45-catalysed cyclopropanation assay for each *para*-functionalised styrene are presented in **figure 85** with yields, stereoselectivities and TTNs presented in **table 4**. The C45-catalysed cyclopropanation assay chromatograms for each respective *para*-functionalised styrene is characterised by a dominant peak around retention time 7.0-7.5 minutes corresponding to the (*R,R*)-enantiomeric product, with many chromatograms exhibiting an additional peak with retention time 9.0-9.5 minutes corresponding to the *(S,S)*-enantiomeric product. The product yields, enantioselectivities, and TTNs for each *para*-functionalised styrene substrate are

presented in **table 4**. From **table 4** there is no consistency in the product yields across the substrate range. The lowest product yield is 36% for *p*-hydroxystyrene. The yields for *p*-chlorostyrene, *p*-fluorostyrene, *p*-vinylphenol, *p*-methoxystyrene, and *p*-*tert*-butylstyrene all are mid-range, with reasonable product yields ranging from 46-57% respectively. The yield for *p*-(trifluoromethyl)styrene is very good, with an average product yield of 79%. From the product yields, there is no immediately apparent relationship between the electronics of the *para*-functionalised substituent and the product yields of the cyclopropanation assays. From electronics alone, it is expected that *p*-*tert*-butylstyrene and *p*-methoxystyrene would increase the reactivity by donating electron density towards the nucleophilic vinyl group of styrene, which is observed, but *p*-trifluoromethylstyrene should exhibit the opposite relationship, a decrease in the reactivity of the substrate as *p*-CF_3 withdraws electron-density from the vinyl nucleophilic group. In addition to electronics, the range of product yields with respect to *para*-functionalisation cannot be immediately accounted for by sterics, although all the yields are lower than for styrene (80%). Irrespective of the ambiguous relationship between styrene functionalisation and product yield, what can be deduced is C45 exhibits good substrate specificity and is capable of turning over an assortment of differing styrene substrates. **Figure 86** presents the LC-MS spectra acquired for each functionalised styrene substrate in addition to the molecular structure of the key fragments observed in each mass spectrum (**figure 87**).

Table 4: The product yields, diastereoselectivities, enantioselectivities and TTN of the C45-catalysed cyclopropanation reaction between EDA and a selection of *para*-functionalised styrene substrates

EDA vs R =	% Yield$_{(R,R)}$	% Yield$_{(S,S)}$	dr$_{(E)}$ %	ee$_{(R,R)}$ %	TTN$_{(R,R)}$
-^{t}Bu	56.73 (± 1.00)	5.52 (± 0.12)	>99	82.27 (± 0.37)	567
-OMe	45.50 (± 1.74)	1.71 (± 0.10)	>99	92.77 (± 0.16)	455
-OH	36.51 (± 5.73)	2.65 (± 0.67)	>99	86.61 (± 1.31)	365
-F	47.00 (± 0.75)	0.83 (± 0.01)	>99	96.72 (± 0.20)	470
-Cl	50.70 (± 3.34)	5.81 (± 0.19)	>99	79.38 (± 0.59)	510
-CN	55.00 (± 8.66)	4.75 (± 1.41)	>99	84.16 (± 0.50)	550
-CF_3	79.65 (± 4.05)	0.14 (± 0.04)	>99	99.65 (± 0.11)	797

Table 4 also provides the calculated values for the enantioselectivities for each assay conducted. In direct contrast to product yield, the range of enantioselectivities are consistent and high. Only *p*-chlorostyrene reports a lower enantioselectivity relative to styrene, but with an approximate deviation of >1%. For the other six substrates investigated the enantioselectivities are consistently higher

relative to styrene, with *p*-methoxy-, *p*-fluoro-, and *p*-trifluoromethyl- styrene each exhibiting exceptional enantioselectivities of >90% (93, 97, 99% respectively).

The origin of the observed stereoselectivity for carbene transfer cyclopropanation has been explored in the literature.[305] It was postulated by Fasan that two key factors govern the stereoselectivity: i) the conformation of the heme-localised metallocarbenoid, and ii) the geometry and orientation of the styrene-carbene transition state which is mediated through a combination of steric effects and various non-covalent intermolecular interactions between the active site residues and both the carbene and styrene substrates. In particular, π-π interactions between styrene and an aromatic residue on the protein were identified as key interactions governing the stereoselectivity (**figure 88**).[305] MD simulations investigating the origin of stereoselectivity in C45-catalysed cyclopropanation reactions concluded that C45 is structurally homogenous, accommodating two possible orientations of the carbene in the active site. The simulations, however, revealed only a minor energetic preference for the enantioselective attack of the substrate on the pro-*R* face of the metallocarbenoid intermediate. It was reported that the orientation of the metallocarbenoid was influenced by the presence of a substrate, reinforcing the hypothesis that the stabilisation of the transition state is more important to the overall stereoselectivity. The simulations also revealed the majority of stabilising interactions between the substrate, carbene and active site are hydrophobic in nature, and demonstrated a lack of hydrogen bonds despite the existence of several hydrogen bonding donors/acceptors. The results of the simulations support the conclusions from Fasan (**figure 88**).[305]

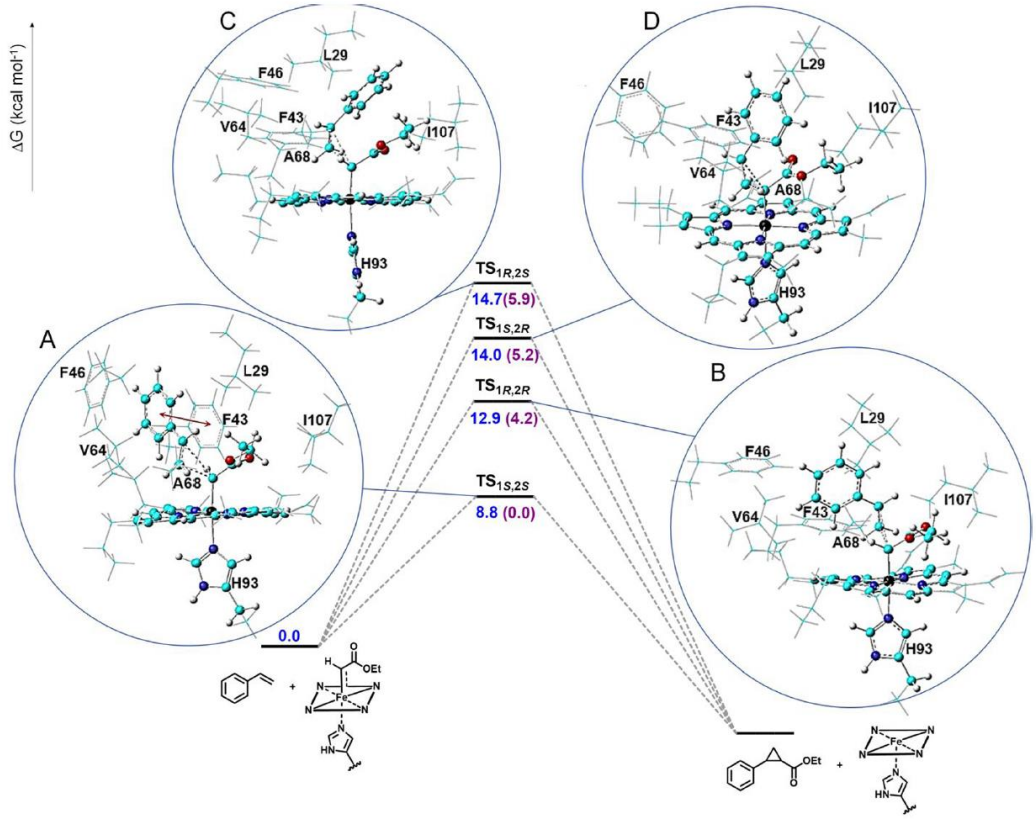


Figure 88: The free energy diagram for the four possible optimised transition states for the Mb(H64V,V68A)-catalysed cyclopropanation reaction between EDA and styrene leading to the four possible stereoisomers: **A)** TS_{1S2S}, **B)** TS_{1R2R}, **C)** TS_{1S2R}, and **D)** TS_{1R2S} (figure taken from Tinoco *et al*).[305]

As mentioned, Fasan highlighted the importance of π-π interactions between the styrene substrate and aromatic groups in the active site of the protein (**figure 88**).[305] In C45, a tryptophan residue in close proximity to the heme cofactor (the tryptophan that the cationic radical of C45-compound I is localised) can provide the necessary π-π interactions with styrene and could account for the stereoselectivity observed. Interestingly, the two highest enantioselectivities observed in the C45-cyclopropanation assays were with *p*-fluoro- and *p*-(trifluoromethyl)- styrene, with enantioselectivities of 99.64 and 96.62% respectively. It is possible the highly polarised and short C-F σ-bond provides stabilising electrostatic interactions and/or possible halogen bonds are provided from the substituents which can either stabilise the transition state and/or stabilise a particular trajectory of approach of the substrate towards the metallocarbenoid. However, the molecular origins of the stereoselectivity observed in the C45-catalysed cyclopropanation assays is currently unclear and therefore additional experiments are imperative. There are three possible experiments that would provide useful information on the parameters governing stereoselectivity: i) using site-directed mutagenesis to introduce select point mutations into C45's active site to identify key amino acids governing the stereoselectivity of the reaction, ii) employing a variety of fluoro-functionalised

styrenes to investigate the nature of the C-F bond in governing stereoselectivity, and iii) additional computational simulations (DFT, MD, QM/MM) to elucidate the key interactions that 1) govern the trajectory of approach of the styrene towards the metallocarbenoid intermediate, and 2) stabilise the metallocarbenoid-substrate transition state.

4.2.4 Spectroscopic characterisation of *Rma*-TDE and Mb(H64V,V68A) metallocarbenoid formation in the presence of receptive olefins – a direct comparison with C45

The UV-VIS spectra of the ferrous and ferric *Rma*-TDE and Mb(H64V,V68A) in the absence and presence of EDA were reported in chapter 3 *section 3.5.1*.4. The spectra for both enzymes were consistent with the formation of a stable, persistent, heme-localised metallocarbenoid species; which was consistent with the X-ray crystallographic structures reported in the literature for *Rma*-TDE and Mb(H64V,V68A) EDA-metallocarbenoid complexes.[219,304] Having successfully demonstrated that C45 i) forms a persistent, stable heme-localised metallocarbenoid complex when introduced to a suitable carbene precursor, and ii) exhibits proficient carbene transfer activity in the presence of suitable olefins, it would be prudent to investigate, spectroscopically, the carbene transfer activity of both *Rma*-TDE and Mb(H64V,V68A) metallocarbenoid complexes in the presence of styrene. A direct comparison between literature reported carbene transferases and C45 would provide a beneficial means of comparing *de novo* and re-engineered natural hemoproteins. The experimental conditions employed to investigate carbene transfer activity for *Rma*-TDE and Mb(H64V,V68A) were identical to the conditions used to investigate C45 activity in the presence of styrene. A solution containing ferrous enzyme (15 µM protein, stochiometric $Na_2S_2O_4$, CHES buffer, pH 8.6) and a solution containing EDA and styrene (100 µM - 1 mM and 200 µM – 3 mM respectively, 20%:80% CHES:EtOH) was prepared in separate compartments (syringes) before a small volume (2x20 µl) of each were concurrently injected into a mixing chamber. The mixing chamber was monitored by UV-VIS spectroscopy over several timescales to monitor any reactions occurring over the duration of the experiment. Final concentrations were 7.5 µM enzyme, stochiometric $Na_2S_2O_4$, 50 µM - 500 µM EDA, 100 µM - 1.5 mM styrene, and 40% EtOH.

The time-course spectra for the stopped-flow experiments for *Rma*-TDE in the presence of EDA (500 µM) and styrene (3 mM) at wavelengths 417 and 433 nm is presented in **figure 89**. 417 nm and 433 nm correspond to ferrous *Rma*-TDE and the *Rma*-TDE-(EDA) metallocarbenoid complex, respectively. In the absence of styrene, after the initial 20 seconds post-mixing there is a gradual decrease in the 417 nm absorbance accompanied by a concurrent increase in the 433 nm absorbance – corresponding to the formation of the *Rma*-TDE-(EDA) metallocarbenoid complex from the ferrous species (chapter 3, *section 3.5.1.4*).

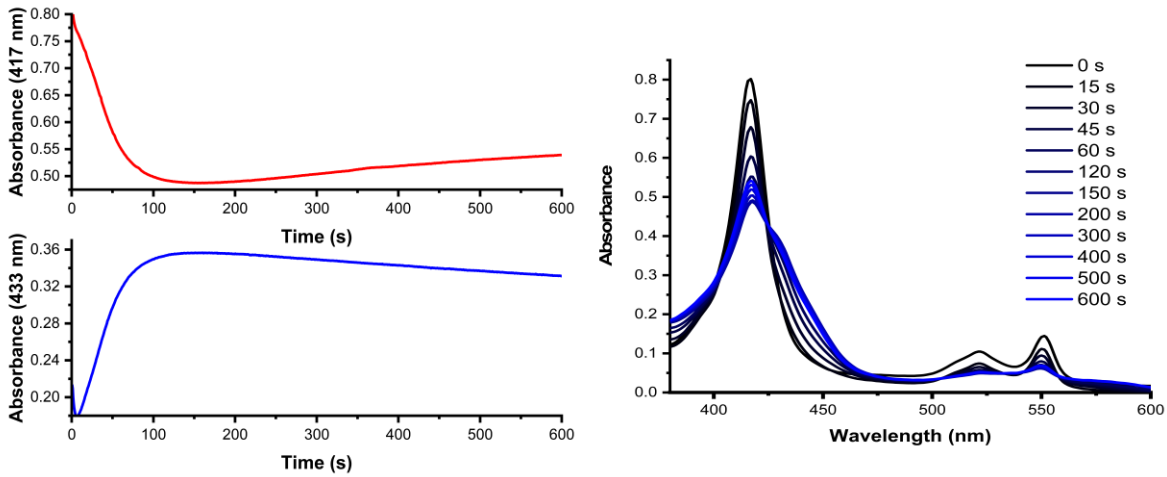


Figure 89: (**left**) Single wavelength traces representing the time course absorbance of ferrous *Rma*-TDE (417 nm, 7.5 µM protein, 10% EtOH) and the metallocarbenoid:*Rma*-TDE complex (433 nm) following rapid mixing of ferrous *Rma*-TDE with 500 µM ethyl diazoacetate at 5^{O}C in the presence of 3 mM styrene. (**right**) UV-VIS traces recorded at various time intervals following the rapid mixing of ferrous *Rma*-TDE with 500 µM ethyl diazoacetate at 5^{O}C in the presence of 3 mM styrene.

The time-dependent spectra in the presence of styrene is presented in **figure 89**. It is apparent that the formation of the *Rma*-TDE-(EDA) metallocarbenoid complex exhibits a lag phase, requiring a duration of approximately 100 seconds before reaching saturation. After 100 seconds, the absorbance at 433 nm begins to gradually decline and is accompanied by a paralleled increase in the absorbance at 417 nm. The observed changes post-100 seconds continue for up to 500 seconds, with no appreciable change in the linear profile of each wavelength. **Figure 89 (right)** is a collection of snapshots of the UV-VIS spectra, taken at various intervals, for the formation of the *Rma*-TDE-(EDA) metallocarbenoid complex in the presence of styrene. **Figure 89 (right)** shows a clear shift towards the metallocarbenoid species (433 nm) in the initial stages of the reaction, but over time the spectra is characterised by the re-emergence of the 417 nm ferrous *Rma*-TDE peak, shouldered by a continuously diminishing *Rma*-TDE-(EDA) metallocarbenoid peak at 433 nm. The data collected from the stopped-flow indicates that *Rma*-TDE is catalytically active towards styrene, catalysing a carbene transfer reaction and returning to the initial active ferrous state. The activity suggested by the UV-VIS spectra also accounts for the lag phase corresponding to the formation of the *Rma*-TDE-(EDA) metallocarbenoid complex. It should be noted that a 1:6 ratio of EDA:styrene was required to observed any spectroscopic activity, a much greater ratio required than for C45 (1:30), suggesting that, although *Rma*-TDE can catalyse the cyclopropanation of styrene, the affinity for EDA is considerably lower and requires elevated concentrations to reach saturation, an observation supported by the reported K_1 values for *Rma*-TDE-(EDA) and C45 (for EDA) in chapter 3 (**table 2**, chapter 3*, section 3.5.1.5*). In addition, the much greater lag phase prior to the decline for the *Rma*-TDE-(EDA)

metallocarbenoid peak suggests that *Rma*-TDE has a much lower affinity for styrene relative to C45. This observation will be returned to shortly.

The time-course spectra for the stopped-flow experiments for Mb(H64V,V68A) in the presence of EDA (100 µM) and styrene (3 mM) at wavelengths 435 and 424 nm is presented in **figure 90**. 435 nm corresponds to the ferrous Mb(H64V,V68A) species and 424 nm corresponds to the Mb(H64V,V68A)-EDA metallocarbenoid complex, respectively. In the absence of styrene the initial 20 seconds post-mixing shows a rapid decrease in the 435 nm absorbance accompanied by a concurrent increase in the 424 nm absorbance – corresponding to the formation of the Mb(H64V,V68A)-EDA metallocarbenoid complex from the ferrous species (chapter 3, *section 3.5.1.4*). This is followed by a 100 second phase marked by a gradual change in both respective absorbances. In the presence of styrene, the initial formation of the Mb(H64V,V68A)-EDA metallocarbenoid complex displays the same initial burst, followed by a major lag phase which persists for over 500 seconds; the 435 peak, corresponding to ferrous Mb(H64V,V68A) does not being to re-emerge over the duration of the experiment, and requiring a duration of approximately 100 seconds before reaching saturation. In contrast, the 424 nm peak, corresponding to the Mb(H64V,V68A)-EDA metallocarbenoid complex, also exhibits a lag phase in the initial formation, and begins to slowly diminish after approximately 200 seconds, which continues for the remainder of the experiment.

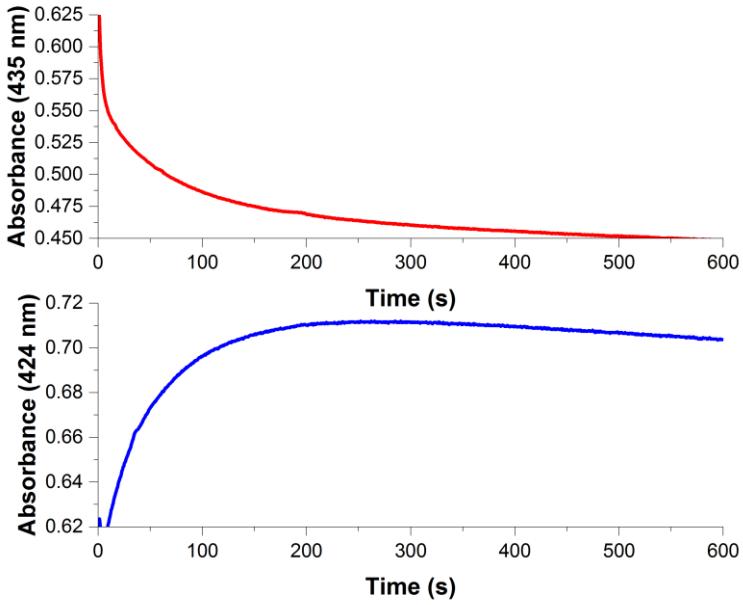


Figure 90: Single wavelength traces representing the time course absorbance of ferrous Mb(H64V,V68A) (417 nm, 7.5 µM protein, 10% EtOH) and the metallocarbenoid:Mb(H64V,V68A) complex (433 nm) following rapid mixing of ferrous Mb(H64V,V68A) with 500 µM ethyl diazoacetate at 5^{O}C in the presence of 3 mM styrene.

The data collected from the stopped-flow indicates that Mb(H64V,V68A) is catalytically active towards carbene transfer and returns to the initial active ferrous state after reacting with substrate. It should be noted that a 1:6 ratio of EDA:styrene was required to observe any spectroscopic activity, an identical ratio required for C45. In addition, the absence of any re-emergence of the ferrous Mb(H64V,V68A) peak (435 nm) indicates that Mb(H64V,V68A) possesses a much higher affinity for EDA relative to either C45 or *Rma*-TDE, which is consistent with the reported K_1 values for metallocarbenoid formation (**table 2**, chapter 3*, section 3.5.1.5*), but the much greater lag phase prior to the decline for the Mb(H64V,V68A)-EDA metallocarbenoid peak suggests that, while Mb(H64V,V68A) does exhibit carbene transfer activity towards styrene, Mb(H64V,V68A) has a much lower affinity for styrene relative to C45.

4.2.5 *Rma*-TDE and Mb(H64V,V68A) catalysed cyclopropanation assays – a comparison with C45

Having demonstrated carbene transferase activity for *Rma*-TDE and Mb(H64V,V68A) spectroscopically, both *Rma*-TDE and Mb(H64V,V68A) were subsequently employed as catalysts in the cyclopropanation assay to quantify reactions yields and stereoselectivities; the values were directly compared to the values acquired for C45. The same reagents and conditions used in the C45-catalysed cyclopropanation assays were employed for *Rma*-TDE and Mb(H64V,V68A). The assays were conducted in 1.5 mL seal-top reaction vials, which were purged with nitrogen prior to conducting the assays. The final reaction volumes were 400 µl, and consisted of 30 mM styrene, 10mM EDA, 10 mM sodium dithionite and 10 µM of *Rma*-TDE in CHES buffer (100 mM KCl, 20 mM CHES, pH 8.6) or 10 µM of Mb(H64V,V68A) in KPi buffer (100 mM phosphate, pH 7). The reactions were stirred for 2 hours, at room temperature, before being quenched (3M HCl) and the products extracted with ethyl acetate for analysis *via* chiral-HPLC and LC-MS.

The normalised chromatogram acquired for the *Rma*-TDE-catalysed cyclopropanation reaction between styrene and EDA vs the normalised chromatogram acquired for commercially acquired enantiomerically-pure (*R,R*) is presented in **figure 91 (top)**. **Figure 91 (bottom)** displays the average chromatogram for the *Rma*-TDE-catalysed cyclopropanation reaction between styrene and EDA, obtained from triplicate repeats.

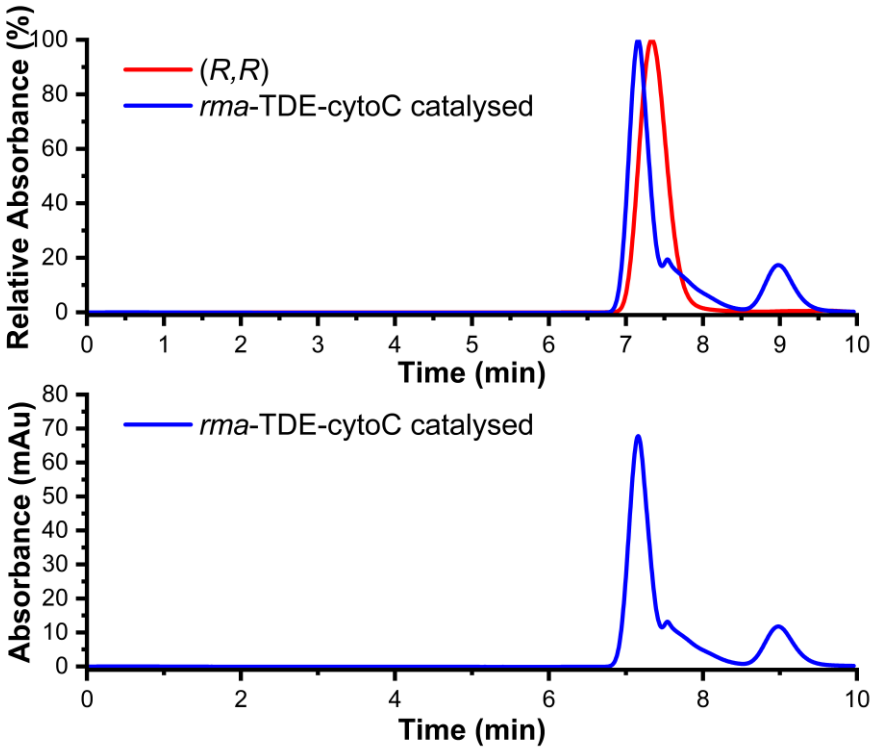


Figure 91: Chiral-HPLC chromatograms for the *Rma*-TDE-catalysed cyclopropanation assays. A polar organic mobile phase (100% MeCN: 0.1% (v/v) TFA:0.1% (v/v) Et_3N) was employed and injection volumes were 2 µl. (**top**) normalised chiral-HPLC chromatograms for commercial (*R,R*)-ethyl 2-phenylcyclopropane-1-carboxylate (280 nm, EtOH, red) and the *Rma*-TDE-catalysed cyclopropanation assays (280 nm, EtOH, blue). (**bottom**) Averaged chromatogram from the *Rma*-TDE (10 µM, 0.1 % catalyst loading) catalyzed cyclopropanation assay between styrene (30 mM) and EDA (10 mM) (CHES buffer, pH 8.6, 280 nm).

The averaged and normalised chromatograms in **figure 91** are both characterised by the presence of a dominant peak with a retention time of 7.1 minutes that is shortly followed by a minor peak with a retention time of 9.0 minutes. The normalised chromatogram **(figure 91, top)** demonstrates a good overlap between the *Rma*-TDE-catalysed cyclopropanation product and the commercially acquired enantiomerically-pure (*R,R*) sample, although there is a noticeable difference in the retention times. The molecular interactions between the stationary phase of the column and the eluting analyte are temperature dependent, and so small deviations in temperature are reflected in retention times;[310–312] because the chiral column is not thermally insulated, the observed shift for *Rma*-TDE-catalysed cyclopropanation product is accounted for by temperature fluctuations.[310–312] Nonetheless, the chromatograms indicate that the peak at 7.1 minutes can be assigned to ethyl (*1R,2R*)-2-phenylcyclopropane-1-carboxylate. From the averaged chromatogram in **figure 91 (bottom)**, an average reaction yield of 73% and enantioselectivity of 70.6% was calculated. In order to unambiguously confirm the identity of the 7.1 peak, the *Rma*-TDE-catalysed cyclopropanation product

was analysed by LC-MS (the acquired spectrum displayed in **figure 94 (bottom right)**). It can be seen that the spectrum possesses the dominant 145 m/z peak and the parent $[M-H]^+$ at 191 m/z, and excluding relevant intensities is practically identical to the spectrum obtained for the commercial sample.

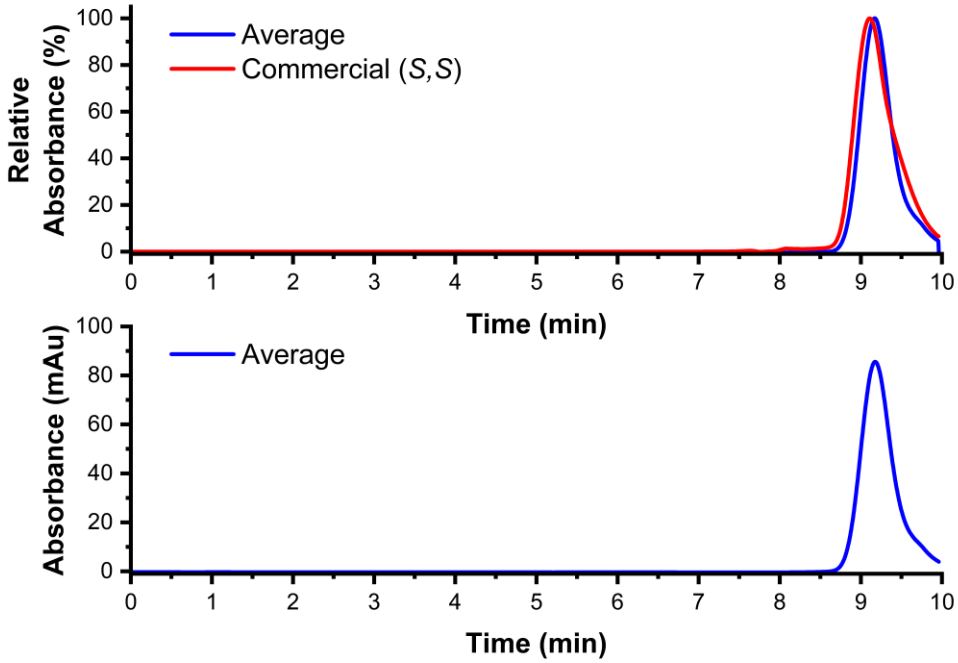


Figure 92: Chiral-HPLC chromatograms for the Mb(H64V,V68A)-catalysed cyclopropanation assays. A polar organic mobile phase (100% MeCN: 0.1% (v/v) TFA:0.1% (v/v) Et_3N) was employed and injection volumes were 2 µl. (**top**) normalised chiral-HPLC chromatograms for commercial (*S,S*)-ethyl 2-phenylcyclopropane-1-carboxylate (280 nm, EtOH, red) and the Mb(H64V,V68A)-catalysed cyclopropanation assays (280 nm, EtOH, blue). (**bottom**) Averaged chromatogram from the Mb(H64V,V68A) (10 µM, 0.1 % catalyst loading) catalyzed cyclopropanation assay between styrene (30 mM) and EDA (10 mM) (CHES buffer, pH 8.6, 280 nm).

The normalised chromatogram acquired for the Mb(H64V,V68A)-catalysed cyclopropanation reaction between styrene and EDA vs the normalised chromatogram acquired for commercially acquired enantiomerically-pure (*S,S*) is presented in **figure 92 (top)**. **Figure 92 (bottom)** displays the average chromatogram for the Mb(H64V,V68A)-catalysed cyclopropanation reaction between styrene and EDA, obtained from triplicate repeats. The averaged and normalised chromatograms in **figure 92** are both characterised by the presence of a single dominant peak with a retention time of 9.0 minutes. The normalised chromatogram (**figure 92, top**) demonstrates an excellent overlap between the Mb(H64V,V68A)-catalysed cyclopropanation product and the commercially obtained enantiomerically-pure (*S,S*)-sample (no additional peaks are present). The chromatograms indicate

the peak at 9.0 minutes can be assigned to ethyl (*1S,2S*)-2-phenylcyclopropane-1-carboxylate, the anticipated product. From the averaged chromatogram in **figure 92 (bottom)**, an average reaction yield of 96% and enantioselectivity of >99.9% was calculated, values identical to those reported in the literature.[195] In order to unambiguously confirm the identity of the peak at 9.0 minutes, the Mb(H64V,V68A)-catalysed cyclopropanation product was analysed by LC-MS (the acquired spectrum is displayed in **figure 94 (bottom left)**). It can be seen that the spectrum possesses the dominant 145 m/z peak and the parent $[M-H]^+$ at 191 m/z, and excluding relevant intensities is practically identical to the spectrum obtained for the commercial sample.

Figures 93 and **94** are comparative normalised chiral-HPLC chromatograms and LC-MS spectra for commercial ethyl (*R,R*)/(*S,S*)-2-phenylcyclopropane-1-carboxylate **(top left)**, the C45-cataysed cyclopropanation product **(top right)**, the *Rma*-TDE-catalysed cyclopropanation product **(bottom left)**, and the Mb(H64V,V68A)-catalysed cyclopropanation product **(bottom right)** respectively; the respective product yields, enantioselectivities, diastereoselectivities, and TTNs for each enzyme are presented in **table 5**.

Table 5: The yields, diastereoselectivities, enantioselectivities, and total turnover numbers for the cyclopropanation reaction between styrene and EDA catalysed by the *de novo* heme-containing maquette C45 and the rationally designed heme-containing proteins Mb(H64V,V68A) and *Rma*-TDE.

Enzyme	% Yield$_{(R.R)}$	% Yield$_{(S.S)}$	dr$_{(E)}$ %	ee$_{(R,R)}$ %	ee$_{(S,S)}$ %	TTN
C45	80.20 (± 8.01)	10.42 (± 1.81)	>99	77.00 (± 2.82)	-	802
Mb(H64V,V68A)	0	95.90 (± 0.51)	>99	-	>99.9	959
***Rma*-TDE**	73.47 (± 5.73)	12.64 (± 0.97)	>99	70.58 (± 2.02)	-	735

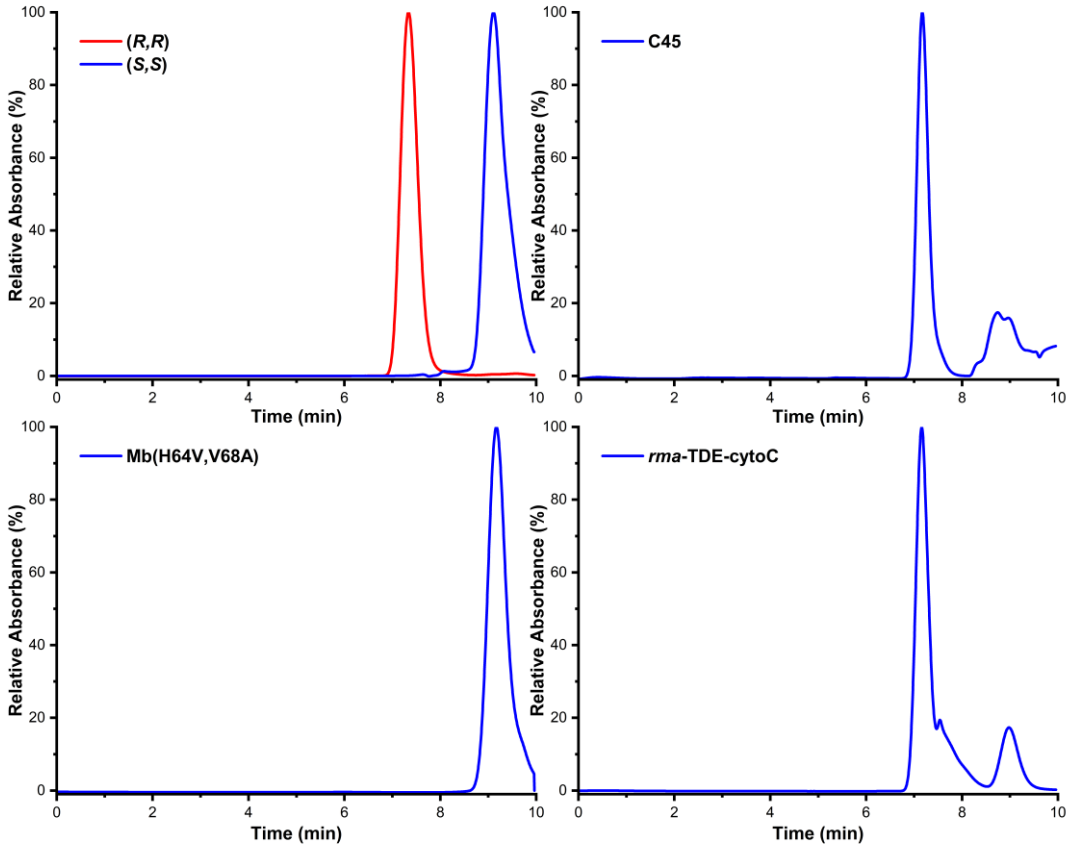


Figure 93: The normalised chiral-HPLC chromatograms for commercial ethyl 2-phenylcyclopropane-1-carboxylate, and the C45, Mb(H64V,V68A) and *Rma*-TDE catalyzed cyclopropanation assay products. (**top left**) commercial (*R,R*)-ethyl 2-phenylcyclopropane-1-carboxylate (in EtOH, red) and (*S,S*)-ethyl 2-phenylcyclopropane-1-carboxylate (in EtOH, blue) (**top right**) C45-catalyzed cyclopropanation assay between styrene (30 mM) and EDA (10 mM). (**bottom left**) Mb(H64V,V68A)-catalyzed cyclopropanation assay between styrene (30 mM) and EDA (10 mM). (**bottom right**) *Rma*-TDE-catalyzed cyclopropanation assay between styrene (30 mM) and EDA (10 mM). A polar organic mobile phase (100% MeCN: 0.1% (v/v) TFA:0.1% (v/v) Et_3N) was employed and injection volumes were 2 µl.

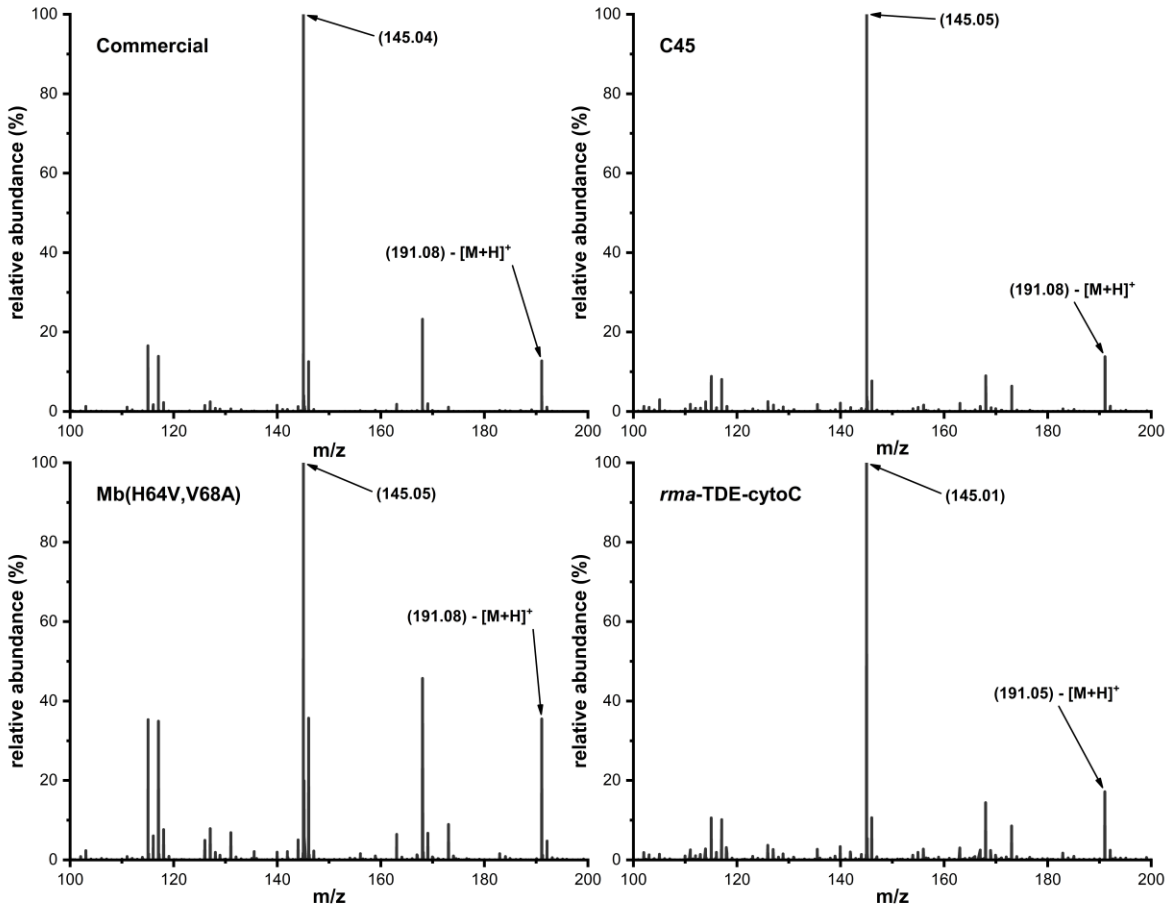


Figure 94: LC-MS spectra of commercial ethyl 2-phenylcyclopropane-1-carboxylate, C45, Mb(H64V,V68A) and *Rma*-TDE catalyzed cyclopropanation assay products. (**top left**) Commercial ethyl 2-phenylcyclopropane-1-carboxylate (in EtOH) exhibiting the dominant oxonium ion fragment at 145 m/z. (**top right**) C45-catalyzed cyclopropanation assay between styrene (30 mM) and EDA (10 mM). (**bottom left**) Mb(H64V,V68A)-catalyzed cyclopropanation assay between styrene (30 mM) and EDA (10 mM). (**bottom right**) *Rma*-TDE-catalyzed cyclopropanation assay between styrene (30 mM) and EDA (10 mM). All spectra were recorded in ES+ mode and monitored at 254 and 280 nm. A C8 column was employed for the LC separation with a gradient mobile phase (95:5:0.1 % v/v water/MeCN/formate 10:90:0.1 % v/v water/MeCN/formate). Injection volumes were 20 µl.

4.2.6 Kinetic characterisation of C45, *Rma*-TDE and Mb(H64V,V68A) catalysed cyclopropanation

The kinetic parameters for C45, *Rma*-TDE, and Mb(H64V,V68A) catalysed cyclopropanation assays between EDA and styrene were determined to gain insight into the differences between the three enzymes. The apparent kinetic parameters, V_{max}, K_1, k_{lim}, and the catalytic efficiencies of the cyclopropanation reaction for each enzyme were determined using stopped-flow spectroscopy. The concentration of EDA was kept constant through the experiments, with the concentration of styrene being varied; the apparent kinetic parameters were determined from the rate of change in absorbance of the Soret peak for the metallocarbenoid intermediate of each respective enzyme. **Figure 95** exhibits

the calculated curves acquired for each enzyme respectively, with each plot displaying concentration-dependent Michaelis-Menten like curves. The apparent kinetic parameters, V_{max}, K_1, k_{lim}, and (k_{lim}/K_1) are presented in **table 6**. The mechanism of cyclopropantion proceed *via* a heme-localised metallocarbenoid involves three distinct identities: i) free metallocarbenoid and substrate, ii) the formation of a three-membered cyclic intermediate, and iii) the elimination of the cyclopropane product to afford the starting ferrous enzyme. There are two underlining assumptions that have been made to assist in the analysis: i) the rate of decomposition of the three-membered cyclic intermediate to the product (k_{lim}) is first order, and ii) the reaction is irreversible (k_2 >>> k_{-2}). The overall process, with associated rate constants, is illustrated in **figure 95 (bottom)**.

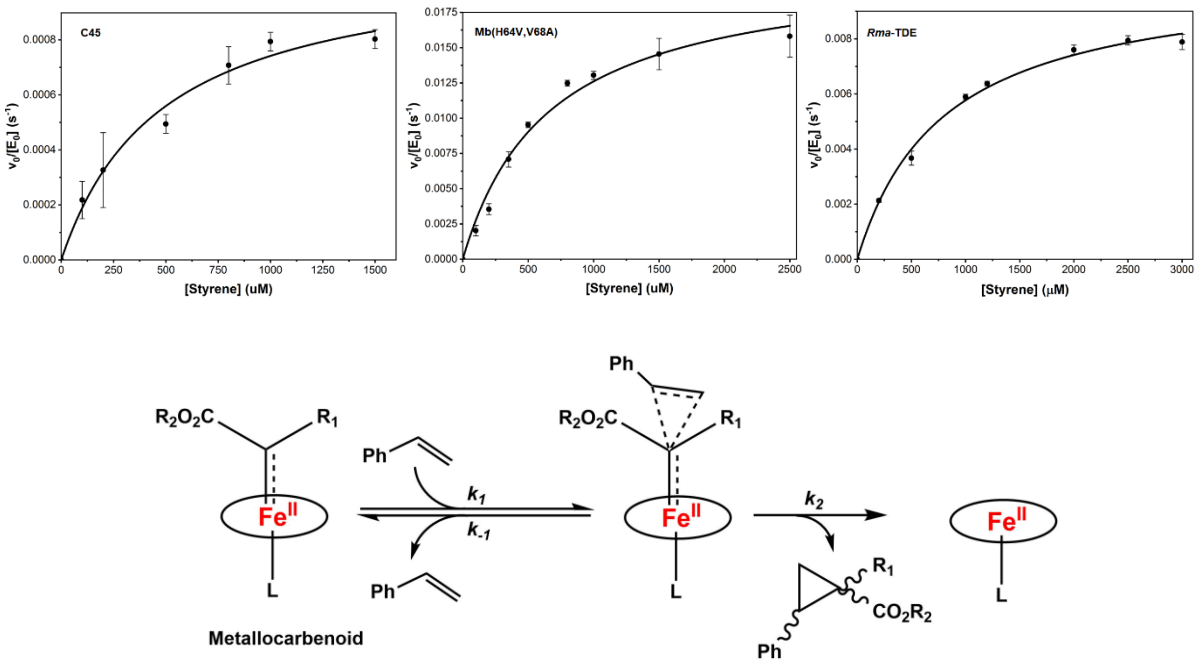


Figure 95: (top) The concentration-dependent plots for the cyclopropanation reaction of the enzyme:EDA metallocarbenoid complex in the presence of styrene for the enzymes (**left**) C45 (7.5 μM, CHES buffer, pH 8.6, 40% (v/v) EtOH); (**middle**) Mb(H64,V68A) (7.5 μM, 0.1M KPi buffer, pH 7.0, 40% (v/v) EtOH); and (**right**) *Rma*-TDE (7.5 μM, CHES buffer, pH 8.6, 40% (v/v) EtOH) Initial rates were determined from the rate of change in absorbance for the enzyme:EDA adduct at 5°C in the presence of styrene. **(bottom)** the mechanism of metallocarbenoid degradation in the presence of styrene showing the respective rate constants. A mechanism involving a tricyclic intermediate/transition state has been postulated and computational studied by Cossio,[313] Zhang,[291] and Fasan.[305]

The K_1 values represent the affinity of the enzyme for the binding of styrene substrate, with lower K_1 values reflecting a higher affinity. The calculated K_1(styrene) values for C45, *Rma*-TDE and Mb(H64V,V68A) are 482.51, 795.92, and 657.7 μM respectively (**table 6**). The K_1 values indicate C45 binds styrene with slightly higher affinity, and is possibility attributed to the simplicity of the maquette scaffold. *Rma*-TDE and Mb(H64V,V68A), although engineered towards carbene transferase activity,

are evolutionarily complex scaffolds harbouring ambiguous interdependency between the amino acids which could hinder the binding of small hydrophobic molecules, such as a styrene. In particular, *Rma*-TDE was engineered to accept bulky silane substrates and is characterised by a dynamic flexibility in the loop region where the V75T, M100D and M103E mutations were introduced. In the absence of any carbene precursor, the amino acid residues Asp-100, Thr-101 and Asp-102 were unresolved in the crystal structure on account of the increase loop flexibility; this suggests the active binding site is also solvent exposed. Additionally, the three mutations introduce a threonine, an aspartic acid, and a glutamic acid, which would introduce electrostatics into the active site of *Rma*-TDE; electrostatic interactions would favour silane binding but would prove detrimental for binding small, hydrophobic molecules.

In contrast, Mb(H64V,V68A) has five key residues in the active site: Leu-29, Phe-43, Val-64, Ala-68, and Ile-107. The leucine, isoleucine, and valine would provide a hydrophobic environment suitable for small molecule binding while the tryptophan, positioned in close proximity to the heme, could stabilise the binding of styrene through favourable π-π interactions. The authors reported the two mutations, H64V and V68A, creates a larger void above the solvent-exposed region of the heme cofactor, facilitating substrate accessibility. In C45, the introduction of three phenylalanine rings provides a stable, hydrophobic core suitable for accommodating small hydrophobic molecules. A tryptophan residue also residues in close proximity to the heme active site, which could provide a stabilising π-π interaction with styrene. The presence of several hydrophobic and aromatic residues in the core of C45 and Mb(H64,V68A) probably accounts for the higher binding affinity of styrene relative to *Rma-TDE*. In addition to active site interactions, the dynamics of the protein scaffold and the degree of solvent exposed most likely play critical roles in determining the K_1(styrene) values.

Table 6: The kinetic parameters for the cyclopropanation reaction between styrene and EDA catalysed by the *de novo* heme-containing maquette C45 and the rationally designed heme-containing proteins Mb(H64V,V68A) and *Rma*-TDE.

Styrene	C45	Mb(H64V,V68A)	*Rma*-TDE
k_{lim} (s^{-1})	0.0011 (± $9x10^{-5}$)	0.021 (± 0.002)	0.0104 (± 0.0004)
V_{max} (μM s^{-1})	0.0083	0.158	0.0078
K_1 (μM)	482.51 (± 108.8)	657.7 (± 110.3)	795.92 (± 82.7)
k_{lim}/K_1 (s^{-1} M^{-1})	2.23	31.93	1.37
$ln(2)/k_{lim}$	10.5 min	33 s	11 min

The k_{lim} values for C45, *Rma*-TDE, and Mb(H64V,V68A) are 0.0011, 0.0104 and 0.021 s^{-1} respectively, which represent the relative rate of product formation following an interaction between styrene and the metallocarbenoid intermediate (**table 6**). The catalytic efficiencies (k_{lim}/K_1) are therefore 2.23, 1.37 and 31.93 μM s^{-1} for C45, *Rma*-TDE and Mb(H64V,V68A) respectively. The data suggests Mb(H64V,V68A) is the most efficient enzyme for cyclopropanation, with a k_{lim} 2-fold higher than *Rma*-TDE and 20-fold higher than C45, respectively. This observation is unsurprising considering Mb(H64V,V68A) was selectively engineered for the cyclopropanation reaction, unlike C45 and *Rma*-TDE. *Rma*-TDE was engineered for Si-H insertion chemistry, with the authors introducing key mutations into the distal loop region to facilitate silane binding and to provide stabilisation for the transition state. C45 was engineered as a simplified oxidoreductase, with an emphasis on oxygen binding, electron transfer, and peroxidase chemistry. The similarities between *Rma*-TDE and C45, such as dynamic flexibility, solvent exposed heme propionates, and a c-type heme cofactor, render the similarities in catalytic efficiency unsurprising, but the kinetic parameters suggest carbene-activated C45 binds styrene more efficiently whereas carbene-activated *Rma*-TDE turns over styrene more efficiently.

At present, it is impossible to infer any conclusions concerning how the active site residues of each enzyme stabilise the transition state or the molecular/dynamic origin of the apparent kinetic parameters. A comparison between the three enzymes k_{lim} is rendered particularly impossible based on the lack of structural information for C45 with respect to metallocarbenoid binding, substrate binding, transition state stabilisation, and substrate turnover. It should also be noted the kinetic parameters discussed above are only apparent kinetic parameters and may not accurately represent the kinetics of the process. This is because the values are derived from the change in absorbance of the metallocarbenoid intermediate Soret peak. As the product does not absorb in the visible region, no new signal can be detected in the electronic spectra which can be followed to acquire the true kinetic parameters. This could be addressed in one of two ways. The first approach would be to develop a colourimetric assay that would allow the emergence of the product and/or disappearance of the substrate to be followed continuously. A rudimentary colourimetric assay was developed by employing indigo as the starting material. Indigo is a fully conjugated aromatic compound with the two fragments electronically connected *via* a C=C bridge. It was rationalised that disrupting that C=C π-bond, *via* the addition of a carbene, would destroy the conjugation and yield a product with no detectable trace in the electronic spectra. The employment of indigo as a substrate, however, proved more difficult than initially expected, and the assay was abandoned. A second approach would be to follow the progress of the reaction discontinuously using HPLC. Although a good idea in theory, discontinuous kinetics are time-consuming, laborious, and subject to human error which would make

comparisons between different enzymes problematic. For the purposes of acquiring a rough, rudimentary understanding of the cyclopropanation reaction catalysed by different enzymes, it was rationalised the approach outlined above (using stopped flow spectroscopy) was sufficient, providing the limitations of the approach are always remembered.

4.3 CONCLUSION

Having demonstrated in chapter 3 that C45 can form stable metallocarbenoid intermediates with three distinctive diazo precursors the question of whether the intermediate was reactive and could be employed in organic transformations was raised. The most established reaction in carbene chemistry is the [2+1] cycloaddition reaction between a reactive carbene and an olefin to yield a cyclopropane-containing product. In the presence of styrene, the spectroscopic profile observed for the C45:EDA metallocarbenoid complex exhibited a significant decay which was concurrent with the re-emergence of ferrous C45. It was subsequently demonstrated, *via* chiral-HPLC and LC-MS, that C45 catalyses the cyclopropanation reaction between EDA and styrene with good product yields (80%), complete diastereoselectivity (99.9%), and good enantioselectivity (77%) in favour of the (*R,R*) product. The reaction was extended to the cyclopropanation reactions between EDA and a variety of *para*-functionalised styrene substrates and between styrene and two additional diazo substrates. Although the reactions mostly afforded lower reaction yields relative to EDA/styrene, the enantioselectivities of the reactions remained good, with C45 catalysing the cyclopropanation reaction between EDA and *p*-methoxystyrene, *p*-fluorostyrene, and *p*-trifluoromethylstyrene with enantioselectivities surpassing 90%. The carbene transferase activity of C45 was also compared against two carbene transferases reported in the literature: *Rma*-TDE and Mb(H64V,V68A). C45 exhibited a similar reactivity to *Rma*-TDE (80% for C45 vs 73.5% for *Rma*-TDE) although which inverted stereoselectivity (77% in favour of the (*R,R*)-product for C45 vs 71% in favour of the (*S,S*)-product for *Rma*-TDE); C45 was, however, inferior to Mb(H64V/V68A) with respect to the cyclopropanation reaction between styrene and EDA (96% product yield and 99.9% enantioselectivity in favour of the (*R,R*)-product). Nonetheless, it has been demonstrated that C45 can form a stable heme-localised and catalytically active metallocarbenoid intermediate capable of reacting with olefins. C45 is a proficient carbene transferase, catalysing cyclopropanation reactions *via* an abiological mechanism (carbene transfer) with numbers equivalent to numbers reported for rationally designed carbene transferases reported in the literature.

CHAPTER 5: SUBSTRATE PROMISCUITY OF A *DE NOVO* CARBENE TRANSFERASE

5.1 SUBSTRATE PROMISCUITY: X-H INSERTION, OLEFIN CARBONYLATION AND RING EXPANSION CHEMISTRY

In chapter 1, *section 1.4.2*, the diverse chemistry of carbene reagents was introduced and included X-H insertion reactions (X = C, N, B, S, Si).[37,59,60,64–66] In organic synthesis, carbene X-H insertions are a versatile and important class of reactions often utilised in the synthesis of novel X-C σ-bonds. For example, N-H insertions allow complex synthetic pathways to be divided into multiple paths, allowing for a non-linear synthesis of complex molecules which can improve overall product yields. In addition to cyclopropanation reactions, several carbene transferases, designed from natural heme-containing protein scaffolds, have been reported in the literature as efficient catalysts in C-H, N-H, B-H, S-H, and Si-H insertions,[17,210,215–217,220,221,288] as was discussed in detail in chapter 2 *section 2.3.3*. Although the reaction mechanism for a carbene transfer X-H insertion has key distinctive differences to the concerted olefin insertions,[39,61] both mechanisms are postulated to proceed *via* an electrophilic metallocarbenoid which is susceptible to nucleophilic attack (e.g. an olefin in cyclopropanation, an amine in N-H insertion). Homologation reactions were also introduced in chapter 1, *section 1.4.3*, as reactions that proceed *via* carbene reagents. As of yet, no reported carbene transferase in the literature has been reported to perform a homologation reaction. Carbonyl olefination reactions, proceeding *via* a metallocarbenoid-phosphonium ylide intermediate, have also been reported in two cases.[229,230]

Enzymes are renowned for exhibiting remarkable selectivity and substrate specificity. In biocatalysis, substrate specificity can be unfavourable because it demands individual biocatalysts be engineered for individual reactions. Enzyme promiscuity, the ability of an enzyme to catalyse multiple reactions, is therefore highly desirable from a synthetic perspective. In light of the activity of C45 towards cyclopropanation chemistry, the question of whether C45 could be employed as a carbene transferase in either X-H insertion, carbonyl olefination, or homologation reactions was raised. It was postulated the lack of any inherent specific substrate recognition sites (SRSs) (unlike the evolved SRSs in P450[314] and myoglobin scaffolds), structural heterogeneity, and flexibility of C45 would favour enzymatic promiscuity. For the X-H insertion, the N-H insertion reaction was initially selected for screening on account of the versatility and ubiquitous distribution of N-H and N-C σ-bonds in synthesis and biological chemistry.[315] The carbonyl olefination reaction was explored using aromatic aldehyde substrates, as was reported in the literature,[229,230] in an attempt to synthesise α,β-unsaturated carbonyl compounds. The homologation reaction was explored using a halogenated α-diazo precursor and pyrrole, a nitrogen-containing heteroaromatic (NHA), in an attempt to synthesise ethyl nicotinate.

If successful, C45 would be the first ever enzyme, rationally designed or *de novo*, to catalyse a homologous ring expansion reaction.

5.2 RESULTS AND DISCUSSION

The C45-catalysed N-H insertion assays were conducted in an analogous way to the cyclopropanation assays, with the notable substitution of the olefin substrate for a suitable amine (**figure 96**). The two amines, piperidine, a secondary amine, and *p*-chloroaniline, a primary amine, were identified as suitable substrates to investigate N-H insertion activity. It is well established that secondary amines, the product of a mono-insertion reaction of a primary amine, can further react with another substrate to yield tertiary amines and ammonium salts. The employment of a primary amine, *p*-chloroaniline, therefore provides additional insight into the specificity of the enzyme towards amine functionalisation. The three diazo compounds employed in the cyclopropanation assays (EDA, BnDA and ^{t}BuDA) were also used in the N-H insertion assays, resulting in a total of six individual reactions to be screened. The reaction conditions employed were identical for the cyclopropanation assays. The reactions were conducted in 1.5 ml sealed, screw-top vials and were purged with nitrogen. The final concentrations of each constituent were 10 mM diazo precursor, 30 mM amine, 10 mM sodium dithionite, and 10 µM C45. The final volume for each reaction was 400 µl, and all reactions were performed in CHES buffer (20 mM CHES, 10 mM KCl, pH 8.6). The reaction mixture was quenched after 2 hours, and the product mixture was subsequently extracted and analysed *via* HPLC (using a chiral-HPLC column) and LC-MS. All reactions were performed in triplicate and the isolated products from each reaction were compared against commercially prepared or synthetically produced samples of the anticipated products.

5.2.1 X-H insertion chemistry

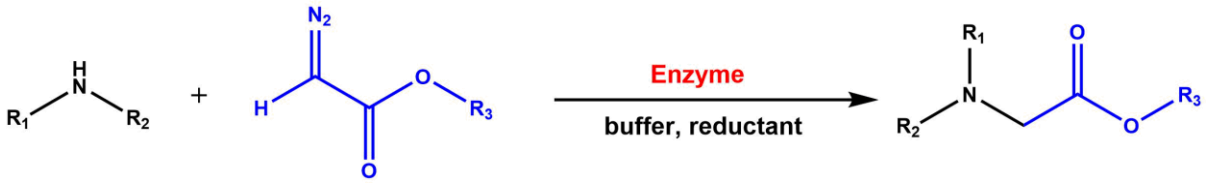


Figure 96: The general reaction scheme for an enzyme-catalysed N-H insertion reaction between a primary/secondary amine and a diazo compound.

5.2.1.1 N-H insertion reaction in the presence and absence of hemin

Initial assays were conducted in i) the absence of any catalyst, and ii) the presence of free hemin (10 µM) for both amine substrates. The experimental conditions were identical to the conditions

employed in the assays using C45. The reactions screened with no catalyst/hemin were the reactions between EDA and i) *p*-chloroaniline and ii) piperidine. The reaction mixtures were extracted with 1.25 ml of n-hexane prior to analysis. The reaction mixtures were analysed *via* HPLC and the resultant chromatograms for all four reactions are presented in **figure 97**. The chromatograms for all four N-H insertion assays conducted with and without hemin present are all dominated by a single major peak with a retention time of 2.1 minutes which corresponds to the starting material EDA; no trace of any reaction product (or starting material) was detectable in any of the assays. The product mixtures were subsequently analysed *via* LC-MS and were also characterised by the lack of any detectable trace of the anticipated products. These assays indicate that no reaction occurs between EDA and piperidine/*p*-chloroaniline in the presence and absence of hemin.

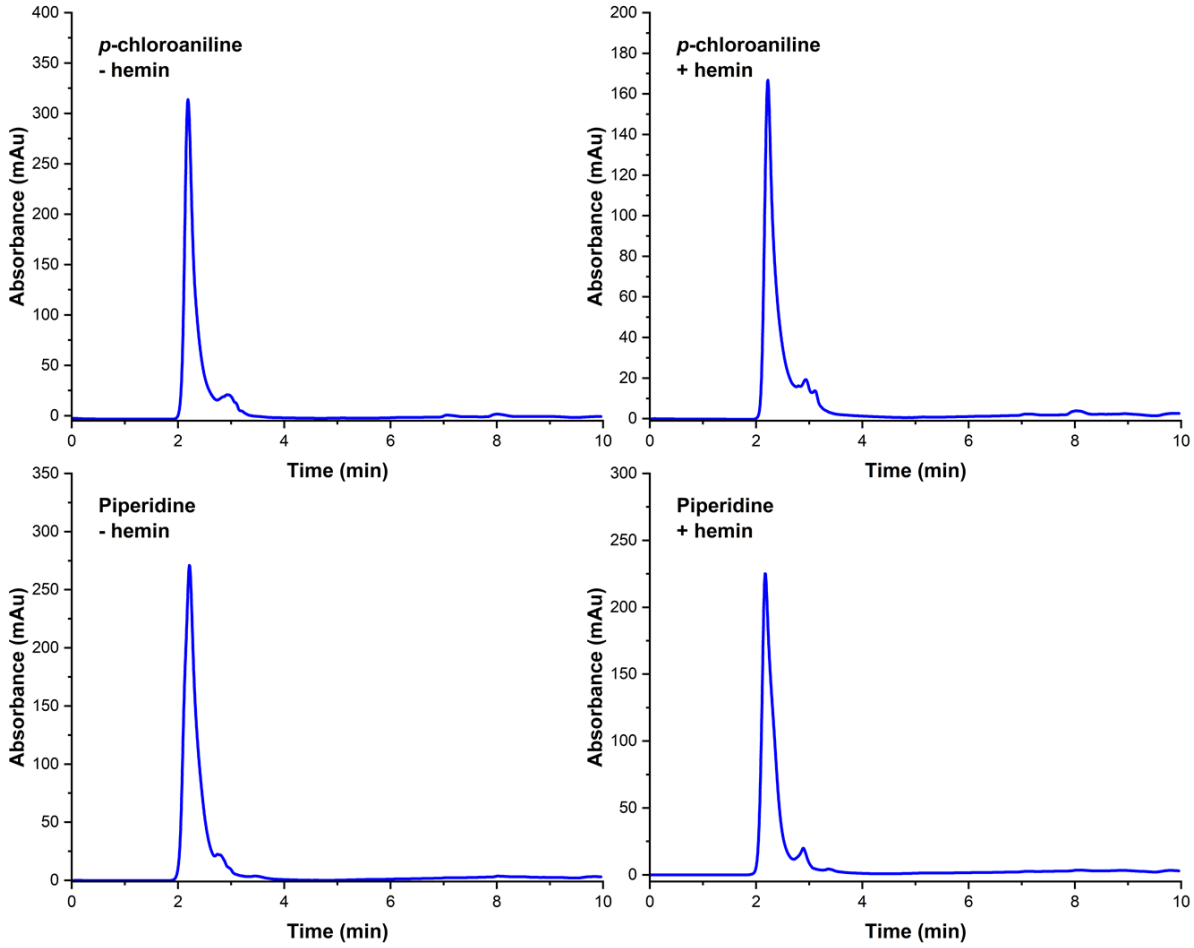


Figure 97: Chiral-HPLC chromatograms for the N-H insertion assays between *p*-chloroaniline/piperidine (30 mM) and EDA (10 mM) (CHES buffer, pH 8.6, 254 nm) in the absence of any catalyst and in the presence of hemin (10 µM, 0.1 % catalyst loading). A polar organic mobile phase (100% MeCN: 0.1% (v/v) TFA:0.1% (v/v) Et_3N) was employed and injection volumes were 2 µl.

5.2.1.2 C45-catalysed N-H insertions: piperidine

The anticipated product from the N-H insertion reaction between EDA and piperidine is ethyl-1-piperidineacetate. The HPLC trace of a commercial sample of ethyl-1-piperidineacetate is dominated by a major peak in the retention time of 3.2 minutes (**figure 98**). The LC-MS spectrum of ethyl-1-piperidineacetate is presented in **figure 99** and is dominated by a major peak at m/z 98, corresponding to the fragmentation of the parent ion to an iminium species, a peak at m/z 144 corresponding to an iminium-acetate complex, and a peak at m/z 172, corresponding to the parent $[M+H]^+$ peak. In addition, the chromatogram for the starting material, piperidine, was recorded (t_R = 4.6 minutes). The external calibrations for ethyl-1-piperidineacetate, *tert*-butyl 2-(piperidin-1-yl)acetate and benzyl 2-(piperidin-1-yl)acetate, used to calculate reaction yields, are provided in **figure 165** (*appendix 1*).

The average chiral-HPLC chromatograms acquired for the C45-catalysed N-H insertion reactions between EDA, BnDA, and ^{t}BuDA and piperidine are presented in **figure 98 (bottom)**. In addition, the normalised average chromatograms acquired for each reaction vs the commercial sample of ethyl-1-piperidineaceate are also presented in **figure 98 (top)**. The average chromatogram for the C45-catalysed N-H insertion reaction between EDA and piperidine is characterised by a retention time of approximately 3.1-3.2 minutes, resulting in a very good overlap between the reaction chromatogram and the commercial product. The average chromatogram for the C45-catalysed N-H insertion reaction between ^{t}BuDA and piperidine is characterised by a slight deviation with respect to the commercial product, with a retention time of approximately 3.4-3.5 minutes relative to 3.2 minutes respectively; a deviation that is expected considering the presence of the bulkier *tert*-butyl substituent. The average chromatogram for the C45-catalysed N-H insertion reaction between BnDA and piperidine is, analogous to EDA/piperidine, characterised by a very good overlap with respect to the commercial product, with a retention time of approximately 3.2 minutes relative to 3.2 minutes for the commercial sample. From the averaged chromatograms the average yields (accounting for dilutions) for each assay (**table 18**, *appendix 2*) were calculated to be 78.90, 62.10, and 88.34% for EDA, ^{t}BuDA, and BnDA, respectively.

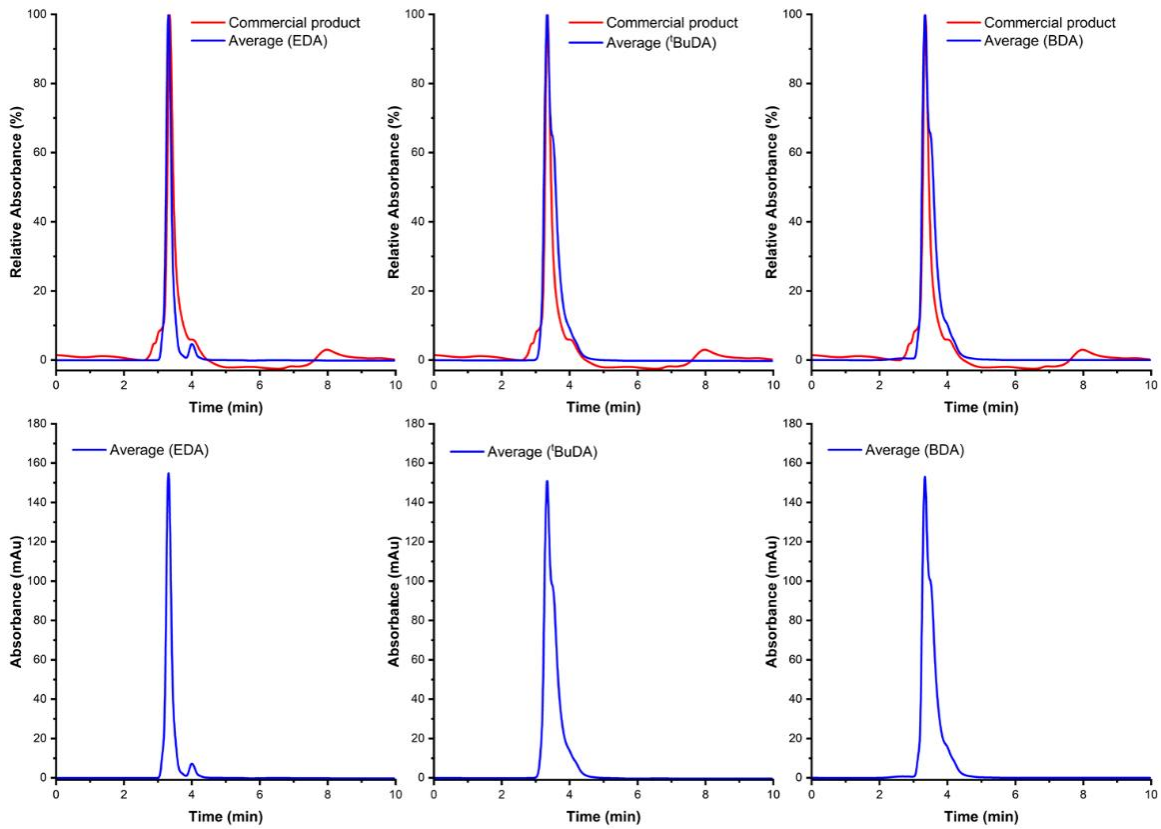


Figure 98: Chiral-HPLC chromatograms for the C45-catalysed (10 µM, 0.1 % catalyst loading) N-H insertion assays between piperidine (30 mM) and EDA (10 mM), ^{t}BuDA (10 mM), and BnDA (10 mM) in (CHES buffer, pH 8.6, 254 nm). (**top row**) The normalised chiral-HPLC chromatograms for the C45-catalysed N-H insertion assays between piperidine and EDA (**top left,** blue), ^{t}BuDA (**top middle,** blue) and BnDA (**top right,** blue) vs the normalised chromatogram of commercial ethyl-1-piperidineacetate (red). (**bottom row**) The average chiral-HPLC chromatograms for the C45-catalysed N-H insertion assays between piperidine and EDA (**bottom left**), ^{t}BuDA (**bottom middle**) and BnDA (**bottom right**). A polar organic mobile phase (100% MeCN: 0.1% (v/v) TFA:0.1% (v/v) Et_3N) was employed and injection volumes were 2 µl.

The LC-MS spectrum acquired for the N-H insertion assay between i) EDA, ii) tBuDA, and iii) BnDA and piperidine are all presented in **figure 99** alongside the LC-MS spectrum acquired for a commercial sample of ethyl-1-piperidineacetate. The mass spectrum for ethyl-1-piperidineacetate presents a fragmentation pattern with dominant peaks at 98, 144 and 172 m/z. The C45-catalysed N-H insertion reaction between EDA and piperidine is dominated by major peaks at 98, 144 and 172 m/z, presenting an identical fragmentation profile to the spectrum acquired for ethyl-1-piperidineacetate. The LC-MS spectrum acquired for the C45-catalysed N-H insertion reaction between ^{t}BuDA and piperidine is characterised by a similar fragmentation profile recorded for the commercial sample, with dominant peaks at 98, 144 m/z, corresponding to the iminium fragment, and the iminium-acetate complex respectively. No peak for the parent ion could be detected. The LC-MS spectrum acquired for the C45-

catalysed N-H insertion reaction between BnDA and piperidine is distinctively different from the EDA/^{t}BuDA-pieridine mass spectrum, and is characterised by a dominant peak at 91 m/z. This peak corresponds to the mass of a tropylium fragment and is produced *via* fragmentation of the benzyl ester. A peak at 234 m/z was detected though, which corresponds to the expected mass of the $[M+H]^{+}$ parent ion. The emergence of a tropylium ion, at 91 m/z, in the mass spectrum for the C45-catalysed N-H insertion reaction between BnDA and piperidine parallels the fragmentation pattern observed for the C45-catalysed N-H insertion reaction between BnDA and *p*-chloroaniline.

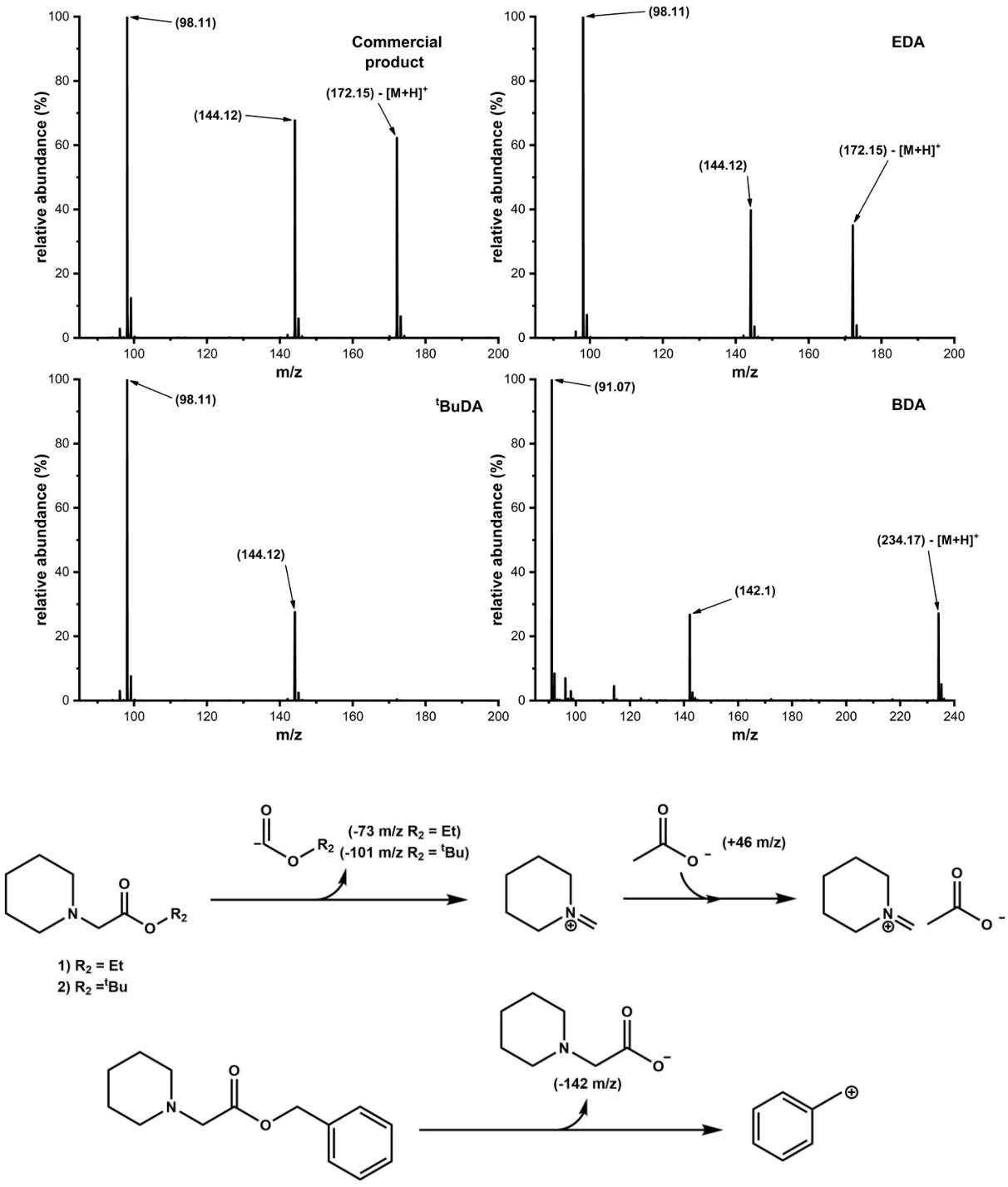


Figure 99: (**top**) LC-MS mass spectra for commercial ethyl-1-piperidineacetate (**top left)** and the C45-catalysed (10 µM, 0.1 % catalyst loading) N-H insertion assays between piperidine (30 mM) and (**top left**) EDA (10 mM), (**bottom left**) ^{t}BuDA (10 mM), (**bottom right**) and BnDA (10 mM) in (CHES buffer, pH 8.6, 254 nm). All spectra were recorded in ES+ mode and monitored at 254 and 280 nm. A C8 column was employed for the LC separation with a gradient mobile phase (95:5:0.1 % v/v water/MeCN/formate 10:90:0.1 % v/v water/MeCN/formate). Injection volumes were 20 µl. (**bottom**) Fragmentation pathway leading to the dominant iminium (98/144) and tropylium (91) ions observed in the mass spectra.

5.2.1.3 C45-catalysed N-H insertions: *p*-chloroaniline

The anticipated product from the N-H insertion reaction between EDA and aniline is *N*-phenylglycine ethyl ester. The HPLC trace of a commercial sample of *N*-phenylglycine ethyl ester is dominated by a major peak in the retention time of range 3.2-3.5 minutes (**figure 100**). The LC-MS spectrum of *N*-phenylglycine ethyl ester is presented in **figure 101** and is dominated by a major peak at 106.8 m/z, corresponding to the fragmentation of the parent ion to an iminium species. The LC-MS spectrum of ethyl (4-chlorophenyl)glycinate is presented in **figure 101** and is dominated by a major peak at 140 m/z, corresponding to the fragmentation of the parent ion to an iminium species with the addition of a chlorine. The peak 140 m/z is followed by a peak at 142 m/z with an approximate ratio of 3:1 in intensity for the two peaks on account of the chlorine. The external calibrations ethyl (4-chlorophenyl)glycinate, *tert*-butyl (4-chlorophenyl)glycinate, and benzyl (4-chlorophenyl)glycinate, used to calculate reaction yields, are provided in **figure 166** (*appendix 1*).

The chiral-HPLC chromatograms acquired for the C45-catalysed N-H insertion reactions between EDA, BnDA, and ^{t}BuDA and *p*-chloroaniline are presented in **figure 100 (bottom)**. In addition, the normalised average chromatograms acquired for each reaction vs a commercial sample of *N*-phenylglycine ethyl ester are also presented in **figure 100 (top)**. The average chromatogram for the C45-catalysed N-H insertion reaction between EDA and *p*-chloroaniline is characterised by a slight deviation of the reaction mixture peak relative to the commercial product, with a retention time of approximately 3.7 minutes relative to 3.2-3.5 minutes. This deviation is accounted for by the chlorine atom in the reaction mixture product, which is absent in the commercial sample. In addition to the single insertion product, a small amount of the double insertion product could be detected. The average chromatogram for the C45-catalysed N-H insertion reaction between ^{t}BuDA and *p*-chloroaniline is also characterised by a very good overlap with respect to the commercial product, with a retention time of approximately 3.2-3.5 minutes relative to 3.2-3.5 minutes. Unlike the assay conducted with EDA, only the single insertion product was detected for ^{t}BuDA. The average chromatogram for the C45-catalysed N-H insertion reaction between BnDA and *p*-chloroaniline is, analogous to ^{t}BuDA/*p*-chloroaniline, characterised by a very good overlap with respect to the commercial product, with a retention time of approximately 3.2-3.5 minutes relative to 3.2-3.5 minutes for the commercial sample. Unlike the assay conducted with EDA, only the single insertion product was detected for BnDA. From the averaged chromatograms (and accounting for dilutions), the average yields for each assay (**table 16**, *appendix 2*) were calculated to be 67.59, 72.50, and 89.44% for EDA, ^{t}BuDA, and BnDA, respectively. For the assay conducted with EDA, an average yield for the double product was 5.29%, resulting in a single/double insertion ratio of 12.7:1.

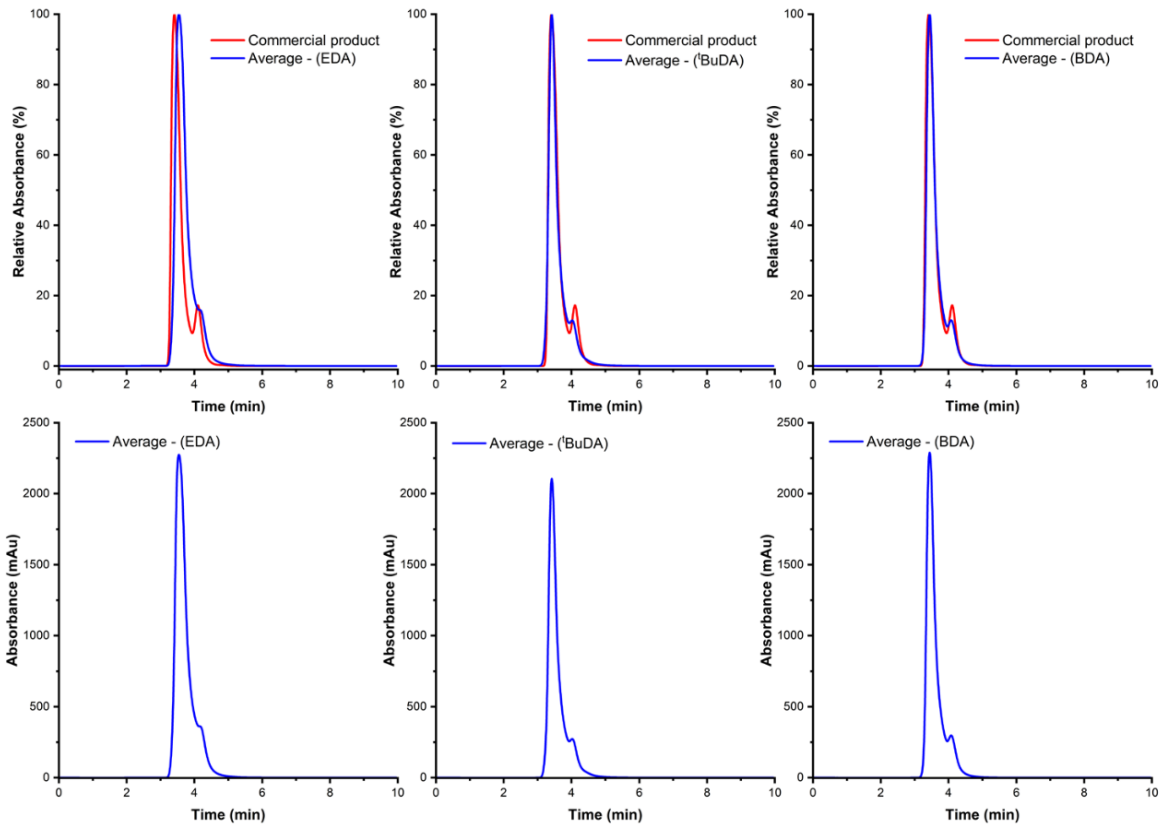


Figure 100: Chiral-HPLC chromatograms for the C45-catalysed (10 µM, 0.1 % catalyst loading) N-H insertion assays between *p*-chloroaniline (30 mM) and EDA (10 mM), ^{t}BuDA (10 mM), and BnDA (10 mM) in (CHES buffer, pH 8.6, 254 nm). (**top row**) The normalised chiral-HPLC chromatograms for the C45-catalysed N-H insertion assays between *p*-chloroaniline and EDA (**top left,** blue), ^{t}BuDA (**top middle,** blue) and BnDA (**top right,** blue) vs the normalised chromatogram of commercial ethyl phenylglycine ester (red). (**bottom row**) The average chiral-HPLC chromatograms for the C45-catalysed N-H insertion assays between *p*-chloroaniline and EDA (**bottom left**), ^{t}BuDA (**bottom middle**) and BnDA (**bottom right**). A polar organic mobile phase (100% MeCN: 0.1% (v/v) TFA:0.1% (v/v) Et_3N) was employed and injection volumes were 2 µl.

After analysis *via* chiral-HPLC, the product mixtures were investigated using LC-MS. The LC-MS spectrum acquired for the N-H insertion assay between i) EDA, ii) ^{t}BuDA, and iii) BnDA and *p*-chloroaniline are all presented in **figure 101** alongside an LC-MS spectrum acquired for commercially prepared *N*-phenylglycine ethyl ester. The mass spectrum for the C45-catalysed N-H insertion reaction between EDA and *p*-chloroaniline is characterised by a dominant peak at 140 m/z corresponding to the mass of the iminium fragment observed in the commercial sample (+ chlorine). The presence of a 1:0.33 ratio between 140:142 m/z corroborates the HPLC chromatogram, indicating a N-H insertion reaction has taken place. In addition, a peak at 214 m/z corresponds to the expected $[M+H]^+$ of the anticipated product of an N-H insertion between EDA and *p*-chloroaniline. A closer inspection of the parent ion reveals the same 1:0.33 ratio between 214/216 m/z, the expected isotopic distribution for

the parent ion mass. The LC-MS spectrum acquired for the C45-catalysed N-H insertion reaction between ^{t}BuDA and *p*-chloroaniline is characterised by a dominant peak at 140 m/z corresponding to the mass of the iminium fragment observed in the commercial sample (+ chlorine) and in the EDA/piperidine mass spectrum. The presence of a 1:0.33 ratio between 140:142 m/z corroborates the HPLC chromatogram, indicating a N-H insertion reaction has taken place. In addition, a peak at 242 m/z corresponds to the expected $[M+H]^{+}$ of the anticipated product of an N-H insertion between ^{t}BuDA and *p*-chloroaniline. A closer inspection of the parent ion reveals the same 1:0.33 ratio between 242/244 m/z, the expected isotopic distribution for the parent ion mass. The LC-MS spectrum acquired for the C45-catalysed N-H insertion reaction between BnDA and *p*-chloroaniline is distinctively different from the EDA-^{t}BuDA/*p*-chloroaniline mass spectrum, and is characterised by a dominant peak at 91 m/z. This peak corresponds to the mass of a tropylium fragment and is produced *via* fragmentation of the benzyl ester. The characteristic 140:142 m/z peaks, characterised by a 1:0.33 isotropic ratio, are also observable in the BnDA/*p*-chloroaniline spectrum, but at a significantly reduced intensity relative to the EDA-^{t}BuDA/*p*-chloroaniline mass spectra. The acquired data indicates that a reaction has occurred between BnDA and *p*-chloroaniline.

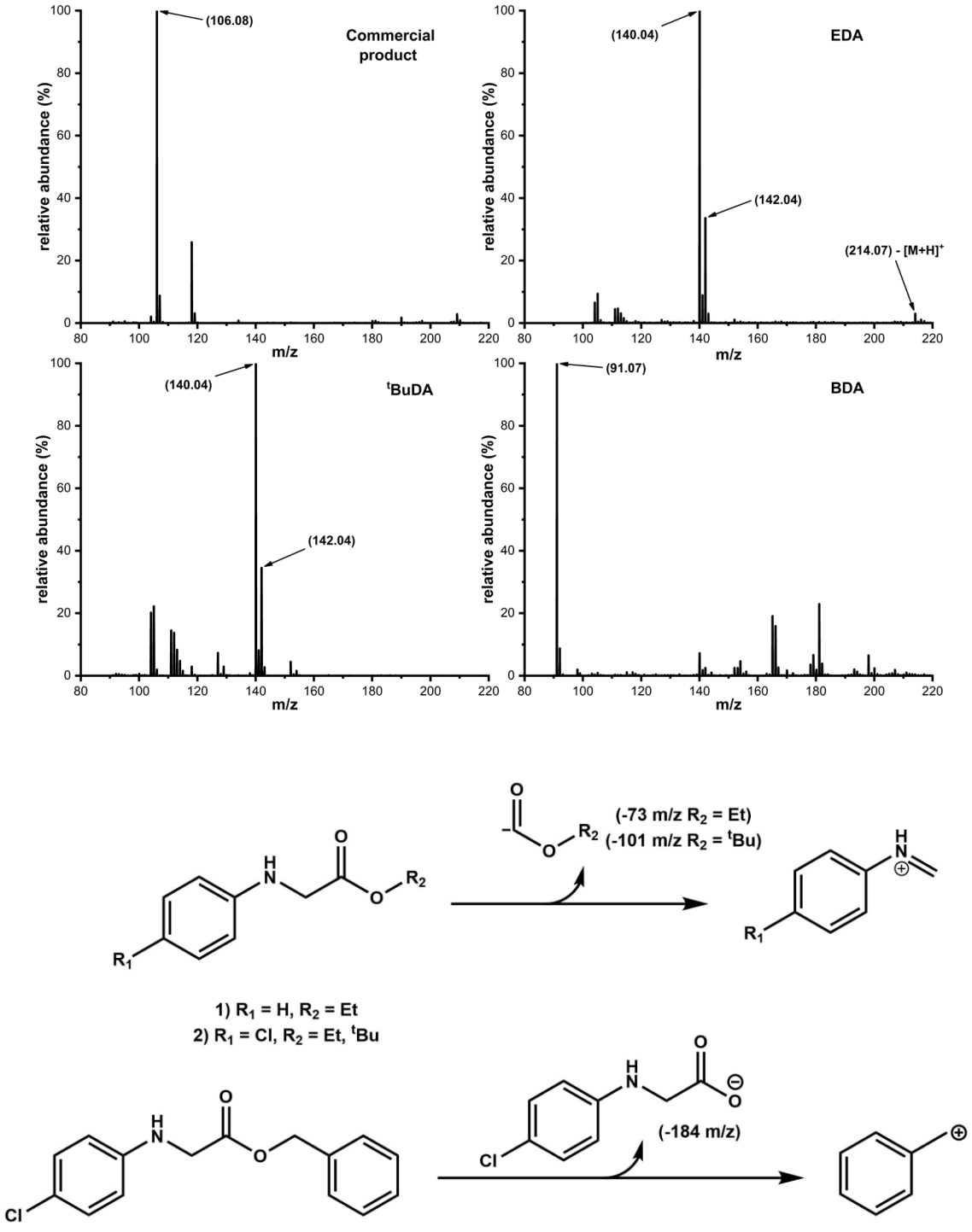


Figure 101: (**top**) LC-MS mass spectra for commercial ethyl phenylglycine ester (**top left)** and the C45-catalysed (10 µM, 0.1 % catalyst loading) N-H insertion assays between *p*-chloroaniline (30 mM) and (**top left**) EDA (10 mM), (**bottom left**) ^{t}BuDA (10 mM), (**bottom right**) and BnDA (10 mM) in (CHES buffer, pH 8.6, 254 nm). All spectra were recorded in ES+ mode and monitored at 254 and 280 nm. A C8 column was employed for the LC separation with a gradient mobile phase (95:5:0.1 % v/v water/MeCN/formate 10:90:0.1 % v/v water/MeCN/formate). Injection volumes were 20 µl. (**bottom**) Fragmentation pathway leading to the dominant iminium (104/140) and tropylium (91) ions observed in the mass spectra.

5.2.1.4 C45-catalysed chemoselective N-H vs O-H insertion: aminophenol regioisomers

Large, complex organic molecules often contain more than one type of functional group. The most common heteroatoms in organic molecules are nitrogen and oxygen, with the functional groups containing nitrogen or oxygen typically being structurally and functionally similar. For example, two extremely common functional groups are the hydroxy (-OH) and primary amine (-NH_2) groups, both of which are characterised by a polarised C-X σ-bond in favour of the heteroatom. These differences in electronegativity imparts nucleophilicity to hydroxy and amine groups. A significant problem frequently encountered in organic synthesis is when a molecule possesses two functional groups with similar reactivity, such as a molecule containing both an amine and a hydroxy group. As an example, consider the reaction between bromomethane and 3-amino-1-propanol; seven products are possible from this reaction (**figure 102**). If only one product is desired, say 3-methoxypropan-1-amine, it would first be necessary to protect the amine group with a suitable protecting group, such as Fmoc-Cl.[316,317] The synthesis of 3-methoxypropan-1-amine now requires three steps, the initial protection of the amine group, the *O*-methylation step, and the subsequent deprotection of the Fmoc-amine group to afford the desired product (**figure 102**).

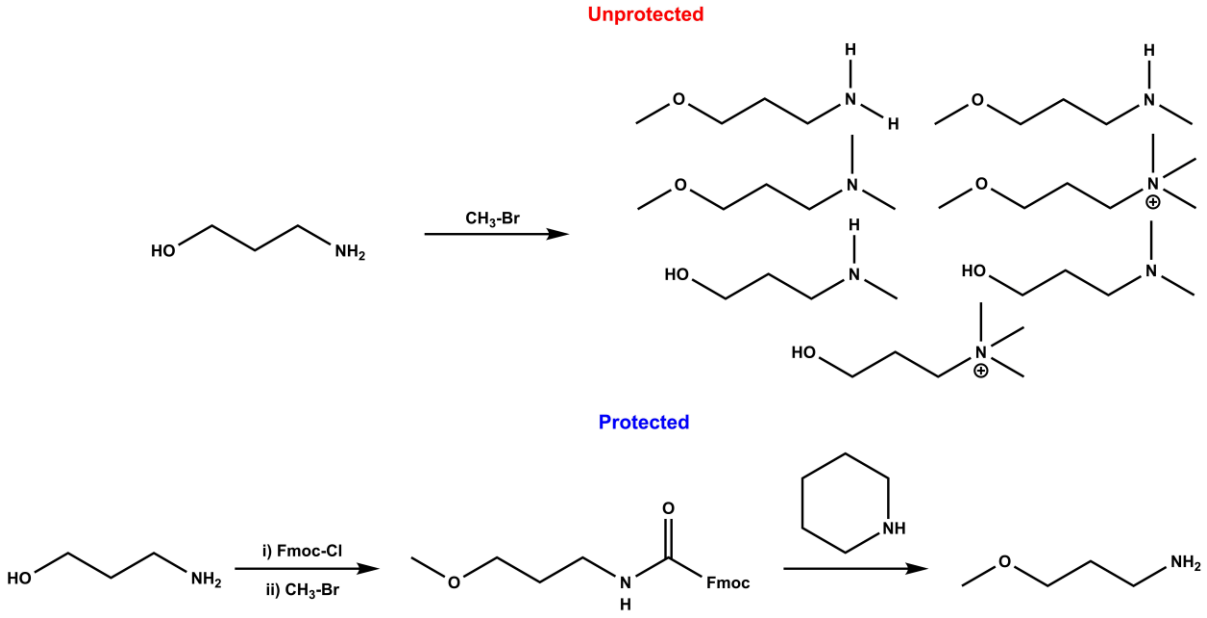


Figure 102: The problem of chemoselectivity. **(top)** the unprotected starting material contains two reactive groups which results in a possible seven products (two possible ammonium salts not shown), and **(bottom)** the amine is protected to afford a single product but requires two additional steps in the synthesis: i) the protection of the amine with Fmoc-chloride, and ii) the deprotection of the Fmoc protecting group to afford the final product.

The problem with the above reaction is a lack of chemoselectivity, defined as a preferential reaction with one functional group over another functional group with similar reactivity. The ability to transform substrates chemoselectivity is of high importance both biologically and technologically in order to avoid complex purification steps and minimize product contamination.[318,319] In addition to stereoselectivity, enzymes often exhibit significant chemoselectivity during biosynthesis,[320–322] selectively transforming one particular functional group in the presence of multiple others. As C45 has already demonstrated proficient N-H carbene insertion chemistry, the question of whether C45 would catalyze the X-H carbene insertion reaction when alkanolamines are employed as substrates was raised. The three regioisomers of aminophenol, 2-aminophenol (2AP), 3-aminophenol (3AP), and 4-aminophenol (4AP) were identified as suitable substrates for exploring the chemoselectivity of X-H insertions (**figure 103**).

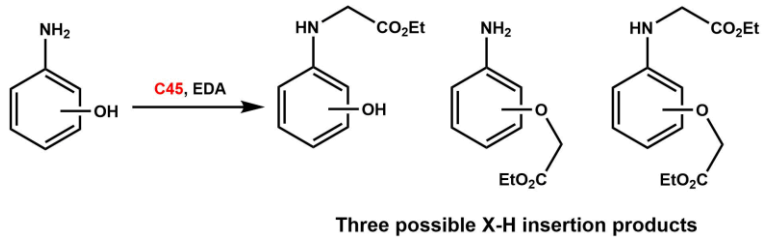


Figure 103: The three possible reaction products following the C45-catalysed O/N-H insertion reaction between a regioisomer of aminophenol and EDA.

Initial reactions were conducted with phenol as the substrate to examine the proficiency of C45 towards catalyzing O-H insertion reactions. The normalized C18-HPLC chromatogram for phenol (red) and for the C45 O-H insertion assay product mixture (blue) are presented in **figure 104**. From the chromatogram, it can be readily identified that no reactivity has occurred when phenol is employed as a substrate, with an almost perfect overlap between the two traces which are dominated by a single major peak with a retention time in the range of 15.2-15.7 minutes. Analysis of the product mixture *via* LC-MS revealed no appreciable O-H insertion product. From the assay it is concluded that C45 does not catalyze the O-H carbene insertion of phenol.

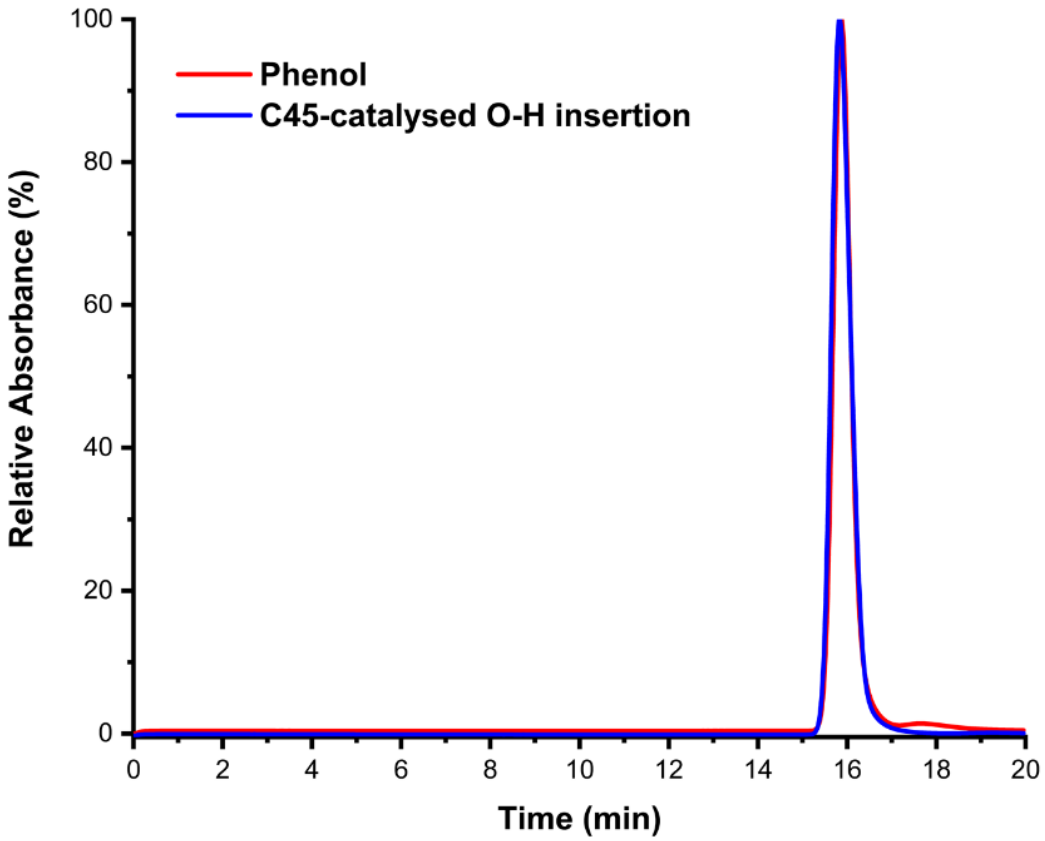


Figure 104: C18-HPLC chromatograms for phenol (red) and the C45-catalysed (10 µM, 0.1 % catalyst loading) O-H insertion assay between phenol (30 mM) and EDA (10 mM) (blue). An isocratic mobile phase (100% CH_3OH: 0.1% w/v NH_4AcO) was employed and injection volumes were 20 µl.

Figure 105 displays the C18-HPLC chromatograms obtained for *N*-phenylglycine ethyl ester, 2AP, 3AP, and 4AP, respectively. *N*-phenylglycine ethyl ester is the product obtained from the reaction between aniline and EDA, and, although structurally different, is similar enough to the anticipated reaction product obtained for the reaction between EDA and xAP (where x = 2,3,4) that is functions as a convenient and suitable comparison. From the chromatograms, at can be readily seen that all three aminophenol regioisomers elute before *N*-phenylglycine ethyl ester, as would be expected given the hydrophobicities and steric considerations of the four compounds.

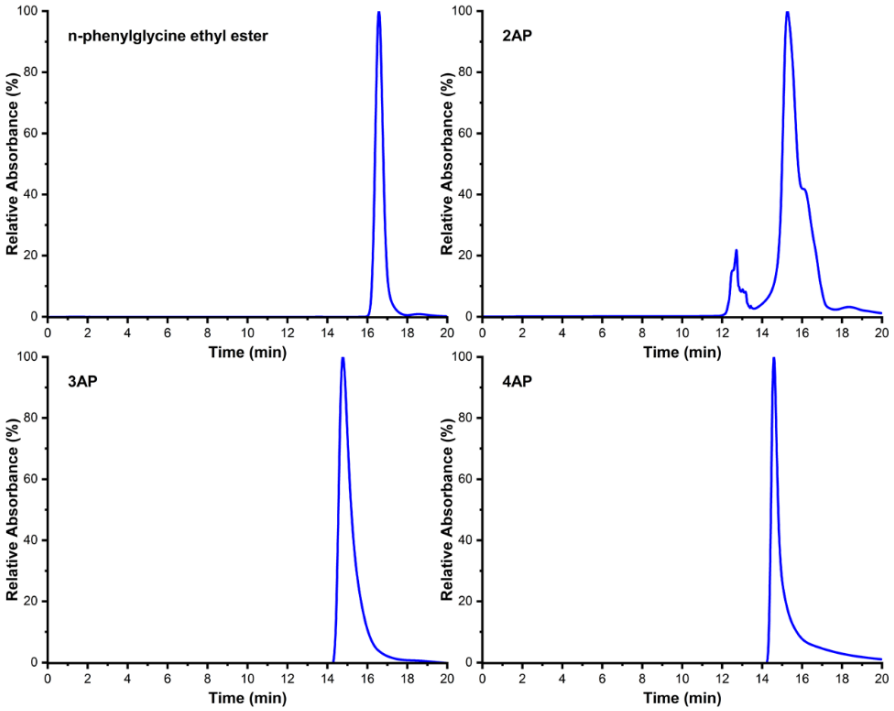


Figure 105: C18-HPLC chromatograms for (**top left**) *N*-phenylglycine ethyl ester, (**top right**) 2-aminophenol, (**bottom left**) 3-aminophenol, and (**bottom right**) 4-aminophenol. An isocratic mobile phase (100% CH_3OH: 0.1% w/v NH_4AcO) was employed and injection volumes were 20 µl.

Figure 106 (top) presents the C18-HPLC chromatograms and the LC-MS mass spectra acquired for the C45-catalysed reaction between 2-aminophenol and EDA, ^{t}BuDA, and BnDA. The chromatogram for the reaction between EDA and 2AP is dominated by a major peak at retention time 16.25 minutes, preluded by a split peak in the range 15.1-15.5 minutes. The split peak preluding the dominant peak is attributed to unreacted 2AP, with the splitting a consequence of the various ionization states accessible by the substrate. The emergence of the new peak at 16.25 minutes corresponds exceedingly well to the retention time reported for *N*-phenylglycine ethyl ester (16.6 minutes), with the deviations probably accounted for by the introduction of a hydroxy group in the *ortho*-position of the product. Subsequent LC-MS analysis of the peak at 16.25 minutes generated a mass spectrum dominated by a peak at 122 m/z, which corresponds to the exact mass expected for the iminium fragment of the 2-AP/EDA single-insertion product ethyl (2-hydroxyphenyl)glycinate. The mass spectrum also possesses a small peak at 195 m/z, corresponding to the expected mass of the $[M+H]^+$ parent ion. The chromatogram for the reaction between ^{t}BuDA and 2AP exhibits a single peak at retention time 15.70 minutes, preluded by a dominant single peak at 15.1 minutes which corresponds to unreacted 2AP. The subsequent LC-MS analysis of the peak at 15.70 minutes generated a mass spectrum dominated by a peak at 122 m/z, which corresponds to the exact mass expected for the

iminium fragment of the 2-AP/^{t}BuDA single-insertion product *tert*-butyl (2-hydroxyphenyl)glycinate. No parent ion could be detected for *tert*-butyl (2-hydroxyphenyl)glycinate, probably on account of the increased probability of fragmentation of a *tert*-butyl group relative to an ethyl group. The chromatogram for the reaction between BnDA and 2AP exhibits a dominant single peak at retention time 16.01 minutes, preluded by a single smaller peak at 15.1 minutes which corresponds to unreacted 2AP. Subsequent LC-MS analysis of the peak at 16.01 minutes generated a mass spectrum dominated by three peaks: i) a peak at 91 m/z corresponding to the tropylium ion fragment previously observed for ESI-MS analysis of benzyl substituents, ii) a peak at 212 m/z, and iii) a peak at 258 m/z which corresponds to the exact mass of the $[M+H]^{+}$ parent ion for benzyl (2-hydroxyphenyl)glycinate.

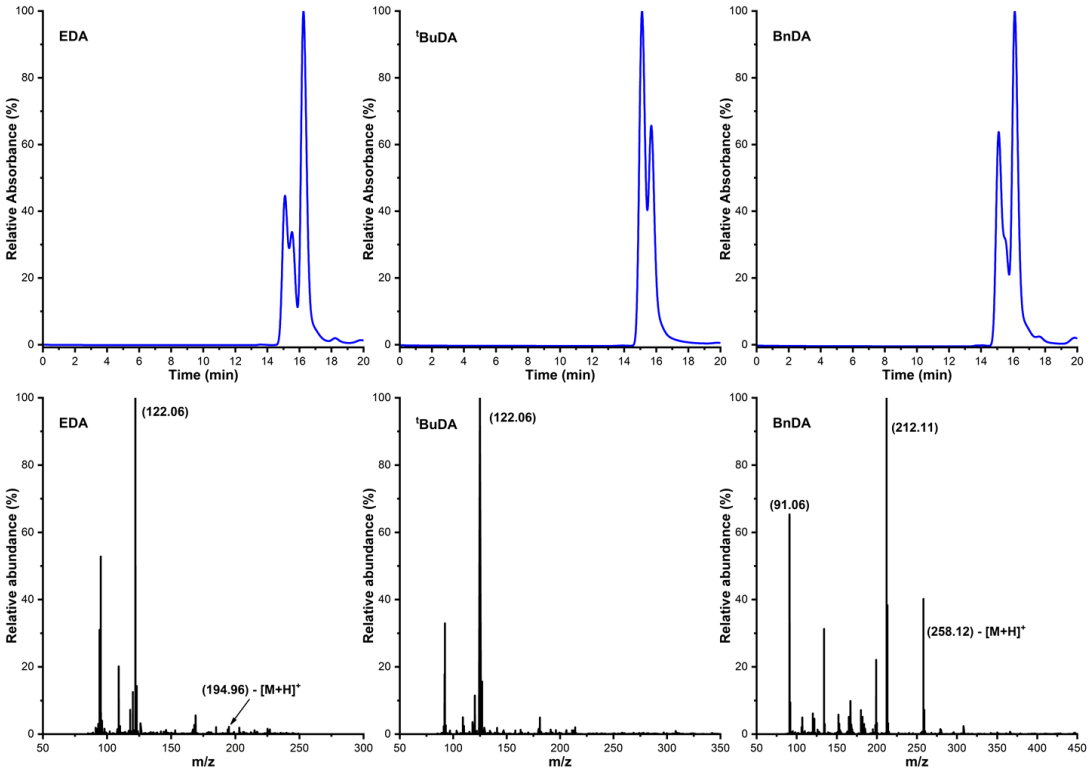


Figure 106: (**top**) C18-HPLC chromatograms for the C45-catalysed (10 µM, 0.1 % catalyst loading) N-H insertion assays between 2-aminophenol (30 mM) and (**top left**) EDA (10 mM), (**top middle**) ^{t}BuDA (10 mM), and (**top right**) BnDA (10 mM) in (CHES buffer, pH 8.6, 254 nm). An isocratic mobile phase (100% CH_3OH: 0.1% w/v NH_4AcO) was employed and injection volumes were 20 µl. (**bottom**) LC-MS spectra for the C45-catalysed (10 µM, 0.1 % catalyst loading) N-H insertion assays between 2-aminophenol (30 mM) and (**bottom left**) EDA (10 mM), (**bottom middle**) ^{t}BuDA (10 mM), and (**bottom right**) BnDA (10 mM). All spectra were recorded in ES+ mode and monitored at 254 and 280 nm. A C8 column was employed for the LC separation with a gradient mobile phase (95:5:0.1 % v/v water/MeCN/formate 10:90:0.1 % v/v water/MeCN/formate). Injection volumes were 20 µl.

The C18-HPLC chromatograms acquired for the C45-catalysed X-H insertion assays employing 2-aminophenol as the substrate revealed the presence of a new, single species being formed in the reaction. The di-insertion product was not detected in the LC-MS analysis, and the fragmentation patterns for all three diazo substrates are consistent with the patterns observed for aniline insertion. The data suggests that C45 is catalyzing the carbene insertion reaction of 2-aminophenol chemoselectively, inserting the carbene exclusively into a N-H bond even in the presence of unprotected O-H groups, a conclusion corroborated by the lack of reactivity exhibited in the phenol O-H insertion assays. From the chromatograms, products yields of 88.17, 66.63, and 70.85% and ($N_{insertion}/O_{insertion}$) ratios of >99.9% were calculated for ethyl (2-hydroxyphenyl)glycinate, *tert*-butyl (2-hydroxyphenyl)glycinate, and benzyl (2-hydroxyphenyl)glycinate respectively (**table 7**). It should be noted that the relative trends observed for product yields previously (EDA > BnDA > ^{t}BuDA) is also observed for 2-aminophenol.

Table 7: The retention times, product yields, and nitrogen/oxygen insertion ratios for the C45-catalysed N-H insertion reactions between the three regioisomers of aminophenol and the diazo compounds EDA, ^{t}BuDA, and BnDA.

	2-aminophenol			3-aminophenol			4-aminophenol		
	t_R (min)	Yield (%)	N/O (%)	t_R (min)	Yield (%)	N/O (%)	t_R (min)	Yield (%)	N/O (%)
AP	15.27	-	-	14.77	-	-	14.60	-	-
EDA	16.25	88.17	>99.9	15.60	73.50	>99.9	15.10	58.12	>99.9
^{t}BuDA	15.70	66.63	>99.9	15.10	>99.9	>99.9	15.42	10.40	>99.9
BnDA	16.01	70.84	>99.9	15.41	53.14	>99.9	15.50	31.36	>99.9

The **figures 107** and **108** present the C18-HPLC chromatograms and the LC-MS mass spectra acquired for the C45-catalysed reaction between 3-aminophenol/4-aminophenol and EDA, ^{t}BuDA, and BnDA, respectively. The trend of chemoselective mono-N-H-insertion is continued across all six reactions. The chromatogram for the reaction between EDA and 3AP is dominated by a major peak at retention time 15.60 minutes, accompanied by a shadowing peak in the range 14.7 minutes corresponding to unreacted 3AP. The subsequent LC-MS analysis of the peak at 16.25 minutes generated a mass spectrum dominated by a peak at 122 m/z, which corresponds to the exact mass expected for the iminium fragment of the 3-AP/EDA single-insertion product ethyl (3-hydroxyphenyl)glycinate. The chromatogram for the reaction between ^{t}BuDA and 3AP exhibits a single peak at retention time 15.10 minutes, with no peak corresponding to 3AP observed (reaction yield >99.9%). Subsequent LC-MS analysis generated a mass spectrum dominated by a peak at 122 m/z, which corresponds to the exact

mass expected for the iminium fragment of *tert*-butyl (3-hydroxyphenyl)glycinate. The chromatogram for the reaction between BnDA and 3AP exhibits a dominant single peak at retention time 15.41 minutes, preluded by a single smaller peak at 14.7 minutes which corresponds to unreacted 3AP. Subsequent LC-MS analysis generated a mass spectrum dominated by two peaks: i) a peak at 91 m/z corresponding to the tropylium ion fragment previously observed for ESI-MS analysis of benzyl substituents, and iii) a peak at 258 m/z which corresponds to the exact mass of the $[M+H]^+$ parent ion for benzyl (3-hydroxyphenyl)glycinate. From the chromatograms, products yields of 73.50, >99.9, and 53.14% and ($N_{insertion}/O_{insertion}$) ratios of >99.9% were calculated for ethyl (3-hydroxyphenyl)glycinate, *tert*-butyl (3-hydroxyphenyl)glycinate, and benzyl (3-hydroxyphenyl)glycinate respectively (**table 7**). The chromatogram for the reaction between EDA and 4AP is dominated by a major peak at retention time 15.10 minutes, with no unreacted 4AP peak detectable (it should be noted that 3M HCl was employed in the purification of assays conducted with 4AP instead of acetic acid. The increased acidity of HCl probably converted most of the unreacted 4-aminophenol to the hydrochloride salt, enhancing its removal from the product mixture upon aqueous-organic work up). Subsequent LC-MS analysis of the peak at 15.10 minutes generated a mass spectrum dominated by a peak at 122 m/z, which corresponds to the exact mass expected for the iminium fragment of the 4-AP/EDA single-insertion product ethyl (4-hydroxyphenyl)glycinate. The chromatogram for the reaction between ^{t}BuDA and 4AP exhibits a single peak at retention time 15.42 minutes, and subsequent LC-MS analysis generated a mass spectrum dominated by a peak at 122 m/z, which corresponds to the exact mass expected for the iminium fragment of *tert*-butyl (4-hydroxyphenyl)glycinate. The chromatogram for the reaction between BnDA and 4AP exhibits a dominant single peak at retention time 15.50 minutes and Subsequent LC-MS analysis generated a mass spectrum dominated by two peaks: i) a peak at 91 m/z corresponding to the tropylium ion fragment previously observed for ESI-MS analysis of benzyl substituents, and iii) a peak at 258 m/z which corresponds to the exact mass of the $[M+H]^+$ parent ion for benzyl (4-hydroxyphenyl)glycinate. The mass spectrum was, however, reasonably unresolved, probably on account of the low reaction yield (10.4%). From the chromatograms, product yields of 58.12, 10.40, and 31.36% and ($N_{insertion}/O_{insertion}$) ratios of >99.9% were calculated for ethyl (4-hydroxyphenyl)glycinate, *tert*-butyl (4-hydroxyphenyl)glycinate and benzyl (4-hydroxyphenyl)glycinate respectively (**table 7**).

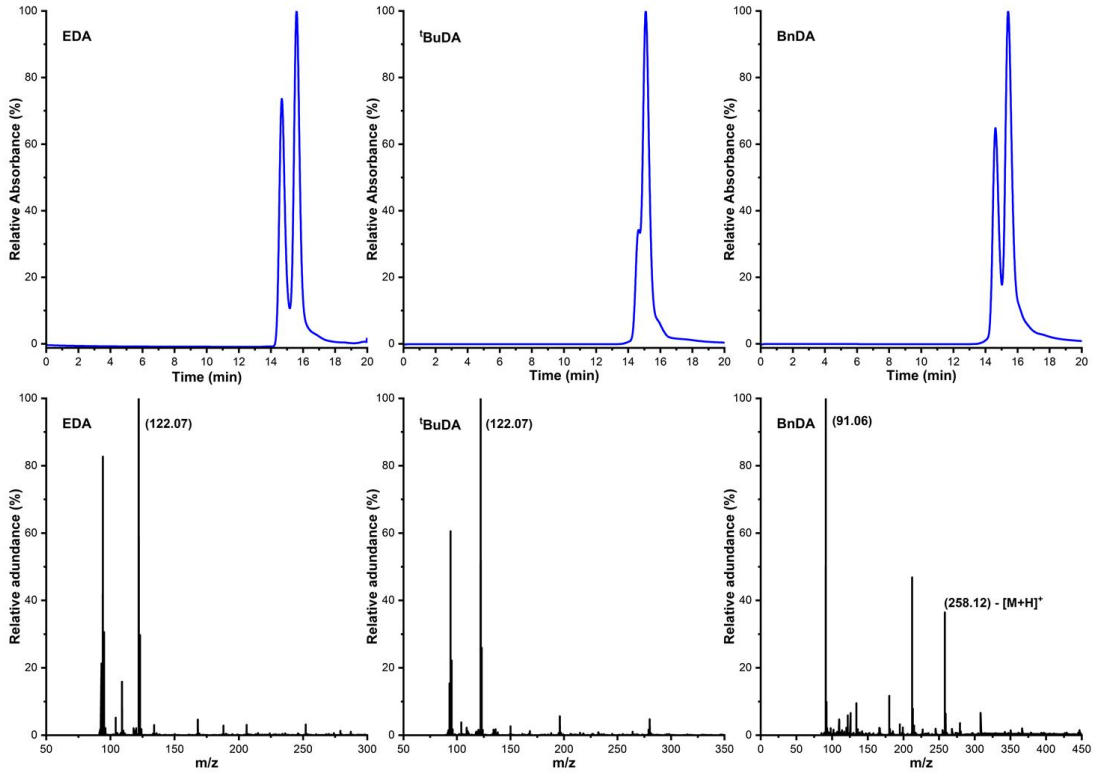


Figure 107: (**top**) C18-HPLC chromatograms for the C45-catalysed (10 µM, 0.1 % catalyst loading) N-H insertion assays between 3-aminophenol (30 mM) and (**top left**) EDA (10 mM), (**top middle**) ^{t}BuDA (10 mM), and (**top right**) BnDA (10 mM) in (CHES buffer, pH 8.6, 254 nm). An isocratic mobile phase (100% CH_3OH: 0.1% w/v NH_4AcO) was employed and injection volumes were 20 µl. (**bottom**) LC-MS spectra for the C45-catalysed (10 µM, 0.1 % catalyst loading) N-H insertion assays between 3-aminophenol (30 mM) and (**bottom left**) EDA (10 mM), (**bottom middle**) ^{t}BuDA (10 mM), and (**bottom right**) BnDA (10 mM). All spectra were recorded in ES+ mode and monitored at 254 and 280 nm. A C8 column was employed for the LC separation with a gradient mobile phase (95:5:0.1 % v/v water/MeCN/formate 10:90:0.1 % v/v water/MeCN/formate). Injection volumes were 20 µl.

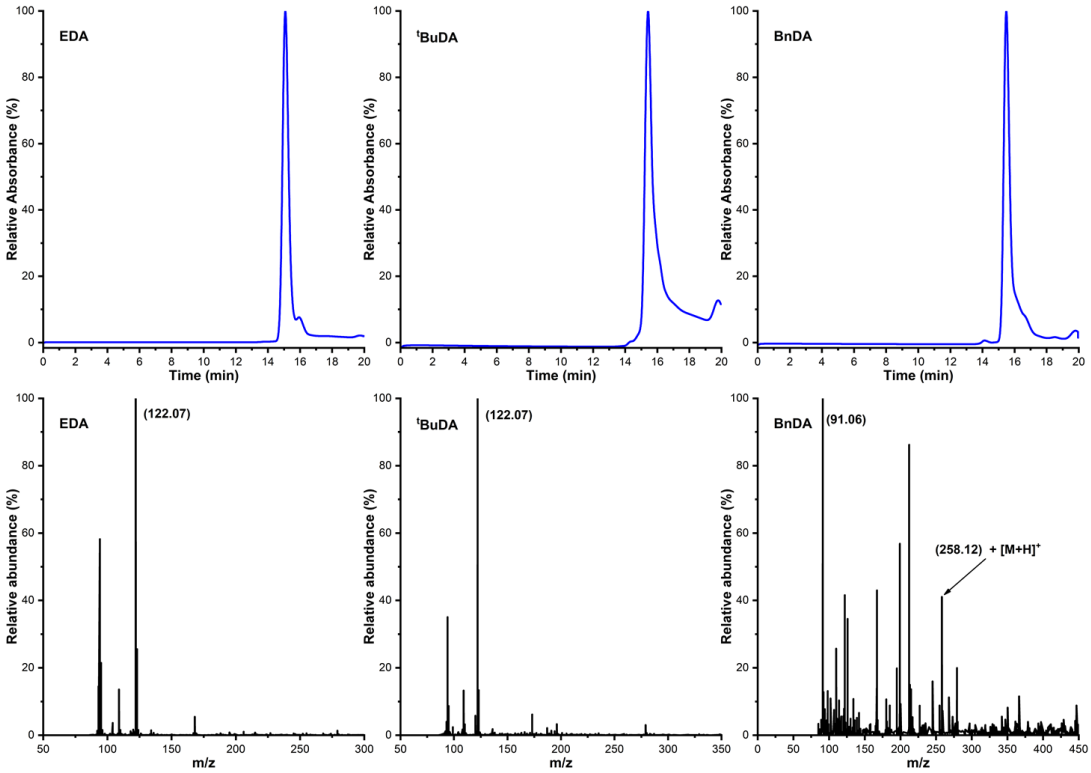


Figure 108: (**top**) C18-HPLC chromatograms for the C45-catalysed (10 µM, 0.1 % catalyst loading) N-H insertion assays between 4-aminophenol (30 mM) and (**top left**) EDA (10 mM), (**top middle**) ^{t}BuDA (10 mM), and (**top right**) BnDA (10 mM) in (CHES buffer, pH 8.6, 254 nm). An isocratic mobile phase (100% CH_3OH: 0.1% w/v NH_4ACO) was employed and injection volumes were 20 µl. (**bottom**) LC-MS spectra for the C45-catalysed (10 µM, 0.1 % catalyst loading) N-H insertion assays between 4-aminophenol (30 mM) and (**bottom left**) EDA (10 mM), (**bottom middle**) ^{t}BuDA (10 mM), and (**bottom right**) BnDA (10 mM). All spectra were recorded in ES+ mode and monitored at 254 and 280 nm. A C8 column was employed for the LC separation with a gradient mobile phase (95:5:0.1 % v/v water/MeCN/formate 10:90:0.1 % v/v water/MeCN/formate). Injection volumes were 20 µl.

From the data, it can be concluded that C45 is a proficient and remarkably chemoselective enzyme for catalyzing the N-H carbene insertion of unprotected alkanolamines. Furthermore, two trends can be identified with respect to product yield: i) the reaction yields follow the trend EDA > BnDA > ^{t}BuDA, and ii) the reaction yields increase with decreasing molecular volume of the substrate (1.38, 1.51 and 1.61 $x10^{-22}$ cm^3 for 2AP, 3AP and 4AP respectively). The first trend has been observed previously for styrene cyclopropanation and *p*-chloroaniline insertion and has been rationalized by the sterics of the three substituents, and possible π-π interactions between the phenyl group of benzyl diazoacetate and aromatic residues in the active site of C45. The increase in product yield with respect to decreasing molecular volume suggests a key factor governing C45's functionality is diffusion and accommodation of a substrate into the active site, a process easier for smaller molecules. The observed general trend

in product yields (ortho > meta > para; 2AP > 3AP > 4AP) can therefore be accounted for, at least partially, by the molecular volume of the substrate (ortho < meta < para; 2AP < 3AP < 4AP), an observation paralleled by the observed rate of diffusion of the three regioisomers of xylene through a zeolitic imidazolate framework: the decreasing size of the molecule (*o*-xylene < *m*-xylene < *p*-xylene) decreases its rate of absorption.[323] No observable relationship was discerned between product yields and the pKa(NH) of the substrates (pKa(NH) = 4.37 (3AP) < 4.78 (2AP) < 5.48 (4AP)), indicating diffusion into the active site, and not the nucleophilicity of the amine, plays a more important role in governing enzymatic activity. One remarkable anomaly, however, is the C45-catalysed N-H insertion reaction between ^{t}BuDA and 3-aminophenol which affords the product with yields consistently >99.9%. The remarkable enzymatic activity exhibited by C45 towards this reaction is likely accounted for by a combination of sterics and electronics. The bulky *tert*-butyl group perhaps disrupts the active-site of C45 upon metallocarbenoid formation, causing select amino acids in the active site to adopt a temporary deviation in conformation. The distortion of the active site could possibly result in a thermodynamic and kinetic effect which accounts for the observed catalytic efficiency. First, a distortion of the active site, by *tert*-butyl, could facilitate the diffusion of a substrate (X-aminophenol) into the active site which would enhance catalysis; however, if the only contributing factor to catalysis was a general distortion of the active site, it would be expected that such a phenomena would translate over to all reactions concerning *tert*-butyl diazoacetate. Therefore, it is conceivable that any deviations away from ideal confirmations to accommodate the bulky *tert*-butyl-substituted metallocarbenoid intermediate could engender an active site topology perfectly arranged to provide selective stabilizing interactions to 3-aminophenol, and/or interactions which modify the pKa of the -NH group, rendering the amine significantly more nucleophilic. Alternatively, it is also possible that such interactions could orientate the 3-aminophenol amine group perfectly for nucleophilic attack on the metallocarbenoid carbon, thus facilitating catalysis. The exact dynamics and electronics governing the reaction between ^{t}BuDA and 3-aminophenol remain unelucidated and requires additional experiments in the future to be fully understood. One possible experiment would involve isotopically labelling 3-aminophenol with ^{18}O and observing any changes in the reaction kinetics. A change in kinetics would indicate the *meta*-hydroxy group plays an important role in the mechanism, probably through an orientating interaction with an active-site amino acids, whereas no appreciable change would eliminate the *meta*-hydroxy groups involvement and would suggest the reason for the enhanced kinetics could be accounted for by protein dynamics and sterics.

5.2.1.5 phenylenediamines: the building blocks for chelating ligands and heteropolymers

In addition to the chemoselective N-H insertion of alkanolamines, another reaction of interest concerns the di-insertion reaction of diamine substrates. The bifunctionality of diamines means this class of molecules are often utilised as cross-linking agents and/or as precursors in heteropolymer synthesis; one prominent example being the aliphatic diamine 1,6-diaminohexane, a precursor in the synthesis of Nylon.[324] The three regioisomers of phenylenediamine are industrially relevant diamines, with all three regioisomers finding applications as precursors to an assortment of industrial useful chemicals.[324] *Ortho*-phenylenediamine is used in the synthesis of chelating ligands and commercial dyes, *meta*--phenylenediamine is employed in the preparation of aramid fibres, epoxy resins and polyurea elastomers and commercial dyes, and *para*-phenylenediamine is employed in the synthesis of Kevlar and Tawron.[324] The proficiency of C45 towards mono N-H carbene insertion reactions raised the question of whether C45 could be used as a catalyst for di-N-H insertions of phenylenediamine regioisomers. If successful, it would afford a new biosynthetic methodology for manufacturing a large variety of novel bifunctional monomers which could find applications in the synthesis of new phenylenediamine-containing heteropolymers.

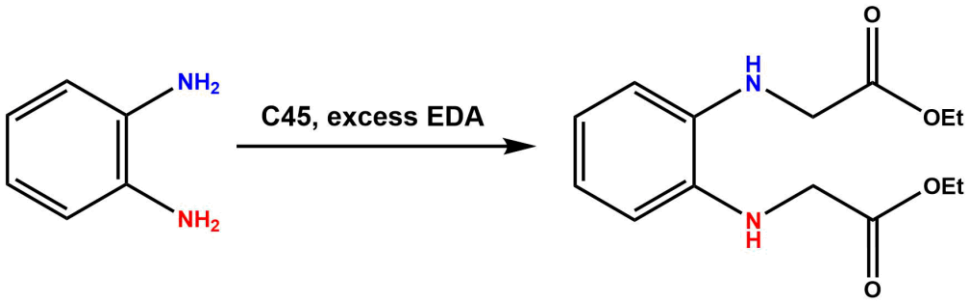


Figure 109: The general reaction scheme for a di-NH-insertion reaction. The di-NH-insertion reaction *ortho*-phenylenediamine and excess EDA catalysed by C45.

Initial small-scale experiments were conducted using *ortho*-phenylenediamine (**figure 109**). In all the N-H insertion assays conducted so far, the ratio of substrate:diazo compound has been 3:1 to encourage mono-insertion and disfavour di-insertion. As the di-insertion product was desired from the C45-catalysed N-H insertion reactive between EDA and *ortho*-phenylenediamine the diamine:EDA ratio was modified to 1:5 to encourage di-insertion. After approximately two hours, the reaction mixture was extracted and analysed *via* LC-MS. The acquired LC-MS mass spectrum for the C45-catalysed reaction between EDA and *o*-phenylenediamine is presented in **figure 110** and is characterised by a dominant peak at 281 m/z, which corresponds to the exact mass of the $[M+H]^+$ parent ion of the double insertion product. The mass spectrum also contains a major peak at 207 m/z, which corresponds to the exact mass of the iminium fragment of the double insertion product, a

fragment detected ubiquitously across all N-H insertion assays and diagnostic of a successful N-H insertion reaction. A negligible quantity of the single insertion product was detected by LC-MS.

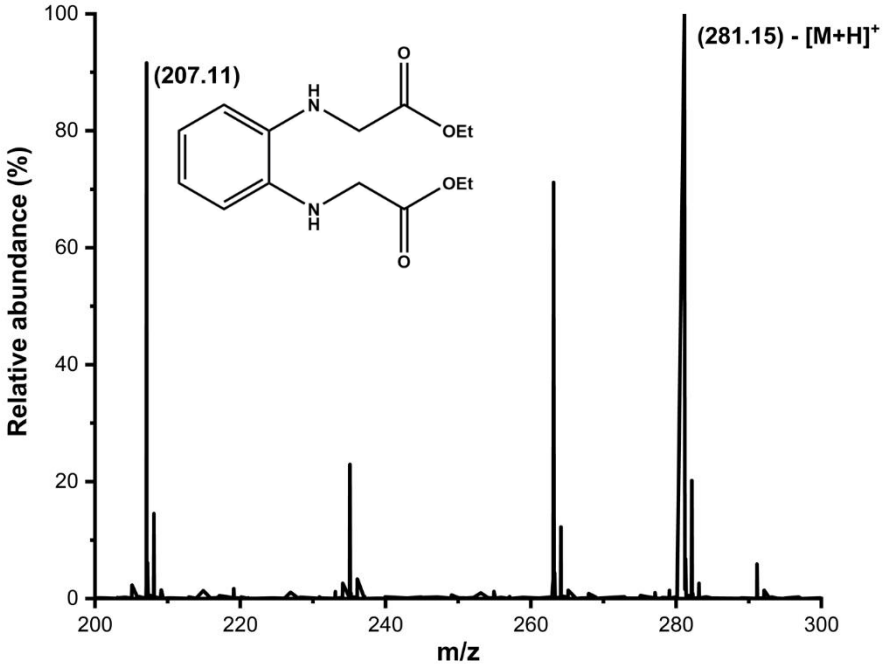


Figure 110: LC-MS spectra for the C45-catalysed (10 µM, 0.1 % catalyst loading) N-H insertion assay between *ortho*-phenylenediamine (10 mM) and EDA (30 mM). The spectrum was recorded in ES+ mode and monitored at 254 and 280 nm. A C8 column was employed for the LC separation with a gradient mobile phase (95:5:0.1 % v/v water/MeCN/formate 10:90:0.1 % v/v water/MeCN/formate). The Injection volume was 20 µl.

The initial assays revealed two interesting points. Firstly, C45 remains catalytically active following an initial carbene insertion, and is capable of reforming the metallocarbenoid intermediate and conducting another round of catalysis. Secondly, C45 can accommodate the single-insertion product into the active site to catalyse the second-insertion reaction. This work will be extended to *meta*-phenylenediamine and *para*-phenylenediamine at a later date. If successfully, the di-insertion products would provide suitable monomers for the synthesis of copolymers *via* a condensation reaction with a diol or peptide bond formation with a diamine.

5.2.1.6 C45-catalysed Si-H insertion – a comparison with *Rma*-TDE

The success of the C45-catalysed N-H insertions for a variety of different substrates raised the question of whether C45 could be used to catalyse a different X-H insertion reaction. The cytochrome *c* mutant *Rma*-TDE, a variant generated in the Arnold lab that is three mutations away from a WT cytochrome *c* isolated from *Rhodothermus marinus*, was engineered for preferential carbene insertion into Si-H σ-

bonds to form new Si-C σ-bonds.[217,219] This work expand the reaction scope of biocatalysis towards the synthesis of novel organosilicon compounds.[217,219] As *Rma*-TDE was employed as a suitable comparison for metallocarbenoid formation and cyclopropanation activity against C45, it was rationalised that the propensity of C45 to also catalyse the carbene insertion reaction of Si-H containing molecules should be investigated and compared to the activity reported for *Rma*-TDE. In the literature, the Arnold lab investigated the catalytic activity of *Rma*-TDE towards Si-H insertion by employing ethyl 2-diazopropanoate (Me-EDA) and dimethylphenylsilane (DMPS). The formation of a new Si-C σ-bond using Me-EDA and DMPS as substrates also results in the emergence of a new stereogenic centre, meaning two possible stereoisomers are possible from this reaction. In the assays conducted using C45, DMPS was employed as the silane substrate, but EDA was employed as the diazo-compound. The selection of EDA as the carbene precursor means that no stereogenic centre arises from the reaction, and hence only the activity, and not the stereoselectivity, of C45 towards Si-H insertion can be compared to *Rma*-TDE (**figure 111**).

Figure 111: The general C45/*Rma*-TDE-catalysed Si-H insertion assay between EDA and DMPS.

The Si-H insertion assays were conducted according to the protocol established for investigating carbene transferase activity. The protocol involves a small-scale (400 µl) reaction between EDA and dimethylphenylsilane (DMPS) catalysed by the enzyme of interest (C45 and *Rma*-TDE) under anaerobic conditions. The reaction is left to mix for 2 hours at room temperature before the product is extracted into an organic solvent and analysed. The analysis of the product mixture was conducted *via* HPLC, employing a C18 Column as the stationary phase. LC-MS was then used to assist in product identification. A C18-HPLC chromatogram of the starting substrate, DMPS (in CH_3CN), and a Si-insertion assay conducted in the absence of any enzyme were acquired initially. The assay was conducted in 1.5 mL seal-top reaction vials, which was purged with nitrogen prior to conducting the assay. The final reaction volume was 400 µl, and consisted of 10 mM DMPS, 10mM EDA and 10 mM sodium dithionite in CHES buffer (100 mM KCl, 20mM CHES, pH 8.6). The reaction vial was allowed to stir for 2 hours, at room temperature, before the reaction was extracted with n-hexane for analysis. The chromatogram for the starting substrate, DMPS, and the Si-H insertion assay conducted in the absence of any enzyme are both presented in **figure 112**.

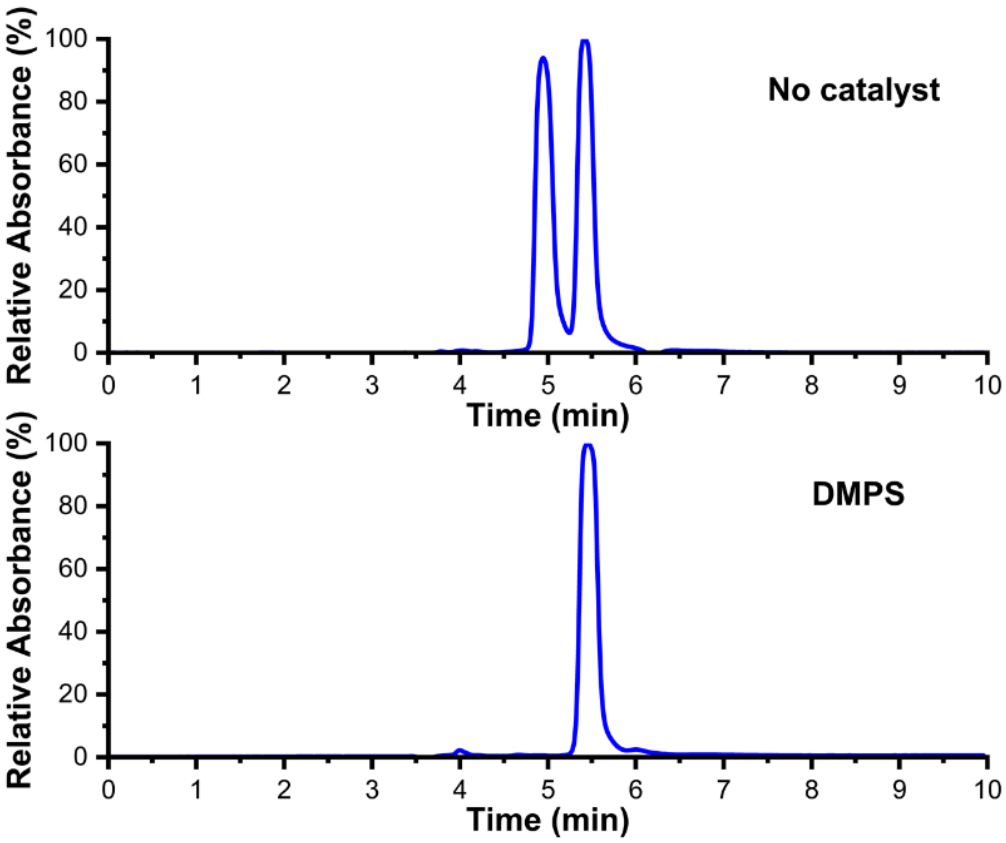


Figure 112: C18-HPLC chromatograms for (**top**) the Si-H insertion reaction between dimethylphenylsilane (30 mM) and EDA (10 mM) in the absence of any catalyst, (**bottom**) of dimethylphenylsilane alone (1 mM, EtOH, 220 nm). An isocratic mobile phase (100% CH_3OH: 0.1% w/v NH_4AcO) was employed and injection volumes were 20 µL.

It is clearly seen in **figure 112** that the chromatogram for DMPS is dominated by a single symmetric peak with a retention time of 5.4 minutes. The chromatogram for the Si-H insertion assay conducted in the absence of any enzyme is dominated by two symmetrical peaks with retention times of 4.9 and 5.4 minutes, which corresponds to EDA and DMPS respectively; no Si-H insertion product was identified in the assays conducted in the absence of any enzyme. The assay was then repeated under identical conditions employed in the controls but in the presence of either *Rma*-TDE or C45. The normalised average chromatograms acquired for the C45-cataysed and *Rma*-TDE-catalysed assays are presented in **figure 113 (top)**.

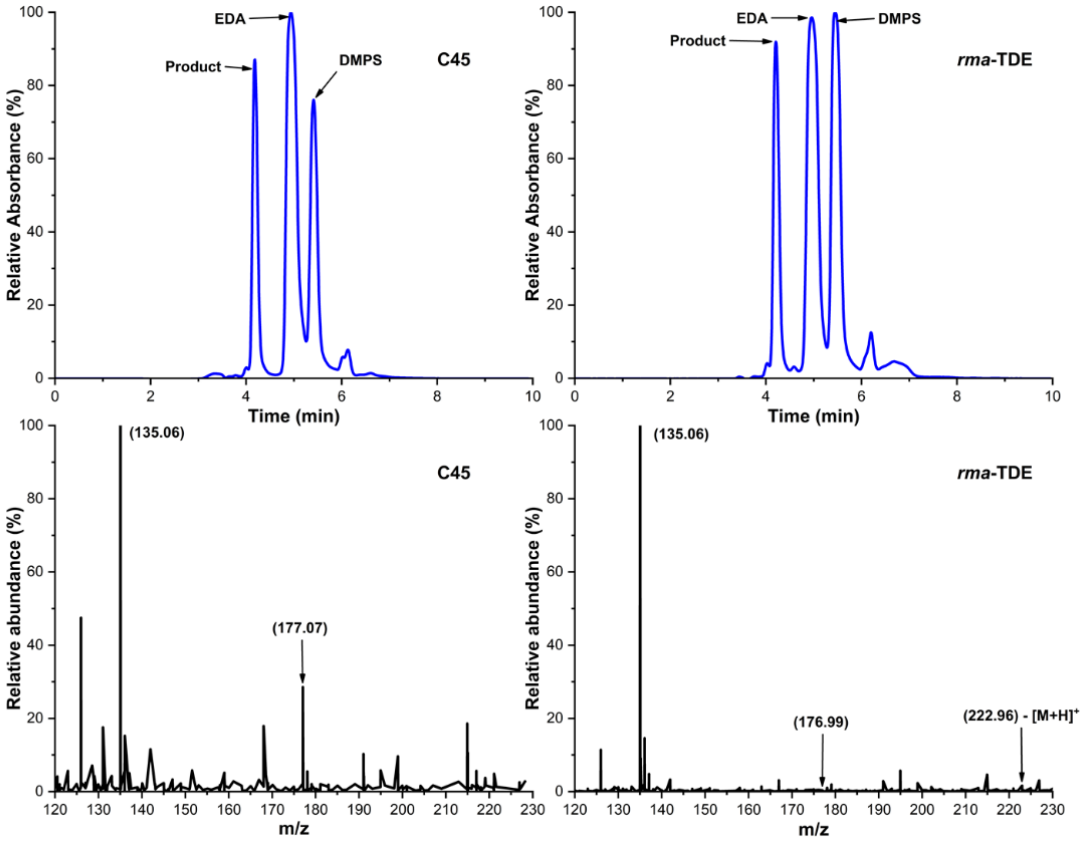


Figure 113: (**top**) C18-HPLC chromatograms for the Si-H insertion assays between dimethylphenylsilane (30 mM) and EDA (10 mM) catalysed by (**top left**) C45 (10 µM, 0.1 % catalyst loading) and (**top right**) *Rma*-TDE (10 µM, 0.1 % catalyst loading). An isocratic mobile phase (100% CH_3OH: 0.1% w/v NH_4AcO) was employed and injection volumes were 20 µL. (**bottom**) LC-MS spectra for the Si-H insertion assays between dimethylphenylsilane (30 mM) and EDA (10 mM) catalysed by (**bottom left**) C45 (10 µM, 0.1 % catalyst loading) and (**bottom right**) *Rma*-TDE (10 µM, 0.1 % catalyst loading). All spectra were recorded in ES+ mode and monitored at 254 and 280 nm. A C8 column was employed for the LC separation with a gradient mobile phase (95:5:0.1 % v/v water/MeCN/formate 10:90:0.1 % v/v water/MeCN/formate). Injection volumes were 20 µl.

The chromatograms for both the C45- and *Rma*-TDE-catalysed Si-H reaction between DMPS and EDA both contain the two major peaks observed in the control assays (EDA, 4.9 minutes; DMPS, 5.4 minutes), but also both contain the emergence of a new peak with a retention time of 4.2 minutes. It has already been reported in the literature that *Rma*-TDE catalyses Si-H insertion reactions with reasonably proficiency, therefore the new peak at 4.2 minutes could be attributed to the Si-H insertion product. In order to confirm the identity of the new peak, the C45- and *Rma*-TDE-catalysed assay mixtures was subsequently analysed by LC-MS; the acquired normalised spectra for both enzymes are displayed in **figure 113 (bottom)**. It can be readily seen that both mass spectra contain a dominant peak at 135 m/z, which corresponds to the $PhMe_2Si^+$ fragment generated from the protonation of the carbonyl oxygen followed by the subsequent heterolytic cleavage of the new Si-C σ-bond to yield a

neutral enolate (which most probably tautomerizes) and the observed fragment. The mass spectrum for C45 (and very marginally in the *Rma*-TDE spectrum) contains an additional peak at 177 m/z, which corresponds to the oxonium ion generated from the elimination of EtOH following the protonation of the parent ion; the diagnostic fragmentation observed in the mass spectra for cyclopropanation assays conducted using EDA. The mass spectrum for *Rma*-TDE also contains a very small peak at 223 m/z, which corresponds to the mass of the protonated parent ion $[M+H]^+$.

The data acquired from the C45- and *Rma*-TDE-catalysed Si-H insertion reactions indicate the new peak can be attributed to the expected Si-H insertion product ethyl 2-(dimethyl(phenyl)silyl)acetate, and is consistent with the literature reports of the assays conducted with *Rma*-TDE. Therefore, it can be reasonably inferred that C45 is also proficient at catalysing carbene insertion reactions into Si-H σ-bonds, opening up the possibility for *de novo* enzymatic synthesis of novel organosilicon compounds. Average product yields of 57.66 (± 4.14) and 62.86 (± 1.95) % were calculated from the chromatograms for the reactions catalysed by *Rma*-TDE and C45 respectively (**table 19**, *appendix 2*).

5.2.2 C45 catalysed carbonyl olefination

One of the most challenging, yet pertinent, reactions in organic chemistry concerns the synthesis of new C-C σ/π-bonds. One prominent example of a C-C bond forming reaction is the Wittig reaction – the reaction between a carbonyl group and a phosphonium ylide to form a new C=C π-bond, with the stereoselectivity of the new C=C bond controlled by the electronics of the phosphonium ylide (**figure 114**). Although the Wittig reaction is a renowned reaction in organic chemistry, there are several limitations concerning stereoselectivity and functional group tolerability which need to be addressed.

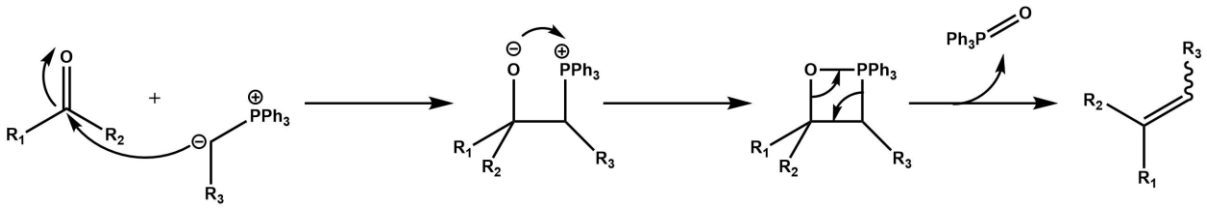


Figure 114: The general reaction scheme for the Wittig reaction.

Following the success described in the C45-catalysed cyclopropanation and N-H insertion assays, and the activity reported by Fasan and Weissenborn, enzymatic carbonyl olefination was identified as another possible reaction to be investigated. The mechanism for enzymatic carbonyl olefination proceeds first *via* the nucleophilic attack of the phosphine (or arsine) species on the electrophilic metallocarbenoid intermediate to yield a reactive ylide (**figure 115**). The ylide subsequently attacks the carbonyl group of a carbonyl species which spontaneously results in the formation of a four-

membered cyclic oxaphosphetane intermediate. A concentred rearrangement of the oxaphosphetane intermediate's C-O and P-C σ-bonds results in the expulsion of a phosphine oxide side product and the desired α,β-unsaturated carbonyl product (**figure 115**). Although the mechanism of hemoprotein-catalysed olefin carbonylation does not proceed *via* a direct carbene transfer, the formation of a heme-localised metallocarbenoid intermediate is a prerequisite for ylide formation. It is therefore conceivable that C45 could exhibit carbonyl olefination activity when presented with a suitable phosphine species.

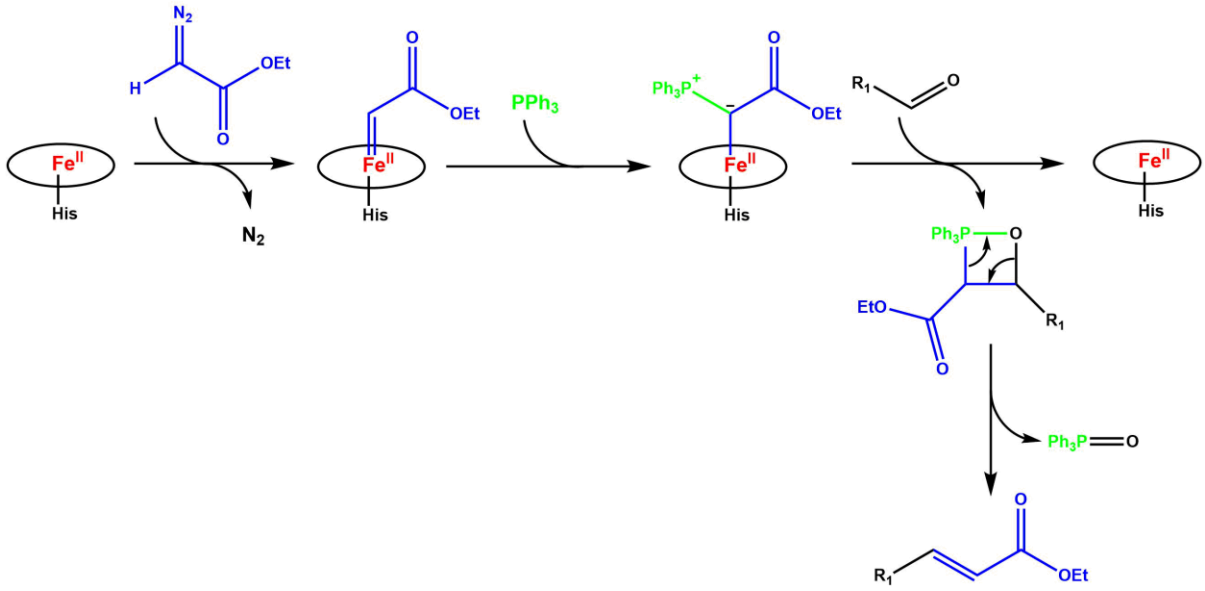


Figure 115: The mechanism for the hemoprotein-catalysed carbonyl olefination between EDA and a carbonyl substrate in the presence of triphenylphosphine.

The C45-catalysed carbonyl olefinations assays were conducted in an analogous way to the previous assays, with two notable modifications: i) an aldehyde-containing substrate was employed, and ii) PPh_3 was introduced. Benzaldehyde was identified as a suitable substrate for the initial assay, and three additional functionalised benzaldehydes were employed to examine substrate promiscuity: i) *p*-cyano, ii) *p*-nitro, and 3,4,5-trimethoxy- benzaldehyde. EDA was employed as the diazo compound. The reaction conditions employed were identical to the previously detailed assays. The reactions were performed in 1.5 ml sealed, screw-top vials and were purged with nitrogen. The final concentration of each constituent was 10 mM diazo precursor, 30 mM aldehyde, 10 mM sodium dithionite, 10 mM PPh_3 and 10 µM C45. The final volume for each reaction was 400 µl, and all reactions were performed in CHES buffer (20 mM CHES, 10 mM KCl, pH 8.6). The reaction mixture was quenched after 2 hours, and the product mixture was extracted and analysed *via* HPLC (using an achiral C18 column) and LC-MS. All the reactions were compared against commercially prepared samples of the anticipated products.

The anticipated reaction product arising from the carbonyl olefination reaction between EDA and benzaldehyde is ethyl cinnamate. The C18-HPLC trace of a commercial sample of (*E*)-ethyl cinnamate is dominated by a major peak with a retention time of 4.9 minutes. The LC-MS spectrum of (*E*)-ethyl cinnamate is dominated by major peaks at m/z 104, 131, and 177 corresponding the loss of EtOH and CO, EtOH only, and the $[M+H]^+$ parent ion, respectively. An external calibration for (*E*)-ethyl cinnamate, showing a linear relationship between sample concentration and the peak height response in the chromatogram, is presented in **figure 167** *appendix 1*.

Initial assays were conducted in i) the absence of any catalyst, and ii) the presence of free hemin (10 µM) for reaction between benzaldehyde and EDA in the presence of PPh_3. The reaction mixtures were quenched and then extracted with 1 ml of ethyl acetate prior to analysis. The reaction mixtures were analysed *via* HPLC and the resultant chromatograms for all reactions are presented in **figure 116**. The chromatograms for the assays conducted with and without hemin are both dominated by two major peaks with retention times of 6.5-6.8 and 8.1-8.3 minutes which corresponds to PPh_3 and benzaldehyde respectively; no reaction product was detectable in either assay. The product mixtures were subsequently analysed *via* LC-MS and also lacked any detectable trace of the anticipated product. These assays indicate that no reaction occurs between EDA, PPh_3, and benzaldehyde in the presence/absence of hemin.

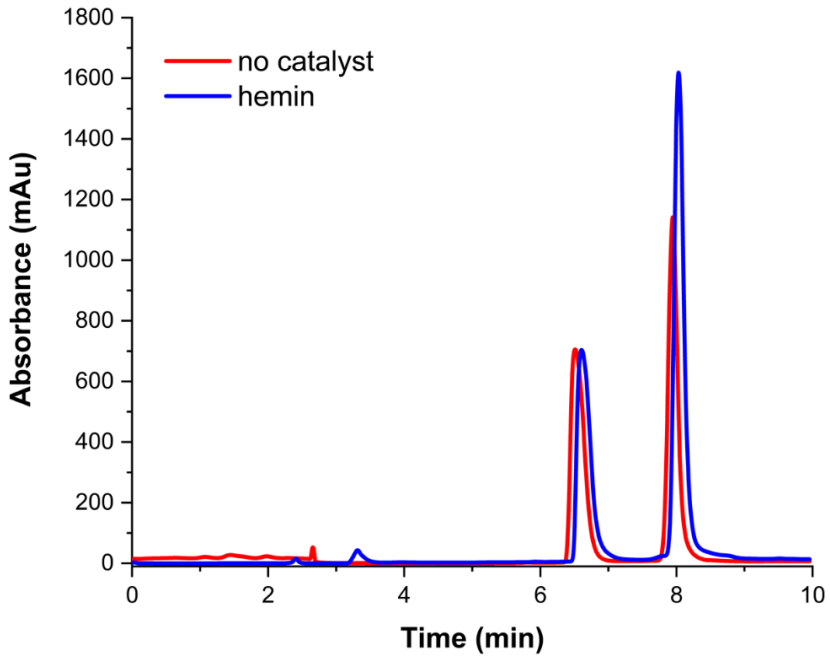


Figure 116: The C18-HPLC chromatogram for the carbonyl olefination reaction between PPh_3 (10 mM), EDA (10 mM) and benzaldehyde (30 mM) in the absence (red) and presence (blue) of hemin. An isocratic mobile phase was employed (100% CH_3CN: 0.1% v/v TFA: 0.1% v/v: Et_3N) and injection volumes were 20 µl.

The C18-HPLC chromatogram obtained from the C45-catalysed carbonyl olefination reaction between EDA and benzaldehyde is presented in **figure 117**. The chromatogram is dominated by a split peak ranging from 4.7-5.3 minutes, with apexes at retention times 4.9 minutes and 5.1 minutes. In comparison to the chromatogram obtained for commercial (*E*)-ethyl cinnamate, a good overlap can be deduced for the peak at 4.9 minutes; however, the presence of another apex 5.1 requires identification. The LC-MS for the assay is presented in **figure 118** and possesses the same fragmentation profile observed for the commercial (*E*)-ethyl cinnamate, indicating a reaction has occurred. The peak at 5.1 minutes could correspond to either i) (*Z*)-ethyl cinnamate, or ii) the dimerization product. From the LC-MS, no dimerization product could be deducted, nor could the presence of any additional impurities; therefore, the peak at 5.1 minutes has been attributed to (*Z*)-ethyl cinnamate. From the chromatogram, the yields for the (*E*)- and (*Z*)- products were calculated to be 2.57% and 3.7% respectively, translating to TTNs of 26 (cis) and 24 (trans) and a cis/trans ratio of 1.1 (**table 8**). The C18-HPLC chromatograms obtained from the C45-catalysed carbonyl olefination reaction between EDA and i) *para*-nitro, ii) *para*-cyano, and iii) 3,4,5-trimethoxy- benzaldehyde are also presented in **figure 117**. All three chromatograms exhibit the same pattern observed in the chromatogram for benzaldehyde, characterised by split peak ranging from 4.7-5.2 minutes, with apexes at retention times 4.9 minutes and 5.1 minutes. From the chromatograms, the yields, TTNs and cis/trans ratios were calculated to be 4.21% (cis) and 4.21% (trans), 42 (cis and trans), and 1 for *para*-nitrobenzaldehyde, 3.16% (cis) and 3.19% (trans), 32 (cis) and 32 (trans), and 0.99 for *para*-cyanobenzaldehyde, and 5.47% (cis) and 5.53% (trans), 55 (cis) and 55 (trans), and 1 for 3,4,5-trimethoxybenzaldehyde respectively (**table 8**).

Table 8: The product yields, total turnover numbers, and cis/trans ratios for the C45-catalysed carbonyl olefination reactions between functionalised benzaldehydes and EDA in the presence of PPh_3.

Substrates (vs EDA)	% Yield$_{(cis)}$	% Yield$_{(trans)}$	cis/trans	TTN$_{cis}$	TTN$_{Trans}$
Benzaldehyde	2.57 (± 0.52)	2.37 (±0.24)	1.1	26	24
para-nitrobenzaldehyde	4.21 (± 1.98)	4.21 (± 1.98)	1	42	42
para-cyanobenzaldehyde	3.16 (± 0.96)	3.19 (± 1.00)	0.99	32	32
3,4,5-trimethoxybenzaldehyde	5.47 (± 3.91)	5.53 (± 4.00)	0.99	55	55

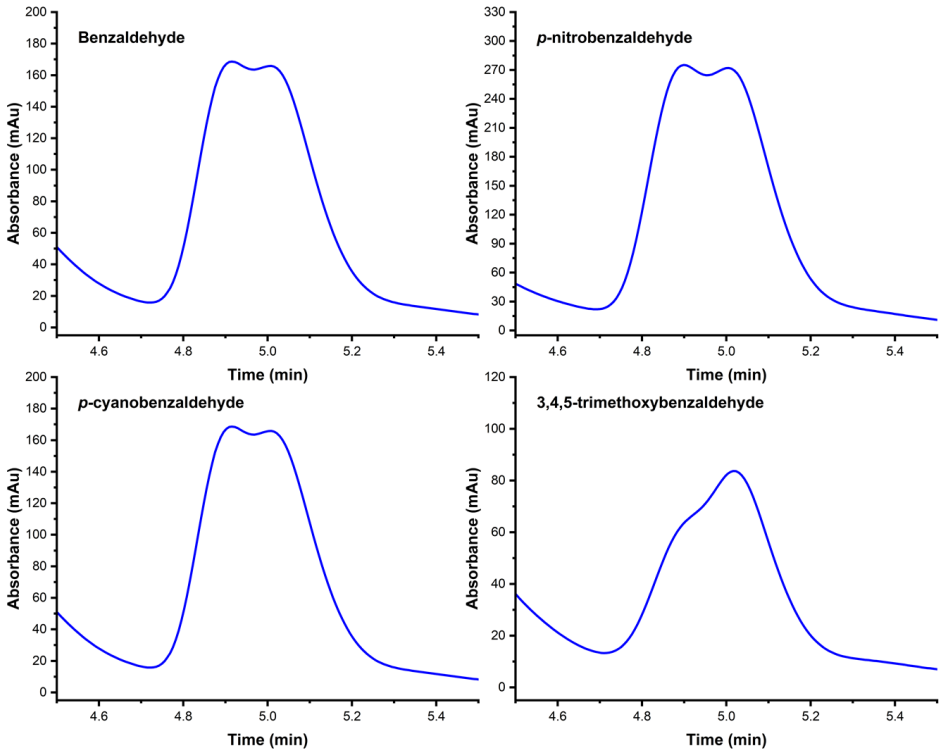


Figure 117: The C18-HPLC chromatogram for the C45-catalsyed (10 µM, 0.1 % catalyst loading) carbonyl olefination reaction between EDA (10 mM) and (**top left**) benzaldehyde (30 mM), (**top right**) *p*-nitrobenzaldehyde (30 mM), (**bottom left**) *p*-cyanobenzaldehyde (30 mM), and (**bottom right**) 3,4,5-trimethoxybenzaldehyde (30 mM) in the presence of PPh_3 (10 mM) (CHES buffer, pH 8.6). An isocratic mobile phase was employed (100% CH_3CN: 0.1% v/v TFA: 0.1% v/v: Et_3N) and injection volumes were 20 µl.

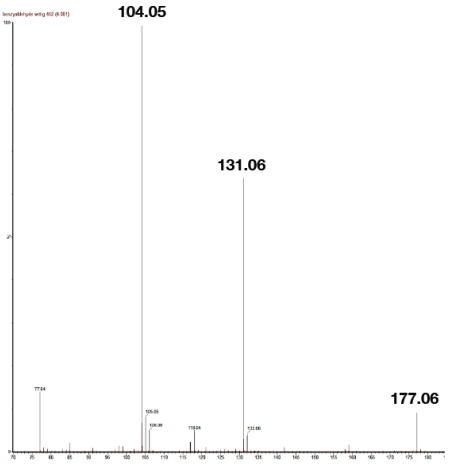


Figure 118: LC-MS spectrum of the reaction product obtained from the C45 catalyzed carbonyl olefination assay between benzaldehyde and EDA in the presence of PPh_3**.** LC-MS spectrum of the C45 catalyzed carbonyl olefination assay products. The mass spectrum was recorded in ES+ mode and monitored at 245 nm. A C8 column was employed for the LC separation with a gradient mobile phase (95:5% H2O:MeCN to 10:90% H2O:MeCN; 0.1% v/v formic acid, 0.25 ml min-1).

Although C45 exhibited activity towards carbonyl olefination, the reactions are consistently characterised by poor product yields and zero-to-negligible selectivity. The poor product yields are undoubtedly a consequence of the need to form an ylide intermediate with the large, bulky PPh_3 prior to reacting with the carbonyl substrate. The binding pocket of C45 would be expected to perform a significant, and energetically expensive, reorganisation to accommodate PPh_3 and the resultant ylide, although favourable π-π interactions between the phosphine phenyl groups and nearby aromatic residues and/or the porphyrin ring could provide some energetic compensation for the reorganisation process. The formation of the ylide is, therefore, most probably the rate-limiting step of the reaction and accounts for the poor yields observed. The poor reaction yields reported in the literature, even after letting the reaction proceed for upwards of 16 hours, supports this conjecture. The absence of any appreciable stereoselectivity can be accounted for in two ways, i) such a significant reorganisation of C45's binding pocket resulting in the lack of any energetically favourable diffusion trajectory for the carbonyl substrate, or ii) dissociation, and expulsion from the binding picket, of the ylide intermediate *prior* to interacting with the carbonyl substrate. The dissociation and expulsion of the ylide into solution would also result in unwanted side-reactions (i.e. solvent hydrolysis) which would also negatively affect the product yields.

From the above experiments it can be concluded that C45 does exhibit some catalytic activity towards carbonyl olefination, but offers no advantage, in terms of yields or selectivity, over established synthetic methodologies. It would be helpful to investigate, in future experiments, if certain key modifications could improve the catalytic performance. For example, employing a less-bulky phosphine substituent (i.e. PMe_3), or by mutating large residues to smaller residues (i.e. valine to glycine) in the binding pocket surrounding the heme to increase the active site volume.

5.2.3 C45 catalysed nitrogen-containing heteroaromatic homologous ring expansion

Nitrogen-containing-heterocycles (NHCs) constitute an enormous class of molecules ubiquitously distributed across all kingdoms of life (**figure 119**). From a biological perspective, the most noteworthy subclass of NHCs are the alkaloids, a diverse group of secondary metabolites that serve an assortment of biologically important functions.[315,325–328] From a technological perspective, the alkaloids are renowned for their various functions, including several important pharmaceutical and therapeutic properties, ranging from antibacterial, antimalarial (i.e. quinine), anticancer (i.e. vinblastine) to vasodilatory (i.e. vincamine), analgesic (i.e. codeine) and antihyperglycemic properties. In addition to direct medical applications, many alkaloids function as molecular blueprints, or starting scaffolds, from which semi-synthetic NHC drugs are developed.[329–333] The application of alkaloids extends beyond medicine, with other alkaloids finding applications in agriculture (pesticides, insecticides etc)

or as stimulants (i.e. caffeine, nicotine). Many alkaloids are also rather infamous for their psychoactive properties, including stimulants (i.e. cocaine), hallucinogens (i.e. psilocybin, ibogaine), although some psychoactive properties have direct medical applications, such as the painkillers morphine and codeine.

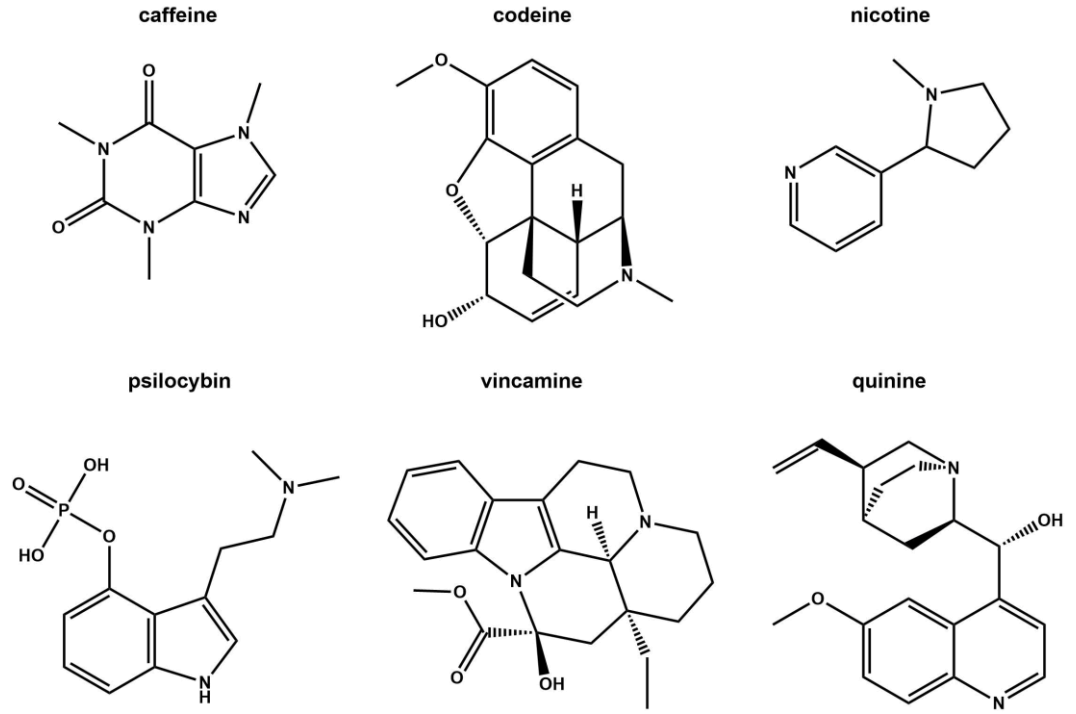


Figure 119: Some important nitrogen-containing heterocycles

The ubiquitous distribution of alkaloids, and the assortment of technologically relevant properties, has naturally resulted in the isolation, extraction and synthesis of alkaloids, semi-alkaloids and other NHCs semi-related to alkaloids becoming pertinent areas of scientific research.[325–331,333] Although many different synthetic methodologies are employed in the synthesis of NHCs, ring expansion reactions are exceptionally useful as they provide a reliable and facile method for acquiring large, expanded ring systems.[334,335] Though currently underused, the homologous ring expansion of nitrogen-containing heteroaromatics (NHAs) could be of considerable use in the synthesis of naturally occurring or semi-synthetic products. One of the most well-established ring expansion reactions is the Buchner ring expansion (**figure 120**): a homologous expansion of a benzene ring to a seven-membered cyclohepta-1,3,5-triene by the insertion of a reactive carbene species.[336,337] Although the reaction does not contain any NHCs or NHAs, it serves as a good introduction to ring expansion chemistry. The reaction proceeds in two steps: i) the initial decomposition of a diazo compound, expelling nitrogen and generating a reactive carbene intermediate. The decomposition of the diazo precursor can be achieved *via* thermolysis or photolysis, or by the application of an organometallic catalyst. The reactive

carbene then undergoes a concerted [2+1] cycloaddition to one of the benzene π-π bonds to yield a bicyclo[4.1.0]heptadiene.[336,337]

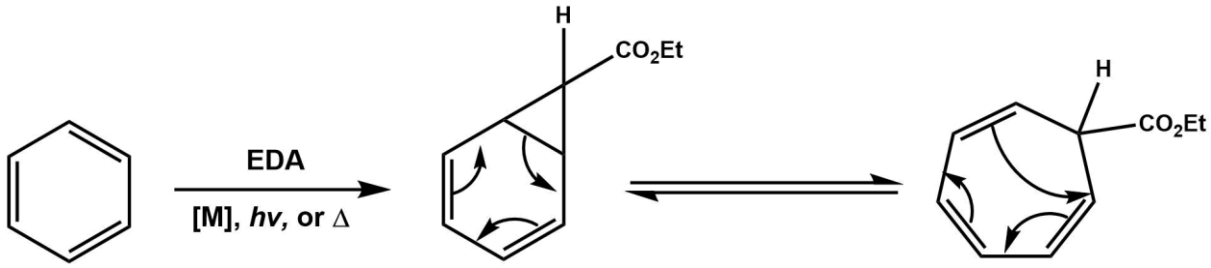


Figure 120: The Buchner ring expansion

The product of the Buchner reaction is a bicyclo[4.1.0]heptadiene species. The cyclopropane fragment of the products is susceptible to ring opening *via* an electrocyclic reaction to yield a cyclohepta-1,3,5-triene, the 7-membered ring expansion product. In solution, the bicyclo[4.1.0]heptadiene and cyclohepta-1,3,5-triene species form a chemical equilibrium, with the position of the equilibrium dependent on steric, electronic, and conformational effects. The equilibrium can be destabilised in favour of the bicyclo[4.1.0]heptadiene by the inclusion of bulky substituents and/or electron-donating groups, while small and/or electron-withdrawing groups stabilise the cyclohepta-1,3,5-triene product and shift the equilibrium towards the ring expansion product. One major limitation of the Buchner ring expansion is the lack of control over the regioselectivity process, meaning multiple regioisomers can feasibly exist in solution at any given time.[336,337]

While the Buchner ring expansion reaction is a good example of a homologous ring expansion, the starting substrate is not an NHC or NHA. The Ciamician-Dennstedt rearrangement, though not as well known, is a good example of a ring expansion reaction involving an NHA species (**figure 123**).[338–340] The original reaction is between the 5-membered NHA pyrrole and chloroform to yield a 6-membered NHA 3-halopyridine product. The reaction is activated by a strong base (i.e. KOH), thermolysis and/or photolysis, which decomposes chloroform to yield the highly reactive carbene species dichlorocarbene. The dichlorocarbene then undergoes a concerted [2+1] cycloaddition reaction with one of pyrrole's π-π double bonds to yield a cyclopropane-containing-bicyclic species, analogous to the mechanism for the Buchner ring expansion reaction.[338–340] If chloroform is used as the carbene precursor, the cyclopropane-moiety of the bicyclic intermediate is functionalised with two chlorine substituents, which are good leaving groups. The cyclopropane-containing-bicyclic species is susceptible to a spontaneous concerted ring opening reaction, which results in the elimination of HCl to yield 3-chloropyridine. The expulsion of HCl from the cyclopropane-containing-bicyclic species to yield 3-chloropyridine is a thermodynamically favourable process and is highly irreversible, meaning

unlike the Buchner ring expansion reaction, the Ciamician-Dennstedt reaction gives a single product and is not an equilibrium in solution.[338–340] The molecular symmetry of pyrrole also renders the reaction regiospecific, yielding the 3-functionalised pyridine (the regiospecificity disappears if an asymmetric pyrrole substrate is employed in the reaction).

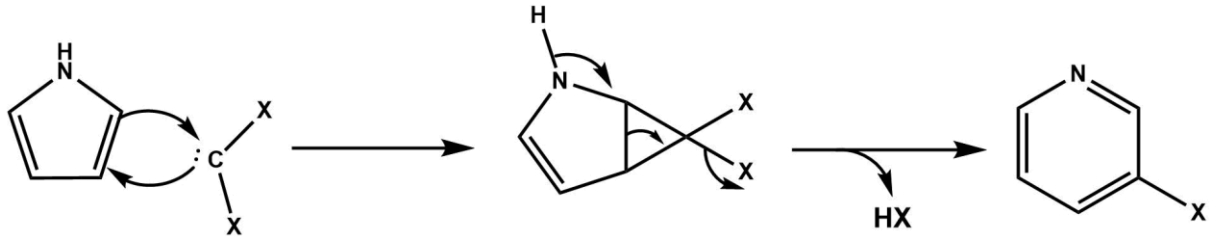


Figure 121: The mechanism of the Ciamician-Dennstedt rearrangement following the [2+1] cycloaddition of a carbene to a nitrogen-containing heteroaromatic molecule (i.e. pyrrole).

It has been demonstrated that C45 can catalyse cyclopropanation reactions between olefins and diazo compounds. It was therefore hypothesised that C45 could catalyse the cyclopropanation of a nitrogen-containing heteroaromatic to produce a cyclopropane-containing bicyclic system. If a suitable halogen-substituted carbene precursor was employed to incorporate a leaving group into the cyclopropane-containing bicyclic compound it would render the product susceptible to spontaneous rearrangement to yield a 3-functionalised pyridine product. If demonstratable, this would be the first enzymatically catalysed homologous ring expansion of an NHA species. The carbene precursor ethyl-2-bromo-2-diazoacetate (Br-EDA) was identified as a suitable substrate, on account of Br-EDA being straightforward to synthesise from EDA (chapter 9, *section 9.7.7.1*) and because the bromine atom would function as a good leaving group, which would hopefully facilitate the spontaneous rearrangement of the cyclopropane-containing-bicyclic intermediate.

Initial assays were conducted in i) the absence of any catalyst, and ii) the presence of free hemin (10 µM) for the ring expansion reaction between pyrrole (10 mM) and Br-EDA (1 mM). The reaction mixtures were quenched and then extracted with ethyl acetate prior to analysis. The reaction mixtures were analysed *via* C18-HPLC and the resultant chromatograms for both reactions are presented in **figure 122**. The chromatograms for the ring expansion assays conducted with and without hemin are both dominated by a single major peak, with a retention time of 2.2-2.5 minutes, which corresponds to Br-EDA respectively; no reaction product was detected. The product mixtures were subsequently analysed *via* LC-MS and were also characterised by the lack of any detectable trace of the anticipated product. These assays indicate that no reaction occurs between pyrrole and Br-EDA in the presence or absence of hemin.

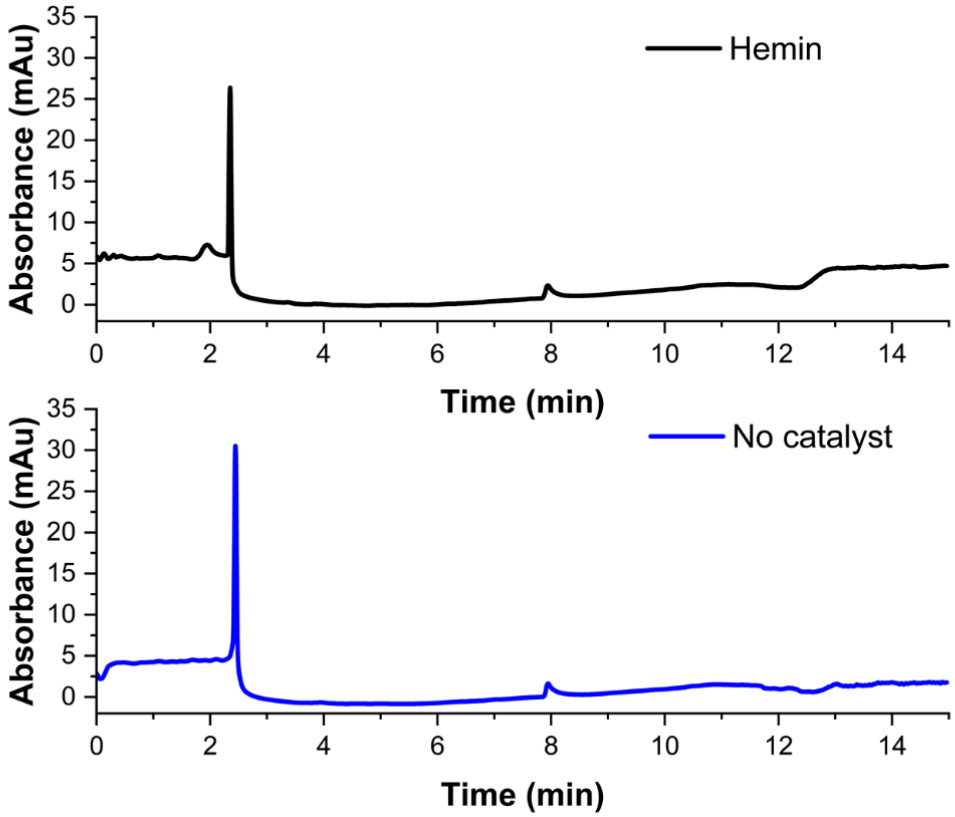


Figure 122: The C18-HPLC chromatogram for the ring expansions reaction between pyrrole (10 mM) and ethyl 2-bromo-2-diazoacetate (1 mM) in the presence (**top**, black, 10 µM, 0.1 % catalyst loading) and absence (**bottom**, blue) of hemin . A gradient mobile phase was employed (70:30% H_2O:CH_3CN to 10:90% H_2O:CH_3CN) and injection volumes were 20 µl.

The anticipated product arising from the homologous ring expansion between Br-EDA and pyrrole is ethyl nicotinate. The C18-HPLC trace of a commercial sample of ethyl nicotinate is dominated by a major peak with a retention time of 8.3-8.6 minutes **(figure 123, bottom)**. The LC-MS spectrum of ethyl nicotinate is presented in **figure 124 (bottom)** and is dominated by major peaks at m/z 106, 124, and 152 corresponding the loss of EtOH, hydrolysis of the ester group, and the $[M+H]^+$ parent ion respectively. An external calibration for ethyl nicotinate is presented in *appendix 1*. Using the same reaction conditions as those described for C45-catalyzed styrene cyclopropanation, it was investigated whether C45 could catalyse the ring expansion of pyrrole with ethyl-2-bromo-2-diazoacetate as the carbene precursor. The C18-HPLC chromatogram obtained from the C45-catalysed ring expansion reaction between pyrrole and Br-EDA is presented in **figure 123 (top)** alongside the C18-chromatogram obtained for commercial ethyl nicotinate under the same conditions **figure 123 (bottom)**. The chromatogram obtained for the C45-catalysed ring expansion reaction between pyrrole and Br-EDA is dominated by a single major peak with a retention time of 8.6 minutes, which overlaps extremely well with the chromatogram acquired for commercial ethyl nicotinate.

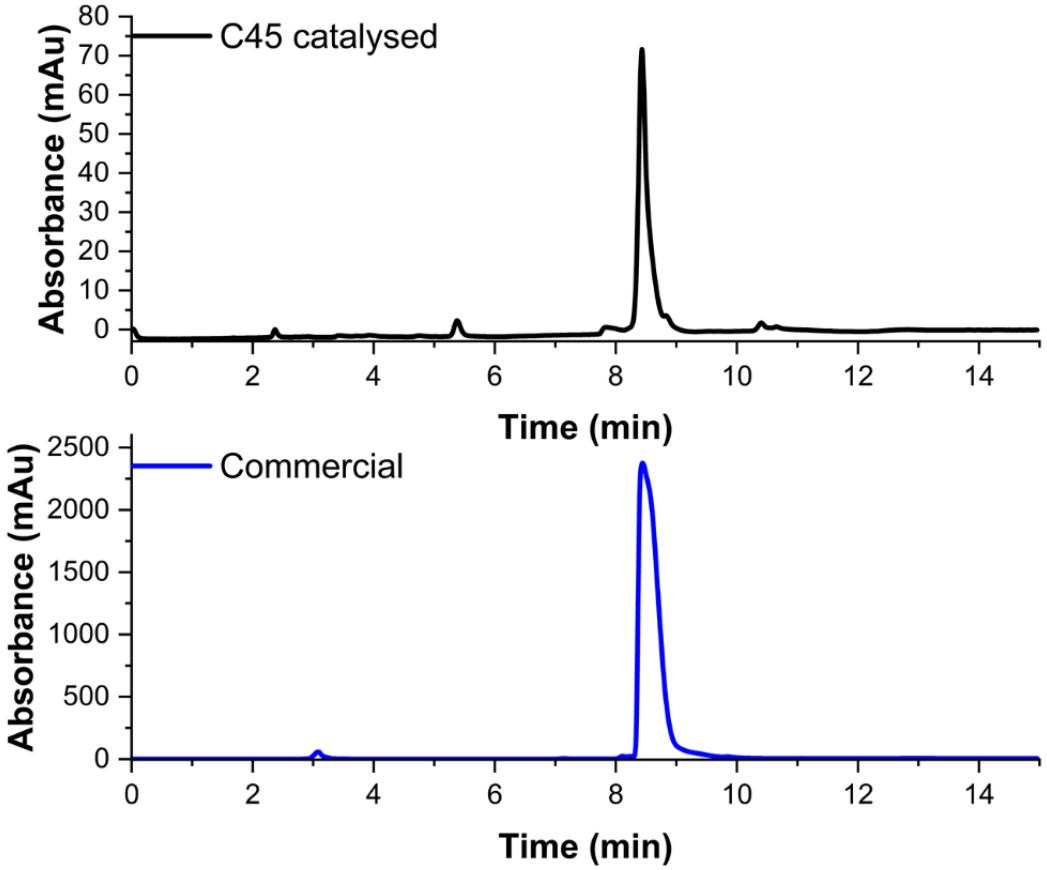


Figure 123: The C18-HPLC chromatogram for the (**top**, black) C45-catalysed (10 µM, 0.1 % catalyst loading) ring expansion reaction between pyrrole (10 mM) and ethyl 2-bromo-2-diazoacetate (1 mM) and (**bottom**, blue) commercial ethyl nicotinate (1 Mm, EtOH). A gradient mobile phase was employed (70:30% H_2O:CH_3CN to 10:90% H_2O:CH_3CN) and injection volumes were 20 µl.

The LC-MS spectrum acquired for the C45-catalysed ring expansion reaction between pyrrole and Br-EDA, is presented in **figure 124 (top)** alongside an LC-MS spectrum acquired for a commercial sample of ethyl nicotinate **figure 124 (bottom)**. The mass spectrum for ethyl nicotinate presents a fragmentation pattern with dominant peaks at m/z 106, 124, and 152 respectively, corresponding the loss of EtOH, hydrolysis of the ester group, and the $[M+H]^+$ parent ion, respectively. The C45-catalysed ring expansion assay is also dominated by major peaks at 106, 124, and 152 m/z respectively, presenting an identical fragmentation profile to the spectrum acquired for commercial ethyl nicotinate. The HPLC chromatogram and LC-MS spectrum acquired for the C45-catalysed ring expansion reaction between pyrrole and Br-EDA strongly suggest that C45 is catalysing the cyclopropanation of pyrrole leading to the bicyclic intermediate, followed by a spontaneous rearrangement reaction, facilitated by the expulsion of HBr, to yield the final product: ethyl nicotinate. From the chromatograms, the average reaction yield was calculated to be 69%, with a TTN of 69 (**table 9**).

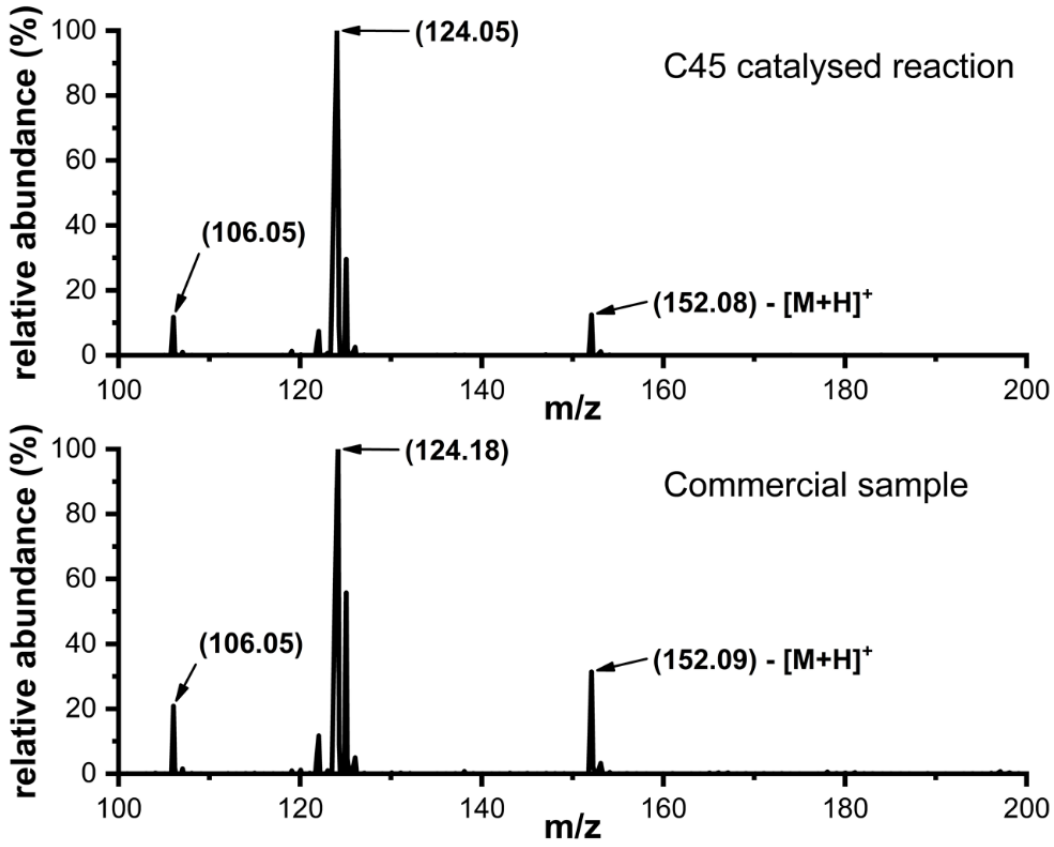


Figure 124: LC-MS spectra for the (**top**) C45-catalysed (10 µM, 0.1 % catalyst loading) ring expansion reaction between pyrrole (10 mM) and ethyl 2-bromo-2-diazoacetate (1 mM) and (**bottom**) commercial ethyl nicotinate (1 Mm, EtOH). All spectra were recorded in ES+ mode and monitored at 254 and 280 nm. A C8 column was employed for the LC separation with a gradient mobile phase (95:5:0.1 % v/v water/MeCN/formate 10:90:0.1 % v/v water/MeCN/formate). Injection volumes were 20 µl.

The C45-catalysed ring expansion of pyrrole to ethyl nicotinate is the first example of an enzyme - natural, engineered or *de novo* - capable of catalysing a homologous ring expansion reaction *via* a carbene transfer mechanism. Interestingly, the product of the reaction, ethyl nicotinate, is an ethyl ester of nicotinic acid (3-pyridinoic acid), biochemically known as niacin, the nicotinamide and NAD(P)H precursor. Subsequent base hydrolysis of ethyl nicotinate was accomplished under basic conditions, with a new trace, corresponding to niacin, appearing in the C18-HPLC chromatogram after 30 minutes of incubation (**figure 125, B, trace 1)**. The ability to synthesise ethyl nicotinate using C45, and subsequently hydrolysing ethyl nicotinate to niacin, raises the possibility of engineering a life sustaining, artificial biochemical pathway from pyrrole to nicotinamide that is reliant on the *in vivo* activity of C45. To further explore the possibility of employing C45 in a new, essential, and life-sustaining pathway from pyrrole to the pyridine nucleotides, the ring-expansion of pyrrole to ethyl

nicotinate in living *E. coli* cells was examined. Using a well-established procedure for carbene transferase activity under such conditions, the ability of *E. coli* bearing the C45-expression plasmid and pEC86 (harbouring the c-type cytochrome maturation apparatus) to perform the ring expansion reaction of pyrrole with ethyl-2-bromo-2-diazoacetate under anaerobic conditions was investigated. Since the maturation apparatus is constitutively expressed and results in the production of several heme containing membrane proteins (CcmC, CcmE & CcmF),[270] 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 (**figure 125, B, traces 3 and 4)**. In contrast, cells that were grown for 3 hours, induced with IPTG and grown for a further 3 hours displayed significant ethyl nicotinate formation with a total yield of 76 % (**table 9**), indicating that under these conditions and *in vivo*, C45 exhibits the catalytic ring expansion activity (**figure 125, B, trace 5)**. It also indicates that both pyrrole and ethyl-2-bromo-2-diazoacetate can cross the outer membrane and access the periplasmically-located C45.

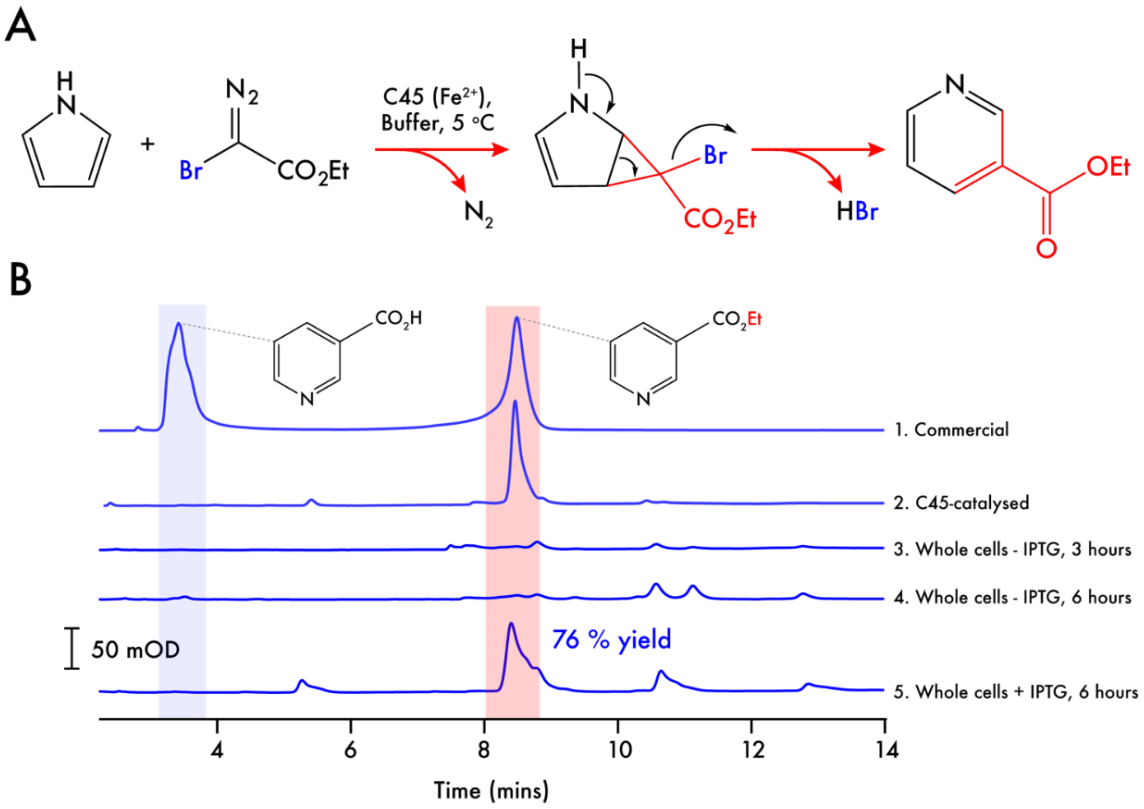


Figure 125: A. Reaction scheme for the ring expansion strategy using ethyl 2-bromo-2-diazoacetate, pyrrole and ferrous C45. Following carbene transfer to the pyrrole, spontaneous rearrangement of the bicyclic ring system leads to elimination of HBr and formation of a 6-membered pyridine ring. **B.** C18 reversed phase HPLC traces of the C45-catalyzed ring expansion of pyrrole to ethyl nicotinate (1% catalyst loading (10 µM C45), 10 mM sodium dithionite, 1 mM ethyl 2-bromo-2-diazoacetate, 10 mM pyrrole in 100 mM KCl, 20 mM CHES, pH 8.6, 5% EtOH). **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 **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.

Table 9: The product yields, total turnover numbers, reaction time, and ITPG parameter for the C45-catalysed ring expansion reaction between pyrrole and Br-EDA *in vivo* and *in vitro.*

Substrates (vs EDA)	% Yield	IPTG	Time	TTN
Pyrrole (free enzyme)	69.41 (± 0.20)	n/a	n/a	69
Pyrrole (whole cell)	0	-	3 hours	n/a
Pyrrole (whole cell)	0	-	6 hours	n/a
Pyrrole (whole cell)	76.10	+	6 hours	n/a

Interestingly, despite the production of ethyl nicotinate by the cells overexpressing C45, there was no observable increase in the concentration 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. In order to determine whether esterase activity could be accomplished in *E.coli* post-ring expansion, a recombinant esterase, *Bacillus subtilis*, that had been expressed in, and purified from *E. coli*, was identified as a suitable candidate for screening for ethyl nicotinate hydrolysis activity.[341] Incubation of this esterase with ethyl nicotinate under near-physiological conditions results in the production of niacin **(figure 126, left)**, highlighting another functional part in a potential novel biosynthetic pathway to pyridine nucleotides from pyrrole through an intermediary niacin.

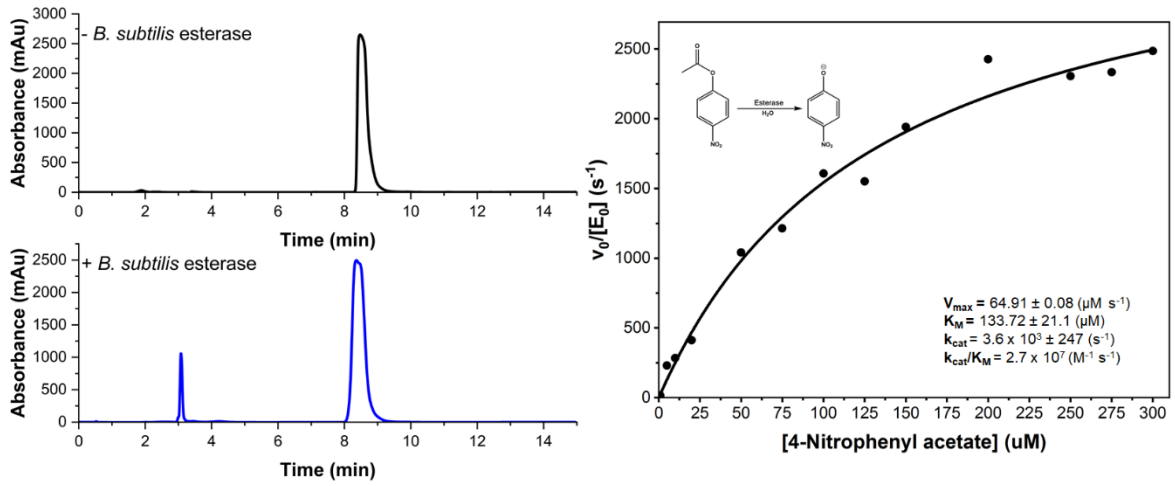


Figure 126: C18-HPLC chromatograms of ethyl nicotinate and esterase-hydrolysed ethyl nicotinate. (**top left**) commercial ethyl nicotinate (200 µl of 5 M stock in DMSO, 19.8 ml CHES buffer, pH 8.6) in the absence of esterase and (**bottom left**) 1 hour after the addition of 2 mg esterase (final esterase concentration is 100 µg/ml, 19.8ml CHES buffer, pH 8.6). The mixture was analysed directly after precipitating the esterase with 3 M trichloroacetic acid. A reverse phase gradient mobile phase (70:30% H_2O:MeCN to 10:90% H_2O:MeCN) was employed and injection volumes were 20 µl; traces were recorded at 265 nm. (**right**) Michaelis-Menten plot and kinetic parameters for the *B subtilis* esterase catalysed hydrolysis of 4-nitrophenylacetate; initial rates were calculated from the change in absorbance of the 4-nitrophenyl acetate anion.

It was speculated 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, then NaAD prior to NAD^+ formation (**figure 127**):[342–345] i) the *de novo* pathway, using L-aspartate as a starting material to produce quinolate which is a NaMN precursor,[343] and ii) the pyridine ring salvage pathway where exogenous niacin or nicotinamide is utilised instead as NaMN precursors, representing the favoured pathway when niacin and nicotinamide are abundant in the environment.[345] Knocking out a key enzyme in the aspartate pathway - e.g. nicotinate-nucleotide diphosphorylase (NadC) - would engender an auxotrophic strain of *E. coli* that, when expressing both C45 and a suitable periplasmically-directed esterase (using a signal sequence such as malE as for C45)[234] in nicotinamide and niacin-lacking media, could harbour the potential for 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 has not been properly investigated, but this strategy highlights the mechanism by which a novel and rationally designed metabolic pathway, reliant on a *de novo* enzyme, to sustain an essential pathway could be engineered in a bacterium (**figure 127**).

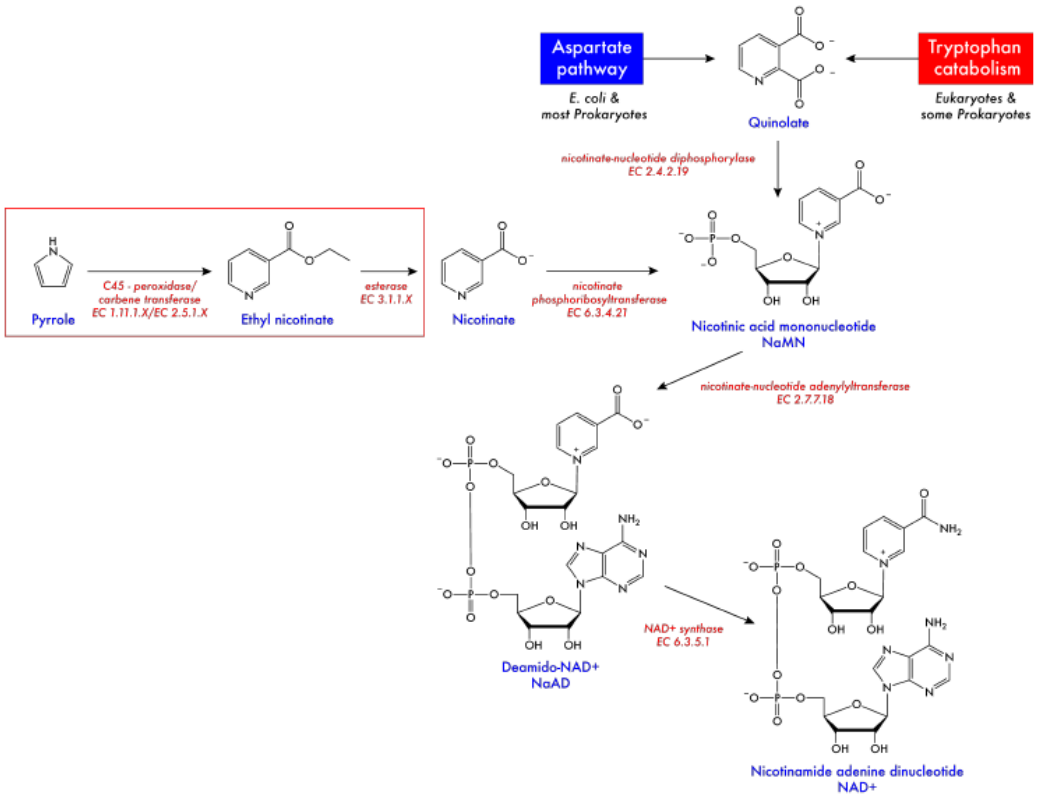


Figure 127: Natural and engineered biosynthetic pathways to NAD+. Steps catalyzed by the *de novo*-designed enzyme C45 and the non-native esterase from *B. subtilis* are displayed in the red box, showing an alternative route from pyrrole to nicotinate. It is hypothesised that deleting the *E. coli* nicotinate nucleotide diphosphorylase and growing esterase- and C45-expressing cells under nicotinate starved conditions with added 2-bromo-2-diazoacetate would result in the life-sustaining biosynthesis of NAD^+. Data and annotations in the figure were obtained from the KEGG database (https://www.genome.jp/kegg/ kegg1.html) and the NC-IUBMB database (http://www.sbcs.qmul.ac.uk/iubmb/enzyme/).

5.3 CONCLUSION

It was demonstrated in chapter 3 that C45 can form stable, heme-localised metallocarbenoid intermediates and it was subsequently demonstrated in chapter 4 that C45 can catalyse the carbene transfer reaction to receptive olefins to yield cyclopropane products. It has been demonstrated that the carbene transferase enzymatic activity of C45 extends beyond cyclopropanation and includes X-H insertion (X = N, Si), chemoselective aromatic N-H insertion, carbonyl olefination (when combined with PPh_3) and the *in vitro* and *in vivo* homologous ring expansion of nitrogen-containing heteroaromatics, the latter being the first enzymatic ring expansion – natural, designed or *de novo* – that proceeds *via* a carbene transfer mechanism. The homologous ring expansion of pyrrole to ethyl nicotinate, catalysed by *E.coli* overexpressing C45, opens up the possibility of engineering a novel biosynthetic pathway to niacin, from pyrrole, which depends on the *in vivo* catalytic activity of a *de novo* carbene

transferase. C45 is not only promiscuous with respect to substrates and reactions, but is also remarkably proficient at the reactions investigated, producing products with up to 90% yields. In addition, unlike what was observed in the cyclopropanation assays, C45 appears to be more accommodating towards bulkier diazo precursors, such as ^{t}BuDA and BnDA, in the N-H insertion reactions; in fact, C45 catalysed the chemoselective N-H insertion reaction between 3-aminophenol and ^{t}BuDA with a product yield of >99%. Overall, the inherent simplicity underpinning the maquette approach to enzyme design, compared to its naturally designed counterparts, renders C45 a versatile, catalytically proficient, and promiscuous enzyme capable of performing an assortment of difficult chemical transformations *in vitro* and *in vivo via* several abiological bioorganometallic intermediates.

CHAPTER 6: NON-AQUEOUS ENZYMOLOGY - ALGINATE-ENCAPSULATED *DE NOVO* HEME-CONTAINING MAQUETTE CATALYSED CARBENE TRANSFERASE CHEMISTRY

Industrial biotechnology is the utilisation of biocatalysts in chemical transformations of industrial importance. Several enzymes have well established reputations as industrial biocatalysts, being used in food processing, pharmaceutical synthesis, biofuel synthesis and materials manufacturing.[83,346–350] Some noteworthy examples include i) xylose isomerase, an isomerase that catalyses the interconversion of D-xylose and D-xylulose and is used in the production of high fructose corn syrup,[85] ii) cellulase, an enzyme that decomposes cellulose to glucose monomers and is used in biofuel production,[346] and iii) penicillin amidase, an enzyme that catalyses the hydrolysis of the amide bond in natural penicillins and is used in the synthesis of semi-synthetic penicillins.[346,351] Enzymes offer several advantages to the production processes when compared to organometallic-based catalysts. A major advantage is the gentle operating conditions required for enzymes to adequately function.[83,346] Most enzymes function best in aqueous conditions at standard atmospheric temperature and pressures, accounting for optimal pH and the inclusion of any necessary cofactors/coenzymes. Although considered an advantage from an environmental and biotechnological perspective, the gentle operating conditions of most enzymes is also limiting the widespread adoption of biocatalysts because slight deviations away from optimal conditions is often detrimental to enzymatic function.[83] This is a severe hinderance in industrial biotechnology as many industrial processes use reaction conditions inhospitable to enzymes, such as high temperatures or high organic solvent contents. A wealth of research has been devoted to engineering enzymes with enhanced thermal stability, inspired by the discovery of several thermostable enzymes from extremophilic microorganisms.[352,353] However, efforts to engineer enzymes with enhanced stability towards organic solvents is still lacking. Several proteins, mainly lipases and select enzymes isolated for extremophiles,[354–357] have remarkable activity in organic solvents, but the vast majority of enzymes are either too insoluble, catalytically inadequate, or both, in organic solvents and exhibit poor catalytic activities compared to aqueous systems.

From a biotechnological perspective, the advantages of non-aqueous enzymology include i) the increased solubility of organic compounds in organic solvents, ii) the suppression of unwanted side reactions, such as hydrolysis, and shifts in chemical equilibriums in favour of desired products, iii) the ease of product extraction and isolation, and iv) the ease of catalyst recoverability relative to water.[356–359] Interestingly, some counterintuitive advantages have also been reported including i) the enhanced thermostability of enzymes in organic solvents, ii) the emergence of novel reactivity unobserved in aqueous solutions, iii) changes in selectivity associated with the physical and chemical properties of

the solvent medium, and iv) the lack of pH dependence in organic solvents.[356–360] These advantages, combined with most industrial reactions using solvents inhospitable to enzymes, means it is pertinent to understand how organic solvents perturb an enzyme's structure, dynamics and function, and to developed chemical strategies to counteract such perturbations in order to stabilise enzymes in organic solvents, and elevate the catalytic activities towards industrially useful numbers.

6.1 ENZYMATIC DEACTIVATION IN ORGANIC SOLVENTS

The majority of enzymes have drastically reduced catalytic activities in organic solvents relative to water.[361] Despite this, the aforementioned advantages of non-aqueous enzymology mean overcoming such barriers are of practical importance. To develop novel methodologies for stabilising enzymes in organic solvents it is necessary to understand how organic solvents interact with enzymes. From a molecular perspective, this question concerns how organic solvents influence protein-protein and protein-water interactions, how organic solvents stabilise/destabilise protein conformations, and how these factors translate to a loss of enzymatic activity. Numerous factors have been identified in governing enzyme functionality in organic solvents, including conformational changes,[356,359–362] protein rigidification,[356,359,360] perturbation of essential noncovalent protein-water interactions,[360,362,363] changes in protein-protein noncovalent interactions,[360] substrate solvation thermodynamics,[356,361] substrate partitioning,[361] solvent penetration into the active site,[356] and the modification of activation energy barriers.[356,361]

A tertiary structure of a protein is stabilised by an intricate network of intraprotein noncovalent-interactions and intermolecular noncovalent-interactions with water molecules. The network of noncovalent hydrogen bonds is essential to the structure, confirmation, and functionality of a protein. It is suggested that water provides a ''lubricating'' effect to enzymes, affording conformational flexibility for optimal catalysis;[356,360] in addition, water molecules also play crucial roles in reaction mechanisms, either directly (e.g. lipases, hydrolases)[364] or indirectly as an acid/base catalyst (e.g. cytochrome P450 compound I formation).[101] Furthermore, water molecules can also provide stability to transition states and reaction intermediates which can lower the effective activation barrier to catalysis.[356,361] Given the importance of water to protein dynamics and catalysis, it is hardly surprising that enzymes are extremely sensitive to any changes in the external environment which perturb its hydrogen bonding network. This includes deviations from optimal pH, elevated temperatures and, in particular, the introduction of organic solvents.

The nature of protein-solvent interactions, and their role in enzymatic function and structure, has received a lot of attention. A study of α-chymotrypsin in mixed water/alcohol, water/glycol and

water/formamide solvent systems found that enzymatic activity was unaffected by the presence of organic cosolvents until a critical concentration was reached, after which even minute additions of the cosolvent proved catastrophic to enzymatic activity.[365] The authors reported a relationship between the critical concentration of residual water (where enzyme inactivation occurs) and the hydrophobicity of a given alcohol solvent.[365] Interestingly, the loss of activity for α-chymotrypsin after surpassing the critical concentration of organic solvent was reversible, with a complete regeneration of activity reported for α-chymotrypsin suspended in alcohol, diol and formamide solvent systems.[365] A similar observation was reported for the hydrolytic activities of three different β-galactosidases dissolved in various 50% (v/v) organic solvent–buffer mixtures.[366] These observations are attributed to a conformational change in the protein, proposing that protein-solvent interactions perturb the solvation shell of enzymes, strip essential water molecules from the active site and strengthen intraprotein noncovalent-interactions; the relationship between the hydrophobicity of a solvent and enzymatic activity supports these claims.[366] It is thought that hydrophilic solvents ''strip'' water molecules from the protein's interior to a greater extent than hydrophobic solvents, while also strengthening the intraprotein hydrophobic interactions.[356,360] Furthermore, the lower dielectric constant of most organic solvents compared to water strengthens the intraprotein electrostatic interactions.[360] The two phenomena result in a rigid, non-flexible protein with less conformational stability. As conformational flexibility is essential for catalysis, the rigidification of an enzyme impairs its function and is one factor accounting for the observed decrease in enzymatic activity in organic solvents. This conclusion is reinforced from reports that additions of small quantities of water to a deactivated enzyme in organic solvents results in a drastic recovery of the catalytic activity and from the elevated thermostability of proteins in organic solvents. A study of moisture content and catalytic activity (k_{cat}/K_M) for α-chymotrypsin hydrolysis revealed a direct correlation between the two factors, in the order octane > toluene > THF > acetone >> pyridine.[367] The addition of 1.5% of water to acetone (v/v) elevated the moisture content of α-chymotrypsin from 1.2 to 2.4% and was accompanied by a 1000-fold recovery in enzymatic activity.[367] The authors concluded that solvent-water interactions were more important to overall catalytic function than protein-solvent interactions. In another study of three different lipases (porcine pancreatic, yeast, and mold) it was shown that all three lipases could efficiency catalyst the transesterification reaction between tributyrin and heptanol in a host of organic solvents.[359] All three lipases obeyed Michaelis-Menten kinetics in the presence of only 0.02 % (v/v) of water.[359]

The conformational changes caused by protein-solvent and solvent-water interactions play an important role in the catalytic activities of enzymes in organic solvents. A governing factor is the influence of organic solvents on protein-water interactions, with hydrophilic solvents exerting a more

pronounced effect on catalytic activities compared to hydrophobic solvents.[356,360] Another factor is solvation thermodynamics, in particular the solvation of the substrate, transition states, and reactive intermediates.[356,361] In an aqueous solution, the substrate is solvated by water molecules to form a solvation shell; the enthalpy of solution, ΔH_{sol}(substrate), is the thermodynamic property that accounts for the change in internal energy of a system following the dissolution of a substrate into a solution, at constant pressure and infinite dilution. The enthalpy of solution can be broken down into the sum of three separate components, the endothermic breaking of solute-solute bonds, the endothermic breaking of solvent-solvent bonds, and the exothermic formation of new solute-solvent bonds (solvation). For organic substrates, diffusion into the active site of an enzyme is accompanied by a breaking of the substrate-solute bonds and the establishment of substrate-enzyme bonds in the hydrophobic core.[368,369] The more hydrophobic the substrate is, the greater the enthalpy change associated with desolvation and diffusion into the active site. The desolvation of hydrophobic solvents/substrates is accompanied by a large entropic change, as the removal of a hydrophobic substrate from an aqueous network liberates the water molecules that were restricted to the solvation sphere of the hydrophobic substrate.[368,369] The net result is that diffusion of the substrate into the hydrophobic cavity of the enzyme is a thermodynamically favourable process, which facilitates catalysis.[369] In organic solvents, both the enthalpy and entropy associated with an organic substrate diffusing into the active site is significantly lower as the perturbation to the solvent-solvent bonds caused by the presence of the organic substrate is significantly lower compared to water. The solute-solvent interactions are also stronger, resulting in a less positive enthalpy of solution for the process. The decrease in both the entropic and enthalpic incentive for substrate diffusion into the active impairs catalysis, and is a factor accounting for the decreased catalytic activities observed by enzymes in organic solvents. Finally, water molecules often establish stabilising noncovalent interactions with transient transition states and reaction intermediates.[368] These interactions lower the overall activation energy of catalysis by reducing the energetic gap between the ground state and the transition state/reaction intermediate. The removal of crucial water molecules in the presence of organic solvents elevates the activation barrier and impairs enzymatic function.[356,360]

The molecular origins of enzyme instability in organic solvents are associated with conformational changes, protein rigidification, and solvation thermodynamics. These phenomena all present unique challenges to performing biocatalysis in non-aqueous systems. To address these problems, research has focused on developing novel chemical strategies to stabilise enzymes in non-aqueous conditions.

6.1.2 Methods for stabilising enzymes in organic solvents: immobilisation

The methods for stabilising enzymes in organic solvents can be separated into three categories: i) enhanced solvent stability through genetic modification, ii) the chemical modification of the solvent environment to lessen the deleterious effects on the protein, and iii) immobilisation techniques.[370] Methods i) and ii) are important areas of research; however, the remainder of this chapter will be exclusively dedicated to method iii). Immobilisation of an enzyme can be broken down into two separate approaches: i) immobilisation of an enzyme onto an inert surface, or ii) encapsulation of an enzyme within an inert matrix (**figure 128**).[370–372]

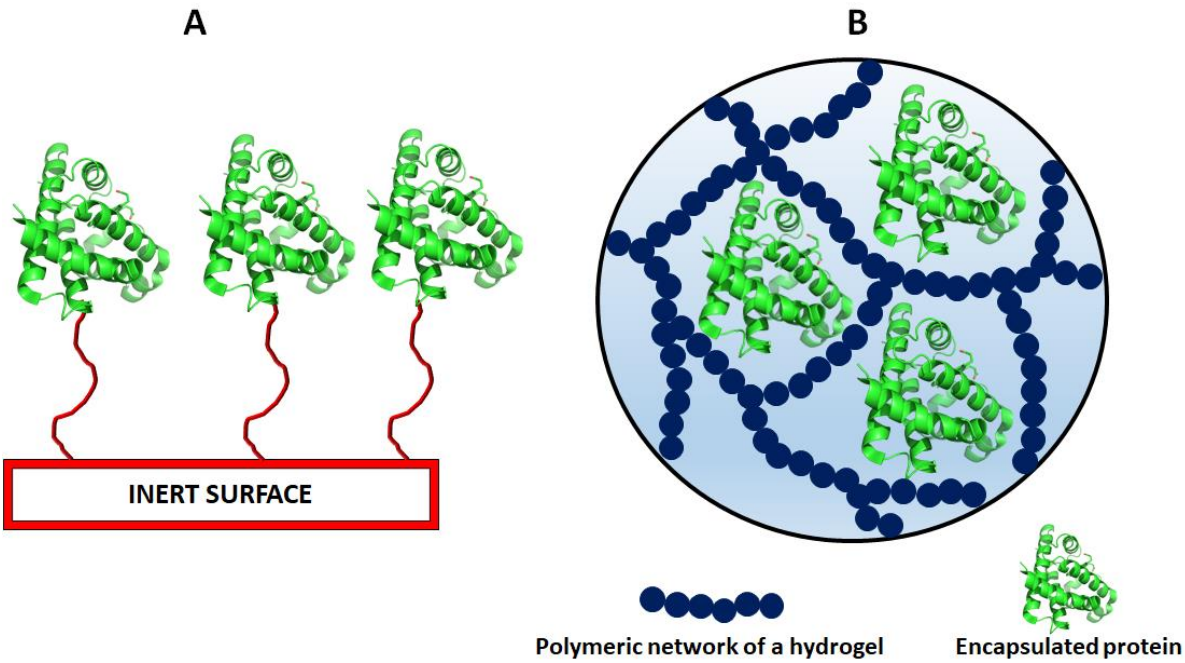


Figure 128: The two major methods of enzyme immobilisation. **A)** The tethering of an enzyme to an inert surface, and **B)** the encapsulation of an enzyme inside an inert matrix.

Immobilisation of an enzyme onto an inert surface can be achieved either *via* non-covalent interactions (adsorption) or covalent binding. Physical adsorption is the simplest strategy for attaching an enzyme to an inert supporting material and can be accomplished by mixing the enzyme with a suitable inert support.[370] For example, several lipases non-covalently immobilised on celite were proficient at the enzymatic interesterification of phosphatidylcholine when suspended in toluene.[373] The authors reported the immobilised enzyme resulted in a 50% regioselective conversion of various fatty acids in the presence of heptadecanoic acid.[373] In another study, immobilised lipase B (from *Candida antarctica*) was employed as the catalyst in the synthesis of (*R*)-α-monobenzoyl glycerol. Immobilized-lipase B catalysed the asymmetric esterification of glycerol and benzoic anhydride in 14

different solvents with reaction yields and enantioselectivities of up to 80 and 50% respectively.[374] Lipase type B, immobilised on a microporous acrylic resin, was also used in the synthesis of 2-substituted-1,3-propanediols.[375] The synthesis was later scaled-up, and afforded kilogram quantities of SCH51048 to be acquired using immobilised lipase type B in pure acetonitrile.[375] A nanoporous sol-gel glass structure was investigated by Wang and Davison as a possible support structure for covalently immobilising enzymes.[376] The authors covalently attached α-chymotrypsin to the glass using a bifunctional trimethoxysilylpropanal ligand; the advantage of the bifunctional ligand is the enzyme can be immobilised *via* multiple covalent linkages instead of the mono-attachment typical of flat surfaces. The transesterification of *N*-acetyl-*L*-phenylalanine ethyl ester and n-propanol, catalysed by immobilised α-chymotrypsin, was conducted in several organic solvents; a 110-fold enhancement in catalytic activity was reported when the reaction was conducted in hexane.[376]

Another common method for immobilising enzymes is entrapment.[370,377,378] In entrapment, the enzyme is encapsulated inside a polymeric network. The encapsulation of an enzyme inside a polymeric network protects the enzyme from the harmful effects of an organic solvent, although problems with mass transfer and substrate/product diffusion can impair overall catalytic rates. Alcohol dehydrogenase, isolated from *Lactobacillus kefir*, was encapsulated inside a polyvinyl alcohol gel bead and employed as the catalyst in the stereoselective reduction of prochiral ketones in pure hexane.[379] The authors reported the entrapped enzyme exhibited higher thermal stability and converted an assortment of prochiral ketones with product yields ranging from 30-100%. In another report, glucose oxidase was encapsulated within polymeric nanocapsules *via* inverse miniemuslion periphery RAFT polymerisation.[380] The encapsulated glucose oxidase retained high activities (71-100%) relative to free glucose oxidase, retaining its activity even when dispersed in toluene and toluene/^{t}BuOH.[380] In another report, it was demonstrated that glucose oxidase follows Michaelis-Menten kinetics for the oxidation of glucose to gluconic acid when encapsulated inside calcium alginate beads; although with a lower V_{max} and higher K_M relative to free glucose oxidase.[381]

6.2 STRUCTURE OF CALCIUM ALGINATE GELS

From all the possible immobilisation strategies, encapsulating an enzyme inside a calcium alginate polymeric network is the simplest and most cost-effective method. The preparation of alginate-encapsulated (AE) enzymes simply involves dissolving the enzyme of interest in a solution of sodium alginate (3% (w/v)) and pipetting the solution dropwise into a stirring mixture of aqueous calcium chloride (100-500 mM). The sodium and calcium instantaneously exchange and result in the spontaneously polymerisation of calcium alginate into spherical gel beads, encapsulating the enzyme inside the porous network. The beads can then be collected *via* vacuum filtration. Calcium alginate

beads are resistant to the vast majority of organic solvents, rendering them versatile heterogeneous platforms for non-aqueous enzymology applications.

Alginates are polysaccharides isolated from brown algae. The polymeric network consists of a linear chain of D-mannuronic acid and L-guluronic acid monomers connected *via* a (1-4) glycosidic linkage (**figure 129**). The (1-4)-linkage between the monomers results in guluronic monomers adopting a 'buckled' confirmation.[382,383] Alginate gels are water-swollen cross-linked polymeric networks and can, therefore, be classified as hydrogels.[383] Hydrogels can be divided into homogenous and heterogeneous hydrogels depending on the degree chain mobility.[384] In homogenous gels, the backbone chains have a large degree of flexibility.[384] In heterogeneous hydrogels, intermolecular cross-linking between chains results in the formation of junction zones, which provide structural integrity to the hydrogel at the expense of flexibility.[384–386] The buckled conformation of the guluronic acid monomers in the polymeric backbone provides suitable binding sites for divalent cations.[383] The introduction of divalent ions to a solution of alginate results in the formation of junction zones (**figure 129**).[383] Junction zones are structural features of a gel resulting from the interpolymer interactions between two polymer chains.[385,386] In alginate gels, the junction zones are formed from the cross-linking of two alginate chains *via* coordination to a divalent ion.[383] The number of junction zones is, therefore, a function of the concentration of the cross-linking divalent ion and the relative strength of the interactions between an alginate chain and divalent cation.

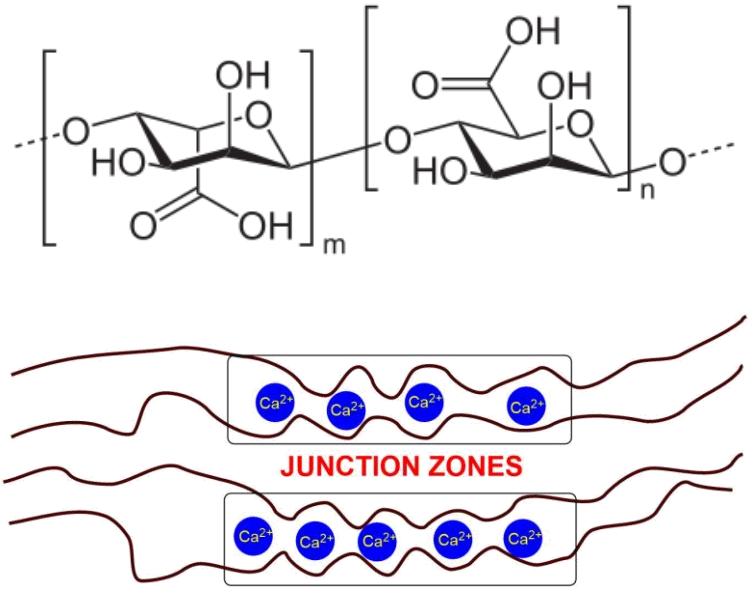


Figure 129: (top) Molecular structure of the repeating monomeric unit of the polysaccharide alginate (figure taken from Wikipedia, https://en.wikipedia.org/wiki/Alginic_acid#/media/File:Algins%C3%A4ure.svg), **(bottom)** A cartoon representation of a junction zone in a calcium alginate hydrogel (black lines = alginate polymer, blue spheres are Ca^{2+} ions).

The defining structural features of hydrogels are the relative pore size and fractional volume of water of the gel.[383] The pore size and fractional volume of water are determined predominantly by the number of junction zones formed by interpolymer interactions, which in alginate hydrogels is governed by the cross-linking divalent ion.[382,383] Homogenous hydrogels naturally have fluctuating pore sizes and water fractions on account of their greater backbone flexibility. In heterogeneous hydrogels, such as calcium alginate, the pore size and fractional volume of water remain constant.[303] In heterogeneous hydrogels, increasing the cross-linking between chains (by increasing Ca^{2+} concentration) reduces the average pore size and fractional volume of water.[303] How changes in pore size and fractional volume of water affect the stability and catalytic activity of an encapsulated enzyme is an important question to consider when preparing encapsulated enzymes for catalytic applications.

In addition, the density of interpolymer interactions also governs the transport of solutes from the bulk solvent phase into and out of the hydrogel.[382,383,387] The magnitude of solute transport is a significant parameter governing the enzymatic activity of encapsulated enzymes; it was reported above that glucose oxidase encapsulated inside calcium alginate retained its catalytic activity, although with significantly reduced catalytic efficiency because of substrate diffusion.[381] It is therefore also imperative to establish methods for quantifying the relationship between solute transport and the observed enzymatic activity of an encapsulated enzyme.

6.3 MASS TRANSFER AND DIFFUSION OF SUBSTRATES THROUGH GELS

In homogenous catalysis, the catalyst (enzyme) and substrate(s) are in the same phase. The encapsulation of the enzyme inside a calcium alginate bead and subsequent suspension of the insoluble beads in a solvent creates a heterogeneous system, where the enzyme and substrates occupy different phases (at t = 0). In the alginate catalyst system, the two phases are the bulk-organic-phase (BOP) and the aqueous-alginate-phase (AAP) (**figure 130**); the two phases being separated by the bulk-alginate-interface (BAI) (**figure 130**). The substrate is initially dissolved in the BOP, and therefore for catalysis to occur the substrate must transfer across the BAI and enter the AAP (**figure 130**). This net movement of molecules across a biphasic interface is described as *mass transfer* and is mathematically defined by the theory of *molecular diffusion* using Fick's laws of diffusion.[388,389]

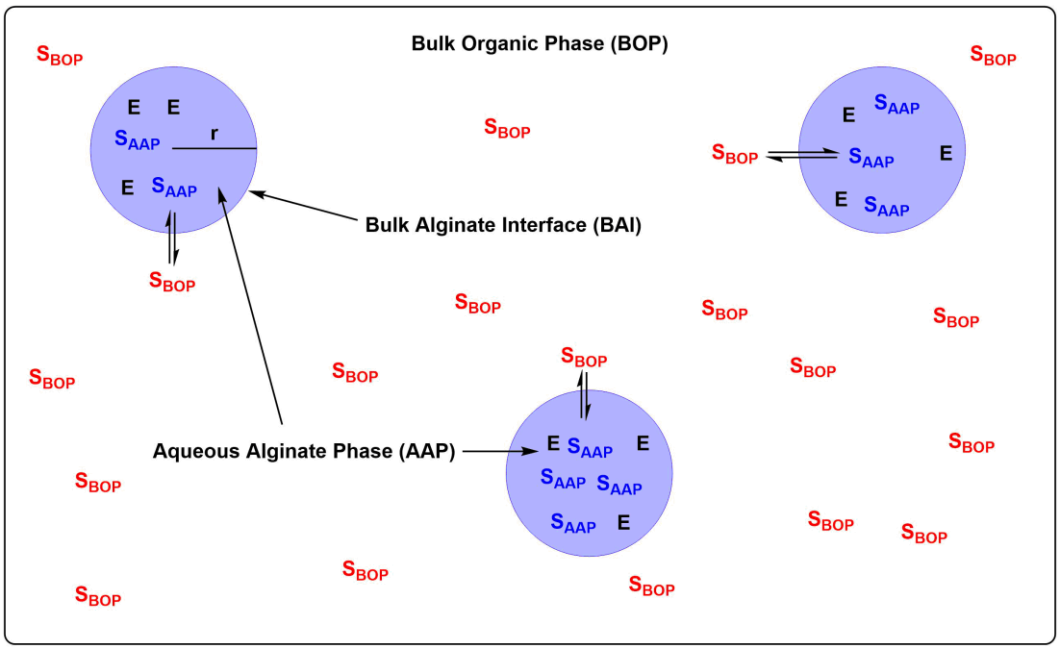


Figure 130: Illustration of substrate diffusion in a heterogeneous system; S_{BOP} = a substrate molecule localised in the bulk organic phase, S_{AAP} = a substrate molecule localised in the aqueous alginate phase, E = an alginate encapsulated enzyme, r = the radius of an alginate bead

In the heterogeneous system described above, the BOP can be labelled phase 1 (P_1) and the AAP system 2 (P_2). At a given time, *t*, the concentration of substrate, S, in P_1 and P_2 is C_{S1} and C_{S2} (when t = 0, C_{S2} = 0). When the system is in a state such that C_{S1} =/= C_{S2} is true a concentration gradient across P_1 and P_2 exists, with the interface located at the BAI; under such conditions the system is in a state of *non-equilibrium*. The chemical potential, μ, is defined as the energy that can be absorbed or realised by a species due to a change in the particle number, N_i, of that given species in the system. Thermodynamically, the chemical potential is the rate of change of the Gibb's free energy of the system, ΔG, with respect to the rate of change of particle number, ΔN_i. Because the chemical potential of species A, $\mu_{i=A}$, defines a rate, it can be mathematically expressed by the partial differential equation

$$\mu_A = \left(\frac{\partial G}{\partial N_i}\right)_{T,P,N_{A\neq B}}$$

Eq 1

And the rate of change of the Gibb's free energy of the entire system consisting of more than one species, at constant pressure and temperature, is expressed as

$$dG = \sum_{i=A}^{n} \mu_i dN_i$$

Eq 2

A system out of equilibrium will tend towards equilibrium, defined thermodynamically as a dynamic system where the Gibb's free energy is at a minimum and is unchanging over time (dG = 0). For the heterogeneous alginate system at time zero the concentration gradient present is thermodynamically defined as $\mu_{P1} > \mu_{P2,}$ where μ_{P1} is the chemical potential of the BOP and μ_{P2} is the chemical potential of the AAP. As systems tend towards energetic minima, a diffusion of substrate A across the BAI will occur until such a point where $\mu_{P1} = \mu_{P2}$ and the system is in thermodynamic equilibrium.[388] At this point the concentration of substrate (S) in the BOP and AAP, defined as C_{S1} and C_{S2}, remains unchanging, with the ratio of the two mathematically defined by the partition coefficient, *K*;[388]

$$K = \frac{[Substrate]_{BOP}}{[Substrate]_{AAP}} = \frac{C_{S1}}{C_{S2}}$$

Eq 3

From a thermodynamic perspective, the partition coefficient contains information on the strength of the solvent(1)-substrate (n_{1s}) and solvent(2)-substrate (n_{2S}) interactions, when solvent(1) and solvent(2) are separated by a well-defined boundary. The rate of mass transport of substrate, S, from P_1 into P_2 as a function of time is expressed by the diffusion equation

$$\frac{\partial \emptyset(\mathbf{r}, t)}{\partial t} = \nabla \cdot [D(\emptyset, \mathbf{r}) \nabla \emptyset(\mathbf{r}, t)]$$

Eq 4

Where $\Phi(\mathbf{r},t)$ is the density of the diffusing material at location **r** and time *t* and $D(\Phi,t)$ is the diffusion coefficient for density Φ at location **r**; ∇ represents the partial differential operator del. If *D* is constant the equation reduces to

$$\frac{\partial \emptyset(\mathbf{r}, t)}{\partial t} = D \nabla^2 \emptyset(\mathbf{r}, t)$$

Eq 5

The diffusion coefficient for a substrate, *D*, is a proportionality constant relating the rate of molecular flux due to molecular diffusion as a function of the concentration gradient of that substrate. The

molecular flux, *J*, defined as the rate at which a molecule crosses a boundary, is mathematically related to the diffusion coefficient by the equation

$$J = -D\frac{dC}{dx}$$

Eq 6

Where *J* is the diffusional flux (quantity of substrate per time), *D* is the diffusion coefficient ($m^2\ s^{-1}$), C is the concentration (mol dm^{-3}) and x is the position (m^2).[388,389] Equation (**6**) is known as *Fick's first law of diffusion* and provides a mathematical description for the driving force of diffusion by relating the rate of diffusion to the concentration gradient ($-\frac{dC}{dx}$). From equation (**6**), it can be deduced that a substrate S will transfer from regions of high concentration into regions of low concentration, until a point where $\mu_{P1} = \mu_{P2}$ is satisfied. The rate of diffusion is a function of a concentration gradient, and so as the concentration gradient is minimised it would be expected that the rate of diffusional flux decreases accordingly. This postulation forms the basis of *Fick's second law of diffusion*, and is expressed mathematically as[389]

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial x^2}$$

Eq 7

In two or more dimensions, equations (**6**) and (**7**) are expressed by equations (**8**) and (**9**)

$$J = -D\nabla C$$

Eq 8

$$\frac{\partial C}{\partial t} = D\Delta C$$

Eq 9

The equations presented above are derived for a system under ideal conditions. The diffusion of a substrate into a gel is a deviation away from ideal conditions, and as such the equations must be modified accordingly. A few considerations need to be discussed to understand why such modifications are justified. Firstly, one way a gel may influence diffusion is by the binding of a substrate, resulting in a situation where the diffusivity of the substrate is characterised by two distinctively different diffusion coefficients, D_{bound} *and* D_{free}. In such a case, the molecular flux would become a function of both diffusion coefficients and would be expressed as[389]

$$J = -D_{bound}\left(\frac{\partial C_{bound}}{\partial x}\right) - D_{free}\left(\frac{\partial C_{free}}{\partial x}\right)$$

Eq 10

Were C_{bound} and C_{free} are the concentrations of bound and free substrate, respectively. In the simplest situations, it is assumed that D_{bound} = 0 and equation (**10**) reduces to the standard equation for Fick's first law. An additional assumption is typically made where the rate of swelling is independent of the diffusion of a substrate into and out of the bead.[389]

The concentration of the substrate can also be defined in two possible ways: i) The concentration of solute per unit volume of gel (C_G), or ii) The concentration of solute per unit void of gel (C_V). However, the concentration of a solution in equilibrium with the gel (C_E), is independent of C_V and C_G. For non-absorbed solutes, C_E is related to C_G through the partition coefficient, K, expressed as[388,389]

$$C_G = (1 - \Phi)C_E$$

Eq 11

The geometric exclusion effect states the fractional void volume available for the centres of large solute molecules (1 – ø) is less than the fractional void volume available for the centres of smaller solvent molecules (1 – β).[389] The porosity of the gel, ε, in terms of the geometric exclusion effect, is defined as (1 – ø); therefore, the porosity of the gel is the constant of proportionality connecting the concentration of solute per unit volume of gel, C_G, to the concentration of the solution in equilibrium with the gel, C_E.[387,389] The molar flux, defined by Fick's first law, can be written as

$$J = -D\frac{\Delta C_G}{\Delta X} = -\varepsilon D\frac{\Delta C_E}{\Delta X} = -D_{eff}\frac{\Delta C_E}{\Delta X}$$

Eq 12

where J is the flux per unit area of gel of thickness ΔX, which separates solutions differing in concentration by amount ΔC_E, ε is the porosity of the gel (which is dependent on the number of junction zones in the gel), and D_{eff} is the *effective diffusion coefficient*[387]

$$D_{eff} = \varepsilon D$$

Eq 13

The equation for Fick's second law of diffusion thus becomes

$$\frac{\partial C_E}{\partial t} = D_{eff}\frac{\partial^2 C_E}{\partial x^2}$$

Eq 14

If the diffusion of a substrate into the spherical bead follows Fickian diffusion, the concentration of substrate inside the bead, C_{S2}, at time, *t*, can be mathematically expressed as[387]

$$C_{S2}(t) = C_{p,\infty}\left[1 + 6(1+\alpha)\sum_{n=1}^{\infty}\frac{e^{\left(\frac{-Dq_n^2 t}{R^2}\right)}}{9 + 9\alpha + q_n^2\alpha^2}\frac{R}{r_p}\frac{\sin\left(\frac{q_n r_p}{R}\right)}{\sin\left(q_n\right)}\right]$$

Eq 15

Where $C_{p,\infty}$ is the substrate concentration inside the bead after an infinite time, r_p is the radial position within the bead, *R* is the radius of the bead, and q_n is the positive non-zero root of the equation[387]

$$\tan(q_n) = \frac{3q_n}{3 + \alpha q_n}$$

Eq 16

Where α is defined by

$$\alpha = \frac{V_b}{V_p K_p}$$

Eq 17

Where V_b is the total volume of the bulk solvent, V_p is the total volume of the alginate beads and K_p is the partition coefficient.[387] If external mas resistance is negligible, equation (**15**) reduces to the non-steady state equation[387]

$$C_{S2}(t) = \frac{\alpha C_{S2}(0)}{1+\alpha}\left[1 + 6(1+\alpha)\sum_{n=1}^{\infty}\frac{e^{\left(\frac{-Dq_n^2 t}{R^2}\right)}}{9 + 9\alpha + q_n^2\alpha^2}\right]$$

Eq 18

Where C_{s2}(t) is the concentration of substrate, *S*, in the bulk at time *t*, C_{s2}(0) is the concentration of substrate, *S*, in the bulk at t = 0, *R* is the radius of the alginate beads, q_n is the position non-zero root of equation (**16**), *t* is time, and *D* is the diffusion coefficient. Equation (**18**) affords a method for readily calculating the diffusion coefficient.

6.4 RESULTS AND DISCUSSION

6.4.1 Chemophysical characterisation of calcium-alginate beads

The synthesis of calcium alginate beads involves the dropwise addition of a 3% (w/v) aqueous sodium alginate solution into an aqueous calcium chloride solution. An ion-exchange between sodium and calcium occurs almost instantaneously, and the calcium alginate chains undergo a spontaneous gelation event to yield spherical, hydrogels. Any substrate dissolved in the sodium alginate solution, such as an enzyme, is entrapped within the newly formed hydrogel, as the rate of gelation upon contact with calcium chloride far exceeds the rate of diffusion through the aqueous calcium chloride solution. It was mentioned above that the physical properties of calcium alginate beads, including the pore size, water retention and strength, are a function of the number of junction zones formed within the alginate polymeric network. In calcium alginate, the number of junction zones is a function of the concentration of calcium chloride, and so the chemophysical properties of a calcium alginate hydrogel can be reasonably tuned by controlling the concentration of calcium chloride.[382,383]

The chemophysical characteristics of calcium alginate beads were first determined prior to conducting any biophysical or catalytic characterisation on C45 encapsulated within an alginate hydrogel (**figure 131**). In particular, the water content and enzyme retention of the calcium alginate beads, as a function of calcium chloride concentration, was investigated. The sodium alginate concentration was chosen to be 3% (w/v), as it was observed that beads made below this concentration were susceptible to degradation in organic media; because calcium ions govern the formation and strength of junction zones, it was not considered necessary to optimise the sodium alginate concentration any further.

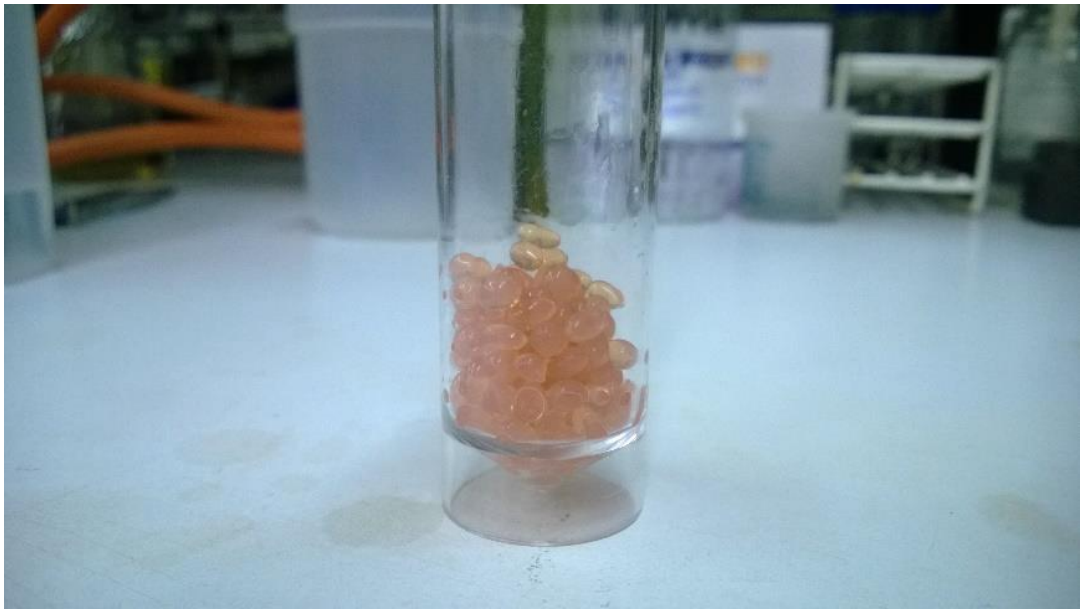


Figure 131: C45 encapsulated inside calcium alginate beads (3% (w/v) alginate, 0.3M $CaCl_2$, CHES buffer, pH 8.6).

The water content and enzyme retention properties of 3% (w/v) alginate hydrogel beads as a function of calcium chloride concentration was initially investigated, with the results tabulated in **table 10**. The water retention was determined by preparing the beads as delineated in chapter 9, *section 9.12*. The beads were then oven-dried at $37^{O}C$ for 2 hours and the diameter was re-measured to determine the change in volume. It can be readily seen from **table 10** that, as expected, the water content of a 3% (w/v) sodium alginate bead decreases linearly with increasing calcium chloride concentration. The overall volume of the bead prior to dehydration is independent of calcium chloride concentration, and so the decreased water content is accounted for by the increased volume occupied by calcium ions inside the alginate beads. As the calcium ions function as the molecular adhesive that establishes junction zones it is probable that a greater proportion of junction zones are formed inside the bead at higher calcium chloride concentrations; an increase in the number of junction zones would reduce the fractional free volume available for water molecules to occupy. Interestingly, although the densities of the calcium alginate beads as a function of calcium chloride were not determined, it was observed experimentally that at values < 0.3M the beads floated to the top of the solvent, whereas for beads > 0.3M the beads sank to the bottom of the suspension. This change in behaviour as a function of calcium chloride concentration indicates that above 0.3M calcium chloride the density of the beads surpasses the density of water, indicative of lower water content inside the alginate beads. The quantity of enzyme retained inside the alginate bead after a certain duration was also measured for two haemoproteins: HRP and cytochrome *c*. As can be seen in **table 10**, the quantity of enzyme retained inside the enzyme after a 24 hour period (as determined by UV-VIS spectroscopy) increased

with increasing calcium chloride concentration, again accounted for by a decrease in pore space and fractional free volume in the bead at higher calcium concentrations.

A calcium chloride concentration of 0.3M was selected to determine the spectroscopic and catalytic properties of C45 encapsulated inside alginate beads (**figure 131**). The rationale is based on two principles: i) the water content of 0.3M $CaCl_2$ is sufficient to provide a hospitable environment to the enzyme while also possessing a sufficient quantity of free volume to accommodate the enzyme, the substrates and the products, ii) at 0.3M calcium chloride, the alginate beads were homogenously distributed when suspended in an aqueous solution (density not measured).

Table 10: The (de)hydrated particle volume, water content/volume, and enzyme retention factors of calcium alginate beads formed from 3% (w/v) sodium alginate (in CHES buffer, pH 8.6) as function of calcium chloride concentration (0.1-0.5M $CaCl_{2(aq)}$.

3% (w/v) Sodium Alginate	0.1M $CaCl_2$	0.2M $CaCl_2$	0.3M $CaCl_2$	0.4M $CaCl_2$	0.5M $CaCl_2$
Particle volume - hydrated (µl ± 0.5)	14.14	14.14	14.14	14.14	14.14
Particle volume - dehydrated (µl ± 0.5)	3.05	4.19	4.85	7.24	11.50
Water content (%)	78.40	70.37	65.70	48.80	18.70
Water volume (µl ± 0.5)	11.10	9.95	9.23	6.90	2.64
HRP retention (% ± 5.1)	11.02	15.25	23.73	27.97	34.32
Cytochrome c retention (% ± 4.3)	19.32	23.59	32.76	40.10	45.60

6.4.2 Spectroscopic characterisation of alginate-encapsulated hemoproteins/maquettes

Initial experiments were concerned with spectroscopically characterising hemoproteins entrapped in alginate beads suspended in several organic solvents. In order to achieve this, thin sheets of protein-encapsulated alginate samples were prepared in a petri dish. A sample of protein was mixed into 3% sodium alginate until fully dispersed and was allowed to sit until all the bubbles had disappeared. Then a thin sheet of enzyme-loaded sodium alginate was loaded into a petri dish. The sodium alginate was submerged in a solution of 0.3M $CaCl_2$ and the samples were left for 1 hour to fully mature. After 1 hour, a scalpel was used to cut out thin strips of the enzyme-encapsulated alginate layers in accordance with the dimensions of the UV-VIS cuvette. In order to optimise this process, initial characterisations were performed with equine cytochrome *c*.

The normalised UV-VIS spectra acquired for cytochrome *c* in CHES buffer, alginate-encapsulated cytochrome *c* in CHES buffer, and alginate-encapsulated cytochrome *c* in anhydrous hexane are presented in **figure 132**. It is immediately apparent that the Soret peak for cytochrome *c*, at 406 nm, remains completely unaffected when cytochrome *c* is encapsulated inside an alginate bead, and when the alginate beads are suspended into anhydrous hexane. In addition, a Q-band, though poorly

resolved, remains observable in the 500-550 nm region for all samples. The poor resolution of the Q-band is probably a consequence of scattering from the alginate beads; nonetheless, the UV-VIS spectra indicate cytochrome *c* remains unperturbed in pure hexane when encapsulated inside an alginate bead.

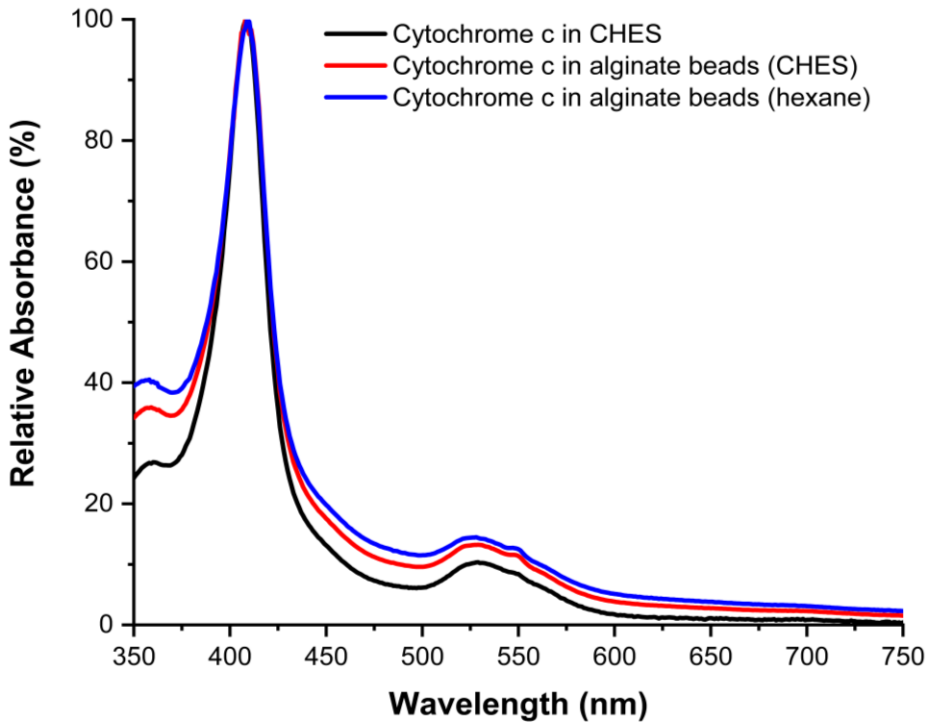


Figure 132: The normalised UV-VIS spectra of cytochrome *c* in CHES buffer (black) and alginate-encapsulated (0.3M $CaCl_2$, 3% (w/v) sodium alginate) cytochrome *c* suspended in CHES (red) and n-hexane (blue).

Having acquired sufficient data for cytochrome *c*, the C45-encapsulated alginate sheets were subsequently investigated spectroscopically. Applying the same protocol for cytochrome *c*, C45-alginate sheets were prepared and examined spectroscopically. The normalised UV-VIS spectra for C45-encapsulated alginate sheets in CHES, MeOH, EtOH, acetone, ethyl acetate, and hexane are presented in **figure 133 (top left)**. The UV-VIS spectra acquired for all solvents systems are all dominated by a Soret peak at 406 nm, as is observed under aqueous conditions. The resolution of the Q-band, however, is poor, probably a consequence of scattering by the alginate bead. All the recorded spectra so far have been for ferric-C45; as it is necessary to reduce ferric-C45 to ferrous-C45 in order to access carbene transferase chemistry, C45-encapsulated alginate sheets were prepared in CHES buffer prior to the addition of excess sodium dithionite. The time-course UV-VIS spectra presented in **figure 133 (top right)** shows that over the course of 90 seconds the ferric-C45 Soret peak at 406 nm disappears and is accompanied by the expected concurrent emergence of a peak at 417 nm,

corresponding to the ferrous-C45 peak. The long duration before appreciable quantities of ferrous-C45 can be detected is a consequence of the slow diffusion rate across the alginate-solvent interface; however, it would be anticipated that reduction would occur more rapidly in the presence of pure hexane, in which sodium dithionite is insoluble. For comparison, the experiment was repeated using *Rma*-TDE. The normalised UV-VIS spectra for *Rma*-TDE in CHES, MeOH, EtOH, acetone, ethyl acetate, and hexane are presented in **figure 133 (bottom left)**, alongside the UV-VIS spectra for the alginate-encapsulated *Rma*-TDE in the presence and absence of sodium dithionite **figure 133 (bottom right)**. The UV-VIS spectra acquired for all the solvent systems are dominated by a Soret peak at 406 nm, as is observed under aqueous conditions. In addition, the UV-VIS spectra recorded in the absence of sodium dithionite is dominated by a Soret peak at 406 nm, correspond to ferric *Rma*-TDE. The UV-VIS spectra in the presence of sodium dithionite, however, is dominated by a Soret peak at 417 nm, corresponding to ferrous *Rma*-TDE. Interestingly, the UV-VIS spectra recorded for alginate-encapsulated ferrous *Rma*-TDE also displays a reasonable well resolved split Q-band with peaks at 525 nm and 551 nm. Overall, the UV-VIS spectra for alginate-encapsulated C45 and *Rma*-TDE demonstrates the enzymes remain spectroscopically distinct and unaffected by the presence of organic solvents, and that reduction from the ferric to the ferrous state is possible and is only impacted by slow diffusion across the alginate-solvent interface.

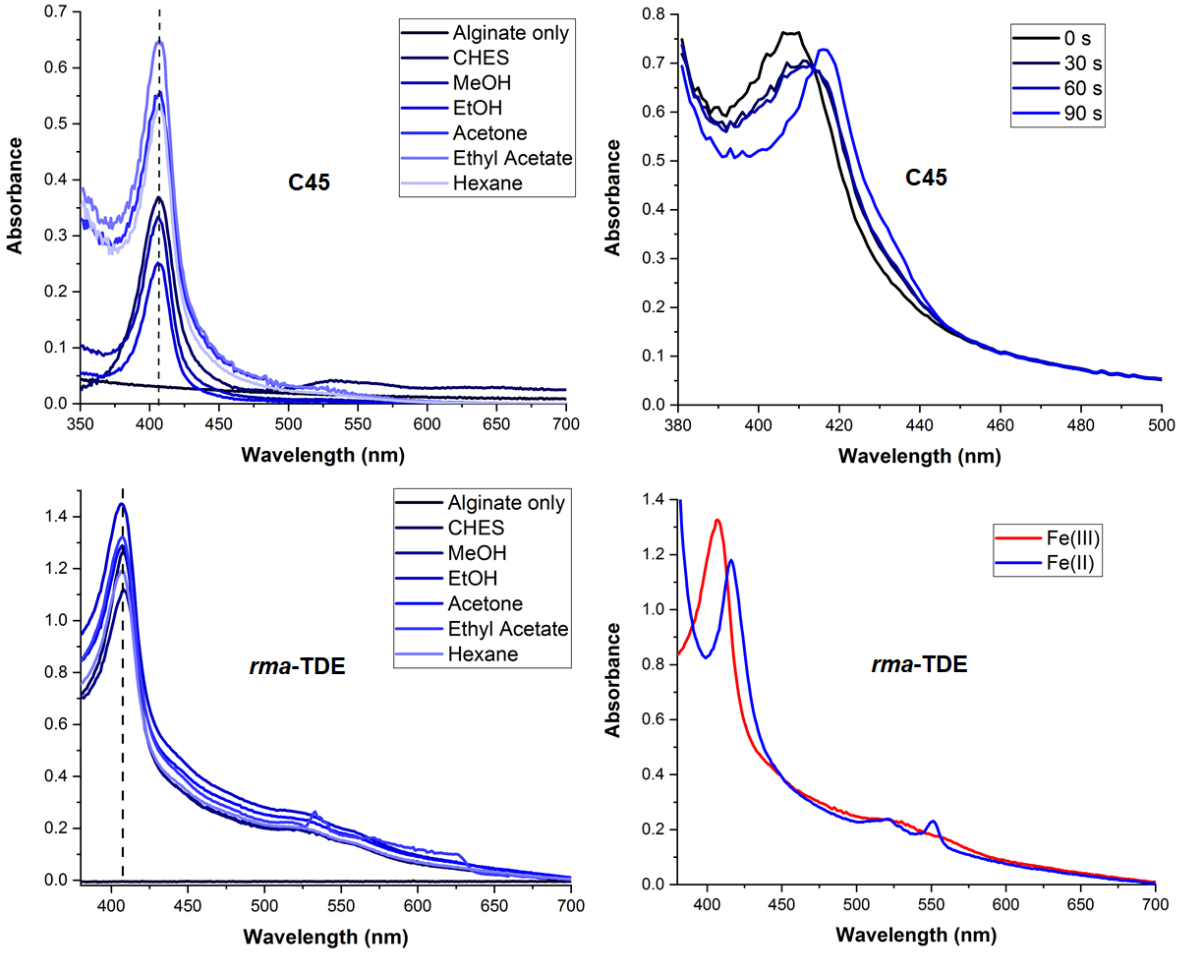


Figure 133: UV-VIS spectra of (**top left**) alginate encapsulated C45 in various organic solvents, (**top right**) alginate encapsulated C45 in the presence of $Na_2S_2O_4$, (**bottom left**) alginate encapsulated *Rma*-TDE in various organic solvents, (**bottom right**) alginate encapsulated *Rma*-TDE in the presence of $Na_2S_2O_4$.

Having demonstrated that C45 retains its spectroscopic profile when encapsulated in calcium alginate beads and can be successfully reduced from the ferric to the ferrous state, the question of whether the formation of the metallocarbenoid intermediate could be identified spectroscopically was raised. Applying the same protocol delineated above, a thin sheet of C45-encapsulated calcium alginate was prepared. The sheet was transferred to a UV-VIS cuvette and submerged in CHES buffer for subsequently UV-VIS analysis. The acquired UV-VIS spectra for ferric C45, ferrous C45, and ferrous C45 in the presence of EDA are presented in **figure 134 (left)**. The UV-VIS of alginate-encapsulated ferric C45 (**figure 134 (left)**; black) is characterised by a large Soret peak at 407 nm, accompanied by a relatively low-intensity, but detectable, Q-band in the 500-600 nm region. After the addition of $Na_2S_2O_4$ a new spectrum emerged (**figure 134 (left)**; red) characterised by an intense Soret peak at 418 nm and two well-defined peaks at 542 nm (α peak) and 584 nm (β peak) composing the Q-band,

with an intensity profile of β > α. The acquired UV-VIS for alginate encapsulated ferrous C45 is consistent with the UV-VIS spectra acquired free ferrous C45. After the addition of EDA to the cuvette, a new red-shifted spectrum emerges (**figure 134 (left)**; blue) which is dominated by an intense Soret peak at 426 nm which is accompanied by Q-band consisting of two red-shifted peaks at 542 nm (α peak) and 584 nm (β peak), with an intensity profile of β > α. **Figure 134 (right)** presents the normalised UV-VIS spectrum acquired for alginate encapsulated ferrous C45 in the presence of EDA (**figure 134 (right)**; blue) compared with the normalised UV-VIS spectra acquired for the C45-EDA metallocarbenoid intermediate in solution (**figure 134 (right)**; red). The two spectra are affectively identical, with both spectra characterised by an intense Soret peak at 426 nm which is accompanied by a Q-band consisting of two peaks at 541 nm (α peak) and 582 nm (β peak), with an intensity profile of β > α. The spectroscopic data acquired clearly indicates that encapsulating C45 inside calcium alginate beads does not impact its ability to form a stable, metallocarbenoid intermediate. The formation of the intermediate is, however, slower when C45 is encapsulated, but this is to be expected on account of partitioning dynamics of EDA across the alginate-solution barrier, and diffusion of EDA through the polymeric network of the alginate bead towards C45.

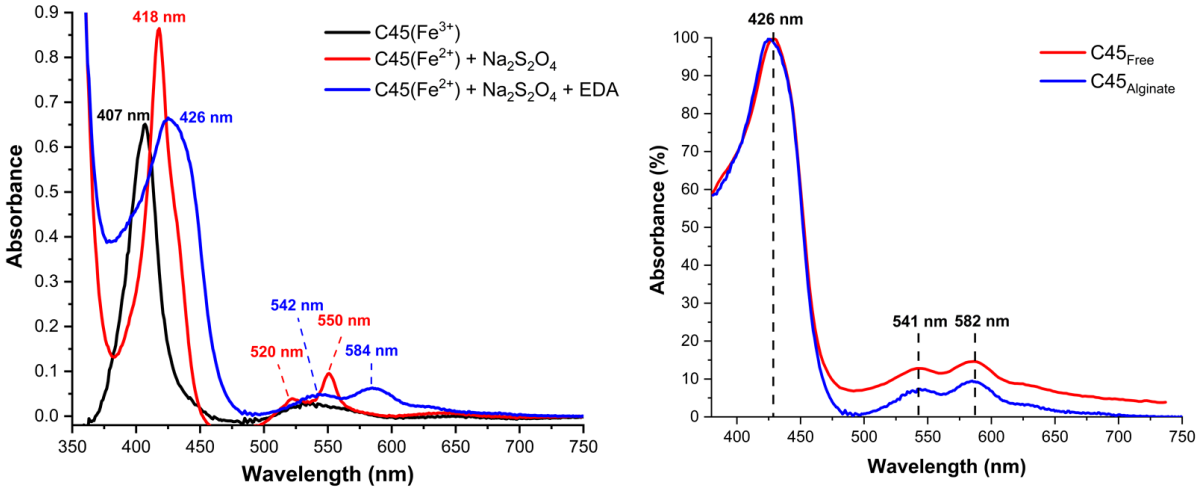


Figure 134: Alginate encapsulated C45 metallocarbenoid formation

6.4.3 Alginate encapsulated-C45/HRP catalysed peroxidase activity and functional recoverability

Having spectroscopically demonstrated that encapsulated hemoproteins remain stable and can be reduced in the presence of suitable reductants, the next question concerned catalytic activity. Although C45 has been demonstrated to be an efficient carbene transferase, C45 was originally designed with the intention of mimicking natural oxidoreductases. C45 is an established *de novo* peroxidase;[235] therefore, the initial experiments investigating the enzymatic activity of alginate-

encapsulated C45 were focused on its peroxidase chemistry. The reasons for investigating peroxidase chemistry prior to investigating carbene transfer chemistry was because i) peroxidase chemistry can be conducted under aerobic conditions, ii) peroxidase assays do not require C45 to be reduced to the ferrous state, and iii) a well-established and colourimetric assay involving ABTS and hydrogen peroxide exists allowing for the peroxidase kinetics to readily obtained much easier compared to carbene transfer chemistry. For comparative purposes, the peroxidase assay between ABTS and hydrogen peroxidase, catalysed by alginate-encapsulated C45, was benchmarked against alginate-encapsulated horseradish peroxidase (HRP), as the kinetics of the peroxidation of ABTS, in aqueous solution, has been previously reported for C45 and was compared against HRP.[235]

6.4.3.1 Horseradish peroxidase and C45 – peroxidase kinetics and recoverability

Horseradish peroxidase is a large monomeric α-helical heme-containing enzymes that catalyses the oxidation of various organic compounds using hydrogen peroxide as a substrate.[101,136] One particular substrate, 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid) (ABTS) is a common reagent used to determine the reaction kinetics of peroxidases. ABTS exists as a colourless dianionic species in solution but can function as a reductant in the presence of a peroxidase, supplying the electrons in the reduction of hydrogen peroxide. The first reduction step generates a green-blue radical cation whose production can be continuously monitored spectroscopically.[235]

Alginate-encapsulated HRP beads were prepared according to the protocol outlined in chapter 9, *section 9.12*. The final beads had an average diameter of 1.4 mm (± 0.1 mm) and an average HRP concentration of 100 µM. A 100 mM solution of ABTS and 1 mM solution of H_2O_2 were prepared in CHES buffer (20 mM CHES, 100 mM KCl, pH – 8.6). All reactions were performed in triplicates, in a 1 ml UV-VIS cuvette with one bead per reaction and a final concentration of 1 mM ABTS. The concentration of hydrogen peroxide was varied across a 1-100 µM range and the reaction progress was monitored using a single-wavelength spectrophotometer set at 680 nm. **Figure 135** shows that the alginate-encapsulated HRP catalysed peroxidation of ABTS at select concentrations of H_2O_2 obeys Michaelis-Menten kinetics, and from **figure 135** values of 8.83 µM and 0.054 s^{-1} were calculated for K_M and k_{cat} respectively, resulting in a calculated catalytic efficiency of k_{cat}/K_m = 6.12 x 10^3 M^{-1} s^{-1}. For reference, the catalytic efficiency for HRP towards in the peroxide activation step in aqueous solutions has been reported as $k_{cat}/K_m = 4.6 \times 10^6\ M^{-1}\ s^{-1}$,[235] meaning encapsulating HRP inside an alginate bead results in a 1000-fold decrease in catalytic efficiency towards the peroxide activation step. For reference, the values for k_{cat} and $K_M(H_2O_2)$ for HRP (100 nM) in solution have been reported as 4100 s^{-1} and 900 µM respectively.[235]

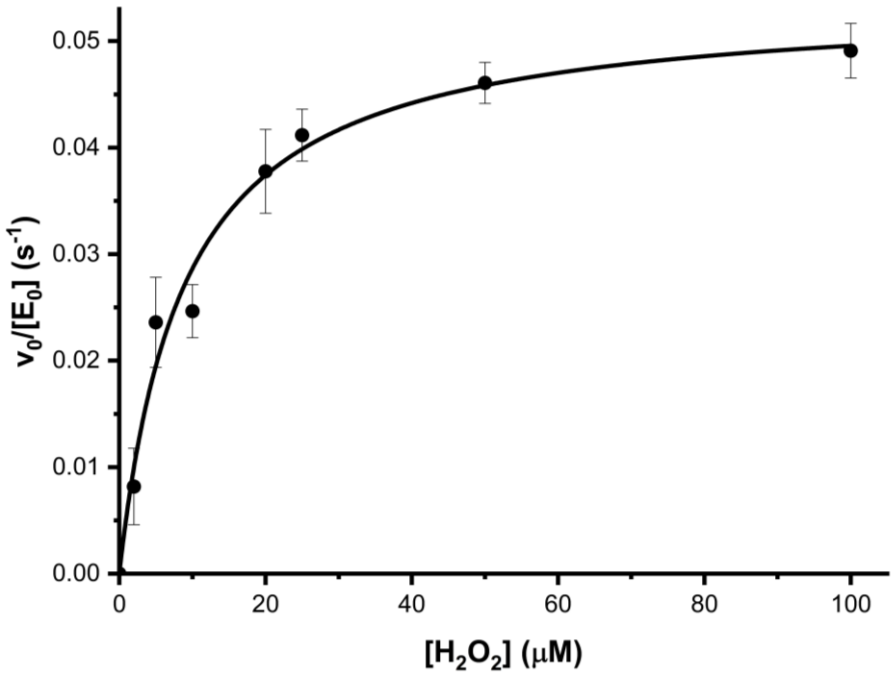


Figure 135: The Michaelis-Menten plot of the AE-HRP catalysed peroxidation of ABTS as a function of hydrogen peroxide concentration. Initial rates were calculated from the rate of change in absorbance of the ABTS cation radical.

Having demonstrated that HRP remains active towards ABTS peroxidation when encapsulated inside an alginate bead, albeit with lower catalytic activity, it was of subsequent interest to whether C45 would also remain catalytically active towards the same chemistry when encapsulated inside alginate beads. Alginate-encapsulated C45 beads were prepared according to the protocol provided in section 9, and the experimental conditions paralleled the ones described for HRP, with exceptions being a greater hydrogen peroxide concentration range (100 – 2500 μM) and C45 concentration. The final beads had an average diameter of 1.4 mm (± 0.1 mm) and an average C45 concentration of 10 μM **Figure 136** shows, as with HRP, that the alginate-encapsulated C45 catalysed peroxidation of ABTS at select concentrations of H_2O_2 obeys Michaelis-Menten kinetics, and values of 686.15 μM and 0.081 s^{-1} were calculated for K_M and k_{cat} respectively, resulting in a calculated catalytic efficiency of k_{cat}/K_m = 1.24 x 10^2 M^{-1} s^{-1}. For reference, the catalytic efficiency for C45 towards the peroxide activation step in aqueous solutions has been reported as $k_{cat}/K_m = 1.3 \times 10^4$ M^{-1} s^{-1},[235] meaning encapsulating C45 inside an alginate bead results in a 100-fold decrease in its catalytic efficiency towards the peroxide activation step. The values for k_{cat} and $K_M(H_2O_2)$ for C45 (100 nM) in solution are 1200 s^{-1} and 94 mM respectively, meaning the significant decrease in k_{cat} is somewhat compensated for by the decrease in K_M.[235]

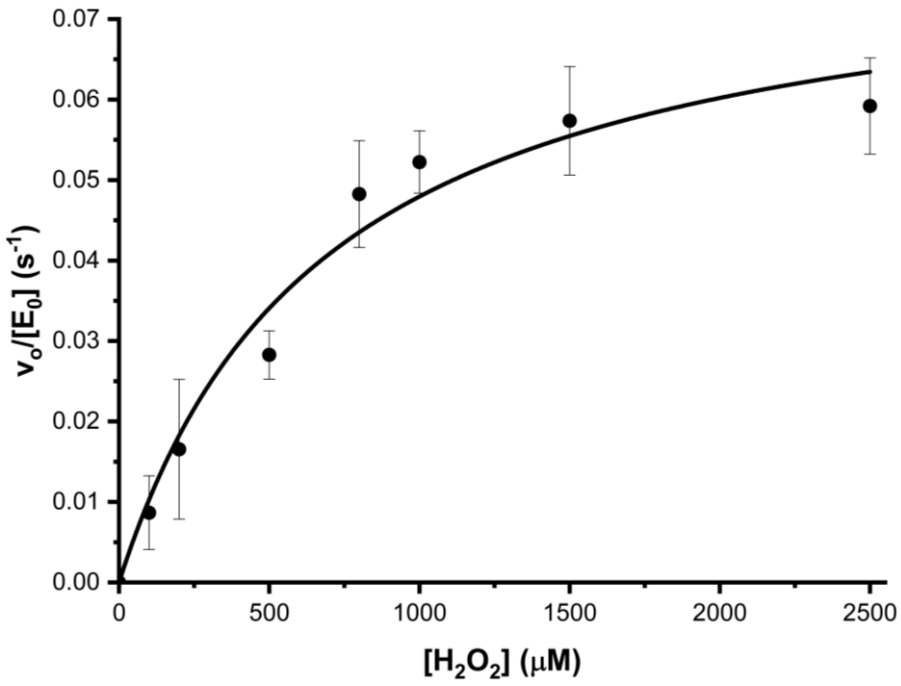


Figure 136: The Michaelis-Menten plot of the AE-C45 catalysed peroxidation of ABTS as a function of hydrogen peroxide concentration. Initial rates were calculated from the rate of change in absorbance of the ABTS cation radical.

The reaction of H_2O_2 and ABTS follows a steady-state ping-pong mechanism typical of peroxidase enzymes.[235] The overall mechanism is broken into two step-wise processes: the initial binding and activation of peroxide followed by the subsequent binding of ABTS prior to an electron transfer. It was previously reported that C45 exhibits a catalytic efficiency of k_{cat}/K_M(ABTS) = 3.2 × 10^6 M^{-1} s^{-1} for the electron transfer between peroxide-activated C45 and ABTS, which is purported to be one of the highest catalytic efficiencies observed for a *de novo* enzyme, possessing a value very close to that for HRP (k_{cat}/K_M(ABTS) = 5.13 × 10^6 M^{-1} s^{-1}).[235] In contrast to the high efficiency of the electron transfer step, the activation of peroxide by C45 is less catalytically efficient than in HRP (C45 k_{cat}/K_M = 1.3 × 10^4 M^{-1} s^{-1}; HRP k_{cat}/K_M = 4.6 × 10^6 M^{-1} s^{-1}), which is attributed to the lack of peroxide-activating amino acid side chains in the vicinity of the heme in C45 relative to HRP, which is reflected in the $K_M(H_2O_2)$ values reported in solution (C45 = 94 mM; HRP = 900 μM).[235] However, when encapsulated inside an alginate bead, HRP exhibits a 1000-fold decrease in catalytic efficiency towards the peroxide step, in contrast to a 100-fold decrease for C45. In order to account for this observation, it is necessary to consider how encapsulating an enzyme inside an alginate bead would affect the dynamics and activity of enzymes which would be reflected in the values of k_{cat} and K_M. The first consideration concerns partitioning of substrates across the bead/solvent interface. As enzymatic processes are dependent on the diffusion of the necessary substrates, the introduction of additional diffusion processes would be expected to negatively affect the overall rate of a reaction, as has been previously reported for

alginate-encapsulated glucose oxidase activity.[381] The effect of substrate diffusion has already been observable in the reduction of ferric alginate-encapsulated C45 and *Rma*-TDE, a process requiring 120 seconds before appreciable quantities of the ferrous species could be detected spectroscopically. The diffusion of substrate(s) across the alginate-solvent barrier is therefore accredited as a major factor contributing to the observed reduction in k_{cat} for both C45 and HRP ABTS peroxidation. The second major consideration concerns the relative sizes, and confirmations, of the enzymes inside the alginate beads.

One major limitation of homogenous biocatalysts is the difficulty associated with recovering the catalyst post-reaction. An enzyme immobilised in an alginate bead creates a heterogeneous biocatalyst that can be readily recovered at the end of the reaction and reused in another catalytic cycle. Therefore, having demonstrated that C45 remains catalytically active towards ATBS peroxidation when encapsulated inside alginate beads, subsequent investigations were concerned with whether C45 would remain catalytically active once recovered after one round of peroxidation chemistry. For comparative purposes, a recoverability assay was initially optimised using HRP. Alginate-encapsulated HRP beads were prepared according to the protocol provided in chapter 9, *section 9.12*. The final beads had an average diameter of 1.4 mm (± 0.1 mm) and an average HRP concentration of 100 µM. A 100 mM solution of ABTS and a 100 mM solution of H_2O_2 were both prepared in CHES buffer (20 mM CHES, 100 mM KCl, pH – 8.6). All the reactions were conducted in a 1 ml UV-VIS cuvette with one bead per reaction and the final concentrations were 1 mM ABTS and 1 mM H_2O_2. The reaction progress was monitored using a single-wavelength spectrophotometer set at 680 nm. After 2 minutes, the bead was recovered from the cuvette, both the cuvette and the bead were washed with 1 ml CHES buffer and another reaction was prepared using the same bead and identical reaction conditions. For the C45 recoverability assays, an identical protocol was followed, with the only exception being the final enzyme concentration was 10 µM. The normalised conversion rates for alginate-encapsulated HRP and C45 ABTS peroxidation recoverability are presented in **figure 137**. For HRP, the enzyme retained 79% and 24% of its initial activity during the 2nd and 3rd runs, respectively. C45 could be recovered and reused three additional times before any lack of appreciable activity could be detected, retaining 52%, 23% and 12.5% of its initial catalytic activity during the 2nd, 3rd and 4th run, respectively.

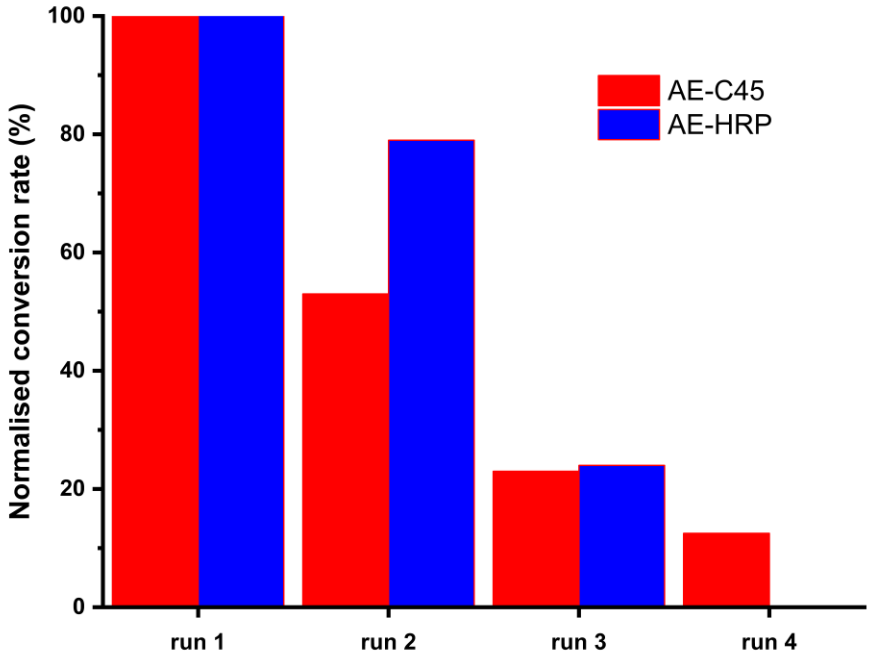


Figure 137: The normalised conversion rate of AE-C45 (red) and AE-HRP (blue) towards the peroxidation of ABTS over successive catalytic cycles (all values are within a 10% error margin).

6.4.4 Alginate encapsulated-C45 catalysed carbene transfer chemistry

Having demonstrated that C45 encapsulated in calcium alginate beads retains both its peroxidase activity and its ability to form a stable metallocarbenoid intermediate, the following step was to explore the carbene transferase activity of alginate-encapsulated C45 (AE-C45) when the loaded beads were submerged in a variety of non-aqueous solvents. As the alginate-encapsulated C45 beads function as a heterogeneous catalyst, coupled with the aforementioned observation that heterogeneous C45 retains its peroxidase activity, it was also of interest to investigate the recoverability of C45 and its carbene transferase proficiency when utilised in subsequent rounds of catalysis.

The cyclopropanation reaction between styrene and EDA, and the N-H insertion reaction between piperidine and EDA were selected as the reactions to investigate the catalytic activity of heterogeneous AE-C45. A selection of 12 varying solvents were used, ranging from polar protic, polar aprotic, and non-polar solvents, and included CHES buffer (20 mM CHES, 100 mM KCl, pH 8.6), methanol, ethanol, isopropanol, acetone, acetonitrile, ethyl acetate, THF, dichloromethane, hexane, petroleum ether, and toluene. All the reactions were performed in 1.5 ml glass vials sealed with an air-tight septum lid. The reaction volumes were 1 ml of neat solvent and each reaction vial contained three C45-containing alginate beads. Stock solutions of $Na_2S_2O_4$ (1 mM in dH_2O), EDA (1 mM in EtOH) and styrene/piperidine (3 mM in EtOH) were prepared and injected into the reaction mixtures, under

a nitrogen atmosphere, to initiate the reactions. The final concentrations of the substrates were 10 mM EDA, 30 mM styrene/piperidine and 10 mM $Na_2S_2O_4$. All the reactions were performed under a nitrogen atmosphere.

6.4.4.1 Styrene cyclopropanation

After leaving the reactions to mix for 2 hours, the reactions were quenched by the addition of 25 µl of a 3M HCl solution. The alginate beads were removed from the vials, and then 1 ml of ethyl acetate was added to each vial before each reaction mixture was transferred to an individual 15 ml falcon tube, vortexed and centrifuged (14,500 rpm; 2 minutes). The organic layer for each reaction mixture was subsequently transferred to a separate 15 ml falcon tube containing 1 ml of 3M NaOH, and the mixture was left stirring for another 2 hours, and TLC was used to monitor the progress of the hydrolysis reaction (7:3 ethyl acetate:hexane, 254 nm). Once no progress could be detected for the hydrolysis reaction, the aqueous layer was extracted and analysed *via* chiral-HPLC and LC-MS.

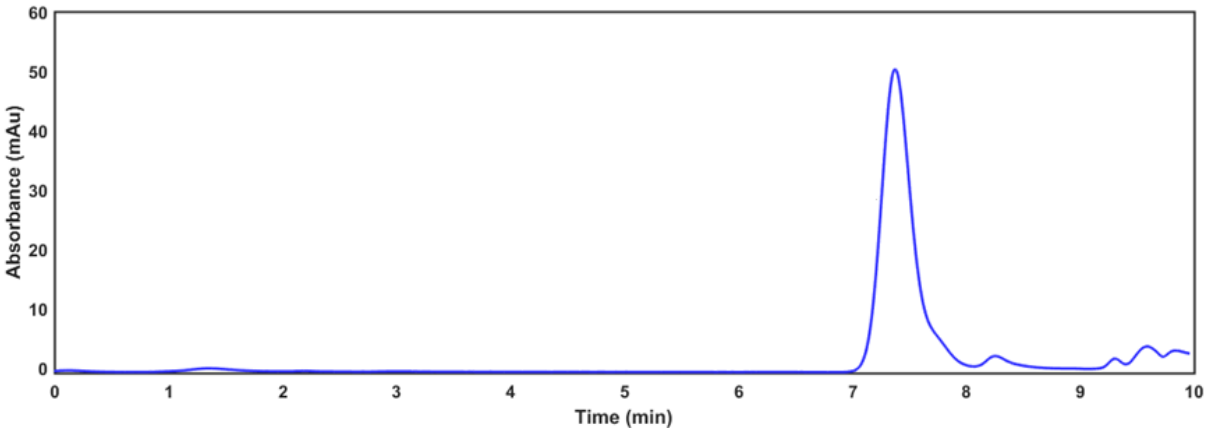


Figure 138: The chiral-HPLC chromatogram for the alginate-encapsulated C45 catalyzed cyclopropanation reaction between styrene (30 mM) and EDA (10 mM) (CHES buffer, pH 8.6, 280 nm). A polar organic mobile phase (100% MeCN: 0.1% (v/v) TFA:0.1% (v/v) Et_3N) was employed and injection volumes were 2 µl.).

Initial assays were conducted using CHES buffer to check for any activity and, if detected, as a suitable comparison to free C45. **Figure 138** presents the chiral-chromatogram for the alginate-encapsulated C45-catalysed cyclopropanation reaction between styrene and EDA in CHES buffer. The chromatogram is dominated by a peak with a retention time of 7.2 minutes, the known retention time of the anticipated (*R,R*)-cyclopropane-product, followed shortly by a small, additional peak with a retention time of 9.2 minutes, the known retention time of the (*S,S*)-cyclopropane-product. The peak at 7.2 minutes was collected and analysed *via* LC-MS to confirm the identity of the compound, with the acquired mass spectrum exhibiting the diagnostic fragmentation profile of ethyl (*1R,2R*)-2-

phenylcyclopropane-1-carboxylate. A reaction yield of 55.30% and enantiomeric excess of 82.81% were calculated from the peak heights. For comparison, the percentage yield and enantiomeric excess for the assays conducted with free C45 are 80.20% and 77.00% respectively. It was previously demonstrated that encapsulating C45 inside an alginate bead does not impair its ability to form a stable metallocarbenoid intermediate, so it is unsurprising that C45's encapsulation does not impair its proficiency towards carbene transfer chemistry; however, the encapsulation does results in an approximate 31% decrease in product yield, which is marginally compensated for by a 7% increase in enantiomeric excess.

The catalytic activity and selectivity of AE-C45 was next investigated using an assortment of diverse organic solvents. The acquired values for product yields and (*R,R*)-enantioselectivity for each solvent is tabulated in **table 11** and presented in **figure 139**. The cyclopropanation reaction catalysed by AE-C45 submerged in the polar protic solvents MeOH, EtOH, and IPA afforded average reaction yields of 59.60, 38.00, and 79.00% and enantiomeric excesses of 69.20, 91.80, and 95.80% respectively. The values reported for MeOH are similar to the values reported for CHES buffer and is probably accounted for by the similarities between the two solvents. For EtOH, the reaction yield is significantly affected, exhibiting a 31.4% decrease, relative to CHES buffer. Isopropanol, however, resulted in an approximate 30% increase in the product yield. The cyclopropanation reaction catalysed by AE-C45 submerged in the polar aprotic solvents CH_3CN, EtOAc, THF, acetone, and DCM afforded average reaction yields of 92, 18.60, 91.00, 91.10, and 49.50% and enantiomeric excesses of 89.40, 74.60, 96.50, 94.80, and 92.70% respectively. The solvents EtOAc and DCM negatively affected the reaction yield, exhibiting a 66.40 and 10.30% decrease in product yield relative to CHES, CH_3CN, THF, and acetone; however, both drastically enhanced the product yields, with an average 39% increase in product yield relative to CHES. The cyclopropanation reaction catalysed by AE-C45 submerged in the two non-polar solvents hexane and pether afforded average reaction yields of 9.40 and 6.50% and enantiomeric excesses of 62.40 and 53.00% respectively. The non-polar solvents most negatively affected the reaction, affording an average 85.70% decrease in product yield and 30.30% decrease in enantiomeric excess, respectively.

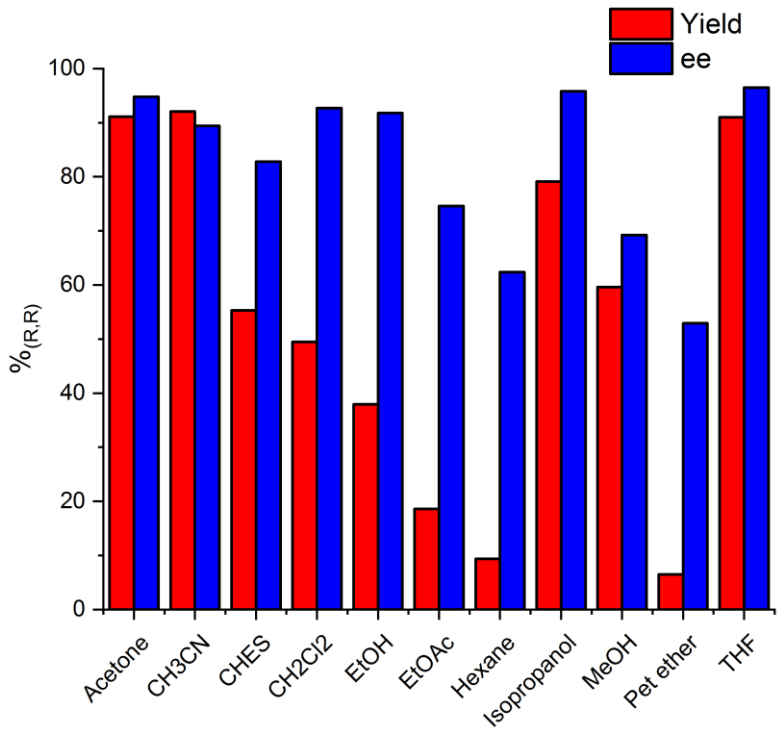


Figure 139: The product yields (red bars) and enantioselectivities (blue bars) of AE-C45 catalysed cyclopropanation reaction between EDA and styrene in various organic solvent.

Table 11: The product yields of the (*R,R*)-cyclopropane product and enantioselectivities for the alginate encapsulated C45 catalysed cyclopropanation reaction between styrene and EDA in an assortment of organic solvents (all values are within a 10% error margin).

	C45	
Solvent	**% yield$_{(R,R)}$**	**% ee$_{(R,R)}$**
Acetone	91.10	94.80
Acetonitrile	92.09	89.40
CHES	55.30	82.80
Dichloromethane	49.50	92.70
Ethanol	37.92	91.78
Ethyl Acetate	18.59	74.57
Hexane	9.40	62.40
Isopropanol	79.09	95.79
Methanol	59.60	69.23
Petroleum ether	6.50	52.94
THF	91.01	96.49

6.4.4.1.1 Product yields and substrate diffusion

Figure 139 presents the product yields (red) and enantioselectivities (blue) for ethyl (*1R,2R*)-2-phenylcyclopropane-1-carboxylate acquired from each AE-C45 solvent assay. From **figure 139**, it can be readily seen that changing the solvent system of the assays affects the enantioselectivity of the reaction, but the solvent induces a much greater influence over the product yields. Unlike the reaction catalysed by free enzyme, the encapsulation of the enzyme inside an insoluble bead introduces additional factors to the overall reaction. The partition of the substrate(s) between the bulk phase and the bead, and the rate of diffusion of both the substrate(s) and product(s) across the bulk-bead interface and through the bead itself introduce novel dynamic process which are intricately entwined with the overall reaction rate. The partition coefficient, as discussed in *section 6.3*, represents the preferential solvation of a substrate molecule across two separate, distinct phases when the system is in equilibrium. In the heterogeneous alginate systems, the two phases are the bulk organic phase (BOP) and the aqueous alginate phase (AAP), which are separated by the bulk-alginate interface (BAI). The partition coefficient is the ratio of the substrate concentration in the AAP relative to the BOP and is a function of the physical and chemical properties of organic solvent constituting the BOP. One important property to consider is the polarity of a solvent, which can be quantified by a solvent's dielectric constant (ε; Farad m^{-1}); the higher the dielectric constant the more polar the solvent. Styrene, being composed exclusively of non-/marginally polar C-C and C-H bonds, is a non-polar molecule with a dielectric constant of approximately 2 F m^{-1}. The dielectric constants of water, being composed of extremely polar O-H bonds, and hexane, being composed of non-polar C-C and C-H bonds, are approximately 80 and 1.68 F m^{-1}, respectively. In a biphasic system where hexane constitutes the BOP it would be anticipated that the partition coefficient ($[styrene]_{hexane}/[styrene]_{bead}$) would be exceedingly high, indicating that styrene overwhelmingly is preferentially solvated by hexane compared to water, resulting in a very low flux of styrene diffusing into the bead to become available for catalysis. **Figure 140** presents the dielectric constant of the various solvents screened in the AE-C45 cyclopropanation assays vs the respective (*R,R*)-product yields for each assay. From **figure 140**, a very tenuous relationship can be inferred between the respective product yield for each AE-C45 cyclopropanation assay and the dielectric constant of the solvent used. This can be accounted for by consideration of the precise information containing in the dielectric constant. The dielectric constant represents the capacitance of a system, the ability of a system to store electrical energy when subject to an electric field. The dielectric constant is a physical quantity and contains no relevant information regarding the solute-solvent molecular interactions, and therefore cannot account for differences observed between separate solvents. For example, DCM, ethyl acetate, and THF all have extremely similar dielectric constants; however, DCM and ethyl acetate are completely immiscible in

water, whereas THF is highly miscible. The differences in the solubilities of THF, ethyl acetate, and DCM are a function of the molecular interactions established between solute-solvent molecules. The oxygen in THF can establish a hydrogen-bond with a water molecule which significantly enhances the miscibility of the two solvents; DCM and ethyl acetate, despite being polar molecules, are incapable of establishing hydrogen-bonds with water which significantly impairs their miscibility. The lack of any molecular information contained with the dielectric constant accounts for the lack of any relationship in **figure 140**.

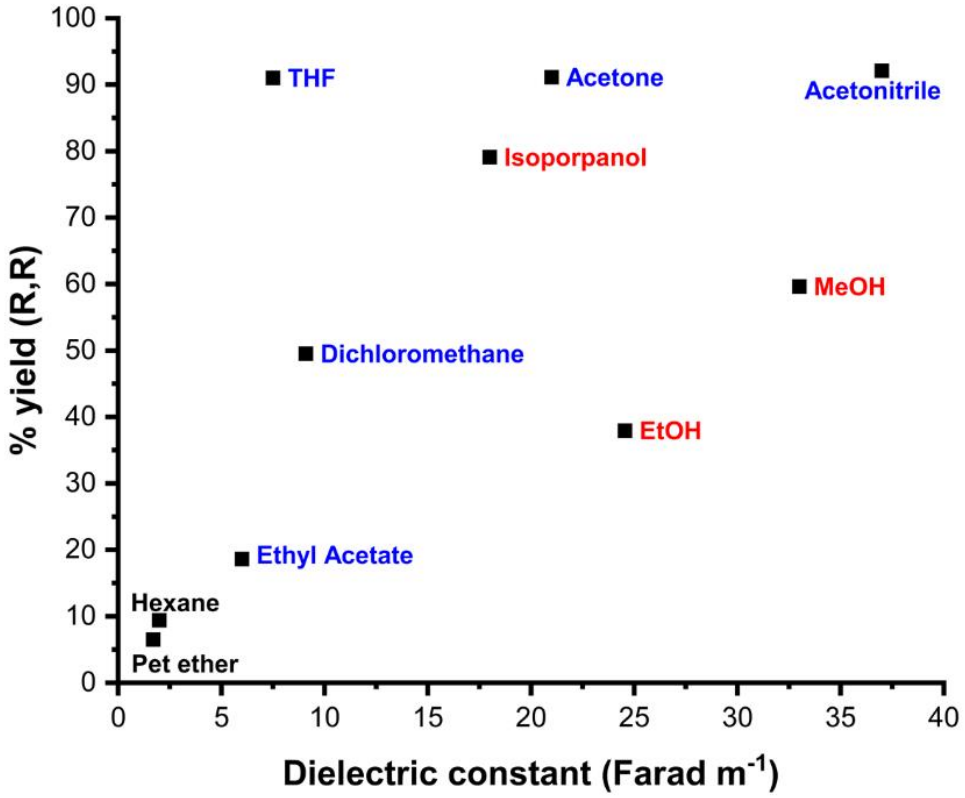


Figure 140: The product yields for the (*R,R*)-cyclopropane product obtained from the AE-C45 catalysed cyclopropanation reaction between EDA and styrene as a function of the dielectric constant of the solvent employed in the reaction.

Unlike the dielectric constant, a partition coefficient is a thermodynamic quantity and contains information on the solute-solvent interactions. The partition coefficient was demonstrated in *section 6.3* to be mathematically and molecularly related to the diffusion coefficient of a substrate and provides information on the rate of flux of a molecule across a permeable barrier which separates two distinct phases. The partition coefficients for styrene in each solvent were determined by C18-HPLC. A 1 ml vial containing three empty alginate beads was submerged with an organic solvent followed by

the introduction of 10 µl of styrene from a 1M stock. The vials were allowed to stir for 2 hours until equilibrium had been achieved, then the BOP was extracted and quantitatively analysed by C18-HPLC. The remaining concentration of styrene in the BOP was calculated with reference to an external calibration of styrene across a concentration range of 500 µM – 10 mM (**figure 167**, *appendix 1*). The partition coefficient for each solvent was calculated using equation (**3**), which were subsequently used to calculate the parameter ''α'' using equation (**17**). The diffusion coefficients for styrene in each solvent were calculated using equation (**18**)and the effective diffusion coefficients were calculated using equation (**13**) (**table 12**; see chapter 9, *sections 9.18 and 9.19* for further details).

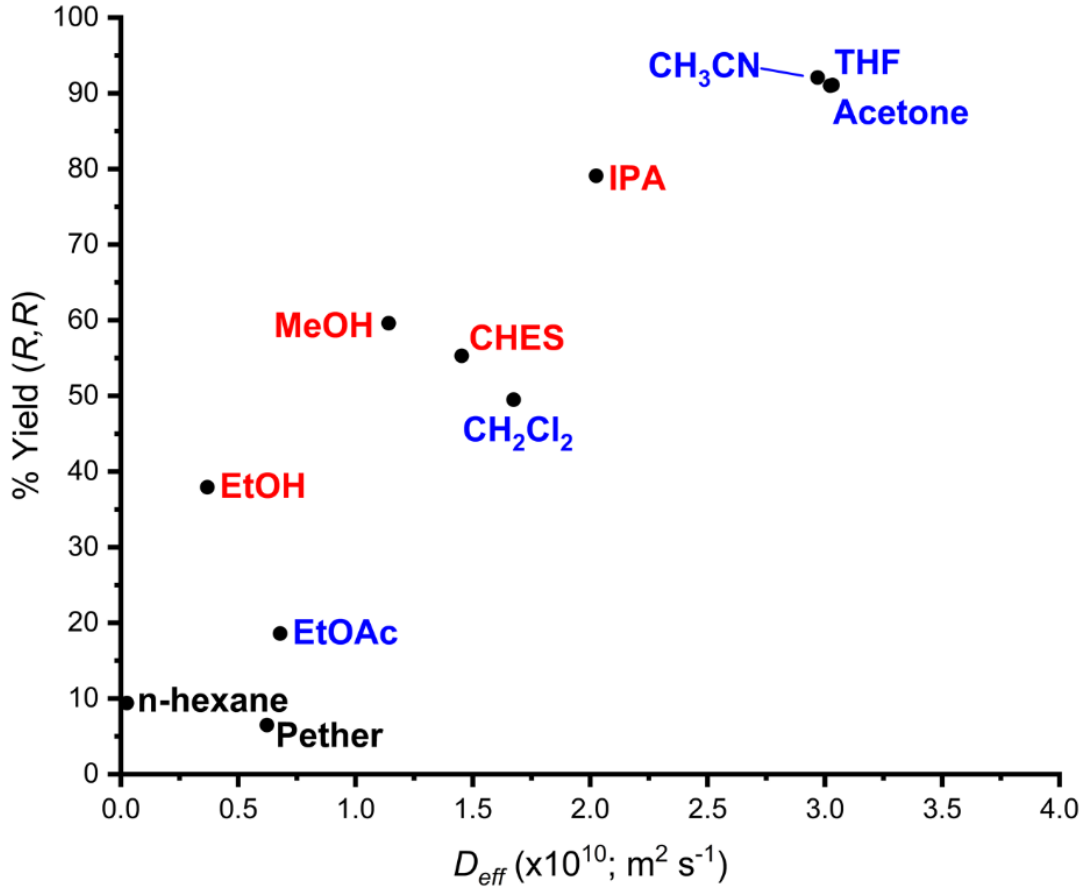


Figure 141: The product yields for the (*R,R*)-cyclopropane product obtained from the AE-C45 catalysed cyclopropanation reaction between EDA and styrene as a function of the effective diffusion coefficient of styrene into the alginate-bead in various organic solvents.

Table 12: Partition coefficient and diffusion coefficients for styrene in various organic solvents suspended in alginate beads prepared from 3% (w/v) sodium alginate and 0.3M $CaCl_2$; R = radius of the alginate beads; alpha parameter is defined by equation (**17**); all values are within a 10% error margin.

Solvent	Partition coefficient, K	α	R (mm)	D (x10^{10}, m^2 s^{-1})	D_{eff} (x10^{10}, m^2 s^{-1})
CHES	0.047	3.366	1.4	1.633	1.452
MeOH	0.088	6.291	1.4	1.284	1.141
EtOH	0.060	4.289	1.4	0.415	0.369
Isopropanol	0.144	10.295	1.4	2.279	2.026
CH3CN	0.252	17.970	1.4	3.340	2.969
EtOAc	0.070	4.965	1.4	0.763	0.679
THF	0.260	18.560	1.4	3.400	3.023
Acetone	0.261	18.668	1.4	3.411	3.033
CH2Cl2	0.118	8.423	1.4	1.883	1.674
Hexane	0.052	3.701	1.4	0.028	0.025
Petroleum ether	0.010	0.736	1.4	0.700	0.623

Figure 141 presents the (*R,R*)-product yield acquired from the AE-C45 catalysed cyclopropanation assays as a function of the calculated effective diffusion coefficient for styrene with respect to the bulk organic solvent. The figure demonstrates a reasonably good linear relationship between the two parameters, indicating that the diffusion rate of styrene across the bulk-alginate interface (BAI) is a key factor governing catalysis. The diffusion coefficient for styrene, with respect to the solvent system, corresponds to the amount of styrene localised inside the alginate bead with respect to time; the higher the diffusion coefficient the greater the general flux of styrene across the BAI from the BOP into the bead. Styrene is highly miscible with n-hexane, pether, ethyl acetate, and ethanol and the favourable solute-solvent interactions established in the bulk phase impairs styrene's diffusion into the bead, as can be induced from **figure 141.** The fact that ethanol would impair styrene's diffusion into the bead is not immediately obvious from the differences in the dielectric constants of the two systems (ethanol is a high outliner in **figure 141**), because it would be anticipated that the high polarity of EtOH, and ethanol's ability to establish hydrogen bonds, would disfavour the dissolution of styrene. However, computational studies of aromatic-water interactions in water-benzene clusters have highlighted the possibility for the formation π-HB interactions.[390] In 2011 Gierszal *et al* combined Raman multivariate curve resolution spectroscopy (Raman-MCR) with quantum and classical computational calculations to investigate the nature of π-hydrogen bond (π-HB) interactions in liquid water.[390] The report indicates that π-HB interactions are enthalpically unfavourable (ΔH > 0) but are entropically favourable (ΔS > 0) as the formation of a π-HB interaction liberates a water molecule.[390] Although further investigations are required into the nature of aromatic-alcohol interactions, the low

diffusion coefficients of EtOH, MeOH, and CHES, coupled with the modest reaction yields, indicate that styrene-solvent solute-solvent interactions when the solvent possess hydrogen bonding donors are stronger than otherwise anticipated merely from polarity considerations. It is also interesting to note that when isopropyl alcohol (IPA) is employed as the solvent both the diffusion coefficient of styrene and the reaction yield increase, a relationship that would not be expected when solely considering dielectric constants ($\varepsilon_{IPA} < \varepsilon_{EtOH} < \varepsilon_{MeOH}$). The observation that alcoholic solvents can stabilise aromatic solutes is supported by the modest changes in the enthalpy of solution (ΔH_{sol}; kcal mol^{-1}) accompanying the dissolution of various aromatic solutes in MeOH, ^{t}BuOH, and 1-octanol, reported by Stephenson and Fuchs, where the magnitude of the enthalpy change upon dissolution increased when a benzene core was functionalised with electron-donating substituents.[391]

The catalytic proficiency of AE-C45 towards cyclopropanation of styrene is significantly enhanced when the assays are conducted in the polar aprotic solvents THF, acetone, and ACN, with all three solvents affording product yields and enantioselectivities of >89%. In contrast, the two other polar aprotic solvents employed, DCM and ethyl acetate impaired the catalytic efficiency of the reaction (although without any significant decrease in enantioselectivity), reporting product yields of approximately 50 and 20% respectively. From **figure 140**, it is clear that absolutely no relationship can be established between the solvent's dielectric constant or the acquired product yield, meanwhile **figure 141** exhibits a clear linear relationship between the diffusion coefficient of styrene in the respective solvent and the acquired product yield: EtOAc < DCM < THF, acetone and ACN. One possible factor influencing this trend is the relative solubilities of the solvents in water; THF, acetone, and ACN are highly miscible with water, whereas DCM and ethyl acetate are highly immiscible. The miscibility of THF, acetone, and ACN could facilitate the diffusion of styrene across the alginate-solvent interface by provide a stabilising 'solvation shell' which encapsulates styrene from the aqueous internal environment of the alginate bead and visa-versa. In addition, THF, acetone, and ACN possess hydrogen-bonding acceptors which would provide stabilising solute-solvent interactions. DCM and ethyl acetate would be incapable of providing a stabilising solvation shell to styrene to facilitate the solvent-alginate interface and assist in catalysis. Ethyl acetate does contain a carbonyl oxygen which is capable of forming hydrogen bonds with water molecules, however, such an interaction is considerably weaker, relative to THF, acetone, and ACN, on account of the hydrophobic alkyl substituents flanked on either side of the carbonyl group (and accounts for the immiscibility of ethyl acetate in water). The bulky hydrophobic chains would favourably solvate styrene (in addition to some possible dipole-π interactions) resulting in styrene preferentially remaining in the bulk organic phase. DCM is incapable of forming hydrogen bonds with water; although, it is possible that DCM could

establish weakly stabilising π-halogen bonds (π-HaloB) with styrene, in a fashion analogous to π-HBs, which may partially account for the position of DCM in **figure 141**.

6.4.4.1.2 Enantioselectivity: protein rigidification vs denaturation

The enantioselectivity of the AE-C45 assays fall within a range of 53-95%, six solvents (acetone, CH_3CN, DCM, EtOH, IPA, and THF) have higher enantioselectivities (>89%) relative to CHES buffer. Interestingly, the four solvents (EtOAc, MeOH, n-hexane, and pether) with lower enantioselectivities compared to AE-C45 catalysed cyclopropanation in CHES buffer do exhibit lower product yields, but DCM and EtOH have enhanced enantioselectivities when compared to CHES buffer despite the appreciable decrease in catalytic efficiency. There is no relationship between the product yield and the enantioselectivity of the AE-C45 catalysed cyclopropanations, indicating the observed alterations of the stereoselectivity is a function of protein dynamics and not substrate diffusion.

It is established in the literature that organic solvents alter protein dynamics either by destabilising the tertiary structure of the protein and encouraging denaturation or by stripping essential water molecules from the internal structure of the protein and causing a rigidification effect (*chapter 6*, *section 6.1*).[361,363,370] If an organic solvent destabilises the tertiary structure of a protein and facilitates denaturation the chiral environment provided by the coordination sphere of the heme cofactor will also be disrupted. In such a situation, any energetic preference associated with the trajectory of approach of a substrate, such as styrene, to the reactive intermediate will also be disrupted. In addition, additional conformations of the metallocarbenoid intermediate that are energetically inaccessible in a well-folded protein will be permitted in the absence of a well-defined active site. In such a case, the changes in the protein flexibility, dynamics, and conformational ensemble caused by the organic solvent would translate to a decrease in observed enantioselectivity, with the magnitude of the decrease being a function of the magnitude of disruption the solvent causes. If the solvent causes a rigidification of the protein, then two scenarios are possible.[363,367,370] The first is that rigidification causes a 'shrinking' effect which enhances the selectivity but decreases the overall activity as a consequence of the active site shrinking, as is observed for EtOH and DCM. The second is the protein preferentially adopts a particular conformation (over all possible conformations adopted in equilibrium in solution) with a well-defined active site. The rigidification causes a 'stiffening effect' which results in a well-defined, less conformationally dynamic active site which translates to enhanced/decreased activity and or enantioselectivity. At present, the relationship between the enantioselectivity observed in the AE-C45 catalysed cyclopropanation reaction and the solvent the reaction is conducted in remains ambiguous. Additional experiments into how different organic solvents influence the structure, stability, and dynamics of C45 are required to elucidate the observed

results. In particular, CD and NMR spectroscopy of C45 in various organic solvents could provide some valuable insights.

6.4.5 Piperidine N-H insertion

The cyclopropanation assays addressed above demonstrated that AE-C45 remains catalytically active and proficient toward not only towards metallocarbenoid formation but also carbene transferase chemistry. In addition, a relationship between the effective diffusion coefficient for a substrate in a given solvent correlates reasonably well with the observed product yield. To further investigate the catalytic activity and relationship between substrate diffusion and product yield, AE-C45 was employed as the catalyst for the N-H insertion reaction between piperidine and EDA; the reaction was chosen because it has already been demonstrated that free C45 can efficiently catalyst the reaction (chapter 5, *section 5.2.1.2*), and because piperidine possess drastically different properties to styrene.

The reaction yields for the N-H insertion assays (determined from C18-HPLC) between piperidine and EDA catalysed by AE-C45, and performed in a variety of different solvents, are presented in **figure 142** and tabulated in **table 13**. It can be readily seen that, for the vast majority of solvent systems, most the product yields are >50%, showing a decrease in overall catalytic activity relative to the activity observed when free C45 is employed as the catalyst. It can also be seen that the best solvents for the styrene cyclopropanation reaction (THF, IPA, CH_3CN) are poor solvents for the N-H insertion reaction (acetone was not used as a solvent for the piperidine N-H insertion assays on account of the favourable formation of the Schiff base intermediate between the two substrates).[392] In direct contrast to the cyclopropanation assays, the product yields acquired for the N-H insertion assays conducted in toluene, pether, and n-hexane are remarkably high, with all three assays consistently exceeding 90% product formation.

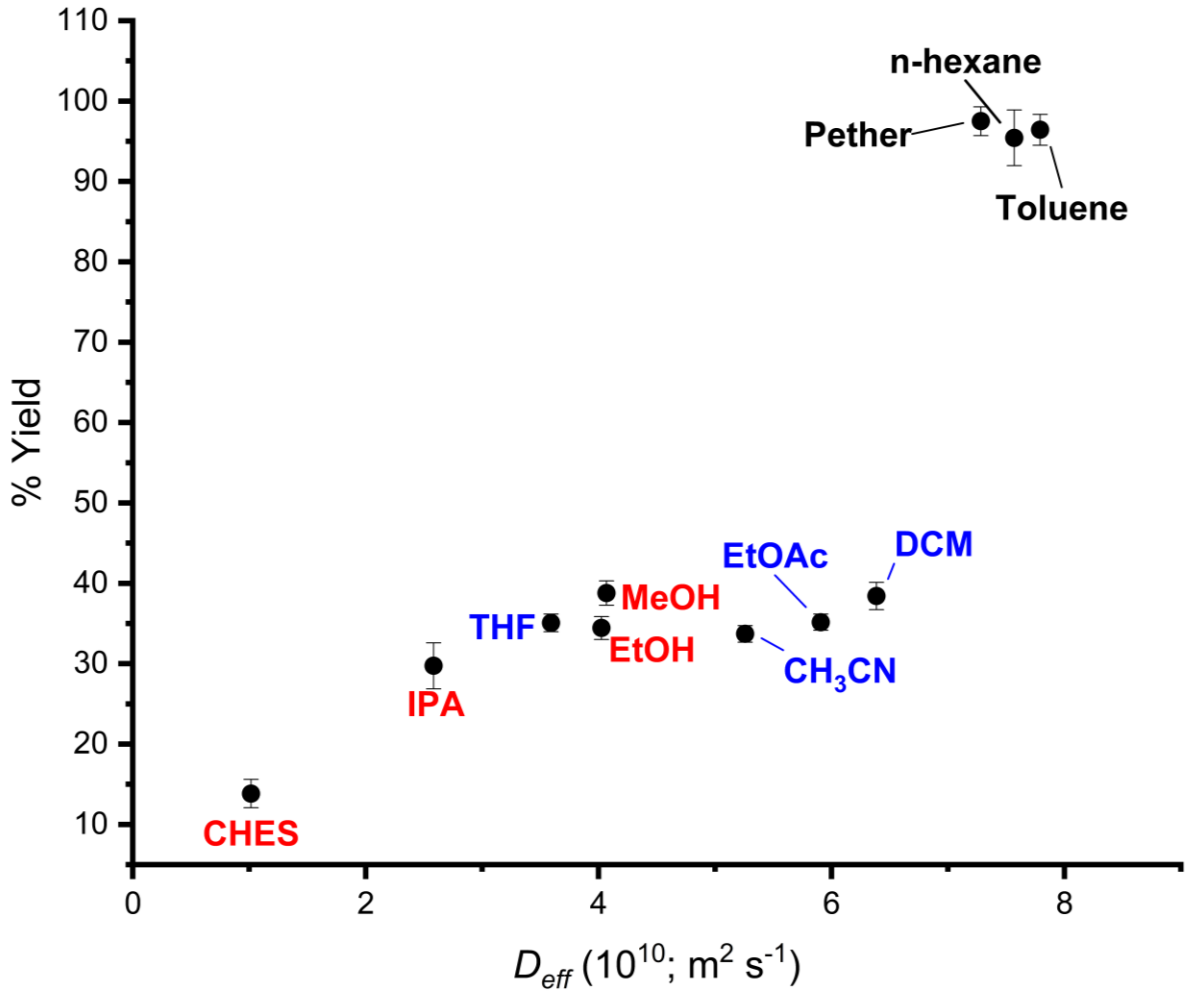


Figure 142: The product yields for the ethyl-1-piperidineacetate product obtained from the AE-C45 catalysed N-H insertion reaction between EDA and piperidine as a function of the dielectric constant of the solvent employed in the reaction.

The calculated effective diffusion coefficient for piperidine in each solvent is tabulated in **table 13** and **figure 142** presents the effective diffusion coefficients vs product yield for each solvent. It can be readily seen that **figure 142** demonstrates a reasonable linear relationship between product yield and D_{eff}, paralleling the observations made in the styrene cyclopropanation assays; however, there are some noticeable differences. For piperidine, the values for the D_{eff} extend across a much broader range (1.0-8.0 x10^{10} m^2 s^{-1}) relative to styrene (0.1-3 x10^{10} m^2 s^{-1}), and is probably a consequence of the N-H groups present in piperidine as multiple ionisation states of piperidine are possible in addition to a hydrogen-bonding acceptor and donor; a conjecture supported by the location of CHES in **figure 142**. The presence of the hydrogen-bonding acceptor and donor in piperidine also provides a possible account for the high yields observed in apolar solvents. The unfavourable interactions between piperidine and an apolar solvent would make the displacement of piperidine into the aqueous bead

more thermodynamically favourable and would increase the rate of diffusion. Overall, the AE-C45 catalysed N-H insertion assays support the postulation that substrate-solvent interactions in both the organic and aqueous phases (which is translated into the partition and diffusion coefficient) play a significant role in heterogeneous enzymatic processes.

Table 13: Reactions yields for the alginate-encapsulated C45 catalysed reaction between piperidine and EDA in various organic solvents, partition coefficient and diffusion coefficients for piperidine in various organic solvents containing in alginate beads prepared from 3% (w/v) sodium alginate and 0.3M $CaCl_2$; R = radius of the alginate beads; alpha parameter is defined by equation (**18**); all values are within a 10% error margin.

Solvent	%Yield	Partition coefficient, K	α	R (mm)	D (x10^{10}, m^2 s^{-1})	D_{eff} (x10^{10}, m^2 s^{-1})
CHES	13.84	0.082	5.887	1.4	1.142	1.015
MeOH	38.79	0.497	35.497	1.4	4.576	4.068
EtOH	34.44	0.483	34.531	1.4	4.528	4.025
IPA	29.74	0.2.00	14.268	1.4	2.906	2.584
ACN	33.71	1.105	78.907	1.4	5.915	5.258
EtOAc	35.17	1.782	127.320	1.4	6.646	5.908
THF	35.08	0.368	26.290	1.4	4.040	3.592
DCM	38.41	2.595	185.349	1.4	7.183	6.386
Hexane	95.42	7.208	514.881	1.4	8.514	7.569
Petroleum ether	97.50	5.567	397.619	1.4	8.192	7.283
Toluene	96.43	8.850	632.143	1.4	8.764	7.791

6.4.6 Enzyme recoverability

The differences observed between the AE-C45 cyclopropanation and N-H insertions also highlight the possibility of affording a degree of control over the observed reactivity. In a multi-step synthesis, which uses the same alginate-encapsulated enzyme for two individual steps, the reaction system could be intelligently engineered to circumvent possible issues associated with chemoselectivity, or to preferentially react with one substrate out of many possible substrates, by carefully selecting a suitable solvent for each reaction step. Although two batches of identical alginate-encapsulated enzymes could feasibly be employed, it would be highly beneficial if the enzyme could be recovered after the first round of catalysis and recycled for a second round of catalysis. This process is called *catalyst/enzyme recoverability* and is highly desirable from a biotechnological and industrial perspective. One of the major advantages of heterogeneous catalysts vs homogenous catalysts is the ease at which heterogeneous catalysts can be recovered on account of occupying a separate phase. In the case of AE-C45, the alginate-beads are insoluble and a wide range of solvents, and recoverability simply entails decanting the reaction mixture after a reaction to isolate/wash/prepare/store/preserve the catalyst for future use. As AE-C45 has already demonstrated a degree of recyclability in the

peroxidase assays, in which AE-C45 could be reused a total of four times, it was particularly pertinent to see whether AE-C45 could be recovered, washed and reused as a carbene transfer heterogeneous catalyst.

The cyclopropanation reaction between EDA and styrene, catalysed by AE-C45, was selected as the initial reaction to screen for AE-C45 recoverability. The three best solvents (acetone, ACN, and THF) were selected for the assays. The reactions were conducted in accordance with the AE-C45 cyclopropanation assays as detailed in chapter 9, *section 9.12.2*. Briefly, the reactions were conducted in 2 ml sealed-top vials. 970 µl of solvent (acetone, ACN, or THF) was added, followed by three alginate beads containing C45. The vials were purged under nitrogen, before 10 µl of $Na_2S_2O_4$ (1M stock in ddH_2O) and 10 µl of styrene (3M stock in EtOH) was added. The mixture was allowed to stir for 1 minute before 10 µl of EDA (1M stock in EtOH) was added to initiate the reaction. After two hours of stirring on a roller, the vials were unscrewed (under a nitrogen atmosphere) and the product mixture was pipetted into a 1.5 ml Eppendorf for purification/analysis. The three alginate beads containing C45 were washed three times with the respective solvent (acetone, ACN, or THF) before the second/third round of catalysis was prepared.

Figure 143 displays the enantioselectivities and product yields for all three runs, for each solvent as a bar chart to facilitate comparison. As anticipated, the first run for each respective solvent afforded decent-to-high product yields (80-90%) and enantioselectivities of >90% determined, in each solvent, for the (*R,R*)-cyclopropane product. In terms of product yields, the second and third runs for each solvent predictably decreased, with yields between 45-65% observed for the second runs and 18-55% for the third runs. Acetonitrile delivered the best performance out of the three systems screened, exhibiting only an approximate 40% decrease in product yield across the three screens. Although the second reactions conducted in THF and acetone resulted in moderate product yields (40-50% respectively), attempts to recover and reuse a third time resulted in low-to-poor yields (15-30%). However, despite the general loss in activity, the initial screens show that AE-C45 remains catalytically active after its initial use and can be easily recovered and successfully redeployed as a carbene transfer catalyst multiple times, allowing 1.5-1.8 times of the desired cyclopropane product to be prepared from the same catalyst relative to a single use catalyst. For the assays conducted in ACN, the enantioselectivity of the reaction was only marginally affected, with an approximate 10% decrease observed after the third reaction. The second and third assays conducted in THF and acetone, however, were characterised by a drastic decrease in enantioselectivity, with moderate values between 50-60% observed during the second runs followed by remarkably low values of ≥20% observed by the third runs. At present, how multiple reuses of AE-C45 affects the enantioselectivity is

unknown. It is possible that swelling of the beads, caused by the diffusion of solvent into the bead, results in the aqueous environment of the bead becoming gradually more organic over time, which could provide an environment more accommodating to hydrophobic solutes and incidentally destabilise the maquette by causing moderate changes to its conformation. Why this would result in such pronounced effects for THF and acetone, but not for ACN, currently remains a mystery. It is also well known that proteins become more rigid in the presence of organic solvents, which would account for the decreasing reactivity observed in each successive run, but does not provide a satisfactory answer for the decrease in enantioselectivity, as it would be expected that protein rigidification would increase the observed enantioselectivity. It is possible a combination of rigidification and conformational changes in the active site result in the observed enzymatic activity and selectivity in each successive run. It is also possible that modification to the alginate structure over time, either *via* swelling, binding of substrate(s), and reacting with substrates, could impart conformational changes to the protein. It is unfortunately absolutely clear that additional experiments are essential before a molecular level explanation can be provided to account for the above observations.

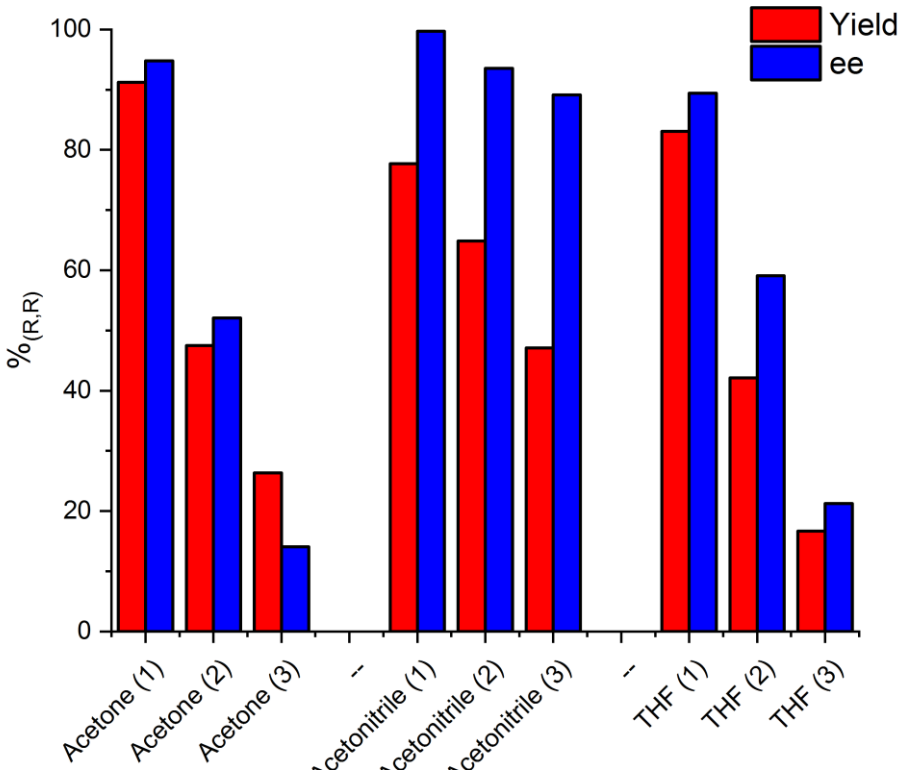


Figure 143: The product yields and enantioselectivities for the successive AE-C45 encapsulated cyclopropanation reactions between styrene and EDA in acetone, acetonitrile and tetrahydrofuran (all errors are within 10%).

After the above assays were performed, a concern to whether the data could be accounted for by ''product leaking'' was raised. It is possible that the diffusion of the cyclopropane product out of the alginate bead is remarkably slow, and in the second/third runs the observed peak corresponds to (*R,R*)-product, which had accumulated inside the bead after the initial run, leaking out of the bead and producing a false positive observation. An additional experiment was conducted, in the same three solvents, in which the first reaction catalysed was the cyclopropanation reaction reported above, followed by the subsequent recovery and re-deployment of the same AE-C45 beads to catalyse the N-H insertion reaction between piperidine and EDA. The reactions were conducted under identical conditions to those previously reported, with the only modification being the reactions were left mixing for one hour instead of two.

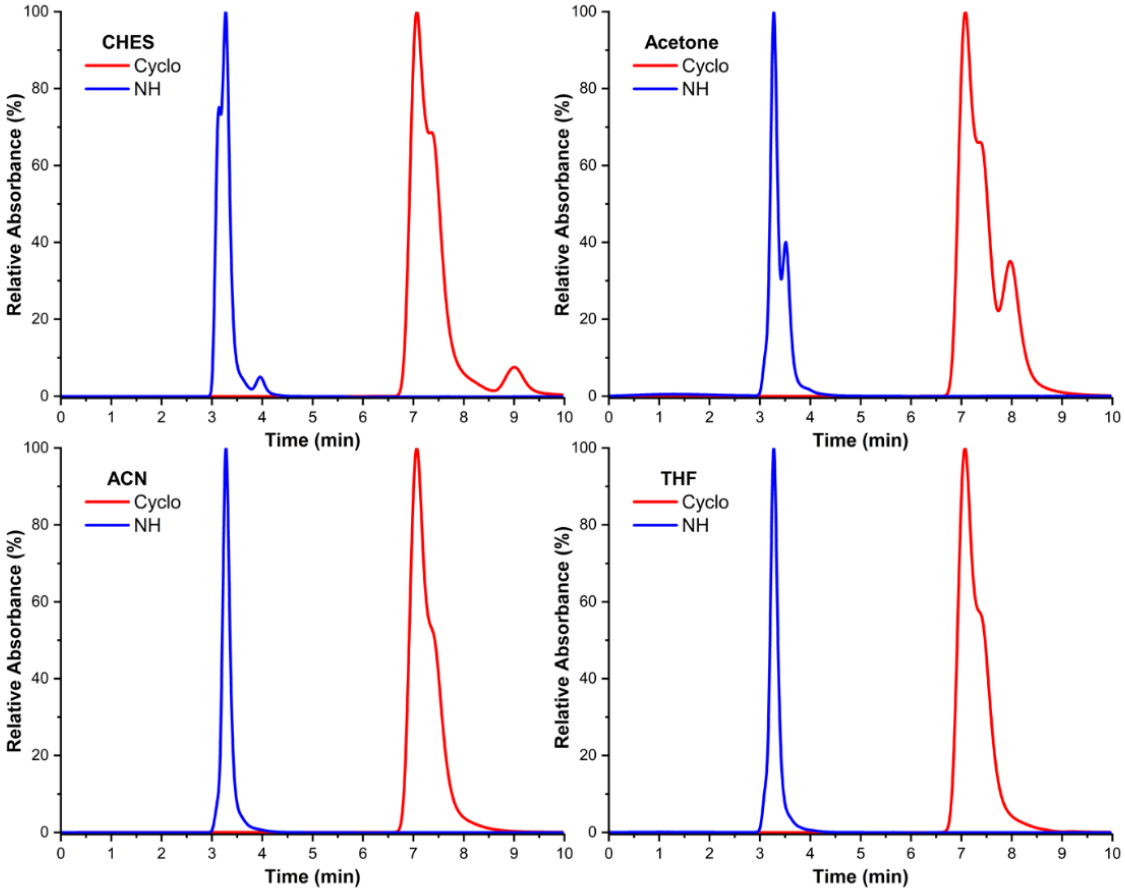


Figure 144: The normalised chiral-HPLC chromatograms for the AE-C45 catalysed cyclopropanation reaction between styrene and EDA (red) followed by the recovery and subsequent use of the same bead in the AE-C45 catalysed N-H insertion reaction between piperidine and EDA (blue) in (**top left**) CHES buffer (pH 8.6), (**top right**) acetone, (**bottom left**) acetonitrile, and (**bottom right**) tetrahydrofuran. A polar organic mobile phase (100% MeCN: 0.1% (v/v) TFA:0.1% (v/v) Et_3N) was employed and injection volumes were 2 µl; all chromatograms are recorded at 254 nm.

The normalised chromatograms for the AE-C45 cyclopropanation reaction between EDA and styrene followed by the recovered AE-C45 catalysed N-H insertion reaction between EDA and piperidine, for four solvent systems (CHES, acetone, ACN, THF), are presented in **figure 144**. The calculated yields for both the cyclopropanation and N-H insertion reactions, and the enantioselectivities of the cyclopropanation reaction, in each solvent system are tabulated in **table 20** (*appendix 2*). It can be seen from **table 20** (*appendix 2*) that all four cyclopropanation assays proceeded with moderately high yields (62-79%) and enantioselectivity (86->99%). After isolating, washing and re-deploying the AE-C45 beads used in the cyclopropanation reaction, a peak in each respective chromatogram, with a retention time similar to the anticipated N-H insertion product, ethyl-1-piperidine acetate, can be readily detected. From the chromatograms, it was calculated that reusing the AE-C45 beads resulted in an average 10% yield of ethyl-1-piperidine acetate across all four solvent systems. No cyclopropanation product was detected in any of the N-H insertion product mixture chromatograms. The ability to detect the N-H insertion product in the second run, although in poor yields, highlights two important things: i) the initial results acquired from the aforementioned cyclopropanation recoverability screens were not a false positive caused by the leaking of cyclopropane product, and ii) conclusively confirms that AE-C45 remains active towards carbene transferase chemistry in successive reactions after being recovered from an initial application.

6.5 CONCLUSION

Throughout chapters 3-5 the catalytical proficiency of C45 towards abiological chemistry, proceeding *via* a metallocarbenoid intermediate, was explored. C45 was demonstrated to be proficient towards cyclopropanation, chemoselective N-H insertions, Si-H insertions, carbonyl olefinations, and ring expansions reactions. In this chapter, the focus was on developing a methodology for expanding the biocatalytic potential of C45 to include reactions conducted in neat organic solvents. The method of entrapment was employed to prepare heterogeneous catalysts for biocatalytic applications. The hydrogel calcium alginate, which is resistant to degradation in organic solvents, was identified as a suitable medium for encapsulating C45. Alginate-encapsulated (AE) C45 was initially subjected to spectroscopic characterisation, in a variety of organic solvents, which revealed the spectroscopic profile of ferric and ferrous C45 remains unperturbed once encapsulated inside calcium alginate hydrogels, or when AE-C45 is suspended in pure organic solvents such as MeOH and n-hexane. In addition, it was demonstrated that AE-C45 retains its ability to form a stable metallocarbenoid intermediate, yielding an electronic spectrum identical to the C45:EDA metallocarbenoid spectrum observed in aqueous solution. It was subsequently demonstrated that AE-C45 remains catalytically active towards catalysing peroxidation chemistry. AE-C45 was also proficient at carbene transferase

chemistry, catalysing the cyclopropanation of styrene and the N-H insertion of piperidine in an assortment of organic solvents. Although the reaction yields were affected by the encapsulation of C45, a relationship between substrate diffusion and product yield was observed, indicating the changes in catalytic performance of AE-C45 relative to C45 can be account for by the inclusion of a substrate/product diffusion process across the solvent:hydrogel interface. It was also demonstrated that AE-C45 can be recovered after one catalytic cycle, and re-employed as the catalyst in subsequent reactions, although at the cost of product yield and enantioselectivity. Overall, in addition to demonstrating the catalytic versatility of C45, it has been demonstrated that AE-C45 can be employed (and recovered for future use) as a biocatalyst to perform important chemical reactions, *via* an abiological metallocarbenoid intermediate, in pure organic solvents.

CHAPTER 7: EVOLUTION OF C45 TOWARDS IMPROVED AND STEREODIVERGENT CARBENE TRANSFERASE ACTIVITY (In collaboration with Jack Steventon)

7.1 INTRODUCTION

In chapter 2 *section 2.3.1* a general outline of a technique used in molecular biology called mutagenesis was introduced. A few mutagenesis methods, including random mutagenesis, site-directed mutagenesis, and saturation mutagenesis, were also introduced and briefly outlined. The technique of *direct evolution* (DE) was then introduced as a method for generating large numbers of mutant libraries. It was explained that DE mimics natural evolution, with the selective pressure being defined by the desired functionality. DE involves iterative rounds of mutagenesis to create mutant libraries that are screened to identify variants that exhibit the desired functionality. The best variant(s) are then amplified to serve as a/the template(s) in subsequent rounds of mutagenesis. The successful redesign of naturally occurring hemoproteins towards carbene transferase activity *via* mutagenesis and DE was subsequently introduced in chapter 2 *section 2.3.1*.

The heme-containing tetra-α-helical maquettes are characterised by a simplified and easily evolvable protein scaffold. In the Anderson lab, Jack Steventon had been using DE to generate novel libraries of tetra-α-helical maquettes in attempts to improve on the peroxidase activity exhibited by C45. This work occurred concurrently alongside research into the carbene transferase activity of C45. Inspired by the success of Arnold, Fasan, Hartwig, and others in redesigned naturally occurring hemoproteins towards a broad range of carbene transferase activities, the carbene transferase activity exhibited by C45, and the pre-existing methodology for generating large libraries of C45 mutants (established by Jack Steventon and Peter Wilson) the first question raised concerned whether the carbene transferase activity (and/or stereoselectivity) of C45 could be enhanced through mutagenesis and/or DE. Alongside using DE to enhance the stereoselectivity of a given reaction, it has also been demonstrated in the literature that select, key mutations can actually change the stereoselectivity exhibited by an enzyme towards a certain reaction. Therefore, whether a novel mutant with differing stereoselectivity towards the cyclopropanation reaction (compared to C45) could be engineered and isolated was the second question raised.

7.2 RESULTS AND DISCUSSION

7.2.1 Carbene transferase screening the first library – JR1

A library of C45 mutants was generated from error-prone PCR by Jack Steventon and were subsequently screened for peroxidase activity using the standard ABTS assay. Three mutants, in

particular, JR1, DH6, and AG5 all reported decrease peroxidase activity relative to C45. As peroxidase activity was the primary functional target influencing the design of C45, and oxygen binding inhibits metallocarbenoid formation, it was rationalised that the three poorest mutants could exhibit enhanced cyclopropanation activity relative to C45. The sequences of the three mutants compared to the sequence of C45 are presented in **figure 145**. JR1 has 2 mutations relative to C45: G67S and G92S. The mutation G67S occurs in the loop region connecting helix B to helix C and the mutation G92S occurs in the loop region connecting helix C to helix B. It was of interest to determine how mutations to the loop region influenced the overall flexibility and conformational dynamics of JR1 and whether the mutations would influence the enantioselectivity of the cyclopropantion reaction. The mutants DH6 and AG5 each possess a single mutation in helix A; The mutation in DH6 is F25Y and the mutation in AG5 is Q5H; both mutations introduce a new hydrogen bonding residue into helix A.

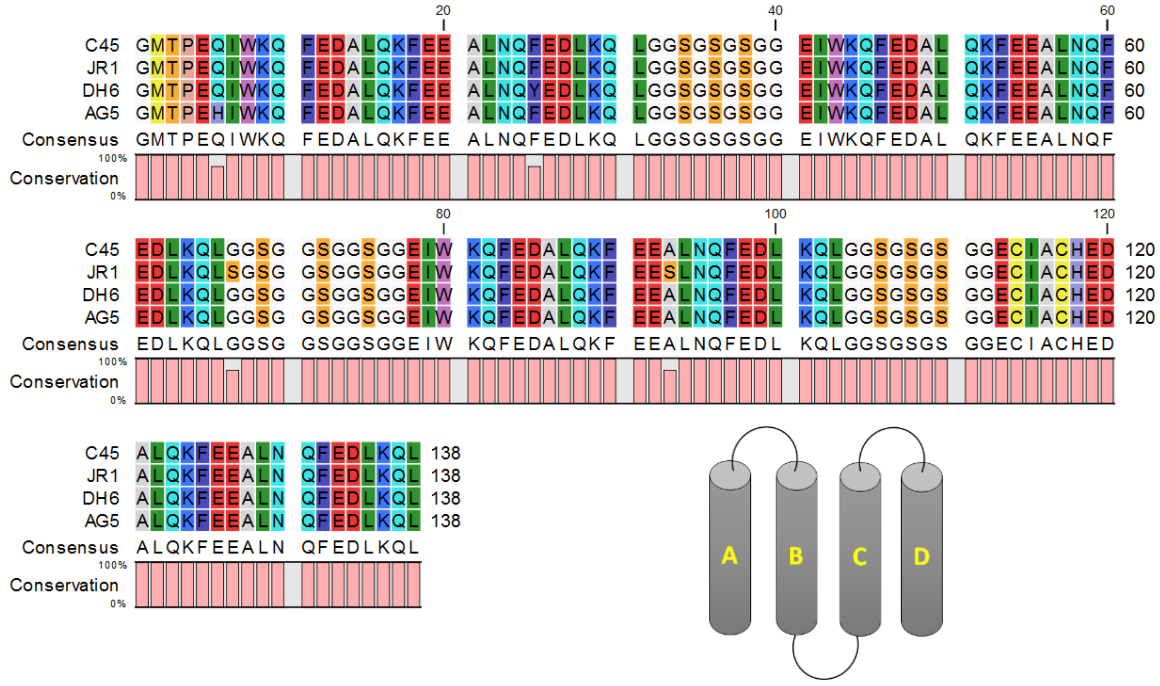


Figure 145: The primary sequences of the four mono-heme containing tetra-α-helical maquettes C45, JR1, DH6, and AG5 and the labelling of the four helices A-D (mono-histidine ligated heme cofactor appended to helix D).

The three mutants were each initially screened for carbene transferase activity using the protocol previously reported in chapter 4 for cyclopropanation, with the average respective yields and enantioselectivities (for the (*R,R*)-cyclopropane product) being tabulated in **table 14**. It can be immediately concluded that neither AG5 or DH6 possess any activity towards cyclopropanation, with reaction yields consistently >10% and negligible to poor enantioselectivities. The exact reason for the lack of activity in AG5 and DH6 still needs to be addressed, but it is possible that the introduction of a new hydrogen bonding residue in the A helix (relative to C45, **figure 145**), which was observed for

both mutants, negatively alters the dynamics of the heme active site or impairs substrate diffusion. The double mutant, JR1, however, exhibits a slightly elevated reaction yield relative to C45 (85.6% vs 80.20%) without any change to the observed enantioselectivity (78% vs 77%).

Table 14: Average product yields, diastereoselectivities, and enantioselectivities for the cyclopropanation reaction between EDA and styrene catalysed by the C45 mutants JR1, AG5, and DH6

	Average			
Mutant	**% Yield$_{(R,R)}$**	**% Yield$_{(S,S)}$**	**dr$_{(E)}$ %**	**ee$_{(R,R)}$ %**
JR1	85.61 (± 4.32)	11.04 (± 4.84)	>99.9	78.00 (± 8.5)
AG5	2.70 (± 0.44)	3.34 (± 4.72)	>99.9	-19.21 (± 27.17)
DH6	5.60 (± 3.42)	4.01 (± 3.63)	>99.9	34.87 (± 61.73)

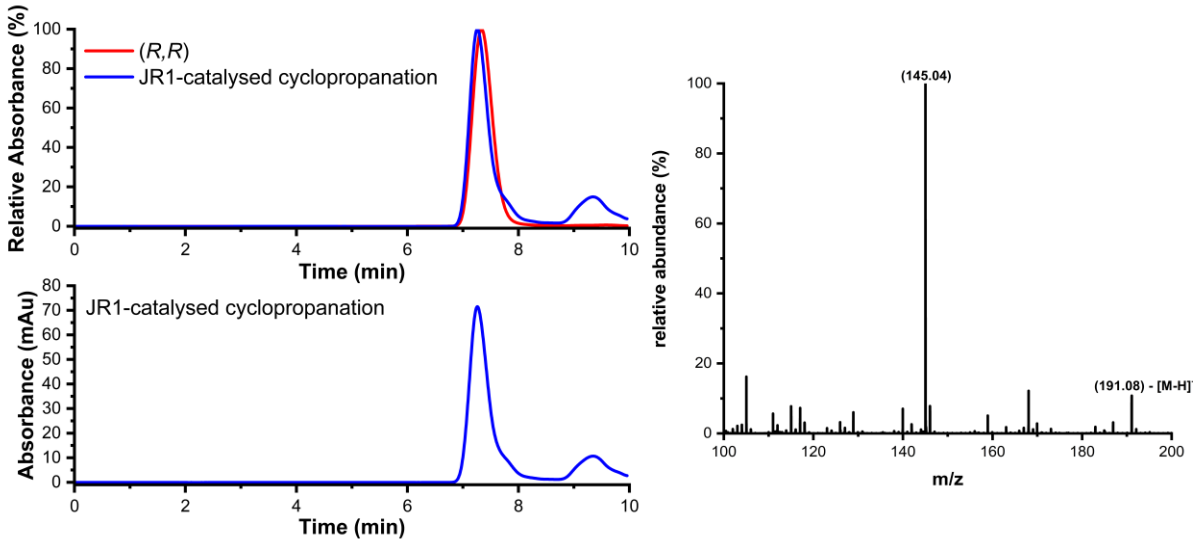


Figure 146: (**left**) Chiral-HPLC chromatograms for the JR1-catalysed (10 µM, 0.1 % catalyst loading) cyclopropanation assays between styrene (30 mM) and EDA (10 mM) (CHES buffer, pH 8.6, 280 nm). A polar organic mobile phase (100% MeCN: 0.1% (v/v) TFA:0.1% (v/v) Et_3N) was employed and injection volumes were 2 µl. (**top left**) normalised chiral-HPLC chromatograms for commercial (*R,R*)-ethyl 2-phenylcyclopropane-1-carboxylate (280 nm, EtOH, red) vs JR1-catalysed cyclopropanation (280 nm, blue); (**bottom left**) Averaged chromatogram from the JR1-catalyzed cyclopropanation assay. (**right**) LC-MS spectra of JR1-catalyzed cyclopropanation assay products exhibiting the dominant oxonium ion fragment at 145 m/z. The spectrum was recorded in ES+ mode and monitored at 254 and 280 nm. A C8 column was employed for the LC separation with a gradient mobile phase (95:5:0.1 % v/v water/MeCN/formate 10:90:0.1 % v/v water/MeCN/formate). Injection volumes were 20 µl.

The ESI-MS of the product extracted from the JR1-characterised cyclopropanation reaction between styrene and EDA is presented in **figure 146** and exhibits the characteristic and well-defined fragmentation profile of the expected cyclopropane product. Although the increase in the reaction

yield is small, analysis of JR1 may still prove useful in providing information on how the structure of the maquette influences the carbene transferase activity, which will undoubtedly aid future engineering efforts. JR1 was therefore subsequently biophysically characterised and the metallocarbenoid formation and cyclopropanation reaction was kinetically characterised.

7.2.1.1 Biophysical characterisation of JR1

JR1 (**figure 147**) was characterised by UV-VIS spectroscopy, redox potentiometry and circular dichroism to determine its UV-VIS absorption profile, standard reduction potential and thermal stability. The information determined was compared against C45. In addition, an MD simulation of JR1 was acquired (**figure 147**).

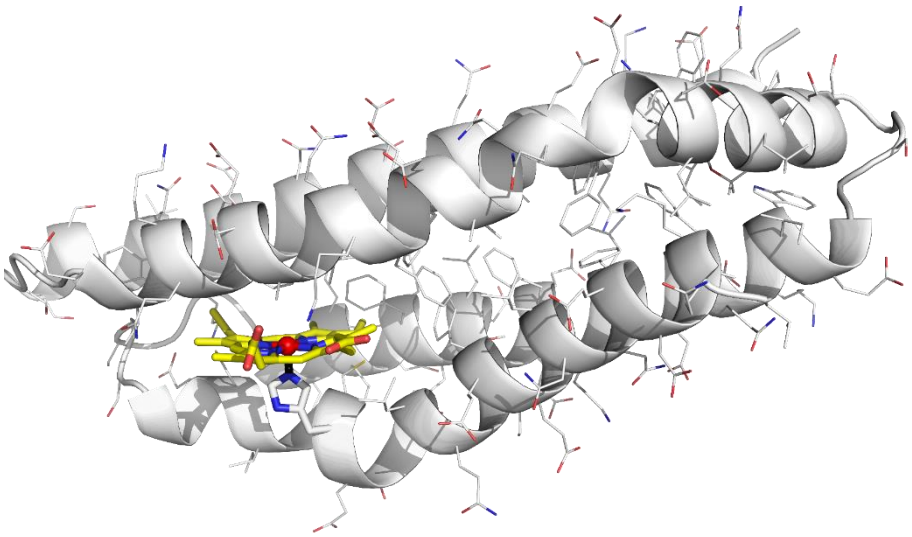


Figure 147: The structure of JR1 acquired after MD (ABMER) simulations.

The UV-VIS spectrum for ferric and ferrous JR1 is presented in **figure 148**. The UV-VIS spectrum for ferric JR1 was dominated by the characteristic Soret peak at 406 nm and was accompanied by a Q-band series of less intense peaks in the range 500-550 nm. Upon the addition of sodium dithionite, a new spectroscopic profile emerged and is attributed to ferrous JR1. The reduced spectrum was dominated by a red-shifted intense Soret peak at 418 nm which is accompanied by a Q-band with two less-intense but well-resolved peaks, the α and β peaks, with the intensity ratio β > α. The UV-VIS spectra for JR1 is consistent with that of c-type mono-histidine ligated tetra-α-helical maquettes, with a spectroscopic profile that parallels the spectra observed for C45 and for *Rma*-TDE. The redox potentiometry for JR1 results in a calculated mid-point potential of -197 mV (vs SEP), meaning JR1 is more reducing than C45 (-176 mV vs SEP)[235] and is possibly a reflection of minor conformational changes to JR1 relative to C45. The more negative mid-point potential could also reflect a small change

in the relative distances of the heme cofactor and the first-coordination sphere amino acids, with a 21 mV decrease suggesting either a lengthening and/or weakening of key bonds governing the reduction potential of JR1 relative to C45. This postulation is corroborated by the 6 ^{O}C decrease in the melting temperature for JR1 relative to C45 (80^{O}C vs 86^{O}C respectively)[235] calculated from circular dichroism **figure 148**. The information extracted from the biophysical analysis of JR1 indicates the combination of glycine to serine mutations in the loops connecting helices A-B and B-C engenders a maquette, JR1, with a slightly increased backbone flexibility.

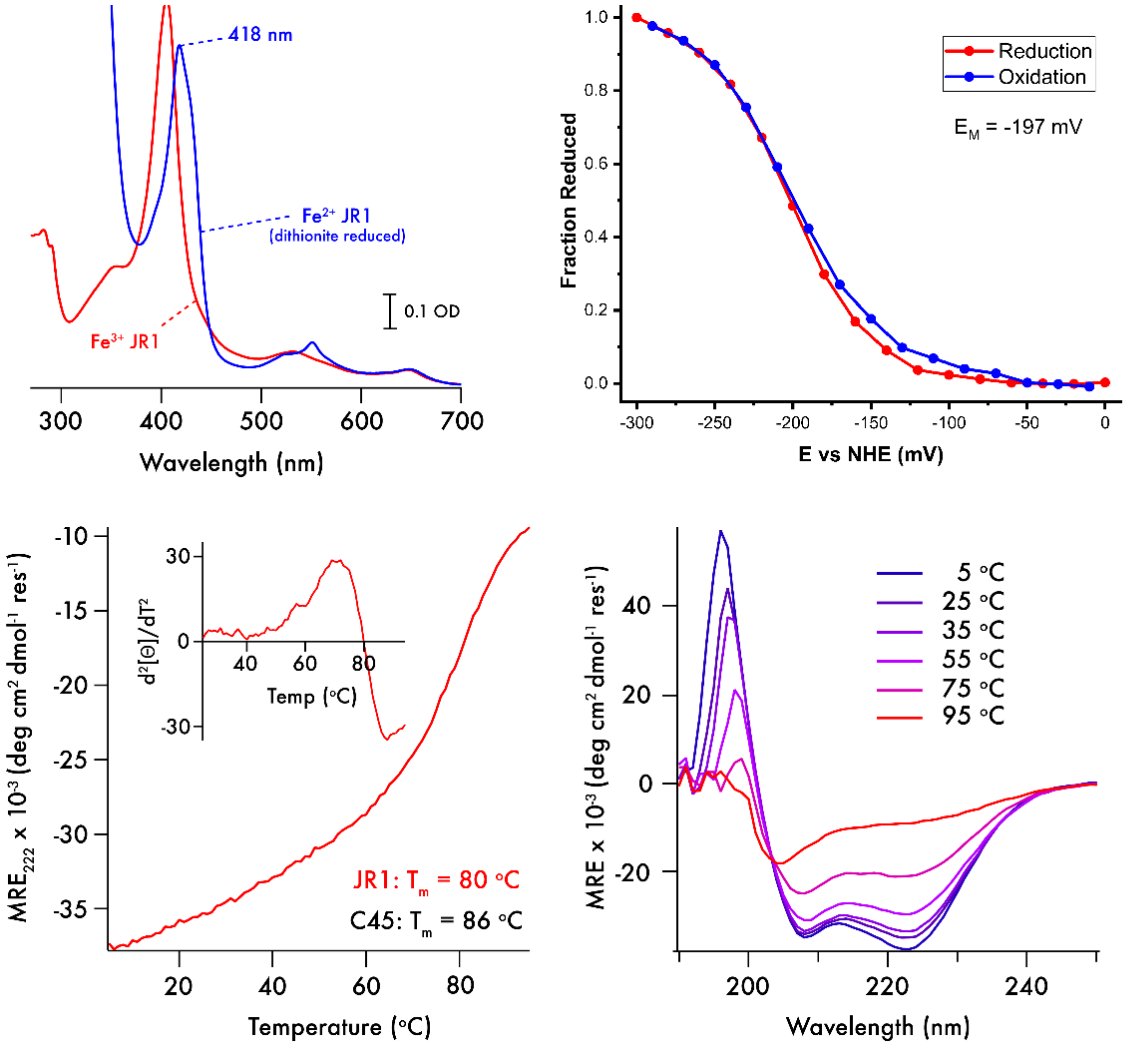


Figure 148: Biophysical characterisation of JR1. (**top left**) The UV-VIS spectrum for ferric (red) and ferrous (blue) JR1, (**top right**) Redox potentiometry of JR1 (100 mM KCl, 50 mM CHES, 10% glycerol, pH 8.6) collected with redox mediators as described in the methods, (**bottom left**) Far-UV circular dichroism spectra of JR1 with varying temperature (100 mM KCl, 20 mM CHES, pH 8.6), (**bottom right**) Temperature dependence of the CD signal monitored at 222 nm during the thermal denaturation of JR1. The inset shows a smoothed second derivative of the thermal melt trace indicating a melting transition (T_m) of 80°C.

7.2.1.2 Spectroscopic and kinetic characterisation of JR1 metallocarbenoid formation

Having demonstrated a 6.25% percentage increase in product yield relative to C45 in the cyclopropanation assay, and after biophysical examination, the formation of the metallocarbenoid intermediate in JR1 was subsequently investigated. Under identical conditions to those employed for C45 (5^{O}C, 40% EtOH, CHES buffer, pH 8.6), in the presence of EDA the expected shifts in the UV-VIS spectrum corresponding to the formation of a terminal metallocarbenoid complex were observed: the emergence of a red-shifted Soret peak at 428-433 nm accompanied by red-shifting and broadening of the Q-band profile (**figure 149**). The time-dependent spectra for the change in absorbance at 417 and 433 nm, corresponding to ferrous JR1 and the JR1-EDA metallocarbenoid intermediate respectively, exhibited a profile analogous to the spectroscopic changes observed for C45 - a rapid decrease in the ferrous peak (417 nm) concurrent with an increase in the metallocarbenoid peak (433 nm) punctuated by a plateauing of both spectroscopic changes upon achieving saturation. The time-course spectra indicate that, analogous to C45, once formed the metallocarbenoid intermediate is stable and persists, exhibiting no appreciable decay over time. The spectra also indicate that JR1 achieves saturation on a time scale similar to C45 (10 s) (**figure 149**).

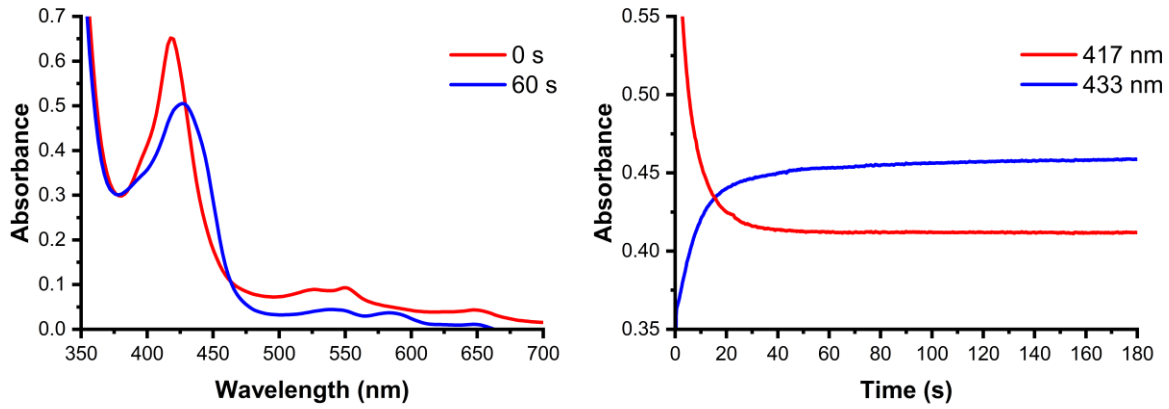


Figure 149: (**Left**) Metallocarbenoid formation and stability in the absence of styrene substrate. The UV-VIS spectra of JR1 (7.5 µM, CHES buffer, pH 8.6, 40% (v/v) EtOH) after 0 (red) and 60 (blue) seconds following rapid mixing of ferrous JR1 with 500 µM ethyl diazoacetate at 5°C. (**Right**) Single wavelength traces represent the time course of ferrous JR1 (417 nm, red, 7.5 µM protein, 10% EtOH) and metallocarbenoid:JR1 adduct (433 nm, blue) following rapid mixing of ferrous JR1 with 500 µM ethyl diazoacetate at 5^{O}C. Once formed, the metallocarbenoid:JR1 adduct persists for the duration of the experiment (1000 seconds).

In order to properly compare metallocarbenoid formation in JR1 to C45, the single turnover kinetic parameters, K_1 and k_{obs} (and the ratio (k_{obs}/K_1)), for the formation of the JR1-metallocarbenoid intermediate must be determined. The kinetics of metallocarbenoid formation in JR1 were

determined using stopped-flow spectroscopy by adopting the protocol used for C45, *Rma*-TDE and Mb(H64V,V68A), which was outlined in chapter 3, section *3.5.1.3*. **Figure 150** exhibits the calculated curve acquired for EDA. As for C45, the analysis of the formation of a metallocarbenoid involves three distinct identities: i) free enzyme and substrate, ii) a tetrahedral transition state, and iii) the metallocarbenoid intermediate. There are three underlining assumptions that have been made to assist in the analysis: i) the rate of decomposition of the tetrahedral transition state to the metallocarbenoid intermediate (k_2) is first-order, ii) the reaction is irreversible ($k_2 >>> k_{-2}$), and iii) a steady-state equilibrium is established between the free enzyme and the tetrahedral transition state. The values for the single turnover kinetic parameters for JR1-metallocarbenoid formation are tabulated, against the same values acquired for C45, in **table 22** (appendix 2).

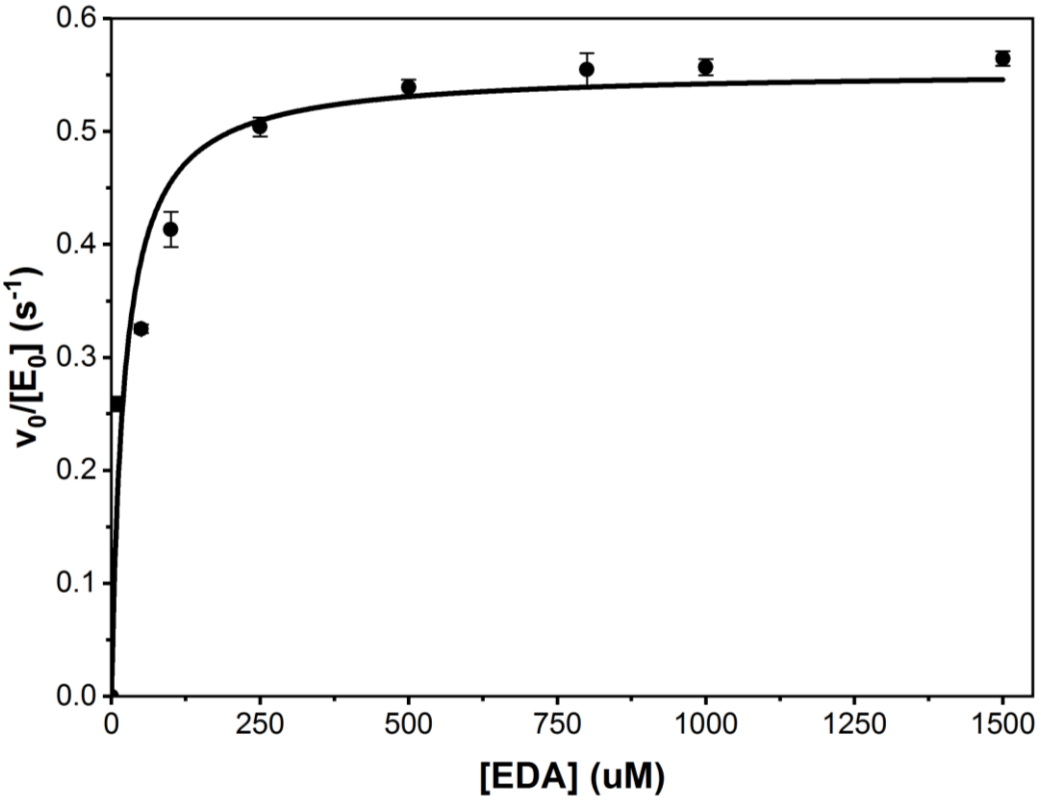


Figure 150: The concentration-dependent kinetic plots of JR1 (7.5 µM, CHES buffer, pH 8.6, 40% (v/v) EtOH) with varying concentrations of EDA. Initial rates were determined from the rate of change in absorbance for the C45:metallocarbenoid adduct at 5°C.

The k_{obs}, K_1, and k_{obs}/K_1 values for JR1 are 0.553 s^{-1}, 21.68 µM and 2.55 x 10^4 s^{-1} M^{-1} respectively, relative to 0.498 s^{-1}, 74.36 µM and 6.70 x 10^3 s^{-1} M^{-1} for C45. The K_1 for JR1 is significantly smaller than for C45, indicating that JR1 possess a binding affinity three-times higher than for C45. Upon reflecting on the biophysical data acquired for JR1, the increased binding affinity is possibly related to a minor increase in backbone flexibility. An increase in flexibility would facilitate the diffusion of EDA into the active site

and would also assist in accommodating the resultant metallocarbenoid intermediate. In addition, the increase in standard electrode potential (-176 vs -196 mV)[235] would make the heme more electron rich relative to C45, which would assist in the hypothesised electron transfer step proceeding metallocarbenoid formation. This is supported by the larger k_{obs} for JR1 metallocarbenoid formation relative to C45 (0.553 vs 0.498 s^{-1} respectively, **table 22** (*appendix 2*)). The change in K_1 and k_{obs} for JR1 alludes to an increased efficiency for the formation of the metallocarbenoid intermediate relative to C45. Although the increase can be partly rationalised with the biophysical data, additional experiments need to be conducted to elucidate the entire picture.

7.2.1.3 Kinetic characterisation of JR1-catalysed cyclopropanation

The formation of the metallocarbenoid intermediate in JR1 is seemingly more efficient relative to C45; however, the single turnover kinetic parameters are limited to the formation of the intermediate and provide no insights or information into the stability and reactivity of the resultant complex. It has already been demonstrated that JR1 catalyses the cyclopropanation reaction between styrene and EDA with a 6.25% increase in reaction yield relative to C45, and so to comprehensively compare the carbene transfer activity of the two maquettes the kinetics of the reaction was studied for JR1. As with metallocarbenoid formation, the kinetic parameters for the JR1-catalysed cyclopropanation reaction were investigated using stopped-flow spectroscopy, adopting the protocol used to determine the analogous C45 kinetic parameters and outlined in chapter 4 *section 4.2.5*.

The experimental conditions employed to investigate carbene transfer activity paralleled the conditions used to investigate metallocarbenoid formation, with a single modification being the introduction of a suitable substrate, styrene. A solution containing ferrous JR1 (15 µM protein, 10 mM $Na_2S_2O_4$, CHES buffer, pH 8.6) and a solution containing ethyl diazoacetate (EDA) and styrene (100 µM -10 mM and 200 µM - 3mM respectively, 20%:80% CHES:EtOH) were prepared in separate compartments (syringes) before a small volume (2x20 µl) of each were injected concurrently into a mixing chamber. The mixing chamber was monitored by UV-VIS spectroscopy over several timescales to monitor any reactions occurring over the duration of the experiment. Final concentrations were 7.5 µM enzyme, stoichiometric $Na_2S_2O_4$, 50 µM - 5 mM EDA, 100 µM - 1.5 mM styrene, and 40% EtOH.

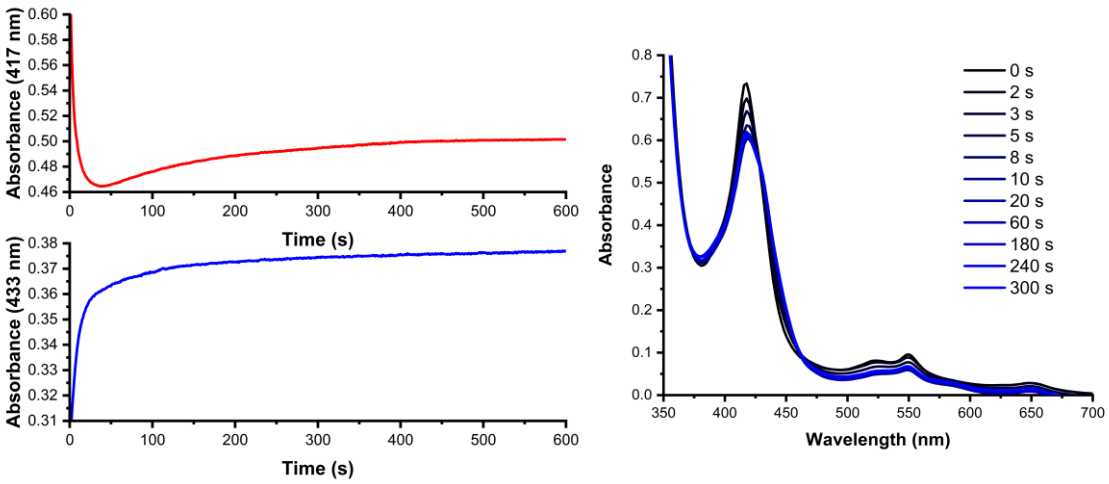


Figure 151: (**left**) Single wavelength traces representing the time course absorbance of ferrous JR1 (417 nm, 7.5 µM protein, 10% EtOH) and the metallocarbenoid:JR1 complex (433 nm) following rapid mixing of ferrous JR1 with 500 µM ethyl diazoacetate at 5^{O}C in the presence of 3 mM styrene. (**right**) UV-VIS traces recorded at various time intervals following the rapid mixing of ferrous JR1 with 500 µM ethyl diazoacetate at 5^{O}C in the presence of 3 mM styrene.

Figure 151 displays the time-course spectra obtained from the stopped-flow experiments for JR1 in the presence of EDA and styrene at wavelengths 417 and 433 nm. 417 nm corresponds to the ferrous JR1 species and 433 nm corresponds to the JR1-EDA metallocarbenoid complex, respectively. The data indicates that in the initial 20 seconds post-mixing there is a drastic decrease in the 417 nm absorbance that is accompanied by a concurrent increase in the 433 nm absorbance – corresponding to the formation of the JR1-EDA metallocarbenoid complex.

It was demonstrated in **figure 149** that the JR1-EDA metallocarbenoid complex is stable and persists over a duration of 600 seconds without any appreciable decay. In the presence of styrene, the initial spectra acquired resembles the spectra acquired in the absence of styrene, corresponding to the formation of the metallocarbenoid intermediate. After approximately 40 seconds, however, the absorbance at 433 nm plateaus and is accompanied by a sharp linear increase in absorbance at 417 nm – indicating the decay of the metallocarbenoid species and the return to the starting ferrous complex. **Figure 151** clearly demonstrates the plateauing of the absorbance changes at 417 and 433 nm in the absence of styrene, corresponding to saturation, is never accomplished in the presence of styrene. UV-VIS traces captured at numerous timepoints over 300 seconds are presented in **figure 151**. In the initial stages of the reaction (1-20 seconds) the UV-VIS exhibits a red-shift corresponding to the formation of the metallocarbenoid intermediate, alongside the concurrent shift and broadening of the Q-band profile. In the UV-VIS spectra acquired in the absence of styrene the red-shift from ferrous to metallocarbenoid intermediate is characterised by sharp Soret peaks as the 417 nm peak

disappears alongside the emergence of a Soret peak at 428-433 nm. In the presence of styrene, the emergence of the 433 nm peak appears as a shoulder alongside the 417 nm, resulting in a broad, split Soret peak, without either the 417 nm or 433 nm peak dominating the spectra. After 20 seconds, the maximum absorbance begins to shift towards the 417 nm peak and is accompanied by the gradual re-emergence of the ferrous Q-band profile. The spectroscopic profile for JR1-metallocarbenoid formation in the presence of styrene parallels the profile exhibited by C45. The analogous spectroscopic profiles mean the kinetic parameters k_{lim}, K_1, and k_{lim}/K_1 for the JR1-catalysed cyclopropanation reaction can be calculated under conditions identical to those employed for C45 (**figure 152**), allowing for comparisons between the two maquettes to be investigated.

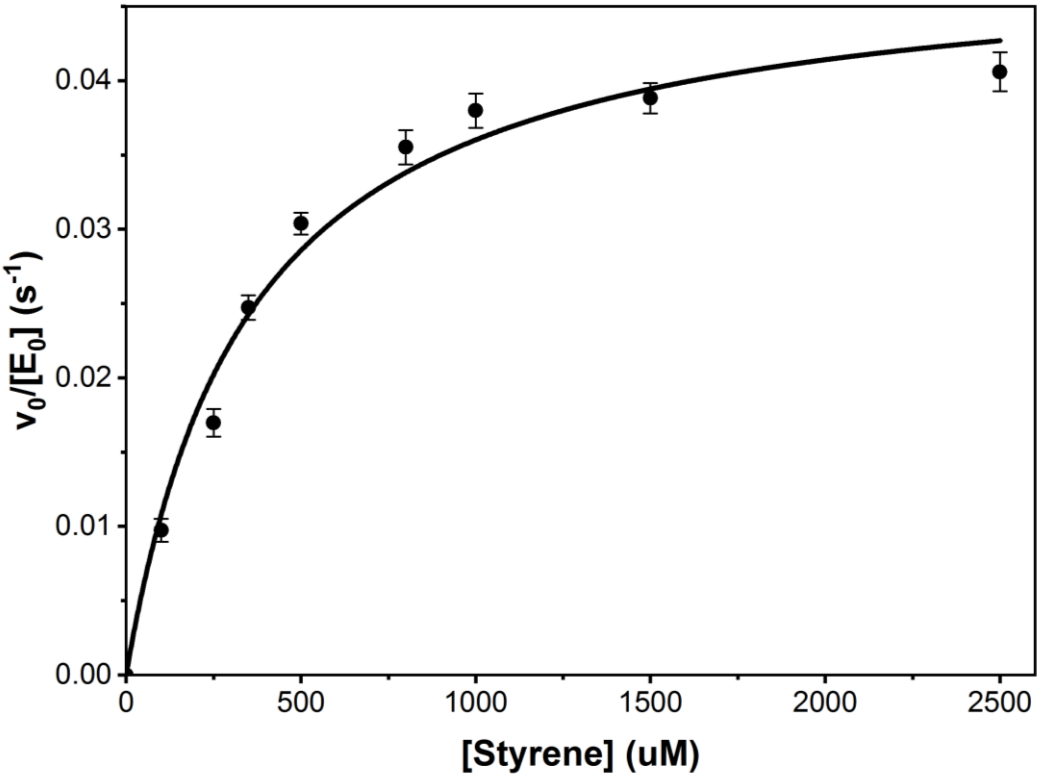


Figure 152: The Michaelis-Menten plots for the cyclopropanation reaction of the JR1:EDA metallocarbenoid complex in the presence of styrene (7.5 µM, CHES buffer, pH 8.6, 40% (v/v) EtOH). Initial rates were determined from the rate of change in absorbance for the JR1:EDA adduct at 5°C in the presence of styrene.

The kinetic parameters k_{lim}, K_1, and k_{lim}/K_1 for the JR1-catalysed cyclopropanation reaction between EDA and styrene are 0.048 s^{-1}, 352,1 µM and 136.32 s^{-1} M^{-1} respectively, relative to 0.001 s^{-1}, 482.51 µM and 2.23 s^{-1} M^{-1} for C45 (**table 23**, *appendix 2*). From the acquired data, it can be calculated that JR1 exhibits at 60-fold increase in cyclopropanation efficiency relative to C45 for the reaction between EDA and styrene. The K_1(styrene) for JR1 shows a 37% decrease relative to C45, indicating the diffusion

and accommodation of styrene in the active site is more favourable with JR1 compared to C45, possibly another reflection of the increased flexibility of JR1 as indicated by the biophysical data. In addition, the flexibility of JR1 relative to C45 could possibly weaken any stabilising interactions afforded to the metallocarbenoid intermediate for the protein backbone. This would decrease the thermodynamic stability of the metallocarbenoid and increase its reactivity, a hypothesis supported by the significantly higher k_{lim} for JR1 (0.048 s^{-1}) vs C45 (0.001 s^{-1}). The kinetic analysis of JR1 indicates the formation of the metallocarbenoid intermediate is enhanced relative to C45, but the rate of decomposition of the metallocarbenoid intermediate when presented with a suitable substrate is also enhanced; both observations can be accounted for partly by an increase in the backbone flexibility.

7.2.2 The second library carbene transferase screening – AP3.2

C45 was initially designed as a simple, heme-containing oxidoreductase and revealed functionality towards peroxidase enzymology. In the Anderson group, C45 was employed as the primary scaffold for robot-assisted directed evolution studies in efforts to generate new maquettes with enhanced peroxidase activity. A novel library of mutants was generated by Jack Steventon, using error-prone PCR, and were classified as APx, (*Artificial peroxidase x,* where x reflects the generation the mutant was identified in). The best mutants from the first generation were taken forward into the second generation of direct evolution. The process resulted in three generational libraries, AP1, AP2 and AP3. It was decided that a random selection of mutants, across all three generations, would be taken from the peroxidase assays and screened for carbene transferase activity. The designated reaction was the cyclopropanation between EDA and styrene so that any modifications in stereoselectivity could be identified alongside any relative changes in enzymatic activity. A total of 17 mutants were expressed and screened for cyclopropanation activity using the standard cyclopropanation protocol (chapter 9, *section 9.7.1*), with the only modification being cell lysate containing the overexpressed mutant was used as the reaction medium (**table 15**). The reaction mixtures were allowed to react overnight, quenched with 3M HCl, and extracted with EtOAc; each reaction mixture was analysed *via* chiral-HPLC to identify and quantify any cyclopropanation activity. The results of the initial screen are tabulated in **table 15**.

Table 15: Average product yields, diastereoselectivities, and enantioselectivities for the cyclopropanation reaction between EDA and styrene catalysed by a library of C45 mutants generated *via* direct evolution. The reactions were catalysed using lysate containing the mutant.

Mutant	% Yield$_{(R,R)}$	% Yield$_{(S,S)}$	dr$_{(E)}$ %	ee$_{(S,S)}$ %
F18K	11.26	21.86	>99.9	14.78
F18R	9.74	29.22	>99.9	32.02
F46S	9.63	35.93	>99.9	50
L50R	*9.31*	*40.26*	*>99.9*	*57.72*
AF11	0	10.39	>99.9	62.45
AG4	0	0	>99.9	-
AG5	0	0	>99.9	-
DP1	0	0	>99.9	-
BG7	0	9.85	>99.9	>99.9
CF6	0	11.15	>99.9	>99.9
DH6	0	0	>99.9	-
DP2	0	11.04	>99.9	>99.9
AP1	12.55	24.68	>99.9	32.56
AP2	27.71	8.44	>99.9	-53.30
AP3.2	*0*	*37.66*	*>99.9*	*>99.9*
DF5	0	26.30	>99.9	>99.9
FG1	*0*	*32.90*	*>99.9*	*>99.9*

From **table 15**, it is readily seen that seven mutants, F18R, F46S, L50R, AP2, AP3.2, DF5, and FG1, catalysed the reaction with >25% product yield. However, more importantly is that all the screened mutants, with the exception of AP2, were stereoselective for the (*S,S*)-cyclopropane product, an inversion to the selectivity exhibited by C45 and JR1; AP3.2, DF5, and FG1 each exhibited an enantioselectivity of >99.9% for the (*S,S*)-product. From the six (*S,S*)-selective mutants, three mutants were selected for additional screening on account of exhibiting the highest stereoselectivities in combination with the highest product yields. AP3.2 and FG1 were selected as both mutants exhibited product yields and (*S,S*)-enantioselectivity of >30% and >99.9% respectively; L50R was chosen over DF5 on account of L50R exhibiting the highest product yield (>40%) from all the initial screens, despite a significantly lower (*S,S*)-enantioselectivity (58%). The sequences of the three mutants, L50R, AP3.2, and FG1, relative to C45 are presented in **figure 153**. FG1 and L50R, both first-generation mutants, are each one mutation away from WT-C45. In FG1, the mutation is a lysine to asparagine (K9N) at position 9, located in helix A. In L50R, the mutation is a leucine to arginine (L50R) at position 50, located in helix B. AP3.2, however, is a third-generation variant and deviates from WT-C45 by eight mutations. The mutations are F11Y in helix A, G39S in the A-B loop, D48Y and F52S in helix B, F83S in helix C, and G111A, F132S and E122G in helix D.

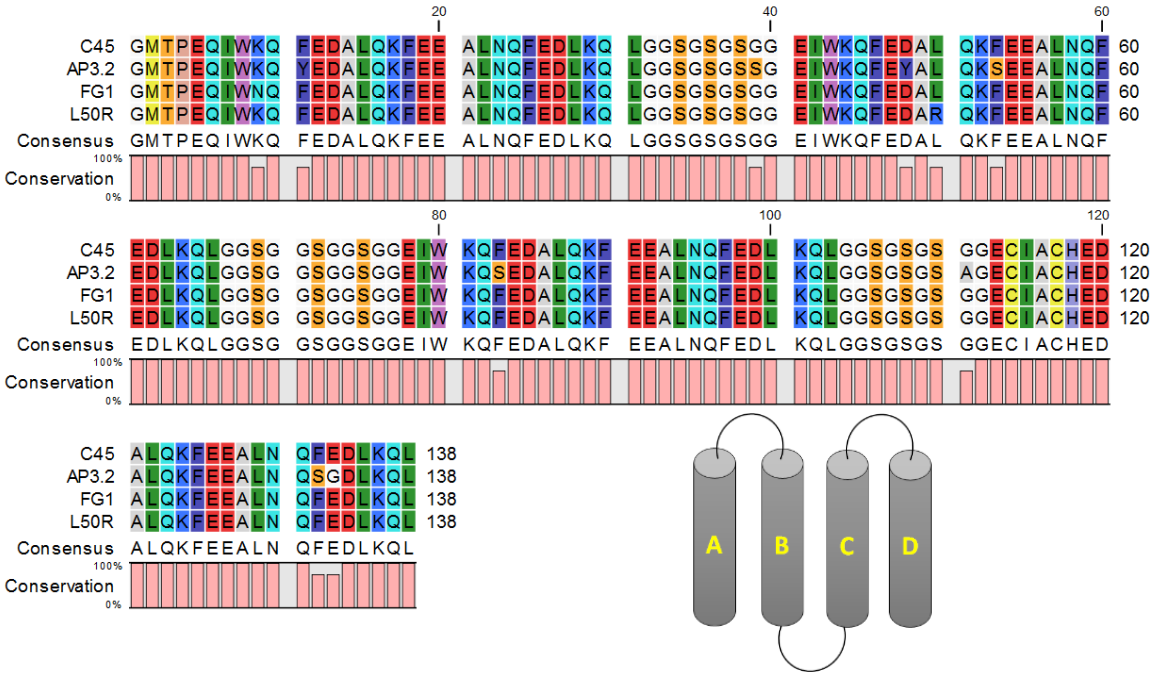


Figure 153: The primary sequences of the four mono-heme containing tetra-α-helical maquettes C45, AP3.2 FG1, and L50R and the labelling of the four helices A-D (the mono-histidine ligated heme cofactor is bound to helix D).

The three selected mutants were expressed and purified and the cyclopropanation assays were repeated with the purified enzyme, using the protocol outlined in chapter 9 *section 9.7.1*. The average product yields and enantioselectivities for the L50R-, AP3.2-, and FG1-catalysed cyclopropanation assays are tabulated in **table 16**. From the table, it can be readily seen that AP3.2 is the best variant, catalysing the reaction with on average 46.6% reaction yield but, more importantly, with a >99.9% enantioselectivity towards the (*S,S*)-cyclopropane product. The resultant normalised chromatogram for the AP3.2-catalysed cyclopropanation reaction and the normalised average chromatogram for the AP3.2-catalysed cyclopropanation reaction plotted against a commercial sample of enantiomerically-pure ethyl (*1S,2S*)-2-phenylcyclopropane-1-carboxylate is presented in **figure 154**. Although the reactivity of AP3.2 is notably lower relative to C45, the absolute inversion in stereochemistry is of significant interest, and therefore AP3.2 was subsequently subjected to biophysical characterisation, alongside the spectroscopic and kinetic characterisation of AP3.2 metallocarbenoid formation and cyclopropanation, in attempts to elucidate the functionality, activity and selectivity of AP3.2 compared to C45 and JR1.

Table 16: Average product yields, diastereoselectivities and enantioselectivities for the cyclopropanation reaction between EDA and styrene catalysed by the C45 mutants FG1, L50R, and AP3.2. The reactions were repeated in triplicate and were catalysed by purified enzyme.

	Average			
Mutant	**% Yield$_{(R,R)}$**	**% Yield$_{(S,S)}$**	**dr$_{(E)}$ %**	**ee$_{(S,S)}$ %**
FG1	0	22.80 (± 1.81)	>99.9	89.67 (± 14.6)
L50R	0	41.45 (± 8.34)	>99.9	95.77 (± 6.0)
AP3.2	0	46.60 (± 4.6)	>99.9	100

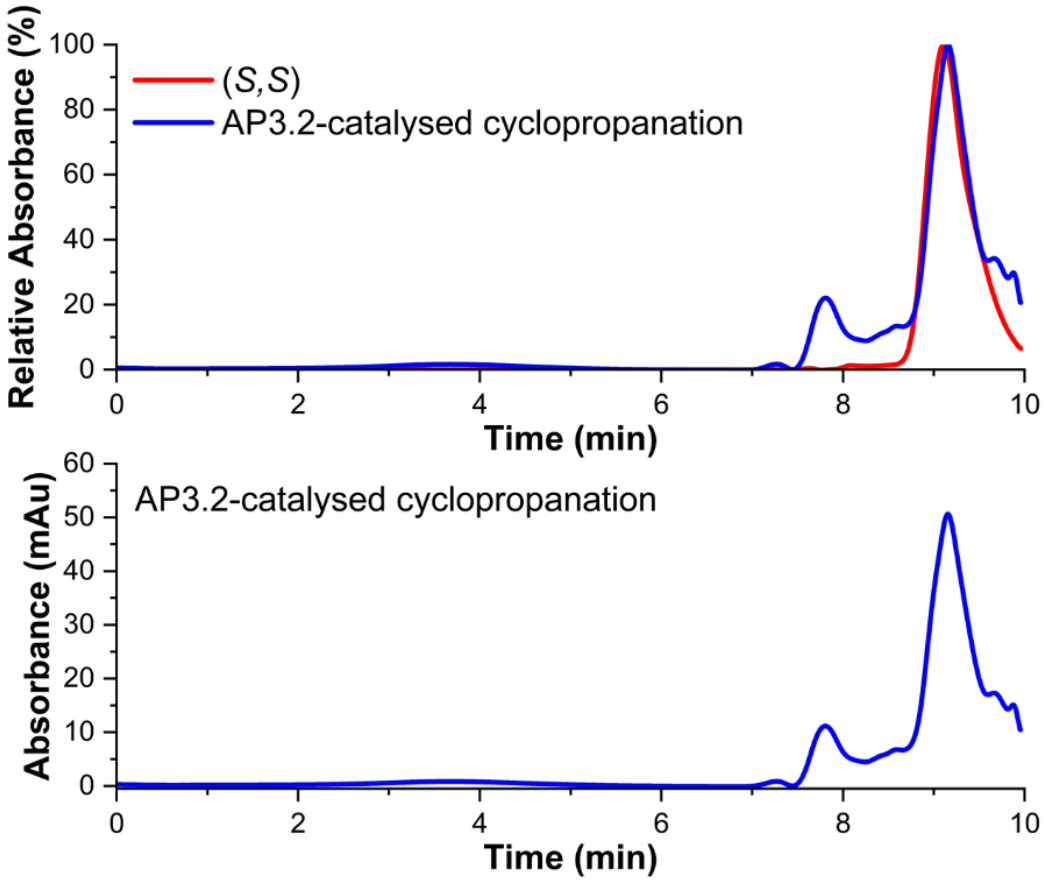


Figure 154: Chiral-HPLC chromatograms for the AP3.2-catalysed (10 µM, 0.1 % catalyst loading) cyclopropanation assays between styrene (30 mM) and EDA (10 mM) (CHES buffer, pH 8.6, 280 nm). A polar organic mobile phase (100% MeCN: 0.1% (v/v) TFA:0.1% (v/v) Et_3N) was employed and injection volumes were 2 µl. (**top**) normalised chiral-HPLC chromatograms for commercial (*S,S*)-ethyl 2-phenylcyclopropane-1-carboxylate (280 nm, EtOH, red) vs AP3.2-catalysed cyclopropanation (280 nm, blue); (**bottom**) Averaged chromatogram from the AP3.2-catalyzed cyclopropanation assay.

7.2.2.1 Biophysical characterisation of AP3.2

AP3.2 (**figure 155**) was characterised by UV-VIS spectroscopy, redox potentiometry and circular dichroism to determine its UV-VIS absorption profile, standard reduction potential and thermal stability. The information determined was compared against C45. In addition, an MD simulation of AP3.2 was acquired (**figure 155**).

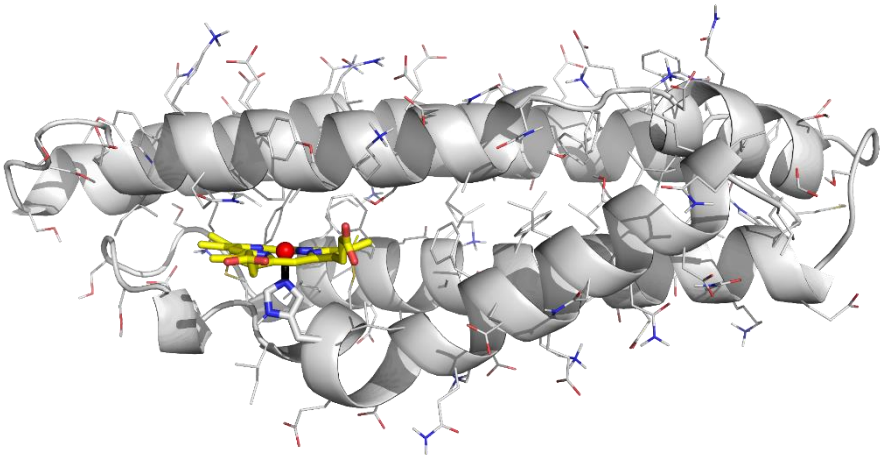


Figure 155: The structure of AP3.2 acquired after MD (ABMER) simulations.

The UV-VIS spectra for ferric and ferrous AP3.2 are presented in **figure 156**. The UV-VIS spectrum for ferric AP3.2 is dominated by the characteristic Soret peak at 408 nm and is accompanied by a Q-band series of less intense peaks in the range 500-550 nm. Upon the addition of sodium dithionite, a new spectroscopic profile emerged and is attributed to ferrous AP3.2. The reduced spectrum is dominated by a red-shifted intense Soret peak at 419 nm which is accompanied by a Q-band with two less-intense but well-resolved peaks, the α and β peaks, with the intensity ratio β > α. The UV-VIS spectra for ferric and ferrous AP3.2 is consistent with that of c-type mono-histidine ligated tetra-α-helical maquettes, with a spectroscopic profile that parallels the spectra observed for C45, JR1 and *Rma*-TDE.

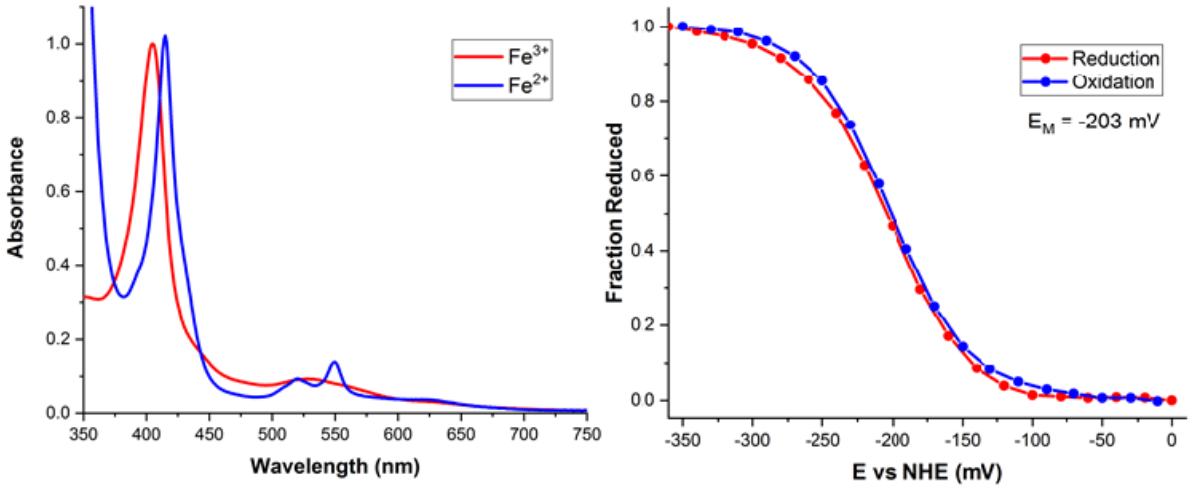


Figure 156: (**left**) The UV-VIS spectrum for ferric (red) and ferrous (blue) JR1, (**right**) Redox potentiometry of AP3.2 (100 mM KCl, 50 mM CHES, 10% glycerol, pH 8.6) collected with redox mediators as described in the methods section.

Redox potentiometry for AP3.2 gives a calculated mid-point potential of -203 mV (vs SEP) relative to C45 (-176 mV vs SEP),[235] indicating a slight stabilization of ferric AP3.2 possibly due to increased solvent accessibility at the iron. AP3.2, however, has a melting temperature (T_M) of 49^{O}C, a substantial 31^{O}C and 37^{O}C decrease in the melting temperature relative to JR1 (T_M = 80^{O}C) and C45 (86^{O}C) respectively,[235] as determined by circular dichroism (**figure 157**), and indicates the maquette scaffold of AP3.2 is significantly destabilised relative to C45/JR1. The significant decrease in T_M for AP3.2 can be largely accounted for by the substitution of three hydrophobic phenylanilines to hydrophilic serines at the helical positions 53, 83 and 132. In C45 and JR1 the residues Phe-53, Phe-83, and Phe-132 all point towards the core of the maquette, providing π-groups which can participate in stabilising inter-residual interactions in addition to enhancing the hydrophobicity of the core; the mutations to serine remove these stabilising interactions and, consequently, facilitate the thermal denaturation of AP3.2 relative to C45/JR1. The introduction of the D48Y mutation in AP3.2 does introduce another aromatic residue, but from the MD simulation of AP3.2 it can be seen the tyrosine is pointing away from the protein, and is orientated above the entrance of the active site parallel to the solvent-exposed propionate groups of the heme cofactor.

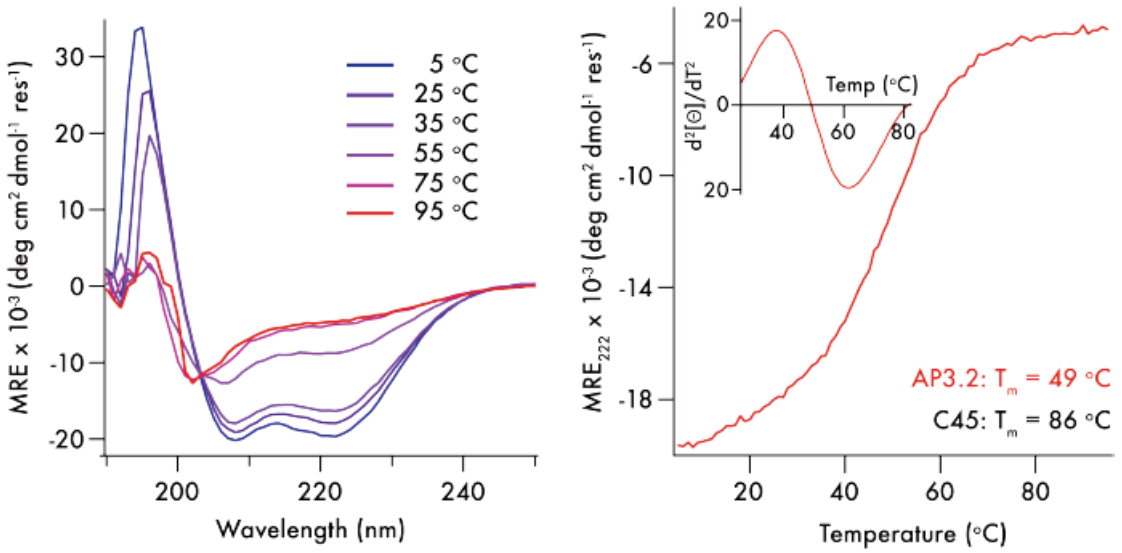


Figure 157: (**left**) Far-UV circular dichroism spectra of AP3.2 with varying temperature (100 mM KCl, 20 mM CHES, pH 8.6) (**right**) Temperature dependence of the CD signal monitored at 222 nm during thermal denaturation of AP3.2. The inset shows a smoothed second derivative of the thermal melt trace indicating a melting transition (T_m) of 49°C.

7.2.2.2 Spectroscopic and kinetic characterisation of AP3.2 metallocarbenoid formation

The formation of the metallocarbenoid intermediate in AP3.2 was investigated under identical conditions to those employed for C45 (5°C, 40% EtOH, CHES buffer, pH 8.6). In the presence of EDA, the expected shifts in the UV-VIS spectra corresponding to the formation of a terminal metallocarbenoid complex where observed (**figure 158**): the emergence of a red-shifted Soret peak at 430-436 nm accompanied by a red-shifted and broadened Q-band profile. The time-dependent spectra for the change in absorbance at 417 and 433 nm, corresponding to ferrous AP3.2 and the AP3.2-EDA metallocarbenoid intermediate respectively, exhibited a profile analogous to the spectroscopic changes observed for C45 - a rapid decrease in the ferrous peak (417 nm) with a concurrent increase in the metallocarbenoid peak (433 nm) punctuated by a plateauing of both absorbances upon saturation. The time-course spectra indicate that once formed the metallocarbenoid intermediate is stable and persists, exhibiting no appreciable decay over time, analogous to C45. The spectra also indicate that AP3.2 achieves saturation on a time scale similar to C45 and JR1 (10 s).

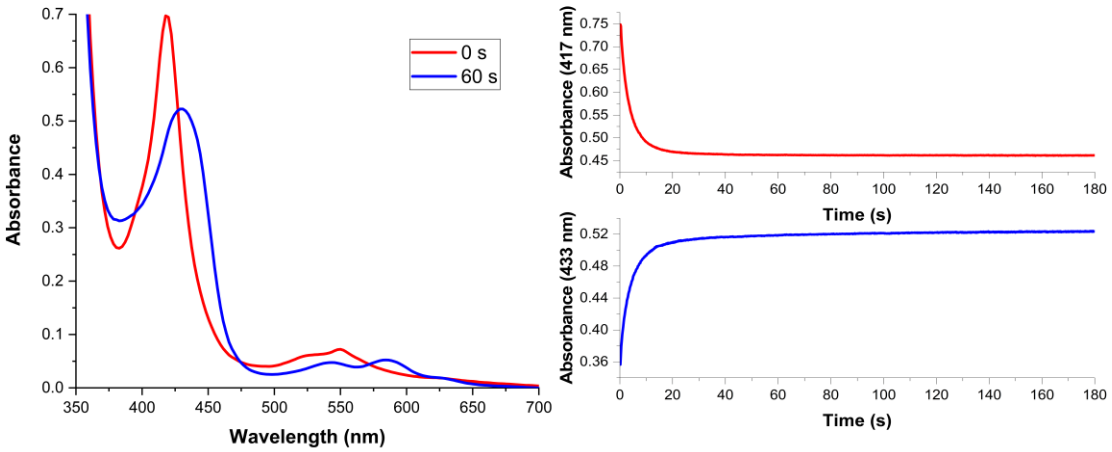


Figure 158: (**left**) Single wavelength traces representing the time course absorbance of ferrous AP3.2 (417 nm, 7.5 µM protein, 10% EtOH) and the metallocarbenoid:AP3.2complex (433 nm) following rapid mixing of ferrous AP3.2 with 500 µM ethyl diazoacetate at 5^{O}C in the presence of 3 mM styrene. (**right**) UV-VIS traces recorded at various time intervals following the rapid mixing of ferrous AP3.2 with 500 µM ethyl diazoacetate at 5^{O}C in the presence of 3 mM styrene.

In order to properly compare metallocarbenoid formation in AP3.2 to JR1 and C45, the single turnover kinetic parameters, K_1 and k_{obs} (and the ratio (k_{obs}/K_1)) for the formation of the AP3.2-metallocarbenoid intermediate were determined. The kinetics of metallocarbenoid formation in AP3.2 were determined using stopped-flow spectroscopy by adopting the protocol used for C45, JR1, *Rma*-TDE, and Mb(H64V,V68A). **Figure 159** exhibits the calculated curve acquired for EDA. As for C45, the analysis of the formation of a metallocarbenoid involves three distinct identities: i) free enzyme and substrate, ii) a tetrahedral transition state, and iii) the metallocarbenoid intermediate. There are three underlining assumptions that have been made to assist in the analysis: i) the rate of decomposition of the tetrahedral transition state to the metallocarbenoid intermediate (k_2) is first-order, ii) the reaction is irreversible ($k_2 >>> k_{-2}$) and iii) a steady-state equilibrium is established between the free enzyme and the tetrahedral transition state. The values for the single turnover kinetic parameters, K_1, k_{obs}, and (k_{obs}/K_1) for AP3.2-metallocarbenoid formation are tabulated, against the same values acquired for C45, in **table 22** (*appendix 10*).

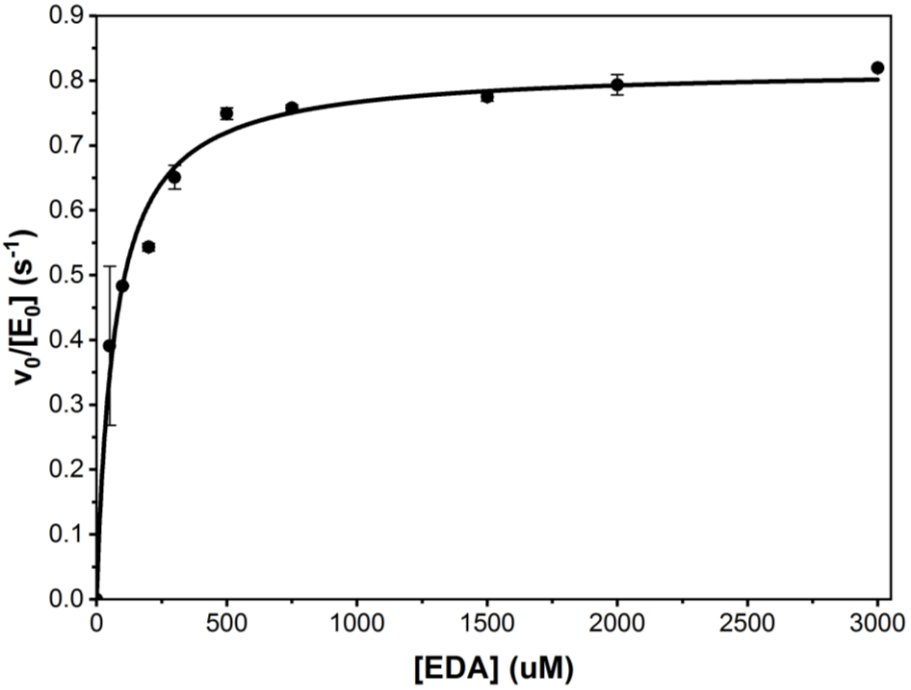


Figure 159: The Michaelis-Menten plots of AP3.2 (7.5 µM, CHES buffer, pH 8.6, 40% (v/v) EtOH) with varying concentration of EDA. Initial rates were determined from the rate of change in absorbance for the AP3.2:metallocarbenoid adduct at 5°C.

The k_{obs}, K_1, and k_{obs}/K_1 values for AP3.2 are 0.82 s^{-1}, 69.26 µM, and 1.18 x 10^4 s^{-1} M^{-1} respectively, relative to 0.498 s^{-1}, 74.36 µM and 6.70 x 10^3 s^{-1} M^{-1} for C45 (**table 22** (*appendix 2*)). The K_1 for AP3.2 is essentially identical to the value for C45, indicating that AP3.2 possesses a binding affinity similar to C45. More importantly, the k_{obs} for the formation of the AP3.2-EDA metallocarbenoid intermediate is the highest value reported across all the enzymes screened (C45, JR1, *Rma*-TDE and Mb(H64V,V68A)). In order to understand the values for K_1 and k_{obs} for AP3.2 it is necessary to understand how the mutations to AP3.2 effect the resultant structure of the maquette scaffold relative to C45. **Figure 160** shows an overlap diagram of the MD generated structures of C45 and AP3.2. The F52S mutation is the only substitution that occurs in the proximity of the heme active site. In conjunction with the decreased secondary structure of the maquette scaffold, the switch from a hydrophobic phenylaniline to a hydrophilic serine would create a cavity in the active site that would facilitate the accommodation of a substrate. In addition, the hydroxy group of serine could possibly participate in a hydrogen bond with the carbonyl group of EDA. The combination of an increase in backbone flexibility and active site volume, and a possible stabilising interaction with serine, could possibly account for the increased k_{obs}.

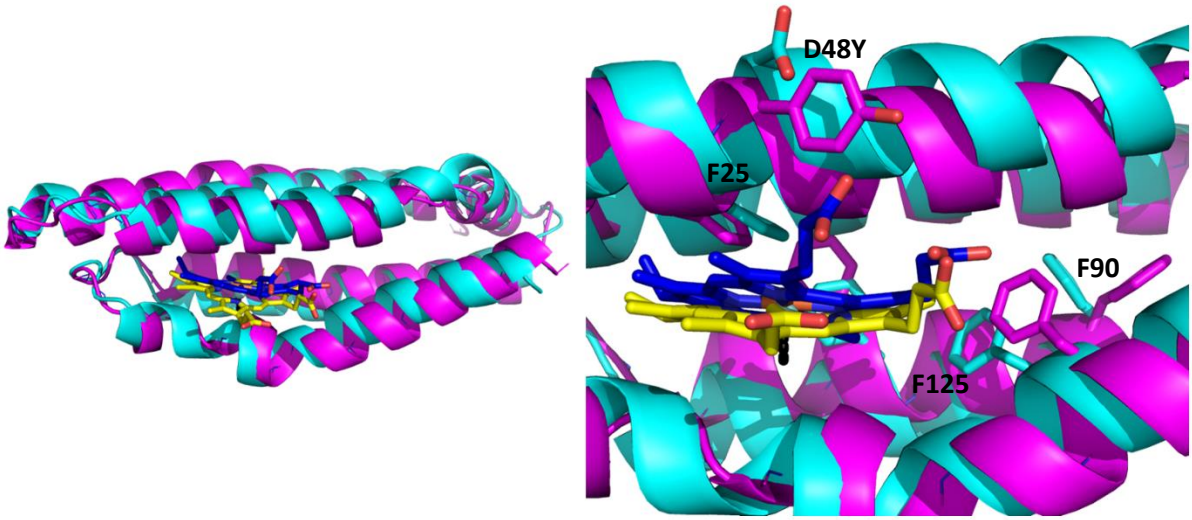


Figure 160: Overlapping MD simulations of the two maquettes C45 (cyan; heme shown in yellow) and AP3.2 (magenta; heme shown in blue), (**left**) overlay of C45 and AP3.2, and (**right**) close up view of the overlap of active sites of C45 and AP3.2 (the D48Y mutation is shown).

7.2.2.3 Kinetic characterisation AP3.2-catalysed cyclopropanation

The experimental conditions employed to investigate carbene transfer activity paralleled the conditions initially used to investigate metallocarbenoid formation, with a single modification being the introduction of a suitable substrate, styrene. A solution containing ferrous AP3.2 (15 µM protein, 10 mM $Na_2S_2O_4$, CHES buffer, pH 8.6) and a solution containing ethyl diazoacetate (EDA) and styrene (100 µM -10 mM and 200 µM - 3mM respectively, 20:80% CHES:EtOH) were prepared in separate compartments (syringes) before a small volume (2x20 µl) of each were injected concurrently into a mixing chamber. The mixing chamber was monitored by UV-VIS spectroscopy over several timescales to monitor any reactions occurring over the duration of the experiment. Final concentrations were 7.5 µM enzyme, stoichiometric $Na_2S_2O_4$, 50 µM - 5 mM EDA, 100 µM - 1.5 mM styrene, and 40% EtOH.

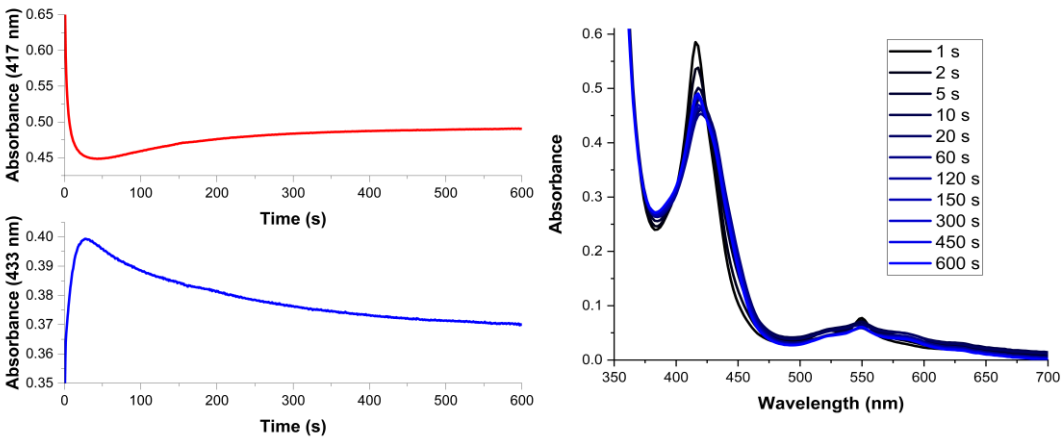


Figure 161: (**left**) Single wavelength traces representing the time course absorbance of ferrous AP3.2 (417 nm, 7.5 µM protein, 10% EtOH) and the metallocarbenoid:AP3.2 complex (433 nm) following rapid mixing of ferrous AP3.2 with 500 µM ethyl diazoacetate at 5^{O}C in the presence of 3 mM styrene. (**right**) UV-VIS traces recorded at various time intervals following the rapid mixing of ferrous AP3.2 with 500 µM ethyl diazoacetate at 5^{O}C in the presence of 3 mM styrene.

Figure 161 displays the time-course spectra obtained from the stopped-flow experiments for AP3.2 in the presence of EDA and styrene at wavelengths 417 and 433 nm. 417 nm corresponds to the ferrous AP3.2 species and 433 nm corresponds to the AP3.2-EDA metallocarbenoid complex, respectively. The data shows that in the initial 20 seconds post-mixing there is a gradual decrease in the 417 nm absorbance which is accompanied by a concurrent increase in the 433 nm absorbance – corresponding to the formation of the AP3.2-EDA metallocarbenoid complex. It was demonstrated above that once formed the AP3.2-EDA metallocarbenoid complex is stable and persists over a duration of 600 seconds without any appreciable decay. In the presence of styrene, the initial spectra acquired for AP3.2 resembles the spectra acquired in the absence of styrene, corresponding to the formation of the metallocarbenoid intermediate. After approximately 30 seconds, however, the absorbance at 433 nm starts linearly decreasing and is accompanied by a concurrent linear increase in absorbance at 417 nm – indicating the decay of the metallocarbenoid species and the return to the starting ferrous complex. UV-VIS traces captured at numerous timepoints over 600 seconds are also presented in **figure 161**. In the initial stages of the reaction (1-60 seconds) the UV-VIS exhibits the red-shift corresponding to the formation of the metallocarbenoid intermediate, alongside the concurrent shift and broadening of the Q-band profile. In the UV-VIS spectra acquired in the absence of styrene the red-shift from ferrous to metallocarbenoid intermediate is characterised by sharp Soret peaks as the 417 nm peak disappears alongside the emergence of a Soret peak at 428-433 nm. In the presence of styrene, the emergence of the 433 nm peak appears as a shoulder alongside the 417 nm peak, resulting in a broad, split Soret peak, without either the 417 nm or 433 nm peak dominating the spectra. After 20 seconds, the

maximum absorbance begins to shift towards the 417 nm peak and is accompanied by the gradual re-emergence of the ferrous Q-band profile. The spectroscopic profile for AP3.2-metallocarbenoid formation in the presence of styrene parallels the profile exhibited by C45 and JR1. The analogous spectroscopic profiles mean the kinetic parameters k_{lim}, K_1, and k_{lim}/K_1 for the AP3.2-catalysed cyclopropanation reaction can be calculated under conditions identical to those employed for C45 and JR1, allowing for comparisons between the maquettes to be investigated (**figure 162**). The kinetic parameters k_{lim}, K_1, and k_{lim}/K_1 for the AP3.2-catalysed cyclopropanation reaction between EDA and styrene are 0.067 s^{-1}, 141.23 µM, and 474.4 s^{-1} M^{-1} respectively, relative to 0.001 s^{-1}, 482.51 µM, and 2.23 s^{-1} M^{-1} for C45 (**table 23**, *appendix 2*). From the acquired data, it can be inferred that AP3.2 possess is a more efficient cyclopropanation carbene transferase relative to C45 for the reaction between EDA and styrene. The kinetic parameters for the AP3.2-catalysed cyclopropanation reaction eclipse all the other enzymes screened thus far, including (in order of decreasing k_{lim}/K_1) JR1, Mb(H64V,V68A), C45, and *Rma*-TDE respectively (**table 23**, *appendix 2*).

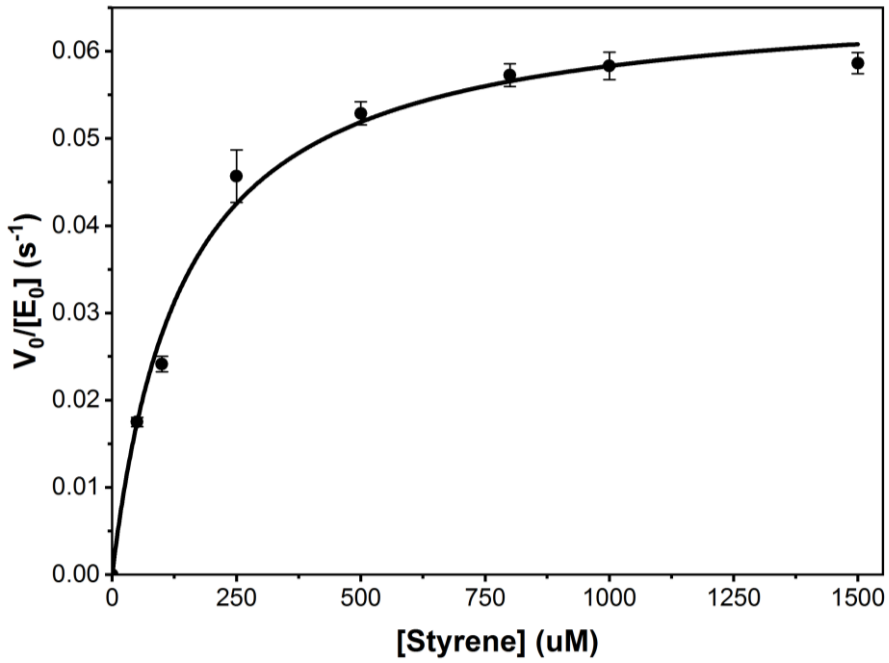


Figure 162: The Michaelis-Menten plots for the cyclopropanation reaction of the AP3.2:EDA metallocarbenoid complex in the presence of styrene (7.5 µM, CHES buffer, pH 8.6, 40% (v/v) EtOH). Initial rates were determined from the rate of change in absorbance for the AP3.2:EDA adduct at 5°C in the presence of styrene.

From the computationally generated C45 and AP3.2 models no immediate explanation can be deduced to account for i) the pronounced change in the observed stereoselectivity, ii) the significantly enhanced cyclopropanation kinetics relative to all other enzymes screened thus far, or iii) the pronounced drop in reaction yield. From the biophysical data, it was shown the three phenylalanine

to serine mutations at key helical positions result in a more destabilised protein scaffold for AP3.2, which was reflected in the drastic decrease in the melting temperature of AP3.2 relative to C45 and JR1 (39^{O}C vs 86^{O}C and 81^{O}C respectively).[235] The destabilisation of the maquette core would engender a more molten globular like structure for AP3.2 which would exist, in solution, in an ensemble of ill-defined tertiary conformations. The F53S is the only mutation that is positioned close to the carbene binding site at the distal face of the heme, and the replacement of a bulky hydrophobic phenylalanine residue for a small hydrophilic serine residue would create a cavity capable of accommodating bulkier substrates, transitions states or intermediates; furthermore, the hydroxy group of serine can function as a hydrogen-bond acceptor/donator, possibly providing some compensatory stabilisation to the substrates and/or enzymatic transitions states/intermediates; the combination of the four phenylalanine to serine mutations in AP3.2 could account for the enhanced kinetic parameters for both the metallocarbenoid formation and its subsequent reaction with styrene. The D48Y mutation in AP3.2 results in a tyrosine positioned above the entrance to the active site, orientated parallel to the solvent-exposed propionate groups of the heme cofactor. The tyrosine residue could possibly provide stabilising π-π interactions to the incoming styrene substrate, functioning to provide a favourable trajectory into the active site, perhaps directing styrene into the cavity generated by the F53S mutation. The directional guiding of styrene by tyrosine, through favourable π-π interactions, into an accommodating space would engender an energetically preferable trajectory of approach towards the metallocarbenoid intermediate and would provide a reasonable account for the inverted and remarkably high stereoselectivity exhibited by AP3.2.

7.3 CONCLUSION

The simplicity and malleability of the four-helix bundle renders maquettes suitable scaffolds for mutagenesis and directed evolution. Error-prone PCR afforded a C45-variant, JR1, which exhibited slightly improved activity towards the cyclopropanation of styrene and EDA. JR1 is only two mutations away from WT-C45 but has a K_1 for the binding of EDA three times lower compared to C45. In addition, despite minor improvements in product yield, JR1 exhibited a k_{lim} for the cyclopropanation of styrene that was almost 50 times higher than C45 (0.048 vs 0.0011 s^{-1} respectively), and twice that of the rationally designed carbene transferase Mb(H64V,V68A) (0.048 vs 0.021 s^{-1} respectively). The biophysical properties of JR1 were largely unaffected by the mutations, exhibiting a UV-VIS profile and possessing values for T_m and E_M that are similar to C45. Another variant, AP3.2 (8 mutations away from C45), was produced using directed evolution. The UV-VIS profile of AP3.2 was consistent with a mono-heme-containing four-helix bundle, but the biophysical data revealed a significant decrease in T_m (39^{O}C vs 86^{O}C respectively) and change in E_M (-203 vs -176 mV vs C45). AP3.2 was catalytically active,

catalysing the cyclopropanation reaction between EDA and styrene, although less catalytically active than its counterparts C45 and JR1 (47% for AP3.2 vs 80/85% for C45/JR1 respectively). Interestingly, AP3.2 exhibited a complete inversion of stereoselectivity towards the reaction, with only the (*S,S*)-cyclopropane product being isolated, in direct contrast to the (*R,R*)-selective maquettes C45 and JR1. Kinetically, AP3.2 outperformed all over variants, reporting the highest k_{obs} for the formation of the metallocarbenoid intermediate (0.820 s^{-1}). AP3.2 also exhibited remarkable kinetic parameters for the cyclopropanation reaction, with a K_1 for styrene of 141.23 µM (relative to 428.51 µM and 657.7 µM for C45 and Mb(H64V,V68A) respectively) and a k_{lim} three times higher than that for Mb(H64V,V68A) (0.067 s^{-1} vs 0.021 s^{-1} respectively). The experiments discussed in this chapter highlight not only the powerful technique of direct evolution, but also how introducing key mutations can have pronounced effects on the kinetics and stereoselectivity of a reaction. At present, it is difficult to elucidate the relationships between the mutations and the observed changes, but nonetheless the maquette scaffold offers a facile platform for engineering novel biocatalysts with tailored functionality, biophysical properties, and reaction selectivity.

CHAPTER 8: CONCLUSION

In chapter 3, the metallocarbenoid formation propensity of the *de novo* tetra-α-helical c-type heme-containing maquette C45 was explored and compared against two artificial carbene transferases reported in the literature: a cytochrome *c* variant engineered in the Arnold lab, *Rma*-TDE, and a sperm whale myoglobin mutant engineered in the Fasan lab, Mb(H64V,V68A). The electronic spectra of C45, *Rma*-TDE, and Mb(H64V,V68A) all exhibited similar profiles. Further spectroscopic characterisation revealed that C45, *Rma*-TDE, and Mb(H64V,V68A) can form stable metallocarbenoid intermediates under the conditions investigated. A metallocarbenoid intermediate of *Rma*-TDE and Mb(H64V,V68A) has been successfully crystallised and reported in the literature, so the similarities between the electronic spectra provides good evidence that C45 is forming stable, terminally bound, heme-localised metallocarbenoid intermediates. This was further supported by experiments conducted using ESI-MS which revealed the new intermediates each had a shifted mass corresponding exactly to the respective carbene precursor (diazo) compound. The kinetics of metallocarbenoid formation revealed that C45 is intermediate between *Rma*-TDE and Mb(H64V,V68A), which supports the notion that *de novo* proteins can compete with rationally redesigned enzymes. In chapter 7, the metallocarbenoid formation of two C45 variants JR1 (2 mutations from WT-C45) and AP3.2 (8 mutations from WT-C45) exhibited superior kinetics relative to C45 (AP3.2 reported a catalytic efficiency superior to Mb(H64V,V68A)) which demonstrates the versatility and evolvability of the C45 tetra-α-helical maquette.

Chapter 4 and 5 explored the catalytic versatility of the C45 maquette towards carbene transfer chemistry. The cyclopropanation of olefins was initially explored in chapter 4, and it was demonstrated that C45 is a proficient, stereoselective carbene transferase, catalysing the reaction between styrene and ethyl diazoacetate with a product yield, diastereoselectivity, and enantioselectivity of 80%, 99% and 77% for the (*R,R*)-cyclopropane product respectively. Investigating the substrate promiscuity of C45 revealed its cyclopropanation activity extended to functionalised diazo esters and *para*-functionalised styrenes, with enantioselectivities surpassing 90% for the cyclopropantion reaction between EDA and *p*-methoxystyrene, *p*-fluorostyrene, and *p*-trifluoromethylstyrene. The kinetics of C45-catalysed cyclopropanation was explored spectroscopically. The catalytically efficiency of C45 was similar to *Rma*-TDE, but inferior to Mb(H64V,V68A). In chapter 7, the cyclopropanation kinetics of JR1 and AP3.2 were also investigated spectroscopically, and it was demonstrated the C45-variants possessed catalytic efficiencies approximately 4-times (136 $s^{-1}\,M^{-1}$) and 12-times (474 $s^{-1}\,M^{-1}$) higher than Mb(H64V,V68A) (32 $s^{-1}\,M^{-1}$) respectively. It was also shown that mutations to the maquette scaffold can afford enantiodivergence, as the maquette AP3.2 catalysed the cyclopropanation reaction

between EDA and styrene with an enantioselectivity surprising 99% in favour of the (*S,S*)-cyclopropane product, directly contrasting the (*R,R*)-selectivity exhibited by C45 and JR1.

In chapter 5, the enzymatic promiscuity of C45 was investigated, and it was demonstrated that C45's carbene transferase activity extends to chemoselective N-H insertions, Si-H insertions, carbonyl olefinations, and *in vivo* and *in vitro* homologous ring expansions. The *in vivo* and *in vitro* homologous ring expansion of pyrrole, catalysed by C45, is the first reported enzymatic ring expansion – natural, designed or *de novo*- – that proceeds *via* a carbene transfer mechanism. With the exception of the carbonyl olefinations, C45 was demonstrated to catalyse the aforementioned reactions with remarkable yields and selectivity; in particular, the C45-catalysed reaction between *tert*-butyl diazoacetate and 3-aminophenol afforded product yields surpassing 99%. Overall, chapters 4 and 5 demonstrated the synthetic utility of small tetra-α-helical maquettes as biocatalysts in a range of important abiological reactions. The synthetic utility of C45 was further expanded in chapter 6, in which it was demonstrated that encapsulating C45 inside an organic-solvent resistant calcium alginate hydrogel afforded a heterogeneous biocatalyst. The spectroscopic profile of alginate-encapsulated C45 (AE-C45) was unperturbed by the encapsulation process, even when the AE-C45 beads were suspended in pure organic solvents such as MeOH and n-hexane. It was demonstrated that AE-C45 also retains its metallocarbenoid formation ability, as the electronic spectra of the AE-C45-EDA complex is identical to the electronic spectra of the C45-EDA complex in aqueous solution. AE-C45 not only retained its ability to form a metallocarbenoid intermediate, but it also retained its catalytic proficiency as AE-C45 was demonstrated to be a proficient catalyst in the i) cyclopropanation reaction between EDA and styrene, and ii) N-H insertion reaction between piperidine and EDA, when the AE-C45 beads were suspended in a range of organic solvents. The product yields of the AE-C45 carbene transfer assays were demonstrated to be dependent on the diffusion of the reacting substrate across the solvent-alginate bead interface. In addition, AE-C45 could be recovered after one round of catalysis and could be successfully re-employed as a carbene transferase in subsequent rounds of catalysis.

The importance of this work is it demonstrates that large and/or complex protein scaffolds are not a prerequisite for engineering novel biocatalytic enzymes towards synthetically and technologically useful chemical reactions. It was demonstrated that C45, a simple, evolutionary naïve tetra-α-helical protein scaffold can catalyse carbene transfer reactions, with a variety of substrates, with yields and selectivities that rival, and even surpass, the carbene transferases, engineered from naturally-occurring heme-containing protein scaffolds reported in the literature.[184,187,288,393] Furthermore, it was demonstrated in chapter 7 that the maquette is also a perfect blueprint scaffold for directed evolution.

The introduction of small changes often proves sufficient for altering a maquette's functionality, which was elegantly highlighted by the aforementioned variant AP3.2. AP3.2 demonstrated a complete inversion of stereochemistry towards the cyclopropanation reaction, has the highest k_{obs} for the formation of the metallocarbenoid intermediate (0.820 s^{-1}), and exhibited remarkable parameters towards the cyclopropanation reaction, with a K_1 for styrene of 141.23 µM (relative to 428.51 µM and 657.7 µM for C45 and Mb(H64V,V68A) respectively) and a k_{lim} three times higher than that for Mb(H64V,V68A) (0.067 s^{-1} vs 0.021 s^{-1} respectively). Overall, a simple, highly engineerable, *de novo* c-type heme-containing protein proficient at i) forming a number of stable, bioorganometallic intermediates that are spectroscopically consistent with the metallocarbenoid intermediates reported for rationally designed carbene transferases, ii) catalysing a wide range of abiological chemical reactions, *via* a metallocarbenoid intermediate, with remarkable chemoselectivity and stereoselectivity, and iii) performing biocatalytic conversions under conditions typically inhospitable to solvents (i.e. neat organic solvents), was reported. These observations collectively reinforce the postulation that re-engineering naturally occurring proteins is unnecessary to achieve abiological function, as the same – if not better – results can be accomplished with much simpler *de novo* proteins.

CHAPTER 9: MATERIALS AND METHODS

All chemicals were purchased from either Sigma or Fisher Scientific, and T7 Express competent cells and Q5 polymerase were purchased from NEB. Cloned synthetic genes for *Rma*-TDE and Mb(H64V/V68A) were purchased from Eurofins Genomics. Error-prone PCR (epPCR) libraries of C45were produced using the GeneMorph II EZClone Domain Mutagenesis Kit. All UV-visible spectra were recorded on an Agilent Cary-60 UV-visible spectrophotometer. OTTLE potentiometry was used to measure heme redox potentials. A JASCO J-815 CD polarimeter was used to measure circular dichroism spectra. All stopped-flow data an SX20 Stopped Flow was acquired using an SX20 Stopped Flow spectrophotometer (Applied Photophysics) housed in an anaerobic glove box under N2 ($[O_2]$ <5 p.p.m.; Belle Technology). All LC-MS mass spectra were acquired using positive electron-spray-ionization (ESI) mass spectrometry (Waters Xevo G2-XS QTof) attached to a C8 reverse column (Grace Vydac, 100 x 21 mm, 5 µm. All High Performance Liquid Chromatography (HPLC) chromatograms were acquired using a Agilent waters HPLC and two separate columns were employed for quantification: a chiral column using either a C18 HPLC reverse phase column (Phenomenex, 150 x 15 mm, 5 µm) or a chiral-HPLC column (Astec CHIROBIOTIC® V, 250 x 21 mm, 5 µm).

9.1 GENERAL & MOLECULAR BIOLOGY, PROTEIN EXPRESSION AND PURIFICATION

All chemicals were purchased from either Sigma or Fisher Scientific and cloned synthetic genes for *Rma*-TDE and Mb(H64V/V68A) were purchased from Eurofins Genomics. All transformations were performed using E. coli T7 Express competent cells (NEB) and by following a standard protocol.[234] For C45, JR1, and AP3.2 a modified pMal-p4x+ plasmid (the maltose-binding protein (MBP) sequence is deleted, leaving the MBP Sec translocon recognition sequence (MKIKTGARILALSALTTMDFSASALAK) that enables the periplasmic translocation of the artificial c-type cytochrome) containing the sequence for C45/JR1/AP3.2 with a C-terminal His_6 tag and a Tobacco Etch Virus (TEV) recognition sequence (ENLYFQ/G) between the His_6 tag and the protein sequence (to allow the removal of the purification tag) and Pec86+ (containing the *ccm* machinery required for incorporation of the c-type heme into C45) expression vectors were used. The vectors pSHT (for periplasmic expression) or pET45b(+) (for cytoplasmic expression) were used for the enzymes *Rma*-TDE (co-expressed with Pec86+) and Mb(H64V,V68A). 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). AP3.2 and APR1 were randomly selected variants from the epPCR C45 library. Each

variant was expressed and purified as described below. 50% glycerol stocks for each cell-line were prepared and frozen for all proteins.

9.2 EXPRESSION PROTOCOL

All of the proteins were expressed and purified *via* the following protocol. All steps were performed under sterile conditions. Overnight starter cultures were prepared by adding 100 µl carbenicillin (50 mg/ml), and 100 µl chloramphenicol (50 mg/ml, Pec86+ containing vectors only), to 100 ml of LB before inoculating the solution with the appropriate glycerol stock. Starter cultures were then incubated overnight at 37^{O}C and 180 rpm. To 1L of LB was added 1ml of the appropriate antibiotics (50 mg/ml) before 50ml of an overnight starter culture was added. Litre cultures were then left to grow until an OD_{600} value between 0.6-0.8 was obtained (usually after 3 hours), at which point 1 ml IPTG (1M stock, 1 mM induction concentration) was added per litre to induce protein expression; the induced solutions were left to express for an addition three hours. After three hours the cultures were collected, centrifuged (4000 rpm, 4^{o}C, 30 mins) and the resultant cell pellets were isolated from the supernatant, resuspended in 10ml lysis buffer (300 mM NaCl, 50 mM sodium phosphate, 20 mM imidazole, pH-8) and stored at -20^{O}C until purification.

9.3 PURIFICATION PROTOCOL

The lysis suspensions were defrosted at room temperature prior to purification. Once thawed, the suspensions were sonicated multiple times in increments of 20 seconds (stored on ice for 20 seconds between increments) to lyse the cells. After sonication, the suspensions were transferred to 50 ml SS34 centrifuge tubes and centrifuged (SS34 centrifuge) at 4^{O}C and a rate of 18000 rpm for 30 minutes. After centrifugation, the supernatant was transferred to a 50 ml Eppendorf tube and the cell pellets discarded. The supernatant was filtered *via* 45 µm and 20 µm membrane filters and transferred onto a nickel column under lysis buffer (300 mM NaCl, 50 mM Sodium phosphate (dibasic), 20 mM imidazole, pH 8). The protein of interest contains a HIS_6-tag, which exhibits a high affinity for nickel; therefore, the HIS_6-tagged protein remained attached to the nickel column as the undesirable contents of the solution (proteins, cell debris etc) was filtered off. After running approximately 25 ml of lysis buffer through the column, the column lines were switched to elution buffer (300 mM NaCl, 50 mM Sodium phosphate (dibasic), 250 mM imidazole, pH 8) which passed through the column, displacing the HIS-tagged protein and eluting it off the column. *The HIS-tagged protein was collected, and subsequently subject to overnight dialysis to re-suspend the protein in TEV protease buffer (0.5 mM EDTA, 50 mM Tris, 5L Milli-Q, pH 8). Once in TEV buffer, excess DDT and TEV protease were added to the mixture (under a nitrogen atmosphere) and the mixture was left to stir for 3-6 hours. The

mixture was then filtered again (45 and 20um membrane filters) and transferred back onto a nickel column under lysis buffer. If the HIS_6-tag had been cleaved successfully, the protein of interest eluted off the nickel column under lysis buffer and can be collected immediately. The remaining HIS-tag was subsequently washed off the column and discarded. If the HIS-tag failed to cleave, the dialysis and TEV protease steps were repeated, but the mixture allowed to mix for longer durations (overnights). After cleaving the HIS-tag, the protein was concentrated down to 5ml using a 10,000-spin-membrane Eppendorf and centrifuging at 4500 rpm and $4^{O}C$ for as long as necessary. The protein was then purified using an AKTA PURE S75 gel filtration column employing CHES buffer (20 mM CHES, 100 mM KCl, pH 8.6) or KPi buffer (100 mM phosphate, pH 7.5, used for Mb(H64V,V68A) only) as the mobile phase. The progress of the protein along the column was monitored spectroscopically, at 280 nm, and the collected fractions containing the desired protein were characterised *via* UV-spectroscopy and an SDS-PAGE gel to examine purity. The desirable fractions were then combined, and the final solution was concentrated (if necessary) and divided into 500 µL aliquots which were stored at $-80^{O}C$ until required.

9.4 UV-VIS SPECTROSCOPY

All UV-VIS spectra were recorded on an Agilent Cary-60 UV-visible spectrophotometer using a 1 mm pathlength quartz cuvette. Reduced spectra for each sample were acquired by the addition of 10 µl of $Na_2S_2O_4$ (1 mM in de-ionised water). All UV-VIS spectra were recorded across a wavelength range of 300-750 nm. The concentrations were determined using the extinction coefficients for the ferric and ferrous samples, respectively. C45: ε_{406} = 147,433 M^{-1} cm^{-1} and ε_{417} = 119,900 M^{-1} cm^{-1}; JR1: ε_{406} = 191,000 M^{-1} cm^{-1} and ε_{417} = 172,000 M^{-1} cm^{-1}; AP3.2: ε_{406} = 62,000 M^{-1} cm^{-1} and ε_{417} = 68,000 M^{-1} cm^{-1}; *Rma*-TDE: ε_{406} = 138,200 M^{-1} cm^{-1} and ε_{417} = 128,000 M^{-1} cm^{-1}; Mb(H64V,V68A): ε_{409} = 156,700 M^{-1} cm^{-1} and ε_{435} = 96,000 M^{-1} cm^{-1}.

9.5 STOPPED-FLOW SPECTROPHOTOMETRY

Stopped-flow kinetics were conducted using an SX20 stopped-flow spectrophotometer (Applied Photophysics) housed in an anaerobic glovebox under N_2 ($[O_2] < 5$ ppm; Belle Technology). In the initial experiments, a solution containing a known concentration of a reduced enzyme (15 µM, 100 mM KCl, 20 mM CHES, pH 8.6, reduced with $Na_2S_2O_4$ for C45, AP3.2, JR1, and *Rma*-TDE; 0.1M phosphate, pH 7.5, reduced with $Na_2S_2O_4$ for Mb(H64V,V68A)) was placed in one syringe and a 5 mM solution of a carbene precursor (either EDA, ^{t}BuDA or BnDA) in ethanol (20-100%) 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, over the course of 180-1000 seconds,

to examine the metallocarbenoid formation and porphyrin degradation pathway, respectively. The formation of the metallocarbenoid was monitored by following the time-course profiles at 417 nm and 428-437 nm, the Soret peak for reduced C45/AP3.2/JR1/*Rma*-TDE and the metallocarbenoid intermediate respectively, and 435 nm and 424 nm, the Soret peak and the metallocarbenoid intermediate for reduced Mb(H64V,V68A). Final concentrations were 7.5 µM ferrous C4 enzyme and 2.5 mM EDA/^{t}BuDA/BnDA (10-50% ethanol). The kinetics of the formation of the metallocarbenoid for each enzyme were determined using the same conditions outlined above but using varying concentrations of EDA (50 µM – 3 mM). The rate constants were calculated at varying substrate concentrations by plotting to a single exponential and then fitting the data to a hyperbolic function (**equation 19**).

$$\text{rate} = k_{obs}[S]/K_1+[S]$$

Eq 19

Additional experiments using stopped-flow spectrophotometry were conducted to study the degradation of the metallocarbenoid intermediates in the presence of a suitable substrate. A solution containing a known concentration of ferrous enzyme (15 µM, 100 mM KCl, 20 mM CHES, pH 8.6, reduced with stoichiometric $Na_2S_2O_4$ for C45, AP3.2, JR1, and *Rma*-TDE; 0.1M phosphate, pH 7.5, reduced with stoichiometric $Na_2S_2O_4$ for Mb(H64V,V68A)) was placed in one syringe and an 80:20% ethanol:water solution containing 1 mM of a carbene precursor (either EDA, ^{t}BuDA or BnDA) and 6 mM of styrene 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, over the course of 180-1000 seconds to examine carbene transfer activity. The reaction progressed was monitored by following the time-course profiles at 428-437 nm and 417 nm for reduced C45/AP3.2/JR1/*Rma*-TDE, and 424 nm and 435 nm for reduced Mb(H64V,V68A). Final concentrations were 7.5 µM reduced enzyme, 500 µM EDA/^{t}BuDA/BnDA, 3 mM styrene (40% ethanol). The kinetics of the degradation of the metallocarbenoid species for each enzyme were determined using the same conditions outlined above but using varying concentrations of styrene (50 µM – 3 mM). The rate constants were calculated at varying substrate concentrations, plotted and then the data were fitted using equation (**20**).

$$k_{obs} = k_{lim}[S]/K_1+[S]. \text{ where } K_1 = (k_{-1} + k_2)/k_1$$

Eq 20

9.6 PROTEIN MASS SPECTROMETRY FOR C45

The formation of the C45-metallocarbenoid intermediate was further examined using positive electron-spray-ionisation (ESI) mass spectrometry (Waters Xevo G2-XS QTof). A 1 ml solution of 350 µM C45 in CHES buffer (100 mM KCl, 20 mM CHES, pH 8.6), in a 1.5 ml sealed-top vial containing a silicon septum, was flushed under scrubbed nitrogen inside an anaerobic glovebox ($[O_2]$ < 5 ppm; Belle Technology) before being sealed, removed from the glovebox and placed on ice for 10 minutes. The sample was then immediately loaded directly into the mass spectrometer, where an isocratic 100 mM ammonium acetate solution was employed as the mobile phase (0.25 ml min^{-1}). 20 µl injections were employed and the progression of the sample into the chamber was monitored spectroscopically at 280 nm. The mass spectrum contained peaks screened across a m/z range of 1500-2700. The dominant peaks were identified, and the charges of the fragments were calculated using the molecular mass of C45 (15168 Da).

A 1 ml solution of 350 µM C45 in CHES buffer (100 mM KCl, 20 mM CHES, pH 8.6), in 1.5 ml sealed-top vials, was then flushed under scrubbed nitrogen inside an anaerobic glovebox ($[O_2]$ < 5 ppm; Belle Technology) followed by the addition of 25 µl of $Na_2S_2O_4$ (400 mM stock in CHES buffer). A separate vial containing the selected diazo compound (EDA, ^{t}BuDA, BnDA, 400 mM stock in EtOH) was deoxygenated alongside the vial containing reduced C45. The vials were sealed, transported out of the glovebox, and cooled in an ice bath for 10 minutes. Once cooled, 50 µl of the selected diazo compound was added *via* a gastight syringe into the vial containing reduced C45 to initiate the reaction. The sample was kept on ice for 1 minute to allow for the formation of the metallocarbenoid intermediate before being directly loaded onto the mass spectrometer. Final reaction concentrations were 350 µM enzyme, 10 mM sodium dithionite, and 20 mM diazo compound. Identical conditions to the experiment with only C45, mentioned above, were employed for each carbene precursor studied

9.7 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) except for the assays conducted with Mb(H64V,V68A) which were performed in KPi buffer (100 mM potassium phosphate, pH 7). The final reaction volumes for all assays were 400 µL unless otherwise stated.

9.7.1 cyclopropanation assays

To 370 µL of a 10 µM C45/AP3.2/*Rma*-TDE/JR1/Mb(H64V,V68A) solution was added 10 µL of $Na_2S_2O_4$ (400 mM stock; de-ionized water) and the mixture was left to stir for 1 minute, ensuring complete reduction of the enzyme 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 in an ice bath. Once cooled, 10 µL of the diazo compound was added *via* a gastight syringe into the reaction vials to initiate the reaction. Final reaction concentrations were 10 µM enzyme (0.1% mol%), 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 (the progress of the ester hydrolysis was monitored using TLC (7:3 ethyl acetate:hexane)). 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) were calculated *via* external calibration with commercial ethyl 2-phenylcyclopropane-1-carboxylate, 2-phenylcyclopropane-1-carboxylic acid, and synthesised standards.

9.7.2 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 and the products 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 were calculated *via* an external calibration with commercial ethyl cinnamate and cinnamic acid.

9.7.3 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 and the products were extracted with 1.25 mL of n-hexane (1.25 mL of hexane as required for sufficient extraction of the product into the organic phase [monitored by TLC]); the products were subsequently analysed by chiral-HPLC and LC-MS as described below. The product yields and total turnover numbers were calculated *via* an external calibration with commercial *N*-phenylglycine ethyl ester and ethyl-1-piperidine acetate (and including a 1.25 multiplication factor to account for product dilutions) and synthesised standards.

9.7.4 Chemoselective N-H insertion assays

Chemoselective N-H insertion assays were conducted under the similar reaction conditions as the N-H insertion assays using *ortho*-, *meta*-, and *para*-aminophenol; the reactions were stirred for2 hours and quenched with 20 µl 10% (w/v) acetic acid (2-aminophenol and 3-aminophenol) or 3M HCl (4-aminophenol). After quenching the products were extracted with 1 mL of ethyl acetate and the products were subsequently analysed by C18-HPLC and LC-MS as described below. The product yield and total turnover numbers were calculated *via* an external calibration with commercial *N*-phenylglycine ethyl ester.

9.7.5 Phenylenediamine N-H insertion assays

Phenylenediamine N-H insertion assays were conducted similarly to the chemoselective N-H insertion assays, with two differences: i) substitution of *ortho*-, *meta*-, *para*-aminophenol for *ortho*-, *meta*- and *para*-phenylenediamine, and ii) final substrate concentrations were 10 mM phenylenediamine and 30 mM EDA.

9.7.6 Si-H insertion assays

Si-H insertion assays were conducted similarly to the N-H insertions assays, with a few notable differences: i) dimethylphenylsilane (DMPS) was used as the substrate in place of an amine, ii) the reactions were quenched by the addition of 1 ml n-hexane, and iii) final substrate concentrations were 20 mM EDA and 20 mM DMPS. The product mixtures were analysed using C18-HPLC and LC-MS

9.7.7 Ring expansion assays

Ring expansion assays were conducted under similar reaction conditions as the cyclopropanation assays described above, but with the following alterations: styrene was substituted for pyrrole (400 mM stock in ethanol); ethyl 2-bromo-2-diazoacetate was used as the carbene precursor (40 mM stock in CH_2Cl_2). The products were extracted with 1 mL of ethyl acetate and subsequently analysed by C18 reverse-phase HPLC and LC-MS as described below. Final reaction concentrations were 10 µM enzyme

(1% mol%), 10 mM sodium dithionite, 1 mM diazo compound, and 10 mM pyrrole. The product yields and total turnover numbers were calculated *via* an external calibration with commercial ethyl nicotinate and niacin.

9.7.7.1 Synthesis of ethyl 2-bromo-2-diazoacetate

Ethyl 2-bromo-2-diazoacetate was synthesised as previously reported. Briefly, *N*-bromosuccinimide (26 mmol) was added to a solution of EDA (20 mmol) and 1,8-Diazabicyclo(5.4.0)undec-7-ene (28 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 Na2S2O3 (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 (~40 mM)

9.8 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 980 µl 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 additional four hours (37 ^{o}C and 180 rpm). After four hours, 980 µl 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 (1 M stock in EtOH) was added and the vials sealed and removed from the glovebox. The vials were cooled on ice before 25 µl of ethyl 2-bromo-2-diazoacetate (40 mM stock in DCM, nitrogen flushed) was added *via* a gastight needle; the final concentration of the reagents was 1 mM ethyl 2-bromo-2-diazoacetate and 10 mM pyrrole. The products were extracted with 1 ml ethyl acetate and analysis was conducted as reported above in the "Ring expansion assays'' section.

9.9 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 (Sigma Aldrich) 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.

9.10 PRODUCT CHARACTERIZATION BY REVERSE PHASE AND CHIRAL HPLC

All the reactions performed were quantified by High Performance Liquid Chromatography (HPLC). Two separate columns were employed for quantification: a chiral column and a reverse phase C18 column. A chiral-HPLC column (Astec CHIROBIOTIC® V, 250 x 21 mm, 5 µm) was used to analytically quantify the cyclopropanation and N-H insertion assays, employing an isocratic mobile phase (100% CH_3CN: 0.1% v/v TFA: 0.1% v/v: Et_3N; 0.1 mL min^{-1} flow rate and 2 µL injection volume for the cyclopropanation assays; 0.2 mL.min^{-1} flow rate and 5 µL injection volume for the N-H insertion assays). Injection volumes of the N-H insertion assay products were 5 µL to account for the 2.5:1 dilution in the extraction step. All elution traces were monitored spectroscopically at 245, 254 and 280 nm. The chiral-HPLC column allowed the retention times of the starting materials and the reaction products to be determined, and to quantify, where necessary, the enantioselectivity of a given reaction. A C18 HPLC reverse phase column (Phenomenex, 150 x 15 mm, 5 µm) was used to analytically quantify the ring expansion and carbonyl olefination assays, employing a gradient as the mobile phase for ring expansion assays (70:30% H_2O:CH_3CN to 10:90% H_2O:CH_3CN; 2 mL min^{-1} flow rate and 20 µL injection volume) and an isocratic mobile phase for the carbonyl olefination assays (100% CH_3CN: 0.1% v/v TFA: 0.1% v/v: Et_3N; 2 mL min^{-1} flow rate and 20 µL injection volume). Injection volumes of the sample mixture were 20 µL to account for the 2.5:1 dilution in the extraction step. All elution traces were monitored spectroscopically at 245, 254, 265 and 280 nm. The C18 column allowed the retention times of the starting materials and the reaction products to be determined, and, where necessary, to quantify the diastereoselectivity of a given reaction. A C18 HPLC reverse phase column (Phenomenex, 150 x 15 mm, 5 µm) was used to analytically quantify the chemoselective N-H insertion and Si-H insertions assays, employing an isocratic mobile phase (100% CH_3OH: 0.1% w/v NH_4AcO, 0.5 mL min^{-1} flow rate and 20 µL injection volume for chemoselective N-H insertion assays; 100% CH_3OH: 0.1% w/v NH_4AcO; 2 mL min^{-1} flow rate and 20 µL injection volume for the Si-H insertion assays). All elution traces were 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 allowed the ($N_{insertion}$/$O_{insertion}$) ratio to be quantified. If appreciable product formation could be detected after the

initial HPLC experiments, the products were subsequently analysed using Liquid chromatography-Mass spectrometry (LC-MS) to assist in product identification (described below).

9.10.1 External calibrations

External calibrations of the anticipated reaction products were conducted using HPLC with commercially available standards or with synthesised standards (see section 9.10.2 for details). The conditions and mobile phases employed in the product analysis for each assay were employed for the calibrations. Each substrate was prepared at multiple concentrations (i.e. 50-500 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 as a function of concentration. A plot of [substrate] vs peak height engendered a straight line which could be used to determine the product yields and TTNs for each assay. Injection volumes for the external calibrations were 2 µL and 8 µL for the chiral-HPLC and C18 columns, respectively. The following equations were subsequently used to calculate total turnover numbers (TTN), enantiomeric excesses (*ee* %), and cis/trans ratios:

$$TTN = \frac{[product]}{[enzyme]}$$

Eq 21

$$ee\ \% = \left(\frac{[R,R] - [S,S]}{[R,R] + [S,S]}\right) \times 100$$

Eq 22

$$cis{:}\,trans = \frac{[Z]}{[E]}$$

Eq 23

9.10.2 Synthesis of external standards

External standards for every reaction (excluding carbonyl olefinations) were synthesised using the following methods.

9.10.2.1 Cyclopropanation synthesis

A 100 ml oven-dried round-bottom flask was purged with nitrogen and 7 ml of DCM, $Rh_2(OAc)_4$ (5% mol) and the appropriate reagent (olefin or amine, 5 equivalents) was added. A 3 ml solution containing the diazo compound (DCM, 1 equivalent) was added dropwise to the solution while stirring

over a 30-minute period. Final concentrations were 250 mM olefin, 50 mM diazo compound and 5% mol catalyst; the final volume was 10 ml. The reaction was left stirring for 1 hour before the solvent was removed under vacuum. The crude mixture was purified by flash column chromatography (9:1 n-hexane:diethyl ether) to yield the pure product. The product identify was confirmed *via* ESI-MS and chiral-HPLC; ***ethyl 2-(4-hydroxyphenyl)cyclopropane-1-carboxylate*:** synthesised following the outlined protocol, colourless oil, % yield (56), LC-MS m/z (% relative intensity): 133.06 (42.48), 207.1 (100); ***ethyl 2-(4-methoxyphenyl)cyclopropane-1-carboxylate*:** synthesised following the outlined protocol, white needle-like crystals, % yield (90), LC-MS m/z (% relative intensity): 147.07 (74.03), 160.05 (63.11), 175.07 (100); ***ethyl 2-(4-fluorophenyl)cyclopropane-1-carboxylate*:** synthesised following the outlined protocol, colourless oil, % yield (67), LC-MS m/z (% relative intensity): 135.05 (40.75), 145.04 (45.70), 163.05 (100), 209.09 (1.14); ***ethyl 2-(4-chlorophenyl)cyclopropane-1-carboxylate***: synthesised following the outlined protocol, colourless oil, % yield (69), LC-MS m/z (% relative intensity): 144.05 (100), 179.02 (51.04), 225.06 (1.47); ***ethyl 2-(4-(trifluoromethyl)phenyl)cyclopropane-1-carboxylate*:** synthesised following the outlined protocol, colourless oil, % yield (61), LC-MS m/z (% relative intensity): 165.01 (100), 183.03 (91.54), 211.06 (57.11), 239.09 (62.33); ***ethyl 2-(4-(tert-butyl)phenyl)cyclopropane-1*-carboxylate:** synthesised following the outlined protocol, colourless oil, % yield (79), LC-MS m/z (% relative intensity): 127.05 (86.37), 145.06 (100), 168.08 (72.42), 191.10 (5.15), 247.16 (9.87); ***ethyl 2-(4-cyanophenyl)cyclopropane-1-carboxylate***: synthesised following the outlined protocol, colourless oil, % yield (47), LC-MS m/z (% relative intensity): 131.05 (42.81), 188.07 (100), 216.1 (10.98); ***tert-butyl 2-phenylcyclopropane-1-carboxylate***: synthesised following the outlined protocol, colourless oil, % yield (62.00), LC-MS m/z (% relative intensity): 127.05 (100), 145.06 (87.13), 168.07 (52.34), 247.16 (16.22); ***benzyl 2-phenylcyclopropane-1-carboxylate***: synthesised following the outlined protocol, colourless oil, % yield (56.00), LC-MS m/z (% relative intensity): 91.05 (100), 165.05 (44.34), 181.09 (75.77).

9.10.2.2 N-H insertion synthesis

A 100 ml oven-dried round-bottom flask was purged with nitrogen and 7 ml of DCM, $Rh_2(OAc)_4$ (5% mol) and the appropriate reagent (amine, 5 equivalents) was added. A 3 ml solution containing the diazo compound (DCM, 1 equivalent) was added dropwise to the solution while stirring over a 30-minute period. Final concentrations were 250 mM olefin, 50 mM diazo compound and 5% mol catalyst; the final volume was 10 ml. The reaction was left stirring for 4 hours before the solvent was removed under vacuum. The crude mixture was purified by flash column chromatography (8:2 pet ether:ethyl acetate) to yield the pure product. The product identify was confirmed *via* ESI-MS and

chiral-HPLC; ***ethyl (4-chlorophenyl)glycinate***: synthesised following the outlined protocol, white crystals, % yield (45), LC-MS m/z (% relative intensity): 140.11 (100), 142.09 (39.56), 214.06 (3.41); ***tert-butyl (4-chlorophenyl)glycinate*****:** synthesised following the outlined protocol, yellow oil, % yield (26), LC-MS m/z (% relative intensity): 140.04 (100), 142.04 (31), 226.05 (10.48), 228.05 (3.45); ***benzyl (4-chlorophenyl)glycinate*****:** synthesised following the outlined protocol, yellow oil, % yield (38), LC-MS m/z (% relative intensity): 91.06 (100), 276.01 (14.9), 278.13 (4.3); ***tert-butyl 2-(piperidin-1-yl)acetate*****:** synthesised following the outlined protocol, yellow oil, % yield (47), LC-MS m/z (% relative intensity): 98.10 (100), 144.10 (57.93); ***benzyl 2-(piperidin-1-yl)acetate***: synthesised following the outlined protocol, yellow oil, % yield (48), LC-MS m/z (% relative intensity): 91.05 (100), 142.08 (17.05), 234.14 (69.87).

9.11 PRODUCT CHARACTERIZATION BY LIQUID CHROMATOGRAPHY-MASS SPECTROMETRY

Liquid chromatography-Mass spectrometry (LC-MS, C8) was employed to identify the products formed in each assay. A C8 reverse column (Grace Vydac, 100 x 21 mm, 5 µm) was used with a 10-minute gradient mobile phase (95:5% H_2O:CH_3CN to 10:90% H_2O:CH_3CN; 0.1% v/v formic acid, 0.25 mL min^{-1}). 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-ionization (ESI) mass spectrometer (Waters Xevo G2-XS QTof). The mass spectrum contained peaks screened across a m/z range of 70-350 (depending on the molecular mass of the product being screened). Retention times and fragmentation patterns were initially determined using commercial samples of the anticipated products.

9.12 PREPARATION OF AN ALGINATE-ENCAPSULATED ENZYME

9.12.1 For UV-VIS spectroscopy

A solution of 3% (w/v) sodium alginate was prepared by mixing 3 g of sodium alginate with 100 ml of CHES buffer (20 mM CHES, 100 mM KCl, pH 8.6). The mixture was left stirring for 1 hour to ensure complete dissolution and homogeneity. Once ready, to 9 ml of the 3% (w/v) sodium alginate solutions were added 1 ml of C45/*Rma*-TDE (100 µM) or cytochrome *c* (1 mM). The mixture was adequately mixed to ensure homogeneity. A thin film of 0.2 M $CaCl_{2(aq)}$ was prepared in a petri dish and the alginate mixture was slowly pipetted over the surface to generate a thin sheet. The sheet was allowed to cure for 30 minutes before the residual $CaCl_{2(aq)}$ was discarded. The sheets were left on the bench to dry for an additional 30 minutes. A scalpel was used to remove a thin rectangle (5 cm x 1 cm x 1

cm) which was transferred into a cuvette for UV-VIS analysis. The UV-VIS trace was then compared to the free enzyme UV-VIS trace.

9.12.2 For non-aqueous enzymology

A solution of 3% (w/v) sodium alginate was prepared by mixing 3 g of sodium alginate with 100 ml of CHES buffer (20 mM CHES, 100 mM KCl, pH 8.6). The mixture was left stirring for 1 hour to ensure complete dissolution and homogeneity. Once ready, to 9 ml of the 3% (w/v) sodium alginate solutions were added 1 ml of C45/*Rma*-TDE (100 µM) or HRP (1 mM). The mixture was adequately mixed to ensure homogeneity, and then pipetted dropwise, from an elevation of approximately 30 cm, into a conical flask equipped with a stirrer bar and containing an excess of 0.2 M $CaCl_{2(aq)}$ solution. The beads formed spontaneously and were allowed to cure for 30 minutes before being collected by vacuum filtration. The beads were allowed to dry for an additional 30 minutes before being collected and stored at $4^{O}C$.

9.13 UV-VIS SPECTROSCOPY OF AN ALGINATE-ENCAPSULATED ENZYME

All UV-VIS spectra were recorded on an Agilent Cary-60 UV-visible spectrophotometer. To a 1 mm pathlength UV-VIS quartz cuvette was added a thin-sheet of alginate-encapsulated C45/*Rma*-TDE/cytochrome *c*. 1 ml of the appropriate solvent was added to the cuvette (CHES buffer, MeOH, EtOH, acetone, ethyl acetate or hexane) and a UV-VIS spectrum was recorded across a wavelength range of 300-750 nm. An initial baseline was acquired using an alginate sheet without any enzyme encapsulated within it. For the reduced spectrum, 10 µl of a 100 mM sodium alginate solution (in de-ionized water) was pipetted into the cuvette containing the solvent/alginate-encapsulated enzyme sheet, and mixing was assisted with the pipette. A UV-VIS spectrum was then recorded every 30 seconds across a wavelength range of 300-750 nm.

9.14 UV-VIS SPECTROSCOPIC IDENTIFICATION OF THE FORMATION OF A METALLOCARBENOID INTERMEDIATE IN AN ALGINATE-ENCAPSULATED ENZYME

All UV-VIS spectra were recorded on an Agilent Cary-60 UV-visible spectrophotometer. To a 1 mm pathlength UV-VIS quartz cuvette was added a thin-sheet of alginate-encapsulated C45. 1 ml CHES buffer (20 mM CHES, 100 mM KCl, pH – 8.6) was added to the cuvette and a UV-VIS spectrum was recorded across a wavelength range of 300-750 nm. An initial baseline was acquired using an alginate sheet without any enzyme encapsulated within it. The cuvette was subsequently sealed with a silicon septum and flushed with N_2 for a nitrogen cylinder. The reduced spectrum was acquired by adding 10 µl of a 100 mM sodium alginate solution (in de-ionized water) into the cuvette containing the

solvent/alginate-encapsulated C45 sheet *via* an airtight syringe. A UV-VIS spectrum was then recorded every 30 seconds across a wavelength range of 300-750 nm. After sufficient reduction could be detected, 10 µl of a stock solution of ethyl diazoacetate (EDA, 400 mM in EtOH) was added into the sealed cuvette *via* an airtight syringe. The solution was left to mix for 1 minute before a UV-VIS spectrum was recorded, across a wavelength range of 300-750 nm, every 30 seconds for 10 minutes.

9.15 CHEMOPHYSICAL CHARACTERISATION OF CALCIUM ALGINATE BEADS

9.15.1 Volumetric analysis

An assortment of varying alginate beads were prepared by pipetting a 1ml of a solution of sodium alginate (3% (w/v) sodium alginate, 20 mM CHES, 100 mM KCl, pH 8.6), dropwise and from a suspension of approximately 30 cm, into varying concentrations of $CaCl_{2(aq)}$ solutions (0.1 – 0.5M). All the beads that prepared were matured for 1 hour in the respective $CaCl_2$ solution and were prepared without any enzyme encapsulated. Once matured, the beads were collected *via* vacuum filtration and subsequently weighted and measured. The beads were placed in a labelled tray and left in an oven set at 37 ^{O}C for 3 hours. After 3 hours, the beads were collected, reweighed and remeasured. The differences in the values pre- and post-oven were used to calculate the water content of each bead (µl and %).

9.15.2 Retention analysis

All UV-VIS spectra were recorded on an Agilent Cary-60 UV-visible spectrophotometer. 10 mg of a select enzyme (horseradish peroxidase, cytochrome *c*, and C45) was dissolved in 10 ml of a sodium alginate solution (3% (w/v) sodium alginate, 20 mM CHES, 100 mM KCl, pH 8.6) and left to stir on a roller to achieve homogeneity. 1 ml of the resultant mixture was added, dropwise and from a suspension of approximately 30 cm, to varying concentrations of $CaCl_{2(aq)}$ solutions (0.1 - 0.5M), resulting in a total of 15 different samples (final enzyme concentration in each sample was 1 mg ml^{-1}). The beads were allowed to mature for 1 hour, before being collected *via* vacuum filtration and resuspended in 50 ml of CHES buffer (20 mM CHES, 100 KCl, pH 8.6). The samples were left mixing on a roller overnight and were analysed the following morning. 1 ml from each respective sample was collected and analysed *via* UV-VIS spectroscopy. The concentration of each protein leaked into the aqueous phase has called from the Soret peak for each respective protein (C45, $\varepsilon_{406\ nm}$ = 147,344 M^{-1} cm^{-1}; HRP, $\varepsilon_{403\ nm}$ = 103,840 M^{-1} cm^{-1}; cytochrome *c*, $\varepsilon_{409\ nm}$ = 98,160 M^{-1} cm^{-1}).

9.15.3 Determination of alginate porosity

The porosity (ε) of the alginate beads was determined following the procedure outlined by Ha.[387] 200 alginate beads (14 µl bead^{-1}) were suspended in a 50 ml solution of 100 mM indigo carmine (C_1) and allowed to stir for 20 minutes. After two hours the beads were collected *via* vacuum filtration and the concentration of indigo carmine remaining in the 50 ml solution was determined *via* UV-VIS spectroscopy ($\varepsilon_{667\ nm}$ = 34145 M^{-1} cm^{-1}). The collected beads were then resuspended in another solution of indigo carmine (17 mM; C_2) and left mixing for 20 minutes, before the beads were collected *via* vacuum filtration and the concentration of indigo carmine in the second solution was determined using UV-VIS spectroscopy (C). The porosity of the alginate beads was calculated using the following equation:[387]

$$\varepsilon = \frac{3V_b(C - C_2)}{4\pi r_p^3 N(C_1 - C)} = \frac{V_b(C - C_2)}{V_p(C_1 - C)}$$

Eq 24

- V_b = the total volume of the bulk solvent (ml)
- V_p = the total volume of the alginate beads (ml)
- C = the bulk concentration of indigo carmine at the end of the experiment (mM)
- C_2 = the concentration of indigo carmine in the second solution (mM)
- C_1 = the concentration of indigo carmine in the first solution (mM)
- r = the radius a single alginate bead (mm)
- N = the total number of alginate beads

9.16 ALGINATE-ENCAPSULATED ABTS PEROXIDATION KINETICS AND RECOVERABILITY

C45 and HRP alginate beads were prepared according to the protocol outlined above. The following protocol was used to measure the peroxidase kinetics of each encapsulated enzyme. To a quartz cuvette was added 980 µl of CHES buffer (20 mM CHES, 100 mM KCl, pH 8.6) and a single alginate bead (r = 1.5 mm). 10 µl of ABTS (100 mM stock in CHES buffer) was added to the cuvette and the mixture was allowed to stir for 1 minute. 10 µl of hydrogen peroxide (250 - 1 mM stocks in CHES) was pipetted into the cuvette to initiate the reaction. The progress of the reaction was monitored using a UV-VIS spectroscopy set to a single wavelength (607 nm). The initial rates were calculated from the slope. The final volume was 1 ml and the kinetics were collected by varying the concentration of hydrogen peroxidase. All reactions were repeated in triplicates, with a fresh bead employed every time. The initial rates were fitted to the Michaelis-Menten equation (equation **20**)). Concentrations of

HRP and C45 in the alginate beads were 100 µM and 10 µM, respectively. The recoverability assays were conducted by using the same bead in multiple runs, with ABTS and hydrogen peroxide concentrations of 1 mM, respectively.

9.17 ALGINATE-ENCAPSULATED CARBENE TRANSFERASE ASSAYS

9.17.1 Styrene cyclopropanation

To 970 µl of solvent (CHES buffer, MeOH, EtOH, isopropanol, acetone, acetonitrile, ethyl acetate, dichloromethane, THF, n-hexane, pet ether)) in a 1.5 ml glass vial was added 3 alginate-enzyme (10 µM, 1.4 mm^3) beads and 10 µl of $Na_2S_2O_4$ (1M stock; de-ionised water) and the mixture was left to stir for 1 minute (complete reduction of C45 from Fe^{3+} to Fe^{2+} could be verified *via* UV/VIS spectroscopy if necessary). 10 µl of a styrene (3M stock; EtOH) was added and the reaction left to mix for 30 seconds. After 30 seconds 10 µl of EDA (1M stock; EtOH) was added into the reaction vials to initiate the reaction. Final concentrations were 30 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 3M HCl. The vials were unscrewed, the beads were removed, and the resultant mixture was transferred to a 15 ml falcon tube containing 1 ml of ethyl acetate, vortexed and then centrifuged (14,500 rpm; 2 minutes). The organic layer was collected and was then slowly added to a 15 ml falcon tube containing 1 ml of an aqueous 3M NaOH solution. The resulting biphasic mixture was left to stir, at room temperature, for 2 hours; the progress of the hydrolysis reaction was monitored by TLC (7:3 ethyl acetate:hexane, 254 nm). After 2 hours, the aqueous layer was isolated analysed by chiral-HPLC and LC-MS. The product yields, enantiomeric excesses, and total turnover numbers (TTN; concentration of product formed/concentration of enzyme) where calculated *via* an external calibration with commercial ethyl 2-phenylcyclopropane-1-carboxylate and 2-phenylcyclopropane-1-carboxylic acid.

9.17.2 Piperidine N-H insertion

To 970 µl of solvent (CHES buffer, MeOH, EtOH, isopropanol, acetone, acetonitrile, ethyl acetate, dichloromethane, THF, n-hexane, pet ether, toluene) in a 1.5 ml glass vial was added 3 alginate-enzyme (10 µM, 1.4 mm^3) beads and 10 µl of $Na_2S_2O_4$ (1M stock; de-ionised water) and the mixture was left to stir for 1 minute (complete reduction of C45 from Fe^{3+} to Fe^{2+} could be verified *via* UV/VIS spectroscopy if necessary). 10 µl of a piperidine (3M stock; EtOH) was added and the reaction left to mix for 30 seconds. After 30 seconds 10 µl of EDA (1M stock; EtOH) was added into the reaction vials to initiate the reaction. Final concentrations were 30 µM enzyme, 10 mM sodium dithionite, 10 mM diazo compound, and 30 mM piperidine. Once mixed, the reactions were stirred on a roller at room

temperature. After 2 hours, the reaction vials were unscrewed, the beads were removed, and the mixture was transferred to a 15 ml falcon tube containing 2 ml of a 1:1 water:ethyl acetate mixture, vortexed and then centrifuged (14,500 rpm; 2 minutes). The organic layer was isolated and subsequently analysed by chiral-HPLC and LC-MS. The product yields and total turnover numbers (TTN; concentration of product formed/concentration of enzyme) where calculated *via* an external calibration with commercial ethyl-1-piperidineacetate.

9.18 DETERMINATION OF SUBSTRATE BULK VS BEADS PARTITION COEFFICIENTS

The partition coefficient, K_p, is defined in equation (**25**).

$$K_p = \frac{[Substrate]_{bead}}{[Substrate]_{bulk}}$$

Eq 25

Where $[Substrate]_{bead}$ is the concentration of a substrate in the alginate-bead and $[Substrate]_{bulk}$ is the concentration of a substrate in the bulk solution. C18-HPLC was used to determine the partition coefficient for styrene and piperidine in a variety of solvents as follows. An external calibration for styrene/piperidine was performed by injecting various concentrations of the sample and recording the peak height in the chromatogram at 280 nm (styrene) or 254 nm (piperidine; the concentration range was 500 µM – 10 mM. A batch of fresh beads were prepared according to the protocol provided in section 9.12 (without the encapsulation of any enzyme). The beads were collected and dried *via* vacuum filtration prior to use. The volume of the beads was determined according to the protocol provided in section 9.15. To a 2 ml screw-top vial was added 990 µl of the respective solvent (CHES, MeOH, EtOH, IPA, ACN, EtOAc, THF, acetone, DCM, n-hexane, pether, toluene), 10 µl of a 3M stock of piperidine/styrene (in EtOH) and three alginate beads; the final concentration of the styrene/piperidine was 10 mM. The vials were left to mix on a roller for 2 hours before the beads were removed from the vials prior to loading onto the HPLC. The concentration of the substrate remaining in the bulk solution was calculated from the peak height and was subsequently used to determine the partition coefficient of the substrate in each respective solvent.

The value, α, was calculated using equation (**26**), and was utilised in the calculations to determine the diffusion coefficients (section 9.19)

$$\alpha = \frac{V_b}{V_p K_p}$$

Eq 26

Where

- *Vb* = the total volume of the bulk solvent (ml)
- *Vp* = the total volume of the alginate beads (ml)
- K_p = the partition coefficient

9.19 DETERMINATION OF SUBSTRATE DIFFUSION COEFFICIENTS AS A FUNCTION OF SOLVENT

The protocol to determine the substrate diffusion coefficients in each respective solvent was identical to the protocol outlined in 9.20. The partition coefficients were calculated using the non-steady state equation (**27**).

$$S_b(t) = \frac{\alpha S_b(0)}{1+\alpha}\left[1 + 6(1+\alpha)\sum_{n=1}^{\infty}\frac{e^{\left(\frac{-Dq_n^2 t}{R^2}\right)}}{9 + 9\alpha + q_n^2\alpha^2}\right]$$

Eq 27

Where

- S_b(t) = The concentration of substrate, *S*, in the bulk at time *t* (mM)
- S_b(0) = The concentration of substrate, *S*, in the bulk at time 0 (mM)
- *R* = the radius of the alginate beads (m^3)
- q_n = the position non-zero roots of equation (**28**)
- *t* = time (s)
- *D* = the diffusion coefficient ($m^2\ s^{-1}$)

$$\tan(q_n) = \frac{3q_n}{3 + \alpha q_n}$$

Eq 28

The non-positive roots for equation 26, and the value of q_n, was solved using the Secant method (equation (**29**)) and by setting the initial guess as $x_0 = 0$ and $x_1 = 1$.

$$x_n = x_{n-1} - f(x_{n-1})\frac{x_{n-1} - x_{n-2}}{f(x_{n-1}) - f(x_{n-2})}$$

Eq 29

The effective diffusion coefficient, D_{eff} (m^2 s^{-1}), for a substrate each solvent was calculated using equation (**30**).

$$D_{eff} = \varepsilon D$$

Eq 30

Where ε is the porosity. For alginate beads prepared from 0.3M $CaCl_2$ and 3% (w/v) sodium alginate the porosity was calculated to be 0.889 (following the method delineated in *section 9.15.3*).

9.20 ALGINATE-ENCAPSULATED CARBENE TRANSFER RECOVERABILITY ASSAYS

The recoverability assays were conducted identically to the protocol outlined *in section 9.17*. After each reaction, the sealed vials were transported into an anaerobic glove box (Belle technology, $[O_2]$ < 5 ppm), the lids were removed, and the product mixture extracted was decanted into an Eppendorf for analysis. The remaining beads were washed with twice with respective solvent before the subsequent reaction.

9.21 CIRCULAR DICHROISM SPECTROSCOPY

Circular Dichroism spectra were collected using a JASCO J-815 CD polarimeter. The enzyme was loaded into a 1 mm pathlength quartz cell at 0.01 mg mL^{-1} in mM KCl, 20 mM CHES, pH 8.58, and far-UV CD spectra were recorded at 100 nm/min with a sensitivity of 50 mdeg. Thermal stability of an enzyme was assessed by monitoring ellipticity at 222 nm while increasing the temperature at a ramp rate of 40 °C/hour with 1 °C intervals. All the raw data was converted to mean residue ellipticity (MRE) and the thermal denaturation transition midpoint (T_m) was assessed by plotting the second derivatives of a smoothed thermal denaturation trace where the x-axis intercept corresponds to the T_m.

9.22 REDOX POTENTIOMETRY

The redox potential of an enzyme was measured as previously described for C45. Briefly, an enzyme (50 µM) was loaded into a spectroelectrochemical cell in 100 mM KCl, 50 mM CHES, pH 8.6, 10% glycerol with the following redox mediators: 20 µM benzyl viologen, 20 µM anthroquinone-2-sulfonate, 20 µM phenazine, 25 µM 2-hydroxy-1,4-napthoquinone. Potential was applied to the spectroelectrochemical cell (platinum working and counter electrodes, Ag/AgCl reference electrode) using a Biologic SP-150 potentiostat in both reductive and oxidative directions to confirm equilibration. The enzyme's heme reduction potential was calculated by plotting the fraction of

reduced protein versus the applied potential, and the data were fitted using the following single electron Nernst function (equation (**31**)):

$$f(x) = (A+B*10^{((E_m-x)/59)})/(1+10^{((E_m-x)/59)})$$

Eq 31

In equation (**31**), A and B are y-axis values at 100% oxidized and reduced heme respectively; E_m is the heme reduction potential.

10 APPENDICES

APPENDIX 1: EXTERNAL CALIBRATIONS

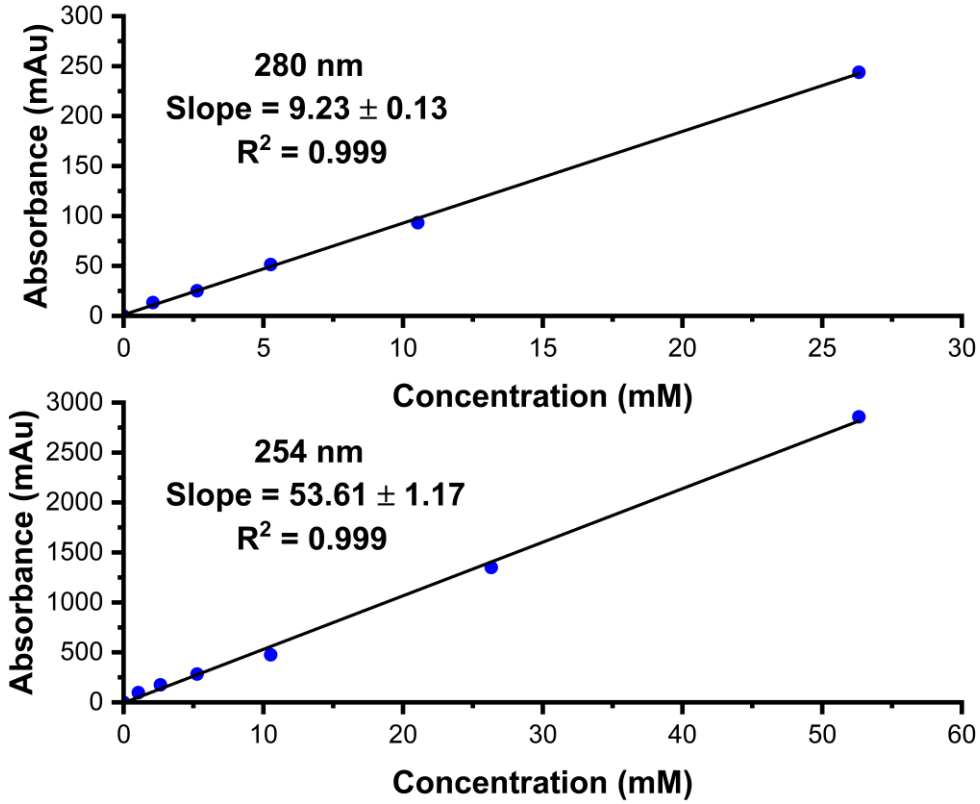


Figure 163: Chiral-HPLC external calibrations for ethyl 2-phenylcyclopropane-1- carboxylate at 280 nm (upper panel) and 254 nm (lower panel). A polar ionic mobile phase (100 % CH_3CN: 0.1 % v/v TFA: 0.1 % v/v: Et_3N) was employed and injection volumes were 2 µl.

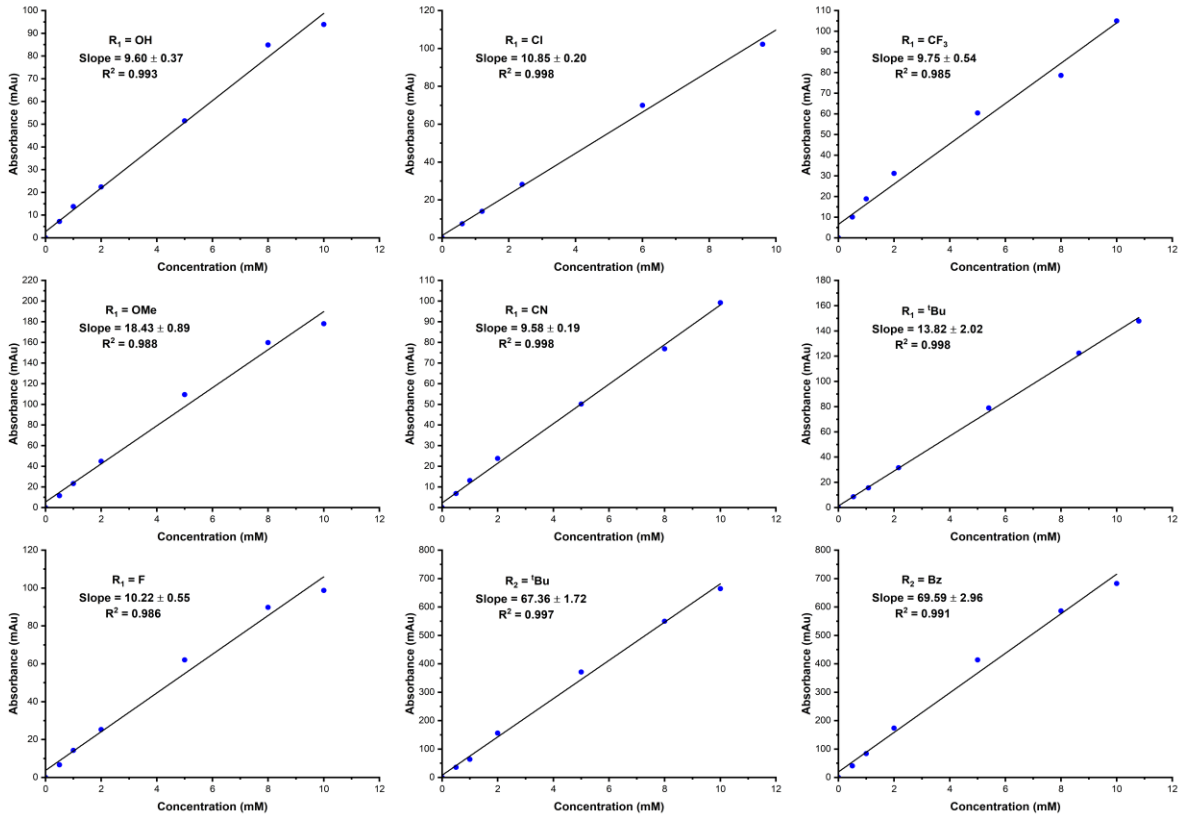


Figure 164: Chiral-HPLC external calibrations for (**top left**) ethyl 2-(4-hydroxyphenyl)cyclopropane-1-carboxylate at 280 nm, (**top middle**) ethyl 2-(4-chlorophenyl)cyclopropane-1-carboxylate recorded at 280 nm, (**top right**) ethyl 2-(4-(trifluoromethyl)phenyl)cyclopropane-1-carboxylate recorded at 280 nm, (**middle left**) ethyl 2-(4-methoxyphenyl)cyclopropane-1-carboxylate recorded at 280 nm, (**middle**) ethyl 2-(4-cyanophenyl)cyclopropane-1-carboxylate recorded at 280 nm, (**middle right**) ethyl 2-(4-(*tert*-butyl)phenyl)cyclopropane-1-carboxylate recorded at 280 nm, (**bottom left**) ethyl 2-(4-fluorophenyl)cyclopropane-1-carboxylate recorded at 280 nm, (**bottom middle**) *tert*-butyl 2-phenylcyclopropane-1-carboxylate recorded at 254 nm, and (**bottom right**) benzyl 2-phenylcyclopropane-1-carboxylate recorded at 254 nm. A polar ionic mobile phase (100 % CH_3CN: 0.1 % v/v TFA: 0.1 % v/v: Et_3N) was employed and injection volumes were 2 µl.

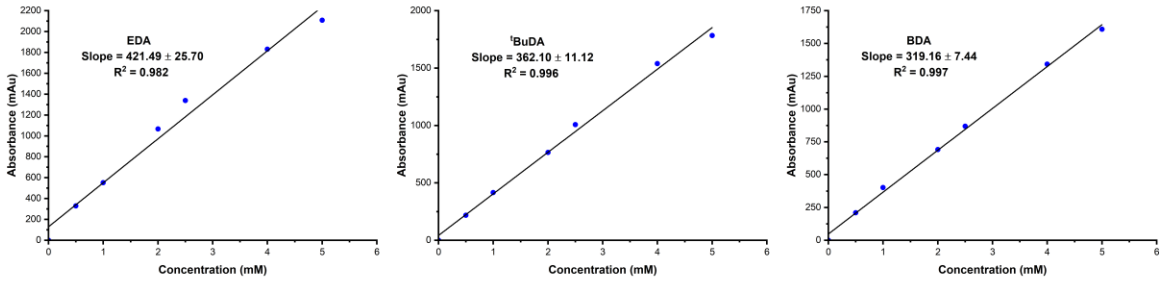


Figure 165: Chiral-HPLC external calibration (**left**) for ethyl (4-chlorophenyl)glycinate at 254 nm, (**middle**) *tert*-butyl (4-chlorophenyl)glycinate at 254 nm, and (**right**) benzyl (4-chlorophenyl)glycinate at 254 nm. A polar ionic mobile phase (100 % CH_3CN: 0.1 % v/v TFA: 0.1 % v/v: Et_3N) was employed and injection volumes were 2 µl.

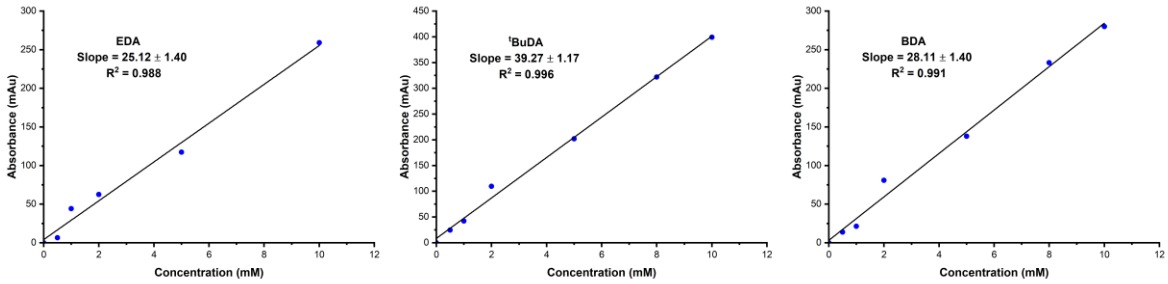


Figure 166: Chiral-HPLC external calibration (**left**) for ethyl-1-piperidineacetate at 254 nm, (**middle**) *tert*-butyl 2-(piperidin-1-yl)acetate at 254 nm, and (**right**) benzyl 2-(piperidin-1-yl)acetate at 254 nm. A polar ionic mobile phase (100 % CH_3CN: 0.1 % v/v TFA: 0.1 % v/v: Et_3N) was employed and injection volumes were 2 µl.

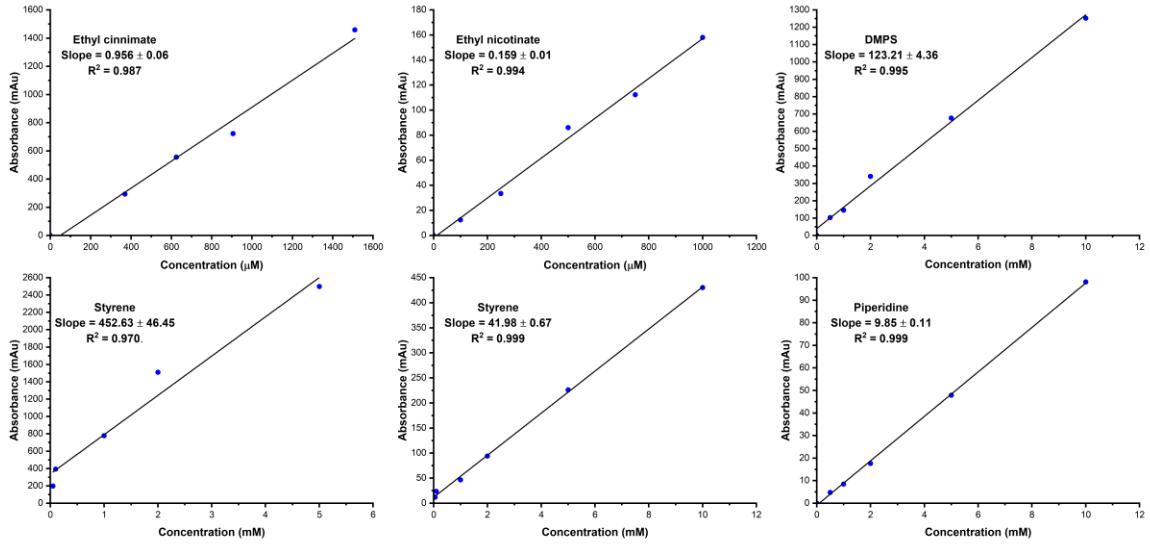


Figure 167: C18-HPLC external calibrations for (**top left**) ethyl cinnamate at 245 nm (a polar ionic mobile phase (100 % CH_3CN: 0.1 % v/v TFA: 0.1 % v/v: Et_3N) was employed), (**top middle**) ethyl nicotinate at 265 nm (a gradient mobile phase (70:30% H_2O:CH_3CN to 10:90% H_2O:CH_3CN) was employed), (**top right** dimethylphenylsilane at 220 nm (a gradient mobile phase (70:30% H_2O:CH_3CN to 10:90% H_2O:CH_3CN) was employed), (**bottom left**) styrene at 254 nm (a polar ionic mobile phase (100 % CH_3CN: 0.1 % v/v TFA: 0.1 % v/v: Et_3N) was employed), (**bottom middle**) styrene at 280 nm (a polar ionic mobile phase (100 % CH_3CN: 0.1 % v/v TFA: 0.1 % v/v: Et_3N) was employed), and (**bottom right**) piperidine at 254 nm (a polar ionic mobile phase (100 % CH_3CN: 0.1 % v/v TFA: 0.1 % v/v: Et_3N) was employed). Injection volumes for all calibrations were 8 µl.

APPENDIX 2: TABLES

Table 17: A non-exhaustive list of rationally designed carbene transferases reported in the literature, including reaction conditions, product yields, enantioselectivities and total turnover numbers for the cyclopropanation reaction between EDA and styrene.

ENZYME	[EDA] / mM	[Styrene] / mM	[E] / µM	% Yield	ee%	TTN
Mb(F43V,V68A)	10.00	30.00	20.00	91.00	67.00	455.00
Mb(F43V,V68F)	10.00	30.00	20.00	99.90	99.90	500.00
Mb(H64V,V68A)	10.00	30.00	20.00	99.90	99.90	500.00
Mb(L29T,H64V)	20.00	10.00	20.00	93.00	59.00	460.00
Mb(H64V,V68S)	20.00	10.00	20.00	82.00	99.00	410.00
P411BM3	8.50	20.00	1.50	50.00	12.00	2900.00
P411BM3-CIS	8.50	20.00	1.30	55.00	96.00	3700.00
P450BM3-T268A	10.00	30.00	20.00	65.00	96.00	323.00
P411BM3-CIS	10.00	30.00	20.00	40.00	94.00	199.00
BM3-CIS-1263A	10.00	30.00	20.00	38.00	91.00	190.00
BM3-CIS-A328G	10.00	30.00	20.00	37.00	52.00	186.00
BM3-CIS-T438S	10.00	30.00	20.00	59.00	97.00	293.00
H2-5-F10	10.00	30.00	20.00	59.00	63.00	294.00
P450BM3-C400S	10.00	20.00	10.00	41.00	rac	411.00
P450CAM-C357S	10.00	20.00	10.00	49.00	rac	487.00
T268-axH	8.70	10.00	1.00	81.00	42.00	710.00
LmrR_M8A	10.00	30.00	11.00	39.00	38.00	391.00

Table 18: The product yields, total turnover numbers, and single/double insertion ratios for the C45-catalysed N-H insertion reactions between piperidine/*p*-chloroaniline and the diazo compounds EDA, ^{t}BuDA, and BnDA.

Substrates	% Yield$_{single}$	% Yield$_{double}$	Single/Double	TTN$_{(single)}$
EDA/Piperidine	78.90 (± 11.90)	-	-	789
EDA/p-chloroaniline	67.59 (± 0.38)	5.29 (± 5.01)	~12.7:1	678
^{t}BuDA/Piperidine	62.10 (± 8.85)	-	-	621
^{t}BuDA/p-chloroaniline	72.50 (± 0.27)	0	n/a	725
BDA/Piperidine	88.34 (± 13.14)	-	-	883
BDA/p-chloroaniline	89.44 (± 0.27)	0	n/a	894

Table 19: The reaction yields and total turnover numbers for the Si-H insertion reaction between EDA and DMPS catalysed by C45 and *Rma*-TDE.

Si-H INSERTIONS		
EDA/DMPS	**% Yield**	**TTN**
***Rma*-TDE**	57.66 (± 4.14)	577
C45	62.86 (± 1.95)	629

Table 20: The product yields and enantioselectivities of the AE-C45 catalysed cyclopropanation reaction between EDA and styrene in various organic solvents, followed by the recovery and re-employment of the same AE-C45 catalyst in the N-H insertion reaction between EDA and piperidine in various organic solvents.

Recoverability: Cyclopropanation + N-H insertion				
Solvent	**% Yield$_{(R,R)}$**	**% Yield$_{(S,S)}$**	**ee$_{(R,R)}$ %**	**%Yield$_{N-H}$**
CHES	62.50 (± 1.50)	4.73 (± 0.78)	86.00 (± 1.90)	9.0 (± 0.36)
Acetone	78.56 (± 1.33)	0.13 (± 0.03)	>99 (± 0.08)	10.46 (± 0.63)
ACN	77.63 (± 1.76)	0.12 (± 0.05)	>99 (± 0.12)	10.37 (± 0.50)
THF	67.96 (± 1.42)	0.14 (± -0.01)	>99 (±0.03)	9.89 (± 0.61)

Table 21: The yields, diastereoselectivities, enantioselectivities, and total turnover numbers for the cyclopropanation reaction between styrene and EDA catalysed by the *de novo* heme-containing maquettes C45, JR1, AP3.2, and the rationally designed heme-containing proteins Mb(H64V,V68A), and *Rma*-TDE.

Enzyme	% Yield$_{(R.R)}$	% Yield$_{(S.S)}$	dr$_{(E)}$ %	ee$_{(R,R)}$ %	ee$_{(S,S)}$ %	TTN
C45	80.20 (± 8.01)	10.42 (± 1.81)	>99	77.00 (± 2.82)	-	802
JR1	85.61 (± 4.32)	11.04 (± 4.84)	>99	78.00 (± 8.5)	-	856
AP3.2	0	41.45 (± 4.6)	>99	-	>99.9	415
Mb(H64V,V68A)	0	95.90 (± 0.51)	>99	-	>99.9	959
***Rma*-TDE**	73.47 (± 5.73)	12.64 (± 0.97)	>99	70.58 (± 2.02)	-	735

Table 22: The kinetic parameters for the metallocarbenoid formation between EDA and the *de novo* heme-containing maquettes C45, JR1, AP3.2, and the rationally designed heme-containing proteins Mb(H64V,V68A), and *Rma*-TDE.

EDA	C45	JR1	AP3.2	Mb(H64V,V68A)	*Rma*-TDE
k_{obs} (s^{-1})	0.498 (± 0.013)	0.553 (± 0.022)	0.820 (± 0.018)	0.224 (± 0.013)	0.144 (± 0.010)
K_1 (μM)	74.36 (± 9.75)	21.68 (± 6.12)	69.26 (± 8.40)	15.04 (± 4.68)	487.94 (± 49.22)
k_{obs}/K_1 (s^{-1} M^{-1})	6.7 x 10^3	2.55 x 10^4	1.18 x 10^4	1.50 x 10^4	294.26
$ln(2)/k_{obs}$	1.39 s	1.25 s	1 s	3.1 s	4.83 s

Table 23: The kinetic parameters for the cyclopropanation reaction between styrene and EDA catalysed by the *de novo* heme-containing maquettes C45, JR1, AP3.2, and the rationally designed heme-containing proteins Mb(H64V,V68A), and *Rma*-TDE.

Styrene	C45	JR1	AP3.2	Mb(H64V,V68A)	*Rma*-TDE
k_{lim} (s^{-1})	0.0011 (± 9x10^{-5})	0.048 (± 0.002)	0.067 (± 0.002)	0.021 (± 0.002)	0.0104 (± 0.0004)
V_{max} (μM s^{-1})	0.0083	0.36	0.503	0.158	0.0078
K_1 (μM)	482.51 (± 108.8)	352.1 (± 52.8)	141.23 (± 17.1)	657.7 (± 110.3)	795.92 (± 82.7)
k_{lim}/K_1 (s^{-1} M^{-1})	2.23	136.32	474.4	31.93	1.37
$ln(2)/k_{lim}$	10.5 min	14.44 s	10.34 s	33 s	11 min

11 BIBLIOGRAPHY

1. Moerdyk, J. P. & Bielawski, C. W. Stable Carbenes. *Contemp. Carbene Chem.* 40–74 (2013). doi:10.1002/9781118730379.ch2

2. Hermann, M. Ueber die bei der technischen Gewinnung des Broms beobachtete flüchtige Bromverbindung. *Justus Liebigs Ann. Chem.* **95**, 211–225 (1855).

3. de Frémont, P., Marion, N. & Nolan, S. P. Carbenes: Synthesis, properties, and organometallic chemistry. *Coordination Chemistry Reviews* **253**, 862–892 (2009).

4. Nef, J. U. Ueber das zweiwerthige Kohlenstoffatom. (Vierte Abhandlung.) Die Chemie des Methylens. *Justus Liebigs Ann. Chem.* **298**, 202–374 (1897).

5. Staudinger, H. & Kupfer, O. Über Reaktionen des Methylens. III. Diazomethan. *Berichte der Dtsch. Chem. Gesellschaft* **45**, 501–509 (1912).

6. Von Doering, W. E. & Knox, L. H. Synthesis of Substituted Tropolones. *J. Am. Chem. Soc.* **75**, 297–303 (1953).

7. Horspool, W. M. & Sutherland, R. G. *The addition of dichlorocarbene to ferrocenylethylenes*. *Chemical Communications (London)* **76**, (UTC, 1966).

8. DeMore, W. B., Pritchard, H. O. & Davidson, N. photochemical experiments in rigid media at low temperatures. II. The reactions of methylene, cyclopentadienylene and diphenylmethylene. *J. Am. Chem. Soc.* **81**, 5874–5879 (1959).

9. Fischer, E. O. & Dipl.-Chem, A. M. On the existence of a tungsten carbonyl carbene complex. *Angew.Chem.Int.Ed* **3**, 580–581 (1964).

10. Tschugajeff, L., Skanawy-Grigorjewa, M., Posnjak, A. & Skanawy-Grigorjewa, M. Über Die Hydrazin-Carbylamin-Komplexe des Platins. *Zeitschrift für Anorg. und Allg. Chemie* **148**, 37–42 (1925).

11. Irikura, K. K., Goddard, W. A. & Beauchamp, J. L. Singlet-Triplet Gaps in Substituted Carbenes CXY (X, Y = H, F, Cl, Br, I, SiH3). *J. Am. Chem. Soc.* **114**, 48–51 (1992).

12. Mueller, P. H. *et al.* Carbene Singlet-Triplet Gaps. Linear Correlations with Substituent π Donation. *J. Am. Chem. Soc.* **103**, 5049–5052 (1981).

13. Harrison, J. F., Liedtke, R. C. & Liebman, J. F. The Multiplicity of Substituted Acyclic Carbenes and Related Molecules. *J. Am. Chem. Soc.* **101**, 7162–7168 (1979).

14. Bauschlicher, C. W., Schaefer, H. F. & Bagus, P. S. Structure and Energetics of Simple Carbenes CH2, CHF, CHCl, CHBr, CF2, and CCl2. *J. Am. Chem. Soc.* **99**, 7106–7110 (1977).

15. Harrison, J. F. Electronic Structure of Carbenes. I. CH2, CHF, and CF2. *J. Am. Chem. Soc.* **93**, 4112–4119 (1971).

16. Feller, D., Thatcher Borden, W. & Davidson, E. R. Dependence of the singlet-triplet splitting in heterosubstituted carbenes on the heteroatom electronegativity and conformation. *Chem. Phys. Lett.* **71**, 22–26 (1980).

17. Sreenilayam, G. & Fasan, R. Myoglobin-catalyzed intermolecular carbene N-H insertion with arylamine substrates. *Chem. Commun.* **51**, 1532–1534 (2015).

18. Cardin, D. J., Cetinkaya, B. & Lappert, M. F. TRANSITION METAL-CARBENE COMPLEXES. *Chem. Rev* **72**, 545–574 (1972).

19. Garrison, J. C. & Youngs, W. J. Ag(I) N-heterocyclic carbene complexes: Synthesis, structure, and application. *Chemical Reviews* **105**, 3978–4008 (2005).

20. Brookhart, M. & Studabaker, W. B. Cyclopropanes from Reactions of Transition-Metal-Carbene Complexes with Olefins. *Chem. Rev.* **87**, 411–432 (1987).

21. Dötz, K. H. & Stendel, J. Fischer carbene complexes in organic synthesis: Metal-assisted and metal-templated reactions. *Chemical Reviews* **109**, 3227–3274 (2009).

22. Öfele, K., Tosh, E., Taubmann, C. & Herrmann, W. A. Carbocyclic carbene metal complexes. *Chemical Reviews* **109**, 3408–3444 (2009).

23. Montgomery, C. D. Fischer and Schrock Carbene Complexes: A Molecular Modeling Exercise. *J. Chem. Educ.* **92**, 1653–1660 (2015).

24. Rowlands, G. Carbenes and Carbene Complexes I. *Http://Www.Massey.Ac.Nz/~Gjrowlan/Adv/Lct6.Pdf* 1–11 (2011).

25. Hofmann, P. & Hämmerle, M. The Mechanism of the Dötz Reaction: Chromacyclobutenes by Alkyne–Carbene Coupling? *Angew. Chemie Int. Ed. English* **28**, 908–910 (1989).

26. Wulff, W. D., Tang, P. C. & Stuart McCallum, J. Regiochemistry of the Reaction of Chromium-Carbene Complexes with Acetylenes. *Journal of the American Chemical Society* **103**, 7677–7678 (1981).

27. Tebbe, F. N., Parshall, G. W. & Reddy, G. S. Olefin Homologation with Titanium Methylene Compounds. *J. Am. Chem. Soc.* **100**, 3611–3613 (1978).

28. Hartley, R. C., Li, J., Main, C. A. & McKiernan, G. J. Titanium carbenoid reagents for converting carbonyl groups into alkenes. *Tetrahedron* **63**, 4825–4864 (2007).

29. Dzik, W. I., Zhang, X. P. & De Bruin, B. Redox noninnocence of carbene ligands: Carbene radicals in (Catalytic) C-C bond formation. *Inorganic Chemistry* **50**, 9896–9903 (2011).

30. Lu, H. *et al.* Experimental evidence for cobalt(III)-carbene radicals: Key intermediates in cobalt(II)-based metalloradical cyclopropanation. *J. Am. Chem. Soc.* **133**, 8518–8521 (2011).

31. Das, B. G., Chirila, A., Tromp, M., Reek, J. N. H. & De Bruin, B. CoIII-Carbene Radical Approach to Substituted 1H-Indenes. *J. Am. Chem. Soc.* **138**, 8968–8975 (2016).

32. Chirila, A., Gopal Das, B., Paul, N. D. & de Bruin, B. Diastereoselective Radical-Type Cyclopropanation of Electron-Deficient Alkenes Mediated by the Highly Active Cobalt(II) Tetramethyltetraaza[14]annulene Catalyst. *ChemCatChem* **9**, 1413–1421 (2017).

33. Cui, X., Xu, X., Jin, L. M., Wojtas, L. & Zhang, X. P. Stereoselective radical C-H alkylation with acceptor/acceptor-substituted diazo reagents via Co(II)-based metalloradical catalysis. *Chem. Sci.* **6**, 1219–1224 (2015).

34. Russell, S. K. *et al.* Synthesis, electronic structure and reactivity of bis(imino)pyridine iron carbene complexes: Evidence for a carbene radical. *Chem. Sci.* **5**, 1168–1174 (2014).

35. Belof, J. L. *et al.* Characterization of tunable radical metal-carbenes: Key intermediates in catalytic cyclopropanation. *Organometallics* **30**, 2739–2746 (2011).

36. Mix, K. A., Aronoff, M. R. & Raines, R. T. Diazo Compounds: Versatile Tools for Chemical Biology. *ACS Chemical Biology* **11**, 3233–3244 (2016).

37. Ford, A. *et al.* Modern Organic Synthesis with α-Diazocarbonyl Compounds. *Chem. Rev.* **115**, 9981–10080 (2015).

38. Khade, R. L. *et al.* Iron porphyrin carbenes as catalytic intermediates: Structures, mössbauer and NMR spectroscopic properties, and bonding. *Angew. Chemie - Int. Ed.* **53**, 7574–7578 (2014).

39. Sharon, D. A., Mallick, D., Wang, B. & Shaik, S. Computation Sheds Insight into Iron Porphyrin Carbenes' Electronic Structure, Formation, and N-H Insertion Reactivity. *J. Am. Chem. Soc.* **138**, 9597–9610 (2016).

40. Khade, R. L. & Zhang, Y. Catalytic and Biocatalytic Iron Porphyrin Carbene Formation: Effects of Binding Mode, Carbene Substituent, Porphyrin Substituent, and Protein Axial Ligand. *J. Am. Chem. Soc.* **137**, 7560–7563 (2015).

41. Taxak, N., Patel, B. & Bharatam, P. V. Carbene generation by cytochromes and electronic structure of heme-iron-porphyrin-carbene complex: A quantum chemical study. *Inorg. Chem.* **52**, 5097–5109 (2013).

42. Ebner, C. & Carreira, E. M. Cyclopropanation Strategies in Recent Total Syntheses. *Chem. Rev.* **117**, 11651–11679 (2017).

43. Xu, L., Doubleday, C. E. & Houk, K. N. Dynamics of carbene cycloadditions. *J. Am. Chem. Soc.* **133**, 17848–17854 (2011).

44. Skatteböl, L. Allenes from gem-dihalocyclopropane derivatives and alkyllithium. *Tetrahedron Lett.* **2**, 167–172 (1961).

45. Von Doering, W. E. & LaFlamme, M. P. A TWO-STEP SYNTHlESIS OF ALLENES FROM OLEFZNS. **2**, 75–79 (1958).

46. Grasse, P. B., Brauer, B. E., Zupancic, J. J., Kaufmann, K. J. & Schuster, G. B. Chemical and Physical Properties of Fluorenylidene: Equilibration of the Singlet and Triplet Carbenes. *J. Am. Chem. Soc.* **105**, 6833–6845 (1983).

47. Wang, J. *et al.* Ultrafast study of 9-diazofluorene: Direct observation of the first two singlet states of fluorenylidene. *J. Am. Chem. Soc.* **129**, 13683–13690 (2007).

48. Sitzmann, E. V., Langan, J. & Eisenthal, K. B. Intermolecular Effects on Intersystem Crossing Studied on the Picosecond Time Scale: The Solvent Polarity Effect on the Rate of Singlet to Triplet Intersystem Crossing of Diphenylcarbene. *J. Am. Chem. Soc.* **106**, 1868–1869 (1984).

49. Davies, H. M. L., Huby, N. J. S., Cantrell, W. R. & Olive, J. L. α-Hydroxy Esters as Chiral Auxiliaries in Asymmetric Cyclopropanations by Rhodium(II)-Stabilized Vinylcarbenoids. *J. Am. Chem. Soc.* **115**, 9468–9479 (1993).

50. Negretti, S., Cohen, C. M., Chang, J. J., Guptill, D. M. & Davies, H. M. L. Enantioselective dirhodium(II)-catalyzed cyclopropanations with trimethylsilylethyl and trichloroethyl aryldiazoacetates. *Tetrahedron* **71**, 7415–7420 (2015).

51. Huang, L., Chen, Y., Gao, G. Y. & Zhang, X. P. Diastereoselective and enantioselective cyclopropanation of alkenes catalyzed by cobalt porphyrins. *J. Org. Chem.* **68**, 8179–8184 (2003).

52. Anding, B. J., Ellern, A. & Woo, L. K. Olefin cyclopropanation catalyzed by iridium(III) porphyrin complexes. *Organometallics* **31**, 3628–3635 (2012).

53. Doyle, M. P. Exceptional selectivity in cyclopropanation reactions catalyzed by chiral cobalt(II)-porphyrin catalysts. *Angewandte Chemie - International Edition* **48**, 850–852 (2009).

54. Kelly, L. F. A synthesis of chrysanthemic ester - An undergraduate experiment. *J. Chem. Educ.* **64**, 1061 (1987).

55. Simmons, H. E. & Smith, R. D. A new synthesis of cyclopropanes from olefins. *J. Am. Chem. Soc.* **80**, 5323–5324 (1958).

56. Furukawa, J., Kawabata, N. & Nishimura, J. Synthesis of cyclopropanes by the reaction of olefins with dialkylzinc and methylene iodide. *Tetrahedron* **24**, 53–58 (1968).

57. Poulter, C. D., Friedrich, E. C. & Winstein, S. The Stereochemistry of the Methylene Iodide-Zinc-Copper Couple Methylenation of Cyclic Allylic Alcohols. *J. Am. Chem. Soc.* **91**, 6892–6894 (1969).

58. 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).

59. Ye, T. & Mckervey ', M. A. Organic Synthesis with a-Diazocarbonyl Compounds. *Chem. Rev* **94**, 1091–1160 (1884).

60. Keipour, H., Jalba, A., Delage-Laurin, L. & Ollevier, T. Copper-Catalyzed Carbenoid Insertion

Reactions of α-Diazoesters and α-Diazoketones into Si-H and S-H Bonds. *J. Org. Chem.* **82**, 3000–3010 (2017).

61. Khade, R. L. & Zhang, Y. C–H Insertions by Iron Porphyrin Carbene: Basic Mechanism and Origin of Substrate Selectivity. *Chem. - A Eur. J.* **23**, 17654–17658 (2017).

62. Gandeepan, P. *et al.* 3d Transition Metals for C-H Activation. *Chem. Rev.* **119**, 2192–2452 (2019).

63. Abrams, D. J., Provencher, P. A. & Sorensen, E. J. Recent applications of C-H functionalization in complex natural product synthesis. *Chem. Soc. Rev.* **47**, 8925–8967 (2018).

64. Doyle, M. P., Duffy, R., Ratnikov, M. & Zhou, L. Catalytic carbene insertion into C-H bonds. *Chem. Rev.* **110**, 704–724 (2010).

65. Gillingham, D. & Fei, N. Catalytic x–h insertion reactions based on carbenoids. *Chem. Soc. Rev.* **42**, 4918–4931 (2013).

66. Doyle, M. P. Perspective on dirhodium carboxamidates as catalysts. *J. Org. Chem.* **71**, 9253–9260 (2006).

67. Demonceau, A., Noels, A. F., Hubert, A. J. & Teyssié, P. Transition-metal-catalysed reactions of diazoesters. Insertion into C-H bonds of paraffins by carbenoids. *J. Chem. Soc. Chem. Commun.* 688–689 (1981). doi:10.1039/C39810000688

68. Ramakrishna, K. & Sivasankar, C. Iridium catalyzed acceptor/acceptor carbene insertion into N-H bonds in water. *Org. Biomol. Chem.* **15**, 2392–2396 (2017).

69. Lee, E. C. & Fu, G. C. Copper-catalyzed asymmetric N-H insertion reactions: Couplings of diazo compounds with carbamates to generate α-amino acids. *J. Am. Chem. Soc.* **129**, 12066–12067 (2007).

70. Arndt, F. & Eistert, B. Ein Verfahren zur Überführung von Carbonsäuren in ihre höheren Homologen bzw. deren Derivate. *Berichte der Dtsch. Chem. Gesellschaft (A B Ser.* **68**, 200–208 (1935).

71. Baratta, W., Del Zotto, A. & Rigo, P. Highly stereoselective formation of cis-enediones from α-diazo carbonyl compounds catalysed by [RuCl(η5-C5H5)PPh32]. *Chem. Commun.* 2163–2164 (1997). doi:10.1039/a706459d

72. Wanzlick, H. W. Aspects of Nucleophilic Carbene Chemistry. *Angew. Chemie Int. Ed. English* **1**, 75–80 (1962).

73. Liu, Y. & Lemal, D. M. Concerning the 'Wanzlick equilibrium'. *Tetrahedron Lett.* **41**, 599–602 (2000).

74. Böhm, V. P. W. & Herrmann, W. A. The 'Wanzlick equilibrium'. *Angewandte Chemie - International Edition* **39**, 4036–4038 (2000).

75. Denk, M. K., Hatano, K. & Ma, M. Nucleophilic carbenes and the Wanzlick equilibrium: A reinvestigation. *Tetrahedron Lett.* **40**, 2057–2060 (1999).

76. Hahn, F. E., Wittenbecher, L., Le Van, D. & Fröhlich, R. Evidence for an equilibrium between an N-heterocyclic carbene and its dimer in solution. *Angew. Chemie - Int. Ed.* **39**, 541–544 (2000).

77. Kirmse, W. The beginnings of N-heterocyclic carbenes. *Angew. Chemie - Int. Ed.* **49**, 8798–8801 (2010).

78. Wanzlick, H.-W. & Schikora, E. Ein neuer Zugang zur Carben-Chemie. *Angew. Chemie* **72**, 494–494 (1960).

79. Wanzlick, H. -W & Schikora, E. Ein nucleophiles Carben. *Chem. Ber.* **94**, 2389–2393 (1961).

80. Lemal, D. M., Lovald, R. A. & Kawano, K. I. Tetraaminoethylenes. The question of dissociation. *J. Am. Chem. Soc.* **86**, 2518–2519 (1964).

81. Winberg, E. H., Downing, J. R., Coffman, D. D. & Brown, M. Tetraaminoethylenes. *J. Am. Chem. Soc.* **87**, 2055–2056 (1965).

82. Liu, Y., Lindner, P. E. & Lemal, D. M. Thermodynamics of a diaminocarbene-tetraaminoethylene equilibrium [1]. *Journal of the American Chemical Society* **121**, 10626–10627 (1999).

83. Reetz, M. T. Biocatalysis in organic chemistry and biotechnology: Past, present, and future. *Journal of the American Chemical Society* **135**, 12480–12496 (2013).

84. Sheldon, R. A. & Pereira, P. C. Biocatalysis engineering: The big picture. *Chem. Soc. Rev.* **46**, 2678–2691 (2017).

85. Schoemaker, H. E., Mink, D. L. & WubboLts, M. G. Dispelling the myths - Biocatalysis in industrial synthesis. *Science* **299**, 1694–1697 (2003).

86. Padmanabhan, R., Jay, E. & Wu, R. Chemical synthesis of a primer and its use in the sequence analysis of the lysozyme gene of bacteriophage T4. *Proc. Natl. Acad. Sci. U. S. A.* **71**, 2510–2514 (1974).

87. Sanger, F., Nicklen, S. & Coulson, A. R. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463–5467 (1977).

88. Fleischmann, R. D. *et al.* Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. *Science (80-.).* **269**, 496–512 (1995).

89. Chien, A., Edgar, D. B. & Trela, J. M. Deoxyribonucleic acid polymerase from the extreme thermophile Thermus aquaticus. *J. Bacteriol.* **127**, 1550–1557 (1976).

90. Cohen, S. N. DNA cloning: A personal view after 40 years. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 15521–15529 (2013).

91. Porter, J. L., Rusli, R. A. & Ollis, D. L. Directed Evolution of Enzymes for Industrial Biocatalysis. *ChemBioChem* **17**, 197–203 (2016).

92. Lin, Y. W. Rational design of metalloenzymes: From single to multiple active sites. *Coordination Chemistry Reviews* **336**, 1–27 (2017).

93. Strohmeier, G. A., Pichler, H., May, O. & Gruber-Khadjawi, M. Application of designed enzymes in organic synthesis. *Chemical Reviews* **111**, 4141–4164 (2011).

94. Yu, F. *et al.* Protein design: Toward functional metalloenzymes. *Chemical Reviews* **114**, 3495–3578 (2014).

95. Huang, P. S., Boyken, S. E. & Baker, D. The coming of age of de novo protein design. *Nature* **537**, 320–327 (2016).

96. Baltzer, L., Nilsson, H. & Nilsson, J. De novo design of proteins - What are the rules? *Chemical Reviews* **101**, 3153–3163 (2001).

97. Lichtenstein, B. R. *et al.* Engineering oxidoreductases: Maquette proteins designed from scratch. in *Biochemical Society Transactions* **40**, 561–566 (2012).

98. Grayson, K. J. & Anderson, J. R. The ascent of man(made oxidoreductases). *Current Opinion in Structural Biology* **51**, 149–155 (2018).

99. Monti, D., Ottolina, G., Carrea, G. & Riva, S. Redox reactions catalyzed by isolated enzymes. *Chemical Reviews* **111**, 4111–4140 (2011).

100. Liu, J. *et al.* Metalloproteins containing cytochrome, iron-sulfur, or copper redox centers. *Chemical Reviews* **114**, 4366–4369 (2014).

101. Poulos, T. L. Heme enzyme structure and function. *Chemical Reviews* **114**, 3919–3962 (2014).

102. Tong, Y. & Guo, M. Bacterial heme-transport proteins and their heme-coordination modes. *Archives of Biochemistry and Biophysics* **481**, 1–15 (2009).

103. Shoji, O. & Watanabe, Y. Peroxygenase reactions catalyzed by cytochromes P450 Topical Issue in honor of Ivano Bertini. Guest editors: Lucia Banci, Claudio Luchinat. *Journal of Biological Inorganic Chemistry* **19**, 529–539 (2014).

104. Smith, A. T. *et al.* Functional divergence of heme-thiolate proteins: A classification based on spectroscopic attributes. *Chemical Reviews* **115**, 2532–2558 (2015).

105. Guo, R., Gu, J., Zong, S., Wu, M. & Yang, M. Structure and mechanism of mitochondrial electron transport chain. *Biomedical Journal* **41**, 9–20 (2018).

106. Liu, Y., Fiskum, G. & Schubert, D. Generation of reactive oxygen species by the mitochondrial electron transport chain. *J. Neurochem.* **80**, 780–787 (2002).

107. Terwilliger, N. B. Functional adaptations of oxygen-transport proteins. *J. Exp. Biol.* **201**, 1085–1098 (1998).

108. Ordway, G. A. & Garry, D. J. Myoglobin: An essential hemoprotein in striated muscle. *J. Exp. Biol.* **207**, 3441–3446 (2004).

109. Gros, G., Wittenberg, B. A. & Jue, T. Myoglobin's old and new clothes: From molecular structure to function in living cells. *J. Exp. Biol.* **213**, 2713–2725 (2010).

110. Drago, R. S. Free radical reactions of transition metal systems. *Coord. Chem. Rev.* **32**, 97–110 (1980).

111. Jensen, K. P. & Ryde, U. How O2 binds to heme. Reasons for rapid binding and spin inversion.

J. Biol. Chem. **279**, 14561–14569 (2004).

112. Matsui, T., Ozaki, S. I., Liong, E., Phillips, G. N. & Watanabe, Y. Effects of the location of distal histidine in the reaction of myoglobin with hydrogen peroxide. *J. Biol. Chem.* **274**, 2838–2844 (1999).

113. Matsui, T., Ozaki, S. I. & Watanabe, Y. On the formation and reactivity of compound I of the His-64 myoglobin mutants. *J. Biol. Chem.* **272**, 32735–32738 (1997).

114. Adamson, J. W. & Finch, C. A. HEMOGLOBIN FUNCTION, OXYGEN AFFINITY, AND ERYTHROPOIETINl. *Annu. Rev. Physiol.* **37**, 351–369 (1975).

115. Ruiz-Larrea, M. B. A simple question to think about when considering the hemoglobin function. *Biochem. Mol. Biol. Educ.* **30**, 235–238 (2002).

116. Lisi, G. P. & Loria, J. P. Solution NMR Spectroscopy for the Study of Enzyme Allostery. *Chemical Reviews* **116**, 6323–6369 (2016).

117. Perutz, M. Mechanisms Regulating The Reactions Of Human Hemoglobin With Oxygen And Carbon Monoxide. *Annu. Rev. Physiol.* **52**, 1–25 (1990).

118. Perutz, M. F., Wilkinson, A. J., Paoli, M. & Dodson, G. G. The stereochemical mechanism of the cooperative effects in hemoglobin revisited. *Annual Review of Biophysics and Biomolecular Structure* **27**, 1–34 (1998).

119. Perutz, M. F. Stereochemistry of cooperative effects in haemoglobin: Haem-Haem interaction and the problem of allostery. *Nature* **228**, 726–734 (1970).

120. Weiss, J. N. The Hill equation revisited: Uses and misuses. *FASEB Journal* **11**, 835–841 (1997).

121. Stefan, M. I. & Le Novère, N. Cooperative Binding. *PLoS Comput. Biol.* **9**, (2013).

122. Hub, J. S., Kubitzki, M. B. & de Groot, B. L. Spontaneous quaternary and tertiary T-R transitions of human hemoglobin in molecular dynamics simulation. *PLoS Comput. Biol.* **6**, 1–11 (2010).

123. Ojima, I., Lin, S., Bernacki, R. J. & Pera, P. *) Ojima, I. Ojima, I. Acc. Chem. Res* **59**, (JAI Press, 1994).

124. Monod, J., Wyman, J. & Changeux, J. P. On the nature of allosteric transitions: A plausible

model. *J. Mol. Biol.* **12**, 88–118 (1965).

125. Koshland Jr, D. E., Nemethy, G. & Filmer, D. *Comparison of experimental binding data and theoretical models in proteins containing subunits*. *Biochemistry* **5**, (1966).

126. Shoji, O. & Watanabe, Y. Peroxygenase reactions catalyzed by cytochromes P450 Topical Issue in honor of Ivano Bertini. Guest editors: Lucia Banci, Claudio Luchinat. *J. Biol. Inorg. Chem.* **19**, 529–539 (2014).

127. Battistuzzi, G., Bellei, M., Bortolotti, C. A. & Sola, M. Redox properties of heme peroxidases. *Archives of Biochemistry and Biophysics* **500**, 21–36 (2010).

128. Baglia, R. A., Zaragoza, J. P. T. & Goldberg, D. P. Biomimetic Reactivity of Oxygen-Derived Manganese and Iron Porphyrinoid Complexes. *Chemical Reviews* **117**, 13320–13352 (2017).

129. Zámocký, M., Furtmüller, P. G. & Obinger, C. Evolution of structure and function of Class I peroxidases. *Archives of Biochemistry and Biophysics* **500**, 45–57 (2010).

130. Zámocký, M. *et al.* Independent evolution of four heme peroxidase superfamilies. *Arch. Biochem. Biophys.* **574**, 108–119 (2015).

131. Calabrese, E. J. & Canada, A. T. Catalase: Its role in xenobiotic detoxification. *Pharmacol. Ther.* **44**, 297–307 (1989).

132. Takaka, K. & Michael, A. K. D. Structure of Beef Liver Catalase. *J. Mol. Rid* **46**, 5–499 (1981).

133. Putnam, C. D., Arvai, A. S., Bourne, Y. & Tainer, J. A. Active and Inhibited Human Catalase Structures: Ligand and NADPH Binding and Catalytic Mechanism. *J. Mol. Biol.* **296**, 295–309 (2000).

134. LOEW, O. A NEW ENZYME OF GENERAL OCCURRENCE IN ORGANISMIS. *Science (80-.).* **11**, 701–702 (1900).

135. Chelikani, P., Fita, I. & Loewen, P. C. Diversity of structures and properties among catalases. *Cellular and Molecular Life Sciences* **61**, 192–208 (2004).

136. Dolphin, D., Forman, A., Borg, D. C., Fajer, J. & Felton, R. H. Compounds I of catalase and horse radish peroxidase: pi-cation radicals. *Proc. Natl. Acad. Sci. U. S. A.* **68**, 614–618 (1971).

137. Rydberg, P., Sigfridsson, E. & Ryde, U. On the role of the axial ligand in heme proteins: A

theoretical study. *J. Biol. Inorg. Chem.* **9**, 203–223 (2004).

138. Ivancich, A., Jouve, H. M., Sartor, B. & Gaillard, J. EPR investigation of-compound I in Proteus mirabilis and bovine liver catalases: Formation of porphyrin and tyrosyl radical intermediates. *Biochemistry* **36**, 9356–9364 (1997).

139. Altschul, A. M., Abrams, R. & Hogness, T. R. Soluble cytochrome c oxidase. *J. Biol. Chem.* **130**, 427–428 (1939).

140. Poulos, T. L. *et al.* The crystal structure of cytochrome c peroxidase. *J. Biol. Chem.* **255**, 575–580 (1980).

141. Volkov, A. N., Nicholls, P. & Worrall, J. A. R. The complex of cytochrome c and cytochrome c peroxidase: The end of the road? *Biochim. Biophys. Acta - Bioenerg.* **1807**, 1482–1503 (2011).

142. Banci, L., Bertini, I., Rosato, A. & Varani, G. Mitochondrial cytochromes c: A comparative analysis. *J. Biol. Inorg. Chem.* **4**, 824–837 (1999).

143. Liu, J. *et al.* Metalloproteins containing cytochrome, iron-sulfur, or copper redox centers. *Chemical Reviews* **114**, 4366–4369 (2014).

144. Lopes, H., Pettigrew, G. W., Moura, I. & Moura, J. J. G. Electrochemical study on cytochrome c peroxidase from Paracoccus denitrificans: A shifting pattern of structural and thermodynamic properties as the enzyme is activated. *J. Biol. Inorg. Chem.* **3**, 632–642 (1998).

145. Karlsen, O. A., Larsen & Jensen, H. B. Identification of a bacterial di-haem cytochrome c peroxidase from Methylomicrobium album BG8. *Microbiology* **156**, 2682–2690 (2010).

146. Becker, C. F., Watmough, N. J. & Elliot, S. J. Electrochemical evidence for multiple peroxidatic heme states of the diheme cytochrome c peroxidase of Pseudomonas aeruginosa. *Biochemistry* **48**, 87–95 (2009).

147. Fülöp, V., Ridout, C. J., Greenwood, C. & Hajdu, J. Crystal structure of the di-haem cytochrome c peroxidase from Pseudomonas aeruginosa. *Structure* **3**, 1225–1233 (1995).

148. Kelly, G. J. & Latzko, E. Soluble ascorbate peroxidase - Detection in plants and use in vitamin C estimation. *Naturwissenschaften* **66**, 617–618 (1979).

149. Groden, D. & Beck, E. H2O2 destruction by ascorbate-dependent systems from chloroplasts.

BBA - Bioenerg. **546**, 426–435 (1979).

150. Patterson, W. R. & Poulos, T. L. Crystal Structure of Recombinant Pea Cytosolic Ascorbate Peroxidase. *Biochemistry* **34**, 4331–4341 (1995).

151. Jones, D. K., Dalton, D. A., Rosell, F. I. & Raven, E. L. Class I heme peroxidases: Characterization of soybean ascorbate peroxidase. *Arch. Biochem. Biophys.* **360**, 173–178 (1998).

152. Raven, E. L. Understanding functional diversity and substrate specificity in haem peroxidases: What can we learn from ascorbate peroxidase? *Natural Product Reports* **20**, 367–381 (2003).

153. Patterson, W. R., Poulos, T. L. & Goodin, D. B. Identification of a Porphyrin π Cation Radical in Ascorbate Peroxidase Compound I. *Biochemistry* **34**, 4342–4345 (1995).

154. Girvan, H. M. & Munro, A. W. Applications of microbial cytochrome P450 enzymes in biotechnology and synthetic biology. *Current Opinion in Chemical Biology* **31**, 136–145 (2016).

155. Guengerich, F. P. & Yoshimoto, F. K. Formation and Cleavage of C-C Bonds by Enzymatic Oxidation-Reduction Reactions. *Chemical Reviews* **118**, 6573–6655 (2018).

156. Wei, Y., Ang, E. L. & Zhao, H. Recent developments in the application of P450 based biocatalysts. *Current Opinion in Chemical Biology* **43**, 1–7 (2018).

157. Acevedo-Rocha, C. G. *et al.* P450-Catalyzed regio- and diastereoselective steroid hydroxylation: Efficient directed evolution enabled by mutability landscaping. *ACS Catal.* **8**, 3395–3410 (2018).

158. Ortiz De Montellano, P. R. Hydrocarbon hydroxylation by cytochrome P450 enzymes. *Chem. Rev.* **110**, 932–948 (2010).

159. Denisov, I. G., Makris, T. M., Sligar, S. G. & Schlichting, I. Structure and chemistry of cytochrome P450. *Chemical Reviews* **105**, 2253–2277 (2005).

160. Perera, R. *et al.* Neutral thiol as a proximal ligand to ferrous heme iron: Implications for heme proteins that lose cysteine thiolate ligation on reduction. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 3641–3646 (2003).

161. Vatsis, K. P., Peng, H. M. & Coon, M. J. Replacement of active-site cysteine-436 by serine

converts cytochrome P450 2B4 into an NADPH oxidase with negligible monooxygenase activity. *J. Inorg. Biochem.* **91**, 542–553 (2002).

162. Onderko, E. L., Silakov, A., Yosca, T. H. & Green, M. T. Characterization of a selenocysteine-ligated P450 compound I reveals direct link between electron donation and reactivity. *Nat. Chem.* **9**, 623–628 (2017).

163. Sevrioukova, I. F. *et al. Structure of a cytochrome P450 –redox partner electron-transfer complex (cytochrome P450 reductase, f lavodoxin protein–protein interaction)*. *Biochemistry* **96**, (1999).

164. Munro, A. W., Girvan, H. M. & McLean, K. J. Cytochrome P450-redox partner fusion enzymes. *Biochimica et Biophysica Acta - General Subjects* **1770**, 345–359 (2007).

165. Miura, Y. & Fulco, A. J. (ω - 2) Hydroxylation of fatty acids by a soluble system from Bacillus megaterium. *J. Biol. Chem.* **249**, 1880–1888 (1974).

166. Miura, Y. & Fulco, A. J. ω-1, ω-2 and ω-3 Hydroxylation of long-chain fatty acids, amides and alcohols by a soluble enzyme system from Bacillus megatyerium. *Biochim. Biophys. Acta (BBA)/Lipids Lipid Metab.* **388**, 305–317 (1975).

167. Ruettinger, R. T. & Fulco, A. J. Epoxidation of unsaturated fatty acids by a soluble cytochrome P-450-dependent system from Bacillus megaterium. *J. Biol. Chem.* **256**, 5728–5734 (1981).

168. Ravichandran, K. G., Boddupalli, S. S., Hasemann, C. A., Peterson, J. A. & Deisenhofer, J. Crystal structure of hemoprotein domain of P450BM-3, a prototype for microsomal P450's. *Science (80-.).* **261**, 731–736 (1993).

169. Munro, A. W. *et al. P450 BM3: The very model of a modern flavocytochrome*. *Trends in Biochemical Sciences* **27**, (2002).

170. Poulos, T. L. Cytochrome P450 P450eryF. *Curr. Biol.* **5**, 767–774 (1995).

171. Li, H. & Poulos, T. L. The structure of the cytochrome p450BM-3 haem domain complexed with the fatty acid substrate, palmitoleic acid. *Nat. Struct. Biol.* **4**, 140–146 (1997).

172. Geronimo, I., Denning, C. A., Heidary, D. K., Glazer, E. C. & Payne, C. M. Molecular Determinants of Substrate Affinity and Enzyme Activity of a Cytochrome P450BM3 Variant. *Biophys. J.* **115**, 1251–1263 (2018).

173. Warman, A. J. *et al. Flavocytochrome P450 BM3: An update on structure and mechanism of a biotechnologically important enzyme*. *Biochemical Society Transactions* **33**, (2005).

174. Munro, A. W. *et al.* P450 BM3: The very model of a modern flavocytochrome. *Trends in Biochemical Sciences* **27**, 250–257 (2002).

175. Modi, S., Sutcliffe, M. J., Primrose, W. U., Lian, L. Y. & Roberts, G. C. K. The catalytic mechanism of cytochrome P450 BM3 involves a 6 Å movement of the bound substrate on reduction. *Nat. Struct. Biol.* **3**, 414–417 (1996).

176. Noble, M. A. *et al.* Roles of key active-site residues in flavocytochrome P450 BM3. *Biochem. J.* **339**, 371–379 (1999).

177. Chen, Z., Ost, T. W. B. & Schelvis, J. P. M. Phe393 Mutants of Cytochrome P450 BM3 with Modified Heme Redox Potentials Have Altered Heme Vinyl and Propionate Conformations. *Biochemistry* **43**, 1798–1808 (2004).

178. Ost, T. W. B. *et al.* Phenylalanine 393 exerts thermodynamic control over the heme of flavocytochrome P450 BM3. *Biochemistry* **40**, 13421–13429 (2001).

179. Ost, T. W. B. *et al.* Structural and spectroscopic analysis of the F393H mutant of flavocytochrome P450 BM3. *Biochemistry* **40**, 13430–13438 (2001).

180. Yeom, H., Sligar, S. G., Li, H., Poulos, T. L. & Fulco, A. J. The Role of Thr268 in Oxygen Activation of Cytochrome 450 BM-3. *Biochemistry* **34**, 14733–14740 (1995).

181. Steitz, J. A. *et al. Crystal Structure of the Cytochrome P-450CAM Active Site Mutant Thr252Ala*. *Prog. Nucleic Acid Res. Mol. Biol* **30**, (1991).

182. Roberts, A. G. *et al.* The role of cytochrome P450 BM3 phenylalanine-87 and threonine-268 in binding organic hydroperoxides. *Biochim. Biophys. Acta - Gen. Subj.* **1860**, 669–677 (2016).

183. Almhjell, P. J., Boville, C. E. & Arnold, F. H. Engineering enzymes for noncanonical amino acid synthesis. *Chemical Society Reviews* **47**, 8980–8997 (2018).

184. Renata, H., Wang, Z. J. & Arnold, F. H. Expanding the enzyme universe: Accessing non-natural reactions by mechanism-guided directed evolution. *Angewandte Chemie - International Edition* **54**, 3351–3367 (2015).

185. Ling, M. M. & Robinson, B. H. Approaches to DNA mutagenesis: An overview. *Analytical Biochemistry* **254**, 157–178 (1997).

186. Reetz, M. T. & Carballeira, J. D. Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. *Nat. Protoc.* **2**, 891–903 (2007).

187. Arnold, F. H. The nature of chemical innovation: New enzymes by evolution. *Quarterly Reviews of Biophysics* **48**, 404–410 (2015).

188. Packer, M. S. & Liu, D. R. Methods for the directed evolution of proteins. *Nature Reviews Genetics* **16**, 379–394 (2015).

189. Mansuy, D. *et al.* Dichlorocarbene Complexes of Iron(II)-Porphyrins–Crystal and Molecular Structure of Fe(TPP)(CCl2)(H2O). *Angew. Chemie Int. Ed. English* **17**, 781–782 (1978).

190. Coelho, P. S., Brustad, E. M., Kannan, A. & Arnold, F. H. Olefin cyclopropanation via carbene transfer catalyzed by engineered cytochrome P450 enzymes. *Science (80-.).* **339**, 307–310 (2013).

191. Coelho, P. S. *et al.* A serine-substituted P450 catalyzes highly efficient carbene transfer to olefins in vivo. *Nat. Chem. Biol.* **9**, 485–487 (2013).

192. Heel, T., McIntosh, J. A., Dodani, S. C., Meyerowitz, J. T. & Arnold, F. H. Non-natural olefin cyclopropanation catalyzed by diverse cytochrome P450s and other hemoproteins. *ChemBioChem* **15**, 2556–2562 (2014).

193. Wang, Z. J. *et al.* Improved cyclopropanation activity of histidine-ligated cytochromeP450 enables the enantioselective formal synthesis of levomilnacipran. *Angew. Chemie - Int. Ed.* **53**, 6810–6813 (2014).

194. Renata, H., Wang, Z. J., Kitto, R. Z. & Arnold, F. H. P450-catalyzed asymmetric cyclopropanation of electron-deficient olefins under aerobic conditions. *Catal. Sci. Technol.* **4**, 3640–3643 (2014).

195. Bordeaux, M., Tyagi, V. & Fasan, R. Highly diastereoselective and enantioselective olefin cyclopropanation using engineered myoglobin-based catalysts. *Angew. Chemie - Int. Ed.* **54**, 1744–1748 (2015).

196. Tinoco, A., Steck, V., Tyagi, V. & Fasan, R. Highly Diastereo- and Enantioselective Synthesis of

Trifluoromethyl-Substituted Cyclopropanes via Myoglobin-Catalyzed Transfer of Trifluoromethylcarbene. *J. Am. Chem. Soc.* **139**, 5293–5296 (2017).

197. Chandgude, A. L. & Fasan, R. Highly Diastereo- and Enantioselective Synthesis of Nitrile-Substituted Cyclopropanes by Myoglobin-Mediated Carbene Transfer Catalysis. *Angew. Chemie - Int. Ed.* **57**, 15852–15856 (2018).

198. Villarino, L. *et al.* An Artificial Heme Enzyme for Cyclopropanation Reactions. *Angew. Chemie - Int. Ed.* **57**, 7785–7789 (2018).

199. Bajaj, P., Sreenilayam, G., Tyagi, V. & Fasan, R. Gram-Scale Synthesis of Chiral Cyclopropane-Containing Drugs and Drug Precursors with Engineered Myoglobin Catalysts Featuring Complementary Stereoselectivity. *Angew. Chemie - Int. Ed.* **55**, 16110–16114 (2016).

200. Knight, A. M. *et al.* Diverse Engineered Heme Proteins Enable Stereodivergent Cyclopropanation of Unactivated Alkenes. *ACS Cent. Sci.* **4**, 372–377 (2018).

201. Bos, J. & Roelfes, G. Artificial metalloenzymes for enantioselective catalysis. *Current Opinion in Chemical Biology* **19**, 135–143 (2014).

202. Hayashi, T. *et al.* Co(ii)/Co(i) reduction-induced axial histidine-flipping in myoglobin reconstituted with a cobalt tetradehydrocorrin as a methionine synthase model. *Chem. Commun.* **50**, 12560–12563 (2014).

203. Oohora, K., Kihira, Y., Mizohata, E., Inoue, T. & Hayashi, T. C(sp3)-H bond hydroxylation catalyzed by myoglobin reconstituted with manganese porphycene. *J. Am. Chem. Soc.* **135**, 17282–17285 (2013).

204. Morita, Y. *et al.* Intraprotein transmethylation via a CH3-Co(iii) species in myoglobin reconstituted with a cobalt corrinoid complex. *Dalt. Trans.* **45**, 3277–3284 (2016).

205. Matsuo, T. *et al.* Preparation and O2 binding study of myoglobin having a cobalt porphycene. *Inorg. Chem.* **44**, 9391–9396 (2005).

206. Hayashi, T. *et al.* Crystal structure and peroxidase activity of myoglobin reconstituted with iron porphycene. *Inorg. Chem.* **45**, 10530–10536 (2006).

207. Cui, X., Xu, X., Wojtas, L., Kim, M. M. & Zhang, X. P. Regioselective synthesis of multisubstituted furans via metalloradical cyclization of alkynes with α-diazocarbonyls:

Construction of functionalized α-oligofurans. *J. Am. Chem. Soc.* **134**, 19981–19984 (2012).

208. Olshansky, L. *et al.* Artificial Metalloproteins Containing Co4O4 Cubane Active Sites. *J. Am. Chem. Soc.* **140**, 2739–2742 (2018).

209. 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).

210. Key, H. M. *et al.* Beyond Iron: Iridium-Containing P450 Enzymes for Selective Cyclopropanations of Structurally Diverse Alkenes. *ACS Cent. Sci.* **3**, 302–308 (2017).

211. Wolf, M. W., Vargas, D. A. & Lehnert, N. Engineering of RuMb: Toward a Green Catalyst for Carbene Insertion Reactions. *Inorg. Chem.* **56**, 5623–5635 (2017).

212. Sreenilayam, G., Moore, E. J., Steck, V. & Fasan, R. Metal Substitution Modulates the Reactivity and Extends the Reaction Scope of Myoglobin Carbene Transfer Catalysts. *Adv. Synth. Catal.* **359**, 2076–2089 (2017).

213. Sreenilayam, G., Moore, E. J., Steck, V. & Fasan, R. Stereoselective Olefin Cyclopropanation under Aerobic Conditions with an Artificial Enzyme Incorporating an Iron-Chlorin e6 Cofactor. *ACS Catal.* **7**, 7629–7633 (2017).

214. 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 (2017).

215. 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).

216. Tyagi, V., Bonn, R. B. & Fasan, R. Intermolecular carbene S-H insertion catalysed by engineered myoglobin-based catalysts. *Chem. Sci.* **6**, 2488–2494 (2015).

217. Kan, S. B. J., Lewis, R. D., Chen, K. & Arnold, F. H. Directed evolution of cytochrome c for carbon-silicon bond formation: Bringing silicon to life. *Science (80-.).* **354**, 1048–1051 (2016).

218. Lewis, R. D. Evolution and characterization of carbene transferases for cyclopropanation and carbon-silicon bond formation. (California institute of technology, 2019).

219. Lewis, R. D. *et al.* Catalytic iron-carbene intermediate revealed in a cytochrome c carbene

transferase . *Proc. Natl. Acad. Sci.* **115**, 7308–7313 (2018).

220. Jennifer Kan, S. B., Huang, X., Gumulya, Y., Chen, K. & Arnold, F. H. Genetically programmed chiral organoborane synthesis. *Nature* **552**, 132–136 (2017).

221. Huang, X. *et al.* A Biocatalytic Platform for Synthesis of Chiral α-Trifluoromethylated Organoborons. *ACS Cent. Sci.* **5**, 270–276 (2019).

222. Alonso-Cotchico, L. & Roelfes, G. A 'broad Spectrum' Carbene Transferase for Synthesis of Chiral α-Trifluoromethylated Organoborons. *ACS Cent. Sci.* **5**, 206–208 (2019).

223. Doyle, M. P. *et al.* Electronic and Steric Control in Carbon-Hydrogen Insertion Reactions of Diazoacetoacetates Catalyzed by Dirhodium(II) Carboxylates and Carboxamides. *J. Am. Chem. Soc.* **115**, 958–964 (1993).

224. Dydio, P. *et al.* An artificial metalloenzyme with the kinetics of native enzymes. *Science (80-.).* **354**, 102–106 (2016).

225. Vargas, D. A., Tinoco, A., Tyagi, V. & Fasan, R. Myoglobin-Catalyzed C–H Functionalization of Unprotected Indoles. *Angew. Chemie - Int. Ed.* **57**, 9911–9915 (2018).

226. Zhang, R. K. *et al.* Enzymatic assembly of carbon–carbon bonds via iron-catalysed sp 3 C–H functionalization. *Nature* **565**, 67–72 (2019).

227. Brandenberg, O. F., Chen, K. & Arnold, F. H. Directed Evolution of a Cytochrome P450 Carbene Transferase for Selective Functionalization of Cyclic Compounds. *J. Am. Chem. Soc.* **141**, 8989–8995 (2019).

228. Zhang, J., Huang, X., Zhang, R. K. & Arnold, F. H. Enantiodivergent α-Amino C-H Fluoroalkylation Catalyzed by Engineered Cytochrome P450s. *J. Am. Chem. Soc.* **141**, 9798–9802 (2019).

229. Tyagi, V. & Fasan, R. Myoglobin-catalyzed olefination of aldehydes. *Angew. Chemie - Int. Ed.* **55**, 2512–2516 (2016).

230. 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).

231. Lichtenstein, B. R. *et al.* Engineering oxidoreductases: maquette proteins designed from

scratch. *Biochem. Soc. Trans.* **40**, 561–566 (2012).

232. Grayson, K. J. & Anderson, J. R. The ascent of man(made oxidoreductases). *Current Opinion in Structural Biology* **51**, 149–155 (2018).

233. Watkins, D. W., Armstrong, C. T. & Anderson, J. L. R. De novo protein components for oxidoreductase assembly and biological integration. *Current Opinion in Chemical Biology* **19**, 90–98 (2014).

234. Watkins, D. W. *et al.* A suite of de novo c-type cytochromes for functional oxidoreductase engineering. *Biochim. Biophys. Acta - Bioenerg.* **1857**, 493–502 (2016).

235. Watkins, D. W. *et al.* Construction and in vivo assembly of a catalytically proficient and hyperthermostable de novo enzyme. *Nat. Commun.* **8**, (2017).

236. Gibney, B. R., Rabanal, F., Skalicky, J. J., Wand, A. J. & Dutton, P. L. Iterative protein redesign. *J. Am. Chem. Soc.* **121**, 4952–4960 (1999).

237. Huang, S. S., Koder, R. L., Lewis, M., Wand, A. J. & Dutton, P. L. The HP-1 maquette: from an apoprotein structure to a structured hemoprotein designed to promote redox-coupled proton exchange. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 5536–5541 (2004).

238. Gibney, B. R. *et al.* Self-assembly of heme A and heme B in a designed four-helix bundle: implications for a cytochrome c oxidase maquette. *Biochemistry* **39**, 11041–11049 (2000).

239. Sharp, R. E., Moser, C. C., Rabanal, F. & Dutton, P. L. Design, synthesis, and characterization of a photoactivatable flavocytochrome molecular maquette. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 10465–10470 (1998).

240. Gibney, B. R., Mulholland, S. E., Rabanal, F. & Dutton, P. L. Ferredoxin and ferredoxin-heme maquettes. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 15041–15046 (1996).

241. Farid, T. A. *et al.* Elementary tetrahelical protein design for diverse oxidoreductase functions. *Nat. Chem. Biol.* **9**, 826–833 (2013).

242. Grzyb, J. *et al.* Empirical and computational design of iron-sulfur cluster proteins. *Biochim. Biophys. Acta* **1817**, 1256–1262 (2012).

243. Norris, J. R. *et al.) White, W. I. In The Porphyrins*. *J. J. Proc. Natl. Acad. Sci. U.S.A* **116**,

(Academic Press, 1994).

244. Shifman, J. M., Gibney, B. R., Sharp, R. E. & Dutton, P. L. Heme redox potential control in de novo designed four-alpha-helix bundle proteins. *Biochemistry* **39**, 14813–14821 (2000).

245. Osuna, S., Jiménez-Osés, G., Noey, E. L. & Houk, K. N. Molecular dynamics explorations of active site structure in designed and evolved enzymes. *Acc. Chem. Res.* **48**, 1080–1089 (2015).

246. Khoury, G. A., Smadbeck, J., Kieslich, C. A. & Floudas, C. A. Protein folding and de novo protein design for biotechnological applications. *Trends in Biotechnology* **32**, 99–109 (2014).

247. Hilvert, D. Design of Protein Catalysts. *Annu. Rev. Biochem.* **82**, 447–470 (2013).

248. Huang, P. S. *et al.* De novo design of a four-fold symmetric TIM-barrel protein with atomic-level accuracy. *Nat. Chem. Biol.* **12**, 29–34 (2016).

249. Obexer, R., Pott, M., Zeymer, C., Griffiths, A. D. & Hilvert, D. Efficient laboratory evolution of computationally designed enzymes with low starting activities using fluorescence-activated droplet sorting. *Protein Eng. Des. Sel.* **29**, 355–366 (2016).

250. Jacobs, T. M. *et al.* Design of structurally distinct proteins using strategies inspired by evolution. *Science (80-.).* **352**, 687–690 (2016).

251. 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).

252. Nanda, V. *et al.* Structural principles for computational and de novo design of 4Fe–4S metalloproteins. *Biochim. Biophys. Acta - Bioenerg.* 1–8 (2015). doi:10.1016/j.bbabio.2015.10.001

253. Lichtenstein, B. R. *et al.* Engineering oxidoreductases: maquette proteins designed from scratch. *Biochem. Soc. Trans.* **40**, 561–566 (2012).

254. Sellés Vidal, L., Kelly, C. L., Mordaka, P. M. & Heap, J. T. Review of NAD(P)H-dependent oxidoreductases: Properties, engineering and application. *Biochimica et Biophysica Acta - Proteins and Proteomics* **1866**, 327–347 (2018).

255. Esser, L., Zhou, F., Yu, C.-A. & Xia, D. Crystal structure of bacterial cytochrome bc 1 in

complex with azoxystrobin reveals a conformational switch of the Rieske iron–sulfur protein subunit . *J. Biol. Chem.* **294**, 12007–12019 (2019).

256. Crofts, A. R. The Cytochrome bc1 Complex: Function in the Context of Structure. *Annu. Rev. Physiol.* **66**, 689–733 (2004).

257. Xia, D. *et al.* Structural analysis of cytochrome bc1 complexes: Implications to the mechanism of function. *Biochimica et Biophysica Acta - Bioenergetics* **1827**, 1278–1294 (2013).

258. DeGrado, W. F. & Ho, S. P. Design of a 4-Helix Bundle Protein: Synthesis of Peptides Which Self-Associate into a Helical Protein. *J. Am. Chem. Soc.* **109**, 6751–6758 (1987).

259. Regan, L. & Degrado, W. F. Characterization of a helical protein designed from first principles. *Science (80-.).* **241**, 976–978 (1988).

260. Robertson, D. E. *et al.* Design and synthesis of multi-haem proteins. *Nature* **368**, 425–432 (1994).

261. Rabanal, F., DeGrado, W. & Dutton, P. Toward the synthesis of a photosynthetic reaction center maquette: A cofacial porphyrin pair assembled between two subunits of a synthetic four-helix bundle multiheme protein. *J. Am. Chem. Soc.* **118**, 473–474 (1996).

262. Gibney, B. R., Rabanal, F., Skalicky, J. J., Wand, A. J. & Dutton, P. L. Design of a unique protein scaffold for maquettes. *J. Am. Chem. Soc.* **119**, 2323–2324 (1997).

263. Skalicky, J. J. *et al.* Sequence-specific resonance assignments for a designed four-α-helix bundle protein. *J. Biomol. NMR* **11**, 227–228 (1998).

264. Walsh, S. T. R., Cheng, H., Bryson, J. W., Roder, H. & Degrado, W. F. Solution structure and dynamics of a de novo designed three-helix bundle protein. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 5486–5491 (1999).

265. Skalicky, J. J. *et al.* Solution structure of a designed four--helix bundle maquette scaffold. *J. Am. Chem. Soc.* **121**, 4941–4951 (1999).

266. Huang, S. S., Gibney, B. R., Stayrook, S. E., Dutton, P. L. & Lewis, M. X-ray structure of a maquette scaffold. *J. Mol. Biol.* **326**, 1219–1225 (2003).

267. Huang, S. S., Koder, R. L., Lewis, M., Wand, A. J. & Dutton, P. L. The HP-1 maquette: From an

apoprotein structure to a structured hemoprotein designed to promote redox-coupled proton exchange. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 5536–5541 (2004).

268. Koder, R. L. *et al.* Design and engineering of an O2 transport protein. *Nature* **458**, 305–309 (2009).

269. Farid, T. A. *et al.* Elementary tetrahelical protein design for diverse oxidoreductase functions. *Nat. Chem. Biol.* **9**, 826–833 (2013).

270. Sanders, C., Turkarslan, S., Lee, D. W. & Daldal, F. Cytochrome c biogenesis: The Ccm system. *Trends in Microbiology* **18**, 266–274 (2010).

271. Allen, J. W. A., Leach, N. & Ferguson, S. J. The histidine of the c-type cytochrome CXXCH haem-binding motif is essential for haem attachment by the Escherichia coli cytochrome c maturation (Ccm) apparatus. *Biochem. J.* **389**, 587–592 (2005).

272. Watkins, D. W., Armstrong, C. T. & Anderson, J. L. R. De novo protein components for oxidoreductase assembly and biological integration. *Current Opinion in Chemical Biology* **19**, 90–98 (2014).

273. Braun, M. & Thöny-Meyer, L. Biosynthesis of artificial microperoxidases by exploiting the secretion and cytochrome c maturation apparatuses of Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 12830–12835 (2004).

274. Müller, H. E. ABTS peroxidase medium as a highly sensitive plate assay for detection of hydrogen peroxide production in bacteria. *J. Microbiol. Methods* **2**, 101–102 (1984).

275. Mansuy, D., Nastainczyk, W. & Ullrich, V. The mechanism of halothane binding to microsomal cytochrome P450. *Naunyn. Schmiedebergs. Arch. Pharmacol.* **285**, 315–324 (1974).

276. Wolf, C. R., Mansuy, D., Nastainczyk, W., Deutschmann, G. & Ullrich, V. The reduction of polyhalogenated methanes by liver microsomal cytochrome P450. *Mol. Pharmacol.* **13**, 698–705 (1977).

277. Mansuy, D., Lange, M. & Chottard, J. C. Reaction of 2,2-Bis(p-chlorophenyl)-1,1,1 - trichloroethane (DDT) with Iron (II) Porphyrins. Isolation of the Vinylidene Carbene Complex, Tetraphenylporphy riniron (II) (C═C(p-Cl-C6H4)2. *Journal of the American Chemical Society* **100**, 3213–3214 (1978).

278. Mansuy, D., Battioni, J. P. & Chottard, J. C. Preparation of a Porphyrin-Iron-Carbene Model for the Cytochrome P 450 Complexes Obtained upon Metabolic Oxidation of the Insecticide Synergists of the 1,3-Benzodioxole Series. *J. Am. Chem. Soc.* **101**, 3971–3973 (1979).

279. Groves, J. T., Nemo, T. E. & Myers, R. S. Hydroxylation and Epoxidation Catalyzed by Iron-Porphine Complexes. Oxygen Transfer from Iodosylbenzene. *J. Am. Chem. Soc.* **4**, 1032–1033 (1979).

280. Groves, J. T. & Nemo, T. E. Epoxidation Reactions Catalyzed by Iron Porphyrins. Oxygen Transfer from Iodosylbenzene. *J. Am. Chem. Soc.* **105**, 5786–5791 (1983).

281. Maux, P. Le, Srour, H. F. & Simonneaux, G. Enantioselective water-soluble iron-porphyrin-catalyzed epoxidation with aqueous hydrogen peroxide and hydroxylation with iodobenzene diacetate. *Tetrahedron* **68**, 5824–5828 (2012).

282. Viski, P. & Groves, J. T. Asymmetric Hydroxylation by a Chiral Iron Porphyrin. *J. Am. Chem. Soc.* **111**, 8537–8538 (1989).

283. Denisov, I. G., Makris, T. M., Sligar, S. G. & Schlichting, I. Structure and chemistry of cytochrome P450. *Chemical Reviews* **105**, 2253–2277 (2005).

284. Ortiz De Montellano, P. R. Hydrocarbon hydroxylation by cytochrome P450 enzymes. *Chem. Rev.* **110**, 932–948 (2010).

285. Ren, X. *et al.* Drug Oxidation by Cytochrome P450BM3: Metabolite Synthesis and Discovering New P450 Reaction Types. *Chem. - A Eur. J.* **21**, 15039–15047 (2015).

286. Wolf, J. R., Kodadek, T., Hamaker, C. G., Djukic, J. P. & Woo, L. K. Shape and Stereoselective Cyclopropanation of Alkenes Catalyzed by Iron Porphyrins. *J. Am. Chem. Soc.* **117**, 9194–9199 (1995).

287. Jeschek, M. *et al.* Directed evolution of artificial metalloenzymes for in vivo metathesis. *Nature* **537**, 661–665 (2016).

288. Brandenberg, O. F., Fasan, R. & Arnold, F. H. Exploiting and engineering hemoproteins for abiological carbene and nitrene transfer reactions. *Current Opinion in Biotechnology* **47**, 102–111 (2017).

289. Gober, J. G. & Brustad, E. M. Non-natural carbenoid and nitrenoid insertion reactions

catalyzed by heme proteins. *Current Opinion in Chemical Biology* **35**, 124–132 (2016).

290. 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).

291. 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).

292. Coelho, P. S. *et al.* Axial-Ligand Influence on P450 Reduction Potentials: Implications for Catalysis. *Nat. Chem. Biol* 485–487 (2013).

293. Kennedy, M. L. *et al.* Model hemoprotein reduction potentials: The effects of histidine-to-iron coordination equilibrium [18]. *J. Am. Chem. Soc.* **123**, 4635–4636 (2001).

294. 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).

295. Gober, J. G., Rydeen, A. E., Schwochert, T. D., Gibson-O'Grady, E. J. & Brustad, E. M. Enhancing cytochrome P450-mediated non-natural cyclopropanation by mutation of a conserved second-shell residue. *Biotechnol. Bioeng.* **115**, 1416–1426 (2018).

296. Reynolds, E. W. *et al.* An Evolved Orthogonal Enzyme/Cofactor Pair. *J. Am. Chem. Soc.* **138**, 12451–12458 (2016).

297. Gober, J. G. *et al.* P450-Mediated Non-natural Cyclopropanation of Dehydroalanine-Containing Thiopeptides. *ACS Chem. Biol.* **12**, 1726–1731 (2017).

298. Lewis, J. C., Coelho, P. S. & Arnold, F. H. Enzymatic functionalization of carbon-hydrogen bonds. *Chemical Society Reviews* **40**, 2003–2021 (2011).

299. Zheng, W., Shan, N., Yu, L. & Wang, X. UV-visible, fluorescence and EPR properties of porphyrins and metalloporphyrins. *Dye. Pigment.* **77**, 153–157 (2008).

300. Gouterman, M. Spectra of Porphyrins. *J. Mol. Spectrosc.* **6**, 138–163 (1961).

301. Dąbrowski, J. M. *et al.* Engineering of relevant photodynamic processes through structural modifications of metallotetrapyrrolic photosensitizers. *Coordination Chemistry Reviews* **325**,

67–101 (2016).

302. Kleingardner, E. C., Asher, W. B. & Bren, K. L. Efficient and flexible preparation of biosynthetic microperoxidases. *Biochemistry* **56**, 143–148 (2016).

303. Aron, J. *et al.* Microperoxidase Sequence. *J. Inorg. Biochem.* **27**, 1–17 (2001).

304. Hayashi, T. *et al.* Capture and characterization of a reactive haem–carbenoid complex in an artificial metalloenzyme. *Nat. Catal.* **1**, 578–584 (2018).

305. Tinoco, A. *et al.* Origin of High Stereocontrol in Olefin Cyclopropanation Catalyzed by an Engineered Carbene Transferase. *ACS Catal.* **9**, 1514–1524 (2019).

306. Talele, T. T. The 'cyclopropyl Fragment' is a Versatile Player that Frequently Appears in Preclinical/Clinical Drug Molecules. *Journal of Medicinal Chemistry* **59**, 8712–8756 (2016).

307. Seubold, F. H. Carbon-carbon bond dissociation energies in the cycloalkanes. *The Journal of Chemical Physics* **22**, 945–946 (1954).

308. Tian, Z., Lis, L. & Kass, S. R. Carbon-hydrogen bond dissociation energies: The curious case of cyclopropene. *J. Org. Chem.* **78**, 12650–12653 (2013).

309. Tian, Z., Fattahi, A., Lis, L. & Kass, S. R. Cycloalkane and cycloalkene C-H bond dissociation energies. *J. Am. Chem. Soc.* **128**, 17087–17092 (2006).

310. Gritti, F. & Guiochon, G. Optimization of the thermal environment of columns packed with very fine particles. *J. Chromatogr. A* **1216**, 1353–1362 (2009).

311. Kaplitz, A. S. *et al.* High-Throughput and Ultrafast Liquid Chromatography. *Analytical Chemistry* 67–84 (2020). doi:10.1021/acs.analchem.9b04713

312. Žuvela, P. *et al.* Column Characterization and Selection Systems in Reversed-Phase High-Performance Liquid Chromatography. *Chemical Reviews* **119**, 3674–3729 (2019).

313. Torrent-Sucarrat, M., Arrastia, I., Arrieta, A. & Cossío, F. P. Stereoselectivity, different oxidation states, and multiple spin states in the cyclopropanation of olefins catalyzed by fe-porphyrin complexes. *ACS Catal.* **8**, 11140–11153 (2018).

314. Dueholm, B. *et al.* Evolution of substrate recognition sites (SRSs) in cytochromes P450 from Apiaceae exemplified by the CYP71AJ subfamily. *BMC Evol. Biol.* **15**, (2015).

315. Spath, E. THE ALKALOIDS*. *Annu. Rev. Biochem.* 513–534 (1937).

316. Jarowicki, K. & Kocienski, P. Protecting groups. *Journal of the Chemical Society. Perkin Transactions 1* **1**, 2109–2135 (2001).

317. Isidro-Llobet, A., Álvarez, M. & Albericio, F. Amino acid-protecting groups. *Chem. Rev.* **109**, 2455–2504 (2009).

318. Shenvi, R. A., O'Malley, D. P. & Baran, P. S. Chemoselectivity: The mother of invention in total synthesis. *Acc. Chem. Res.* **42**, 530–541 (2009).

319. Afagh, N. A. & Yudin, A. K. Chemoselectivity and the curious reactivity preferences of functional groups. *Angewandte Chemie - International Edition* **49**, 262–310 (2010).

320. Lairson, L. L., Henrissat, B., Davies, G. J. & Withers, S. G. Glycosyltransferases: Structures, Functions, and Mechanisms. *Annu. Rev. Biochem.* **77**, 521–555 (2008).

321. Liang, D. M. *et al.* Glycosyltransferases: Mechanisms and applications in natural product development. *Chemical Society Reviews* **44**, 8350–8374 (2015).

322. Dalziel, M., Crispin, M., Scanlan, C. N., Zitzmann, N. & Dwek, R. A. Emerging principles for the therapeutic exploitation of glycosylation. *Science* **343**, (2014).

323. Peralta, D. *et al.* The separation of xylene isomers by ZIF-8: A demonstration of the extraordinary flexibility of the ZIF-8 framework. *Microporous Mesoporous Mater.* **173**, 1–5 (2013).

324. Eller, K., Henkes, E., Rossbacher, R. & Höke, H. Amines, Aliphatic. *Ullmann's Encyclopedia of Industrial Chemistry* (2000). doi:doi:10.1002/14356007.a02_001

325. Ziegler, J. & Facchini, P. J. Alkaloid Biosynthesis: Metabolism and Trafficking. *Annu. Rev. Plant Biol.* **59**, 735–769 (2008).

326. Debnath, B. *et al.* Role of plant alkaloids on human health: A review of biological activities. *Mater. Today Chem.* **9**, 56–72 (2018).

327. Amirkia, V. & Heinrich, M. Alkaloids as drug leads - A predictive structural and biodiversity-based analysis. *Phytochem. Lett.* **10**, xlviii–53 (2014).

328. Kochanowska-Karamyan, A. J. & Hamann, M. T. Marine indole alkaloids: Potential new drug

leads for the control of depression and anxiety. *Chem. Rev.* **110**, 4489–4497 (2010).

329. Ehrenworth, A. M. & Peralta-Yahya, P. Accelerating the semisynthesis of alkaloid-based drugs through metabolic engineering. *Nature Chemical Biology* **13**, 249–258 (2017).

330. French, C. E. *et al.* Biological production of semisynthetic opiates using genetically engineered bacteria. *Bio/Technology* **13**, 674–676 (1995).

331. Hoskins, W. M. & Crout, D. H. G. Pyrrolizidine alkaloid analogues. Preparation of semisynthetic esters of retronecine. *J. Chem. Soc. Perkin Trans. 1* **35**, 538–544 (1977).

332. Gelbaum, L. T., Gordon, M. M., Miles, M. & Zalkow, L. H. Semisynthetic Pyrrolizidine Alkaloid Antitumor Agents. *J. Org. Chem.* **47**, 2501–2504 (1982).

333. Petroski, R. J., Yates, S. G., Weisleder, D. & Powell, R. G. ISolation, semi-synthesis, and nmr spectral studies of loline alkaloids. *J. Nat. Prod.* **52**, 810–817 (1989).

334. Clarke, A. K. & Unsworth, W. P. A happy medium: The synthesis of medicinally important medium-sized rings: Via ring expansion. *Chemical Science* **11**, 2876–2881 (2020).

335. Kantorowski, E. J. & Kurth, M. J. Expansion to seven-membered rings. *Tetrahedron* **56**, 4317–4353 (2000).

336. Buchner, E. & Curtius, T. Synthese von Ketonsäureäthern aus Aldehyden und Diazoessigäther. *Berichte der Dtsch. Chem. Gesellschaft* **18**, 2371–2377 (1885).

337. Lebel, H., Marcoux, J. F., Molinaro, C. & Charette, A. B. Stereoselective cyclopropanation reactions. *Chem. Rev.* **103**, 977–1050 (2003).

338. Ciamician, G. L. & Dennstedt, M. Ueber die Einwirkung des Chloroforms auf die Kaliumverbindung Pyrrols. *Berichte der Dtsch. Chem. Gesellschaft* **14**, 1153–1163 (1881).

339. Alexander, R. E., Herrick, A. B. & Roder, T. M. The Formation of 3-Substituted Pyridines from Pyrrole. *J. Am. Chem. Soc.* **72**, 2760–2761 (1950).

340. Parham, W. E. & Wright, C. D. Formation of Naphthalenes from Indenes. IV. The Effect of Substitution at the Ethylenic Double Bond. *J. Org. Chem.* **22**, 1473–1477 (1957).

341. Henke, E., Bornscheuer, U. T., Schmid, R. D. & Pleiss, J. A molecular mechanism of enantiorecognition of tertiary alcohols by carboxylesterases. *ChemBioChem* **4**, 485–493

(2003).

342. 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).

343. Kurasov, O. *et al.* NAD Biosynthesis: Identification of the Tryptophan to Quinolinate Pathway in Bacteria NAD Biosynthesis: Identification of the Tryptophan to Quinolinate Pathway in Bacteria. *Chem. Biol.* **10**, 1195–1204 (2003).

344. Wang, X. *et al.* Engineering Escherichia coli nicotinic acid mononucleotide adenylyltransferase for fully active amidated NAD biosynthesis. *Appl. Environ. Microbiol.* **83**, (2017).

345. Bouvet, O., Bourdelier, E., Glodt, J., Clermont, O. & Denamur, E. Diversity of the auxotrophic requirements in natural isolates of Escherichia coli. *Microbiol. (United Kingdom)* **163**, 891–899 (2017).

346. Sheldon, R. A. & Woodley, J. M. Role of Biocatalysis in Sustainable Chemistry. *Chem. Rev.* **118**, 801–838 (2018).

347. Bornscheuer, U. T. The fourth wave of biocatalysis is approaching. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **376**, (2018).

348. Huisman, G. W. & Collier, S. J. On the development of new biocatalytic processes for practical pharmaceutical synthesis. *Current Opinion in Chemical Biology* **17**, 284–292 (2013).

349. Truppo, M. D. Biocatalysis in the Pharmaceutical Industry: The Need for Speed. *ACS Med. Chem. Lett.* **8**, 476–480 (2017).

350. Brenna, E., Fuganti, C., Gatti, F. G. & Serra, S. Biocatalytic methods for the synthesis of enantioenriched odor active compounds. *Chemical Reviews* **111**, 4036–4072 (2011).

351. Patel, R. N. Biocatalysis for synthesis of pharmaceuticals. *Bioorganic and Medicinal Chemistry* **26**, 1252–1274 (2018).

352. Vieille, C. & Zeikus, G. J. Hyperthermophilic Enzymes: Sources, Uses, and Molecular Mechanisms for Thermostability. *Microbiol. Mol. Biol. Rev.* **65**, 1–43 (2001).

353. Modarres, H. P., Mofrad, M. R. & Sanati-Nezhad, A. Protein thermostability engineering. *RSC Advances* **6**, 115252–115270 (2016).

354. Salihu, A. & Alam, M. Z. Solvent tolerant lipases: A review. *Process Biochemistry* **50**, 86–96 (2015).

355. Gupta, A. & Khare, S. K. Enzymes from solvent-tolerant microbes: Useful biocatalysts for non-aqueous enzymology. *Critical Reviews in Biotechnology* **29**, 44–54 (2009).

356. Doukyu, N. & Ogino, H. Organic solvent-tolerant enzymes. *Biochemical Engineering Journal* **48**, 270–282 (2010).

357. Sellek, G. A. & Chaudhuri, J. B. Biocatalysis in organic media using enzymes from extremophiles. *Enzyme and Microbial Technology* **25**, 471–482 (1999).

358. Serdakowski, A. L. & Dordick, J. S. Enzyme activation for organic solvents made easy. *Trends in Biotechnology* **26**, 48–54 (2008).

359. Zaks, A. & Klibanov, A. M. *Enzyme-catalyzed processes in organic solvents*. *Biochemistry* **82**, (1985).

360. Klibanov, A. M. Improving enzymes by using them in organic solvents. *Nature* **409**, 241–246 (2001).

361. Klibanov, A. Why are enzymes less active in organic solvents than in water? *Trends Biotechnol.* **15**, 97–101 (1997).

362. Klibanov, A. M. Enzymatic catalysis in anhydrous organic solvents. *Trends in Biochemical Sciences* **14**, 141–144 (1989).

363. Hudson, E. P., Eppler, R. K. & Clark, D. S. Biocatalysis in semi-aqueous and nearly anhydrous conditions. *Current Opinion in Biotechnology* **16**, 637–643 (2005).

364. Gupta, R., Gupta, N. & Rathi, P. Bacterial lipases: An overview of production, purification and biochemical properties. *Applied Microbiology and Biotechnology* **64**, 763–781 (2004).

365. MOZHAEV, V. V. *et al.* Catalytic activity and denaturation of enzymes in water/organic cosolvent mixtures: α-Chymotrypsin and laccase in mixed water/alcohol, water/glycol and water/formamide solvents. *Eur. J. Biochem.* **184**, 597–602 (1989).

366. Yoon, J. H. & Mckenzie, D. A comparison of the activities of three β-galactosidases in aqueous-organic solvent mixtures. *Enzyme Microb. Technol.* **36**, 439–446 (2005).

367. Zaks, A. & Klibanov, A. M. Enzymatic catalysis in nonaqueous solvents. *J. Biol. Chem.* **263**, 3194–3201 (1988).

368. Cannon, W. R. & Benkovic, S. J. Solvation, Reorganization Energy, and Biological Catalysis* Do Unusually Strong Interactions Exist in ES ‡ ? Downloaded from. *J. Biol. Chem.* **273**, 26257–26260 (1998).

369. Marchetti, L. & Levine, M. Biomimetic catalysis. *ACS Catalysis* **1**, 1090–1118 (2011).

370. Stepankova, V. *et al.* Strategies for stabilization of enzymes in organic solvents. *ACS Catalysis* **3**, 2823–2836 (2013).

371. Sheldon, R. A. & Woodley, J. M. Role of Biocatalysis in Sustainable Chemistry. *Chem. Rev.* **118**, 801–838 (2018).

372. Sheldon, R. A. & van Pelt, S. Enzyme immobilisation in biocatalysis: Why, what and how. *Chem. Soc. Rev.* **42**, 6223–6235 (2013).

373. Svensson, I., Adlercreutz, P. & Mattiasson, B. Interesterification of phosphatidylcholine with lipases in organic media. *Appl. Microbiol. Biotechnol.* **33**, 255–258 (1990).

374. Xu, J. H., Kato, Y. & Asano, Y. Efficient preparation of (R)-α-monobenzoyl glycerol by lipase catalyzed asymmetric esterification: Optimization and operation in packed bed reactor. *Biotechnol. Bioeng.* **73**, 493–499 (2001).

375. Morgan, B., Dodds, D. R., Zaks, A., Andrews, D. R. & Klesse, R. Enzymatic Desymmetrization of Prochiral 2-Substituted-1,3-propanediols: A Practical Chemoenzymatic Synthesis of a Key Precursor of SCH51048, a Broad-Spectrum Orally Active Antifungal Agent. *J. Org. Chem.* **62**, 7736–7743 (1997).

376. Wang, P., Dai, S., Waezsada, S. D., Tsao, A. Y. & Davison, B. H. Enzyme stabilization by covalent binding in nanoporous sol-gel glass for nonaqueous biocatalysis. *Biotechnol. Bioeng.* **74**, 249–255 (2001).

377. Ogino, H. & Ishikawa, H. Enzymes which are stable in the presence of organic solvents. *Journal of Bioscience and Bioengineering* **91**, 109–116 (2001).

378. Wachtmeister, J. & Rother, D. Recent advances in whole cell biocatalysis techniques bridging from investigative to industrial scale. *Current Opinion in Biotechnology* **42**, 169–177 (2016).

379. Temiño, D. M. R. De, Hartmeier, W. & Ansorge-Schumacher, M. B. Entrapment of the alcohol dehydrogenase from Lactobacillus kefir in polyvinyl alcohol for the synthesis of chiral hydrophobic alcohols in organic solvents. *Enzyme Microb. Technol.* **36**, 3–9 (2005).

380. Ishizuka, F. *et al.* Polymeric Nanocapsules for Enzyme Stabilization in Organic Solvents. *Macromolecules* **51**, 438–446 (2018).

381. Blandino, A., Macías, M. & Cantero, D. Immobilization of glucose oxidase within calcium alginate gel capsules. *Process Biochem.* **36**, 601–606 (2001).

382. Blandino, A., Macías, M. & Cantero, D. Formation of calcium alginate gel capsules: Influence of sodium alginate and CaCl2 concentration on gelation kinetics. *J. Biosci. Bioeng.* **88**, 686–689 (1999).

383. Puguan, J. M. C., Yu, X. & Kim, H. Diffusion characteristics of different molecular weight solutes in Ca-alginate gel beads. *Colloids Surfaces A Physicochem. Eng. Asp.* **469**, 158–165 (2015).

384. Amsden, B. Solute diffusion within hydrogels. Mechanisms and models. *Macromolecules* **31**, 8382–8395 (1998).

385. Benguigui, L., Busnel, J. P. & Durand, D. Study of junction zones in gelatin gels through selective enzymatic digestion. *Polymer (Guildf).* **32**, 2680–2685 (1991).

386. Mel'Nichenko, Y. B., Gomza, Y. P., Shilov, V. V & Osipov, S. I. On the nature of junction zones in gelatin gels. *Polym. Int.* **25**, 153–157 (1991).

387. Ha, J., Engler, C. R. & Lee, S. J. Determination of diffusion coefficients and diffusion characteristics for chlorferon and diethylthiophosphate in Ca-alginate gel beads. *Biotechnol. Bioeng.* **100**, 698–706 (2008).

388. Øyaas, J., Storrø, I., Svendsen, H. & Levine, D. W. The effective diffusion coefficient and the distribution constant for small molecules in calcium-alginate gel beads. *Biotechnol. Bioeng.* **47**, 492–500 (1995).

389. Muhr, A. H. & Blanshard, J. M. V. Diffusion in gels. *Polymer* **23**, 1012–1026 (1982).

390. Gierszal, K. P. *et al.* π-hydrogen bonding in liquid water. *J. Phys. Chem. Lett.* **2**, 2930–2933 (2011).

391. Stephenson, W. K. & Fuchs, R. Enthalpies of interaction of nitrogen base solutes with organic solvents. *Can. J. Chem.* **63**, 2540–2544 (1985).

392. Erkkilä, A., Majander, I. & Pihko, P. M. Iminium catalysis. *Chem. Rev.* **107**, 5416–5470 (2007).

393. McIntosh, J. A., Farwell, C. C. & Arnold, F. H. Expanding P450 catalytic reaction space through evolution and engineering. *Current Opinion in Chemical Biology* **19**, 126–134 (2014).