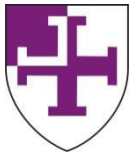
NMR SPECTROSCOPY STUDIES OF ENZYME FUNCTION AND INHIBITION

A thesis submitted to the Board of the Faculty of Physical Sciences
at the University of Oxford in partial fulfilment of the requirements
for the degree of Doctor of Philosophy

Amjad Khan
St Cross College, University of Oxford
Hilary Term 2016



ST CROSS COLLEGE
UNIVERSITY OF OXFORD

NMR spectroscopy studies of enzyme function and inhibition

Amjad Khan *DPhil Thesis*

St Cross College, University of Oxford *Hilary Term 2016*

Abstract

The work described in this Thesis focuses on the application of NMR spectroscopy methods in understanding the function and inhibition of two protein systems; these are γ-butyrobetaine hydroxylase (BBOX) and the bacterial potassium ion efflux (Kef) system. BBOX belongs to the super family of enzymes called the 2-oxoglutarate (2OG) and Fe^{II} dependent oxygenase and is involved in the biosynthesis of L-carnitine in humans and other prokaryotes. BBOX is a current drug target for the treatment of myocardial infarction. Kef is a ligand-gated system that protects bacteria from toxic electrophilic species. Kef is inhibited by the binding of cytoplasmic glutathione (GSH) to KTN (K^+ transport and nucleotide) binding domains and activated by glutathione-S-conjugates (GS-X). Since bacterial Kef activation during electrophilic exposure is a critical determinant of their survival, perturbation of Kef activity is potentially a novel target for the development of antibiotic drugs.

^{1}H NMR direct ligand-observation was employed to study the binding interaction of the natural substrate γ-butyrobetaine (GBB) and co-substrate 2OG with BBOX. A ^{1}H NMR-based dual-reporter ligand displacement method was developed to assess the nature of inhibitor binding to BBOX i.e to determine whether an inhibitor competes with GBB or 2OG or both. The method was exemplified with a set of isoquinoline-based inhibitors; the results reveal `cystallographically unexpected ´structure-activity relationship with some inhibitors competing 2OG only and some competing both 2OG and GBB. Using ^{1}H NMR spectroscopy, a simple and efficient BBOX inhibition assay was developed for inhibitor IC_{50} measurement. Similarly, ^{1}H NMR-based assays were applied to demonstrate that the cation-π interaction between the substrates and aromatic cage residues of BBOX play a critical role in BBOX substrate recognition.

^{1}H NMR spectroscopy was applied to show that in the absence of a 2OG oxygenase, ascorbate in the assay mixture is slowly degraded by the dissolved oxygen to yield H_2O_2 which simultaneously leads to 2OG breakdown into succinate. It is proposed that in the assays of 2OG oxygenases, the apparent increase in the level of ˜uncoupled˝ 2OG turnover with ascorbate over time could possibly be due to the artifacts of the ascorbate induced-2OG breakdown instead of being due to enzyme catalysis. Other reducing agents were also found to oxidise identically by the dissolved oxygen as ascorbate in the mixture and result in 2OG breakdown.

In the Kef system, ^{1}H NMR direct ligand-observation was applied to investigate the influence of each functional group of the Kef activating ligand glutathione-S-*N*-tertiary butylsuccinimide on its binding interaction (K_D 0.4 μM) with Kef-QCTD (Q-linker carboxy terminal domain; a KTN domain) from *Shewanella denitrificans* (*sd*) with the aim of developing novel non-peptidic ligands (antibacterial agents) of Kef. In addition, ^{19}F NMR was employed to develop an efficient ligand-observed binding assay for Kef that was used for ligand screening as well as measuring their binding dissociation constant value from a single NMR spectrum. Finally, a ^{1}H NMR technique was applied to confirm that the electron density found in the nucleotide binding pocket in the crystal structure of *apo-sd*Kef-QCTD is unambiguously an AMP molecule that is naturally bound to the protein and has a role in stabilising the dimeric form of KTN domains (Kef proteins).

Table of Contents

Acknowledgments

First of all, I would like to take this opportunity to sincerely thank my supervisor Prof. Tim Claridge for his support and guidance throughout in the last four years, especially for giving me the freedom to explore the applications of NMR spectroscopy in enzymology and for the opportunity to participate in a variety of different projects. I find Tim Claridge a very nice, helpful and cooperative person.

I am also very grateful to my collaborator Prof. Chris Schofield for his guidance and support, particularly for introducing me into the BBOX cation-π interaction project, for allowing me to use his laboratory resources and proof reading this thesis.

I would also like to thank Dr. Ivanhoe Leung for introducing me into the BBOX ligand binding studies and ascorbate/2OG projects. I must say that I received a huge amount of support and help almost on a day to day basis in the early stages of my research.

Thanks also go to my collaborators Dr Sam Grayer, Dr Anthony Chan and Prof. Stuart Conway for their collaboration in the Kef structure-activity relationship studies, slow exchange ^{19}F NMR binding assay and AMP projects. I also want to thank my collaborators Prof. Jasmin Mecinovic and Jos Kamps for their contribution in the BBOX cation-π interaction project. I would like to thank Dr Anna Rydzik for providing me the isoquinoline-ligands, Dr Jurgen Brem for producing BBOX and Robert Lesniak for performing the BBOX fluorescence assays.

I also want to thank Dr Barbara Odell, Tina Jackson and Maria Marshall for their kind help and support in the NMR lab.

Sincere thanks also go to Prof. Ihsan Ali, Vice Chancellor, Abdul Wali Khan University Mardan, Pakistan and Prof. Saeed Anwar for helping me in the scholarship to pursue my D.Phil studies at Oxford.

Finally, I would like to thank my parents, wife Ruqia Begum and my brother Mohammad Iqbal for their continuous love, support and encouragement throughout my studies.

Abbreviations and acronyms

2,4-D	2,4-dichlorophenoxyacetic acid
2,4-PDCA	2,4-pyridinedicarboxylic acid
2-HG	2-hydroxyglutarate
2OG	2-oxoglutarate
5Cac	5-carboxycytosine
5fc	5-formylcytosine
5-HT	5-hydroxytryptamine
5mc	5-methylcytosine
8-HQ	8-hydroxyquinoline
ACC	1-aminocyclopropane-1-carboxylic acid
ACCO	1-aminocyclopropane-1-carboxylic acid oxidase
ADH	alcohol dehydrogenase
ADH9	aldehyde dehydrogenase 9
ADP	adenosine diphosphate
AMBA	3-amino-4-(methylamino)butanoic acid
AMP	adenosine monophosphate
ANS	anthocyanadin synthase
BH_4	tetrahydrobiopterin
BLAST	basic local alignment search tool
BSA	bovine serum albumin
DSBH	double stranded β-helix
c.p.s	cycles per second
CarC	carbapenem synthase
CAS	clavaminate synthase
CHELPG	charges from electrostatic potentials using a grid-based method
CNBD	cyclic nucleotide binding domain
Co-A	co-enzyme A
CODD	carboxy terminal oxygen-dependent degradation domain
CPMG	Carr-Purcell-Meiboom-Gill
CPT1	carnitine palmitoyl transferase 1
CSA	chemical shift anisotropy

CT1	cardiotrophin-1
DAOCS	deacetoxycephalosporin C synthase
DCPIP	2,6-dichlorophenolindophenol
DHA	dehydroascorbic acid
DHQ	dihydroquercetin
DKG	diketogulonic acid
DMF	dimethylformamide
DPN	diphosphopyridine nucleotide
DSF	differential scanning fluorimetry
DTNB	5,5′-dithiobis-2-nitrobenzoic acid
DTT	dithiothreitol
DβM	dopamine-β-monooxygenase
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
EFE	ethylene forming enzyme
ePHOGSY	enhanced protein hydration observed through gradient spectroscopy
ESG	*N*-ethylsuccinimido-S-glutathione
FAXS	fluorine chemical shift anisotropy and exchange for screening
FIH	factor inhibiting HIF
FKBP	FK6 binding proteins
FMN	flavin mononucleotide
FtmOx1	fumitremorgin B endoperoxidase
FTO	fat mass and obesity associated protein
GBB	γ-butyrobetaine
GBBH	γ-butyrobetaine hydroxylase
H6H	hyoscyamine 6 β-hydroxylase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIF	hypoxia inducible factor
HIV	human immunodeficiency virus
hPDI	human protein disulphide isomerase
HPLC	high performance liquid chromatography
HPP	4-hydroxyphenylpyruvic acid
HPPD	4-hydroxyphenylpyruvic acid dioxygenase

HSA	human serum albumin
HTML	3-hydroxy-N^{ϵ}- trimethyllysine
HTMLA	3-hydroxy-N^{ϵ}- trimethyllysine aldolase
Hz	hertz
IC_{50}	half maximal inhibitory concentration
IHD	ischemic heart diseases
IPAA	isopropylidine-ascorbic acid
IPNS	isopenicillin N synthase
ITC	isothermal titration calorimetry
JMJD	jumonji domain containing protein
K_A	association constant
K_{cat}	catalytic constant
K_D	dissociation constant
KDM	histone (lysine) demethylase
K_I	inhibition constant
K_M	Michaelis constant
k_{off}	dissociation rate constant
k_{on}	association rate constant
KTN	K^+ transport and nucleotide (binding domain)
LFCA	long chain fatty acid
LS-MS	liquid chromatography-mass spectrometry
Mc	mega cycles
Me	methyl
MES	2-(*N*-morpholino)ethanesulfonic acid
MOPS	3-(*N*-morpholino)propanesulfonic acid
MW	molecular weight
NADH	nicotinamide adenine dinucleotide (reduced)
NAD^+	nicotinamide adenine dinucleotide (oxidised)
NODD	N-terminal oxygen-dependent degradation domain
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser effect spectroscopy
NOG	*N*-oxalyl glycine
hOCTN2	human organic cation/carnitine transporter 2

PAH	phenylalanine hydroxylase
PCA	protocatechuic acid
PDB	protein data bank
PHD2	prolyl hydroxylase domain containing enzyme isoform 2
PHDs	prolyl hydroxylase domains
PIPES	piperazine-*N*,*N*′-bis(2-ethanesulfonic acid)
POCA	2-[5-(4-chlorophenyl)pentyl]oxirane-2-carboxylate
ppm	parts per million
PRE	paramagnetic relaxation enhancement
ps	*pseudomonas*
QCTD	Q-linker carboxy-terminal domain
R_2	transverse relaxation rates
RCK	regulator of conductance of K^+
ROS	reactive oxygen species
SAM	*S*-adenosyl-L-methionine
SAR	structure-activity relationship
sd	*Shewanella denitrificans*
SHMT	serine hydroxymethyl transferase
SIIS	solvent induced isotopic shift
SLG	*S*-lactoyal glutathione
SPLASTIC	spin label attached to a protein side chain to identify interacting compounds
STD	saturation transfer difference
T_1	longitudinal relaxation time constant
T_2	transverse relaxation time constant
TCA	tricarboxylic acid
TCEP	tris(2-carboxyethyl)phosphine
TDG	thymine DNA glycosylase
TEAB	tetraethyl ammonium bromide
TEMPO	(2,2,6,6-tetramethylpiperidin-1-yl)oxyl
TES	2-[tris(hydroxymethyl)-methylamino]-ethanesulfonic acid
TET	ten-eleven translocation
TFA	1,1,1 trifluoroacetic acid
T_h	temperature of hydrophobic exposure

THDH	4,5,5,6-tetrahydroxy-2,3-diktohexanoic acid
THP	3-(2,2,2-trimethylhydraziniumyl)propionate
T_m	melting temperature
TMABADH	4-*N*-trimethylaminobutyraldehyde dehydrogenase
TML	N^{ε}-trimethyllysine
TMO	N^{ε}-trimethyl ornithine
TOCSY	total correlation spectroscopy
Tricine	*N*-(tri(hydroxymethyl)methyl)glycine
Tris	tris(hydroxymethyl)aminomethane
tr-NOE	transferred-NOE
t_c	correlation time
V_{max}	maximum velocity
WaterLOGSY	water-ligand observed *via* gradient spectroscopy

List of Figures

1 Introduction

1.1 Historical Applications of Nuclear Magnetic Resonance Spectroscopy (NMR) in Enzymology

Nuclear magnetic resonance (NMR) spectroscopy is a well-established technique, which has long been used to study biochemical systems, the history of which goes back to early 1950s. In the earlier studies, NMR spectroscopy had been mostly applied for assignment and characterization of biologically important small molecules and products of enzyme catalysed reactions.

In 1956, for the first time, ^{1}H NMR spectroscopy was applied in a biological system to investigate the effect of pH on the absorption frequencies of protons of amino acids glycine, alanine, β-alanine, cysteine, proline, hydroxyproline, and peptides glycyl-glycine, and alanyl-alanine on a 40.1 Mc (megacycles) instrument.[1] In this study it was observed that the absorption frequency of protons of the amino and carboxyl groups, which are readily exchanged with water protons, depends highly on the acidity of the medium, while protons directly attached to the carbon are much less affected. In a subsequent work, a systematic study on the effect of pH on the absorption frequencies of all the biologically most important amino acids was then conducted.[2]

In 1957, the first ^{1}H NMR spectrum of protein pancreatic ribonuclease (Figure 1.1) was also published.[3] In this spectrum four broad peaks were observed for the protein. Based on the chemical shifts, peaks I and IV were tentatively assigned to aromatic hydrogens and to hydrogens on aliphatic carbon atoms attached only to other aliphatic carbons.

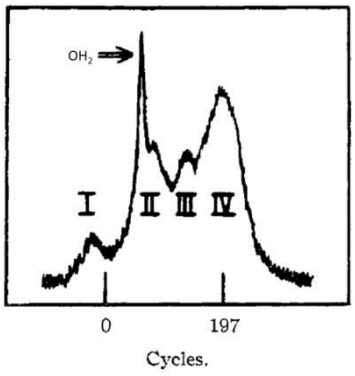


Figure 1.1 The first ^{1}H NMR spectrum of a protein (ribonuclease) published in 1957. (Reproduced from reference [3]).

In a subsequent study by a different group[4] the spectrum was reanalysed. In this study, peak II was assigned to the 123 α-CH protons of all amino acids, the 30 β-protons of serine the 10 β-protons of threonine which were expected to occur at + 90 to + 110 c.p.s (cycles per second).[2] The 6 glycine protons and 8 protons of the CH_2 groups in arginine which were expected to occur at higher fields of + 117 to + 120 c.p.s[2] were also assigned to be in a region in between the second and third peaks. Peak III (+ 135 to + 145 c.p.s) was assigned to arise from CH_2 protons from 16 cysteine, 8 methionine, 32 aspartic acid, 20 lysine (ε-CH_2) 10 proline, 8 histidine, 6 phenylalanine and 12 tyrosine. Peak IV was assigned to be a combination of two peaks; one possibly between + 180 to + 190 c.p.s and the other between + 210 to + 220 c.p.s. The former peak was assigned to the aliphatic CH groups of leucine (2 protons), isoleucine (3) and valine (9) and CH_2 groups of leucine (4 hydrogens), isoleucine (6), proline (20) lysine (60) and arginine (16) and the CH_3 group of methionine (12 protons). The latter peak was assigned to the CH_3 groups of alanine (36 hydrogens), valine (54), leucine (12), isoleucine (18) and threonine (30).

In 1957, for the first time the stereochemistry of the product of an enzyme catalysed reaction was established by ^{1}H NMR technique.[5] The reaction studied was the fumarase catalysed hydration of fumarate into L-malate. Based on the steric relation of the fumarate protons on carbons 2 and 3 in the L-malate, it was shown that the addition of water to the double bond of fumarate occurs in a *cis* manner.

In 1963, the binding interaction of a small molecule ligand, diphosphopyridine nucleotide (DPN) to the enzyme yeast alcohol dehydrogenase (ADH) was observed by ^{1}H NMR direct ligand observation.[6] In this assay, the ligand NMR resonances were broadened and attenuated after the addition of protein. This study showed that when a ligand binds to a protein, it behaves as if it were part of the protein and hence its NMR resonances become broad due to its faster transverse relaxation rates (R_2) in the bound form. (This technique is described further in section 1.4.1)

In 1968, ^{1}H NMR spectroscopy was used to study kinetics of an enzyme catalysed reaction. In this study, the hydrolysis of acetylcholine catalysed by horse serum enzyme cholinesterase was studied.[7] The rate of the reaction was followed from the decay of the acetate peak of acetylcholine and formation of sodium acetate signal. Both these peaks occurred at different absorption frequencies in the spectrum. The K_M for the substrate and the K_I values

determined for various inhibitors were found in good agreement to the values obtained by other biophysical techniques.

During the 1960-1970 period, significant developments occurred in the applications of NMR techniques; these include Overhauser effect (NOE) between nuclear spins,[8] pulse-fourier transform NMR,[9] and the introduction of two-dimensional (2D) NMR experiments.[10]

In 1972, the transferred NOE (tr-NOE) technique was applied to study the binding of lysine-vasopressin and peptide analogs to bovine neurophysin II. In this technique, the binding interaction of the ligands with the protein was observed in the form of decreasing in intensity of the ligand proton signals after irradiation of specific protein resonances. This is due to the intermolecular Overhauser effect which is transferred from the protein to the bound-ligand.[11-13] (The tr-NOE binding technique is discussed further in section 1.4.7)

In 1982, the strategy of sequential assignment of protein resonances for protein solution (by NMR spectroscopy) structure determination was proposed.[14] This technique was based on the experimental observation in 2D NMR experiments (COSY, NOESY) for complete through-bond and through-space proton-proton connectivity. This allowed for sequential resonance assignments of each individual amino acid in sequence and to characterize the 3D structure of the polypeptide chain by a combination of semi-quantitative, intramolecular distance constraints. Based on this technique, in 1985, the first solution structure of a small protein proteinase inhibitor IIA (57 amino acids with a molecular weight of ~6 kDa), was published.[15] In 1990, the strategy of three-dimensional ^{1}H,^{15}N,^{13}C triple resonance NMR spectroscopy techniques using ^{15}N and ^{13}C-labelled protein for protein structure determination emerged.[16] This technique was demonstrated on uniformly double labelled (^{15}N and ^{13}C) protein calmodulin (16.7 kDa). In this technique, the sequential connectivity was established from one-bond scalar (J) couplings. All these 3D NMR experiments involve ^{1}H detection to provide higher sensitivity.

In 1996, the technique of SAR (structure-activity relationship) by NMR emerged as the first protein-observed NMR technique for screening large chemical libraries of small molecules that bind to target proteins.[17] In this technique the binding interaction of ligands to the protein FKBP was demonstrated by changes in the chemical shift of protein resonances in 2D ^{1}H-^{15}N HSQC experiment after the addition of ligand (chemical shift perturbation). FKBP inhibits calcineurin (a serine-threonine phosphatase) and blocks T cell activation[18] when it is complexed to FK506. (The SAR by NMR is described in detail in section 1.3.1)

In 1999, the saturation transfer difference (STD) NMR spectroscopy emerged as a ligand-observed NMR technique for studying protein-ligand binding interactions.[19] In this technique, the detection of a binding interaction was based on the saturation transfer to ligand molecules in direct contact with a protein. This technique also allowed in identifying parts (hydrogens) of the ligand that remain in close contact with the protein during binding (ligand binding epitopes).[19,20] (The STD NMR technique is described further in section 1.4.8)

In 2000, the technique of water-ligand observation *via* gradient spectroscopy (waterLOGSY) NMR spectroscopy was developed as another ligand-observed NMR method for studying protein-ligand binding interactions.[21] This technique was based on the transfer of magnetization from bulk water to protein and then onto bound ligand. This technique is now widely used as a primary screening method in medicinal chemistry, especially as part of fragment-based drug discovery. (This technique is explained in section 1.4.9)

In the same 2000s period, NMR competition experiments utilizing either ^{1}H, [22-24] more details in section 1.4.2) or ^{19}F, [25-27] (described further in section 1.4.5) ^{31}P, [28] or ^{13}C [29] were introduced as another NMR ligand-based technique for studying protein-ligand binding interactions. In this technique, changes in the NMR parameter of an initially bound ligand after the addition of competitive ligand, demonstrates the binding of the competitive ligand to the protein. These experiments are now widely used for detecting site specific ligand binding interaction to the proteins.

1.2 NMR spectroscopy methods for studying protein-ligand binding

A precise knowledge of protein-ligand recognition events is the key for understanding the mechanism of the biological function of a protein. Understanding the binding interaction of a ligand with a protein at an atomic level under near physiological conditions has become possible by a combination of solution NMR spectroscopic techniques and x-ray crystallographic studies. The NMR assays are extensively applied in the fields of chemical biology, medicinal chemistry and drug discovery and provide a diverse range of information such as detection of protein-ligand binding interactions, binding kinetics, measurements of binding dissociation constants, protein 3D structure determination, protein conformational changes and dynamics, enzyme catalysis and inhibition studies.[30-39]

In general, the NMR experiments applied to study protein-ligand binding interactions can be grouped into two categories; protein-observed experiments and ligand-observed experiments

(Figure 1.2). The protein-observed methods involve monitoring changes in the NMR parameters of a protein after the addition of a ligand as a probe of the binding interaction, whereas, the ligand-observed methods involve monitoring changes in the NMR parameters of a ligand after the addition of protein. Theoretically, all NMR spectroscopic parameters may be used as a probe to detect the binding activity of a ligand to a protein, however in practice, the parameters that can be measured easily and with high sensitivity are of most significance. For example chemical shift changes, changes in transverse relaxation rates, changes of diffusion constants and changes of NOEs may serves as probe of binding interaction.[20]

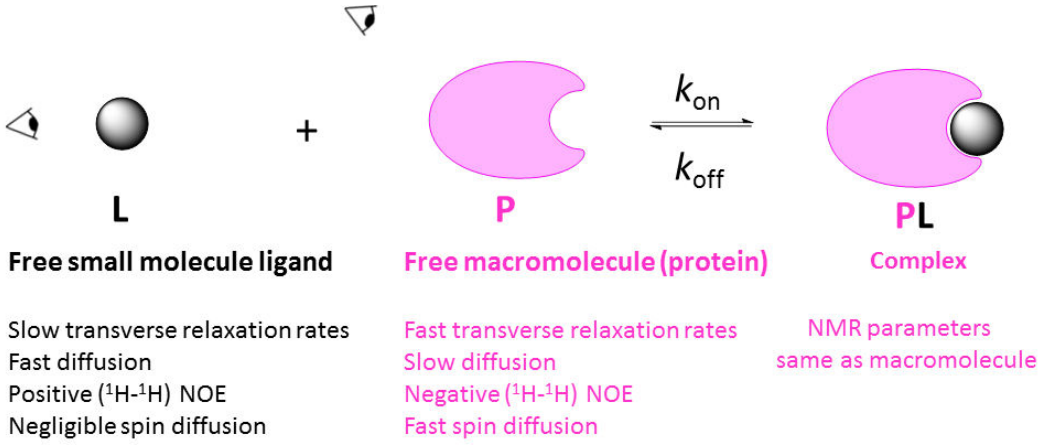


Figure 1.2 NMR methods for protein-ligand binding.
The NMR methods for studying the binding interaction of a ligand (L) with a protein (P) are of two types; protein-observed methods and ligand-observed methods. The ligand-observed methods follow a ligand NMR parameter for binding, whist; the protein-observed methods follow the NMR parameter of protein. Due to difference in molecular weight, the NMR properties of a small molecule ligand are different from those of a protein molecule or of a protein-bound small molecule ligand.

1.2.1 NMR chemical exchange in protein-ligand binding studies

The NMR chemical exchange system (which by definition refers to any process in which a nucleus exchanges between two or more chemical environments) in which its NMR parameters (chemical shift, relaxation rate, scalar coupling, dipolar coupling etc.) differ, also provides useful information about many aspects of protein-ligand binding interactions. The NMR chemical exchange regime may be considered as of three types namely, slow, intermediate (peak coalescence) and fast (Figure 1.3). In general, in a slow exchange system on NMR time scale (NMR time scale means to how fast an event happens relative to the NMR observables), the frequency of the interconversion of the two forms of a system is lower than the difference in their frequencies on chemical shift scale, hence the populations of the two forms appear as two separate resonances in the NMR spectrum. In an intermediate

or coalescence NMR exchange system, the frequency of the interconversion of the two forms of system is approximately equal to their frequency difference on chemical shift scale. Hence the separate lines for the two forms are no longer discernible. In a NMR fast exchange system the frequency of the interconversion of the two forms of a system is higher than their frequencies difference on chemical shift scale, hence for such a system a single resonance is observed in the spectrum whose chemical shift is the population weighted average of the chemical shifts of the two individual states.

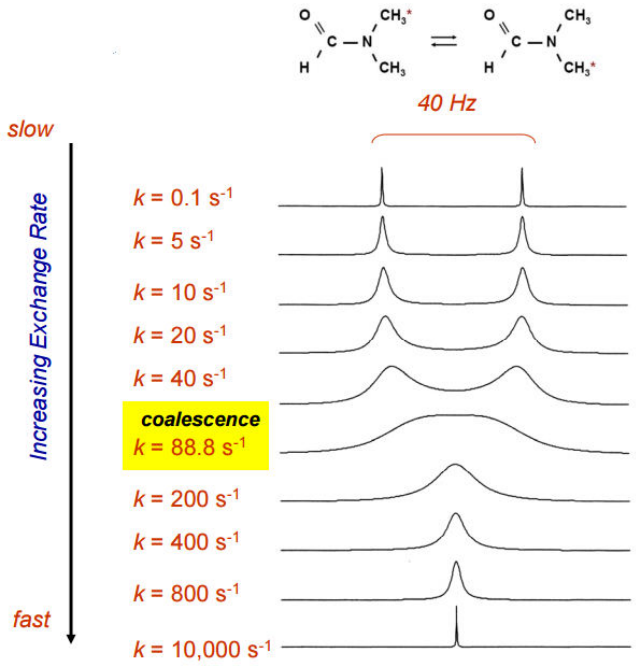


Figure 1.3 Illustration of NMR chemical exchange system.
The restricted rotation about the partial double bond results in the two forms of the dimethylformamide (DMF) as demonstrated by ^{1}H NMR.

The NMR chemical exchange plays a key role in studying protein-ligand binding interaction. For example, in the NMR ligand-observed binding techniques (Figure 1.2), it is the average signal of the ligand between the bound and free forms that gives the binding information; hence, the detection of a binding interaction primarily relies on the fast/intermediate exchange of the ligand between its bound and free form (fast on/off rates). Due to their requirement for fast on-off ligand exchange, the ligand-observed NMR binding methods can be insensitive to the binding of high-affinity ligands. Similarly, the NMR (^{19}F) slow exchange binding regime is also highly important, for example, it can used to measure ligand binding dissociation constant (further details in section 1.4.4) as well as for the detection of binding interaction of another competitive ligand (described further in section 1.4.5).

1.2.2 Basic theory of protein-ligand binding equilibria

Biological processes such as physiological and pathological mechanisms are governed by molecular interactions between enzymes and substrates, and between receptors and their ligands. The basic concept of molecular interactions was described by Emil Fischer's classic ″lock-key″ analogy (1894) and Daniel E. Koshland's ″induced fit theory″ (1958).[40] The ″lock-key″ model is based on appropriate shape and volume of the interacting molecules, while the ″induced fit theory″ is based on the dynamic rearrangements of the molecules that enable the fit of the ligand within the macromolecule. The complex formed is stabilized by non-covalent hydrogen-bonding, ionic, van der Waals, hydrophobic and π-interactions and in some instances by covalent interactions.

The dynamic equilibrium in the binding interaction of a single small molecule ligand (L) with a macromolecule such as protein (P) at a single binding site is generally represented as, Figure 1.4.

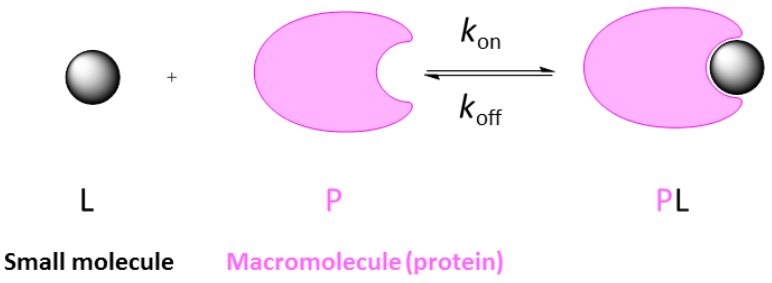


Figure 1.4 Protein-ligand binding equilibria.
Cartoon representation of the binding equilibrium of a small molecule ligand (L) with a macromolecule such as a protein (P).

The strength of the binding interaction can be represented either by the equilibrium association constant K_A, or the dissociation constant K_D. More commonly, the term K_D is used to describe the binding affinity of a ligand to a protein. The values of both K_A and K_D can be directly calculated from the equilibrium concentrations of the free protein [P], free ligand [L] and the protein-ligand complex [PL] as below,

$$K_A = \frac{[PL]}{[P][L]} \qquad (1)$$

$$K_D = \frac{[P][L]}{[PL]} \qquad (2)$$

The stronger P and L bind to one another, the more PL there will be at equilibrium. Conversely, the weaker the binding of L to P, the more unbound P and unbound L there will be at equilibrium. Thus, according to eq. 2, very strong binding is described by very small K_D values.

The values of [P] or [L] can be calculated from,

$$[P_T] = [P] + [PL]$$

Or,
$$[P] = [P_T] - [PL] \qquad (3)$$

Similarly,

$$[L_T] = [L] + [PL]$$

Or,
$$[L] = [L_T] - [PL] \qquad (4)$$

Where, $[P_T]$ represents the total concentration of protein and $[L_T]$ the total concentration of ligand present in solution.

The equilibrium constant may also be expressed as the ratio of the rates of association and dissociation of the complex as,

$$K_A = \frac{k_{on}}{k_{off}} \qquad (5)$$

$$K_D = \frac{k_{off}}{k_{on}} \qquad (6)$$

Where, k_{on} and k_{off} are the rate constants for the association and dissociation reactions, respectively.

In principle, if the concentrations of protein-ligand complex [PL], free protein [P] and free ligand [L] are known, then the binding dissociation constant can be directly calculated from equation 2. However, experimentally, the concentration of [PL] is typically difficult to measure. Instead, it is measured as a fraction bound out of total protein or total ligand concentration in the assay i.e $\frac{[PL]}{[P_T]}$ or $\frac{[PL]}{[L_T]}$

The fraction of the bound species can be determined in the form of a response (i.e change in the NMR parameter) from titration of one binding partner to a fixed concentration of the other binding partner. In experimentally measuring the fraction of the bound species, the following two cases are considered;

(1) In situation where one binding partner is in large excess over the other e.g if the concentration of $[L_T] >> [P_T]$, then one can make the approximation that [PL] is negligible, hence in such case, $[L_T] = [L] + [PL]$ becomes, $[L] = [L_T]$, i.e free ligand concentration equals to total ligand concentration.

Hence, by putting the value of $[L] = [L_T]$ and the value of [P] from eq. 3 into eq. 2, we get,

$$K_D = \frac{[L_T]([P_T] - [PL])}{[PL]}$$

After rearranging,

$$[PL] = \frac{[L_T][P_T]}{(K_D + [L_T])}$$

This eq. in terms of the fraction of $[L_T]$ bound in the [PL] complex ca be written as,

$$\frac{[PL]}{[L_T]} = \frac{[P_T]}{(K_D + [L_T])} \qquad (7)$$

From eq. 7, it may be concluded that if one could measure the fraction of ligand-bound as a function of total ligand concentration $[L_T]$ by titration experiments, the data would map out a curve that could be fit to yield a value for K_D.

(2) In a more general case, when the concentrations of the two binding partners are comparable, then in such system, the fraction of the bound species can be determined as,

Eq. 2, after inserting the values of [L] and [P] becomes,

$$K_D = \frac{([P_T] - [PL])([L_T] - [PL])}{[PL]}$$

Or, $$[PL]K_D = [P_T][L_T] - [PL][L_T] - [PL][P_T] + [PL]^2$$

After multiplying it out and rearranging,

$$[PL]^2 - ([L_T] + [P_T] + K_D)([PL] + [L_T][P_T]) = 0 \qquad (8)$$

Which is the form of quadratic eq. as,

$$ax^2 + bx + c = 0$$

Where, $a = 1,\ b = -([L_T] + [P_T] + K_D),\ c = [L_T][P_T]$

Thus, $$[PL] = \left(\frac{([L_T] + [P_T] + K_D) - \sqrt{([L_T] + [P_T] + K_D)^2 - 4[L_T][P_T]}}{2} \right)$$

Multiplying both sides by a common factor $\frac{1}{[L_T]}$ to get the equation in terms of the fraction of the ligand bound in total ligand.

$$\frac{[PL]}{[L_T]} = \left(\frac{([L_T] + [P_T] + K_D) - \sqrt{([L_T] + [P_T] + K_D)^2 - 4[L_T][P_T]}}{2[L_T]} \right)$$

Or,

$$\frac{\Delta_{obs}}{\Delta_{max}} = \left(\frac{([L_T] + [P_T] + K_D) - \sqrt{([L_T] + [P_T] + K_D)^2 - 4[L_T][P_T]}}{2[L_T]} \right) \qquad (9)$$

Where, Δ_{obs} represents a change in any observable NMR parameter at a single point in the titration, and Δ_{max} indicates the maximum change of the monitored NMR parameter i.e the saturation point.

1.3 Protein-observed NMR methods for binding studies

In general, in the protein-observed NMR binding methods, a change in the chemical shift of the protein resonances after the addition of a ligand demonstrates the binding interaction of the ligand with the protein. Some of these methods are discussed below.

1.3.1 2D ^{1}H-^{15}N HSQC or 2D ^{1}H-^{13}C HSQC binding experiments

The commonly used NMR protein-observed method for ligand binding studies involves studying the chemical shift perturbation of protein backbone amide protons with the addition

of a ligand in 2D ^{1}H-^{15}N HSQC experiment (Figure 1.5).[20,30,38,41] This technique also referred as SAR (structure-activity relationship) by NMR, typically requires uniformly ^{15}N-labelled protein. ^{15}N-enriched protein is most often produced by expression of the host-cells in minimal medium supplemented with ^{15}N-ammonium salts.[42] Each resonance (cross peak) in the ^{1}H-^{15}N HSQC spectrum represents a single amino acid (except proline in which the N does not have attached H) in the sequence. In addition, two peaks for each NH_2 group from the side chains of Asn, Gln and Trp are also observed in the spectrum.

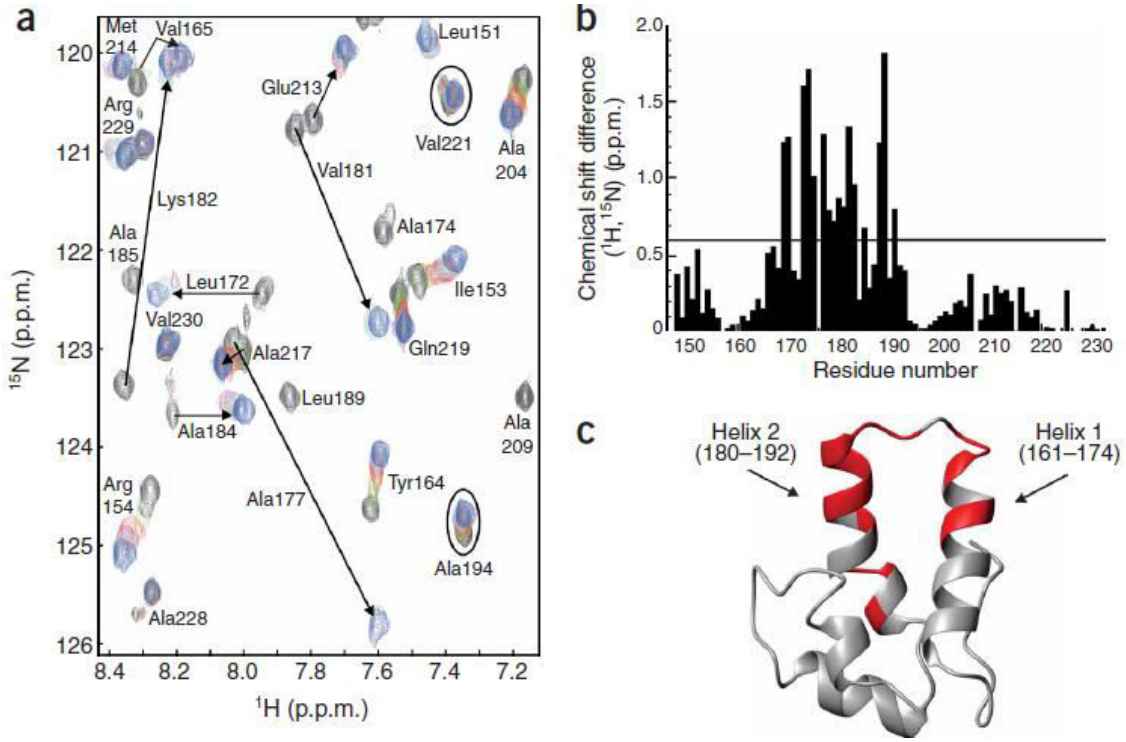


Figure 1.5 Illustration of protein- ligand binding by chemical shift perturbation in ^{1}H-^{15}N–HSQC NMR.
(a) Overlay of a series of 2D ^{1}H-^{15}N–HSQC spectra obtained for viral c-terminal capsid domain (C-$CA_{W184A/M185A}$) (130 μM) in the absence (130 μM black) and presence of increasing concentrations of CAI inhibitor (inhibitor/protein ratio as follows: 0.5:1 (green), 1:1 (red), 2:1 (magenta), 4:1 (cyan), 7.3:1 (blue)) in NMR fast exchange binding regime. Examples of peaks shifting upon CAI titration are shown by arrows indicating their involvement in inhibitor binding, and unaffected peaks are circled. The K_D of the interaction was determined to be 15 ± 7.2 μM by fractional-shift fitting. For clarity, only part of the spectrum is shown. (b) ^{1}H-^{15}N chemical shift perturbations plotted for each assigned amino acid in the protein after the addition of ligand. The horizontal line denotes a substantial chemical shift perturbation (0.6 ppm). (c) Ribbon 3D structure of the protein. Regions of the protein with residues showing chemical shift perturbations > 0.6 ppm upon inhibitor binding are coloured red. (Reproduced from reference [43])

The spectral dispersion of the ^{1}H-^{15}N-HSQC peaks is characteristically different for proteins with a defined three-dimensional structure (folded or active form), compared with proteins with random coil conformations.[44] The lack of dispersion and sharpness of the peaks are typical of unfolded protein, whereas, the well-dispersed and broader peaks are the characteristics of folded proteins (Figure 1.6).

Because the amide chemical shifts are highly sensitive to perturbations in the local chemical environment, they are good reporters of ligand binding, mapping ligand-binding site, protein conformational changes, protein dynamics and exposure of the protein surface to the solvent. Upon ligand binding, residues in or close to the ligand binding pocket undergo large chemical shift perturbation than those away from the ligand binding pocket.

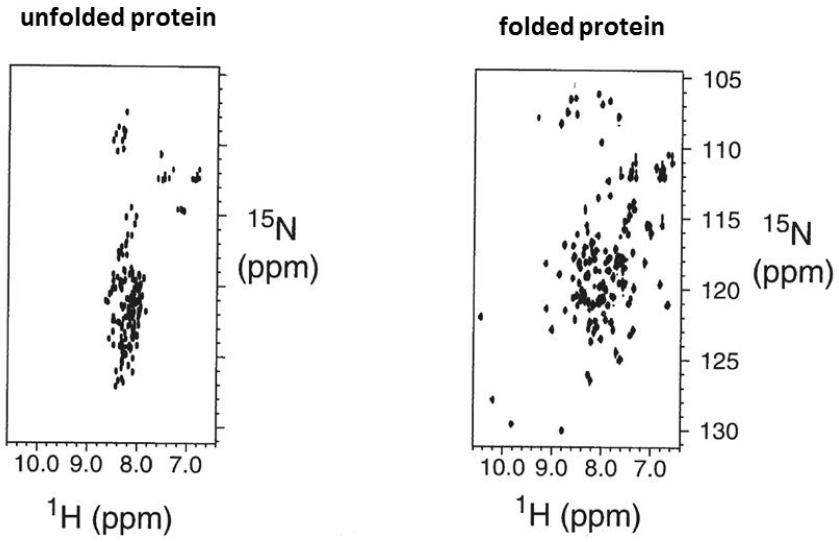


Figure 1.6 ^{1}H-^{15}N HSQC spectra of unfolded (left) and folded (right) apomyoglobin.
For the folded form of the protein the resonances are well-dispersed in contrast to the unfolded form. (Reproduced from reference [45])

Typically perturbations are considered significant if $\Delta\delta > 0.1$ ppm for a minimum of two peaks in the spectrum,[46] where,

$$\Delta\delta = \sqrt{[(\Delta\delta_H)^2 + \frac{1}{6}(\Delta\delta_N)^2]}$$

Where $\Delta\delta_H$ and $\Delta\delta_N$ are the differences between the HN and ^{15}N chemical shifts, respectively, of the fully saturated and free protein. The ligand binding dissociation constant (K_D) can also be measured by this technique, by titrating ligand to a solution containing fixed protein concentration (Figure 1.7).[47-49] For a fast exchanged binding system, the chemical shift perturbation is measured as a function of ligand concentration and the K_D value is calculated by fitting the fractional shift according to quadratic equation.

In order to map which specific residues in the sequence in the protein structure are involved in the binding interaction with the ligand, the assignment of these resonances is the ^{1}H-^{15}N HSQC spectrum is required. For small-sized proteins (MW < 10 kDa), the sequential

assignment strategy may be used.[50] In this technique first a 3D heteronuclear-edited correlation spectra (such as TOCSY-^{1}H-^{15}N HSQC) experiment is recorded to identify amino acid spin systems by tracing through-bond scalar interactions along the amino acid side chain. The next step involves the analysis of through-space dipolar interactions, usually through 3D heteronuclear-edited experiments such as the NOESY-^{1}H-^{15}N HSQC for sequential NOE analysis. For larger-sized proteins (MW ~ 10-25 kDa), the triple resonance assignment strategy is used.[51,52] This approach requires uniformly ^{15}N and ^{13}C (backbone amide carbonyl)-

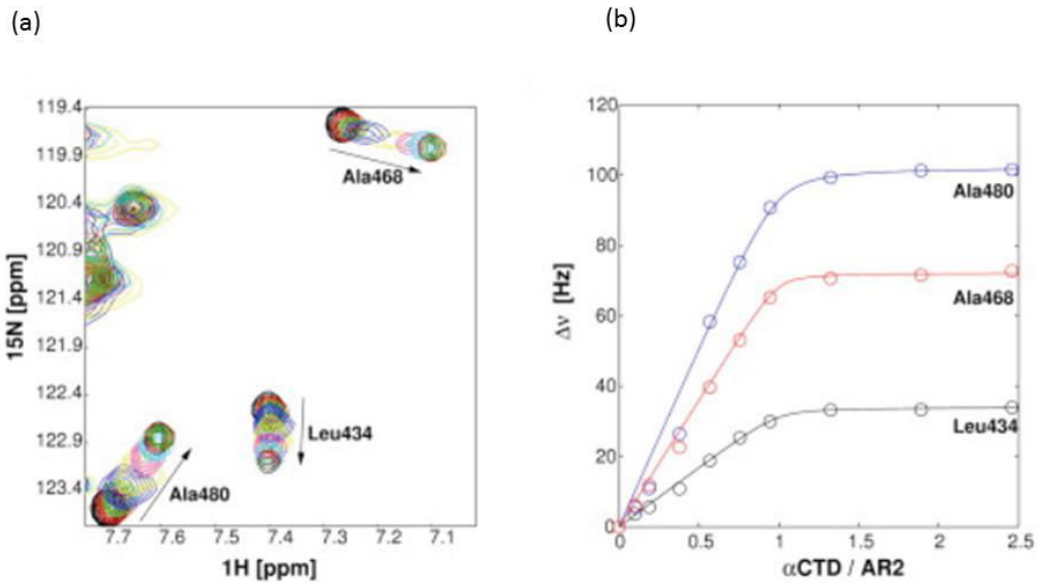


Figure 1.7 Ligand binding constant (K_D) measurement by ^{1}H-^{15}N HSQC chemical shift perturbation.
(a) To a fixed concentration of protein, ligand is titrated and the change in the chemical shift of the protein resonances is measured as a function of ligand concentration. (b) Binding curves of (a). (Reproduced from reference [53])

labelled protein, but for protein above 25 kDa, triple labelling (^{15}N, ^{13}C, ^{2}H) is typically required.[54] The triple resonance experiments depends primarily on the relatively large ^{1}J (and in some cases ^{2}J) couplings between ^{1}H, ^{15}N and ^{13}C spins in the protein backbone sequence for coherent magnetization transfer (Figure 1.8a). The set of triple resonance experiments that are recorded for sequence specific backbone assignments most commonly include the HNCA, HN(CO)CA, HNCACB, HN(CO)CACB, HNCO and HN(CA)CO (Figure 1.8b). The naming of these experiments describes the magnetization transfer pathway, with the chemical shifts of the nuclei in parenthesis is not being mapped. The experiments are recorded in pairs with one experiment gives correlation information to both the residue itself (residue *i*) and the previous residue (residue *i-1*) and its partner experiment provides only the inter-residue (*i-1*) correlation.

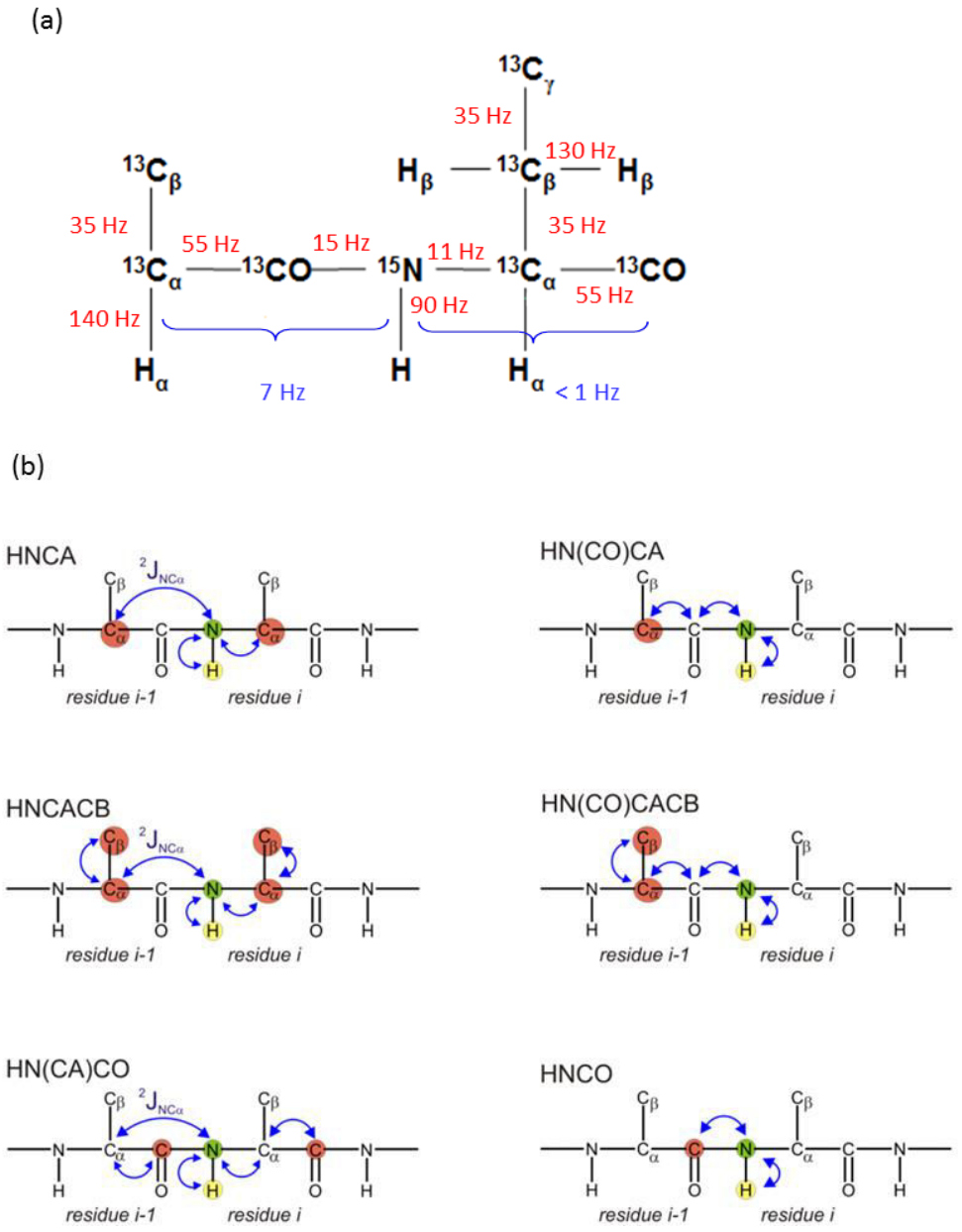


Figure 1.8 Typical NMR triple resonance experiments for the sequential assignment of protein backbone resonances. (a) Typical $^{1}J/^{2}J$ coupling constants in a protein backbone spin system (Reproduced from reference [51]). (b) Different types of triple resonance experiments recorded for the sequential assignment of protein backbone resonances. The HNCA experiment gives the connection between the amide proton of a residue with its own Cα and with Cα of the preceding residue. The HN(CO)CA experiment connects the amide proton of a residue to the Cα of the preceding residue. The HNCACB gives connectivity between the amide proton of one residue to its own Cα and Cβ and Cα and Cβ of the preceding residue. The HN(CO)CACB experiment gives correlation between the amide proton of a residue to the Cα and Cβ of the preceding residue. The HN(CA)CO correlates the amide proton of a residue to its own amide carbonyl carbon and the amide carbonyl carbon of the preceding residue. The HNCO experiment gives correlation between the amide proton of a residue to the amide carbonyl of the preceding residue. Each correlation gives rise to a cross peak in the 3D spectrum. The ^{13}C chemical shifts of α and β carbons then identifies the amino acid type.

From the chemical shift perturbation, it is also possible to distinguish between specific and non-specific binding interactions. This technique also does not rely on the ligand fast

exchange rate to retrieve the information from the bound state, making it possible to detect both high and low affinity interactions.

Although the ^{1}H-^{15}N HSQC technique is very useful in characterizing ligand binding events, the technique also has some limitations. For example, it does not always report ligand binding due to hydrophobic interactions, which has a small effect on backbone NH chemical shifts. Similarly, it also does not report ligand binding involving proline resides. In addition, in this technique, as the size of the protein increases (~ 40 kDa), the resonances in the ^{1}H-^{15}N HSQC spectrum become broader and more abundant, which may sometimes lead to signals overlap, and hence it may not be possible to detect chemical shift perturbation in the presence of ligand. The sequential resonance assignment for such proteins also sometimes become challenging in terms of data analysis, resources and instrument time. The problem of signal overlap may be overcome by using site-specifically ^{15}N-labelled protein (e.g. histidine) or by using ^{13}C-labelled protein in 2D ^{1}H-^{13}C HSQC experiments (Figure 1.9).[20] Each cross peak in the ^{1}H-^{13}C HSQC spectrum represents a directly bounded pair of ^{13}C, ^{1}H nuclei in the protein. The ^{13}C-enriched protein is produced by expression of the host cells in minimal medium most commonly supplemented with ^{13}C-glucose.[42] This approach is however not generally applied as compared to ^{1}H-^{15}N HSQC technique, as amino acids contain multiple ^{1}H-^{13}C correlations *via* CH, CH_2 and CH_3 groups which results in protein ^{1}H-^{13}C HSQC spectrum being rather more complex than ^{1}H-^{15}N HSQC spectrum. This complexity leads to spectral overlap and difficulty in interpretation. Also the ^{13}CH chemical shifts differences are usually smaller than those observed for ^{15}NH and generally more difficult to see.[55]

The problem of signals overlap may be somewhat overcome by the labelling of only methyl groups in residues such as valine, leucine or isoleucine.[56-59] In this approach the sensitivity of the experiments also increases by 3-fold due to the presence of 3 protons in methyl compared with one proton of amide group in ^{1}H-^{15}N HSQC experiments. Consequently, the consumption of protein in the assay is also reduced. In addition, the favourable relaxation properties of methyl groups also allow the applicability of this technique to larger sized proteins. The applicability of selectively ^{13}C(methyl)-labelling approach however clearly depends on the presence of valine, leucine, or isoleucine residues in the active site of the protein under investigation.

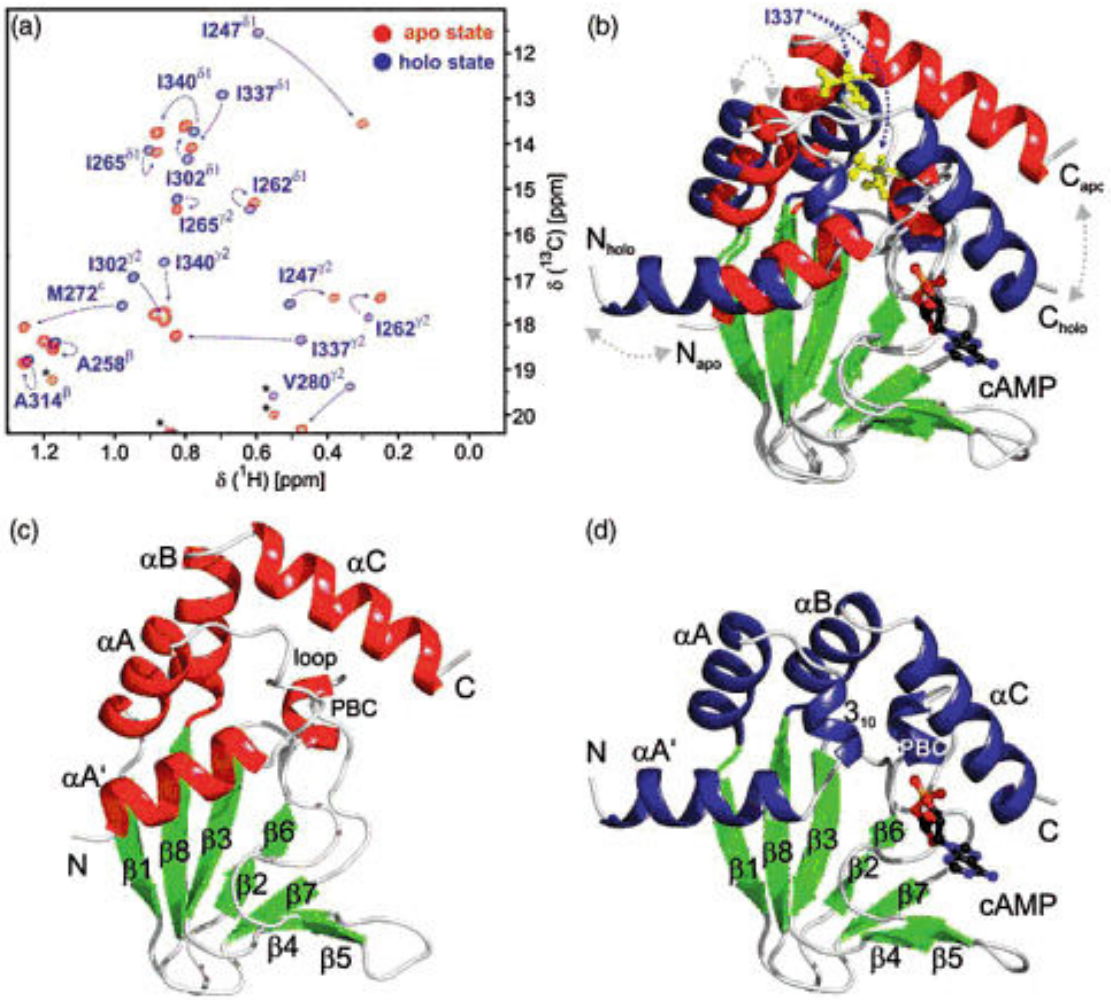


Figure 1.9 Illustration of protein- ligand binding by chemical shift perturbation in ^{1}H-^{13}C HSQC NMR.
Example shown is the binding of cAMP to MloK1 cyclic nucleotide binding domain (CNBD). (a) Superimposition of 2D ^{1}H-^{13}C HSQC spectra (parts shown) of 500 µM cAMP-free (red) and cAMP-bound (blue) [*U*-^{13}C, ^{15}N]-labelled MloK1 CNBD. Chemical shifts of all CNDB cross-peaks changed upon binding of cAMP (dashed-arrow) (b) Superimposition of cAMP-free and cAMP-bound solution structures shows that the β-roll region (green) is almost identical, whereas the helical portion undergoes significant structural rearrangement (dashed arrows) upon cAMP binding. Different side chain positions of I337 in the cAMP-free and cAMP-CNDB are exemplary shown (yellow) (c) Ribbon representation of cAMP-free (*apo*-form) and (d) cAMP-bound (*holo*-form) CNDB solution structures. The cAMP molecule is shown as ball-and-stick model. (Reproduced from reference [60])

Over and above the inherent resonance assignment problem in the ^{1}H-^{15}N HSQC/^{1}H-^{13}C HSQC, the requirement for isotopic enrichment is an additional issue. The production of ^{15}N and/or ^{13}C-labelled protein in a suitable expression host is resource and labour intensive work. Due to these limitations, the protein-observed ^{1}H-^{15}N HSQC/^{1}H-^{13}C HSQC techniques may not be readily implemented for studying ligand binding events to protein.

1.3.2 1D ^{19}F NMR protein-observed binding technique

Standard 1D ^{19}F NMR protein-observed experiments are finding increasing use to detect protein-ligand binding interactions, protein dynamics and conformational changes that arise due to ligand binding events.[61-65] The ^{19}F nucleus has a comparable NMR sensitivity to that of a ^{1}H i.e 0.83 of a proton (1), has a 100% natural abundance, and has a large chemical shift range of up to 800 ppm.[66] ^{19}F nucleus is much more sensitive to changes in the local chemical environment as compared to NH group in ^{1}H -^{15}N HSQC/^{1}H-^{13}C HSQC and therefore acts as a good NMR probe to detect changes in the local chemical environment occurring due to binding events. Fluorine incorporation into a protein is generally nonperturbing, especially when substituted for hydrogen in an amino acid side chain. The atomic size of fluorine is almost similar to hydrogen atom (covalent radii of 1.35 Å and 1.2 Å, respectively); therefore fluorine substitution to hydrogen has little steric effect.

For the ^{19}F-NMR protein-observed experiments, fluorinated probes are incorporated into the protein, and when a ligand binds in close proximity (up to 25 Å) to the fluorinated-probe, it (probe) undergoes chemical shift changes. This technique can also be used to calculate ligand binding affinities usually by titrating ligand to fixed protein concentration and monitoring changes in the chemical shift of the fluorine signal of protein as a function of ligand concentration. The ^{19}F nucleus chemical shifts in proteins are governed primarily by the fluorine lone-pair electrons, which give a large paramagnetic term in the shielding formula. The chemical shift of a ^{19}F NMR resonance therefore, is exquisitely sensitive to changes in the local van der Waals environment, as well as to local electrostatic fields.

Importantly, ^{19}F nucleus does not occur in the proteins naturally, therefore the spectrum is free from interfering background signals, which simplifies spectral interpretation and allows high sensitivity. This technique is also therefore particularly useful for *in vivo* binding studies.

In the protein-observed 1D ^{19}F NMR binding technique, various fluorinated amino acids have been used as an NMR probe for binding studies, with most of these unnatural fluorinated amino acids are commercially available.[67] In most cases “lightly fluorinated” amino acids are biosynthetically incorporated into the protein in place of their natural forms, as they are not discriminated by the cognate actyl-tRNA synthetases. These include fluorinated aromatic amino acids such as 3-fluorophenylalanine (3-fPhe),[68,69] 4-fluorophenylalanine (4-fPhe),[70-72] 5-fluorotryptophan (5-fTrp),[73-77] 6-fluorotryptophan (6-fTrp),[71,78] 3-fluorotyrosine (3-fTyr);[79-

[88] fluorinated aliphatic amino acids including 5-fluoroleucine (5-fLeu),[89] trifluoroethylglycine (tfeGly),[90,91] trifluoromethionine (tfMet),[92-94] difluoromethionine,[95] and 2-fluorohistidine (2-fHis).[96]

Several techniques for the incorporation of fluorinated amino acids into protein have been reported in the literature.[97-100] Generally, lightly fluorinated amino acids are incorporated into proteins by simply adding the fluorinated amino acid to the expression growth medium using a suitable auxotrophic bacterium. The incorporation of highly fluorinated amino acids is however relatively complex. Several techniques have been used to assign ^{19}F NMR resonances of proteins. Most assignments are made based on site-directed mutagenesis to replace or to "nudge" each specific fluorine label in the spectrum (Figure 1.10). In this technique the residue of interest is substituted with a residue with most similar side chain, thereby causing the corresponding resonance in the ^{19}F NMR spectrum to disappear while the remaining resonances in the spectrum are unperturbed (Figure 1.10B).[69,101-106] The second approach uses the "nudge" analogy[102,103] and applies to proteins with known crystal structure. This technique identifies a sidechain (a residue) in van der Waals contact with the target residue for conservative replacement, thereby generating a local nudge (Figure 1.10C).

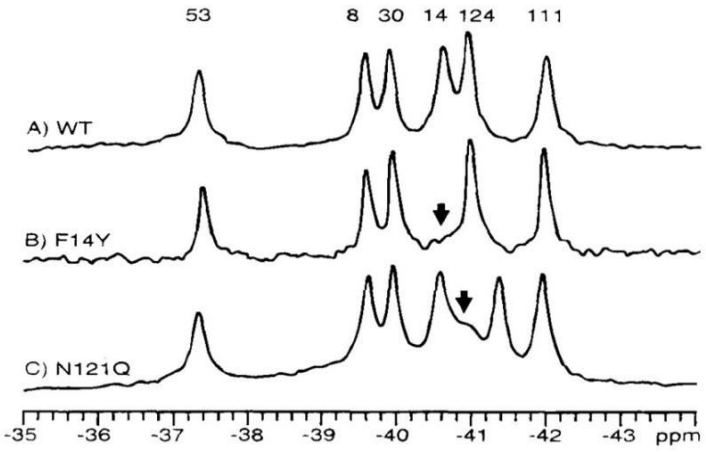


Figure 1.10 Strategies for the assignment of fluorinated residues of a protein.
^{19}F NMR spectra illustrating the two protein engineering techniques applied to assign the 4-F-Phe resonances of the phospho-signalling protein CheY. (A) The protein is labelled at its six phenylalanine positions with 4-F-Phe, giving rise to six distinct resonances in the ^{19}F NMR spectrum. (B) Direct mutation of the target residue (C) Nudge approach. (Reproduced from reference [102])

The resonance assignments can be further confirmed by chemical shift changes due to ligand binding to specific site, solvent exposure or chemical modification of an amino acid close to the fluorinated residue. Solvent exposure can be detected by solvent-induced isotopic shift

(SIIS). The SIIS effect can cause up to a 0.25 ppm chemical shift of a ^{19}F resonance of a residue exposed to aqueous solution when the solvent changes from H_2O to D_2O.[67,87] In contrast, a ^{19}F resonance of a residue from the interior of the protein does not experience solvation and the concomitant chemical shift changes, unless the residue interacts with buried water molecules in the protein interior.

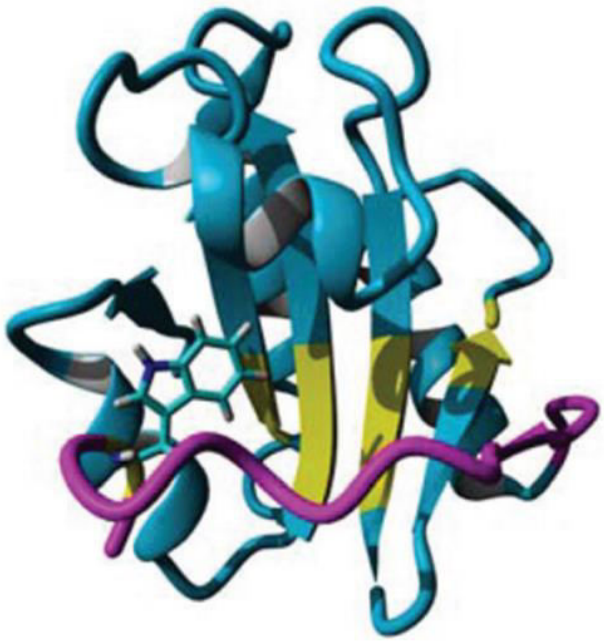
Site specific incorporation of unnatural fluorinated amino acids can be performed by employing an orthogonal aminoacyl-tRNA synthase/tRNA pair.[97] This protocol has been reported to be very useful for the incorporation of particularly fluorinated aromatic amino acids such as trifluoromethylphenylalanine (tfmPhe).[82,83,107,108] For small-sized proteins, solid phase chemical synthesis is commonly employed, because it allows more flexibility in terms of both the identity and position of the noncanonical amino acid. Together with native chemical ligation methods, large proteins containing specifically-labelled residues can potentially be synthesized.[109-111] This protocol has recently been used to produce affibody proteins (small proteins engineered for binding to target proteins) that contain perfluorinated phenyl rings.[112]

Fluorinated probes can also be incorporated into the proteins by chemical modification, by employing fluorinated "tags" such as 2,2,2-trifluoroethanethiol (TFET) that are reactive toward nucleophilic side chains, most commonly cysteinyl residues to form disulphide linkages.[65,113,114] Other fluorinated probes such as 3-bromo-1,1,1-trifluoroacetone[115] and fluorinated α-haloacetamides[116] have also been used for chemical conjugation, which react with the thiol group forming comparatively more stable thio-ether linkage. This method provides the simplest approach of introducing highly fluorinated probes into large proteins.

An example of the ligand-induced chemical shift changes in the protein as demonstrated by protein-observed 1D ^{19}F NMR assay is discussed in Figure 1.11.[117] Human protein disulphide isomerase (hPDI) is an enzyme that is required for the native disulphide bonds formation in proteins. hPDI is a multifunctional protein that catalyses the oxidation and isomerisation of disulphide bonds, as well as to bind to unfolded proteins and act as a molecular chaperone.[118] The structure of hPDI consists of four domains two of which have redox-active catalytic sites (a and a′) and two (b and b′) don't. The domains order is a-b-b′-x-a′-c, where x is a 19 residue linker and c is the c-terminal acidic tail;[119,120] b and b′ have been shown as essential for ligand binding to the hPDI.[121] The b′ domain binds both small and large peptide ligands. Investigating the underlying mechanism of substrate binding to the b′

domain is of significant interest that may provide useful information about the role of b′ and x in these proteins. Recent studies have shown that the hPDI b′ and x domains are structurally connected to each other. The b′x construct has been reported to accept ligand binding[119]. Similarly, x has also been shown to moderate homodimerisation in b′x and bb′x contructs.[122] The crystal structure studies of a b′x mutant construct have also shown that b′ contains the thioredoxin-fold with x closing the ligand binding site.[119,123] Based on these results, it is believed that b′x likely exists in two possible conformational states with the x-linker 'capping' or 'uncapping' the ligand binding site and that 'uncapping' allows the opportunity for homodimerisation. This capping and uncapping of the ligand binding site by x has been proposed as a potential regulation mechanism for the protein in *vivo*.[122] Initially ^{1}H -^{15}N HSQC technique was applied to probe these local environment changes in the x region of b′x upon ligand binding, however the backbone resonances in the linker itself were highly difficult to observe and track due to significant broadening. The hPDI b′x construct contains only a single tryptophan residue within the x-linker at position 347 and is located directly above the reported ligand binding site as shown in (Figure 1.11a). Hence incorporating a fluorine atom in this residue was a good probe that can potentially reveal these local environment changes of x b′x upon ligand binding in ^{19}F NMR protein-observed experiment. In the assay to a fixed concentration of hPDI 5-F-Trp-b′x, somatostatin peptide ligand was titrated in increasing amounts (Figure 1.11b).

(a)



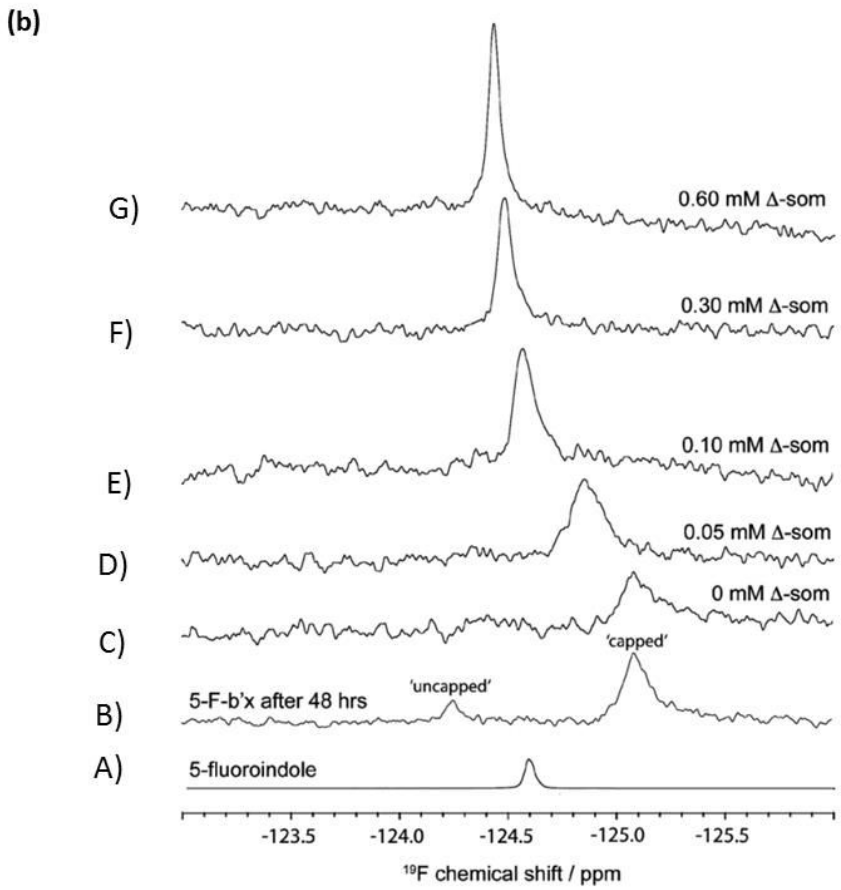


Figure 1.11 Studying protein-ligand binding by protein-observed ^{19}F NMR.
(a) Crystal structure of hPDI b′x I272A (PDB ID: 3bj5) with the 'capped' x-linker shown in magenta and the ligand-binding site in yellow. The x-linker Trp 347 is shown in stick form. (b) ^{19}F NMR spectra of 0.3 mM hPDI 5-F-Trp-b′x with increasing concentration of somatostatin (som) ligand (C-G). B: ^{19}F NMR spectra of 48 h 'aged' hPDI 5-F-trp-b′x showing both the 'capped' and 'uncapped' forms of the hPDI 5-F-trp-b′x. hPDI 5-F-trp-b′x is prone to dimerise slowly over time. All these experiments were however run in 30 min time during which no dimerization was observed. A: ^{19}F NMR reference spectrum of 5-fluoroindole. (Reproduced from reference [117]).

It was observed that as the ligand concentration increases, the fluorine signal of 5-F-Trp-b′x line-width decreases (becomes sharper) with concomitant changes in its chemical shift demonstrating the displacement of X-linker due to ligand binding to the protein. This supports the capping and uncapping of ligand binding site by the x-linker in hPDI b′x.

Although, the protein-observed ^{19}F NMR assays are very useful for studying ligand binding events to proteins, the technique still requires the production of ^{19}F-labelled protein and the assignment of the ^{19}F resonances. Sometimes, the specific ^{19}F labelling is also not possible due to issues with protein stability.

1.4 Ligand-observed NMR methods for binding studies

Compared to the NMR protein-observed binding methods, ligand-based binding methods are more readily implemented because these do not require labelled-protein. In the ligand observed NMR binding experiments, a change in the NMR properties, such as chemical shift, peak-height, line-width, relaxation rates, NOE or diffusion coefficient of the ligand that arise after the addition of protein, demonstrates the binding interaction of the ligand with the protein. In contrast to the protein-observed methods, the ligand-observed methods are not restricted to the size of protein but instead benefit from the large size of the protein, which increases the level of change in the NMR properties of ligand that occur due to binding with the protein, hence increasing the NMR sensitivity of binding interactions.

Most commonly the ligand-observed binding experiments are performed on ^{1}H [6,124] and ^{19}F [124-127] nuclei due to their high NMR sensitivity; other nuclei such as ^{13}C [127] and ^{31}P [28] have also been used but rarely due to their low intrinsic NMR sensitivity.

In general, the ligand-observed methods based on changes in the ligand transverse relaxation rates (or line-shape) and methods based on ligand NOEs are most commonly used for studying small molecules binding interaction with the protein.

1.4.1 ^{1}H and CPMG edited ^{1}H NMR direct ligand observation

Direct ligand observation based on monitoring changes in the ligand line-shape for broadening and/or attenuation after the addition of protein in standard 1D ^{1}H NMR[6] experiments as a probe of binding interaction, is the most sensitive and readily implemented NMR technique for studying small molecules binding interaction with the proteins (Figure 1.12).

A small molecule (MW ~100-1,000 Da) generally tumbles fast in solution, hence has a large T_2 relaxation time constant or low transverse relaxation rates (R_2) which translates into sharp resonances in the NMR spectrum. On the other hand, a macromolecule (MW 10s of kDa), tumbles slowly in solution, thus has a shorter T_2 relaxation time constant or higher transverse relaxation rates (R_2) and gives rise to broad resonances in the NMR spectrum. When a small molecule binds with a macromolecule (protein), then it behaves as if it were part of the macromolecule, and as a result its NMR resonances are broadened and/or attenuated (Figure 1.12). As this technique is based on monitoring the signals of the bulk (free) ligand in

solution for binding interaction, therefore the bulk ligand in solution must be in rapid exchange with the bound ligand in solution.

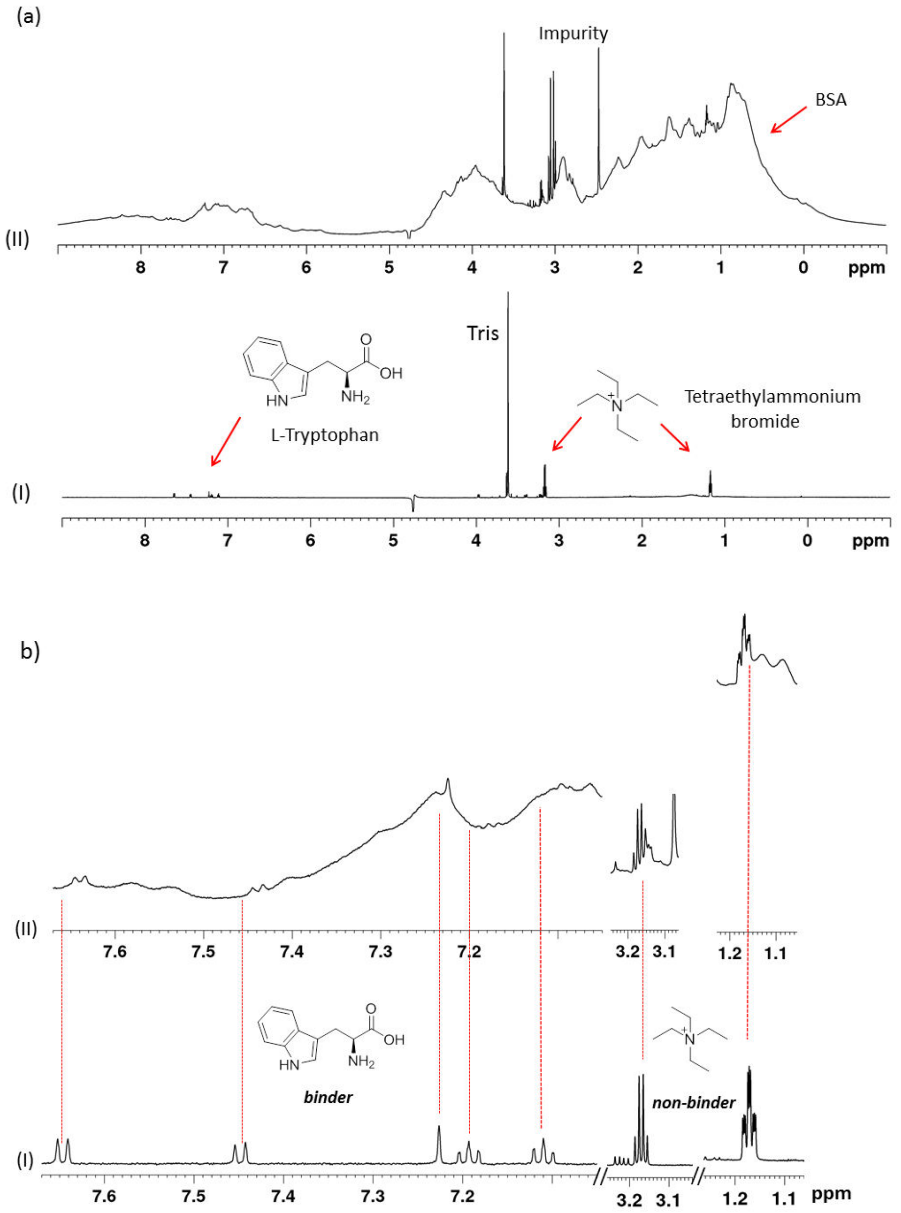


Figure 1.12 Studying Protein-ligand binding by ^{1}H NMR direct ligand observation.
The broadening and/or attenuation of ligand signals with the addition of protein demonstrate the binding interaction of the ligand with the protein. (a) I: solution of 50 µM L-tryptophan and 50 µM tetraethylammonium bromide in 50 mM Tris-D_{11}, pH 7.5, in D_2O. II: after the addition of 500 µM bovine serum albumin (BSA) protein to I. (b) Zoom in view of (a). In the spectrum after the addition of BSA (II), the signals of the L-tryptophan are broadened and attenuated, demonstrating its fast exchange binding interaction with BSA, whereas the signals of the tetraethylammonium bromide remained unaffected after the addition of BSA, suggesting that it doesn't bind with BSA. Note that in ^{1}H NMR direct ligand observation, the spectrum is dominated by the protein resonances.

If the bulk ligand in solution is in fast to intermediate chemical exchange with the bound-ligand (short residence time on protein or weak interaction) on NMR time scale, then in such

cases the total broadness in the ligand signals is the sum of broadness due to binding to the protein and broadness due to weighted-average ligand populations between the bound and free forms. In contrast, if the bulk (free) ligand in solution is in slow chemical exchange with the bound-ligand on NMR time scale (long residence time on protein or strong interaction), then in such binding regimes, two separate signals are observed for the ligand in the spectrum; one for the free population of the ligand (sharp resonances) and one for the bound population of the ligand (broad signal). In a ^{1}H NMR spectrum, however the signal for the bound population of the ligand is not usually observable as it is buried under the protein resonances, hence only the signal of the residual free ligand left is observed and followed for binding interaction.

In this technique, two spectra are recorded, one for the ligand in the absence of protein and another in the presence of protein. The ligand signals in the two spectra are compared for broadening and/or attenuation, as a detection of binding interaction. A non-binder compound is also sometimes added to the sample as an internal NMR standard, to demonstrate that the ligand˝binder˜peak broadening and/or attenuation with the addition of protein is due to its binding interaction with the protein and not due to poor sample preparation, ligand-aggregation or field inhomogeneity (poor shimming); the signals of internal NMR control remain sharp with full intensity both in the absence and presence of protein (Figure 1.12).

As an alternative and often beneficial approach to the conventional ^{1}H NMR is to record the transverse relaxation edited or Carr-Purcell-Meiboom-Gill (CPMG)[128,129] (Figure 1.13) edited ^{1}H NMR experiments (Figure 1.14). The CPMG spin echo pulse cycles of $(t\text{-}180°\text{-}t)_n$ exploits the differential T_2 relaxation rate of large molecules such as protein and small molecules ligands such that during the $2t$ period (also called total echo time/filter time or mixing time) the intensity of signals with short T_2 (broad lines) diminishes much more quickly than that for signals with long T_2 (sharp lines). The selective removal of the protein background broad resonances from the spectrum by CPMG filter allows more clear observation of the ligand NMR resonances with the addition of protein. In addition, the CPMG filter also improves the sensitivity to ligand binding by enhancing the attenuation of the broad signal of ligand.

Figure 1.13 shows the pulse sequence of a spin echo; the average net magnetisation from a group of nuclei with similar spin (*isochromats*) such as ^{1}H in a molecule at equilibrium is represented by a vertical red arrow along z-axis (A). The net magnetisation is excited by applying a 90° pulse along horizontal (x-y) plane (B).

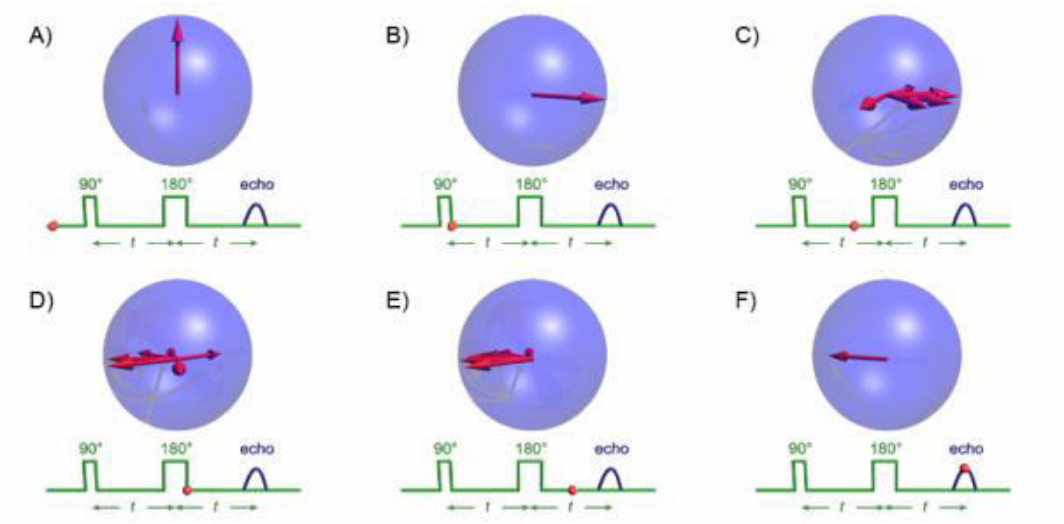


Figure 1.13 CPMG spin echo pulse sequence.
(A) The average spin from a group of nuclei with similar spins (^{1}H) in a molecule at equilibrium is represented as vertical red arrow. (B) A 90° pulse along x-axis flips this net magnetisation into the x-y plane (spins excitation). (C) Due to difference in chemical shifts (frequencies), the spins (isochromats) fan out during *t*. (D) A 180° pulse along y-axis is now applied to invert the magnetisation. (E) At time *t*, the fast spins catch up with the average spin and the slow moments drift back toward the average spin. (F) Complete refocusing occurs at total echo time 2*t*. This 180° pulse is repeated several times in order to freeze the evolution of J coupling.

During the time period *t*, the spins evolve according to their chemical shift differences, such that some spins have higher frequency than the mean spin vector, whereas, some have lower frequency than the average spin vector (C). A 180° pulse is then applied along the y-axis, which inverts the spins (D). Following the inversion pulse, another delay (*t*) is applied, during which all the magnetization vectors begin to refocus (E), such that at the end of 2*t* period all the magnetisation is completely refocused (F) and the FID (free induction decay) is collected. In the CPMG spin echo, the sequence (*t*-180°-*t*) is repeated several times (*n*) in order to remove the artefacts in the 180° pulse to prevent distortion in the signal phasing. During this 2*t* period, it is this sufficiently large difference in the T_2 spin relaxation times of protein and small molecule ligand, that results in the preferential rapid decay of the protein resonances with the CPMG echo sequence due its shorter T_2 (fast relaxing spins) compared with the resonances of the small molecules, which are less affected due to larger T_2 (slow relaxing spins) (Figure 1.15). In this technique, in order to freeze the J coupling evolution (distortion in the signal multiplicity) during the 2*t* period, the value of *t* is kept smaller than the value of spin coupling constant J ($t << 1/J$). The total echo time used in the CPMG edited ^{1}H NMR experiments is usually from a few tens of milli seconds to a few hundred milli seconds. This technique is also commonly used in the study of metabolomics biofluids to attenuate the background signals from large molecules such as proteins and lipids.

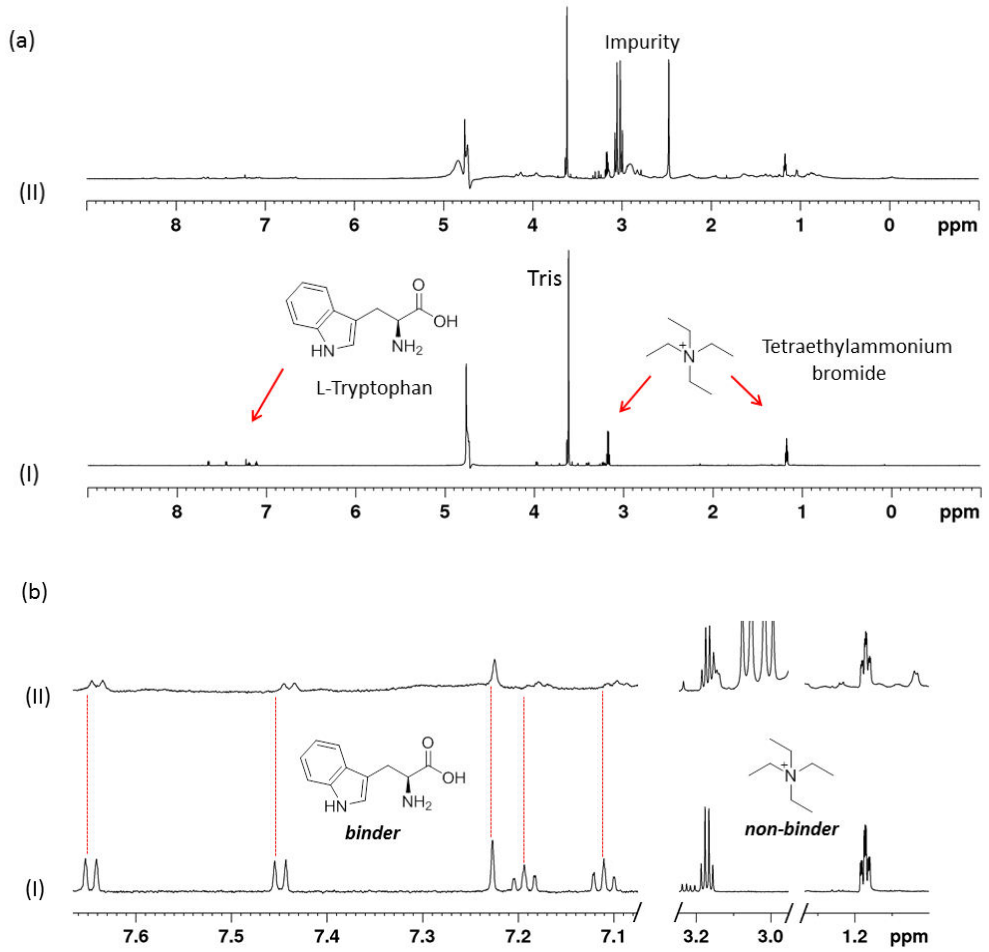
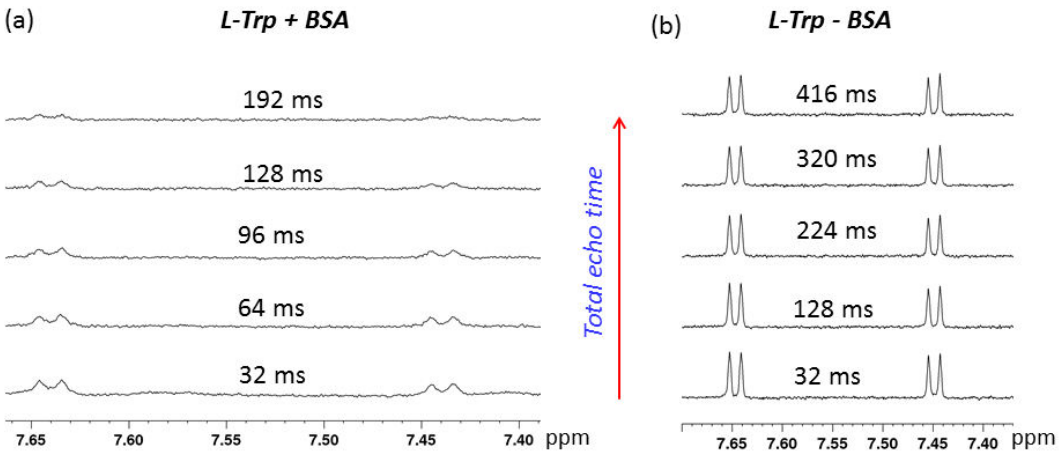


Figure 1.14 Studying Protein-ligand binding by CPMG edited ^{1}H NMR direct ligand observation.
(a) I: spectrum of 50 µM L-tryptophan and 50 µM tetraethylammonium bromide in 50 mM Tris-D_{11} , pH 7.5, in D_2O. II: spectrum after the addition of 500 µM BSA to (I). Note that the CPMG filter has completely removed the protein resonances from the spectrum. (b) Zoom in view of the corresponding spectra of (a). Note that after the removal of protein resonances from the spectrum by CPMG filter, it become easy for the ligands signal to be followed for binding interaction. The total filter time used was 32 ms.



(c)

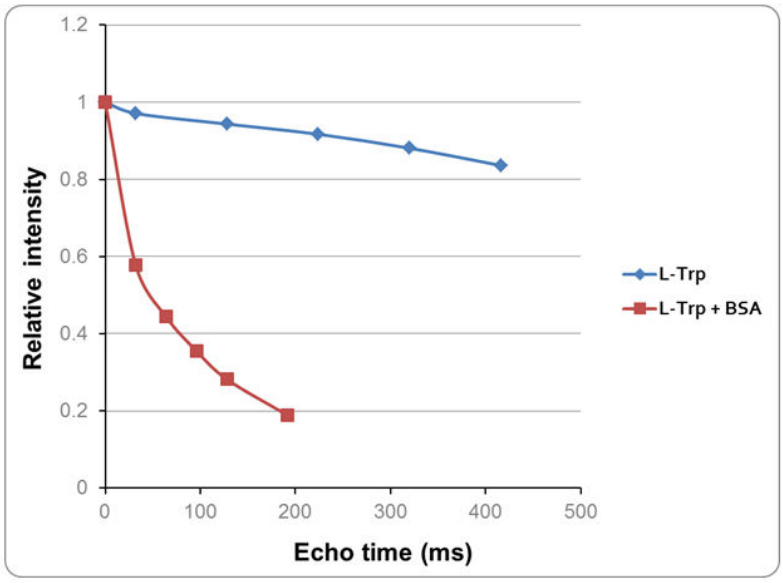


Figure 1.15 The effect of CPMG filter time on the signal intensity of free and bound ligand.
The effect of CPMG total echo time on the L-Trp signals in the (a) presence of BSA (b) absence of BSA (c) The signals decay profiles of (a) and (b) as a function of total echo time. The CPMG filter improves the sensitivity to ligand binding by enhancing signal attenuation of broad resonances (a), whereas, it has a negligible effect on signals of free ligand (b). The concentrations used are 50 µM L-Trp, 50 µM tetraethylammonium bromide and 500 µM BSA. The assay samples were prepared in 50 Mm Tris-D_{11}, pH 7.5 in D_2O.

The direct ligand observation by CPMG edited ^{1}H NMR also offers one of the simplest NMR approaches to measure the ligand binding dissociation constant (K_D), by titrating protein into a fixed ligand concentration, measuring the decrease/attenuation in the ligand signal intensity relative to the blank ligand (without protein) spectrum, as a function of protein concentration (Figure 1.16). The peak-height/area of the ligand signals in the presence of protein for binding can be simply measured by direct integration. The titration data points are fitted using appropriate fitting software and the binding constant value is extracted using the quadratic K_D equation. Using a single concentration, binders can also be qualitatively ranked on the basis of level of attenuation and/or broadening of their NMR resonances in the presence of protein in a single shot experiment.

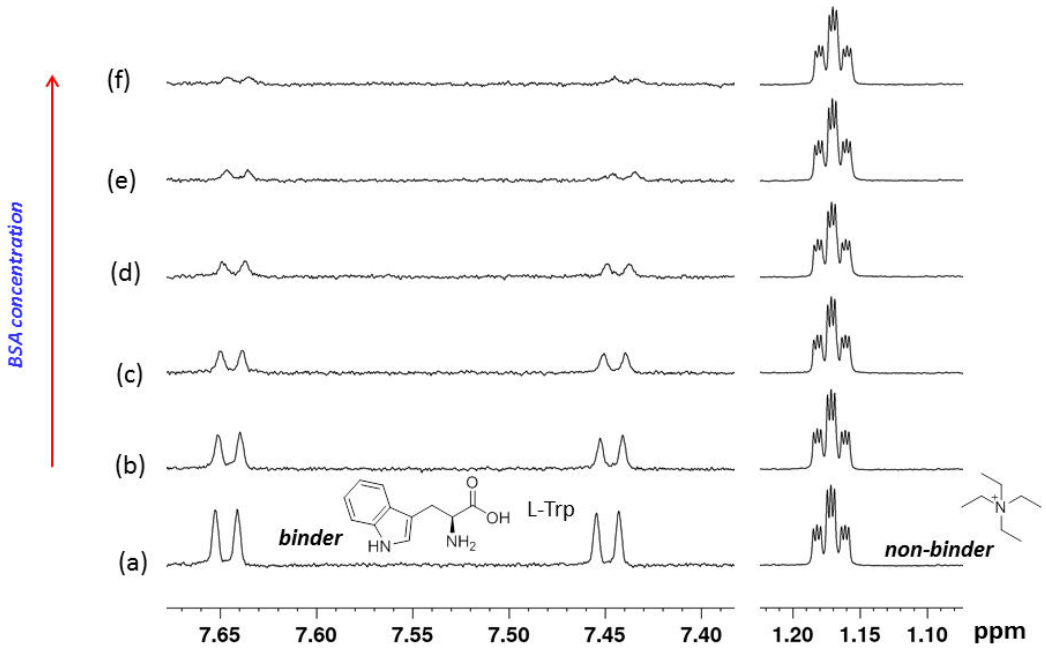


Figure 1.16 Ligand binding constant measurement by direct ligand observation in CPMG edited ^{1}H NMR.
BSA titration (b-f; 100 µM, 200 µM, 300 µM, 400 µM, 500 µM) to fixed concentration of L-tryptophan (50 µM) and tetraethylammonium bromide (50 µM) solution. The total echo time used was 192 ms. The splitting of each line in the triplet of tetraethylammonium bromide (4 CH_3 groups) into further triplets is due to their J coupling with the ^{14}N quadrapolar nucleus (spin 1, ~ 100% natural abundance).

In general, the direct ligand observation is an operationally simple, sensitive, the data are easy to interpret and requires the amounts of material comparable to other biophysical techniques. Provided that high field instrument (cryoprobe) is available, the ligand concentration can be as low as 5 µM, thus allowing to overcome the potential problems of ligand solubility or aggregation, possibly encountered when used at high concentration.

Due to the potential problem of signal overlap, this technique may not be suitable to screen a mixture of compounds in a single spectrum. In this technique since to detect the binding interaction, the off rates of the binding kinetics have to be higher than the ligand spin relaxation rates, ligands that may for example have very low affinity e.g nM or less may not be released into solution during the spin relaxation period and hence such interactions may not be possible to detect by this method. In case of very weak affinity interaction (e.g K_D mM), the concentration of [PL] complex in the solution is very low due to the ligand short residence time on the protein, influences the ligand NMR parameters as seen. In order to enhance the response of ligand binding, either the protein concentration needs to be increased or the ligand concentration needs to be reduced. Hence in such cases, this technique may not

be applicable due to limited resources. This technique also does not differentiate between specific and non-specific bonding and allosteric binding interactions.

1.4.2 ^{1}H NMR ligand competition assay for binding studies

The NMR ligand competition experiment is another ligand-observed NMR procedure used for detecting site-specific small molecule binding interactions with proteins. In this process changes in the NMR parameter (such as line-width, peak-height, chemical shift, NOE etc.) of an initially bound ligand observed after the addition of another competitive ligand is used to demonstrate the binding interaction of the competitive ligand with the protein. In this method the initially bound ligand is generally referred as a reporter ligand or spy molecule.

Although in this process any ligand observed method can be used, but the technique based on monitoring the reporter line shape in CPMG edited ^{1}H NMR spectrum is one of the most sensitive methods for ligand screening studies (Figure 1.17 & Figure 1.18).

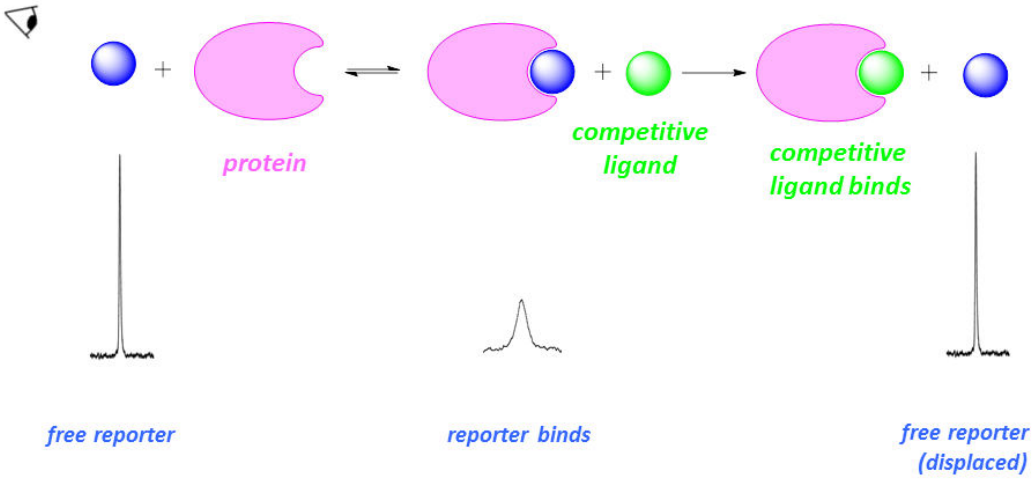


Figure 1.17 Studying protein-ligand binding by ligand-based NMR competition experiments.
This competition experiment is based on monitoring the reporter line-shape after the addition of the competitive ligand to the solution of reporter ligand and protein. From Left to right: Reporter ligand (blue sphere) in the absence of protein, showing typical sharp resonances in the spectrum. After the addition of protein to the sample, the reporter signal is broadened and attenuated demonstrating its fast exchange binding interaction with the protein. The competitive ligand (green sphere) is added to the solution of reporter ligand and protein. The competitive ligand displaces the reporter from its binding pocket into solution as demonstrated by the recovery of reporter's NMR resonances. The displacement of the reporter ligand demonstrates the binding of the competitive ligand to the protein. The signals of the competitive ligand are not considered for binding.

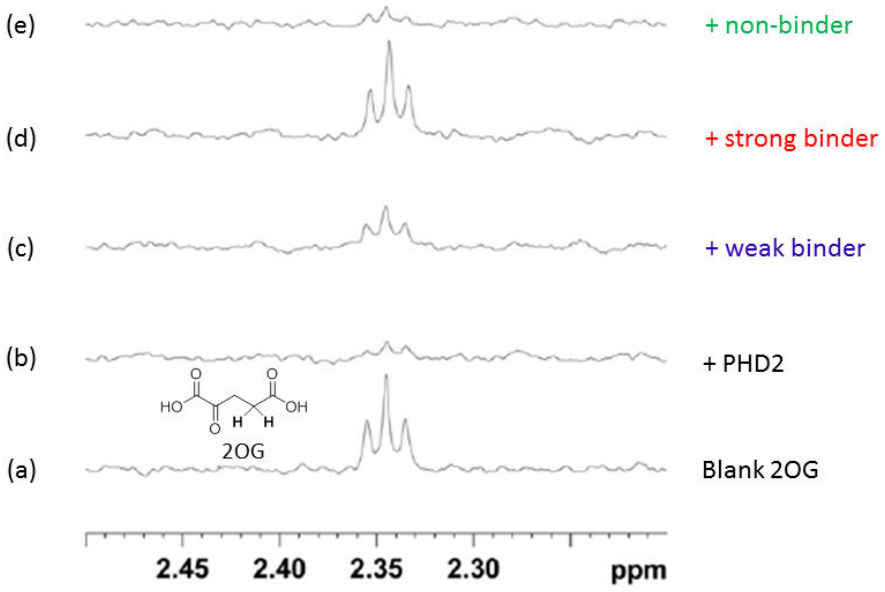


Figure 1.18 Illustration of ^{1}H NMR (CPMG edited)-based reporter ligand displacement assay for ligand binding studies. (a) Spectrum of the reporter molecule 2-oxoglutarate (2OG); triplet (for protons shown as bold) at 2.35 ppm is shown. (b) Spectrum after the addition of Zn^{II}-PHD2 protein. The 2OG signal is broadened/attenuated due to its binding interaction with the Zn^{II}-PHD2 protein. (c) After the addition of a weak competitive binding ligand to (b). The 2OG signal is only slightly recovered. (d) After the addition of a strong competitive binding ligand to (b). The 2OG signal is fully recovered (e) After the addition of a non-binder ligand; no effect on 2OG signal. (Reproduced from reference [130])

In this technique when the competitive ligand is added to the initially bound reporter ligand, then if the competitive ligand binds to the same site as reporter, it will displace the reporter ligand free into solution and as result the reporter NMR resonances will become sharper and recover in intensity[22] (Figure 1.17 & Figure 1.18). Therefore the signals of the competitive ligands themselves are not considered for binding interactions. If the competitive ligand is a non-binder, its addition will have no effect on the NMR resonances of the reporter molecule in the spectrum (Figure 1.18).

Compared to direct ligand observation, this assay has several advantages; for example,

1) Both high and low affinity interactions can be detected
2) The competitor does not need to be ″NMR-active‟
3) Mixtures can be screened for high throughput
4) Systems are optimised for the reporter molecule only
5) Reporter can be isotopically labelled e.g ^{19}F (see below) for clear response.
6) The competitive binding compounds can be easily ranked qualitatively from the level of reporter displacement using a single concentration of the competitive ligand. A high affinity competitive ligand will effectively compete with the reporter, while a

low affinity competitive ligand will weakly compete with the reporter molecule (Figure 1.18).

This assay can also be used to measure the binding dissociation constant of the competitive ligand, by titrating the competitive ligand to a solution containing a fixed concentration of reporter ligand and protein.[130] The recovery (commonly signal intensity) of the reporter NMR resonances is measured as a function of competitive ligand concentration. The titration data points are fitted in appropriate fitting software and the apparent binding dissociation constant value of the reporter molecule is determined. From the knowledge of the actual and apparent binding constant values of the reporter ligand, the binding constant of the competitive ligand can be calculated.[131]

Although, this assay is widely used for ligand screening studies, it also has some limitations; for example, a preliminary knowledge of the location of the binding site of the reporter ligand is essential. Similarly, it may also not differentiate between the reporter displacement site specifically by the competitive ligand and the reporter displacement due to protein conformational changes produced due to the binding of the competitive ligand to an allosteric site (false positive). Also, if the competitive ligand binds non-specifically to the protein, it may not displace the reporter ligand. Therefore allosteric binding of the competitive ligand also cannot be determined by this technique. Furthermore, in case of signal overlap of the reporter ligand with those of competitive ligand, the binding of the competitive ligand also cannot be assessed through this assay.

1.4.3 1D ^{19}F NMR ligand-observed binding experiments

Due to its comparable NMR sensitivity of ^{19}F (0.83) to ^{1}H (1) and 100% natural abundance, 1D ^{19}F NMR ligand-observed experiments are increasingly used to study small molecule binding interaction with a target protein. These experiments require the presence of a fluorine atom in the ligand. The high sensitivity of ^{19}F nucleus to changes in its local chemical environment and high chemical shift range (~ 800 ppm)[66] makes it a better choice over ^{1}H for binding studies.

Most commonly, ^{19}F NMR direct ligand observation and reporter ligand displacement experiments are used for studying protein-ligand binding interactions.

1.4.4 1D ^{19}F NMR Direct ligand observation

^{19}F NMR direct ligand observation based on monitoring the broadening and/or attenuation of the ligand NMR resonances with the addition of protein is an effective NMR technique for studying small molecules binding events with proteins (Figure 1.19). In this technique, the large chemical shift anisotropy (CSA) of ^{19}F nucleus also contributes additional broadening to the broad resonances of a fast exchange binding ligand,[67,132-134] hence results in improved ligand binding sensitivity as compared to ^{1}H NMR direct ligand observation. The broadening of a fast exchange binding ligand further increases with the size of the protein and with the square of the magnetic field strength.[135] High magnetic fields can lead to extremely broad signals with line-width up to ~ 200 Hz. The fluorine signal is usually scalar coupled with the protons and therefore for sensitivity improvement it is useful to record the spectra with proton decoupling during acquisition where possible. Due to its high sensitivity to changes in the local chemical environment, changes in the ligand ^{19}F signal chemical shift can also be used as a probe of its binding interaction with the protein.

One of the most important advantages of 1D ^{19}F NMR ligand based methods over 1D ^{1}H NMR ligand-based techniques is the absence of signal overlap in the spectrum. Therefore this technique allows screening a mixture of fluorinated compounds in a single shot experiment.

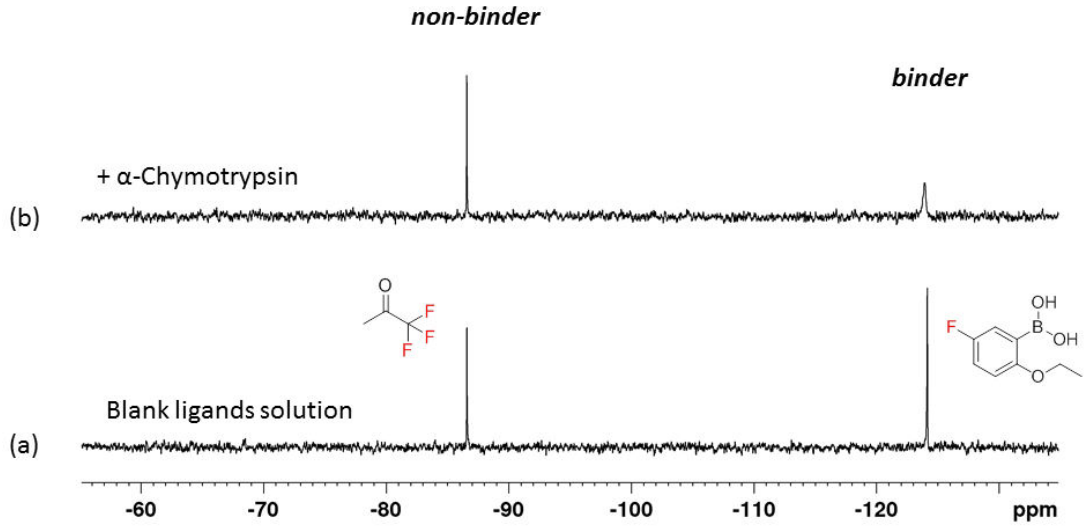


Figure 1.19 Studying protein-ligand binding by ^{19}F NMR direct ligand observation.
(a) ^{19}F NMR spectrum of a solution of 2-ethoxy-5-flurophenylboronic acid (200 µM) and 1,1,1 trifluoroacetone (50 µM) in 50 mM Tris-D_{11}, pH 6.6, in D_2O. (b) After the addition of α-chymotrypsin (200 µM) to (a). The fluorine signal of 2-ethoxy-5-flurophenylboronic acid (-124.1 ppm) is attenuated after the addition of protein, demonstrating its binding interaction with α-chymotrypsin, whereas, the fluorine signal of 1,1,1 trifluoroacetone (-86.5 ppm) remains unaffected, demonstrating that it doesn't bind with α-chymotrypsin.

Similar to ^{1}H NMR, in this technique also two experiments are recorded, one for the ligand(s) solution in the absence of protein and one after the addition of target protein (Figure 1.19). The two spectra are compared for signal attenuation and/or broadening as a probe of binding interactions. To the assay mixture a fluorinated internal NMR control molecule is also necessarily added, whose line-shape and chemical shift will remain unaffected after the addition of protein.

Ligand binding dissociation constants can also be measured by this technique, usually by titrating protein to a solution containing a fixed ligand concentration. The attenuation in the ligand signal intensity is measured as a function of protein concentration. The titration data points are fitted and the ligand binding constant is measured. Hits can also be ranked qualitatively from single shot experiment based on the level of broadening and/or attenuation using a single concentration.

The second most interesting and valuable information that is obtained from ^{19}F NMR direct ligand observation and which is not usually possible in ^{1}H NMR direct ligand observation is the observation of the signal of bound population of a binder ligand in the spectrum in NMR slow exchange regime (Figure 1.20).[136-139] Since in this spectrum, the concentrations of free ligand and protein-ligand complex are known, therefore the binding dissociation constant of the ligand can be readily calculated using equation 2 as described in section 1.2.2.[138]

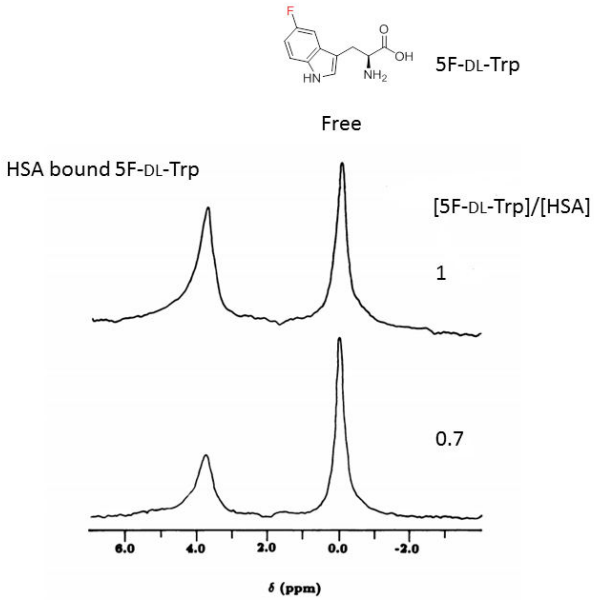


Figure 1.20 Illustration of NMR slow exchange ligand binding system as observed by ligand-based ^{19}F NMR. Monitoring the binding of 5F-DL-Trp to human serum albumin (HSA) protein by ^{19}F NMR direct ligand observation. The spectra show the signals of both free and HSA bound 5F-DL-Trp in a NMR slow exchange ligand binding system. (Reproduced from reference [138])

Another advantage of the ^{19}F NMR observation over ^{1}H NMR observation is the possibility to perform the screening even in the presence of protonated solvents and buffers instead of their deuterated forms. The deuterated buffers are usually much more expensive compared to their non-deuterated forms. Hence the ^{19}F NMR direct ligand observation is economical over ^{1}H NMR direction ligand observation experiments.

1.4.5 1D ^{19}F NMR reporter-ligand displacement assay for ligand binding studies

1D ^{19}F NMR reporter-ligand based displacement experiments based on monitoring the recovery of reporter NMR resonances after the addition of competitive ligand are also commonly used for detecting site-specific ligand binding interaction with the proteins (Figure 1.21).[26,132,140,141]

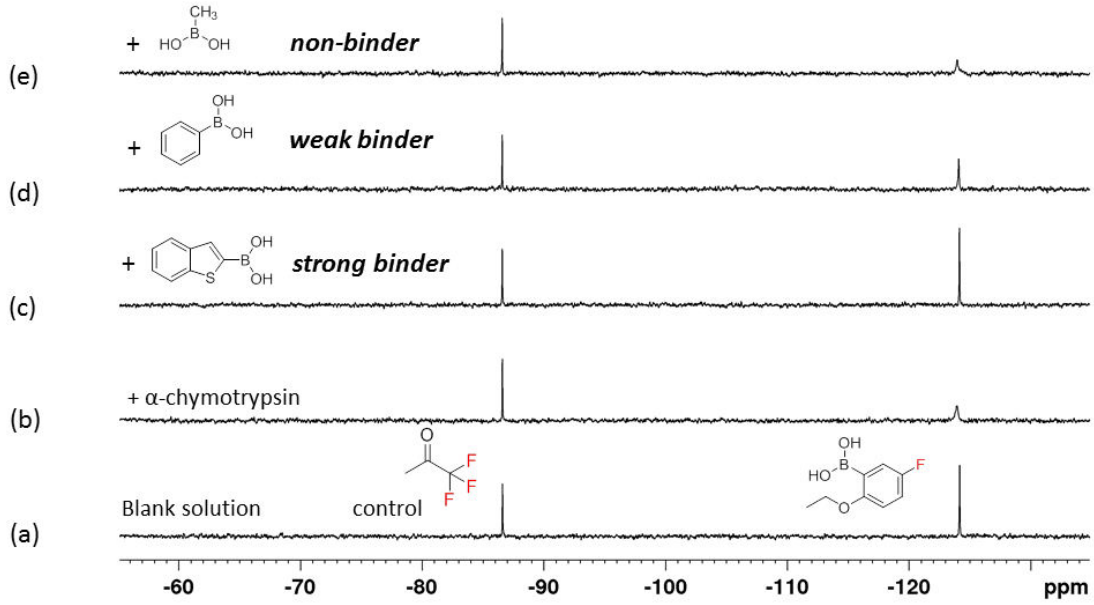


Figure 1.21 ^{19}F NMR reporter ligand displacement assay for ligand screening.
(a) ^{19}F NMR spectrum of a solution of 2-ethoxy-5-fluoro-phenylboronic acid and 1,1,1 trifluoroacetone in 50 mM Tris-D_{11}, pH 6.6, in D_2O. 2-Ethoxy-5-fluoro-phenylboronic acid is used as a reporter molecule and 1,1,1 trifluoroacetone is used as an internal NMR standard. (b) The reporter signal is broadened/attenuated with the addition of α-chymotrypsin due to its binding interaction. (c) The addition of benzo[b]thien-2-ylboronic acid to (b) results in full recovery of signal of the 2-Ethoxy-5-fluoro-phenylboronic acid demonstrating that benzo[b]thien-2-ylboronic acid is a stronger binder of α-chymotrypsin binding in the 2-ethoxy-5-fluoro-phenylboronic acid binding site. (d) The addition of phenyl boronic acid to (b) results in a small recovery of signal of 2-ethoxy-5-fluoro-phenylboronic acid, demonstrating that phenyl boronic acid is a weak binder of α-chymotrypsin. (f) The addition of methyl boronic acid to (b) has no effect on the signal of 2-ethoxy-5-fluoro-phenylboronic acid, suggesting that methyl boronic acid doesn't bind α-chymotrypsin. The concentrations used are 2-ethoxy-5-fluoro-phenylboronic acid 200 µM, 1,1,1 trifluoroacetone 50 µM, α-chymotrypsin 200 µM and competitive boronic acid ligand 750 µM.

Generally, in this technique a spy molecule containing a CF_3 moiety is preferred because of the advantage of high sensitivity of fluorine signal from chemically equivalent three fluorine

atoms hence allowing the reporter ligand to be used in lowest possible concentration in the assay.

A slow exchange reporter ligand binding system is particularly more useful in this technique. In such systems, the signal of the bound population of the reporter ligand can also be used as a probe (in addition to monitoring the signal of free ligand in solution) to detect the binding interaction of the competitive ligand with the protein (Figure 1.22).[138] If the competitive ligand binds to the same site as the reporter ligand, then the signal of the bound population of the reporter ligand would be decreased in intensity due to displacement of the reporter ligand from its binding pocket into free in solution. From this experiment the apparent binding constant of the reporter ligand can be directly calculated using equation 2 (section 1.2.2). From the knowledge of the actual and apparent binding constant values of the reporter ligand, the binding constant of the competitive ligand can be calculated.[131] This technique offers one of the most efficient NMR assays for screening ligands and measurement of their binding dissociation constant from a single shot experiment.

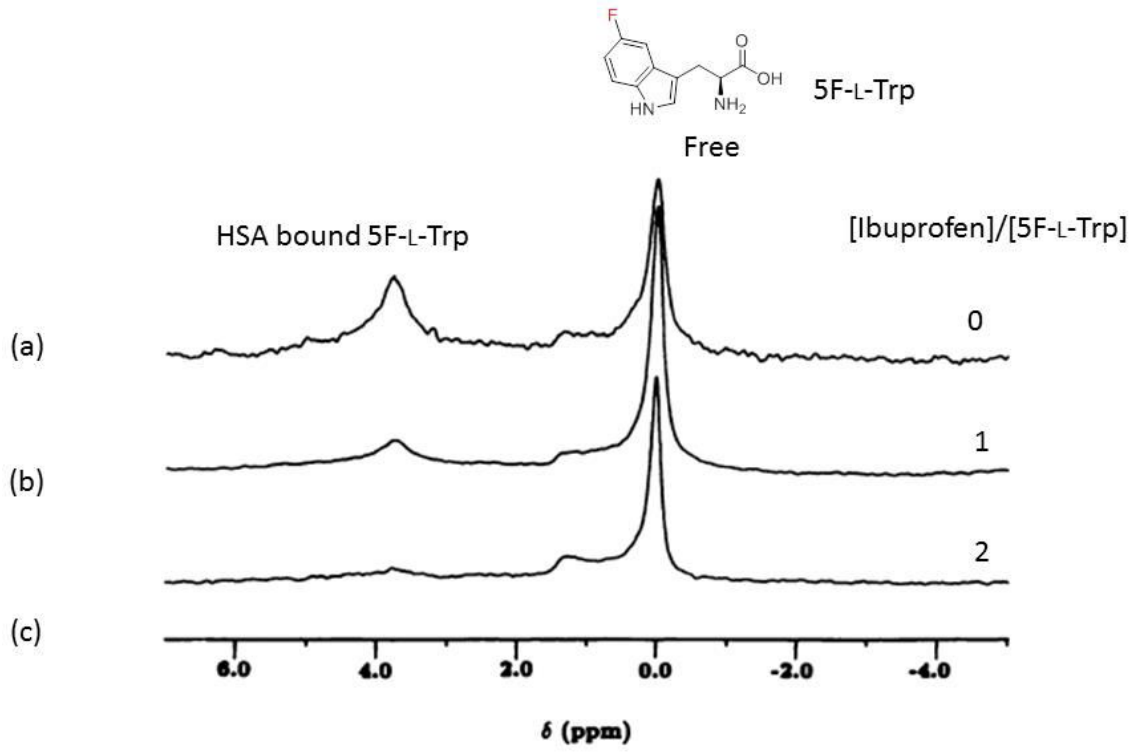


Figure 1.22 ^{19}F NMR competition assay with reporter ligand in slow chemical exchange.
^{19}F NMR spectra demonstrating the binding interaction of 5F-L-Trp and ibuprofen to HSA (0.730 mM). (a) Binding of 5F-L-Trp to HSA by NMR direct ligand observation in NMR slow exchange binding regime. The signals of both free and HSA-bound 5-F-L-Trp are seen. (b-c) The addition of competitive ligand ibuprofen in increasing amounts to the sample of (a) results in displacement of 5-F-LTrp as demonstrated by decrease in intensity of its bound population signal and concomitant increase in intensity of its free population signal. (Reproduced from reference [138])

The large chemical shift anisotropy (CSA) of ^{19}F nucleus which contributes additional broadening to the broad resonances of a fast exchange binding ligand[135] is particularly more useful for reporter displacement experiment using the CPMG edited ^{19}F NMR in the so-called technique of FAXS[126,141] (fluorine chemical shift anisotropy and exchange for screening). The CPMG pulse-sequence improves the NMR sensitivity to reporter binding by enhancing the attenuation of its broadened resonances, and thus increases the difference in ligand signal intensity between its bound and free forms. This technique is widely used for screening of mixture of large chemical libraries in pharma industries (Figure 1.23).

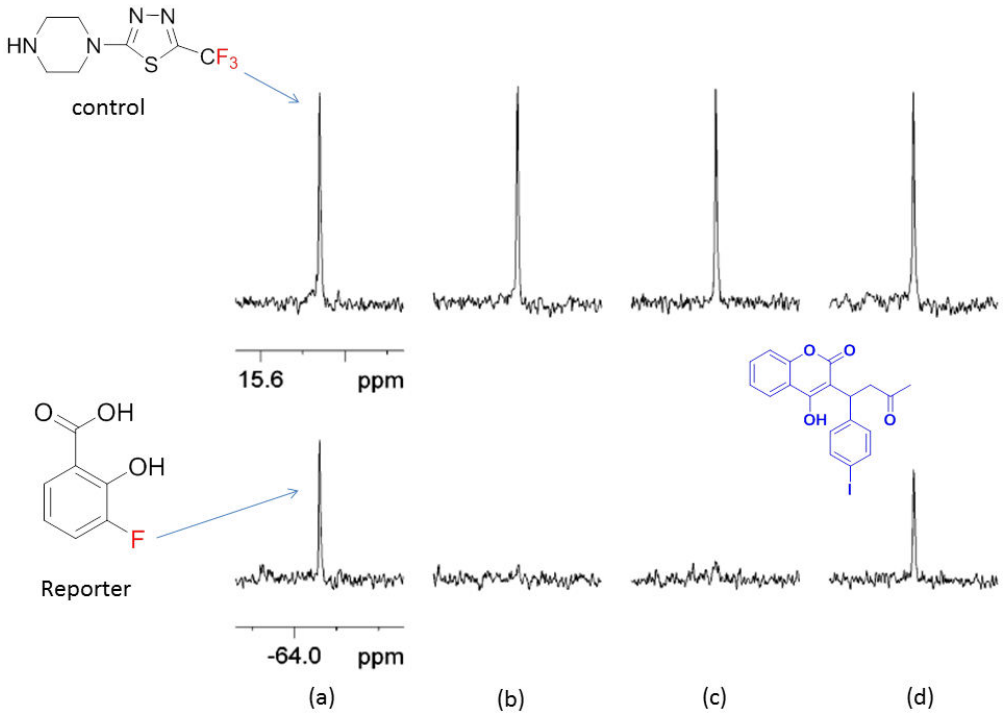


Figure 1.23 Illustration of the FAXS NMR technique for ligand screening.
In this assay 2-hydroxy 3-fluorobenzoic acid (bottom structure) is used as a spy molecule and 1-[5-(trifluoromethyl)1,3,4-thiadiazol-2-yl]piperazine (top structure) as a control molecule. Only parts of the spectra are shown for clarity. The spectra were recorded with a total spin-echo period of 160 ms. The concentration of spy molecule and control was 50 and 25 µM, respectively. (a) Spectra of the reporter and control in the absence of protein (b) Spectra of the reporter and control in the presence of 600 nM HSA. (c) Spectra of the reporter, control, HSA, 50 µM 5-CH_3 DL-Trp and sucrose. (d) Spectra of the reporter, control, HSA, 50 µM 5-CH_3 DL-Trp, sucrose and 4-hydroxy-3-[1-(p-iodophenyl)-3-oxobutyl]coumarin (blue colour, a warfarin derivative). The results of (c) identifies 5-CH_3 DL-Trp and sucrose as non-binders, whereas the results of (d) identifies 4-hydroxy-3-[1-(p-iodophenyl)-3-oxobutyl]coumarin as a hit. The chemical shifts are referenced to trifluoroacetic acid.(Reproduced from reference [141])

For a reporter ligand in fast exchange binding regime, the apparent binding constant can be determined by competitive ligand titration and measuring the recovery of the reporter resonances.[130] The competitive ligand can also be qualitative ranked on the basis of single shot displacement experiment from the level of reporter displacement at single concentration.

1.4.6 Ligand-Nuclear Overhauser effect (NOE)-based NMR binding techniques

Some NMR experiments based the ligand NOE effect are also employed to study protein-ligand binding interactions. These include transferred NOE (tr-NOE),[11-13,20,142] saturation transfer difference (STD)[19,20] and water ligand-observed *via* gradient spectroscopy (waterLOGSY) experiments.[20,21] The detection of binding interactions by these methods primarily relies on the fast exchange of the ligand between the bound and the free forms and the rapid transfer of magnetisation within proteins *via* spin diffusion that is mediated by the NOE. These techniques are discussed below.

1.4.7 Transferred NOEs (tr-NOE)-based binding technique

Small or medium molecular-weight ligands (MW <1000–2000 Da) free in solution have a short correlation time t_c and thus exhibit small and positive intra-ligand NOEs, no NOEs, or very small negative NOEs, depending on their molecular weight, shape, and magnetic field strength. In contrast, large molecules (MW ~ >2000 Da) generally have longer rotational correlation time and thus exhibit strong negative NOEs (Figure 1.24). As a rule-of-thumb, the correlation time of a molecule can be estimated from its molecular weight i.e t_c [ps] > 0.5 Da (e.g. a MW of 1000 Da corresponds to ca. 0.5 ns), however, other factors also influence a molecule correlation time, for example, temperature (the higher the temperature, the shorter t_c), solvent viscosity (the more viscous, the longer t_c).

In the transferred NOE (tr-NOE) binding technique, the binding interaction of a ligand with a protein is detected by the observation of the NOE that generates within the ligand (intra-ligand NOE) whilst bound to the protein. These change sign for bound state and may differ from ligand NOEs if conformation changes on binding. The binding interaction of a ligand to a protein can thus simply be detected by looking at the sign and size of the observed tr-NOEs in the spectrum. The tr-NOE signal detected in the spectrum is an average ligand NOE between its free and bound states. In this technique either 1D or 2D transient NOE experiments may be recorded to monitor the binding interaction of a ligand to a target protein. The detection of binding interactions by this technique can be viewed as a heat transfer process. For example, if a ligand bind to a hot object (protein), then upon binding, this heat is transferred from hot object (protein) onto the bound ligand through spin diffusion and when the ligand dissociates from the hot object (protein), it takes this heat (NOE) into solution, which is detected as a response of its binding interaction with the protein.

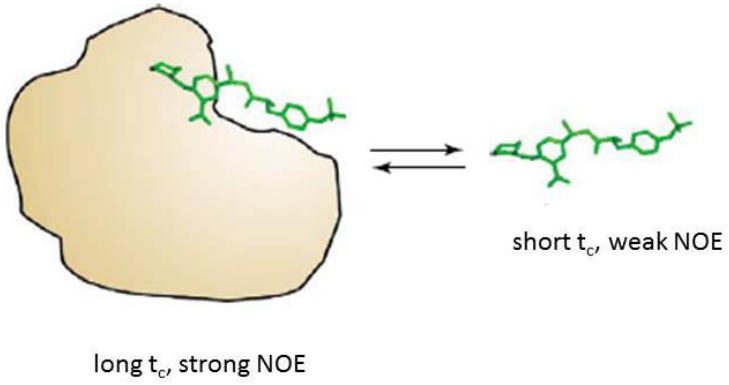


Figure 1.24 Cartoon representation of transferred NOE (tr-NOE) ligand-observed NMR binding technique.

This ligand-based NMR technique is however not commonly used for studying protein-ligand binding interactions due to the requirements of its stringent experimental conditions such as large excess of ligand concentration relative to the protein concentration, longer experiment time, false positives. Furthermore, this technique is also very sensitive to correct on/off rates of the ligand to give good NOE response.

1.4.8 Saturation-Transfer-Difference (STD) NMR Spectroscopy

Saturation transfer difference (STD) NMR (Figure 1.25) is also an NOE-based ligand-observed technique applied to study protein-ligand binding interaction (Figure 1.26 & Figure 1.27).[19,20] As the name indicates, this technique is based on the transfer of saturation from protein onto bound-ligand, which is then detected as a response of binding interaction in the signals of the free ligand in solution. In this technique, the protein resonances are irradiated at a selected region in the spectrum (where their intensity is high e.g methyl groups region, 0 to – 1 ppm and sufficiently far away (e.g couple of hundred Hz) from the closest ligand signal) for a defined period of time (typically 1-5 s) in order to saturate. This ^{1}H NMR spectrum is called an on-resonance spectrum. By spin diffusion through the NOE, all protein resonances in the spectrum also undergo rapid saturation, which in turn is passed onto the bound ligand. Hence, if a ligand binds to the protein, then this magnetisation can also be transferred intermolecularly to the bound ligand *via* spin diffusion in the on-resonance experiment. This transfer of magnetisation to the ligand also results in partial saturation of ligand resonances. The degree of ligand saturation depends on the ligand residence time inside the protein binding pocket. Thus, upon dissociation of the ligand from protein into solution, this partial saturation received can be detected as a reduction in its signal intensity in the spectrum.

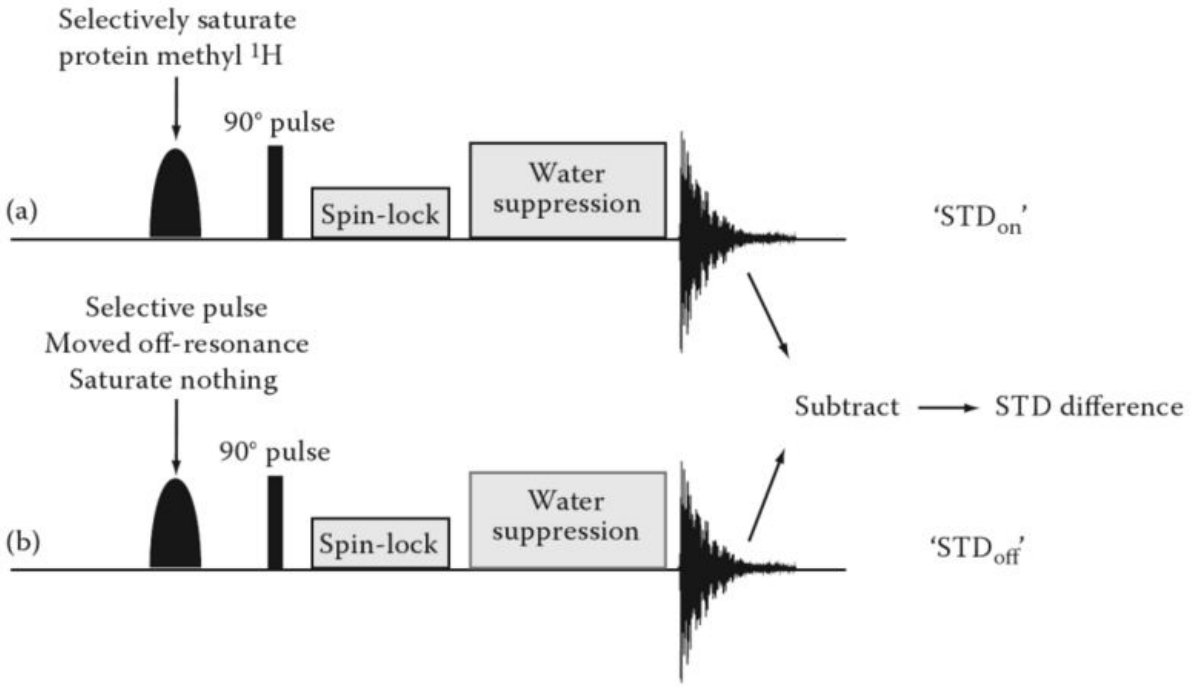


Figure 1.25 General pulse-sequence for a 1D ^{1}H STD NMR experiment.
The sequence involves the recording of on-resonance and off-resonance spectra. Subtraction of on-resonance spectrum from off-resonance spectrum produce STD difference spectrum. In order to achieve the desired selectivity and to avoid any side-band irradiation, shaped pulse is used for the saturation of the protein signals. The water suppression is required for the samples made in aqueous media. The optimised spin lock is used to attenuate the broad background protein resonances to provide a relatively flat base-line for ligand resonance detection. (Reproduced from reference [143])

Next, a region of the spectrum far away (e.g 40 ppm) from the protein and ligand resonances is irradiated (basically a standard ^{1}H NMR). This ^{1}H NMR spectrum is called an off-resonance spectrum, also a ˜control ″or˜reference″spectrum. Subtraction of the on-resonance spectrum from the off-resonance spectrum produces a STD difference spectrum in which signals of the ligand that are attenuated by partial saturation transfer appear (Figure 1.26 & Figure 1.27). Other small molecules that may be present in the assay mixture but which do not bind to the protein will not receive magnetisation transfer from protein (hence no partial saturation) and therefore their signals will be of equal intensity in both the on-resonance and the off-resonance spectra and, consequently after subtraction no signals will appear in the STD difference spectrum from the non-binding ligands (Figure 1.26 & Figure 1.27). In this technique it is highly important to irradiate the region of the protein resonances which is far away from the ligand resonances to avoid generating false positive results arising from direct saturation of the ligand resonances.

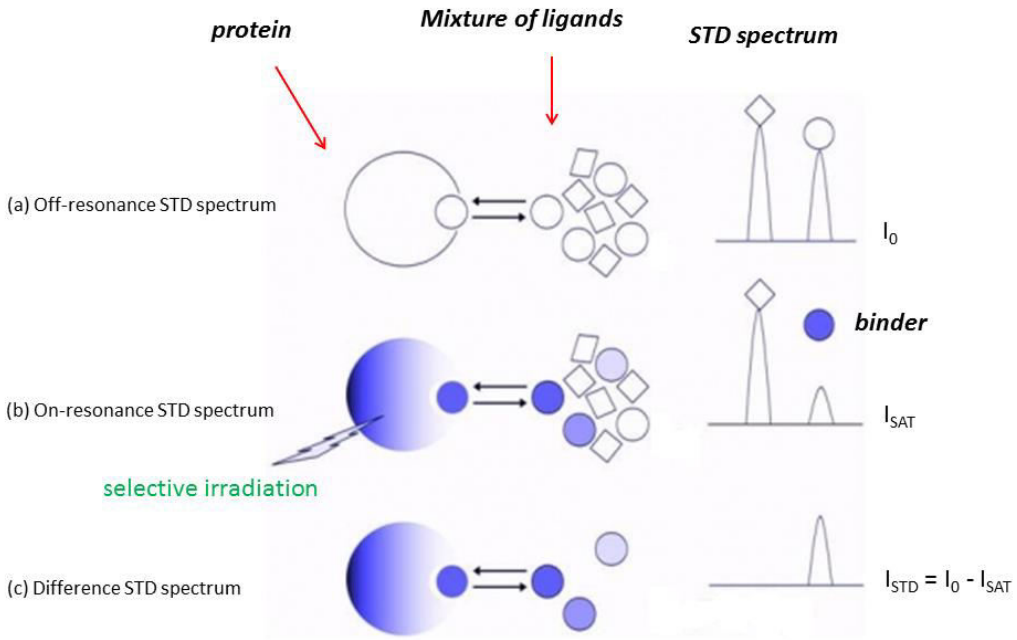
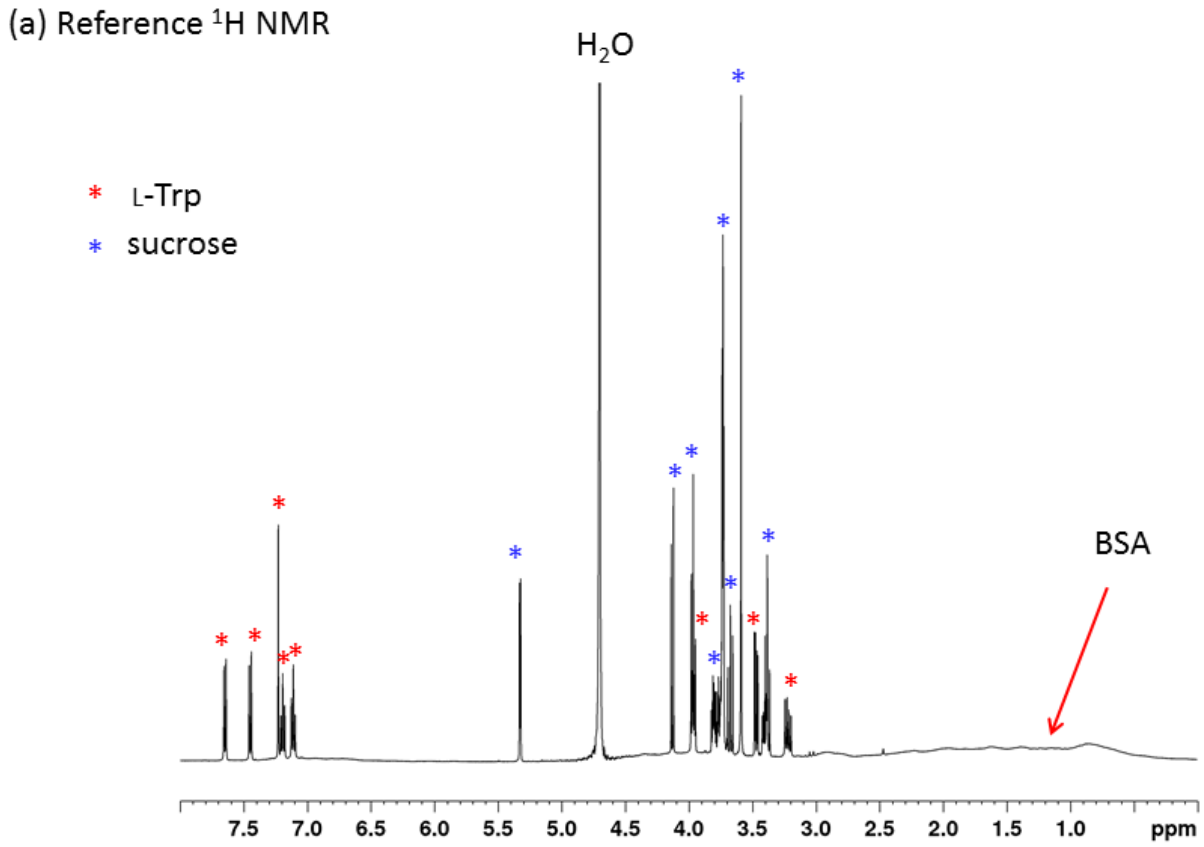


Figure 1.26 Cartoon representation of the STD NMR binding technique.
I_0, I_{SAT} and I_{STD} represent the ligand signal intensity in the off-resonance, on-resonance and difference-STD spectrum respectively.



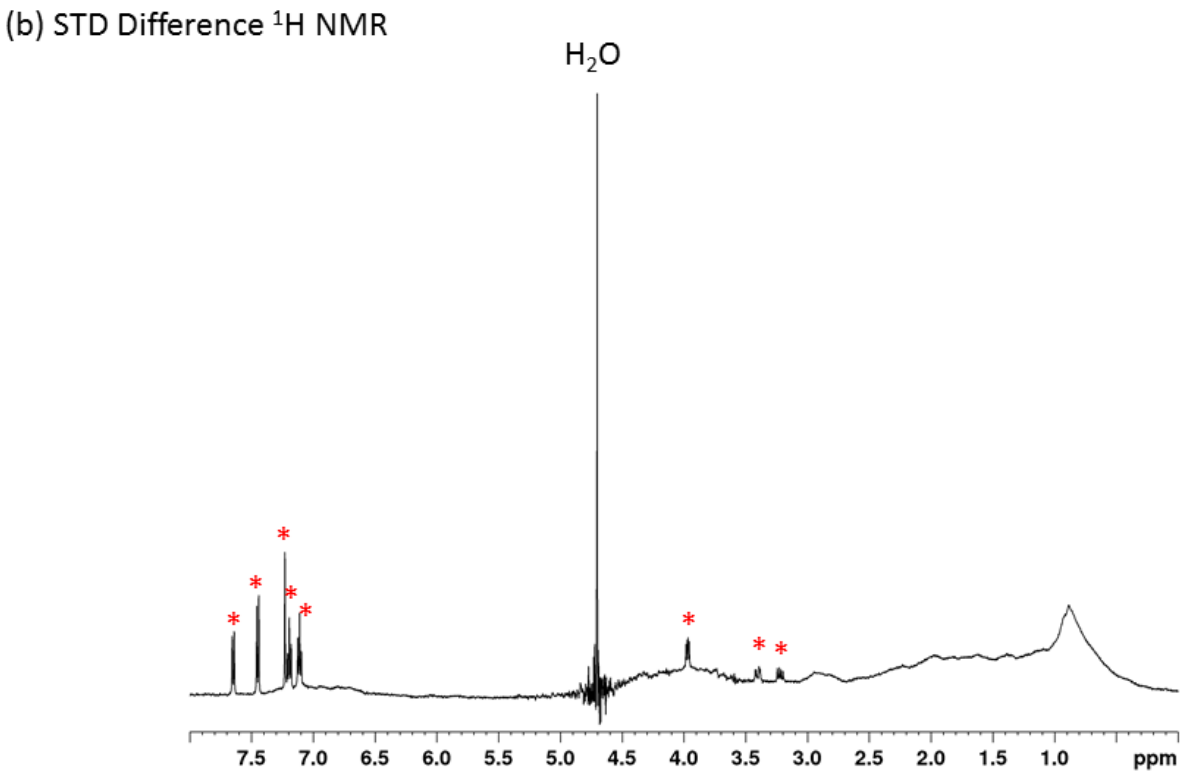


Figure 1.27 Example of protein-ligand binding by STD ^{1}H NMR.
Monitoring the binding of L-Trp and sucrose to BSA in a mixture by STD NMR technique. (a) Reference ^{1}H NMR spectrum. The red and blue asterisks represent the signals of L-Trp and sucrose respectively. (b) STD difference ^{1}H NMR spectrum. In the STD difference spectrum, only the signals of L-Trp are seen, suggesting that only L-Trp binds to BSA. The assay mixture contained 10 mM L-Trp, 10 mM sucrose, 100 µM BSA in D_2O. The NMR experimental conditions were; selective on-resonance irradiation was applied at 0.86 ppm and off resonance irradiation was applied at 50 ppm. The saturation time used was 2 seconds and the relaxation delay used was 2 seconds.

The binding interaction by STD NMR is usually quantified as,

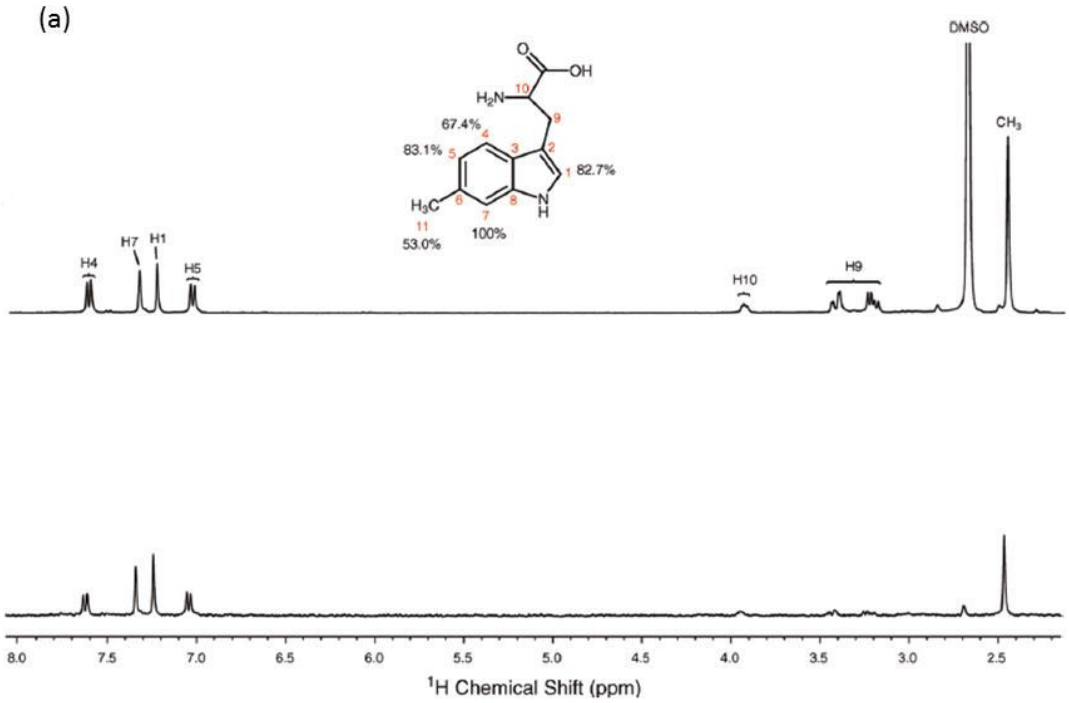
$$I_{STD}\ (\%) = (I_0 - I_{SAT} / I_0) \times 100$$

Where, I_0 is the intensity of one signal (of ligand) in the off-resonance spectrum, I_{SAT} is the intensity of the corresponding signal in the on-resonance spectrum and I_{STD} denotes the intensity of the corresponding signal in STD difference spectrum. The STD binding effect is sometimes also described by the STD amplification factor (A_{STD}),[20] obtained by multiplying the relative STD effect (I_{STD}) of a given hydrogen at a given ligand concentration ($[L_T]$) with the molar ratio of ligand in excess relative to the protein $[L_T]/[P_T]$.

$$A_{STD} = (I_{STD}) \times [L_T]/[P_T]$$

Where L_T and P_T represents the total concentration of the ligand and total concentration of the protein used in the experiment. For a given saturation time, the A_{STD} can also be defined as an average number of ligand molecules saturated per molecule of protein.

In this technique, the saturation of the protein occurs very fast (~ 100 ms). Therefore ligands with only fast off-rates can transfer this saturation into solution. Thus if a large excess of ligand is used over protein, then one binding site can saturate many ligand molecules in few seconds. Ligands in solution loose their information by T_1 relaxation processes, which for small molecules are usually in the order of about seconds. Thus during the saturation time, the level of the saturated ligands increases in solution and so as the binding information resulting from protein saturation is also amplified. The intensity of the STD signal depends on the length of irradiation time and excess of ligand concentration over protein. Large excess of ligand concentration and longer irradiation time leads to stronger STD signal (Figure 1.28). In general a 100 fold ligand excess has been found to produce a good STD binding response. The large ligand excess allows the progressive build-up of the ligands carrying saturation into solution enhances the STD signal. From the higher ligand: protein ratios, it is evident that only a low amount of protein is used in the assay.



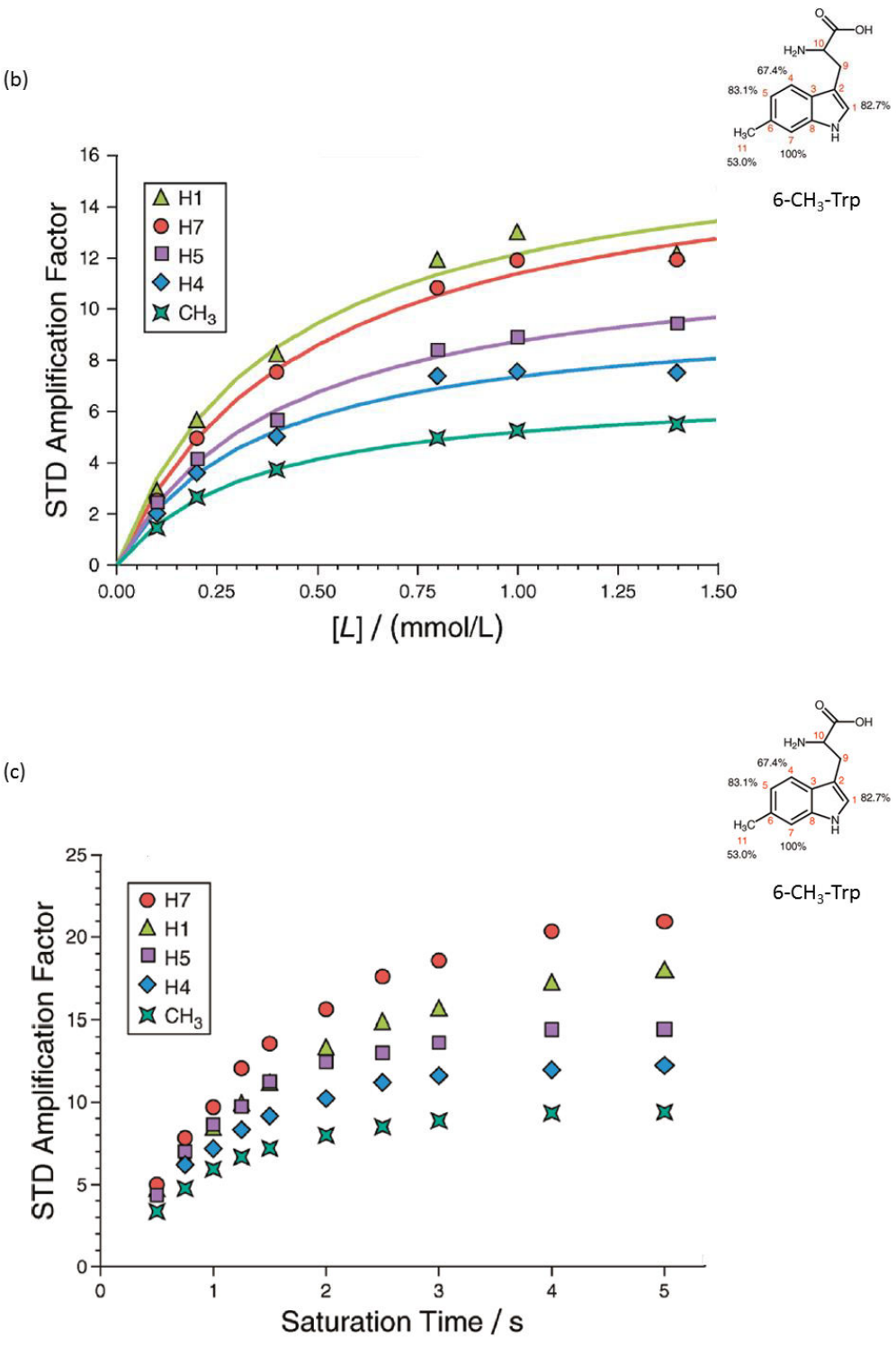


Figure 1.28 Demonstration of the effect of ligand concentration and saturation time on the STD effect.
Example shown is the binding of DL-6-CH_3-Trp with HSA.. (a) Top; reference ^{1}H NMR spectrum of DL- 6-CH_3-Trp (2.0 mM) with HSA (20 μM). Bottom: STD NMR spectrum of DL- 6-CH_3-Trp with HSA (saturation time 1.50 s). The relative STD effects of the individual hydrogens are normalized against hydrogen H7. (b) The effect of ligand concentration on the STD response. High ligand concentration is required to enhance the STD effect (c) The effect of saturation time on STD response. From both (a) and (b) it is also apparent that different hydrogens of the ligand produce different STD response, this is possibly due to their proximity in the protein binding pocket (described below in the text). (Reproduced from reference [144])

In this technique the response of binding relies on the fast exchange of the ligand between the free and bound form, hence if the binding interaction is very tight and consequently the off rates are in the range 0.1- 0.01 Hz (K_D < 1 nM),[19] the saturation transfer to the bulk ligand in solution will not be very efficient and hence such a tight binding interaction may not be detected by this technique. In generally, if the K_D values are 100 nM or larger, the fast exchange of free and bound ligands results in very efficient build-up of the saturation of the ligand molecules in solution. Similarly, the STD technique may also not detect very low affinity interaction because the residence time of the ligand inside the protein binding pocket is too short for the transfer of magnetisation to occur from the protein to the bound ligand.

In addition to detecting the protein-ligand binding interaction, the STD NMR technique is also particularly useful is identifying the ligand ″group epitope mapping‟, meaning which chemical moieties of the ligand are important for molecular recognition in the binding site (Figure 1.28 & Figure 1.29). According to the NOE theory, the magnetisation transfer from protein to ligand protons by intermolecular NOE depends on the inverse sixth power of their distance in the bound state. Thus, the shorter the protein-ligand proton-proton distance (bound state) the stronger the intensity of the corresponding STD signal. So by normalising all the measured STD intensities (arbitrarily assigned a value of 100%) against the most intense signal the″group epitope mapping‟ is obtained (in percentage). The binding epitopes are generally obtained for a given saturation time with the assumption that the resulting ligand‟ group epitope mapping″ is independent on the chosen saturation time. However, significantly different R_1 relaxation rates of the ligand protons can generate artefacts in the‟ group epitope mapping″ definition. Especially, for protons with slower R_1 enables a more efficient accumulation of saturation in solution such that their relative STD intensities may be significantly overestimated at long saturation times, thus also overrating the proximity of those protons to the protein surface. Indeed, the structural information that the binding epitope provides can be influenced by several factors including the differences in the R_1 relaxation rates of the ligand protons, the extent of saturation received in the first place and the kinetics of binding.

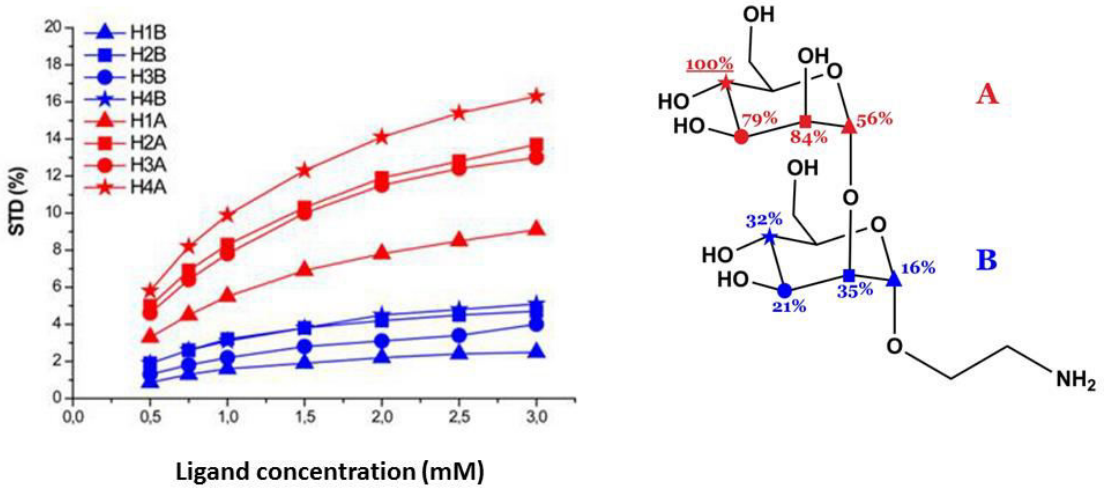


Figure 1.29 Illustration of ligand binding epitope mapping by STD NMR.
Left- STD build up rates curves for the binding of Man-α(1,2)-Man-α-O(CH_2)NH_2 (right) with the anti-HIV-1 human antibody 2G12. Right: Binding epitope for Man-α(1,2)-Man-α-O(CH_2)NH_2 ligand. Note that proton H_4A (marked with star) receives the highest saturation (100%; arbitrarily assigned) and is used as reference to calculated the STD percentages of the other protons. The results of the STD build up rates for the different protons in the ligand suggest that **A** part of the ligand makes the main contacts with the protein in the bound state. (Reproduced from reference [145])

It may also be possible to determine a ligand binding dissociation constant value by STD NMR technique by titrating ligand to a fixed protein concentration, monitoring the STD growth rates as a function of ligand concentration.[20]

Although the STD technique is useful to study small molecule binding events to proteins, but it still has some limitations, such as its dependency on ligand high dissociation rates, spin relaxation time and false positives due to nonspecific binding or ligand aggregation.

1.4.9 Water-ligand observed *via* gradient spectroscopy (waterLOGSY) NMR

The water-ligand observed *via* gradient spectroscopy (waterLOGSY) NMR experiment is another NOE-based 1D ligand-observed technique widely applied for studying protein-ligand binding interaction.[21,146] WaterLOGSY begins with the selective irradiation (strictly inversion) of bulk water by radio-frequency pulses; this magnetisation then transfers (through spin diffusion) to the bound ligand *via* the water–protein–ligand complex and upon dissociation from the protein, the ligand carries this magnetisation into solution as is detected in its signals in the form of large negative NOE (Figure 1.30 & Figure 1.31). In contrast, small molecules that only interact with bulk water and not the protein will experience much faster tumbling, which translates into a positive NOE. Therefore, binders and non-binders can be easily differentiated as opposite-phased resonances in the waterLOGSY spectrum.

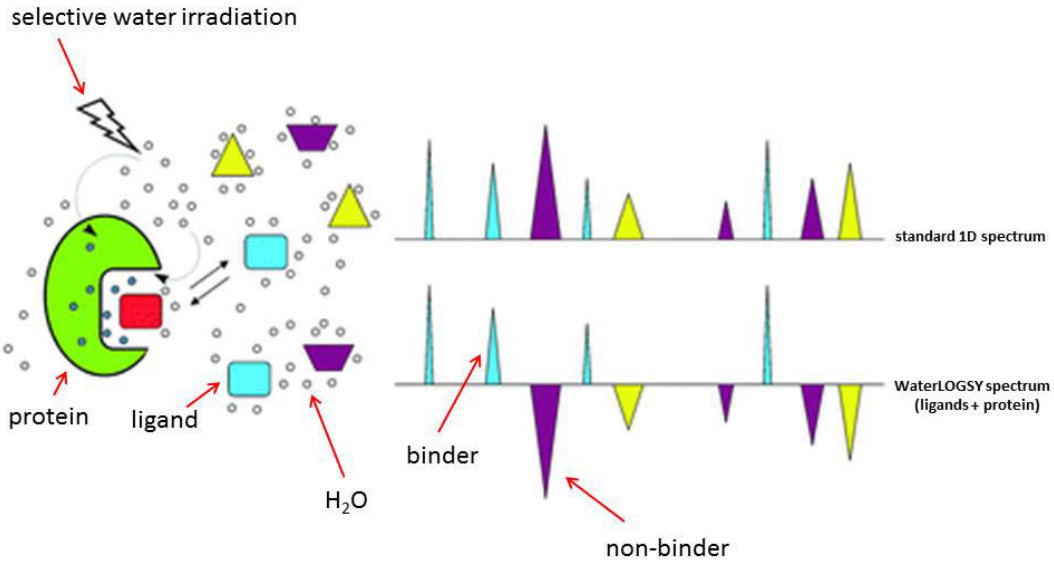
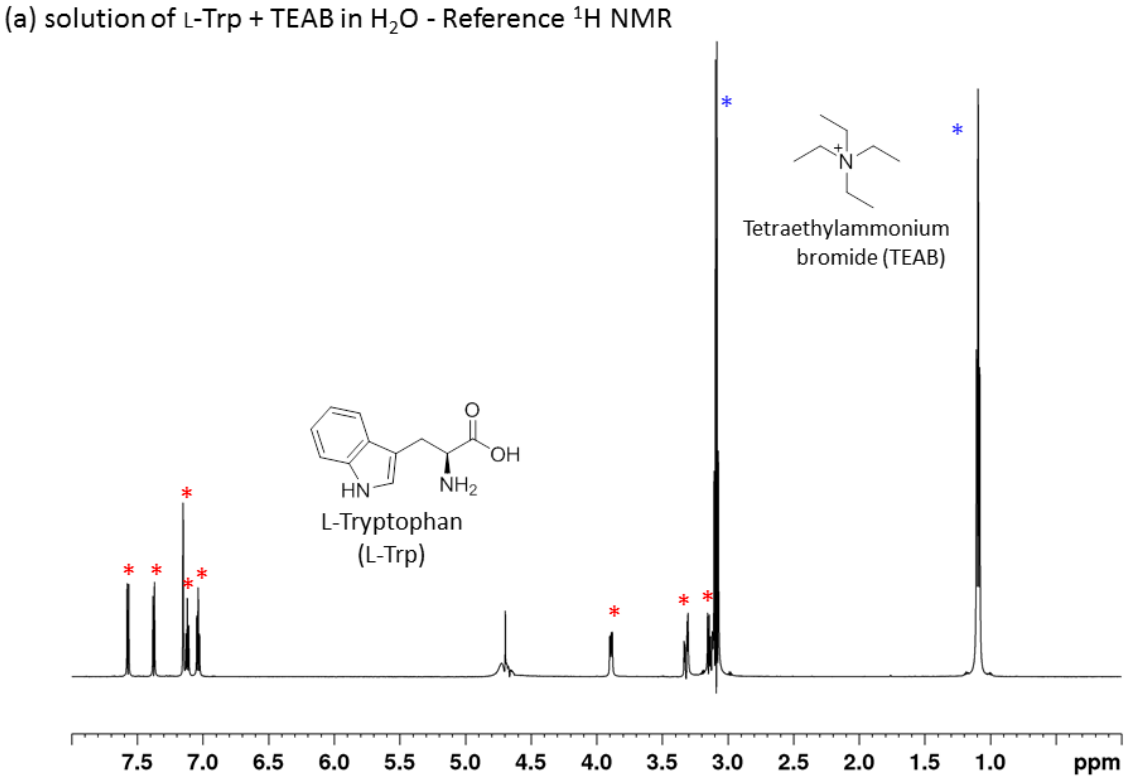
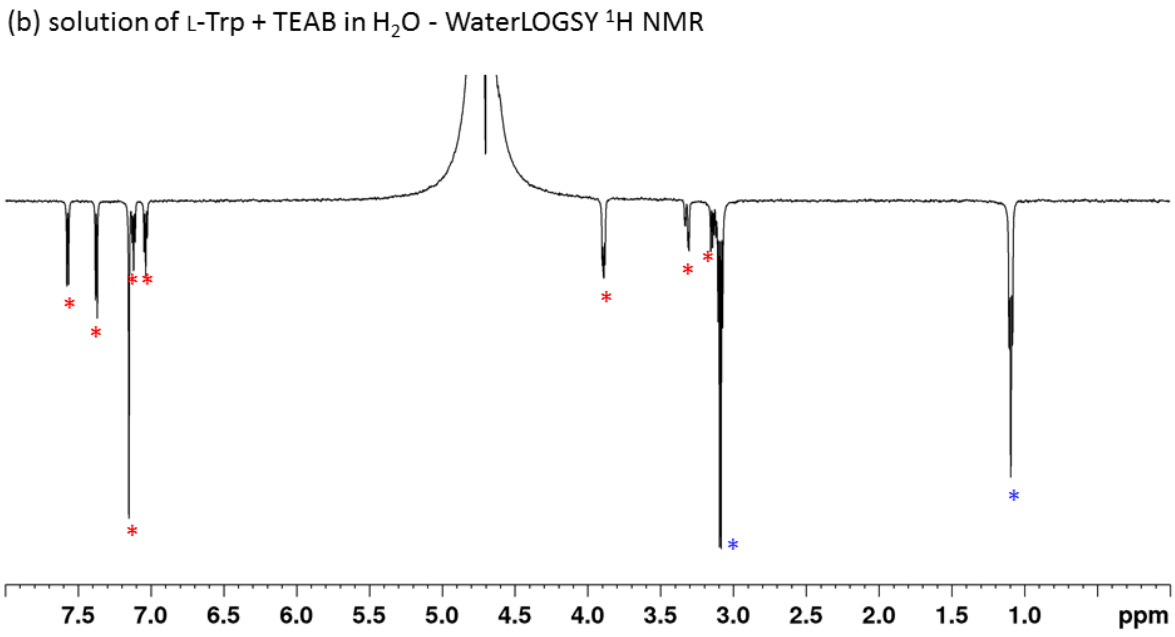


Figure 1.30 Illustration of the waterLOGSY NMR binding technique.
(Left): is a mixture of a protein (green object) and three small molecules (cyan, blue-violet, yellow) in water (small spheres). The bulk water is first selectively irradiated with radio-frequency pulses. This magnetisation then transfers to the protein and then to the bound-ligand (red coloured; it is the cyan colour ligand, redness denote the magnetisation it receives from the protein). This ligand upon dissociation from protein takes this magnetisation into solution and is detected in its NMR resonances in the form of large negative NOE (from protein-ligand complex). The blue-violet and yellow small molecules do not receive magnetisation from protein. (Right); binder (cyan) and non-binders (blue-violet and yellow) are seen as opposite-phased resonances in the waterLOGSY spectrum. (Reproduced from reference [147])





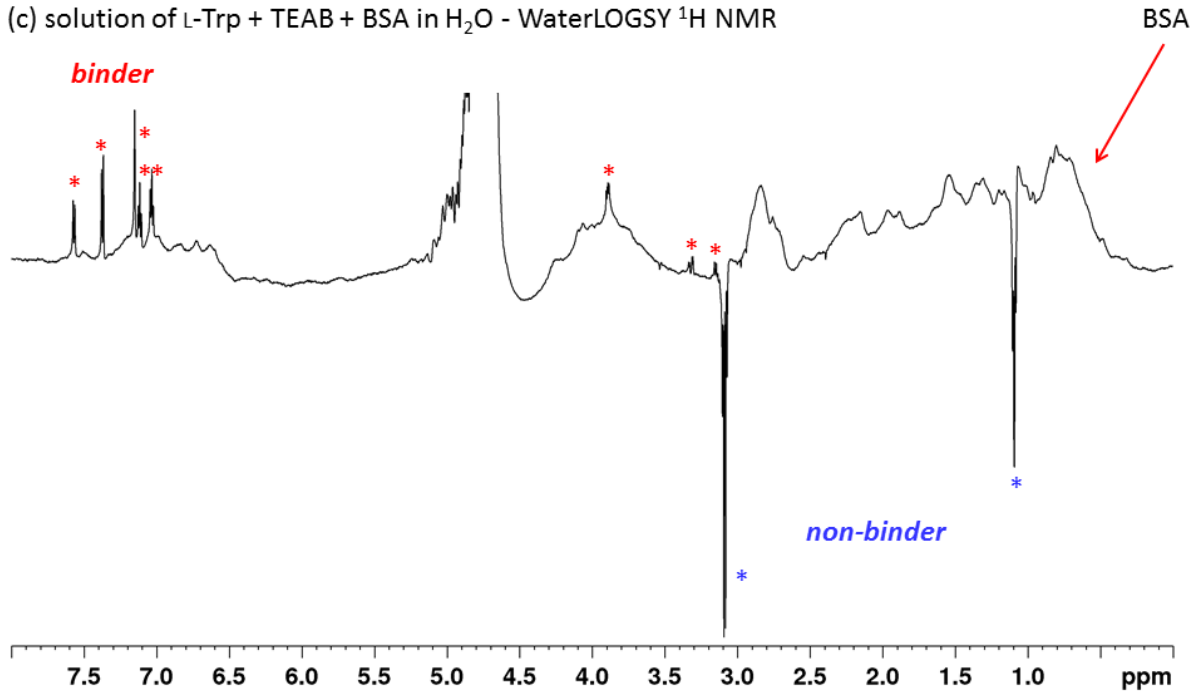


Figure 1.31 Example of protein- ligand-binding by WaterLOGSY ^1H NMR.
Binding of L-tryptophan with BSA as demonstrated by WaterLOGSY ^1H NMR technique. The assay mixture contained 2 mM L-Trp, 2 mM tetraethylammonium bromide (TEAB), 100 µM BSA in H_2O (10% D_2O). Selective on-resonance irradiation was applied at 4.7 ppm. The mixing time was 1 second and the relaxation delay was 2 seconds. (a) Standard ^1H NMR spectrum of a solution of L-Trp (signals with red asterisks) and tetraethylammonium bromide (TEAB) (signals with blue asterisks). (b) WaterLOGSY spectrum of (a). (c) WaterLOGSY spectrum after the addition of BSA to the sample of (a). Note that in this spectrum, the signals of the L-Trp and TEAB are opposite phased. Comparison of this spectrum with (b) shows that L-Trp has received the large negative NOE from BSA, hence is a binder, whereas the TEAB has not and is a non-binder in the mixture.

The waterLOGSY signal detected for the binding ligand is an average of its NOE between the bound form and free form (Figure 1.30 & Figure 1.31). In this technique since bulk water is used as the source for the transfer of magnetisation, the assay solution is therefore prepared in 90% H_2O and 10% D_2O (for locking).

In this technique first a reference standard 1D and a WaterLOGSY spectrum of the ligands in the absence of protein are recorded. A second WaterLOGSY spectrum is recorded following the addition of protein. The two WaterLOGSY spectra are compared; the signals of binder and non-binders will be either opposite-phased or attenuated in intensity (depending upon competition between positive and negative NOE) (Figure 1.31). The waterLOGSY technique typically requires a ligand:protein concentration ratio of 100:1. The protein concentration usually ranges from few µM to few tens of µM. The temperature of the experiment can be decreased (up to 15 °C) to favour the spin diffusion through the protein.

Ligand binding constants can also be measured by WaterLOGSY NMR technique, by titrating ligand to a fixed concentration of protein, measuring the intensity of the ligand signal as a function of protein concentration.[146] Although Fielding *et al.*[148] have shown that the accuracy of the K_D value may be influenced by protein concentration. The exact reason of this correlation remains unknown.

In waterLOGSY experiment several techniques have been used for selectively and efficiently exciting the water signal.[149-152] Amongst these, the ePHOGSY (enhanced protein hydration observed through gradient spectroscopy) (Figure 1.32) technique, developed by Dalvit[153,154] which is based on a water-selective 180° refocusing pulse between two pulsed field gradients, is one of the most readily implemented and robust schemes to obtain selective water excitation (Figure 1.32). This technique effectively defocuses the magnetization of all resonances that are not near the water chemical shift, hence provides the basis for a low level of artifacts. The optimum mixing time used in the WaterLOGSY experiment depends on the size of the complex, but generally a mixing time of 1 to 3 s is used (Figure 1.32). During the mixing time, the water magnetization that has transferred to the protein is migrated to the ligand *via* spin-diffusion and NOE. In large proteins the transfer of magnetization *via* spin-diffusion is very efficient due to the fast flip-flop transitions[155,156] and hence a short mixing times are sufficient to spread the magnetization through the entire protein.

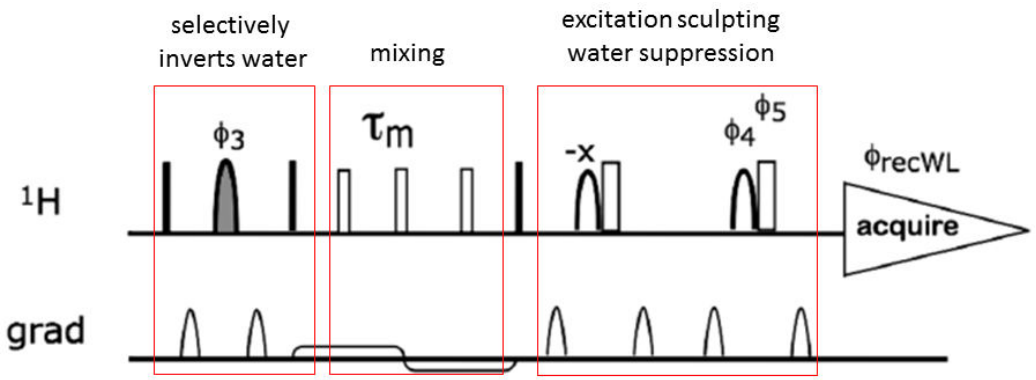


Figure 1.32 Pulse sequences of the homonuclear 1D ePHOGSY NMR.
(Reproduced from reference [157]) The solid bars and boxes represent hard 90°and 180° radio frequency pulses, respectively, whereas, the shaded shaped pulse indicate gauss pulse and the unshaded represent the 180° shape pulse. The shaded shaped pulse selectively inverts the water magnetisation. During the mixing period (τ_m) the water magnetization transfers to the protein (*via* spin-diffusion and NOE) and then to the bound-ligand *via* protein-ligand complex. During the mixing period a weaker bipolar gradient is applied to suppress radiation dumping. In the final section the bulk water signal is suppressed by excitation sculpting. (For more detailed explanation see the original reference [146])

Similar to STD, this technique also relies on the fast exchange of the ligand between the bound and free form, therefore this may not work for very strong binding ligands. Similarly, for very weak affinity interaction, the shorter residence time of the ligand in the protein binding pocket may not be sufficient for efficient transfer of magnetisation and hence such interaction may also not be detected. Work by Dalvit *et al.* suggests that the binding interaction of ligands with K_D values as low as 100 nM can be detected by this technique.[146]

1.5 Objectives

The main objective of the work reported in this Thesis was to apply NMR spectroscopy to gain a better understanding of the function and inhibition of two enzyme systems, γ-butyrobetaine hydroxylase (BBOX), which belongs to the super-family of enzymes called the 2-oxoglutarate (2OG) dependent oxygenases and the bacterial potassium ion efflux protein (Kef).

In Chapter 2, NMR spectroscopy was applied to study the binding interaction of natural substrates GBB and 2OG with BBOX by ^{1}H NMR direct ligand observation. ^{1}H NMR was also applied to develop a dual-reporter ligand displacement assay, using GBB and 2OG simultaneously as reporter ligands for assessing the nature of inhibitor binding with BBOX. In this assay, the application of NMR paramagnetic relaxation enhancement (PRE) effect was also explored. NMR was also applied to develop an inhibition assay for BBOX for inhibitor IC_{50} measurement. ^{1}H NMR-based assays were also applied to demonstrate that the cation-π

interaction between the positive charge of the substrates and the aromatic cage residues of the BBOX play a critical role in BBOX substrate recognition. Finally, using ^{1}H NMR spectroscopy, the binding interaction of clinical drug mildronate with BBOX was also studied.

In Chapter 3, ^{1}H NMR spectroscopy was applied to investigate the chemistry of uncoupled 2OG decarboxylation into succinate by L-ascorbate degradation in the absence of a 2OG oxygenase.

In Chapter 4, ^{1}H NMR direct ligand observation was applied to investigate the role of each functional group of the Kef activating ligand glutathione-S-*N*-tertiary butylsuccinimide on its binding interaction with Kef for developing non-peptidic antibacterial agents of Kef. ^{1}H NMR was also used to establish the identity of the electron density found in the nucleotide binding pocket of Kef protein. ^{19}F NMR spectroscopy was applied to develop an efficient assay for Kef that can be used for ligand screening and measurement of ligand binding dissociation constant values. Finally, with the aid of ^{19}F NMR ligand-observed assay, the conformational changes that occur in Kef due to ligand binding were investigated.

In Chapter 6, experimental procedures are described.

1.6 References

1. Takeda, M.; Jardetzky, O. Proton magnetic resonance of simple amino acids and dipeptides in aqueous solution. *J. Chem. Phys.* **1957**, 26, 1346-1347.
2. Jardetzky, O.; Jardetzky, C. D. Proton magnetic resonance spectra of amino acids. *J. Biol. Chem.* **1958**, 233, 383-387.
3. Saunders, M.; Wishnia, A.; Kirkwood, J. G. The nuclear magnetic resonance spectrum of ribonuclease 1. *J. Am. Chem. Soc.* **1957**, 79, 3289-3290.
4. Jardetzky, O.; Jardetzky, C. D. An interpretation of the proton magnetic resonance spectrum of ribonuclease. *J. Am. Chem. Soc.* **1957**, 79, 5322-5323.
5. Farrar, T.; Gutowsky, H.; Alberty, R.; Miller, W. The mechanism of the stereospecific enzymatic hydration of fumarate to L-malate 1. *J. Am. Chem. Soc.* **1957**, 79, 3978-3980.
6. Jardetzky, O.; Wade, N. G.; Fischer, J. J. Proton magnetic resonance investigation of enzyme-coenzyme complexes. *Nature* **1963**, 197, 183-184.
7. Kato, G. Studies on serum cholinesterase kinetics by nuclear magnetic resonance spectroscopy. *Mol. Pharmacol.* **1968**, 4, 640-644.
8. Kaiser, R. Use of the nuclear Overhauser effect in the analysis of high-resolution nuclear magnetic resonance spectra. *J. Chem. Phy.* **1963**, 39, 2435-2442.
9. Ernst, R. R.; Anderson, W. Application of fourier transform spectroscopy to magnetic resonance. *Rev. Sci. Instrum.* **1966**, 37, 93-102.
10. Aue, W.; Bartholdi, E.; Ernst, R. R. Two-dimensional spectroscopy. Application to nuclear magnetic resonance. *J. Chem. Phy.* **1976**, 64, 2229-2246.
11. Albrand, J.; Birdsall, B.; Feeney, J.; Roberts, G.; Burgen, A. The use of transferred nuclear Overhauser effects in the study of the conformations of small molecules bound to proteins. *Int. J. Biol. Macromol.* **1979**, 1, 37-41.
12. Balaram, P.; Bothner-By, A.; Dadok, J. Negative nuclear Overhuaser effects as probes of macromolecular structure. *J. Am. Chem. Soc.* **1972**, 94, 4015-4017.
13. Balaram, P.; Bothner-By, A.; Breslow, E. Localization of tyrosine at the binding site of neurophysin II by negative nuclear Overhouser effects. *J. Am. Chem. Soc.* **1972**, 94, 4017-4018.
14. Wüthrich, K.; Wider, G.; Wagner, G.; Braun, W. Sequential resonance assignments as a basis for determination of spatial protein structures by high resolution proton nuclear magnetic resonance. *J. Mol. Biol.* **1982**, 155, 311-319.

15. Williamson, M. P.; Havel, T. F.; Wüthrich, K. Solution conformation of proteinase inhibitor IIA from bull seminal plasma by ^{1}H nuclear magnetic resonance and distance geometry. *J. Mol. Biol.* **1985**, 182, 295-315.

16. Kay, L. E.; Ikura, M.; Tschudin, R.; Bax, A. Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *J. Magn. Reson.* **1990**, 89, 496-514.

17. Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. Discovering high-affinity ligands for proteins: SAR by NMR. *Science* **1996**, 274, 1531-1534.

18. Liu, J.; Farmer, J. D.; Lane, W. S.; Friedman, J.; Weissman, I.; Schreiber, S. L. Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP-FK506 complexes. *Cell* **1991**, 66, 807-815.

19. Mayer, M.; Meyer, B. Characterization of ligand binding by saturation transfer difference NMR spectroscopy. *Angew. Chem., Int. Ed.* **1999**, 38, 1784-1788.

20. Meyer, B.; Peters, T. NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. *Angew. Chem., Int. Ed.* **2003**, 42, 864-90.

21. Dalvit, C.; Pevarello, P.; Tatò, M.; Veronesi, M.; Vulpetti, A.; Sundström, M. Identification of compounds with binding affinity to proteins via magnetization transfer from bulk water. *J. Biomol. NMR* **2000**, 18, 65-68.

22. Jahnke, W.; Floersheim, P.; Ostermeier, C.; Zhang, X.; Hemmig, R.; Hurth, K.; Uzunov, D. P. NMR reporter screening for the detection of high-affinity ligands. *Angew. Chem., Int. Ed.* **2002**, 41, 3420-3423.

23. Dalvit, C.; Flocco, M.; Stockman, B.; Veronesi, M. Competition binding experiments for rapidly ranking lead molecules for their binding affinity to human serum albumin. *Comb. Chem. High Throughput Screen.* **2002**, 5, 645-650.

24. Dalvit, C.; Flocco, M.; Knapp, S.; Mostardini, M.; Perego, R.; Stockman, B. J.; Veronesi, M.; Varasi, M. High-throughput NMR-based screening with competition binding experiments. *J. Am. Chem. Soc.* **2002**, 124, 7702-7709.

25. Stockman, B. J. 2-Fluoro-ATP as a versatile tool for ^{19}F NMR-based activity screening. *J. Am. Chem. Soc.* **2008**, 130, 5870-5871.

26. Dalvit, C. Ligand-and substrate-based ^{19}F NMR screening: principles and applications to drug discovery. *Prog. Nucl. Magn. Reson. Spectrosc.* **2007**, 51, 243-271.

27. Dalvit, C.; Ardini, E.; Flocco, M.; Fogliatto, G. P.; Mongelli, N.; Veronesi, M. A general NMR method for rapid, efficient, and reliable biochemical screening. *J. Am. Chem. Soc.* **2003**, 125, 14620-14625.

28. Manzenrieder, F.; Frank, A. O.; Kessler, H. Phosphorus NMR spectroscopy as a versatile tool for compound library screening. *Angew. Chem., Int. Ed.* **2008**, 47, 2608-2611.
29. Araç, D.; Murphy, T.; Rizo, J. Facile detection of protein-protein interactions by one-dimensional NMR spectroscopy. *Biochemistry* **2003**, 42, 2774-2780.
30. Bertini, I.; McGreevy, K. S.; Parigi, G. *Interactions of Metallodrugs with DNA, in NMR of Biomolecules: Towards Mechanistic Systems Biology*. Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, **2012**.
31. Powers, R. Advances in nuclear magnetic resonance for drug discovery. *Expert Opin Drug Discov.* **2009**, 4, 1077.
32. Pellecchia, M.; Bertini, I.; Cowburn, D.; Dalvit, C.; Giralt, E.; Jahnke, W.; James, T. L.; Homans, S. W.; Kessler, H.; Luchinat, C. Perspectives on NMR in drug discovery: a technique comes of age. *Nat. Rev. Drug Discov.* **2008**, 7, 738-745.
33. Jahnke, W. Perspectives of biomolecular NMR in drug discovery: the blessing and curse of versatility. *J. Biomol. NMR* **2007**, 39, 87-90.
34. Lepre, C. A.; Moore, J. M.; Peng, J. W. Theory and applications of NMR-based screening in pharmaceutical research. *Chem. Rev.* **2004**, 104, 3641-3676.
35. Jahnke, W.; Widmer, H. Protein NMR in biomedical research. *Cell. Mol. Life Sci.* **2004**, 61, 580-599.
36. Blommers, M. J.; Flörsheimer, A.; Jahnke, W. Strategies for drug discovery using NMR. In *BioNMR in Drug Research*, Wiley-VCH Weinheim, Germany: **2002**; pp 439-457.
37. Heller, M.; Kessler, H. NMR spectroscopy in drug design. *Pure Appl. Chem.* **2001**, 73, 1429-1436.
38. Harding, S. E.; Chowdhry, B. Z. *Protein-ligand interactions, structure and spectroscopy: a practical approach*. Oxford University Press: Oxford, UK, **2001**.
39. Otting, G. Experimental NMR techniques for studies of protein-ligand interactions. *Curr. Opin. Struct. Biol.* **1993**, 3, 760-768.
40. Koshland, D. E. Application of a Theory of Enzyme Specificity to Protein Synthesis. *Proc. Natl. Acad. Sci. U.S.A.* **1958**, 44, 98-104.
41. Williamson, M. P. Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, 73, 1-16.

42. Cai, M.; Huang, Y.; Sakaguchi, K.; Clore, G. M.; Gronenborn, A. M.; Craigie, R. An efficient and cost-effective isotope labeling protocol for proteins expressed in *Escherichia coli*. *J. Biomol. NMR* **1998**, 11, 97-102.

43. Sticht, J.; Humbert, M.; Findlow, S.; Bodem, J.; Muller, B.; Dietrich, U.; Werner, J.; Krausslich, H.-G. A peptide inhibitor of HIV-1 assembly in vitro. *Nat. Struct. Mol. Biol.* **2005**, 12, 671-677.

44. Yee, A.; Chang, X.; Pineda-Lucena, A.; Wu, B.; Semesi, A.; Le, B.; Ramelot, T.; Lee, G. M.; Bhattacharyya, S.; Gutierrez, P.; Denisov, A.; Lee, C. H.; Cort, J. R.; Kozlov, G.; Liao, J.; Finak, G.; Chen, L.; Wishart, D.; Lee, W.; McIntosh, L. P.; Gehring, K.; Kennedy, M. A.; Edwards, A. M.; Arrowsmith, C. H. An NMR approach to structural proteomics. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, 99, 1825-30.

45. Dyson, H. J.; Wright, P. E. Equilibrium NMR studies of unfolded and partially folded proteins. *Nat. Struct. Biol.* **1998**, 5 Suppl, 499-503.

46. Hajduk, P. J.; Boyd, S.; Nettesheim, D.; Nienaber, V.; Severin, J.; Smith, R.; Davidson, D.; Rockway, T.; Fesik, S. W. Identification of novel inhibitors of urokinase via NMR-based screening. *J. Med. Chem.* **2000**, 43, 3862-6.

47. Fielding, L. NMR methods for the determination of protein–ligand dissociation constants. *Prog. Nucl. Magn. Reson. Spectrosc.* **2007**, 51, 219-242.

48. Fielding, L. NMR methods for the determination of protein-ligand dissociation constants. *Curr. Top. Med. Chem.* **2003**, 3, 39-53.

49. Hajduk, P.; Sheppard, G.; Nettesheim, D.; Olejniczak, E.; Shuker, S.; Meadows, R.; Steinman, D.; Carrera, G.; Marcotte, P.; Severin, J. Discovery of potent nonpeptide inhibitors of stromelysin using SAR by NMR. *J. Am. Chem. Soc.* **1997**, 119, 5818-5827.

50. Wuthrich, K. *NMR of proteins and nucleic acids*. John Wiley & Sons: New York, **1986**.

51. Sattler, M.; Schleucher, J.; Griesinger, C. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Magn. Reson. Spectrosc.* **1999**, 34, 93-158.

52. Ikura, M.; Kay, L. E.; Bax, A. A novel approach for sequential assignment of proton, carbon-13, and nitrogen-15 spectra of larger proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin. *Biochemistry* **1990**, 29, 4659-4667.

53. Kamps, J. J. A. G.; Khan, A.; Choi, H.; Lesniak, R. K.; Brem, J.; Rydzik, A. M.; McDonough, M. A.; Schofield, C. J.; Claridge, T. D. W.; Mecinović, J. Cation–π Interactions Contribute to Substrate Recognition in γ-Butyrobetaine Hydroxylase Catalysis. *Chem. Eur. J.* **2016**, 1270-1276.

54. Gardner, K. H.; Kay, L. E. The use of ^{2}H, ^{13}C, ^{15}N multidimensional NMR GTO study of the structure and dynamics of proteins. *Annu. Rev. Biophys. Biomol. Struct.* **1998**, 27, 357-406.

55. Carbajo, R. J.; Kellas, F. A.; Runswick, M. J.; Montgomery, M. G.; Walker, J. E.; Neuhaus, D. Structure of the F1-binding domain of the stator of bovine F1 Fo-ATPase and how it binds an α-subunit. *J. Mol. Biol.* **2005**, 351, 824-838.

56. Hajduk, P. J.; Augeri, D. J.; Mack, J.; Mendoza, R.; Yang, J.; Betz, S. F.; Fesik, S. W. NMR-based screening of proteins containing ^{13}C-labeled methyl groups. *J. Am. Chem. Soc.* **2000**, 122, 7898-7904.

57. Goto, N. K.; Gardner, K. H.; Mueller, G. A.; Willis, R. C.; Kay, L. E. A robust and cost-effective method for the production of Val, Leu, Ile (δ1) methyl-protonated ^{15}N-, ^{13}C-, ^{2}H-labeled proteins. *J. Biomol. NMR* **1999**, 13, 369-374.

58. Gardner, K. H.; Kay, L. E. Production and incorporation of ^{15}N, ^{13}C, ^{2}H (1H-δ1 methyl) isoleucine into proteins for multidimensional NMR studies. *J. Am. Chem. Soc.* **1997**, 119, 7599-7600.

59. Rosen, M. K.; Gardner, K. H.; Willis, R. C.; Parris, W. E.; Pawson, T.; Kay, L. E. Selective methyl group protonation of perdeuterated proteins. *J. Mol. Biol.* **1996**, 263, 627-636.

60. Mannhold, R.; Kubinyi, H.; Folkers, G.; Gohlke, H. *Protein-ligand Interactions*. Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, **2012**.

61. Arntson, K. E.; Pomerantz, W. C. Protein-observed ^{19}F NMR: A bioorthogonal approach for small molecule discovery. *J. Med. Chem.* **2015**, DOI: 10.1021/acs.jmedchem.5b01447.

62. Marsh, E. N. G.; Suzuki, Y. Using ^{19}F NMR to probe biological interactions of proteins and peptides. *ACS Chem. Biol.* **2014**, 9, 1242-1250.

63. Yu, J. X.; Hallac, R. R.; Chiguru, S.; Mason, R. P. New frontiers and developing applications in ^{19}F NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, 70, 25-49.

64. Chen, H.; Viel, S.; Ziarelli, F.; Peng, L. ^{19}F NMR: a valuable tool for studying biological events. *Chem. Soc. Rev.* **2013**, 42, 7971-82.

65. Kitevski-LeBlanc, J. L.; Prosser, R. S. Current applications of ^{19}F NMR to studies of protein structure and dynamics. *Prog. Nucl. Magn. Reson. Spectrosc.* **2012**, 62, 1-33.

66. Emsley, J. W.; Phillips, L. Fluorine chemical shifts. *Prog. Nucl. Magn. Reson. Spectrosc.* **1971**, 7, 1-520.

67. Gerig, J. Fluorine NMR of proteins. *Prog. Nucl. Magn. Reson. Spectrosc.* **1994**, 26, 293-370.

68. Larda, S. T.; Simonetti, K.; Al-Abdul-Wahid, M. S.; Sharpe, S.; Prosser, R. S. Dynamic equilibria between monomeric and oligomeric misfolded states of the mammalian prion protein measured by ^{19}F NMR. *J. Am. Chem. Soc.* **2013**, 135, 10533-41.

69. Luck, L. A.; Falke, J. J. Fluorine-19 NMR studies of the D-galactose chemosensory receptor. 1. Sugar binding yields a global structural change. *Biochemistry* **1991**, 30, 4248-4256.

70. Li, H.; Frieden, C. Observation of sequential steps in the folding of intestinal fatty acid binding protein using a slow folding mutant and ^{19}F NMR. *Proc. Natl. Acad. Sci. U.S. A.* **2007**, 104, 11993-8.

71. Li, H.; Frieden, C. Comparison of C40/82A and P27A C40/82A barstar mutants using ^{19}F NMR. *Biochemistry* **2007**, 46, 4337-47.

72. Li, H.; Frieden, C. NMR studies of 4-^{19}F-phenylalanine-labeled intestinal fatty acid binding protein: evidence for conformational heterogeneity in the native state. *Biochemistry* **2005**, 44, 2369-77.

73. Chadegani, F.; Lovell, S.; Mullangi, V.; Miyagi, M.; Battaile, K. P.; Bann, J. G. ^{19}F nuclear magnetic resonance and crystallographic studies of 5-fluorotryptophan-labeled anthrax protective antigen and effects of the receptor on stability. *Biochemistry* **2014**, 53, 690-701.

74. Evanics, F.; Bezsonova, I.; Marsh, J.; Kitevski, J. L.; Forman-Kay, J. D.; Prosser, R. S. Tryptophan solvent exposure in folded and unfolded states of an SH3 domain by ^{19}F and ^{1}H NMR. *Biochemistry* **2006**, 45, 14120-8.

75. Anderluh, G.; Razpotnik, A.; Podlesek, Z.; Macek, P.; Separovic, F.; Norton, R. S. Interaction of the eukaryotic pore-forming cytolysin equinatoxin II with model membranes: ^{19}F NMR studies. *J. Mol. Biol.* **2005**, 347, 27-39.

76. Luck, L. A.; Vance, J. E.; O'Connell, T. M.; London, R. E. ^{19}F NMR relaxation studies on 5-fluorotryptophan- and tetradeutero-5-fluorotryptophan-labeled *E. coli* glucose/galactose receptor. *J. Biomol. NMR* **1996**, 7, 261-72.

77. Peersen, O. B.; Pratt, E. A.; Truong, H. T.; Ho, C.; Rule, G. S. Site-specific incorporation of 5-fluorotryptophan as a probe of the structure and function of the membrane-bound D-lactate dehydrogenase of *Escherichia coli*: a ^{19}F nuclear magnetic resonance study. *Biochemistry* **1990**, 29, 3256-62.

78. Bann, J. G.; Pinkner, J.; Hultgren, S. J.; Frieden, C. Real-time and equilibrium ^{19}F NMR studies reveal the role of domain-domain interactions in the folding of the chaperone PapD. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, 99, 709-14.

79. Zigoneanu, I. G.; Pielak, G. J. Interaction of alpha-synuclein and a cell penetrating fusion peptide with higher eukaryotic cell membranes assessed by ^{19}F NMR. *Mol. Pharm.* **2012**, 9, 1024-9.

80. Pomerantz, W. C.; Wang, N.; Lipinski, A. K.; Wang, R.; Cierpicki, T.; Mapp, A. K. Profiling the dynamic interfaces of fluorinated transcription complexes for ligand discovery and characterization. *ACS Chem. Biol.* **2012**, 7, 1345-50.

81. Schlesinger, A. P.; Wang, Y.; Tadeo, X.; Millet, O.; Pielak, G. J. Macromolecular crowding fails to fold a globular protein in cells. *J. Am. Chem. Soc.* **2011**, 133, 8082-5.

82. Wang, G. F.; Li, C.; Pielak, G. J. ^{19}F NMR studies of alpha-synuclein-membrane interactions. *Protein Sci.* **2010**, 19, 1686-91.

83. Li, C.; Wang, G. F.; Wang, Y.; Creager-Allen, R.; Lutz, E. A.; Scronce, H.; Slade, K. M.; Ruf, R. A.; Mehl, R. A.; Pielak, G. J. Protein ^{19}F NMR in *Escherichia coli*. *J. Am. Chem. Soc.* **2010**, 132, 321-7.

84. Li, C.; Lutz, E. A.; Slade, K. M.; Ruf, R. A.; Wang, G. F.; Pielak, G. J. ^{19}F NMR studies of alpha-synuclein conformation and fibrillation. *Biochemistry* **2009**, 48, 8578-84.

85. Evanics, F.; Kitevski, J. L.; Bezsonova, I.; Forman-Kay, J.; Prosser, R. S. ^{19}F NMR studies of solvent exposure and peptide binding to an SH3 domain. *Biochim. Biophys. Acta* **2007**, 1770, 221-30.

86. Eccleston, J. F.; Molloy, D. P.; Hinds, M. G.; King, R. W.; Feeney, J. Conformational differences between complexes of elongation factor Tu studied ^{19}F NMR spectroscopy. *Eur. J. Biochem.* **1993**, 218, 1041-7.

87. Hull, W. E.; Sykes, B. D. Fluorine-19 nuclear magnetic resonance study of fluorotyrosine alkaline phosphatase: the influence of zinc on protein structure and a conformational change induced by phosphate binding. *Biochemistry* **1976**, 15, 1535-46.

88. Hull, W. E.; Sykes, B. D. Fluorotyrosine alkaline phosphatase. ^{19}F nuclear magnetic resonance relaxation times and molecular motion of the individual fluorotyrosines. *Biochemistry* **1974**, 13, 3431-7.

89. Feeney, J.; McCormick, J. E.; Bauer, C. J.; Birdsall, B.; Moody, C. M.; Starkmann, B. A.; Young, D. W.; Francis, P.; Havlin, R. H.; Arnold, W. D. ^{19}F nuclear magnetic resonance chemical shifts of fluorine containing aliphatic amino acids in proteins: studies on lactobacillus casei dihydrofolate reductase containing (2*S*, 4*S*)-5-fluoroleucine. *J. Am. Chem. Soc.* **1996**, 118, 8700-8706.

90. Suzuki, Y.; Buer, B. C.; Al-Hashimi, H. M.; Marsh, E. N. G. Using fluorine nuclear magnetic resonance to probe changes in the structure and dynamics of membrane-active peptides interacting with lipid bilayers. *Biochemistry* **2011**, 50, 5979-5987.

91. Buer, B. C.; Chugh, J.; Al-Hashimi, H. M.; Marsh, E. N. G. Using fluorine nuclear magnetic resonance to probe the interaction of membrane-active peptides with the lipid bilayer. *Biochemistry* **2010**, 49, 5760-5765.

92. Suzuki, Y.; Brender, J. R.; Soper, M. T.; Krishnamoorthy, J.; Zhou, Y.; Ruotolo, B. T.; Kotov, N. A.; Ramamoorthy, A.; Marsh, E. N. Resolution of oligomeric species during the aggregation of Abeta1-40 using ^{19}F NMR. *Biochemistry* **2013**, 52, 1903-12.

93. Holzberger, B.; Rubini, M.; Möller, H. M.; Marx, A. A highly active DNA polymerase with a fluorous core. *Angew. Chem., Int. Ed.* **2010**, 49, 1324-1327.

94. Duewel, H.; Daub, E.; Robinson, V.; Honek, J. F. Incorporation of trifluoromethionine into a phage lysozyme: implications and a new marker for use in protein ^{19}F NMR. *Biochemistry* **1997**, 36, 3404-3416.

95. Vaughan, M. D.; Cleve, P.; Robinson, V.; Duewel, H. S.; Honek, J. F. Difluoromethionine as a novel ^{19}F NMR structural probe for internal amino acid packing in proteins. *J. Am. Chem. Soc.* **1999**, 121, 8475-8478.

96. Eichler, J. F.; Cramer, J. C.; Kirk, K. L.; Bann, J. G. Biosynthetic incorporation of fluorohistidine into proteins in *E. coli*: a new probe of macromolecular structure. *Chembiochem* **2005**, 6, 2170-2173.

97. Wang, L.; Xie, J.; Schultz, P. G. Expanding the genetic code. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, 35, 225-249.

98. Link, A. J.; Mock, M. L.; Tirrell, D. A. Non-canonical amino acids in protein engineering. *Curr. Opin. Biotechnol.* **2003**, 14, 603-609.

99. Chin, J. W.; Cropp, T. A.; Chu, S.; Meggers, E.; Schultz, P. G. Progress toward an expanded eukaryotic genetic code. *Chem. Biol.* **2003**, 10, 511-519.

100. Cotton, G. J.; Muir, T. W. Peptide ligation and its application to protein engineering. *Chem. Biol.* **1999**, 6, 247-256.

101. Hoeltzli, S. D.; Frieden, C. ^{19}F NMR spectroscopy of [6-^{19}F] tryptophan-labeled Escherichia coli dihydrofolate reductase: equilibrium folding and ligand binding studies. *Biochemistry* **1994**, 33, 5502-5509.

102. Danielson, M. A.; Biemann, H. P.; Koshland, D. E., Jr.; Falke, J. J. Attractant- and disulfide-induced conformational changes in the ligand binding domain of the chemotaxis aspartate receptor: a ^{19}F NMR study. *Biochemistry* **1994**, 33, 6100-9.

103. Drake, S. K.; Bourret, R. B.; Luck, L. A.; Simon, M. I.; Falke, J. J. Activation of the phosphosignaling protein CheY. I. Analysis of the phosphorylated conformation by ^{19}F NMR and protein engineering. *J. Biol. Chem.* **1993**, 268, 13081-8.

104. Ropson, I. J.; Frieden, C. Dynamic NMR spectral analysis and protein folding: identification of a highly populated folding intermediate of rat intestinal fatty acid-binding protein by ^{19}F NMR. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, 89, 7222-6.

105. Li, E.; Qian, S. J.; Yang, N. C.; d'Avignon, A.; Gordon, J. I. ^{19}F nuclear magnetic resonance studies of 6-fluorotryptophan-substituted rat cellular retinol binding protein II produced in *Escherichia coli*. An analysis of four tryptophan substitution mutants and their interactions with all-trans-retinol. *J. Biol. Chem.* **1990**, 265, 11549-54.

106. Rule, G. S.; Pratt, E. A.; Simplaceanu, V.; Ho, C. Nuclear magnetic resonance and molecular genetic studies of the membrane-bound D-lactate dehydrogenase of *Escherichia coli*. *Biochemistry* **1987**, 26, 549-56.

107. Loscha, K. V.; Herlt, A. J.; Qi, R.; Huber, T.; Ozawa, K.; Otting, G. Multiple-site labeling of proteins with unnatural amino acids. *Angew. Chem., Int. Ed.* **2012**, 51, 2243-6.

108. Jackson, J. C.; Hammill, J. T.; Mehl, R. A. Site-specific incorporation of a ^{19}F-amino acid into proteins as an NMR probe for characterizing protein structure and reactivity. *J. Am. Chem. Soc.* **2007**, 129, 1160-6.

109. Kent, S. B. Total chemical synthesis of proteins. *Chem. Soc. Rev.* **2009**, 38, 338-51.

110. Dawson, P. E.; Kent, S. B. Synthesis of native proteins by chemical ligation. *Annu. Rev. Biochem.* **2000**, 69, 923-60.

111. Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. Synthesis of proteins by native chemical ligation. *Science* **1994**, 266, 776-9.

112. Spokoyny, A. M.; Zou, Y.; Ling, J. J.; Yu, H.; Lin, Y. S.; Pentelute, B. L. A perfluoroaryl-cysteine S(N)Ar chemistry approach to unprotected peptide stapling. *J. Am. Chem. Soc.* **2013**, 135, 5946-9.

113. Liu, J. J.; Horst, R.; Katritch, V.; Stevens, R. C.; Wuthrich, K. Biased signaling pathways in beta 2-adrenergic receptor characterized by ^{19}F NMR. *Science* **2012**, 335, 1106-10.

114. Klein-Seetharaman, J.; Getmanova, E. V.; Loewen, M. C.; Reeves, P. J.; Khorana, H. G. NMR spectroscopy in studies of light-induced structural changes in mammalian rhodopsin: applicability of solution ^{19}F NMR. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, 96, 13744-9.

115. Kim, T. H.; Chung, K. Y.; Manglik, A.; Hansen, A. L.; Dror, R. O.; Mildorf, T. J.; Shaw, D. E.; Kobilka, B. K.; Prosser, R. S. The role of ligands on the equilibria between functional states of a G protein-coupled receptor. *J. Am. Chem. Soc.* **2013**, 135, 9465-74.

116. Manglik, A.; Kim, T. H.; Masureel, M.; Altenbach, C.; Yang, Z.; Hilger, D.; Lerch, M. T.; Kobilka, T. S.; Thian, F. S.; Hubbell, W. L.; Prosser, R. S.; Kobilka, B. K. Structural insights into the dynamic process of beta 2-adrenergic receptor signaling. *Cell* **2015**, 161, 1101-11.

117. Curtis-Marof, R.; Doko, D.; Rowe, M. L.; Richards, K. L.; Williamson, R. A.; Howard, M. J. ^{19}F NMR spectroscopy monitors ligand binding to recombinantly fluorine-labelled b'x from human protein disulphide isomerase (hPDI). *Org. Biomol. Chem.* **2014**, 12, 3808-12.

118. Hatahet, F.; Ruddock, L. W. Substrate recognition by the protein disulfide isomerases. *FEBS J* **2007**, 274, 5223-34.

119. Byrne, L. J.; Sidhu, A.; Wallis, A. K.; Ruddock, L. W.; Freedman, R. B.; Howard, M. J.; Williamson, R. A. Mapping of the ligand-binding site on the b' domain of human PDI: interaction with peptide ligands and the x-linker region. *Biochem. J.* **2009**, 423, 209-17.

120. Pirneskoski, A.; Klappa, P.; Lobell, M.; Williamson, R. A.; Byrne, L.; Alanen, H. I.; Salo, K. E.; Kivirikko, K. I.; Freedman, R. B.; Ruddock, L. W. Molecular characterization of the principal substrate binding site of the ubiquitous folding catalyst protein disulfide isomerase. *J. Biol. Chem.* **2004**, 279, 10374-81.

121. Klappa, P.; Ruddock, L. W.; Darby, N. J.; Freedman, R. B. The b' domain provides the principal peptide-binding site of protein disulfide isomerase but all domains contribute to binding of misfolded proteins. *EMBO J.* **1998**, 17, 927-35.

122. Wallis, A. K.; Sidhu, A.; Byrne, L. J.; Howard, M. J.; Ruddock, L. W.; Williamson, R. A.; Freedman, R. B. The ligand-binding b' domain of human protein disulphide-isomerase mediates homodimerization. *Protein Sci.* **2009**, 18, 2569-77.

123. Nguyen, V. D.; Wallis, K.; Howard, M. J.; Haapalainen, A. M.; Salo, K. E. H.; Saaranen, M. J.; Sidhu, A.; Wierenga, R. K.; Freedman, R. B.; Ruddock, L. W.; Williamson, R. A. Alternative conformations of the x region of human protein disulphide-isomerase modulate exposure of the substrate binding b' domain. *J. Mol. Biol.* **2008**, 383, 1144-1155.

124. Dalvit, C.; Gossert, A. D.; Coutant, J.; Piotto, M. Rapid acquisition of ^{1}H and ^{19}F NMR experiments for direct and competition ligand-based screening. *Magn. Reson. Chem.* **2011**, 49, 199-202.

125. Tengel, T.; Fex, T.; Emtenäs, H.; Almqvist, F.; Sethson, I.; Kihlberg, J. Use of ^{19}F NMR spectroscopy to screen chemical libraries for ligands that bind to proteins. *Org. Biomol. Chem.* **2004**, 2, 725-731.

126. Dalvit, C.; Flocco, M.; Veronesi, M.; Stockman, B. Fluorine-NMR competition binding experiments for high-throughput screening of large compound mixtures. *Comb. Chem. High Throughput Screen.* **2002**, 5, 605-611.

127. Sastry, M. K.; Swamy, M. J.; Surolia, A. Analysis of dynamics and mechanism of ligand binding to Artocarpus integrifolia agglutinin. A ^{13}C and ^{19}F NMR study. *J. Biol. Chem.* **1988**, 263, 14826-14831.

128. Hajduk, P. J.; Olejniczak, E. T.; Fesik, S. W. One-dimensional relaxation- and diffusion-edited NMR methods for screening compounds that bind to macromolecules. *J. Am. Chem. Soc.* **1997**, 119, 12257-12261.

129. Meiboom, S.; Gill, D. Modified spin-echo method for measuring nuclear relaxation times. *Rev. Sci. Instrum.* **1958**, 29, 688-691.

130. Leung, I. K. H.; Demetriades, M.; Hardy, A. P.; Lejeune, C.; Smart, T. J.; Szollossi, A.; Kawamura, A.; Schofield, C. J.; Claridge, T. D. W. Reporter ligand NMR screening method for 2-oxoglutarate oxygenase inhibitors. *J. Med. Chem.* **2013**, 56, 547-55.

131. Dalvit, C. Theoretical analysis of the competition ligand-based NMR experiments and selected applications to fragment screening and binding constant measurements. *Concepts in Magn. Reson. Part A* **2008**, 32A, 341-372.

132. Dalvit, C.; Flocco, M.; Veronesi, M.; Stockman, B. J. Fluorine-NMR competition binding experiments for high-throughput screening of large compound mixtures. *Comb. Chem. High Throughput Screen.* **2002**, 5, 605-11.

133. Gerig, J. Fluorine nuclear magnetic resonance of fluorinated ligands. *Methods Enzymol.* **1989**, 177, 3.

134. Hull, W. E.; Sykes, B. D. Fluorotyrosine alkaline phosphatase: internal mobility of individual tyrosines and the role of chemical shift anisotropy as a ^{19}F nuclear spin relaxation mechanism in proteins. *J. Mol. Biol.* **1975**, 98, 121-53.

135. Glaser, S. D. Nuclear magnetic resonance: concepts and methods. In John Wiley & Sons: Chichester, New York, Brisbane, Toronto, Singapore, **2000**.

136. Matei, E.; Andre, S.; Glinschert, A.; Infantino, A. S.; Oscarson, S.; Gabius, H. J.; Gronenborn, A. M. Fluorinated carbohydrates as lectin ligands: dissecting glycan-cyanovirin interactions by using ^{19}F NMR spectroscopy. *Chemistry (Easton)* **2013**, 19, 5364-74.

137. Goudreau, N.; Coulombe, R.; Faucher, A. M.; Grand-Maitre, C.; Lacoste, J. E.; Lemke, C. T.; Malenfant, E.; Bousquet, Y.; Fader, L.; Simoneau, B.; Mercier, J. F.; Titolo, S.; Mason, S. W. Monitoring binding of HIV-1 capsid assembly inhibitors using ^{19}F ligand-and ^{15}N protein-based NMR and X-ray crystallography: early hit validation of a benzodiazepine series. *ChemMedChem* **2013**, 8, 405-14.

138. Jenkins, B. G.; Lauffer, R. B. Detection of site-specific binding and co-binding of ligands to human serum albumin using ^{19}F NMR. *Mol. Pharmacol.* **1990**, 37, 111-8.

139. Dugad, L. B.; Gerig, J. T. NMR studies of carbonic anhydrase-4-fluorobenzenesulfonamide complexes. *Biochemistry* **1988**, 27, 4310-4316.

140. Dalvit, C.; Mongelli, N.; Papeo, G.; Giordano, P.; Veronesi, M.; Moskau, D.; Kümmerle, R. Sensitivity improvement in ^{19}F NMR-based screening experiments: theoretical considerations and experimental applications. *J. Am. Chem. Soc.* **2005**, 127, 13380-13385.

141. Dalvit, C.; Fagerness, P. E.; Hadden, D. T. A.; Sarver, R. W.; Stockman, B. J. Fluorine-NMR experiments for high-throughput screening: Theoretical aspects, practical considerations, and range of applicability. *J. Am. Chem. Soc.* **2003**, 125, 7696-7703.

142. Ni, F. Recent developments in transferred NOE methods. *Prog. Nucl. Magn. Reson. Spectrosc.* **1994**, 26, 517-606.

143. Rowe, M. L.; Wagstaff, J. L.; Howard, M. J. NMR in Ligand Binding Studies. In *Modern NMR Techniques for Synthetic Chemistry*, CRC Press: **2014**; pp 63-124.

144. Viegas, A.; Manso, J.; Nobrega, F. L.; Cabrita, E. J. Saturation-transfer difference (STD) NMR: a simple and fast method for ligand screening and characterization of protein binding. *J. Chem. Educ.* **2011**, 88, 990-994.

145. Enriquez-Navas, P. M.; Marradi, M.; Padro, D.; Angulo, J.; Penades, S. A solution NMR study of the interactions of oligomannosides and the anti-HIV-1 2G12 antibody reveals distinct binding modes for branched ligands. *Chemistry (Easton)* **2011**, 17, 1547-60.

146. Dalvit, C.; Fogliatto, G.; Stewart, A.; Veronesi, M.; Stockman, B. WaterLOGSY as a method for primary NMR screening: practical aspects and range of applicability. *J. Biomol. NMR* **2001**, 21, 349-359.

147. Unione, L.; Galante, S.; Díaz, D.; Jiménez-Barbero, J. NMR and molecular recognition. The application of ligand-based NMR methods to monitor molecular interactions. *Med. Chem. Comm.* **2014**, 5, 1280-1289.

148. Fielding, L.; Rutherford, S.; Fletcher, D. Determination of protein–ligand binding affinity by NMR: observations from serum albumin model systems. *Magn. Reson. Chem.* **2005**, 43, 463-470.

149. Melacini, G.; Kaptein, R.; Boelens, R. Editing of chemical exchange-relayed NOEs in NMR experiments for the observation of protein-water interactions. *J. Magn. Reson.* **1999**, 136, 214-8.

150. Melacini, G.; Boelens, R.; Kaptein, R. Band-selective editing of exchange-relay in protein-water NOE experiments. *J. Biomol. NMR* **1999**, 13, 67-71.

151. Wider, G. Technical aspects of NMR spectroscopy with biological macromolecules and studies of hydration in solution. *Prog. Nucl. Magn. Reson. Spectrosc.* **1998**, 32, 193-275.

152. Otting, G. NMR studies of water bound to biological molecules. *Prog. Nucl. Magn. Reson. Spectrosc.* **1997**, 31, 259-285.

153. Dalvit, C. Homonuclear 1D and 2D NMR experiments for the observation of solvent–solute interactions. *J. Magn. Reson; Ser B* **1996**, 112, 282-288.

154. Dalvit, C.; Hommel, U. Sensitivity-improved detection of protein hydration and its extension to the assignment of fast-exchanging resonances. *J. Magn. Reson; Ser B* **1995**, 109, 334-338.

155. Stoesz, J. D.; Redfield, A. G.; Malinowski, D. Cross relaxation and spin diffusion effects on the proton NMR of biopolymers in H_2O: Solvent saturation and chemical exchange in superoxide dismutase. *FEBS Lett.* **1978**, 91, 320-324.

156. Kalk, A.; Berendsen, H. Proton magnetic relaxation and spin diffusion in proteins. *J. Magn. Reson.* **1976**, 24, 343-366.

157. Antanasijevic, A.; Ramirez, B.; Caffrey, M. Comparison of the sensitivities of WaterLOGSY and saturation transfer difference NMR experiments. *J. Biomol. NMR* **2014**, 60, 37-44.

2 Catalysis and inhibition studies of γ-butyrobetaine hydroxylase (BBOX)[†]

2.1 2OG Oxygenases

The 2-oxoglutarate (2OG) and ferrous ion (Fe^{II}) dependent oxygenases (Table 2.1) encompass a super-family of structurally related non-haem enzymes, which are widely found in all life forms.[1-5] These enzymes use the tricarboxylic acid (TCA) cycle intermediate 2-oxoglutarate (2OG), also termed as α-ketoglutarate and molecular oxygen as co-substrates and Fe^{II} as co-factor in catalysis.

The 2OG oxygenases catalyse a broad range of biochemical reactions that is amongst the most diverse of any known enzyme family (Figure 2.1). In animals, these enzymes catalyse hydroxylation and demethylation reactions (proceeding *via* hydroxylation),[1,6,7] and in plants and microorganisms, these enzymes catalyse not only hydroxylations, but also a broad range of other oxidation reactions including epimerisations, desaturations, ring expansions, ring closure and halogenation reactions (Figure 2.1).[8-11] The substrates of 2OG oxygenases range from small molecules to macromolecules including proteins/peptides/amino acids, nucleic acids/bases, lipids and natural products including antibiotics and signalling molecules.[1,6]

In human physiology, the 2OG oxygenases play critical roles in a diverse range of important biological processes including biosynthesis (e.g L-carnitine biosynthesis, collagen biosyn-

---

[†] *The work described in this chapter was designed and conducted in collaboration with Robert Lesniak, Dr Anna M. Rydzik, Dr Jurgen Brem, Jos J. A. Kamps and Dr Ivanhoe K. H. Leung.*

R.L performed the fluorescence binding assay and fluoride ion release assay. A.M.R provided the isoquinoline-based compounds. J.B produced the psBBOX. J.J.A.K synthesised the phospha, arsa and carba GBB analogue ligands and performed the LC-MS assays. I.K.H.L provided the inhibitors other than iso-quinoline-based compounds.

Parts of the work from this chapter have been published as described:

*1. A. Khan, R. K. Leśniak, J. Brem, A. M. Rydzik, H. Choi, I. K. H. Leung, M. A. McDonough, C. J. Schofield and T. D. W. Claridge; Development and applications of ligand-based NMR screening assays for γ-butyrobetaine hydroxylase, Med. Chem. Commun., 2016, DOI: 10.1039/c6md00004e. (**Hot Article**)*

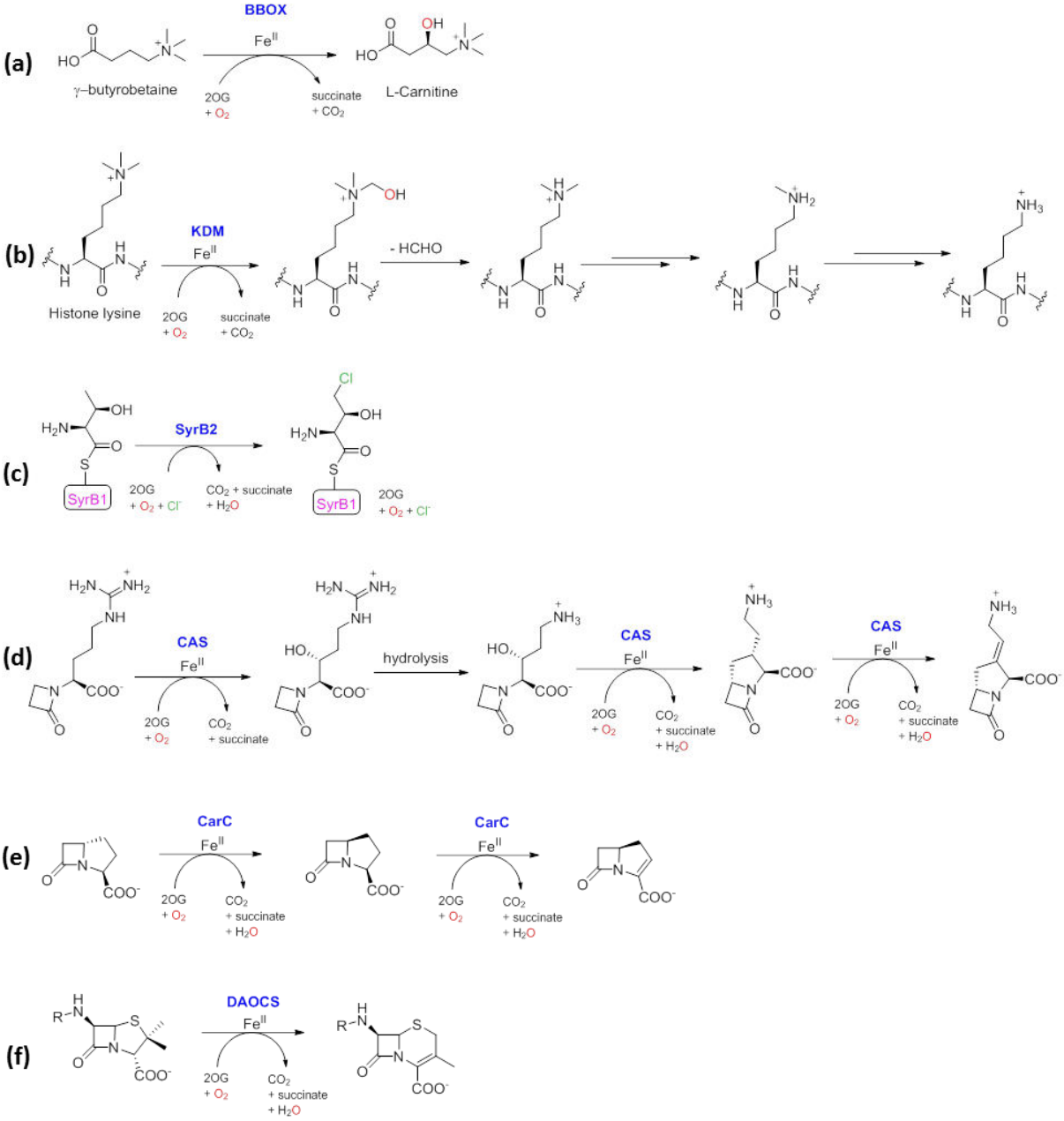
*2. Kamps, J. J.; Khan, A.; Choi, H.; Lesniak, R. K.; Brem, J.; Rydzik, A. M.; McDonough, M. A.; Schofield, C. J.; Claridge, T. D.; Mecinovic, J. Cation-pi Interactions Contribute to Substrate Recognition in γ-Butyrobetaine Hydroxylase Catalysis. Eur. Chem. J.; 2016, 22, 1270-1276, DOI: 10.1002/chem.201503761. (**Hot Article**)*

thesis in prokaryotes to gibberellin biosynthesis in plants.[4,7]

Table 2.1 List of known human 2OG oxygenases with substrates (Reproduced from reference [6]).

Gene name	GeneID	Chemical function if assigned
ASPH (aspartyl/asparaginyl β-hydroxylase)	444	Asn/Asp 3*R*-hydroxylase, EGF-like domains
ASPHD2 (hypothetical protein LOC57168)	57168	
ASPHD1 (hypothetical protein LOC253982)	253982	
C17orf101 (PKHD domain-containing transmembrane protein C17orf101)	79701	
LEPRE1, LEPRE1-like, LEPRE2 (leucine proline-enriched proteoglycan (leprecan))	64175 55214 10536	Prolyl-3*S*-hydroxylase
P4H TM (hypoxia-inducible factor prolyl 4-hydroxylase isoform a, transmembrane (endoplasmic reticulum))	54681	Prolyl hydroxylation
P4HA1, P4HA2, P4HA3 (procollagen-proline, 2-oxoglutarate 4-dioxygenase)	5033 8974 283208	Prolyl-4*R*-hydroxylase (collagen)
PLOD1, PLOD3, PLOD2 (procollagen-lysine, 2-oxoglutarate 5-dioxygenase)	8985 5351 5352	Lysyl-5*R*-hydroxylase (collagen)
JMJD4 (Jumonji domain containing 4)	65094	
JMJD6/PTDSR (Jumonji domain containing 6/phosphatidylserine receptor)	23210	Lysyl-5-hydroxylase
JMJD5 (Jumonji domain containing 5/hypothetical protein FLJ13798)	79831	Histone H3K36me2 demethylase
JMJD8 (hypothetical LOC339123)	339123	
TYW5/C2orf60 (chromosome 2 open reading frame 60)	129450	Wybutosine hydroxylase (modified RNA)
FIH (factor inhibiting hypoxia-inducible factor)	55662	Asparaginyl hydroxylase (HIF CTAD and ankyrin repeat domain containing proteins)
PASS1/HSPBAP1	79663	
JMJD7/PLA2gIVB (phospholipase A2, group IVB)	8681	
NO66	79697	Proposed histone demethylase
MINA53B (MYC induced nuclear antigen)	84864	
JMJD3/KDM6B (Jumonji domain containing 3/histone lysine demethylase 6B)	23135	Histone H3K27me3 demethylase
UTX/KDM6A (ubiquitously transcribed tetratricopeptide repeat, X chromosome/histone lysine demethylase 6B)	7403	Histone H3K27me3 demethylase
UTY (ubiquitously transcribed tetratricopeptide repeat protein, chromosome Y)	7404	
JARID1B/PLU-1/KDM5B (Jumonji, AT rich interactive domain 1B (RBP2-like))	10765	Histone H3K4me3 demethylase
JARID1A/RBBP2/KDM5A (retinoblastoma binding protein 2)	5927	Histone H3K4me3 demethylase
JARID1C/SMCX/KDM5C (Smcx homolog, X chromosome)	8242	Histone H3K4me3 demethylase
JARID1D/SMCY/KDM5D (Smcy homolog, Y-linked)	8284	Histone H3K4me3 demethylase
JMJD2A/JHDM3A/KDM4A (Jumonji domain containing 2A/histone lysine demethylase 4A)	9682	Histone H3K9me3/K36me3 demethylase
JMJD2C/GASC1/KDM4C (Jumonji domain containing 2C/gene amplified in squamous cell carcinoma 1/histone lysine demethylase 4C)	23081	Histone H3K9me3/K36me3 demethylase
JMJD2B/KDM4C (Jumonji domain containing 2B/histone lysine demethylase 4B)	23030	Histone H3K9me3/K36me3 demethylase
JMJD2D/KDM4D (Jumonji domain containing 2D/histone lysine demethylase 4D)	55693	Histone H3K9me3 demethylase
JMJD2E/KDM4E (Jumonji domain containing 2E/histone lysine demethylase 4E)		Histone H3K9me3 demethylase
FBXL10/JHDM1B/KDM2B (F-box and leucine-rich repeat protein 10/histone lysine demethylase 2B)	84678	Histone H3K36me2 demethylase
FBXL11/JHDM1A/KDM2A (F-box and leucine-rich repeat protein 11/histone lysine demethylase 2A)	22992	Histone H3K36me2 demethylase
KIAA1718/JHDM1D/KIAA1718	80853	Histone H3K9me2/K27me2 demethylase
PHF8/KIAA1111 (PHD finger protein 8)	23133	Histone H3K9me2 demethylase
PHF2/JHDM1E/GRC5 (PHD finger protein 2)	5253	
HR (hairless)	55806	
JMJD1A/KDM3A/TSGA (Jumonji domain containing 1A/histone lysine demethylase 3A)	55818	Histone H3K9me2 demethylase
JMJD1B/KDM3B/5qNCA (Jumonji domain containing 1B/histone lysine demethylase 3B)	51780	Histone H3K9me2 demethylase
JMJD1C/TRIP8/KIAA1380 (Jumonji domain containing 1C/histone lysine demethylase 3C)	23081	Histone H3K9me2 demethylase
JARID2/JMJ (JARID2/original Jumonji protein)	3720	Missing iron binding residue
PHD1 (HIF prolyl-4-hydroxylase, N-terminal domain disordered)	112398	Prolyl-4*R*-hydroxylase
PHD2 (HIF prolyl-4-hydroxylase, N-terminal MYND)	54583	Prolyl-4*R*-hydroxylase
PHD3 (HIF prolyl-4-hydroxylase, No N-terminal domain)	112399	Prolyl-4*R*-hydroxylase
ABH1 (alkylated DNA repair protein alkB homolog ABH1)	8846	3-meC demethylase (DNA/RNA)
ABH2 (alkylated DNA repair protein alkB homolog ABH2)	121642	1-meA/3-meC demethylase (DNA/RNA)
ABH3 (alkylated DNA repair protein alkB homolog ABH3)	221120	1-meA/3-meC demethylase (DNA/RNA)
ABH4 (alkylated DNA repair protein alkB homolog ABH4)	54784	
ABH5 (alkylated DNA repair protein alkB homolog ABH5)	54890	
ABH6 (alkylated DNA repair protein alkB homolog ABH6)	84964	
ABH7 (alkylated DNA repair protein alkB homolog ABH7)	84266	
ABH8 (alkylated DNA repair protein alkB homolog ABH8)	91801	5-Methoxycarbonyl-methyluridine (*S*)-hydroxylase (modified RNA)[77]

ABH3 (alkylated DNA repair protein alkB homolog ABH3)	221120	1-meA/3-meC demethylase (DNA/RNA)
ABH4 (alkylated DNA repair protein alkB homolog ABH4)	54784	
ABH5 (alkylated DNA repair protein alkB homolog ABH5)	54890	
ABH6 (alkylated DNA repair protein alkB homolog ABH6)	84964	
ABH7 (alkylated DNA repair protein alkB homolog ABH7)	84266	
ABH8 (alkylated DNA repair protein alkB homolog ABH8)	91801	5-Methoxycarbonyl-methyluridine (*S*)-hydroxylase (modified RNA)[77]
FTO (fat mass and obesity associated)	79068	3-meT demethylase (DNA/RNA)
TET1, TET2, TET3 (ten-eleven translocation 1 gene protein)	80312	5-meC hydroxylase (DNA)
	54790	
	200424	
PAHX (phytanoyl-CoA hydroxylase precursor)	5264	Phytanoyl-CoA 2-*threo*-hydroxylase
PHYHD1	254295	
OGFOD1 (2OG, Fe dependent oxygenase domain 1)	55239	
OGFOD2 (2OG, Fe dependent oxygenase domain 2)	79676	
GBBH/BBOX (γ-butyrobetaine hydroxylase)	8424	γ-Butyrobetaine 3*S*-hydroxylase
TMLH (trimethyllysine hydroxylase)	55217	Trimethyllysine 5*R*-hydroxylase



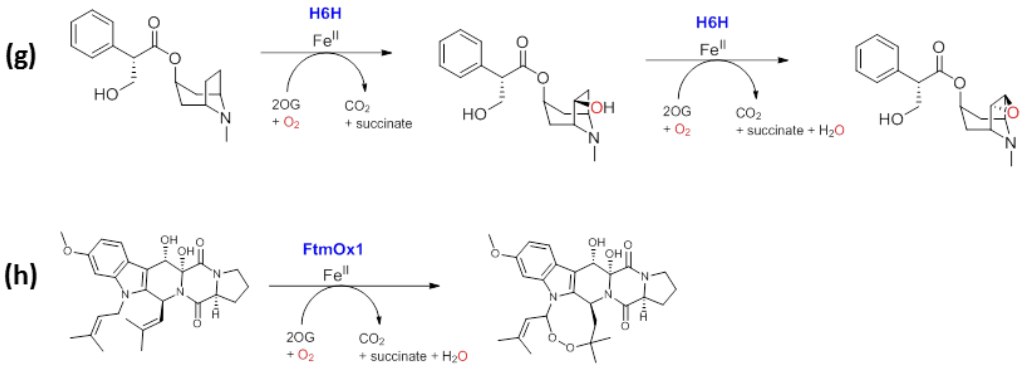


Figure 2.1 The diverse range of reactions catalysed by 2OG and Fe^{II} dependent oxygenases.
(a) BBOX catalysed hydroxylation of γ-butyrobetaine to give L-carnitine. (b) Histone lysine demethylation reaction catalysed by histone lysine demethylases (KDM). The demethylation proceeds *via* hydroxylation. (c) Reaction catalysed by SyrB2, a model halogenase. The substrate of this enzyme is L-threonine in thio-ester linkage to a phosphopantetheine group bound to the SyrB1 protein. (d) The trifunctional enzyme CAS (clavaminate synthase). In addition to its oxidative ring formation reaction (third reaction), CAS also catalyses hydroxylation and desaturation reaction. (e) Reactions catalysed by CarC (carbapenem synthase); example of epimerisation (first reaction) and desaturation (second reaction). (f) Ring expansion reaction catalysed by DAOCS (deacetoxycephalosporin C synthase), R = δ-(L-α-aminopropyl). (g) Sequential hydroxylation and epoxidation reaction catalysed by H6H (hyoscyamine 6 β-hydroxylase). (h) Formation of endperoxide catalysed by FtmOx1 (fumitremorgin B endoperoxidase).

.

The total number of 2OG oxygenases found in biology is extremely large. Humans and other animals are believed to contain about 70 such enzymes.

2.1.1 2OG oxygenase catalysed oxidation

2OG oxygenases catalyse oxidation reactions by incorporating a single oxygen atom from molecular oxygen into their substrates. The substrate oxidation is always coupled with 2OG oxidation into succinate and carbon dioxide (Figure 2.2).[12-14]

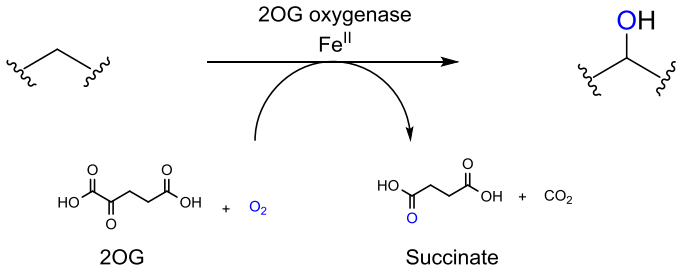


Figure 2.2 2OG oxygenase catalysed oxidation (hydroxylation) reaction.
The co-substrates are 2OG and molecular oxygen and the co-products are succinate and carbon dioxide. 2OG oxygenases use Fe^{II} as co-factor.

Although the 2OG oxygenases catalyse the oxidation of a wide range of substrates, they are believed to follow a consensus ordered sequential catalytic mechanism as shown in Figure 2.3.[12-14] The catalytic cycle begins with the binding of 2OG to the enzyme in the resting state

(with Fe^{II} bound in the active site) by displacing two water ligands (step 1).[15,16] 2OG coordinates to the active site Fe^{II} in a bidentate manner *via* its C-1 carboxylate oxygen and C-2 keto oxygen. In the next step, the substrate binds the enzyme in close proximity to the metal centre (but does not directly coordinate to the metal) inside the active site (step 2). The substrate binding cause a conformational change in the enzyme active site which leads to the loss of the last bound water molecule.[17] This creates a free co-ordination site for the binding of an oxygen molecule to the active site metal. In the next step, molecular oxygen binds to the enzyme, likely in a position directed towards the substrate molecule (step 3). The Fe^{III}-superoxide specie then attacks on the C-2 carbonyl carbon of 2OG forming a peroxide (step 4), which then collapse to form CO_2, Fe^{IV}-bound succinate and highly oxidising Fe^{IV}-oxo species (step 5).[18-21]

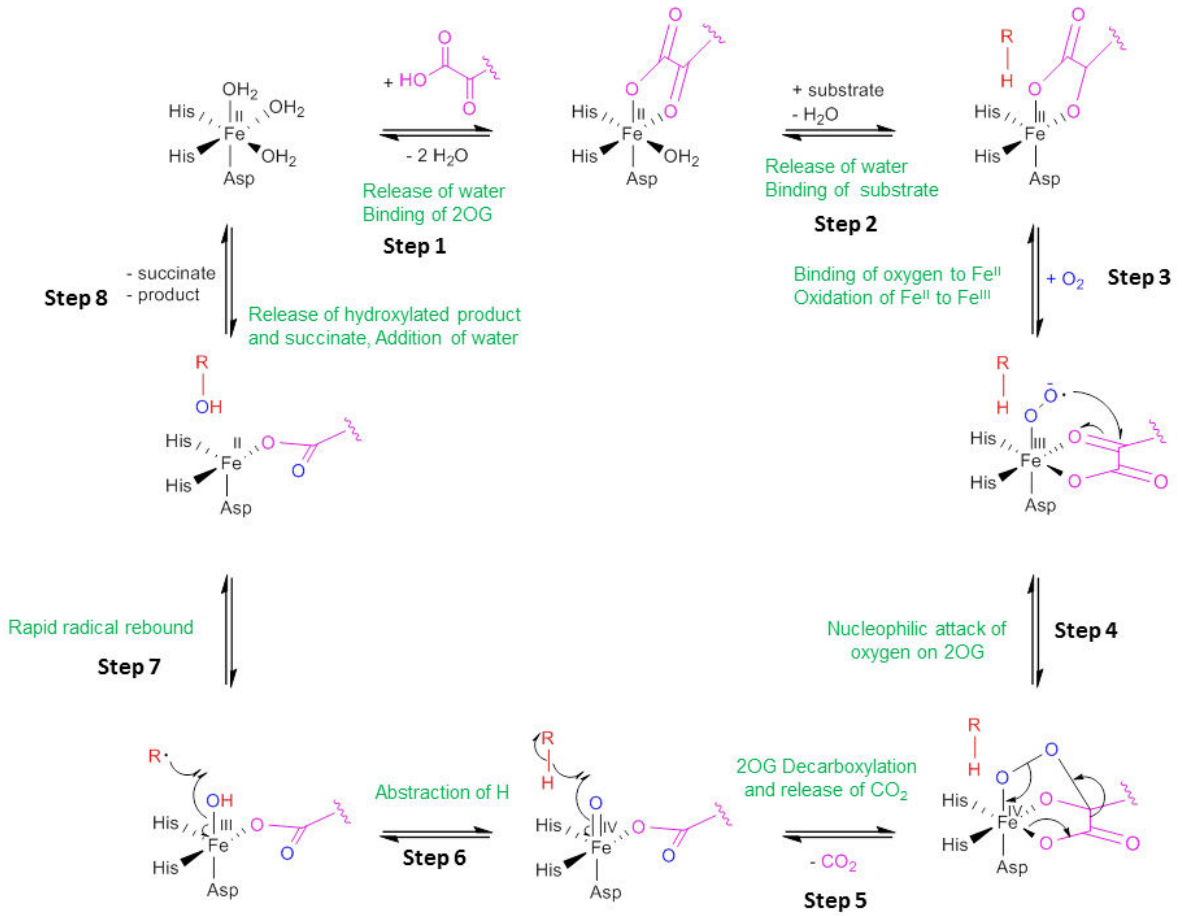


Figure 2.3 Proposed consensus mechanism for 2OG oxygenase catalysed hydroxylation.
In the resting state of the enzyme, the active site Fe^{II} is octahedrally co-ordinated mainly by a conserved HXD...H motif. The remaining 3 co-ordination sites are occupied by 3 water molecules. In catalysis, the binding of 2OG (pink) (which occurs as a result of displacement of two bound water molecules) and then substrate (red) occurs sequentially in the active site. The substrate binding results in a conformational change in the enzyme active site, which results in the loss of third water molecule. This creates a coordination site for molecular oxygen to bind. After oxygen binding, decarboxylation of 2OG occurs and a highly reactive Fe^{IV}=O intermediate is created. This ferryl intermediate then carrys out the hydroxylation of the substrate. Finally, the hydroxylated product and succinate are released from the active site, allowing water molecules to recoordinate to the metal centre. Blue oxygen atoms represent those derived from molecular oxygen.

This Fe^{IV}-oxo intermediate is positioned adjacent to the target C–H bond of the substrate, such that there is either direct insertion of an oxygen atom into this C–H bond or the hydrogen atom is abstracted by the Fe^{IV}-oxo intermediate (step 6). A rapid re-bound of the resulting hydroxyl group onto the generated carbon radical results in substrate hydroxylation (step 7). Finally, the hydroxylated product and succinate are released from the enzyme active site and the vacant co-ordination sites are re-occupied by water molecules (step 8). Thus, in this enzyme catalysed oxidation, one of the oxygen atoms of molecular oxygen is incorporated into succinate, and the other into the substrate,[22] with the reactive Fe^{IV} intermediate reduced to Fe^{II}.[23]

There are however, many examples of 2OG oxygenases which also catalyse the uncoupled 2OG decarboxylation. This usually occurs either in the absence of a substrate or in the presence of a sub-optimal substrate.[2] In such cases, it has been proposed that the Fe^{IV}-oxo specie produced is reduced back to Fe^{II} by either a reducing agent used in the reaction assay (such as L-ascorbate),[24] water (similar to Fenton chemistry),[25,26] buffer[27] or by the protein residues close to the metal centre inside the active site (self-hydroxylation).[28,29]

Although most 2OG oxygenases are believed to follow the above sequential and consensus mechanism, (Figure 2.3) a relatively different mechanistic cycle has been proposed for the catalysis by the bacterial 2OG oxygenase deacetoxycephalosporin C synthase (DAOCS) for two sub-optimal substrates, penicillin G and ampicillin, based on x-ray crystal structures (PDB IDs: 1UOB and 1UNB, respectively).[30] According to the proposed mechanism, first 2OG and then molecular oxygen binds to the enzyme active site metal. However, due to the overlap of the 2OG and substrate binding sites, the binding of the substrate to the enzyme in the presence of 2OG is prevented. Hence, oxidation of 2OG without substrate occurs and the reactive Fe^{IV}-oxo intermediate is generated. In the next step, succinate and carbon dioxide are released from the active site, which creates space for the substrate to bind. The enzyme then oxidises the substrate and with the release of the product, the enzyme is recycled to its resting form. However, in recent work by Tarhonskaya *et al.* using detailed kinetics and binding studies suggests that a consensus mechanism similar to other 2OG oxygenases also likely operates in DAOCS catalysis.[31]

2.1.2 Co-factor and co-substrate specificity of 2OG oxygenases

The presence of bound Fe^{II} inside the active site is an absolute requirement for 2OG oxygenases catalysis. Many transition metal ions including Mn^{II}, Zn^{II}, Ni^{II}, Cu^{II} and Co^{II}

binds in the Fe^{II} active site of many 2OG oxygenases in a similar mode and affinity as Fe^{II},[32] but, however, none of these metal ions has been reported to replace Fe^{II} for 2OG oxygenase catalysis.[33-35] After catalysis, when 2OG oxygenase is recycled to its resting form, Fe^{II} is retained inside the active site. It is still unclear whether an active shuttle system exists for the delivery of iron from cells into the active site of 2OG oxygenases or the enzymes just simply compete for the iron in cells. The catalytic activity of many 2OG oxygenases in *in vitro* assays is also enhanced in the presence of reducing agent such as L-ascorbate.[24,36] It is proposed that the reducing agent reduces Fe^{III} into Fe^{II}, thus maintaining a pool of Fe^{II} in solution and promote maximum enzyme catalytic activity.

2OG oxygenases also have an absolute requirement for 2OG as an oxidisable co-substrate. So far, only 2-ketoadepic acid and α-bromo 2OG have been reported to act as 2OG alternatives in catalysis by 2OG oxygenases, but with significantly reduced activity compared to 2OG.[37-39] Recently, it has also been reported that 2-hydroxyglutarate (2-HG) can enables hydroxylation reaction for some 2OG oxygenases, however in these reactions, the 2-hydroxyglutarate is first non-enzymatically oxidised into 2OG by Fe^{II} and ascorbate likely *via* Fenton chemistry, and then the 2OG enables hydroxylation.[40] No direct binding of 2-hydroxyglutarate to any of these 2OG oxygenases have been reported.

2.1.3 Structural aspects of 2OG oxygenases

Crystallographic studies have shown that 2OG oxygenases are characterised by the presence of a conserved double-stranded β-helix (DSBH) core fold, which is formed of eight antiparallel β-strands, which form a squashed barrel type structure, open at one end (Figure 2.4).[5,41-43] Inside the DSBH core fold, is a highly conserved 2-His-1-carboxylate (HXD/E...H) amino acid residue triad. In the resting state, the active site Fe^{II} is coordinated by two histidine residues and one aspartic acid or glutamic acid residue (Figure 2.4),[44] while the other three sites are typically occupied by the oxygens of three water ligands. Studies have shown that these two histidine residues are important for the binding of Fe^{II}.[45] 2OG coordinates to the active site metal in a bidentate manner its *via* C-1 carboxylate oxygen and C-2 keto oxygen (Figure 2.4). The 2OG C-5 carboxylate group projects towards inside of the active site pocket and is always held by electrostatic (usually with an arginine or lysine) and hydrogen bonding interactions (e.g with serine or tyrosine residues) (Figure 2.4). Crystallographic studies have shown that the substrate carbon atom undergoing oxidation is observed to be proximate (~4.5 A˚) to the iron. The substrates of 2OG oxygenases are bound

by a combination of hydrogen bonding, hydrophobic, and electrostatic interactions in such a way that the carbon atom undergoing oxidation projects towards the iron. In many cases it is observed that the binding of substrate in the active site region involves an induced fit binding. For example substantial enzyme conformational changes have been observed in the hypoxia-inducible factor (HIF) prolyl hydroxylase, PHD2, a human oxygen sensing 2OG oxygenase, isopenicillin N synthase (IPNS), a microbial 2OG oxygenase and the *Escherichia coli* alkylation repair 2OG oxygenase AlkB.[41,42] The correct positioning of 2OG and the substrate to the active site Fe^{II} is important in order to avoid the uncoupled oxidation of 2OG to primary substrate oxidation and oxidation of other atoms including self-hydroxylation.

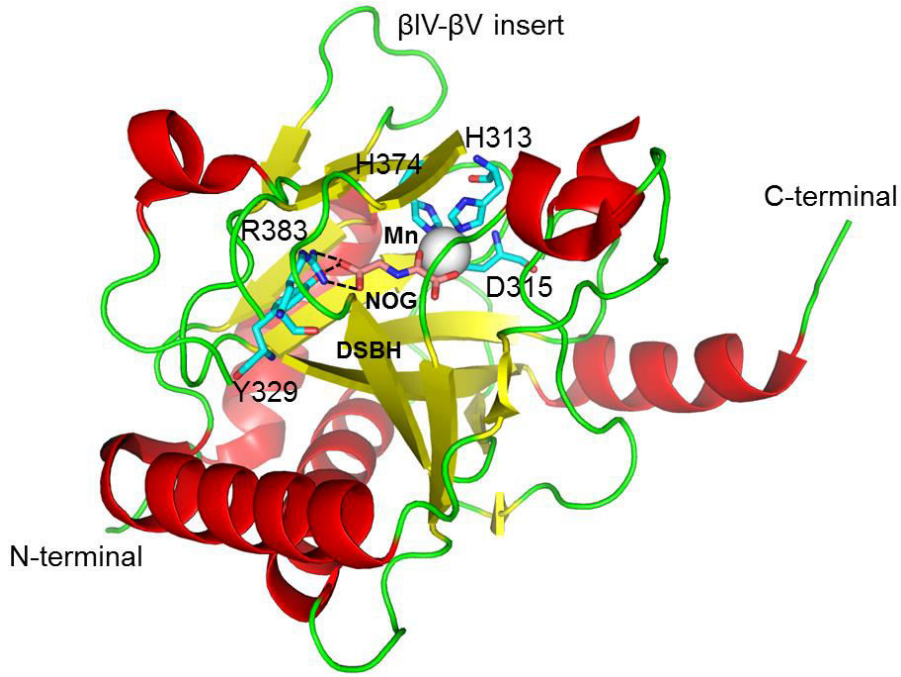


Figure 2.4 General structural features of 2OG oxygenases.
View from a crystal structure of the catalytic domain of hypoxia inducible factor prolyl hydroxylase PHD2 (PDB ID: 3HQR), showing the DSBH fold (yellow), binding of Mn^{II} (grey sphere; unreactive substitute for Fe^{II}) and NOG (candy pink; *N*-oxalylglycine, unreactive substitute for 2OG) and an insert in β-sheets of DSBH fold.

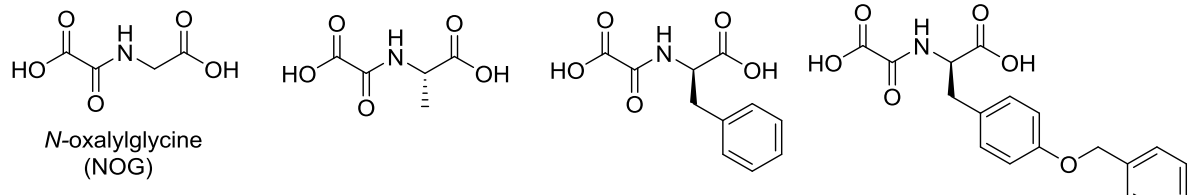
In addition to DSBH fold, 2OG oxygenases are also characterised by other semi-conserved secondary structural features.[46] These include α-helices and β-strands at the N-terminus of the DSBH and in many cases an α-helix at the C-terminus of the DSBH. Discrete domains are also located either at the N-/C-terminal regions. Sub-families also have characteristic inserts frequently occur in the loops between the β-strands of the DSBH, particularly between strands βIV and βV. This insert ranges from a small loop of a few amino acids to a whole protein domain (Figure 2.4).[5,41-43] These additional structural elements are believed to help in substrate/co-substrate binding or in stability of the enzyme active conformation.

2.1.4 Inhibition of 2OG oxygenases

Many members of 2OG oxygenases are currently being targeted for therapeutic and commercial purposes. For example, human 2OG oxygenases, such as the hypoxia inducible factor (HIF) hydroxylases which are involved in human oxygen sensing system, are the inhibition targets for the treatment of diseases including cancer, ischemia, inflammation and anaemia.[47-50] γ-Butyrobetaine hydroxylase is targeted for the treatment of myocardial infarction.[51] Similarly, plant 2OG oxygenases such as gibberellin C-20 oxidase are an inhibition targets for the development of commercially used herbicides.[52]

The inhibitors of 2OG oxygenases can be divided into two categories. The first categories of inhibitors are 2OG competitors (Figure 2.5) and the second categories of inhibitors are substrates competitors (Figure 2.6). However, most of the reported small molecule inhibitors of 2OG oxygenases are 2OG competitors binding to the Fe^{II} inside the enzyme active site to prevent catalysis.[6] These inhibitors also chelate the free Fe^{II} in solution and thus prevent catalysis by stopping the Fe^{II} from binding to the enzyme.[53] Based on the structure of the core skeleton, these 2OG mimic inhibitors of 2OG oxygenases are generally of several types.[6] These include *N*-oxalyl amino acids, 2,x-pyridine dicarboxylic acids, bipyridyl compounds, hydroxyquinoline/isoquinolinyl derivatives and *N*-hydroxythiazoles derivatives (Figure 2.5a). Although all these 2OG-based inhibitors contain different functional groups, they all however, appear to bind to the enzyme active site Fe^{II} [6] (Figure 2.5b-d) (and free Fe^{II} in solution)[53] in a bidentate manner similar to 2OG and inhibit catalysis by competing with 2OG.[6]

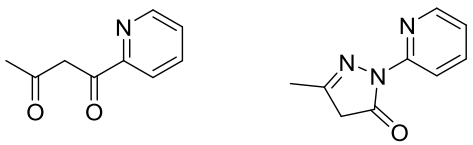
(a)



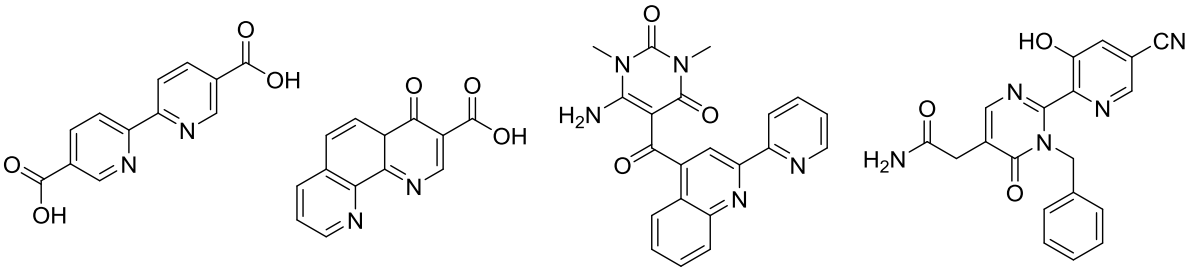
N-oxalylglycine
(NOG)

N-oxalyl-amino acid-based inhibitors of 2OG oxygenases

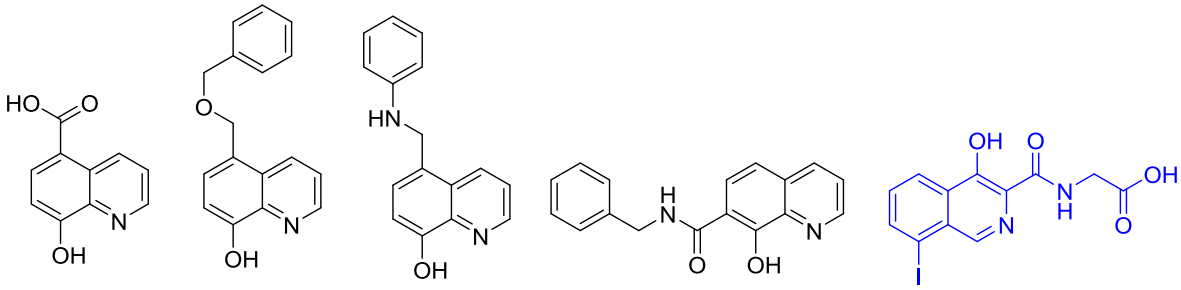
2,4-Pyridinedicarboxylic acid



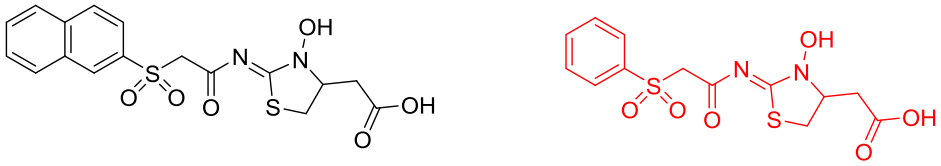
Pyridine carboxylic acid core containing inhibitors of 2OG oxygenases



Bipyridyl core containing inhibitors of 2OG oxygenases

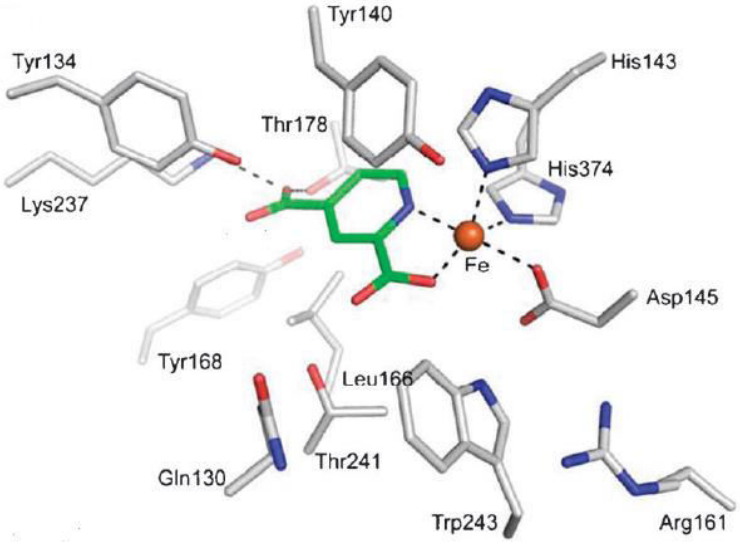


Hydroxyquinoline/isoquinoline-based inhibitors of 2OG oxygenases

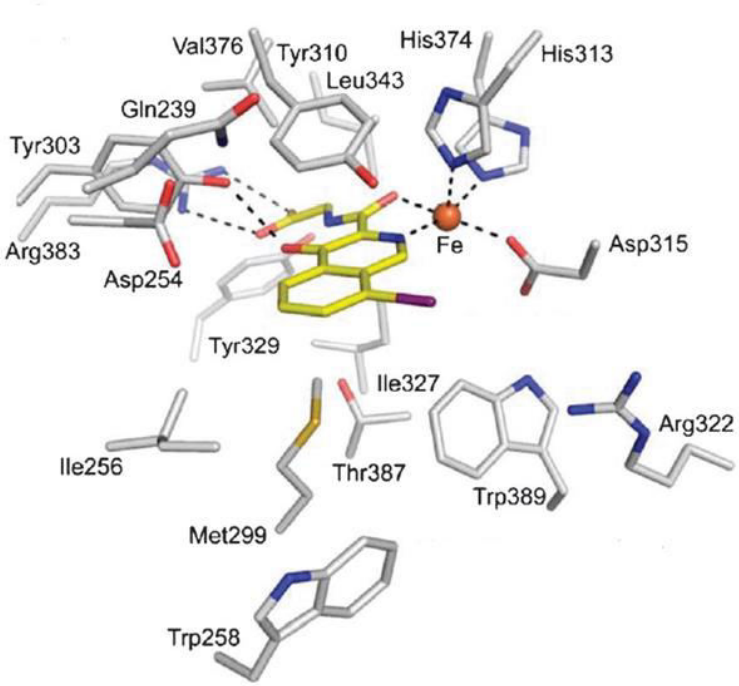


N-hydroxythiazole-based inhibitors of 2OG oxygenases

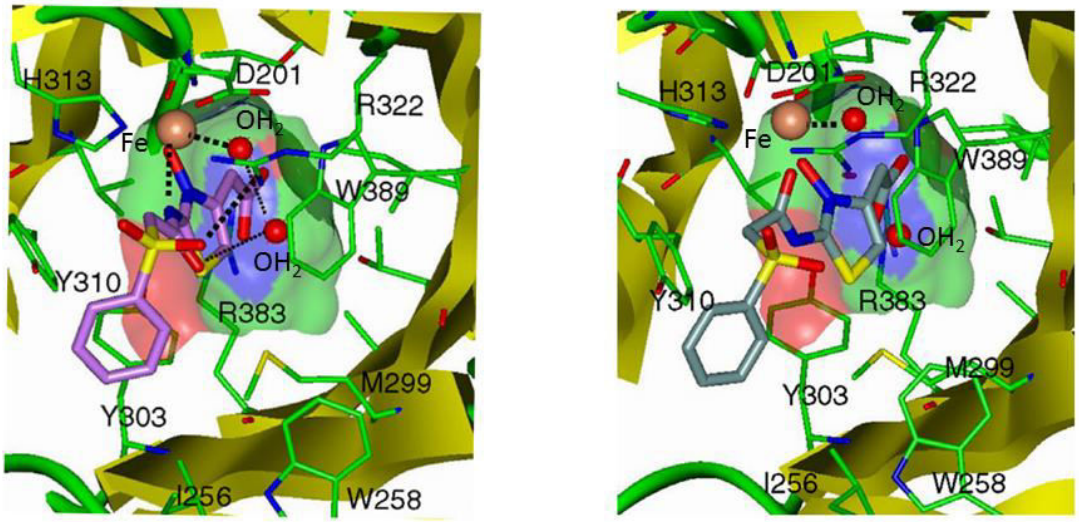
(b) C-P4H.Fe^{II}.2,4-PDCA



(c) PHD2.Fe^{II}. glycinamide 3-hydroxy-7-iodo-isoquinolinyl-2-(*N*-glycinyl)-carboxamide



(d) PHD2.Fe^{II}. 2-(3-hydroxy-2-((2-(phenylsulfonyl)acetyl)imino)-2,3-dihydrothiazol-4-yl)acetic acid



binding mode 1 binding mode 2

Figure 2.5 Different types 2OG-based inhibitors of 2OG oxygenases.
(a) Structure of the common 2OG mimic inhibitors with different core skeletons. (b) View from a crystal structure of collagen prolyl-4hydroxylase (C-P4H) in complex with 2,4-PDCA (coloured green in **a**) and Zn^{II} (PDB ID: 2JIG).(c) View from a crystal structure of PHD2 in complex with the bicyclic glycinamide 3-hydroxy-7-iodo-isoquinolinyl-2-(*N*-glycinyl)-carboxamide (coloured blue in **a**) (PDB ID: 2G19) (d) (Reproduced from reference [6]). Docked structures of 2-(3-hydroxy-2-((2-(phenylsulfonyl)acetyl)imino)-2,3-dihydrothiazol-4-yl)acetic acid (coloured red in **a**) in the active site of PHD2 (PDB ID: 2G19), showing two possible active site metal chelation modes.(d: reproduced from reference [54])

The functionalization of the core skeleton of these inhibitors can enable selectivity and potency of inhibition towards a particular 2OG oxygenase.[6] Amongst these inhibitors, *N*-oxalyl amino acids and 2,x-pyridine dicarboxylic acids derivatives are rather generic 2OG oxygenases inhibitors as they closely mimic 2OG. *N*-Oxalylglycine (NOG) and 2,x-pyridine dicarboxylic acids are the simplest inhibitors amongst these two classes. Selectivity of NOG for different 2OG oxygenases has been achieved by substituting the glycine moiety with other L/D amino acids.[55] NOG chelates the active site Fe^{II} *via* its amido and C-1 carboxylate oxygen, while the pyridine 2,x dicarboxylic acids inhibitors chelate the active site Fe^{II} through their pyridinyl nitrogen and 2-carboxylate group.[6]

Bipyridyl derivatives are also established 2OG oxygenase inhibitors.[6] Their selectivity amongst various 2OG oxygenases have been achieved by functionalisation of the bipyridyl core skeleton.[56,57] They inhibit the activity of many 2OG oxygenases such as the collagen P4H, the HIF hydroxylases and the histone demethylases.[56,57]

Similarly, the hydroxyquinoline/isoquinolinyl derivatives[6,58] and the structurally related *N*-hydroxythiazoles-based compounds[54,59] are very potent inhibitors of many 2OG oxygenases including the PHD2 and the histone demethylases. X-ray crystallographic studies suggest that in addition to competing with 2OG, these inhibitors also tend to compete with the enzyme substrate. For the hydroxyquinoline/isoquinoline derivatives, their selectivity for different 2OG oxygenases has been achieved by functionalization of the core hydroxyquinoline/isoquinoline skeleton and the side chain amino acid.[60]

There are a fewer substrate mimic inhibitors of 2OG oxygenases (Figure 2.6). For example some of the inhibitors of PHD2 are based on its substrate HIF-1α (hypoxia inducible factor) peptide fragment *C*-terminal oxygen-dependent degradation domain (CODD; HIF-1α$_{556-574}$) (Figure 2.6). In these inhibitors, the proline residue that undergoes the PHD2 catalysed hydroxylation (coloured blue/underlined) is replaced by unhydroxylable proline analogues such as L-3,4-dehydroproline and L-4-thioproline (Figure 2.6). Both these CODD-based inhibitors selectively inhibit PHD2 over FIH, another HIF hydroxylase.[61] Mildronate[62] and methyl-GBB[63] are also substrates-based inhibitor of γ-butyrobetaine hydroxylase (BBOX) (Figure 2.6), competing with the natural substrate GBB for C-3 stereoselective hydroxylation. In mildronate, the GBB C-4 position CH_2 group is replaced by a NH group, whereas, in methyl-GBB, one of the methyl groups in the GBB trimethylammonium moiety is extended to an ethyl group.

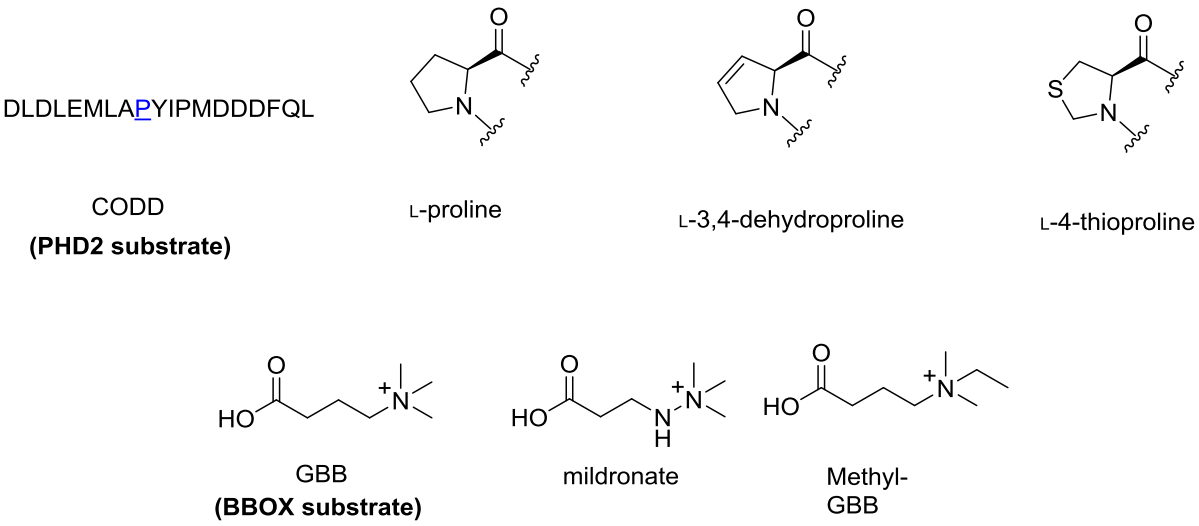


Figure 2.6 Examples of substrate-based inhibitors of 2OG oxygenases.

In addition to the 2OG-based and substrate-based inhibitors of 2OG oxygenases, many metals ions such as Zn^{II}, Ni^{II}, Co^{II} etc. are inhibitors of 2OG oxygenases.[33-35] These metal ion inhibitors compete with Fe^{II} for binding in the enzyme active site and thus prevent catalysis.

Some 2OG oxygenases have also been reported to contain structural zinc ions. These zinc binding domains have been found as highly critical for the stability of the folded (active) conformation of these enzymes. Such 2OG oxygenases include histone demethylases (Figure 2.7) and γ-butyrobetaine hydroxylase (BBOX) (details in section 2.2.11). Therefore specific compounds termed as zinc ejectors have been developed to irreversibly remove the structurally important Zn^{II} from the zinc binding site of 2OG oxygenases (Figure 2.7). As a result of zinc ejection, the enzyme unfolds and become inactive. Amongst such compounds, selenium and sulfur containing compounds have been reported as effective zinc ejectors, due to their demonstrated ability to interact with the cysteine-residues involved in Zn^{II} binding in many proteins.[64-66] Examples of these zinc ejectors are disulfiram and ebselen (Figure 2.7), which have been shown to irreversibly inactivate histone demethylases[65] and γ-butyrobetaine hydroxylase[67] by `removing´ the structurally important Zn^{II} from the zinc binding domains of these enzymes.

Oxygen mimic compound such as nitric oxide, has been shown to inhibit PHD2. It is believed that nitric oxide competes with oxygen and thus prevents catalysis.[68,69] In addition, common metal ion chelators such as ethylenediaminetetraacetic acid (EDTA) also removes the active

site Fe^{II} from most 2OG oxygenases including P4H[70] and bacterial AlkB[71] and thus inhibit their catalysis.

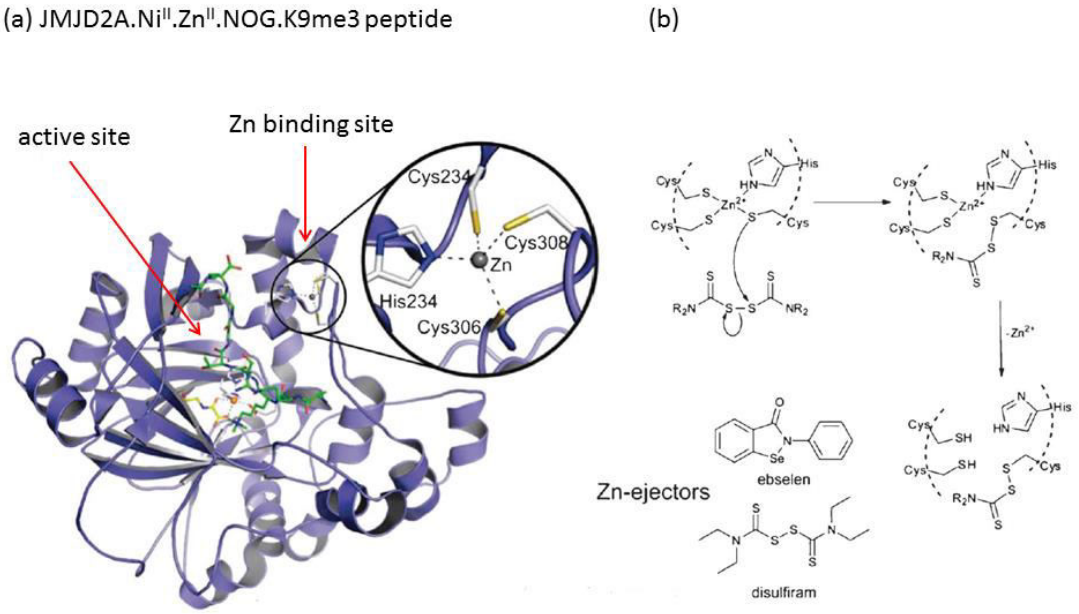


Figure 2.7 Inhibition of the histone demethylase JMJD2A by ejection of structural Zn^{II}.
(a) View from a crystal structure of JMJD2A in complex with Ni^{II}, Zn^{II} and $H_3K_9me_3$ peptide substrate (PDB ID: 2OQ6). Zn^{II} is coordinated by three cysteines and one histidine residues (zoom in view; shown in the circle) serving to hold together the two loops of the enzyme. (b) This structurally critical Zn^{II} is irreversibly removed from the enzyme by 'Zn-ejectors' such as disulfiram and ebselen through possible mechanism shown. (Reproduced from reference [6])

2.2 γ-Butyrobetaine hydroxylase (BBOX)

2.2.1 L-Carnitine

L-Carnitine (Figure 2.8) is an essential metabolite,[72-75] found in almost all animal species, numerous micro-organisms and plants.[76-79] In humans, its known primary function is to act as a shuttle system for the transport of the activated long-chain fatty acids (LCFA) from cytosol into mitochondrial matrix (Figure 2.9).[80-83] In mitochondria, these long chain fatty acids undergoes β-oxidation to produce metabolic energy.[84,85]

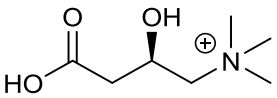


Figure 2.8 Structure of L-Carnitine

In humans, under normal conditions, the energy consumed by heart, liver and skeletal muscles mostly comes from fatty acid metabolism.[86,87] Other functions of L-carnitine include,

transportation of products of peroxisomal β-oxidation including acetyl-CoA to the mitochondria for oxidation to H_2O and CO_2 in the TCA cycle,[88,89] modulation of the acyl-CoA/CoA ratio,[88,90] storage of energy as acetylcarnitine[75,91] and the modulation of the toxic effects of poorly metabolised acyl group by excreting them as carnitine esters.[92] All these biochemical roles of L-carnitine involve reversible esterification of its C3-hydroxyl group.[93,94] The D-form of carnitine is reported to be biologically inactive.[95]

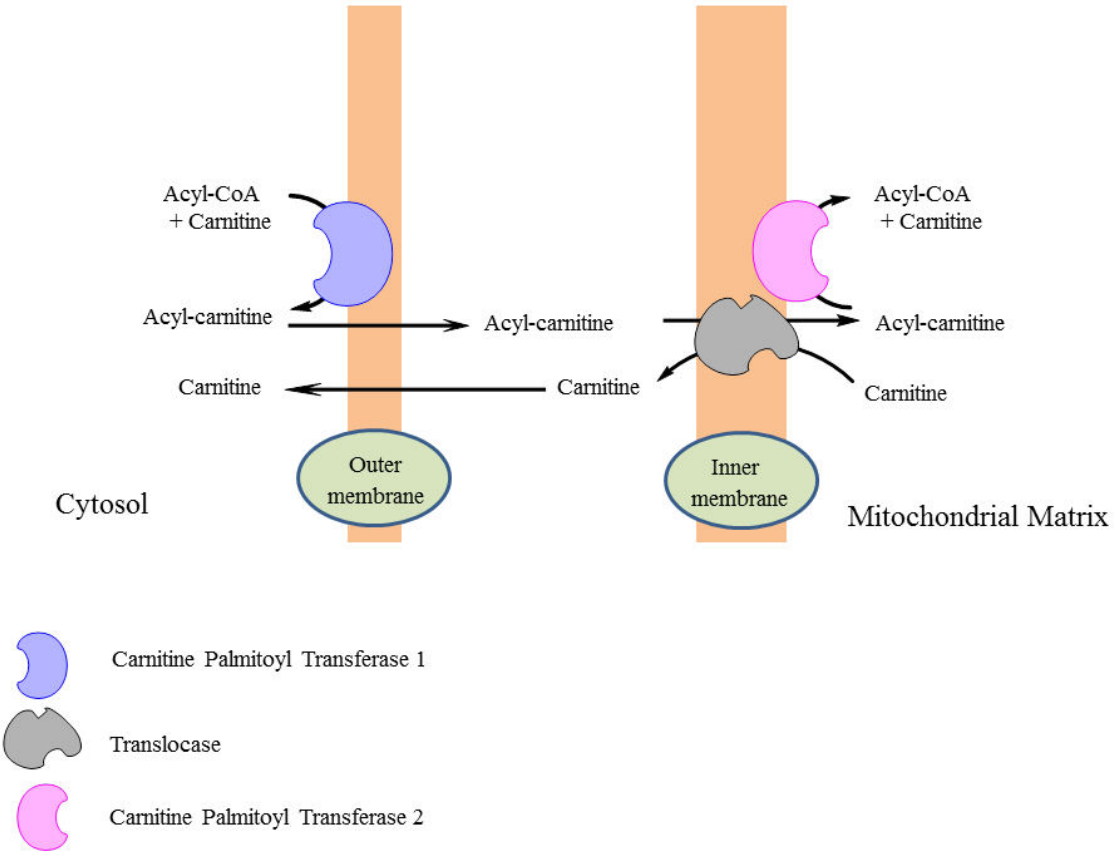


Figure 2.9 L-Carnitine shuttle system of long chain fatty acids from cytosol into mitochondria.
Fatty acids break down occurs in mitochondrial matrix producing metabolic energy (ATP) *via* the TCA cycle. The free fatty acid in the cytosol is attached with a thioester bond to coenzyme A (CoA). Acyl-CoA is transferred to the hydroxyl group of carnitine by carnitine palmitoyl transferase 1, located on the outer mitochondrial membrane. Acyl-carnitine is shuttled inside by a translocase. Acyl-carnitine is converted into Acyl-CoA and carnitine by a carnitine palmitoyl transferase 2, located on the inner mitochondrial membrane. The carnitine liberated in the mitochondrial matrix diffuses back to cytosol by a translocase, whereas the Acyl-CoA undergoes β-oxidation and is converted into Acetyl-CoA. Finally, Acetyl-CoA enters the TCA cycle and produces ATP (metabolic energy).

In humans, it is believed that ~ 75% of the total carnitine in the body is obtained from the diet, while ~ 25% comes from endogenous synthesis.[72,77,96] Red meat, fish and dairy products are rich sources of carnitine. Health conditions like, muscle weakness, hypoglycaemia, cardiomyopathy and metabolic acidosis have been associated with the deficiency of L-

carnitine.[97] An active sodium dependent carnitine transporter has been reported in humans (called hOCTN2) and in rat intestine (called CT1 or rat OCTN2).

2.2.2 L-Carnitine biosynthesis

An active and highly conserved L-carnitine biosynthetic pathway exists in many eukaryotes (mammals, plants) and prokaryotes (bacteria).[80,98-100] The overall L-carnitine biosynthesis consists of four enzyme catalysed steps (Figure 2.10).

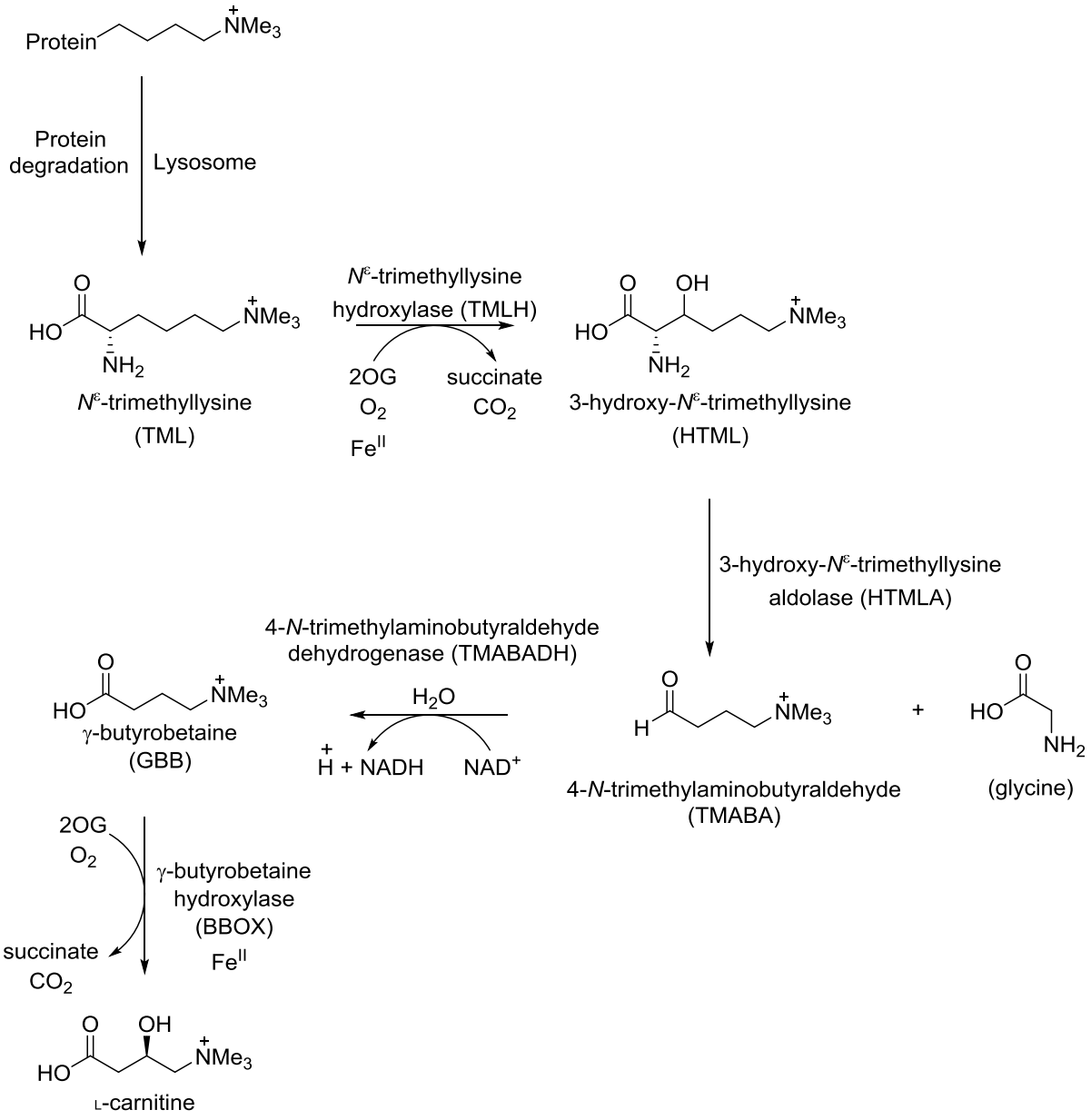


Figure 2.10 Carnitine biosynthesis

The bioynthsis of L-carnitine starts from N^{ϵ}-trimethyllysine,[101-103] which in proteins is formed from amino acids lysine and methionine. Lysine forms the carbon backbone of L-carnitine,[104,105] while the 4-*N*-methyl groups originate from methionine.[106] Many mammalian proteins, such as histones, myosin, actin and cytochrome c contain N^{6}-trimethyllysine residues.[107] *N*-methylation of these lysine residues occurs as a result of post translation modification of these proteins.[108,109] This reaction is catalysed by methyltransferases, which utilize S-adenosylmethionine as a methyl donor.[107] Lysosomal hydrolysis of these proteins eventually releases trimethyllysine (TML).

In the first step of L-carnitine biosynthesis, trimethyllysine hydroxylase (TMLH; EC 1.14.11.8) catalyses the hydroxylation of N^{ϵ}-trimethyllysine to form 3-hydroxy-N^{ϵ}-trimethyllysine (HTML).[102,110] TMLH is a 2-oxoglutarate (2OG) and Fe^{II}-dependent oxygenase,[102,110] with similar activity as proline hydroxylases. This is the only mitochondrial enzyme involved in L-carnitine biosynthesis. Due to difficulties in production, this enzyme has not been studied in detail. It was only until 2014 that this enzyme was successfully expressed and purified from *Escherichia coli*.[111]

In the second step, 3-hydroxy-N^{ϵ}-trimethyllysine is cleaved into 4-*N*-trimethylaminobutyraldehyde (TMABA) and glycine. This reaction is catalysed by hydroxy-N^{ϵ}-trimethyllysine aldolase (HTMLA; EC 4.1.2.‘ X ’).[77,102] In mammals, no pathway involving hydroxy-N^{ϵ}-trimethyllysine aldolase enzyme has so far been detected. A study on yeast *Candida albicans* suggests that HTMLA might belong to L-threonine aldolase class of enzymes.[112] In another study, serine hydroxymethyltransferase (SHMT), isolated from rat liver has also shown the activity of TMABA.[113] It is possible that L-threonine and SHMT have closed links together.

In the third step, the dehydrogenation of 4-*N*-trimethylaminobutyraldehyde results in the formation of 4-*N*-trimethylaminobutyrate (γ-butyrobetaine or GBB). This reaction is catalysed by cytosolic enzyme 4-*N*-trimethylaminobutyraldehyde dehydrogenase (TMABA-DH; EC 1.2.1.47).[114-116] Most cells can synthesize γ-butyrobetaine.[117] TMABA-DH has been isolated and purified both from rat liver[115] and bacterium *pseudomonas* (here after denoted by ps).[114] Rat TMABA-DH is closely related to human aldehyde dehydrogenase 9 (ADH9),[115,116,118] which catalyses the dehydrogenation of γ-amino butyraldehyde into γ-amino butyric acid.[119,120]

In the fourth and final step, 4-*N*-trimethylaminobutyrate is stereospecifically[121] hydroxylated on C-3 position to give the final product L-carnitine.[122] This reaction is catalysed by γ-butyrobetaine hydroxylase (GBBH or BBOX; EC 1.14.11.1), which is a 2OG and Fe^{II}-dependent oxygenase.[123-127] BBOX is a cytosolic enzyme found in only certain tissues, in humans, for example its high concentration is found in kidneys, liver and to a small extent in brain.[123,124] BBOX has been isolated and purified from humans kidney,[125] rat-liver,[128-130] calf liver[131] and bacterium *pseudomonas*.[100,132]

2.2.3 Inhibition of BBOX for the treatment of myocardial infarction

Amongst all the four enzymes involved in the overall carnitine biosynthesis pathway, γ-butyrobetaine hydroxylase or BBOX (encoded by BBOX1 gene in humans) is the only enzyme that has been studied in details and structurally well-characterised. BBOX is a current drug target for the treatment of myocardial infarction.[51,133-136]

Ischemic heart disease (IHD) is a serious health condition, resulting in disability and deaths of millions of people each year across the globe.[137] The pathological consequences of IHD results from the detrimental effects of the accumulation of the long chain fatty acids and the products of their metabolic activation (long chain acyl-CoA and acyl-carnitine).[138,139] Thus, the pharmacological intervention that targets the LCFA accumulation has been proposed for the development of novel treatment strategies to improve the clinical outcomes of patients with IHD.

Under normoxia, a major part of the energy required by heart tissues comes from fatty acids metabolism.[139] However, under ischemic-reperfusion conditions, a reduced use of fatty acids and a greater use of glucose oxidation as the energy source lead to better functional recovery of the myocardium (heart). It has been proved that under ischemic (low-oxygen than normal) conditions, the survival chances of the cardiac muscle cells is greatly increased by glucose oxidation, because it consumes less oxygen than fatty acid oxidation.[140]

It is possible that this shift in myocardium energy source from fatty acid metabolism to glycolysis may be achieved by a controlled decrease in the amount of bioavailable L-carnitine or inhibition of L-carnitine biosynthesis or inhibition of L-carnitine mediated fatty acids transport from cytosol into mitochondria.[141]

In earlier work, inhibitors of carnitine palmitoyl transferase-1 (CPT1) including oxophenycine (*S*-4-hydroxyphenyl glycine)[142-144] and POC (2-[5-(4-

chlorophenyl)pentyl]oxirane-2-carboxylate)[145,146] were tested. Although, these inhibitors resulted in inhibition of fatty acid β-oxidation and reduction in demand of oxygen by heart muscles, they have substantial limitations resulting directly from the mechanism of their action.

One of the more promising drugs proved so far in treatment of IHD, is mildronate or 3-(2,2,2-trimethylhydrazinium)propionate (THP), also known as MET-88, meldonium, quaterine, is a clinically used inhibitor of γ-butyrobetaine hydroxylase (BBOX).[62] It was developed by Ivars Kalvins and his associates at the Latvian institute of Organic Synthesis (OSI) in 1970. Mildronate is a structural analogue of γ-butyrobetaine (GBB), with a NH group replacing the CH_2 at the C-4 position of γ-butyrobetaine (Figure 2.11). It is the most extensively studied inhibitor of BBOX. It is marketed as a commercially available drug in Latvia, Russia, Ukraine, Georgia, Kazakhstan, Azerbaijan, Byelorussia, Uzbekistan, Moldova, and Kyrgyzstan , for the treatment of myocardial infarction,[51,133-136] diabetes,[147] and neurological disorders.[148,149] A recent study has also shown that mildronate may be beneficial for enhancing memory and learning capabilities.[150]

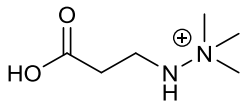


Figure 2.11 Structure of 3-(2,2,2-trimethylhydrazinium)propionate (THP).
THP is also called mildronate, an inhibitor of human BBOX, is clinically used as a cardioprotective drug.

The proposed mode of action of mildronate is believed to involve inhibition of L-carnitine biosynthesis[135] and inhibition of L-carnitine mediated transportation of activated long chain fatty acids from cytoplasmic matrix into mitochondria.[151-153] Such inhibition is proposed to prevent accumulation of cytotoxic intermediate products of fatty acid β-oxidation in ischemic tissues and to block this highly oxygen-consuming process, thus shifting the myocardial energy source from fatty acids oxidation to glucose utilization.

Initially, THP was suggested as a non-competitive inhibitor of BBOX employing enzyme kinetics studies,[62] however, study by another group under modified experimental conditions proved that it is in fact a competitive inhibitor of BBOX.[154] More recently, it has also been shown that THP acts a substrate for hBBOX forming malonic acid semialdehyde, dimethylamine, formaldehyde and 3-amino-4-(methylamino)butanoic acid (AMBA) (Figure 2.12).[155] The reported crystal structure of BBOX in complex with THP suggests that it binds

in the BBOX active site exactly in the same pocket and same mode as natural substrate GBB (Figure 2.13).[156]

Although, the use of THP has so far been proved successful in treatment of IHD, however its required dose in currently very high (1 g/day), possibly be due to its low BBOX inhibition efficiency (IC_{50} 62 µM). Therefore there is a need to develop more potent and selective hBBOX inhibitors.

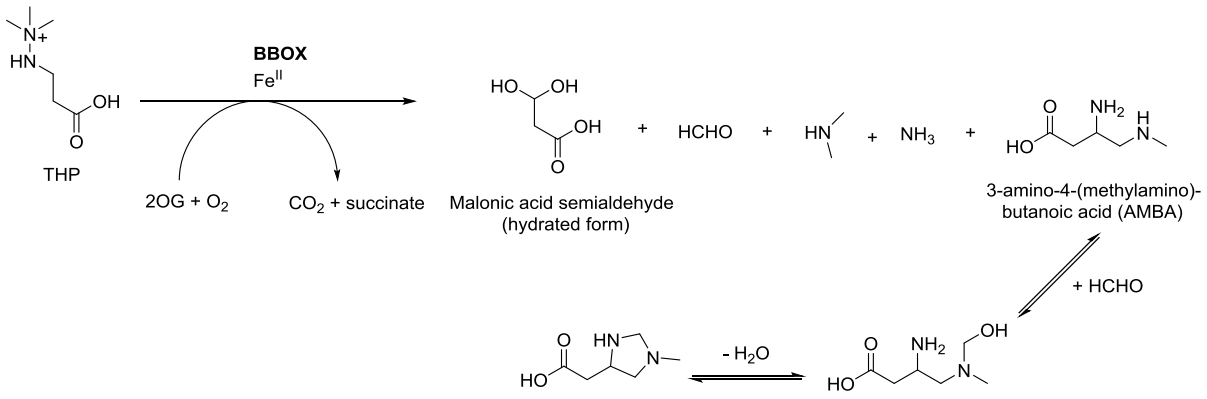


Figure 2.12 BBOX catalysed oxidation of THP.
For detailed mechanism see the reference [155].

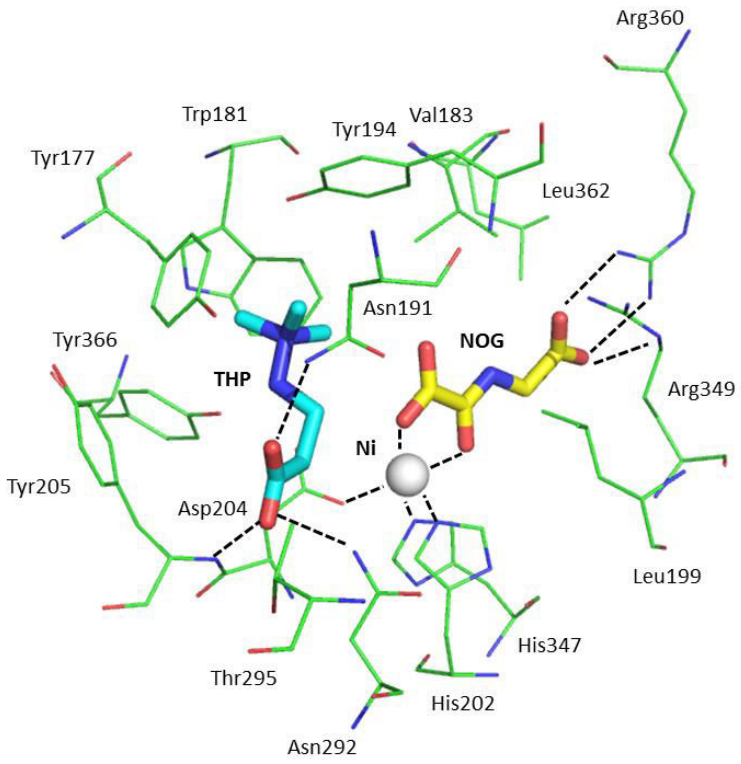
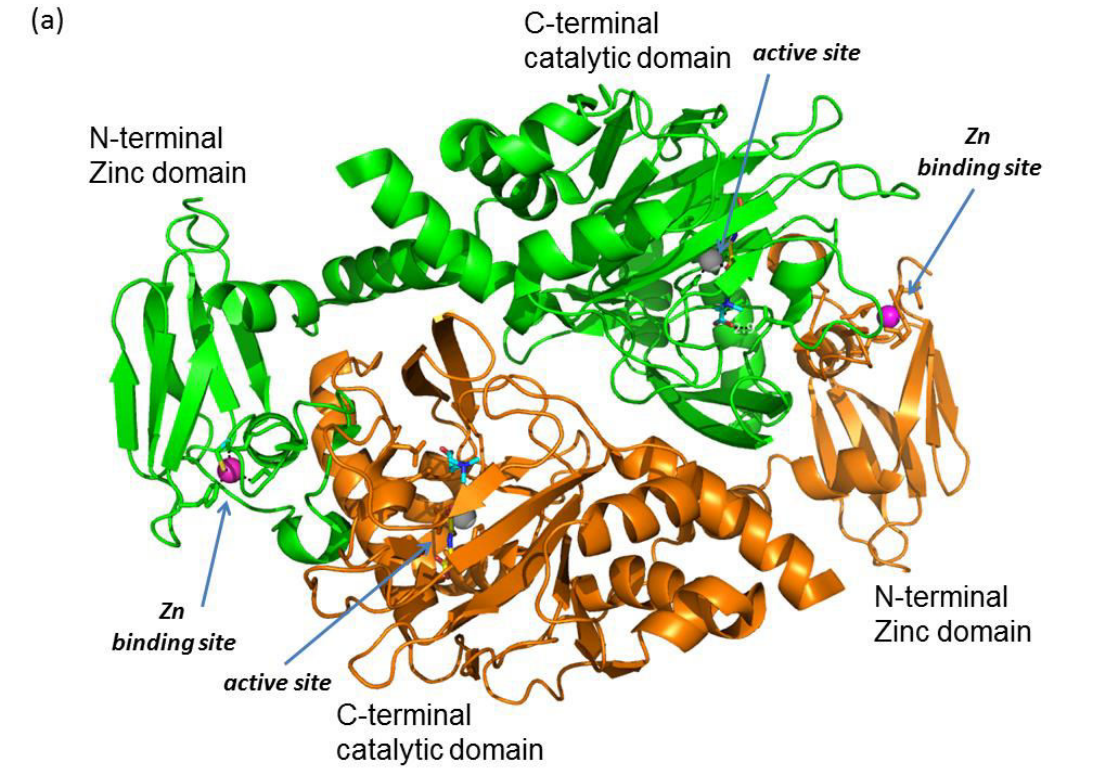


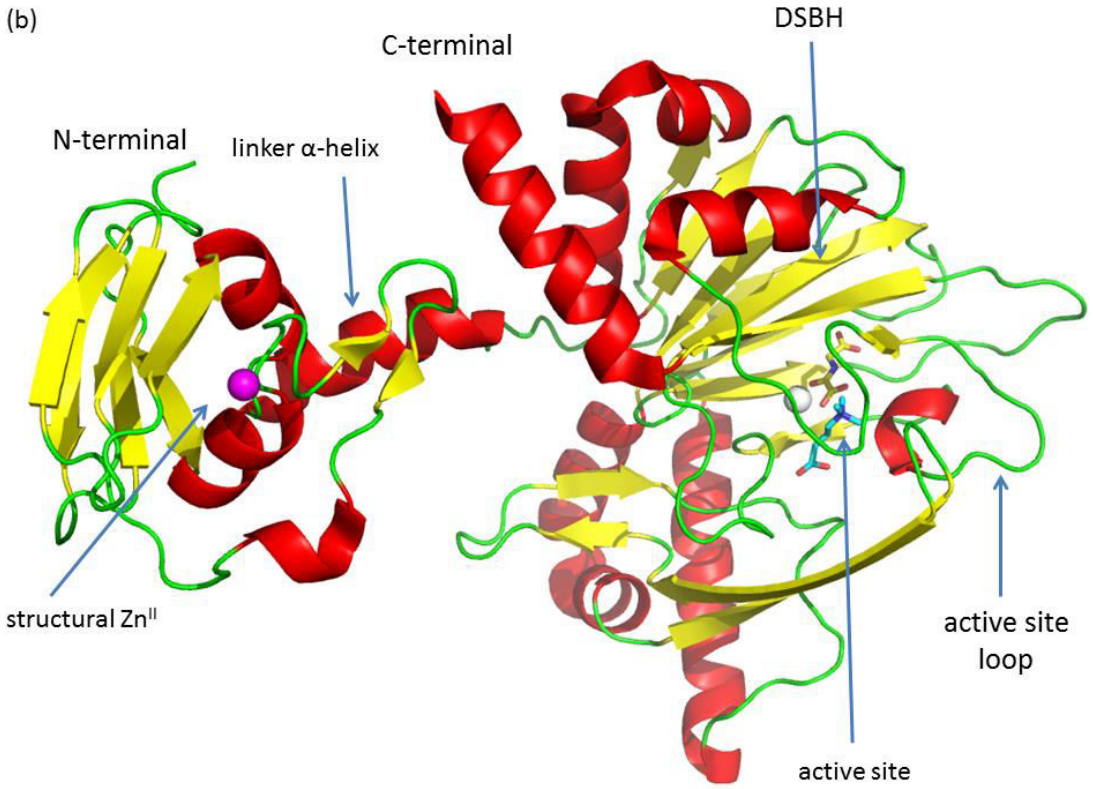
Figure 2.13 Human BBOX crystal structure in complex with THP.
View from a crystal structure of human BBOX in complex with THP (cyan), NOG (yellow colour), Ni^{II} (grey sphere) and Zn^{II} (in zinc binding site; not shown) (PDB ID: 3MS5). The THP trimethylammonium group is located in an aromatic cage formed by the side chains of Tyr-205, Tyr- 366, Tyr-177, Trp-181 and Tyr-194, while, its carboxylate group is positioned to interact with the side chains of Asn-292, Asn-191 and the backbone amide of Tyr- 205. This image was made using pymol.

2.2.4 Crystal structure of hBBOX

Amongst all the four enzymes involved in the biosynthesis of L-carnitine, crystal structures have only been solved for BBOX (Figure 2.14).[156-160] BBOX exists as a 43.3 kDa homodimer (Figure 2.14a),[128,161,162] consisting of 387 residues. Each monomer consists of two distinct domains, an N-terminal zinc binding domain and a C-terminal catalytic domain. The two domains are connected by an α-helix (residues range 97-110) (Figure 2.14b). Dimerization occurs *via* its N-terminal zinc binding domains to interact with the opposing subunits oxygenase domains (Figure 2.14a). BBOX represents the first example of dimerization of a 2OG oxygenase *via* an N-terminal domain. In other 2OG oxygenases, structurally characterised, dimerization has been observed to occur *via* the C-terminal domains. Although several other 2OG oxygenases have more than one domain, however, none of them resemble the N-terminal domain of BBOX. Based on sequence comparisons (26% identity), both the N-terminal zinc and C-terminal catalytic domains of BBOX appear to be similar to those of TMLH. The residues involved in the binding of zinc are Cys-38, Cys-40, Cys-43 and His-82 (Figure 2.14c). The N-terminal zinc binding domain of BBOX is identical to the N-terminal domain of a protein with no known functions from *Acidithiobacillus ferrooxidans* ATCC 23270 (PDB ID: 3LUU). A recent study suggests that the Zinc binding domain is crucial for the overall stability of BBOX fold.[67] The C-terminal domain of BBOX hosts the catalytic active site (Figure 2.14d) and consists of a double-stranded β-helix (DSBH) fold, which is a characteristic feature of all 2OG oxygenases.[42] The opening of the tunnel leading toward the active site cavity is bridged by the side chains of Val-192 and Ala-294, thus likely facilitating the binding of O_2 and release of CO_2 (Figure 2.15). Due to the small size of the tunnel opening, it is believed that conformational changes by loop movements are likely involved in the binding of substrates and release of the products. Similar to other 2OG oxygenases,[42] the BBOX DSBH core supports the conserved HXD...H motif, consisting of His-202, Asp-204, and His-347, which coordinates active site metal (Figure 2.14d). Several specific interactions can explain the binding mode of substrate and the stereochemical course of substrate hydroxylation. *N*-Oxalylglycine (NOG), a close analogue of 2OG, coordinates to the active site metal in a bidentate manner *via* a single oxygen of its 1-carboxylate group (*trans* to His-347) and its C-2-carbonyl oxygen (*trans* to Asp-204) (Figure 2.14d). The terminal carboxylate (C-5) of NOG is positioned to form interactions with a basic residue (Arg-360) and a hydroxyl group (Ser-229, not shown). Additionally, Arg-349 also interacts with the

terminal carboxylate of NOG. The binding of a water molecule completes a Fe^{II} octahedral binding coordination geometry.





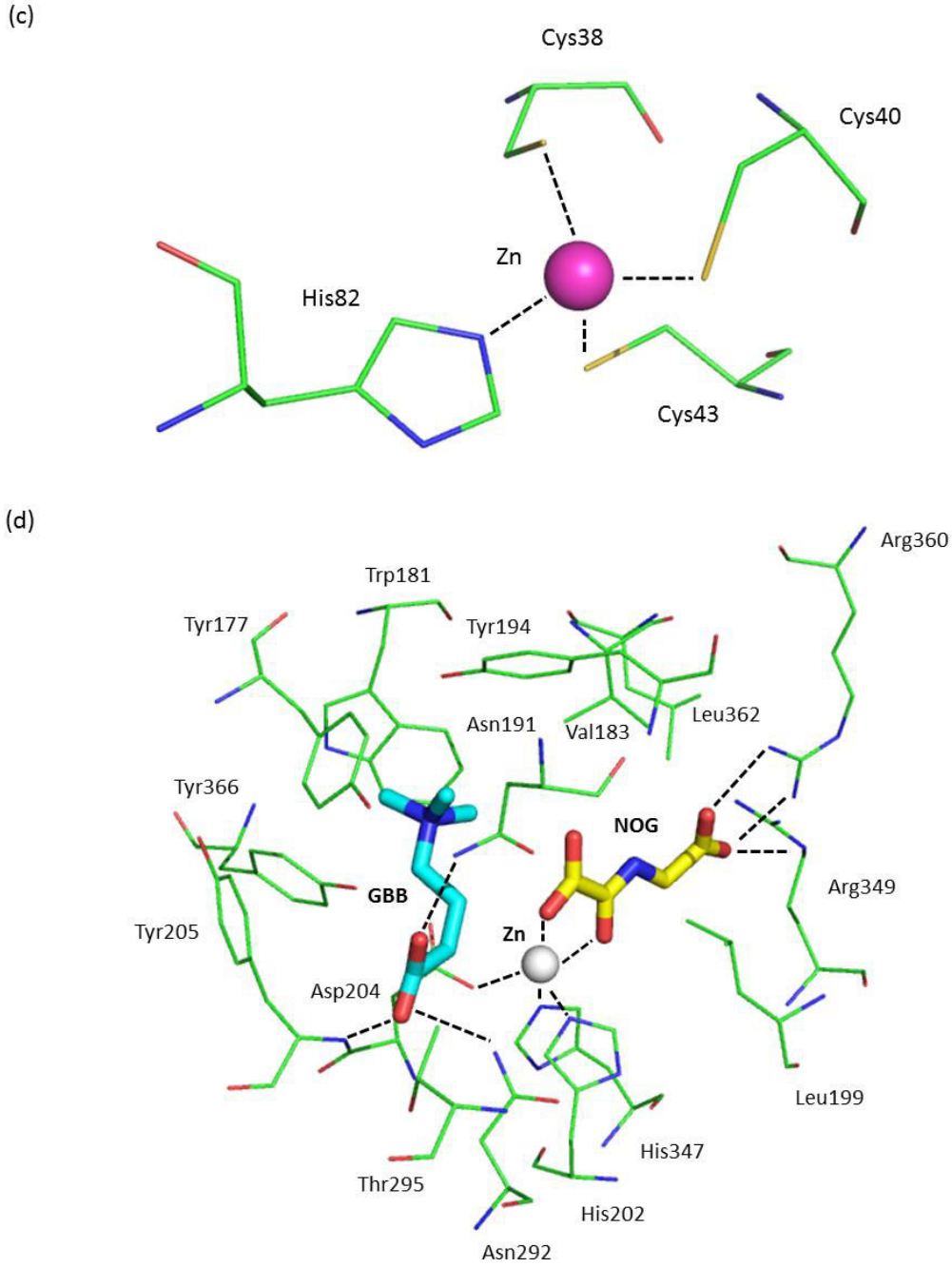


Figure 2.14 View from the crystal structures of human BBOX.GBB.NOG.Zn^{II}.Zn^{II} complex (PDB ID: 3O2G).
(a) BBOX dimer; the two monomers are shown as green and golden colours with the catalytic C-terminal domains containing active site metal Zn^{II} (grey spheres; substituting for Fe^{II}), ligands GBB (cyan sticks), NOG (yellow sticks) and N-terminal zinc binding domains containing the structurally important Zn^{II} (pink spheres). The catalytic C-terminal domain of green (monomer) interacts with the N-terminal zinc-binding domain of golden monomer. Similarly, the catalytic C-terminal domain of golden (monomer) interacts with the N-terminal zinc-binding domain of green monomer. (b) View of a monomer; showing the catalytic C-terminal domain (right) containing the active site metal (Zn^{II} substituting for Fe^{II}, grey sphere), ligands GBB (cyan stick), NOG (yellow stick), the 2OG oxygenases characteristics DSBH core fold and N-terminal zinc binding domain (left) containing the natural structural zinc (pink sphere). (c) Zoom in view of the structural zinc binding site in the zinc binding domain. The Zn^{II} (pink sphere) is chelated by Cys38, Cys40, Cys43 and His82. (d) Zoom in view of the C-terminal catalytic active site. The GBB (cyan colour) trimethylammonium group is located in an aromatic cage formed by the side chains of Tyr-205, Tyr- 366, Tyr-177, Trp-181 and Tyr-194, while its carboxylate group is positioned to interact with the side chains of Asn-292, Asn-191 and the backbone amide of Tyr- 205. The *N*-oxalyl glycine (NOG, yellow colour), an unreactive 2OG analogue, binds to the active site metal (grey sphere) in a bidentate manner *via* its C-1 carboxylate oxygen and C-2 keto oxygen. The terminal carboxylate of NOG forms interactions with Arg360 and Arg349.

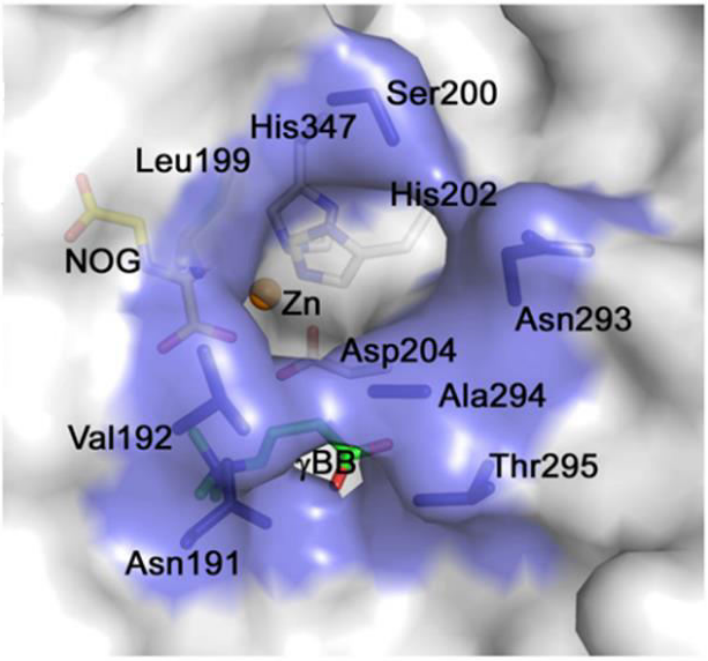


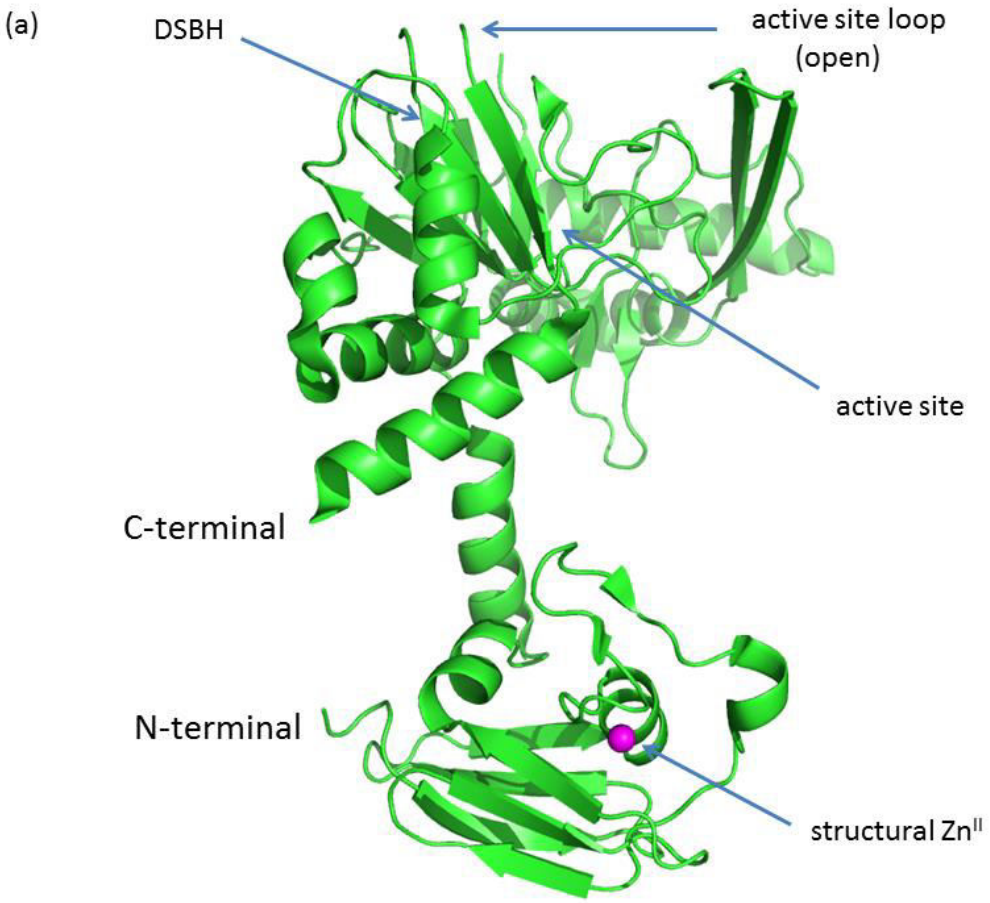
Figure 2.15 Surface view of the crystal structure of human BBOX.GBB.NOG.Zn^{II}.Zn^{II} complex (PDB ID: 3O2G).
Solvent accessible surface of BBOX (gray and blue) showing the opening of the active site tunnel bridged by residues Val-192 and Ala-294. The tunnel leads to the active site and is lined by Val-192, Ala-294, Ala-293, Asn-191, Thr-295, Leu-199 (shown as blue surface and sticks). NOG is shown as yellow stick, γBB or GBB as green stick and active site metal (Zn^{II}) as sphere. (Reproduced from reference [163])

Water forms hydrogen bonding interactions with O^{ε} of Gln-215 (2.7 Å) and the carboxylate of Asp-204 (2.8 Å), suggesting that both of these residues may be involved in the O_2 binding, the rest of the residues around this putative O_2 binding site are hydrophobic (Trp-181, Leu-217, Leu-362). The metal co-ordination site closest to C-3 of the substrate is occupied by the NOG C1-carboxylate. This pattern imply that O_2 may coordinate to the active site metal at the position occupied by water,[164] thus facilitating the catalytic ferryl intermediate to be located adjacent to the oxidised bond.[2,18] GBB binds adjacent to the metal such that its C-3 *pro*-R C-H bond is positioned close to the catalytic metal (C-3 GBB to metal: 4.7 Å).[122] This arrangement thus results in BBOX catalysed stereoselective C3-hydroxylation of the substrate.

The positively charged, *N*-trimethyl ammonium group of GBB is located in an electron rich hydrophobic cage formed by the side chains of Tyr177, Tyr-194, Trp-181, Tyr-205 and Tyr-366 of BBOX (Figure 2.14d). The cation-π interaction between the positively trimethylammonium group and the aromatic cage residues is believed to be critical in substrate recognition. Tyr-366 is also positioned to form a hydrogen bond with the side chain of Asp-204, suggesting that it may be involved in coordinating GBB binding and reaction at

the metal centre.[163] The carboxylate of GBB is positioned to form hydrogen bonding interactions with the side chains of Asn-292 (2.9 Å), Asn-191 (2.9 Å) and the backbone amide of Tyr-205 (2.9 Å). Based on the position of C-3 atom of GBB, its *pro-R* hydrogen projects towards the catalytic metal, whereas, its *pro-S* hydrogen points away. This mode of binding of GBB can explain the retention of configuration after H-abstraction.

In its *apo* form, the active site region of human BBOX is opened with the loop above the active site is disordered (Figure 2.16a), whereas, in the holo-form of human BBOX (Figure 2.16b), the loop closes the active site, suggesting that an induced fit mechanism enables the binding of active site metal, GBB/2OG for catalysis.



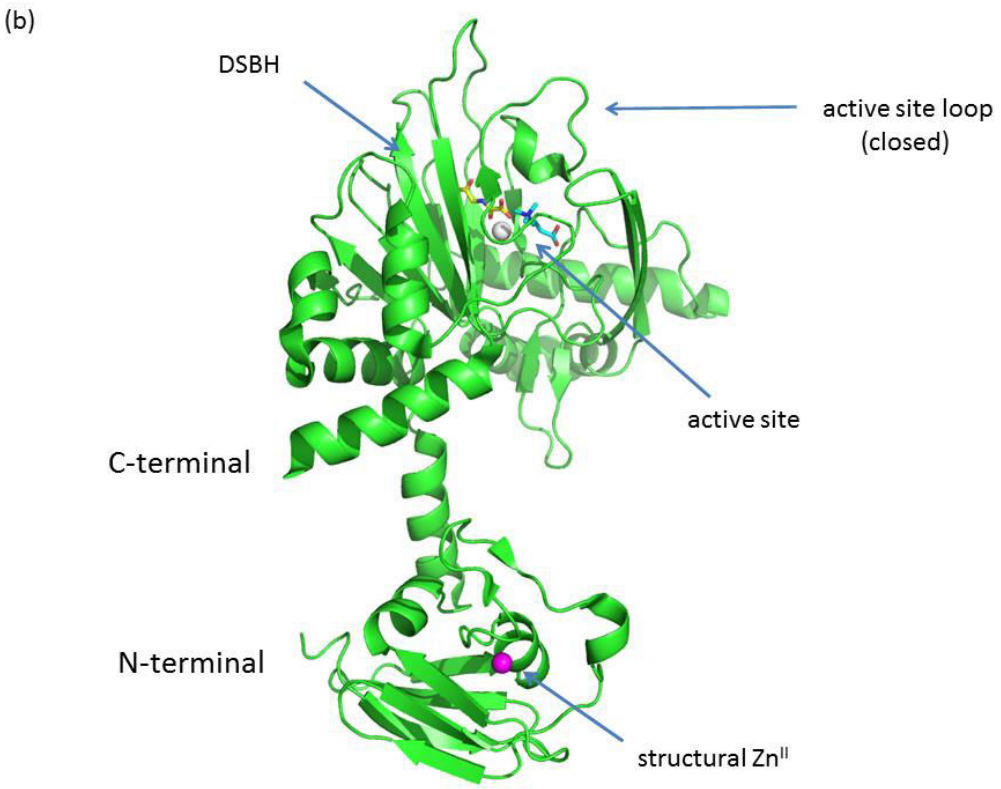


Figure 2.16 Demonstration of conformational changes in human BBOX during catalysis.
View from a crystal structure of *apo* (PDB ID: 3N6W) and *holo* (PDB ID: 3O2G) forms of human BBOX. The ligands (GBB; shown as cyan stick, NOG; shown as yellow stick and active site metal; grey sphere) binding at the C-terminal domain is likely to occur *via* an induced fit mechanism (loop above active site is closed in **b**). No such structural changes occur in the zinc binding domain (bottom part in **a**, and **b**).

2.2.5 Substrate specificity of BBOX

The substrate specificity studies of BBOX (Figure 2.17) suggest that the carboxylic group of GBB is essential for activity as substrate of BBOX. *N*,*N*,*N*-trimethylethanaminium (**I**), *N*,*N*,*N*-trimethylbutan-1-aminum (**II**) and *N*,*N*,*N*-trimethylhexan-1-aminum ions (**III**) do not show any activity towards recombinant human BBOX.[163] Similarly, replacement of the positively charged trimethyl ammonium group of GBB with an isopropyl group results in a complete loss of substrate activity towards recombinant human BBOX (**IV**).[163] However, a separate study using a different technique suggests that analogue ligand with a tertiary butyl group is a very poor BBOX substrate (**V**) forming the corresponding C-3 hydroxylated product (**V**$^{\prime}$) .[165] GBB analogue with the trimethyl ammonium group replaced with a dimethyl ammonium group (**VI**) is also accepted as BBOX substrate forming the corresponding C-3 hydroxylated product (**VI**$^{\prime}$), however with a poor activity compared to the positively charged trimethylammonium group (GBB).[163] Hence, the presence of positive charge of trimethylammonium group appears to improve substrate activity towards BBOX. Substitution

of a hydrogen atom of any of methyl group of trimethylammonium group of GBB with a fluorine (**VII**), also does not affects its activity towards recombinant human BBOX, also forming the corresponding C-3 hydroxylated product (**VII′**).[166]

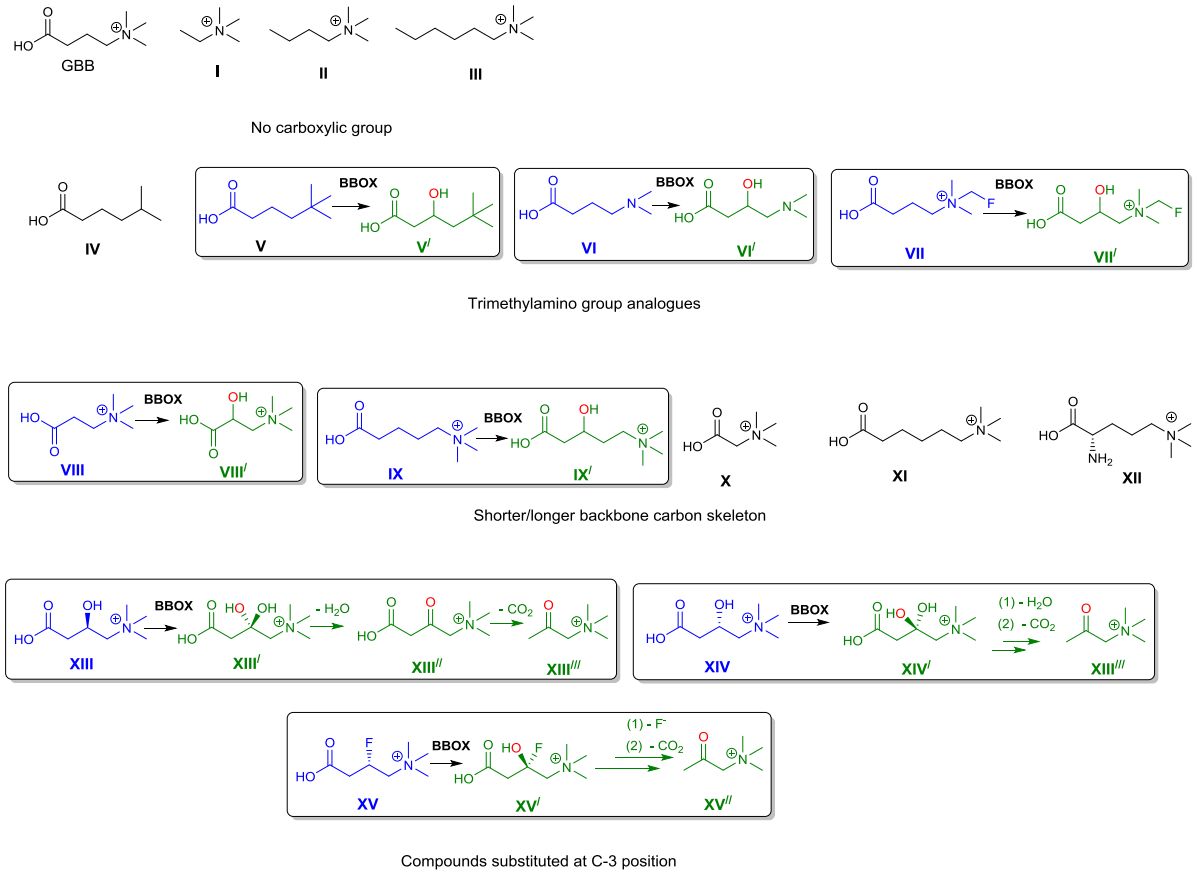


Figure 2.17 Human BBOX substrate specificity.
Structures coloured blue are accepted as BBOX substrates for oxidation to give the green coloured products. In case of XIII and XIV, the corresponding C-3 hydroxylated product loses a water molecule at C-3, forming di-keto GBB (XIII″), which upon subsequent decarboxylation forms trimethylaminoacetone (XIII‴). Both the dehydration and decarboxylation reaction occurs non-enzymatically. The same final product is also formed from XV.

Similarly, GBB analogues with one fewer carbon (3-(trimethylammonio)propanoate) (**VIII**) or one extra carbon (5-(trimethylammonio)pentanoate) (**IX**) in the backbone skeleton are also found to be hydroxylated by recombinant human BBOX at C-2 (**VIII′**) and C-3(**IX′**) respectively,[163] but with lesser activity compared to GBB, whereas analogues with two fewer carbons (trimethylammonio)acetate (**X**) or two extra carbons 6-(trimethylammonio)hexanoate (**XI**) do not show any activity towards recombinant human BBOX.[163] Similarly, N^{ε}-trimethyllysine (**XII**) also does not show any activity towards recombinant human BBOX.[163] Similarly GBB analogues substituted at C3 position e.g with a hydroxyl group (L, (**XIII**) and

D-carnitine (**XIV**))[163] or with a fluorine (*S*-stereochemistry) (**XV**)[167] are also accepted as substrates of recombinant human BBOX forming the corresponding C-3 hydroxylated products (**XIII**$^{/}$, **XIV**$^{/}$ and **XV**$^{/}$, respectively). Their activity is however lower as compared to GBB, possibly because there is little room for the C-3 substitution in the active site cavity. In case of L- and D- carnitine, their C-3 hydroxylated product (**XIII**$^{/}$, **XIV**$^{/}$, respectively) loses the C-3 water molecule forming 3-keto GBB (**XIII**$^{//}$), which subsequently undergoes decarboxylation in the assay mixture to give trimethylaminoacetone (**XIII**$^{///}$).[167] Both the hydration and decarboxylation reactions are believed to occur non-enzymatically in the assay mixture. Similarly, the C-3 hydroxylated product (**XV**$^{/}$) of 3-(*S*)-fluoro-GBB (**XV**) loses the fluorine atom also forming 3-keto GBB (**XIII**$^{//}$), which on decarboxylation forms trimethylaminoacetone (**XIII**$^{///}$).[167,168]

2.2.6 Co-substrates studies of BBOX

Like other 2OG and Fe^{II} dependent oxygenases, BBOX requirement for 2OG is also highly specific. To date, there is very little literature on the co-substrate scope for human BBOX. Work by Robert *et al.* on rat liver BBOX suggests that β-bromo-α-ketoglutarate is able to support L-carnitine formation,[38] however, with a 4-fold lower activity compared with 2OG. He also showed that human hypoxia inducible factor (HIF) prolyl hydroxylase domain containing enzyme (PHD) isoforms 1, 2 and 3, which is also a 2OG and Fe^{II} dependent oxygenase, also accepts β-bromo-α-ketoglutarate as an alternative co-substrate to 2OG for catalysis.

2.2.7 Stimulation of BBOX by potassium ions

Potassium ions markedly increase BBOX activity with a K_M of ~ 100 mM.[38] Similar to other potassium ion-activated enzymes,[169] BBOX activity is stimulated by ammonium and rubidium ions.[38] Similarly, alkali metals of significantly smaller or larger ionic radii are reported to either slightly inhibit (Li^+) or do not affect BBOX activity.[38] It appears that this activation is cation specific and not a general ionic strength effect. Several theories have been proposed to explain the mechanism by which monovalent cation stimulate enzyme activities.[170] Some studies suggest that these may have a role in inducing enzyme conformational changes,[171] while some studies suggests that these cation may be forming a bridge between the active site components and the reactants,[172,173] while some studies propose a combination of these two modes.[174]

2.2.8 Dependence of L-ascorbate for BBOX activity

Like other 2OG and Fe^{II} dependent oxygenases,[5,175,176] recombinant hBBOX activity is also stimulated by the addition of L-ascorbate in the assays. It is generally believed that L-ascorbate acts as a reducing agent in these assays by reducing Fe^{III} to Fe^{II}, thus maintaining a pool of Fe^{II}, and hence promotes maximum enzyme activity. To date, no direct binding of L-ascorbate either to BBOX or other 2OG oxygenases have been reported.

2.2.9 γ- Butyrobetaine from *Pseudomonas* specie AK1

In contrast to humans, a complete carnitine biosynthetic pathway has not so far been reported in bacteria. In bacteria, the physiological role of L-carnitine still remains unknown. The reported literature about carnitine in bacteria is limited to its degradations mechanisms.[76,79] Some studies suggest that it may function as an osmoprotectant[177] or growth stimulant,[178] while other suggest that several microorganisms utilize carnitine as a source of nitrogen and carbon.[79]

Several species of bacterial, such as those from *Pseudomonas* AK1 produce proteins homologous to human BBOX,[179] which has been shown to convert γ-butyrobetaine (GBB) into L-carnitine.[132,180,181]

Most of the early studies conducted on characterization and inhibition of BBOX were performed either on rat or psBBOX AK1 due to unavailability of an optimized protocol for producing recombinant hBBOX. psBBOX AK1 has been used as a substitute for hBBOX for assays development and inhibition studies, because it can be prepared on a large scale.[40] *Pseudomonas* BBOX is a homodimer with 383 residues per sub-unit.[180,181] BLAST analyses revealed that human and psBBOX AK1 share ~ 30% sequence identity (Figure 2.18).

To date, no crystal structure has been solved for psBBOX AK1. Based on recombinant human BBOX crystal structure (PDB ID: 3O2G), a generated homology model structure for psBBOX AK1 reveals that both the enzymes are almost similar in overall structure (Figure 2.19a). Like hBBOX, psBBOX AK1 also has a N-terminal zinc binding domain and a C-terminal catalytic domain (Figure 2.19a). An α-helix connects the N-terminal domain to the C-terminal domain. In addition to similarity in the overall folds of both the enzymes, all the crucial residues involved in the active site metal chelation and GBB/2OG binding in psBBOX AK1 are similar to those of hBBOX (Figure 2.19b). The active site metal is coordinated by two histidines i.e His350, His209, and an aspartate residue (Asp211), 2OG C-

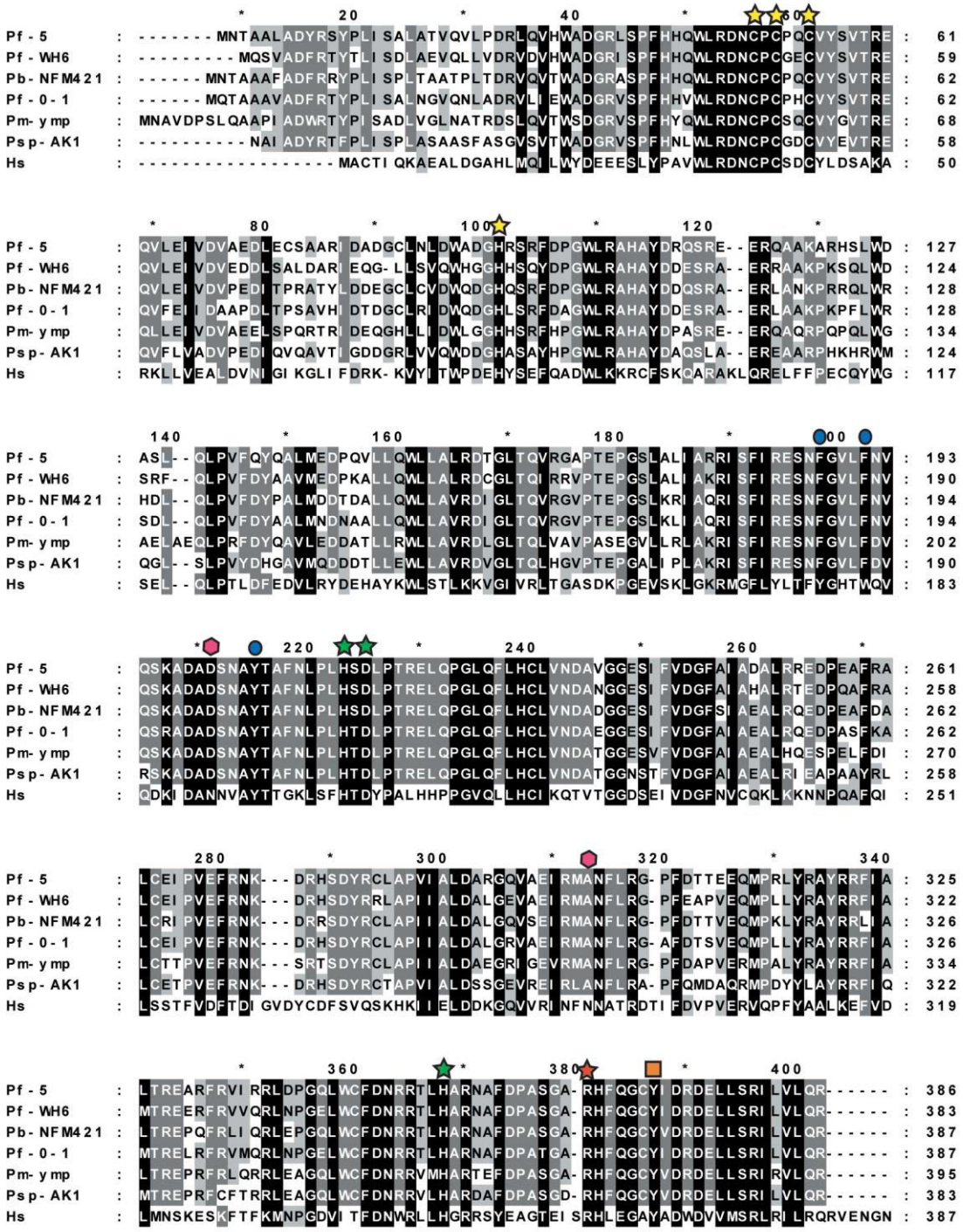
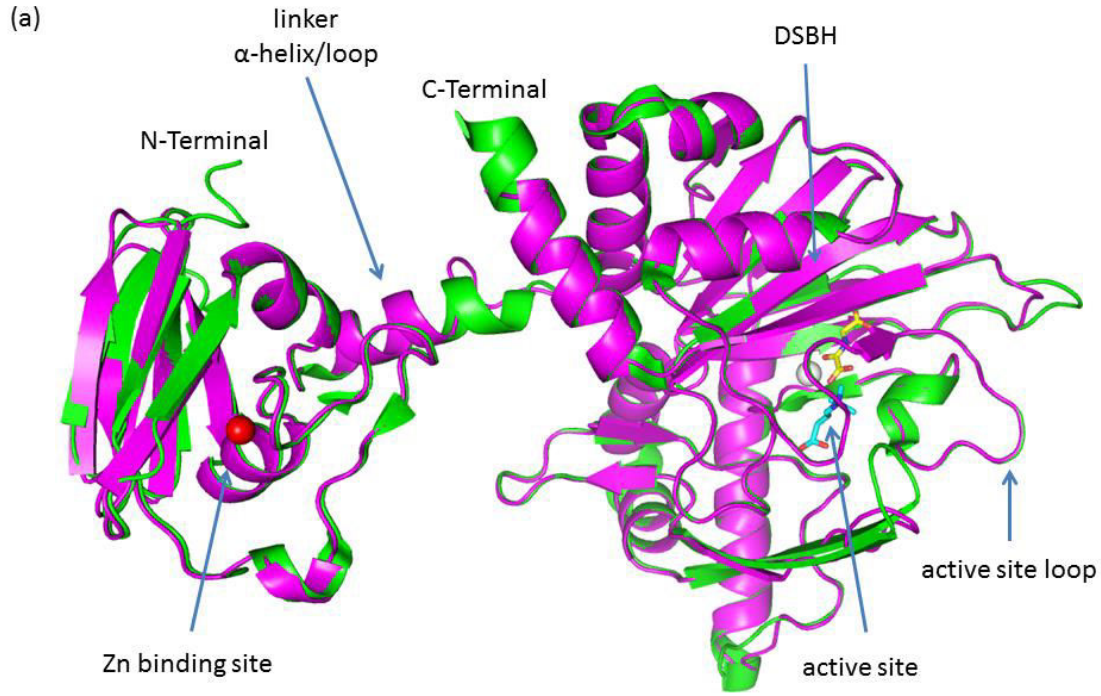


Figure 2.18 Sequence alignment of human and various species of Pseudomonas BBOX.
Sequence alignment of BBOX homologues from human (Hs, gi|158261239) and Pseudomonas (*Pseudomonas fluorescens* Pf-5, gi|70733041; *Pseudomonas fluorescens* WH6, gi|312963522; *Pseudomonas brassicacearum subsp. brassicacearum* NFM421, gi|330812266; *Pseudomonas fluorescens* Pfo-1, gi|77461455; *Pseudomonas mendocina* ymp, gi|146305540; *Pseudomonas* sp. AK1, gi|231642). The Zn^{II}-binding domain is present in both hBBOX and the different subspecies of psBBOX and the residues involved in Zn^{II} binding (yellow stars) are highly conserved in both cases. The HxD/E...H – Fe^{II} binding motif, which is a characteristic all the 2OG oxygenases (green stars) is conserved both in human and pseudomonas BBOX. Similarly, in all 2OG oxygenases, the terminal C-5 carboxylate of 2OG forms interaction with an arginine binding (red stars), these arginine are also conserved in human and psBBOX. Blue circles represent residues forming the aromatic cage, which forms hydrophobic and π-cation interactions with the trimethylammonium moiety of GBB. For both hBBOX and psBBOX these residues have aromatic side chains, i.e. phenylalanine and tyrosine (psBBOX) or tyrosine and tryptophan (hBBOX). Residues forming interactions with the GBB carboxylate group in hBBOX (orange circles) are not conserved in psBBOX. This may result in relatively weak binding of GBB in psBBOX compared with hBBOX.

1-carboxylate oxygen and C2-carbonyl oxygen atoms (Figure 2.19b). In hBBOX, these residues are His347, His202 and Asp204, respectively. The 2OG C-5-carboxylate interacts with an arginine (Arg362, hBBOX Arg360), which is a characteristic feature of all 2OG oxygenases. Additionally, Arg352 (hBBOX Arg349) also appears to interact with the 2OG C-5-carboxylate group. Some differences however, exist in the aromatic residues forming the GBB binding hydrophobic cage, and the residues interacting with carboxylate of GBB in both the enzymes (Figure 2.19b). In psBBOX AK1, Trp181 and Tyr177 of the aromatic cage present in hBBOX structure are replaced with phenylalanines (Phe188 and Phe184, respectively), whereas, Asn191 and Asn292 responsible for GBB carboxylate interactions in hBBOX are replaced by Ser198 and Ala296 in psBBOX AK1, respectively (Figure 2.19b). Similarly, in the zinc binding site, Zn^{II} is also co-ordinated by three cysteines (Cys46, Cys48, Cys51) and a histidine (His91), identical to hBBOX (Cys38, Cys40, Cys43 and His82, respectively) (Figure 2.19c).



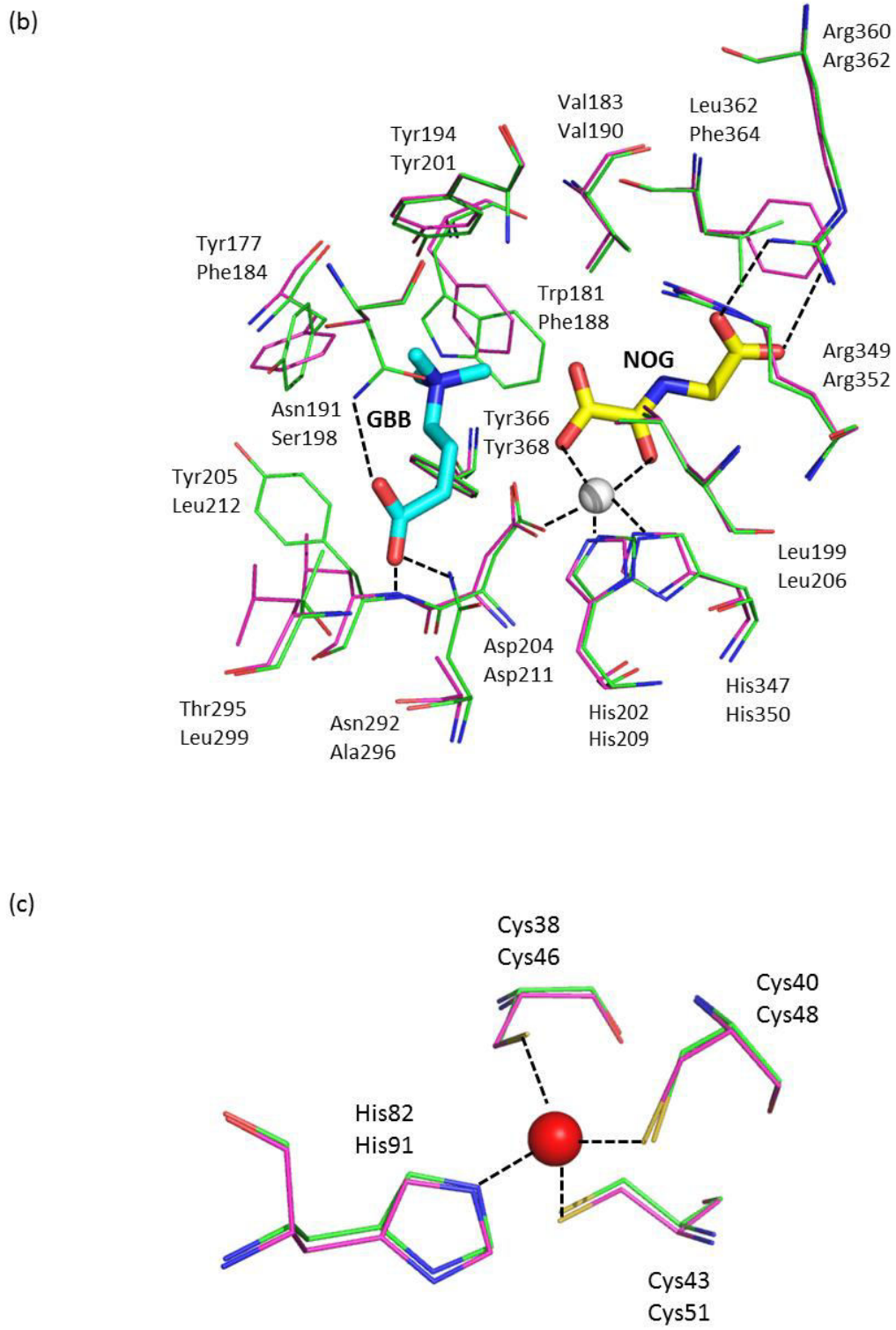


Figure 2.19 Comparison of the crystal structure of hBBOX with a generated model of psBBOX.
(a) Overlay of the crystal structure of recombinant human BBOX (green, PDB ID: 3O2G) and a generated homology mode structure of psBBOX AK1 (pink), showing that the overall folds of both enzymes are very similar. (b) Zoom in view of the active site in (a), showing comparison (hBBOX; green sticks, upper residues numbers, psBBOX; pink sticks, lower residue numbers) of the residues involved in the binding of active site metal (grey sphere) GBB (cyan) and NOG (yellow). (c) Zoom in view of the Zn^{II} binding site in (a), showing comparison of the residues involved in the binding of structural zinc (red sphere) (hBBOX; green sticks, upper residues numbers, psBBOX; pink sticks, lower residue numbers).

2.2.10 Comparison of substrate preference of human and *Pseudomonas* BBOX AK1

In general most of the compounds accepted as substrates by the hBBOX, are also accepted as substrates by psBBOX AK1, but, with varying degree of activity (Figure 2.20).[167] The variation in the nature of some of the residues involved in GBB binding in both the enzymes (hBBOX; Trp181, Tyr177, Asn191 and Asn292 vs psBBOX; Phe188, Phe184, Ser198 and Ala296, respectively) (Figure 2.19b) is reflected in the different preferences of these enzymes towards certain substrates. Amongst the similarities for example, GBB, C-3 substituted GBB analogues including 3-(*S*)-fluoro-GBB, D-carnitine, GBB analogue with one extra carbon in the backbone chain (GBB-5) and GBBNF (Figure 2.20A) are accepted as substrates by hBBOX; are also the substrates for psBBOX, forming the corresponding C-3 hydroxylated products. Similarly, GBB analogues with two carbons shorter (GBB-2) or two carbons longer (GBB-6) in the backbone chain, N^{ε}-trimethyllysine (GBB-6 with C-2 substituent), and GBB analogues with C-2 substituent including N^{ε}-trimethyl ornithine (TMO) derivatives such as, (*S*)-TMO and (*R*)-TMO, GBB analogues without free carboxylate such as GBB methyl ester (GBB-OMe), keto-GBB derivative (GBB-CH_3), acetylcholine and its structural analogues such as thioacetylcholine, carbachol and phosphocholine (Figure 2.20B), are all neither the substrates for hBBOX nor the psBBOX.[167]

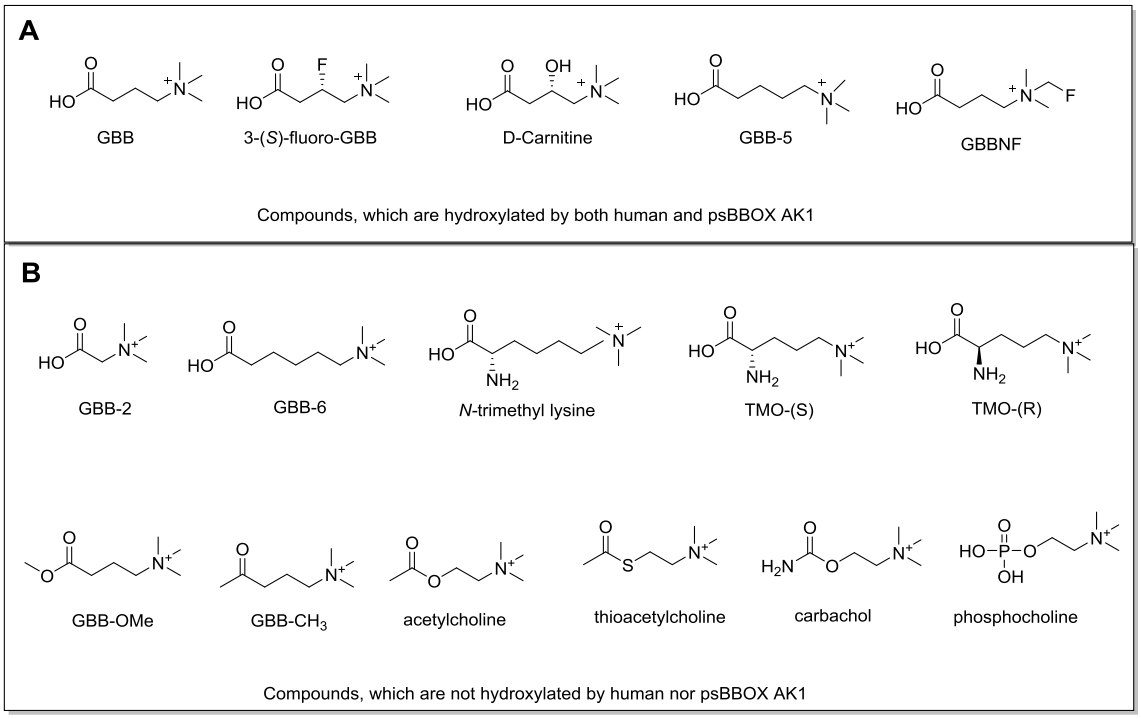


Figure 2.20 Similarity of catalytic activity of human and psBBOX AK1.

Amongst the co-substrate (2OG) analogues, none of succinic semialdehyde, pyruvate, 2-ketobutyrate, 2-ketovalerate, p-hydroxyphenylpyruvate, oxalacetate, 2-ketoadipate, 2-ketopimelate, glutarate, 2-hydroxyglutarate, glutamate, citrate, isocitrate, cis-aconitate, succinate, fumarate, or malate have been able to replace 2OG as a co-substrate in catalysis by both human and psBBOX AK1.[132]

Amongst the differences, psBBOX AK1 does not hydroxylate GBB-3 and L-carnitine (Figure 2.21A), both are poor substrates for recombinant human BBOX. In contrast, both the (*R*) and (*S*) enantiomers of 2-amino GBB and (*S*)-2-hydroxy GBB (Figure 2.21B) are all accepted by psBBOX for C3-hydroxylation, but not by recombinant hBBOX, presumably there is lesser room for GBB-C2 substituent in the active site cavity of hBBOX.

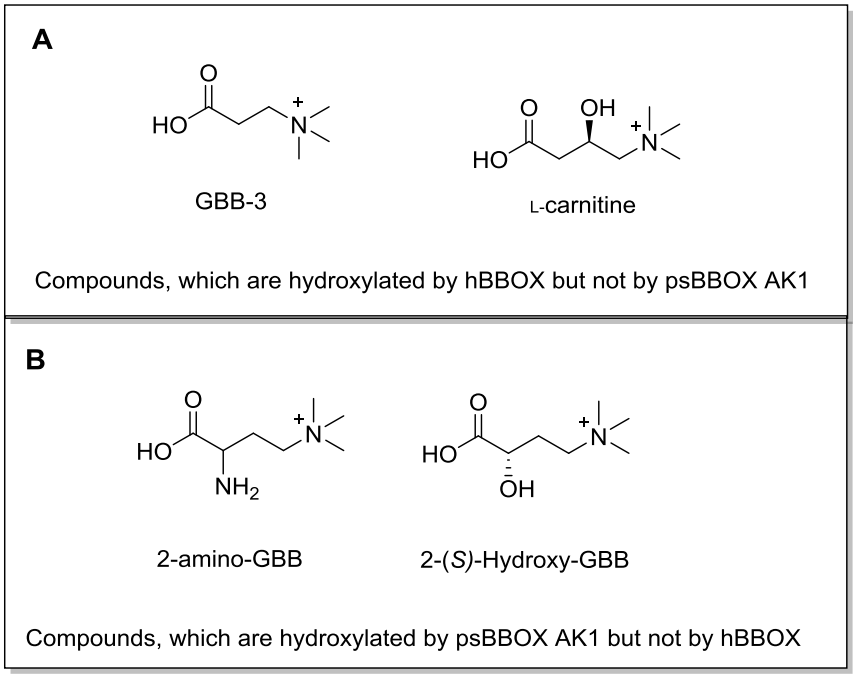
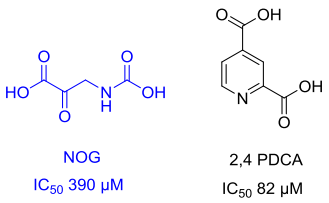


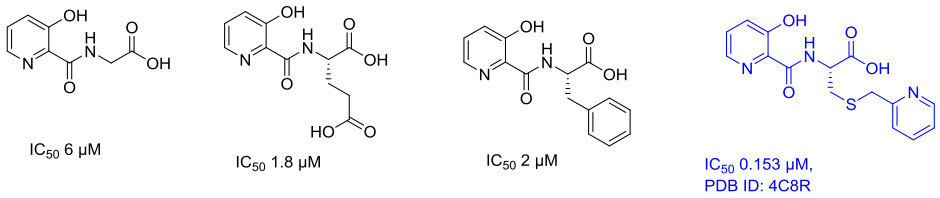
Figure 2.21 Difference of catalytic activity between human and psBBOX AK1.

2.2.11 Inhibitors of BBOX

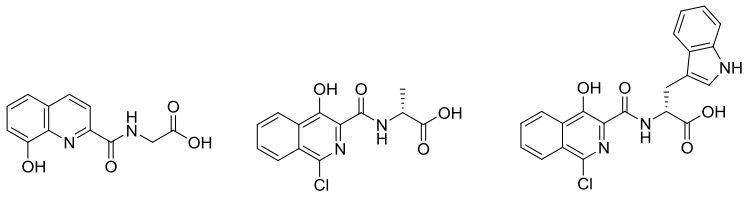
Most of the reported inhibitors of recombinant hBBOX fall into two types. The first type predominantly acts as 2OG competitors (Figure 2.22).[159,168] These inhibitors include, NOG, 2,4-PDCA, 3-hydroxypyridine, 8-hydroxyquinoline (8-HQ) and 3-hydroxyisoqionoline derivatives (Figure 2.22). The second category of inhibitors are GBB analogues (Figure 2.23).[157,159,165,168]



Generic 2OG-based recombinant human BBOX inhibitors



Hydroxypyridine-core containing recombinant human BBOX inhibitors



Hydroxypyridine-quinoline/isoquinoline containing recombinant human BBOX inhibitors

Figure 2.22 Different types of 2OG-based inhibitors (with IC_{50} values) of recombinant human BBOX.
Reported in references [159,168]. Blue coloured inhibitors are those for which the crystal structure of hBBOX has been reported.

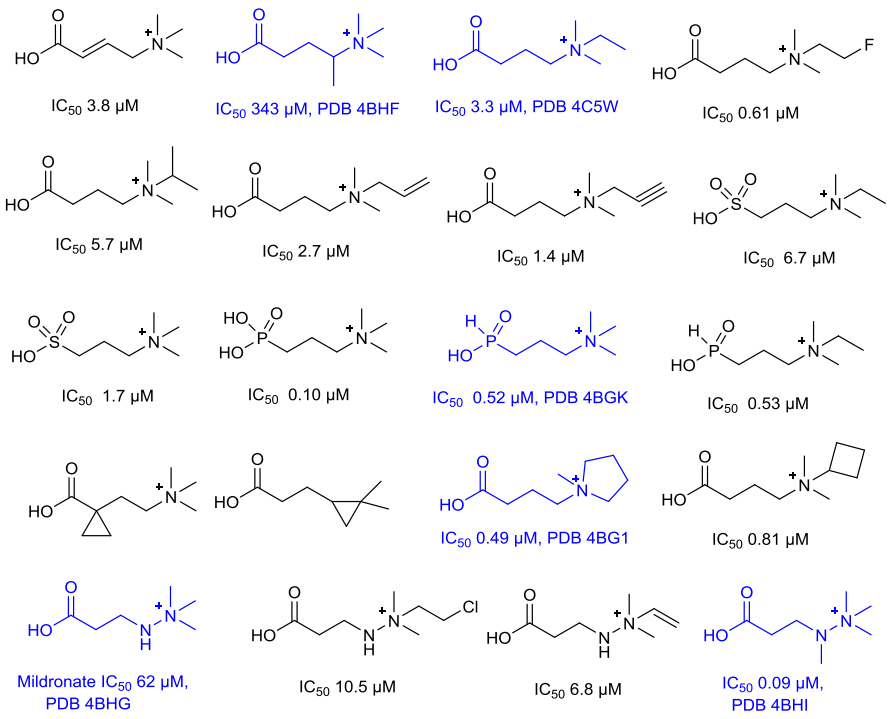


Figure 2.23 GBB mimic inhibitors of recombinant hBBOX.
Reported in reference [157]. Blue coloured inhibitors are those for which the crystal structure of BBOX has been reported.

Several metal ions are also reported as potent BBOX inhibitors, including Co^{II}, Ni^{II}, Cu^{II}, Zn^{II}, Cd^{II} and Hg^{II}.[132] These metal ion inhibitors likely compete with the active site Fe^{II} during catalysis.

Several selenium and sulfur based compounds are also reported as inhibitors of recombinant hBBOX (Figure 2.24).[67] Their mode of inhibition of BBOX activity is however different from 2OG, GBB or Fe^{II} mimicking inhibitors as described above. These selenium and sulphur-based compounds are effective zinc ejectors.[64-66] These compounds are proposed to irreversibly eject the structurally critical Zn^{II} from the N-terminal domain of the BBOX, and thus disrupt the stabilised and active (folded) conformation of the enzyme. [67]

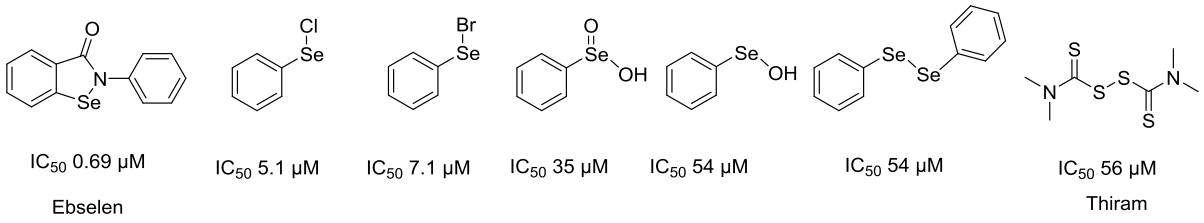


Figure 2.24 Selenium and sulphur based structural inhibitors of recombinant human BBOX.

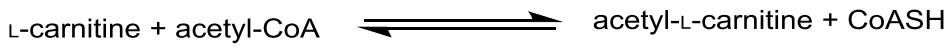
2.2.12 Assays for BBOX

A wide variety of enzyme activity-based techniques have been employed for BBOX substrate screening (whether a compound is hydroxylated by BBOX) and inhibitor screening (whether the presence of another compound can inhibit BBOX catalysed substrate hydroxylation) studies.

Most of the early (1970s) studies on BBOX were carried out by using either [*methyl*- $^{14}C_3$]-γ-butyrobetaine (GBB), monitoring the formation of [*methyl*- $^{14}C_3$]-L-carnitine [132] or [1-$^{14}C_1$]-2OG, monitoring the formation of $^{14}CO_2$.[181] In the [*methyl*- $^{14}C_3$]-γ-butyrobetaine assay, the BBOX activity is assessed by incubating the BBOX, 2OG, and cofactors (Fe^{II}, L-ascorbate) with labelled-GBB for a defined period of time. The labelled-carnitine formed is separated from the assay mixture by ion-exchange chromatography and then its concentration is determined by measurements of its radioactivity. Similarly, in the 2OG-degradation activity assay, BBOX, GBB and cofactors are incubated with [1-$^{14}C_1$]-2OG and the resulting labelled-carbon dioxide formed in the reaction is trapped on a piece of filter paper attached to a wire in the rubber stopper of the test tube. After diffusion of $^{14}CO_2$ for a defined length of time, the

filter papers are transferred to a scintillation mixture and then its radioactivity is measured in a liquid scintillation spectrometer.

An enzyme-coupled assay has also been used to detect L-carnitine concentrations as low as 20-2000 picomolar.[182-184] This method is based on the conversion of [1-^{14}C]- acetyl-CoA to [1-^{14}C]-acetylcarnitine in the presence of carnitine acetyltransferase (EC 2.3.1.7). The formation of radioactive acetylcarnitine was shown to be proportional to the carnitine concentration over a wide range.



The acetylcarnitine formed is separated from the labelled acetyl-coenzyme A by passing the reaction mixture through a anion exchange chromatographic column and its isotope content is then determined in a liquid scintillation counter.[184] The Coenzyme A formed in this reaction is determined by reaction (trapping) with 5,5′-dithiobis-2-nitrobenzoic acid (DTNB) or by enzymic methods (for details see reference [185]) The rapid separation of [1-^{14}C]-acetylcarnitine from [1-^{14}C]-acetyl-CoA has also been carried out by the selective adsorption of the latter on charcoal in the presence of an acid and ethanol.[183]

Mass spectrometry (Figure 2.25)[155,157,163] and NMR spectroscopy (Figure 2.26)[155,156,168,186] turnover assays have also been employed for BBOX substrate and inhibitor screening.

A fluorescence-based technique has also been employed for BBOX inhibitor screening.[168] This technique is based on the BBOX catalysed C-3 hydroxylation of 3(*S*)-fluoro-GBB (called as GBBF) and the subsequent release of fluoride ion from the unstable hydroxylated intermediate (Figure 2.27). The fluoride ion released is detected by *tert*-butyldimethylsilyl-protected fluorescein. When *tert*-butyldimethylsilyl-protected fluorescein is subjected to increase fluoride ion concentration, a direct proportional increase in fluorescence is observed following the removal of the silyl protective groups.[187] By employing this assay, standard enzyme kinetics time course can be readily obtained from changes in the fluorescence intensity.

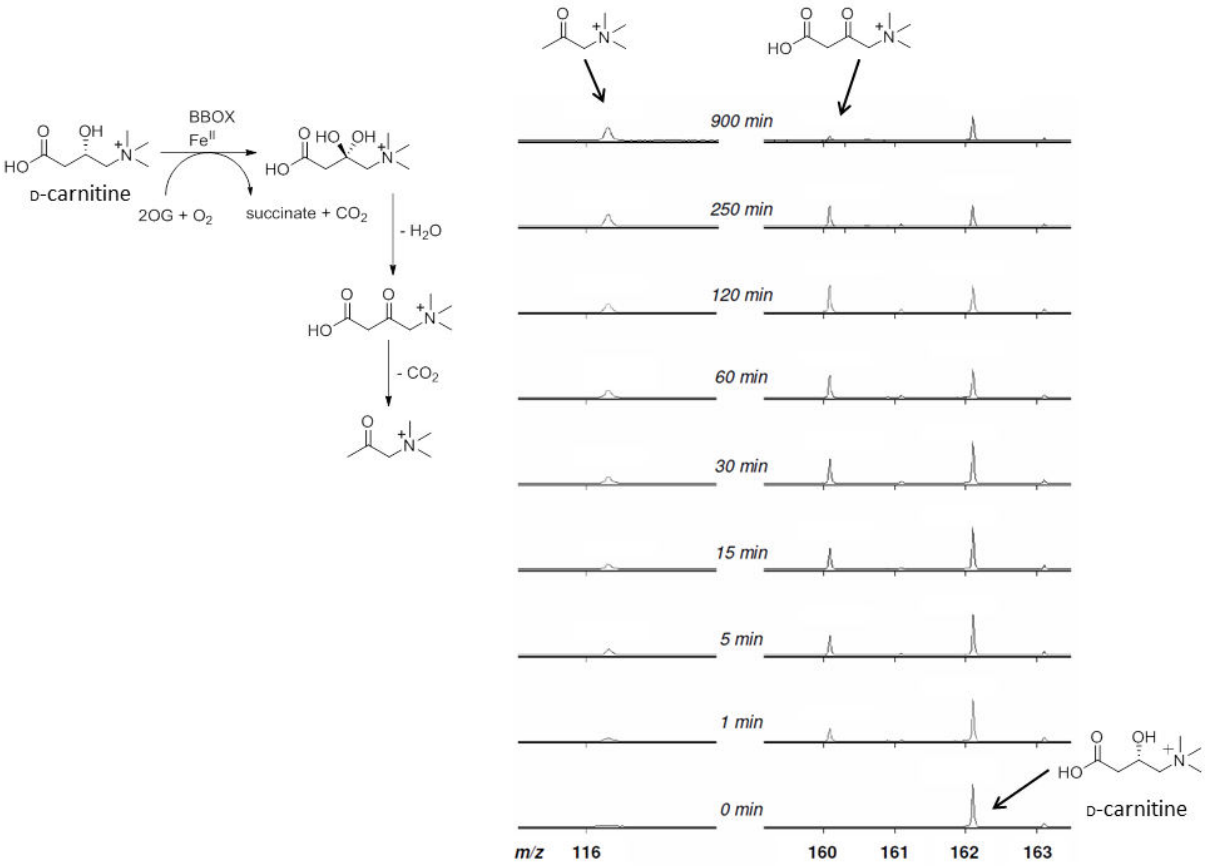


Figure 2.25 MS (mass spectrometry) analysis of hBBOX catalysed hydroxylation of D-carnitine. (Reproduced from reference [188])

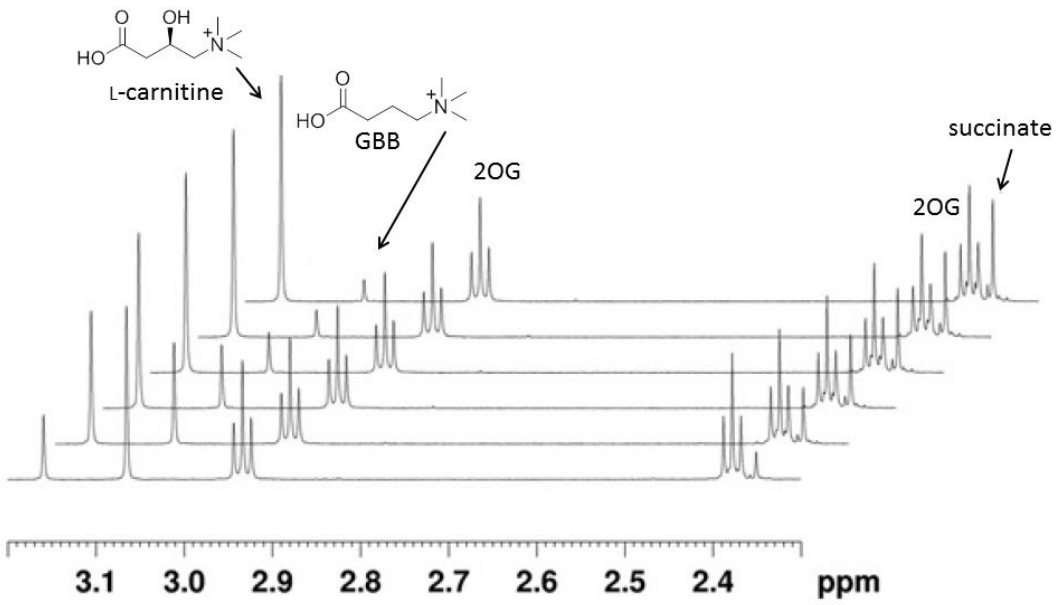


Figure 2.26 hBBOX catalysed GBB hydroxylation into L-carnitine as demonstrated by ^{1}H NMR. (Reproduced from reference [163]).

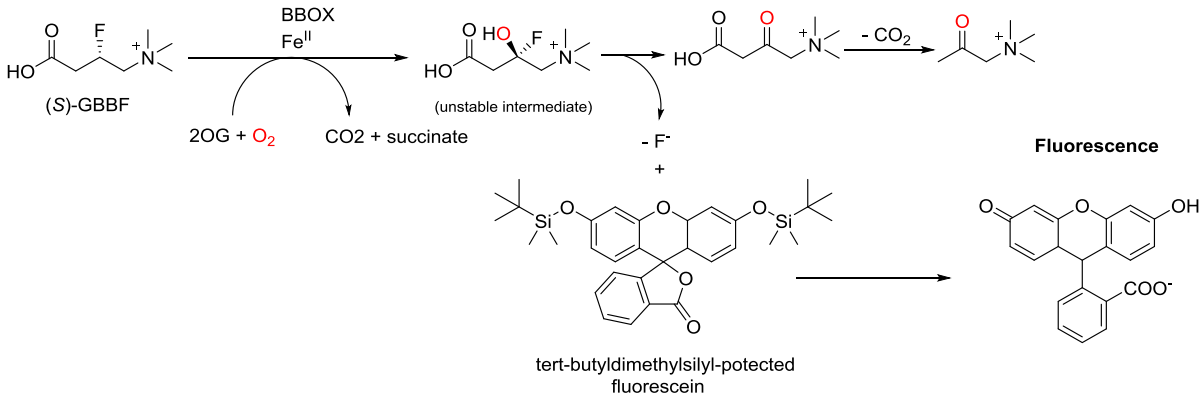


Figure 2.27 Fluoride ion release assay for BBOX inhibitors screening.
The method is based on BBOX catalysed hydroxylation of 3-(*S*)-fluoro-GBB with subsequent fluoride ion release. The fluoride ion released is detected by using a tertbutyldimethylsilyl (TBS) protected fluorescein chemical probe.

In addition to the above enzyme activity-based assays, only a handful of binding assays have also been reported for BBOX. These include an isothermal calorimetric titration-based method[159] and a BBOX intrinsic tryptophan fluorescence quenching-based assay.[67] These methods have been used to measure the binding dissociation constant (K_D) of small molecules ligands with BBOX.[67,157] In the latter assay, the quenching of intrinsic tryptophan fluorescence of BBOX after the addition of a ligand to the BBOX assay mixture, assays for binding to BBOX (Figure 2.28). The ligand binding dissociation constant can be measured by titrating a ligand to a fixed BBOX concentration and measuring the relative change in the BBOX intrinsic tryptophan fluorescence i.e $\Delta F_{obs}/F_{max}$ (ΔF_{obs} is the decrease in observed fluorescence and F_{max} is the observed fluorescence signal without any inhibitor present) as a function of ligand concentration.

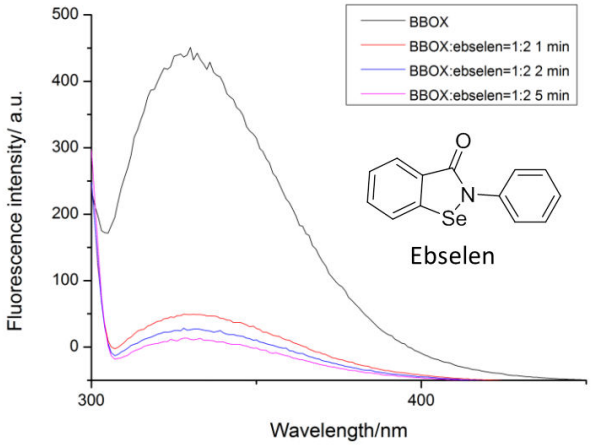


Figure 2.28 BBOX intrinsic tryptophan quenching binding assay.
Treatment of BBOX with ebselen causes quenching of BBOX intrinsic tryptophan fluorescence in a time dependent manner, demonstrating the binding interaction of ebselen with BBOX. BBOX concentration used was 5 μM. (Reproduced from reference [67])

Non-denaturing MS analyses have also been used to investigate small molecules binding interaction (covalent) with BBOX.[67] This assay was employed to study the binding of sulphur and selenium-based compound (zinc ejectors) to BBOX, which forms covalent linkage with the cysteine thiol groups involved in the binding of structural zinc in BBOX (N-terminal zinc binding domain).[67]

2.3 Aims of the project

The aims of the work described in this chapter were primarily to gain a better understanding of BBOX catalysis and its mode of inhibition. I aimed to develop solution-based NMR binding assays for BBOX, that could be readily implemented, sensitive, site specific, not require any specific probe, be operationally simple and which can be used to rank the hits with a relatively fast approach compared to measuring their binding dissociation constant value by titration experiments. Understanding the full picture of an inhibitor binding in the active site of BBOX (mode of BBOX inhibition) is crucial for the development of its selective inhibitors over other 2OG oxygenases of biological interests. In principle, a compound can inhibit BBOX activity either by competing with 2OG or GBB or both, however, the current reported inhibitor screening assays for BBOX do not readily and reliably give this information. As observed in the case of PHD2, the crystallographic studies do not always reflect the physiological conditions i.e in contrast to crystal structure, a different mode of inhibitor binding was observed by solution NMR studies.[189] It may be possible to determine whether a compound inhibits BBOX activity competitively with respect to either 2OG or GBB by enzyme kinetics studies. However, such studies are labour intensive. Moreover, the kinetics studies for determining the mode of enzyme inhibitions are known to be prone to errors arising from the design of the experimental conditions. For instance, in early work on investigating the mode of inhibition of BBOX by mildronate, based on kinetics studies, one group reported it as a non-competitive inhibitor of BBOX.[62] Subsequently, however, another group used a different approach, and demonstrated that mildronate in fact is a competitive inhibitor of BBOX.[154] Due to complications in the assay conditions, enzyme kinetics assays are not commonly used in determining the mode of an enzyme inhibition by an inhibitor. Hence, there is need to develop efficient and reliable solution-based assays which can demonstrate the nature of BBOX inhibition. Hence, I aimed to develop efficient NMR assays that can demonstrate whether a compound inhibits BBOX catalytic activity by competing with its natural substrate GBB, co-substrate 2OG or both.

I also aimed to investigate how important is the positive charge on the nitrogen of the trimethylammonium moiety of GBB in BBOX substrate recognition. Though the hBBOX crystal structure (PDB ID: 3O2G) shows that the GBB trimethylammonium group is located in an aromatic cage, it is still unknown whether under solution physiological conditions it is the hydrophobic interactions between the trimethyl groups of GBB and aromatic cage residues, or whether the cation-π interaction between the positively charged nitrogen of trimethylammonium group of GBB and aromatic cage residues that hold GBB in the aromatic cage.

All the work described in this chapter was conducted on psBBOX AK1(produced by Dr Jurgen Brem), because it was more simple to be produced on a large scale using an optimised reported protocol[159] compared with recombinant human BBOX, and the results are widely applicable for method development works.

2.4 ^{1}H NMR direct ligand observation binding assay for BBOX

^{1}H NMR direct ligand-observation monitoring the broadening and/or attenuation of ligand NMR resonances after the addition of a target protein is the most sensitive and readily implemented binding technique (described in section 1.4.1) as compared to other ligand-observed binding methods. We begin by applying this technique to first investigate the binding interaction of natural substrate GBB to psBBOX (1:5 ligand:protein ratio) in the absence of both co-substrate 2OG and metal co-factor. The high sensitivity of the Me_3 signal (singlet at 3.05 ppm) of GBB allowed the use of its lowest possible concentration in the assay (5 µM) in order to reduce the consumption of protein in the assay. No change in the GBB signals intensity or line-width occurred with the addition of psBBOX (Figure 2.29b) suggesting that in its *apo*-form, psBBOX does not appear to form stabilized and active conformation required for substrate binding (catalysis). It is possible that the *apo*-form of psBBOX may also has a disordered active site loop region as observed for *apo*-hBBOX (PDB ID: 3N6W) (as discussed in Figure 2.16, Section 2.2.4). To the assay mixture, 2OG (300 µM i.e ~ 2 times its K_M for hBBOX,[188] believed to be sufficient to saturate all the psBBOX molecules) was then added, but still no changes in the GBB NMR resonances were observed (Figure 2.29c) implying that BBOX-2OG complex may also not appear to be stable and active to allow substrate binding. This prompted us to study the binding of GBB to psBBOX in the presence of both metal co-factor and 2OG in the assay solution. In the assay, we employed Zn^{II} instead of native catalytic Fe^{II} to prevent catalysis. The concentration of Zn^{II}

used was 150 µM, which is 6 times the K_M for Fe^{II} for hBBOX.[188] In the presence of both 2OG and Zn^{II}, it was observed that the GBB signals broaden/attenuate after the addition of psBBOX (Figure 2.29d) demonstrating its binding interaction with the fully *holo*-form of psBBOX. These results support the proposed sequential ordered catalytic cycle of metal and 2OG stabilized active BBOX conformation required prior to substrate binding for catalysis.[1,6]

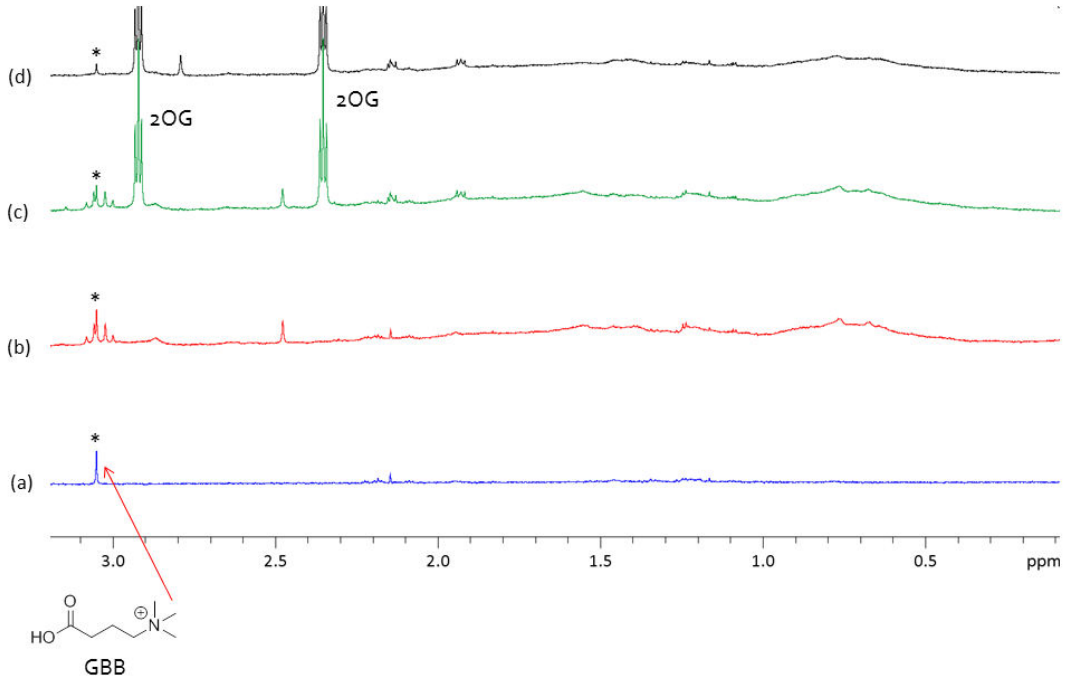
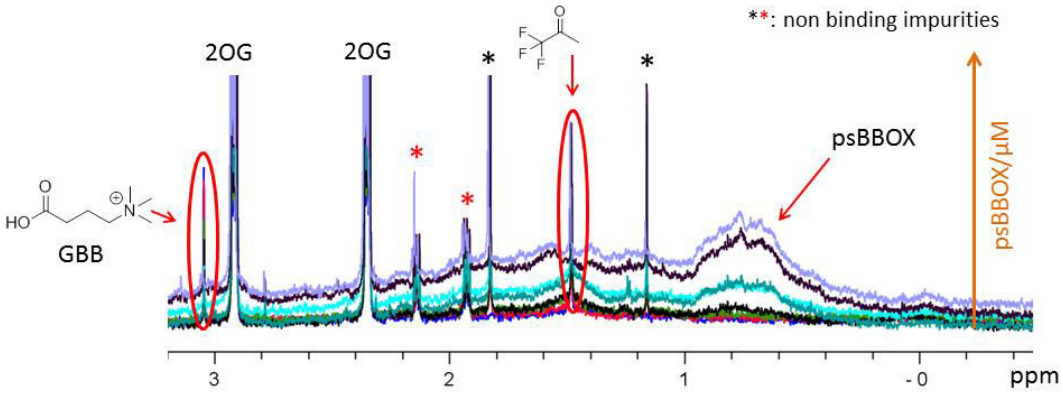


Figure 2.29 Investigating the binding of GBB with psBBOX by ^{1}H NMR direct ligand observation.
The GBB Me_3 peak at 3.05 ppm marked with asterisk is followed. (a) Solution of GBB in buffer. (b) After the addition of psBBOX to (a). The broad lump represents the signals of psBBOX. Signals of small molecule impurities from protein solution are also seen. (c) After the addition of 2OG to (b). (d) After the addition of Zn^{II} to (c). The GBB signals were attenuated only in the presence of both 2OG and metal co-factor Zn^{II}. The concentrations used were 5 µM GBB, 25 µM BBOX, 300 µM 2OG, 80 mM KCl, 150 µM Zn^{II}. The assay solution was prepared in 50 mM Tris-D11, pH 7.5 in H_2O.

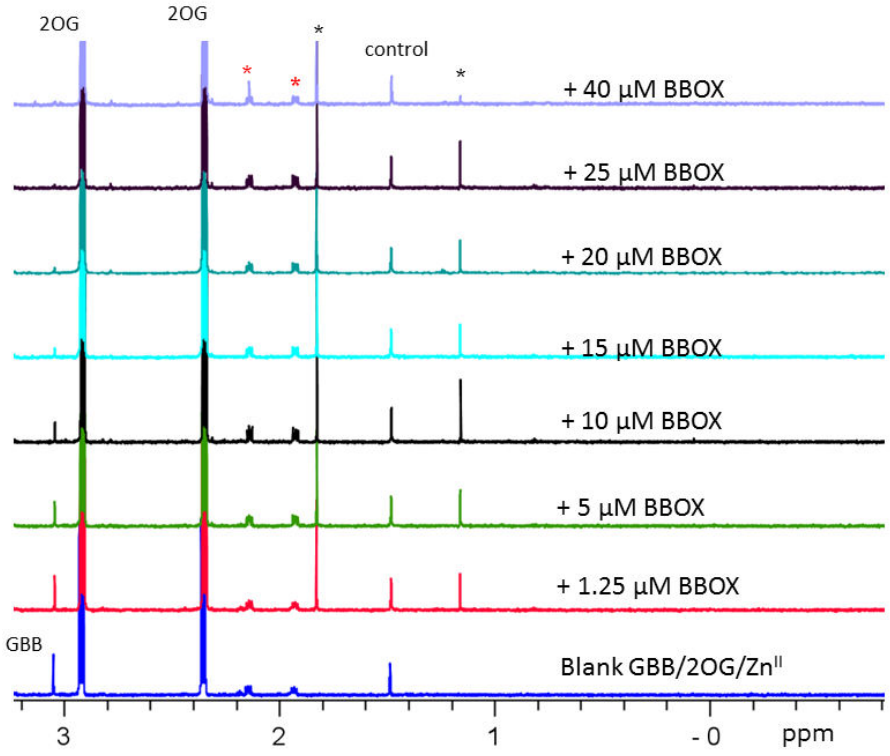
In order to measure the binding affinity of GBB to psBBOX, psBBOX was titrated to a solution containing a fixed concentration of GBB, 2OG and Zn^{II} (Figure 2.30). Both standard ^{1}H (Figure 2.30a) and CPMG edited ^{1}H NMR (Figure 2.30b) spectra were recorded. The removal of the protein resonances by the CPMG filter clearly showed that the GBB Me_3 signal broaden and attenuates as the BBOX concentration increases (Figure 2.30c), demonstrating its fast exchange binding interaction with BBOX. The loss in the GBB Me_3 signal was measured as a function of BBOX concentration (Figure 2.30c). Fitting of the titration data points using Origin Pro. according to the quadratic equation (details in experimental section 5) calculated a binding constant (K_D) value of 5 ± 1 µM for GBB (Figure 2.30d).

(a)



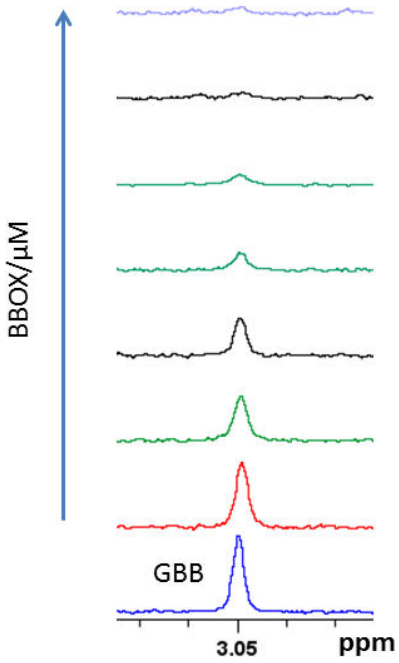
^{1}H NMR

(b)



CPMG edited ^{1}H NMR

(c)



psBBOX / μM	I_P	$\Delta I\,(\%) = (\frac{I_P - I_B}{I_B}) * 100$
0	1	0
1.25	0.8945	-10.55
5	0.6738	-32.62
10	0.4747	-52.53
15	0.2764	-72.36
20	0.2138	-78.62
25	0.1141	-88.59
40	0.0998	-90.02

(d)

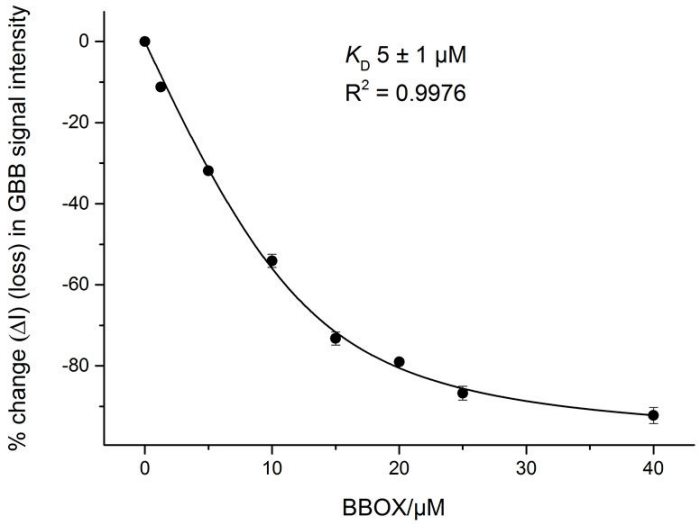


Figure 2.30 Measurement of the binding constant (K_D) of GBB to psBBOX-Zn^{II}-2OG complex by CPMG edited ^{1}HNMR-based direct ligand observation.

Titration of psBBOX to a solution containing a fixed GBB, 2OG and Zn^{II} concentration. (a) Overlay of the ^{1}H NMR spectra of the assay solution with psBBOX concentration increases from bottom to the top (broad resonances). The red and black stars are signals of the small molecules impurities from buffer and protein solutions respectively. (b) The corresponding CPMG edited ^{1}H NMR spectra. Note that the CPMG filter has selectively removed the protein background resonances. The GBB Me_3 at 3.05 ppm is conveniently seen as attenuating with BBOX titration, demonstrating its binding interaction. The signal of the control molecule (1,1,1 trifluoroacetone) at 1.48 ppm remains unaffected with BBOX titration. A filter time of 32 ms was used in these experiments. (c) Left: A zoomed in view of the GBB Me_3 signal in (b), showing GBB fast exchange binding with BBOX. Left: Quantification of loss of GBB Me_3 signal intensity as a function of BBOX concentration. I_B and I_P represent GBB signal intensity in the absence and presence of protein (BBOX) respectively. (d) Fitting curve of titration data points from (c) into Origin Pro. A K_D of 5 ± 1 μM was obtained. The error bars represent standard deviation from three separate measurements. The assay mixture contains 10 μM GBB, 250 μM Zn^{II}, 700 μM 2OG, 80 mM KCl, 50 μM 1,1,1 trifluoroacetone (internal NMR standard) in 50 mM Tris-D_{11}, pH 7.5, in 90% H_2O:10% D_2O.

2.4.1 NMR paramagnetic relaxation enhancement

After establishing that ^{1}H NMR direct ligand observation is a suitable technique that can be applied to study small molecules binding interaction with psBBOX, we begin to optimise the assay conditions to simultaneously detect the binding of both GBB and 2OG in a single spectrum, potentially for the development of an NMR competition assay for ligand screening (discussed in section 1.4.2), whereby both 2OG and GBB can simultaneously be employed as reporter molecules. However, we observed that using Zn^{II} as co-factor, high psBBOX concentrations are required to enhance the sensitivity of GBB/2OG binding, which is ideally required for NMR reporter displacement assay. We envisaged that greater sensitivity to reporters binding might be achieved by the application of paramagnetic relaxation enhancement (PRE).[190-194] Nuclear spins relaxation enhancement by paramagnetic spin-labels has long (1960s) been used as a valuable tool to characterise metalloenzyme-ligand binding interactions.[195] This technique is based on the fact that the magnetic moment of an unpaired electron is 658 times greater than that of a proton nuclear spin,[196] hence generating strong local magnetic field and therefore acts as efficient source for nuclear spin relaxation. Thus, when a ligand binds with a protein in the vicinity of a paramagnetic centre (within 15-20 Å), its nuclear transverse spin relaxation rates are significantly enhanced or in other words the ligand NMR signals are significantly broadened. This technique was originally developed by covalently tagging a paramagnetic spin-label[191,192] such as TEMPO (2,2,6,6-tetramethyl-1-piperidine-*N*-oxyl) to proteins (Figure 2.31). This technique is termed as SPLASTIC (spin labels attached to a protein side chain to identify interacting compounds). In addition to the spin labels, paramagnetic metals such as Co^{II} [197] and Mn^{II} [198,199] Ni^{II}, low spin Fe^{III} and most lanthanides (with the exception of Gd^{IV}) have also been used to produce PRE effect.[191]

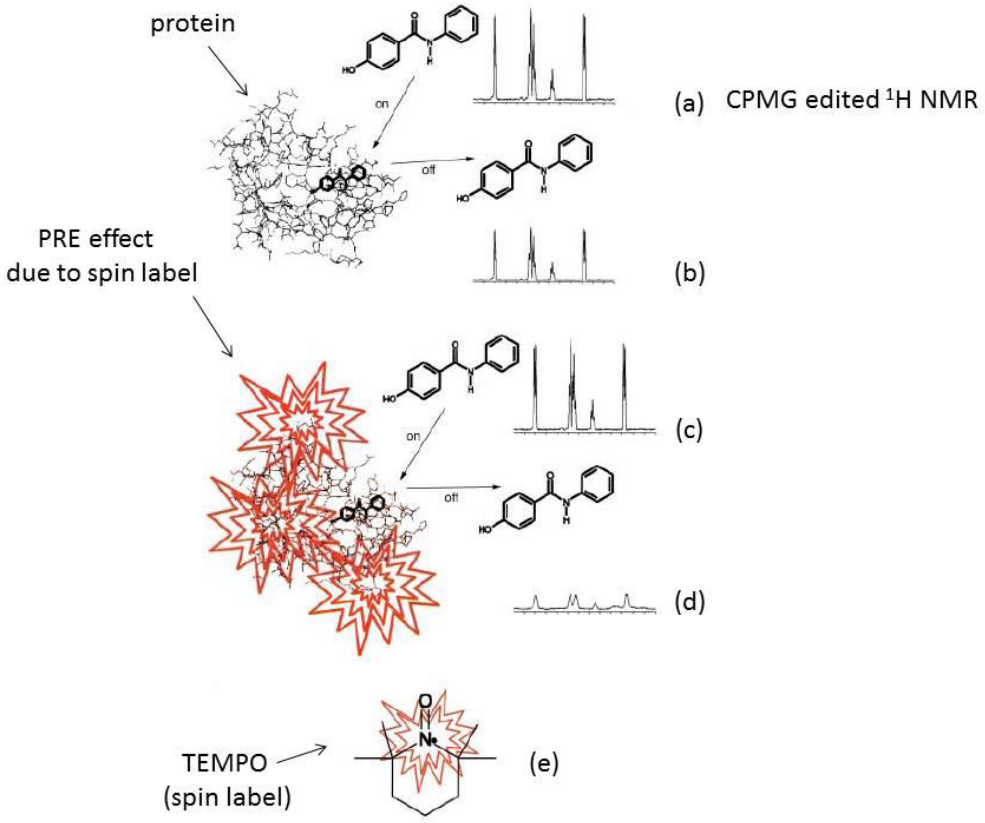


Figure 2.31 Illustration of application of NMR PRE effect in ^{1}H NMR-based direct ligand observation binding technique. (a) CPMG edited ^{1}H NMR spectrum of 4-hydroxy-*N*-phenylbenzamide. (b) After the addition of protein to (a). The ligand signals are attenuated (though only slightly), demonstrating its binding interaction with the protein. (c) CPMG edited ^{1}H NMR spectrum of 4-hydroxy-*N*-phenylbenzamide. (d) After the addition of protein tagged with spin label TEMPO (2,2,6,6-tetramethyl-1-piperidine-Noxyl) to (c). Note that the due to the PRE effect of the spin label, the transverse relaxation rates of the bound ligand are significantly enhanced and as a result the ligand NMR resonances are significantly broadened/attenuated. (e) Spin-label TEMPO. (Reproduced from reference [191])

The magnitude of the paramagnetic relaxation enhancement (PRE) caused by the dipolar interactions depends on the square of the gyromagnetic ratios of both the spins involved, the inverse sixth power of the interspins distance, and the correlation time. The transverse relaxation rate enhancement, R_{2para}, is described by below the equation[200],

$$R_{2\text{para}} = \frac{1}{15} S(S+1) \frac{\gamma_I^2 g^2 \beta^2}{r^6} \left(4\tau_c + \frac{3\tau_c}{1+\omega_I^2+\tau_c^2} \right)$$

Where, S denotes the electron spin, γ_I the proton gyromagnetic ratio, g the electronic g factor, β the Bohr magneton, r the distance between the electron spin and the nuclear spin, ω_I the resonance frequency of protons, and τ_c the correlation time of the vector linking the electron and the nuclear spins. A similar equation holds for the longitudinal relaxation rate enhancement, R_{1para}.

Hence, taking the advantage that the active site of psBBOX does have a metal binding site to which 2OG directly binds in a bidentate manner and which is also in close approximation (~ 3 Å) to the location of the GBB binding site (Figure 2.32), we envisaged that the simplest approach to produce the PRE effect in the psBBOX active site would be to replace the diamagnetic Zn^{II} in the assay mixture with paramagnetic Mn^{II}. Hence, we first investigated the binding of Mn^{II} to psBBOX by a reported NMR competition-based assay based on measuring the water longitudinal relaxation rates ($R_1 = 1/T_1$) in the presence of paramagnetic Mn^{II}.[198,199] In the absence of any metal chelating ligand and provided that access is possible, water molecules can chelate a paramagnetic metal in a solution and as a result their T_1 is decreased their due to the PRE effect. These Mn^{II}-bound water molecules undergo exchange with water molecules in the bulk solvent matrix. Hence, if Mn^{II} binds in the active site of a protein, then the titration of the protein to the assay solution would result in reduction of the free Mn^{II} in solution, which in turn would reduce the water molecules bound to Mn^{II}, and this would ultimately result in increase of the bulk water T_1. We thus titrated psBBOX to a solution containing fixed Mn^{II} concentration and it was observed that water T_1 increased as a function of psBBOX concentration demonstrating the binding interaction of the Mn^{II} to psBBOX (Figure 2.33). A binding constant value of 3 ± 1 µM was obtained for the binding of Mn^{II} to psBBOX, likely forming a complex in a 1:1 ratio (from saturation point).

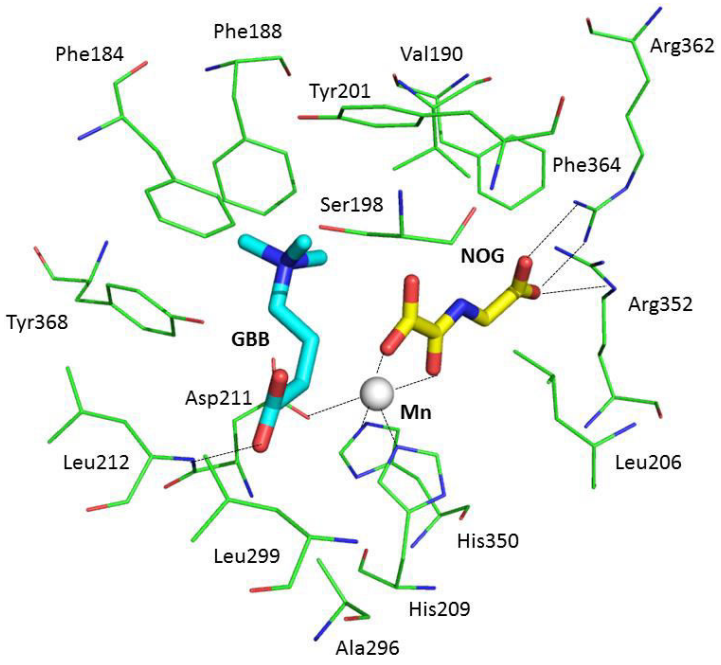


Figure 2.32 View from the active site of psBBOX in a model structure based on hBBOX crystal structure (PDB ID: 3O2G). The PRE effect of Mn^{II} (grey sphere) drastically increase the transverse relaxation rates (line-shape) of the GBB (cyan) (see below) and 2OG (yellow) NMR resonances when they bind to psBBOX.

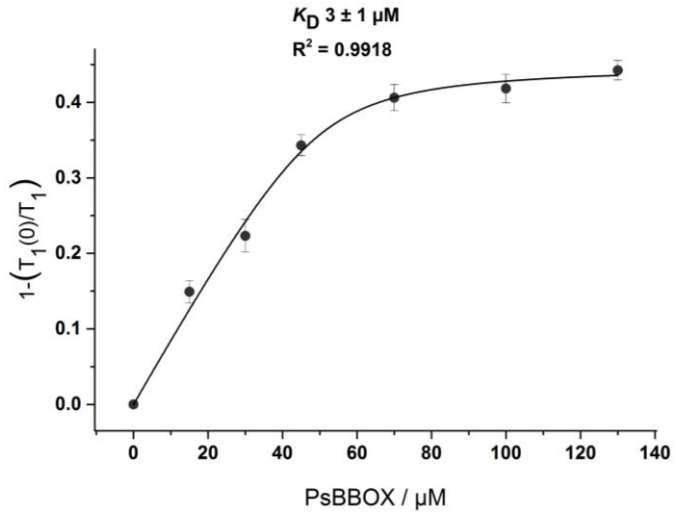
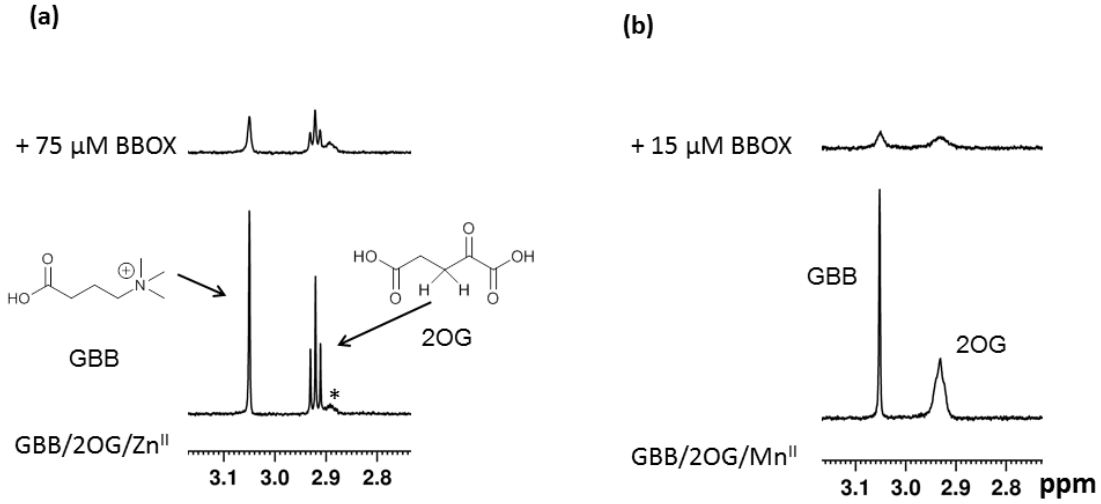


Figure 2.33 Measurement of the binding constant of Mn^{II} to psBBOX by water T_1 measurement assay.
By titrating psBBOX into 50 µM Mn^{II} in a solution containing 1.2 mM 2OG and 100 mM KCl in 50 mM Tris-D_{11} in 20% H_2O/80% D_2O, we demonstrated the formation of a psBBOX-Mn^{II} complex. A binding constant of ~ 3 ± 1 µM was obtained for Mn^{II} binding to psBBOX. The error bars represent standard deviation from three separate measurements.

After establishing the binding of Mn^{II} to psBBOX, we then just simply replaced the diamagnetic Zn^{II} in the assay solution with paramagnetic Mn^{II}. Interestingly, we observed that when GBB and 2OG (reporter ligands) binds to psBBOX with Mn^{II} in the active site, then their NMR resonances are drastically broadened/attenuated in the presence of significantly low amount of psBBOX as compared to diamagnetic Zn^{II} (Figure 2.34). This is due to their significantly higher transverse relaxation rates (R_2) caused by the PRE effect of bound Mn^{II} in the active site of psBBOX.



*: impurity

Figure 2.34 Application of PRE effect in GBB/2OG binding with psBBOX by CPMG edited ^{1}H NMR-based direct ligand observation.
GBB **Me_3** singlet at 3.05 ppm and 2OG C**H_2** triplet at 2.91 ppm are shown. (a) GBB and 2OG binding to BBOX-Zn^{II} (b) GBB and 2OG binding to BBOX-Mn^{II}. The assay mixture contains 25 GBB, 300 µM 2OG, 150 µM M^{II}, 80 mM KCl, in 50 mM Tris-D_{11} buffer, pH 7.5, in D_2O. The spectra were run on a 700 MHz instrument.

Control experiments (without psBBOX) revealed that at the concentrations employed, free Mn^{II} in solution did not affect the GBB line-shape in the spectrum (Figure 2.35 & Figure 2.36), likely because it doesn't chelate the Mn^{II}, it did, however, broaden the resonances of 2OG, since this directly ligates Mn^{II} (Figure 2.35). However, despite this broadening, the resonance intensities (^{1}H NMR peak areas) remained similar to those prior to the addition of the metal (Figure 2.35 & Figure 2.37). Hence, based on these results, for a reporter (GBB, 2OG) displacement assay it would be more appropriate to follow the reporters peak attenuation as a probe of their binding with psBBOX intead of following their line-widths.

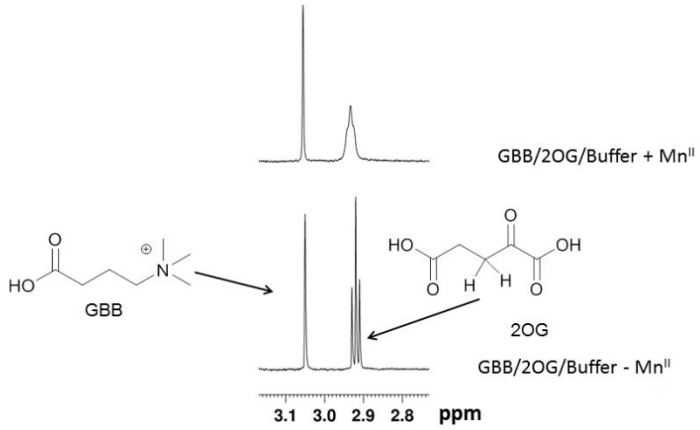


Figure 2.35 Monitoring the effect of Mn^{II} in solution on the line shape of free GBB and 2OG by ^{1}H NMR.
Following the addition of Mn^{II}, the GBB signals (singlet at 3.05 ppm) are unaffected, whereas the 2OG signals (triplet at 2.91 ppm) are broadened; however the absolute intensity of the latter remains the same with and without Mn^{II}. The final assay mixture contained 25 µM GBB, 300 µM 2OG, 150 µM Mn^{II}, 80 mM KCl, in 50 mM Tris-D_{11} buffer, pH 7.5, in D2O.

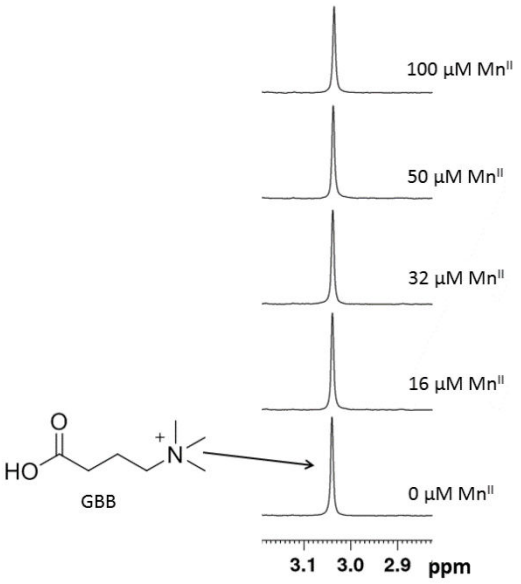


Figure 2.36 Mn^{II} in solution does not affect the line-shape of GBB (^{1}H NMR).
The concentration of GBB was 100 µM. The solutions were made in 50 mM Tris-D_{11}, pH 7.5, in 90% H_2O and 10% D_2O.

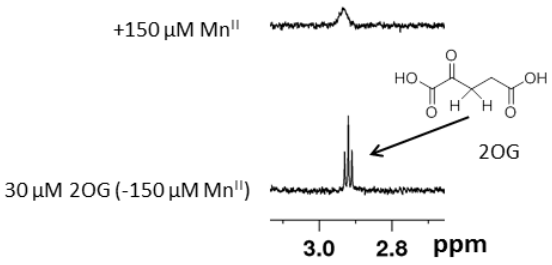


Figure 2.37 The effect of Mn^{II} on 2OG in solution.
2OG C**H**$_2$ signal at 2.91 ppm shown. Under the experimental conditions used Mn^{II} only broadens the 2OG resonances due to direct chelation, but the resonances absolute intensities remain the same with and without Mn^{II}.

In another control experiment, the protein in the sample of GBB/2OG/Mn^{II}/psBBOX was thermally denatured (heated at 100 °C for 5 minute), and it was observed that signals of the 2OG and GBB recovers (Figure 2.38), demonstrating that the attenuation of their NMR resonances after the addition of psBBOX was due to their binding interaction with the psBBOX.

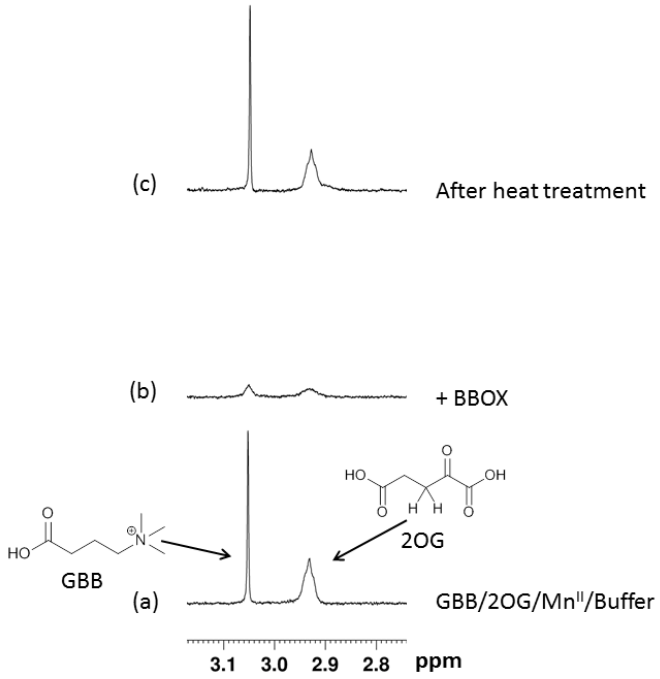


Figure 2.38 psBBOX denaturation control experiment for the binding of GBB and 2OG by CPMG edited ^{1}H NMR direct ligand observation.
(a) Solution of GBB/2OG with Mn^{II} in buffer. (b) After addition of BBOX to (a); the GBB and 2OG signals are attenuated due to binding with BBOX. (c) After heat treatment of (b); denaturation of BBOX results in the recovery of signals of both GBB and 2OG, demonstrating that the signals attenuation of both these molecules in the presence of psBBOX is due to binding with the BBOX. The final assay mixture contained 25 µM GBB, 300 µM 2OG, 150 µM Mn^{II}, 80 mM KCl, 15 µM psBBOX in 50 mM Tris-D_{11} buffer, pH 7.5, in D_2O.

These results thus clearly demonstrate that PRE is a useful technique that increases the NMR sensitivity to ligand binding in the presence of significantly low amount of protein.

2.5 ^{1}H NMR dual-reporter displacement assay for psBBOX inhibitor binding studies

In principle, an inhibitor can inhibit psBBOX catalysis by competing with co-substrate 2OG, or substrate GBB or both. Keeping in view that investigating the mode of psBBOX inhibition by enzyme kinetics studies (V_{max} and K_M determination with and without an inhibitor) is a labour intensive process, we envisaged that in an alternative approach this can be achieved more readily and reliably by investigating the mode of inhibitor binding to psBBOX i.e whether an inhibitor binds psBBOX by competing with 2OG or GBB or both. This information can be more readily obtained *via* a NMR dual-reporter displacement assay, whereby both GBB and 2OG can be employed as reporter molecules. If an inhibitor binds to psBBOX competitively with 2OG, it will displace only 2OG from its binding site, if it competes with only GBB, then only GBB will be displaced from its binding pocket, and if it competes with both 2OG and GBB, then simultaneously both 2OG and GBB will be displaced. The displacement of these reporter ligands can be conveniently observed as recovery of their NMR resonances in the spectrum. In this technique paramagnetic Mn^{II} can be used to enhance the NMR sensitivity of binding of the two reporter ligands.

As a proof of principle, we first analysed the binding of known 2OG competitive ligands, 2,4-PDCA, which is generic 2OG oxygenase inhibitor[6] including hBBOX[168] and AMG-1957661, which is a PHD2 inhibitor,[54] to psBBOX using this technique. It was observed that after the addition of 2,4-PDCA to the sample of 2OG/GBB/Mn^{II}/psBBOX, only the signals of the 2OG recover, suggesting that 2OG is no longer bound to the psBBOX, but is now free in solution i.e it is displaced from its binding site by 2,4-PDCA (Figure 2.39c). This demonstrates that 2,4-PDCA binds to psBBOX by competing only with 2OG and not GBB. In case of AMG-1957661, we noticed that the signals of both 2OG and GBB recover, implying that AMG1957661 binds psBBOX by competing with both GBB and 2OG (Figure 2.39d). The data also imply that after displacement, the splitting (multiplicity) of the 2OG resonances also become apparent, suggesting that the 2OG molecules are no longer chelating the free Mn^{II} in solution, i.e the inhibitors also chelate free Mn^{II} in solution in addition to the chelating the active site Mn^{II}.

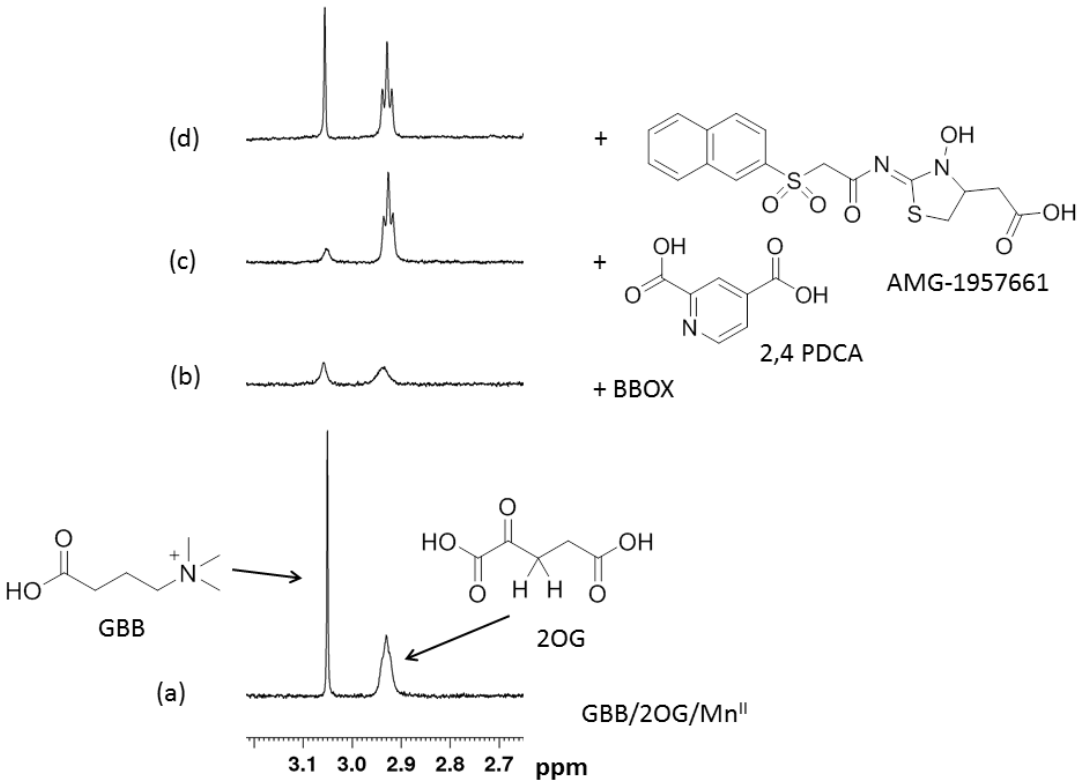


Figure 2.39 Illustration of the ^{1}H NMR (CPMG edited) based dual-reporter ligands displacement assay for ligand binding studies to psBBOX.
Both GBB and 2OG are simultaneously used as reporter molecules. (a) The spectrum (part shown for clarity) of a solution of GBB, 2OG and Mn^{II} in buffer. (b) After the addition of psBBOX to the sample of (a). The reporters signals are attenuated due to their binding to psBBOX. (c) The addition of 2,4-PDCA to the sample of (b) results in recovery of only 2OG resonances, showing that 2,4-PDCA binds psBBOX by displacing only 2OG.(d) The addition of AMG-1957661 to (b) results in the recovery of both 2OG and GBB NMR resonances, implying that AMG1957661 binds psBBOX by competing with both 2OG and GBB. The assay mixture contained 25 µM GBB, 300 µM 2OG, 150 µM Mn^{II}, 15 µM psBBOX, 80 mM KCl, competitive ligand 700 µM in 50 mM Tris-D_{11}, pH 7.5 in D_2O.

2.5.1 Binding of isoquinoline-based inhibitors to psBBOX by ^{1}H NMR-dual reporter displacement assay- A structure-activity relationship study

After establishing the proof of principle of NMR dual-reporter displacement assay for psBBOX inhibitor binding studies, we then applied this technique to investigate the nature of binding of a series of isoquinoline derivatives (provided by Dr Anna M. Rydzik) to psBBOX. All these compounds possessed the same isoquinoline-core skeleton but varied in the side chain amino acid (Figure 2.40). The isoquinoline-based compounds are known inhibitors of many 2OG oxygenases including hBBOX.[6,159] Crystal structures of several 2OG oxygenases in complex with isoquinoline-derivatives have also been reported.[159,189,201-203] These compounds are primarily designed as 2OG competitors. We were interested to test these compounds for two potential reasons (1) There is no reported crystal structure for psBBOX in

complex with this series of inhibitors and (2) in case of PHD2, the crystal structure in complex with an analogue (PDB ID: 2G19) containing the glycine side chain suggests that the inhibitor binds in the 2OG binding pocket by chelating the active site metal *via* the pyridine nitrogen and amide carbonyl oxygen.[189] However, solution-based NMR studies revealed that the inhibitor binds in the 2OG binding site in two different modes with equal population, one mode of metal chelation was similar to the one reported in the crystal structure, whereas, in the new mode, the inhibitor chelates the active site metal *via* the hydroxyl oxygen and amide carbonyl oxygen.

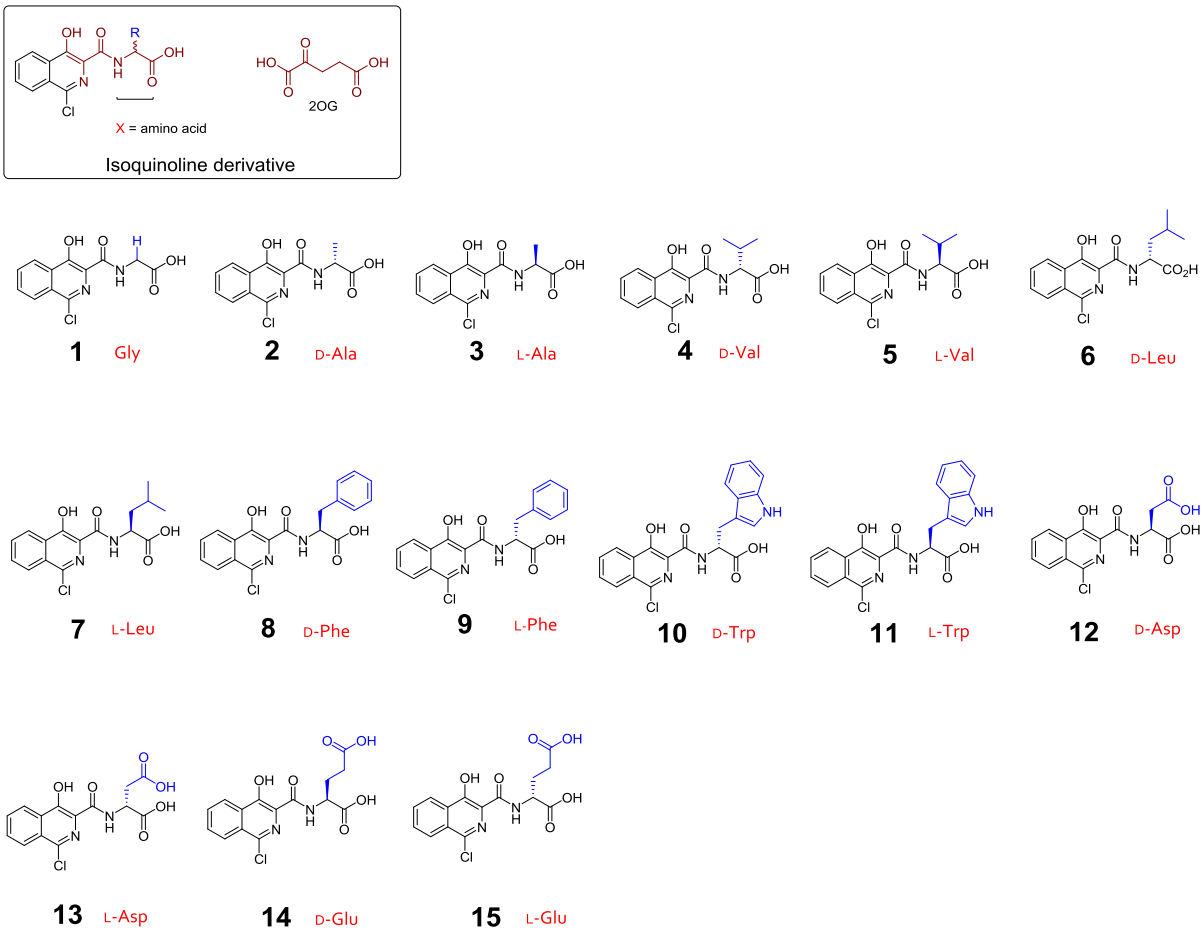


Figure 2.40 Structures of the isoquinoline-based compounds tested in the study.
In box: Core skeleton of the isoquinoline-based inhibitor. The brown colour represents part of the inhibitor that mimics 2OG. (Below the box): structures of the isoquinoline-based compounds tested for their binding interaction with psBBOX by CPMG edited ^{1}H NMR dual reporter displacement assay.

Exemplary results from the assay are presented in Figure 2.41 for two of these isoquinoline-based ligands, one possessing glycine (**1**) and the other a L-Trp (**11**) amino acid functionality. In this assay, the spectrum of the reporter molecules (GBB and 2OG) with Mn^{II} was first recorded in the absence of psBBOX (Figure 2.41a). A second spectrum of the reporter

molecules with Mn^{II} was then recorded following the addition of psBBOX (Figure 2.41b). In this spectrum, the reporter resonances are broadened and attenuated due to their binding interaction with psBBOX. To the same sample, isoquinoline-ligand (**1**) was added and a spectrum was recorded (Figure 2.41c). It can be seen that upon the addition of **1** only the signal of 2OG recovers, demonstrating that **1** binds to psBBOX by displacing (competing) only 2OG. Similarly, to another sample, isoquinoline-ligand (**11**) was added and it was observed that the signals of both 2OG and GBB recover (Figure 2.41d), demonstrating that (**11**) binds to psBBOX and competes with both 2OG and GBB.

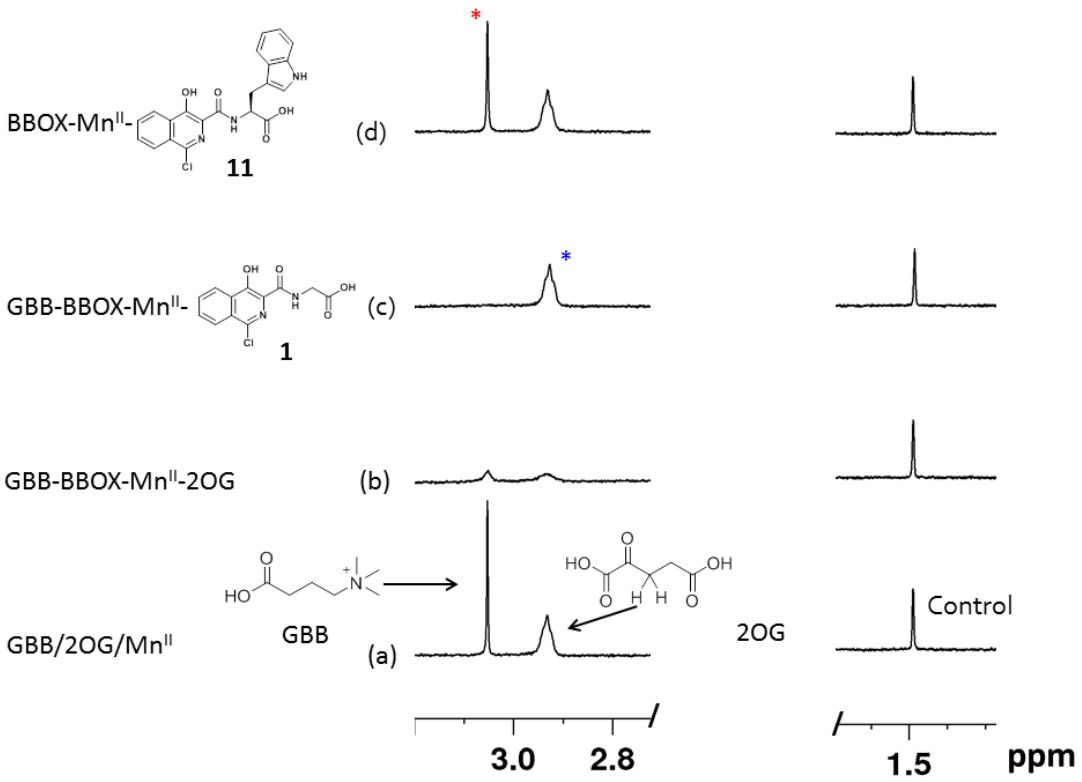


Figure 2.41 Studying the binding of isoquinoline-ligands (**1**) and (**11**) with psBBOX by CPMG edited ^{1}H NMR-based dual-reporter displacement assay.
GBB signal at 3.05 ppm, 2OG signal at 2.91 ppm, and 1,1,1 trifluoroacetone (internal NMR control) Me_3 peak at 1.49 ppm, are shown. (a) Reporter ligands GBB and 2OG, control molecule with Mn^{II} in solution in the absence of psBBOX (b) After the addition of psBBOX (c) Addition of (**1**) to sample of (b). (d) Addition of (**11**) to sample of (b). The asterisks represent the signal of the reporter ligand (2OG: blue, GBB: red) displaced after the addition of competitive ligand. The final assay mixture contained 25 µM GBB, 300 µM 2OG, 150 µM Mn^{II}, 15 µM BBOX, 80 mM KCl and 25 µM isoquinoline-based ligand in 50 mM Tris-D11 buffer, pH 7.5, in D_2O.

The results for all the isoquinoline-ligands tested in the series are summarised in Figure 2.42. All these ligands were screened at a single concentration (25 µM) in order to produce a contrast amongst the nature of their binding with psBBOX. The chart compares the signal intensity of the reporter molecules GBB (green colour) and 2OG (orange colour) in the

absence of psBBOX (1st left set of column; intensity set to unit intensity) with their intensity after the addition of psBBOX (2nd left set of column) and with their intensity after the addition of isoquinoline-based inhibitors (**1**)-(**15**) with varying side chain amino acid (X). The threshold line represents the relative intensity of the reporters in the presence psBBOX. After the addition of isoquinoline ligand, the reporters intensity above this line represents the level of their displacement, whereas their intensity below this line represents further enhancement of their binding. Asterisks highlight those isoquinoline-based ligands whose addition results in a negligible influence on the reporters under these conditions.

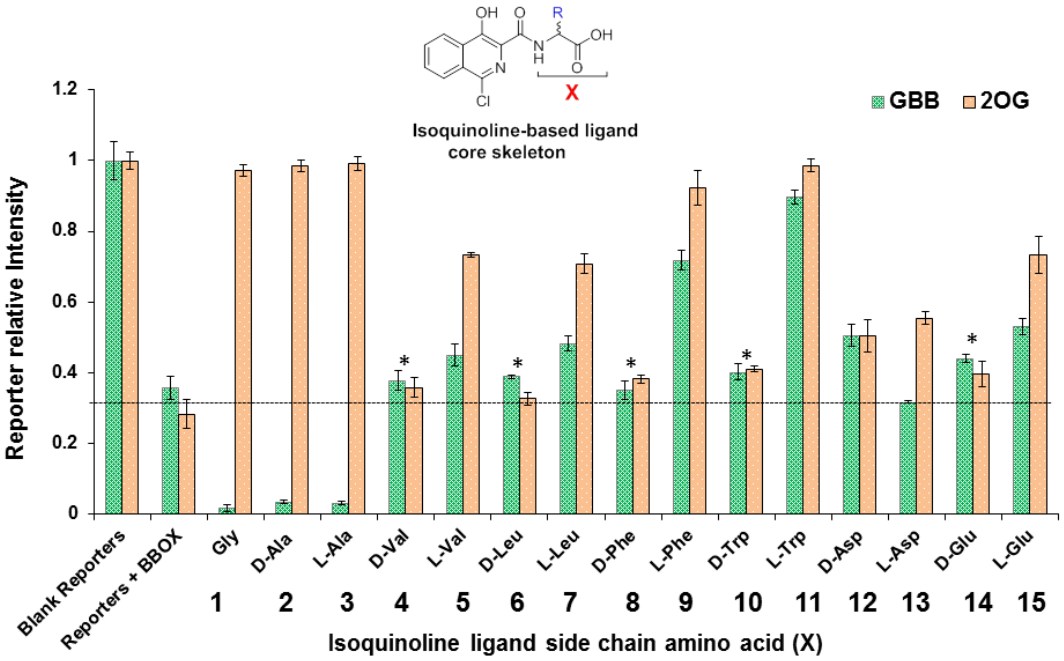


Figure 2.42 Structure-activity relationship study of isoquinoline-based compounds with psBBOX by CPMG edited ^{1}H NMR-based dual-reporter displacement assay.
The chart compares the signal intensity of GBB (green) and 2OG (orange) in the absence of psBBOX (1st left set of column; intensity set to unit intensity) to their intensity after the addition of psBBOX (2nd left set of column) and that after the addition of isoquinoline-based inhibitors **1**-**15** with varying amino acid (x-axis). The threshold line represents the relative intensity of the reporters in the presence of psBBOX. After the addition of isoquinoline ligand, the level above this line represents displacement; the level below represents enhancement of binding. Asterisks highlight ligands whose addition results in a negligible difference on GBB/2OG under these conditions. The error bars represent standard deviation from three separate measurements.

From these results it may be concluded that isoquinoline-ligand containing glycine, (**1**) or alanine side chains (both D (**2**) and L (**3**) forms) appears to bind psBBOX by competing with only 2OG. The full displacement of 2OG by these ligands suggests that they are likely strong binders of psBBOX, binding *via* active site metal chelation. The data further suggest that the

binding of these compounds enhances the binding of GBB relative to its binding in the presence of 2OG as demonstrated by further attenuation of GBB signals.

However, with an increase in size of the amino acid side chain from H (**1**) and CH_3 (**2**), (**3**) to isopropyl (**5**) and isobutyl (**7**) (both in L-forms), the ligand appears to bind less favourably as demonstrated by relatively lower level of 2OG displacement compared to (**1**), (**2**) and (**3**). The ligands (**5**) and (**7**) also compete with GBB, but only weakly, as shown by low level of displacement of GBB. The displacement of both 2OG and GBB by these ligands suggests that these ligands may access to occupying the 2OG binding pocket, whilst also blockin or occupying, the GBB binding pocket. The D forms of these ligands i.e (**4**) and (**6**), appear not to be capable of binding in the psBBOX active site, likely because they cannot occupy a suitable orientation that enables coordination to the active site metal. The introduction of a more bulky aromatic hydrophobic group such as phenyl (**9**) or 3-methyl-indole (**11**), as the L-isomers enables the binding of the ligand to again yield full displacement of 2OG as for (**1**)-(**3**). These ligands also appear to prevent the binding of GBB, (in the order (**11**) > (**9**), to much greater extent than (**5**) and (**7**). It is possible that these isoquinoline- ligands bind psBBOX in a mode such that their side chain group is located in the aromatic cage of the GBB binding pocket that normally holds the trimethylammonium group of GBB and thus form hydrophobic interaction with the aromatic cage residues. In the case of (**11**) and (**9**) this would be more pronounced when compared with (**5**) and (**7**). Similar to (**4**) and (**6**), the D-forms of these ligands i.e (**8**) and (**10**), also appear not to be effective at binding in the psBBOX active site.

The influence of hydrophilic side chain groups on ligand binding was also investigated. It was observed that ligands with Asp (both D (**12**) and L (**13**) forms) are weak 2OG competitors (as demonstrated by low level of 2OG displacement) relative to (**5**), (**7**), (**9**) and (**11**) (all ligands with hydrophobic group). It is possible that ligands (**12**) and (**13**) may be binding in an orientation, such that their side chain group is projecting towards the aromatic cage of GBB, producing unfavourable repulsion between the π-electrons and the carboxylate group. Interestingly, (**12**) compete weakly with both 2OG and GBB, whereas, (**13**) competes with only 2OG, suggesting these two ligands may bind in different modes relative to each other. Ligand (**14**) (D-Glu), appears to compete only very weakly with 2OG as compared to (**12**) and (**13**). In contrast its L-form (**15**) is again a relatively moderate 2OG competitor which also moderately competes with GBB.

The results of this qualitative structure-activity assay were then validated by measuring the binding dissociation constants of these isoquinoline-based ligands by the tryptophan intrinsic fluorescence quenching assay (performed by Robert K. Lesniak) previously applied to study ligand binding to human BBOX (described in section 2.2.12). The results of this assay (Table 2.2) also demonstrated that ligands with glycine (**1**), D-Ala (**2**), L-Ala (**3**), L-Phe (**9**) and L-Trp (**11**) are the strongest binders in the series, in agreement with the results of the ^{1}H NMR dual reporter displacement assay. Similarly, ligands with L-Val (**5**), L-Leu (**7**) and L-Glu (**15**) are considered medium affinity ligands in the series.

Table 2.2 ^{1}H NMR and fluorescence-based assays results of the isoquinoline-based compounds (Figure 2.40) tested in the NMR dual reporter displacement assay (Figure 2.42)

Compound no.	Amino acid (X)	K_D (μM) by Fluorescence assay
1	Gly	22 ± 4
2	D-Ala	20 ± 4
3	L-Ala	15 ± 3
4	D-Val	144 ± 6
5	L-Val	55 ± 3
6	D-Leu	103 ± 4
7	L-Leu	55 ± 3
8	D-Phe	82 ± 3
9	L-Phe	26 ± 4
10	D-Trp	168 ± 37
11	L-Trp	29 ± 5
12	D-Asp	101 ± 1
13	L-Asp	91 ± 2
14	D-Glu	89 ± 6
15	L-Glu	64 ± 2

The similar binding affinities of these ligands are also qualitatively reflected by similar levels of 2OG displacement by these ligands in the ^{1}H NMR dual-reporter displacement assay. Similarly, the fluorescence binding assay also demonstrated that the D-configured ligands, (**4**), (**6**), (**8**), (**10**), (**12**), (**13**) and (**14**) are the weakest binders in the series, consistent with the ^{1}H NMR dual reporter displacement assay. Thus, the results of this fluorescence based assay further validate the applicability by ^{1}H NMR dual reporter displacement assay for qualitative ligand screening of psBBOX.

Based on the results of both these binding assays, it appears that the level of 2OG displacement can be correlated with the binding affinity of these isoquinoline-based ligands. For example isoquinoline ligands (**1**), (**2**), (**3**), (**9**) and (**11**) all fully displace 2OG and have binding affinities also close to each other (22, 20 15, 26 and 29 µM, respectively). Similarly, ligands (**5**), (**7**) and (**15**), all compete relatively moderately with 2OG compared to ligands (**1**), (**2**), (**3**), (**9**) and (**11**), and again have binding affinities (55, 55 and 64 µM) that are weaker than those of ligands (**1**), (**2**), (**3**), (**9**) and (**11**). Similarly, ligands (**12**) and (**13**), both of which displace 2OG to the same extent have similar affinities (101, 91 µM respectively), are also weak binders compared with ligands (**5**), (**7**) and (**15**). Thus, whilst the level of 2OG displacement correlates well with the absolute binding affinity, (which follows from the fact that the isoquinoline core is designed for chelation of the active site metal), the NMR assay further provides evidence for disruption of GBB binding.

Although, we could not solved a crystal structure for psBBOX with an isoquinoline-based inhibitor to assess the nature of binding of these inhibitor; manual docking and energy minimisation studies of a model of psBBOX generated based the crystal structure of human BBOX (PDB ID: 3O2G) suggests a possible structural explanation for the observed trends (Figure 2.42) in this study. Studies by Rydzik *et al*[159] suggests that for hBBOX, the L-configured ligands (**3**) and (**11**) are better inhibitors than their corresponding D-forms (**2**) and (**10**) (IC_{50} 6 µM *vs* 73 µM, and 11 µM *vs* 33 µM, respectively), in agreement with the trend in our studies with psBBOX. The docking studies revealed that (**1**) binds in the active site of psBBOX in a similar mode as observed in the crystal structures of other 2OG oxygenases, i.e. PHD2 (PDB ID: 2G1M)[203] and the fat mass and obesity protein FTO (PDB ID: 4IE6)[204], likely by chelating the active site metal *via* its pyridinium nitrogen and amide carbonyl oxygen. Although more than one chelation modes such as *via* ring hydroxyl group and amide oxygen are also possible, as reported in studies by Rydzik *et al* on hBBOX[159], which suggest that removal of ring hydroxyl group results in significant loss in inhibitor potency. In our proposed binding mode for (**1**), the phenyl ring of the isoquinoline functionality is positioned towards the GBB binding aromatic cage, thereby leading to slight movement of Phe184, but it does not disrupt the aromatic cage pocket to the extent so as the GBB binding is ablated (Figure 2.43), which agrees with the NMR binding mode for (**1**)-(**3**), i.e (**1**)-(**3**) displaces only 2OG but not GBB.

The observation that inhibitors (**5**), (**7**), (**9**), (**11**) (ligands with larger hydrophobic side chains) compete with both co-substrate 2OG and substrate GBB, is interesting. Manual docking and

energy minimisation studies of (**11**) suggest that the bicyclic ring system of these inhibitors is tilted with respect to that of the docked/minimised compound (**1**), while retaining its metal coordination. The chlorine and phenyl ring of the inhibitor projects deeper into the GBB binding aromatic cage, resulting in movement of Phe184 to approximately 2.5 Å closer to the aromatic inhibitor thus altering the shape of the aromatic cage and also sterically blocking binding of the GBB trimethylamino group.[159] This observation is consistent with the NMR binding mode i.e. inhibitor competes with both 2OG and GBB.

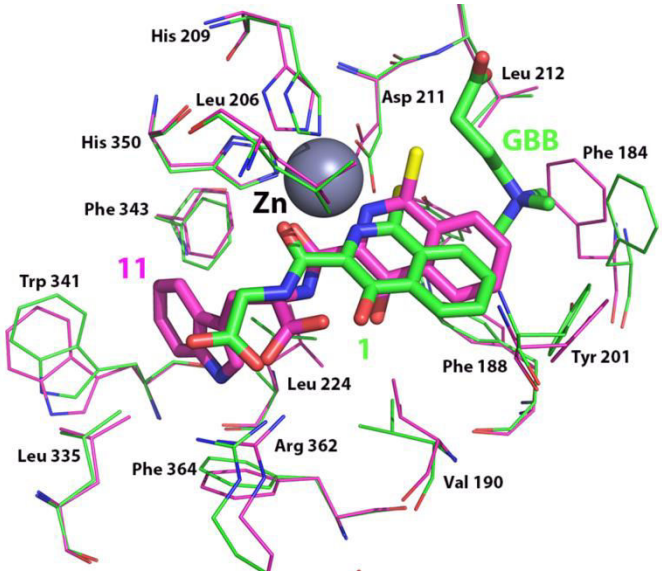


Figure 2.43 Overlay of the active site residues of human and psBBOX showing possible modes of binding (**1**) and (**11**). psBBOX (model generated; pink residues with numbers) and human BBOX (green residues) (PDB ID: 3O2G). Docking simulation implying possible conformational changes in the GBB binding pocket residues of psBBOX as a result of binding of inhibitors **1** (green sticks, Cl in yellow) and **11** (pink sticks, Cl in yellow).

Although, the possibility of alternative binding/chelation modes also cannot be ruled out, the modelling and energy minimization studies suggest that the larger side chains bind in a large hydrophobic cavity adjacent to the 2OG binding pocket (formed by the side chains of Val-190, Ala-200, Leu-206, Ser-236, Leu-224, Leu-335, Trp-341, Phe-343, Arg-352 and Phe-364). It also implies that in this pocket the binding of L- forms of inhibitors are preferred as compared to the D- forms as observed (Table 2.2, Figure 2.43), due to a potential clash between the carboxylate of (**10**) and Leu-206. The psBBOX model structure suggests that that a loop comprising of residues 190-205 involved in GBB binding also forms part of the hydrophobic pocket; it is also possible that the binding of inhibitor (**11**) may elicit changes in the loop and prevents GBB binding. Biophysical studies have revealed the occurrence of substantial conformational changes in human BBOX catalysis.[159,160,163]

2.6 Development of ^{1}H NMR assay for IC_{50} measurements for BBOX inhibitors

To determine whether a small molecule ligand that binds to psBBOX do indeed inhibit psBBOX catalysis, it is added to the assay solution of psBBOX and see if it stops the GBB turnover into L-carnitine. This can be easily observed from a standard ^{1}H NMR spectrum as both GBB and L-carnitine gives separate NMR resonances. Due to high sensitivity, the Me_3 peak; GBB at 3.05 ppm, L-carnitine 3.14 ppm can be conveniently followed to monitor the inhibition of psBBOX catalytic activity in the presence of an inhibitory ligand. The reaction can be initiated by adding psBBOX to the assay solution, usually contained in an eppendorf tube.

Like hBBOX, psBBOX was also found to be very active, required only nM concentrations to observe a reasonable amount of turnover. In monitoring psBBOX catalysed GBB turnover into L-carnitine, we noted that the speed/amount of L-carnitine formation from incubation the reaction mixture in a 1.5 mL Eppendorf tube was higher as compared to the reaction in a 5 mm diameter NMR tube (psBBOX added to the assay mixture already in the NMR tube), whereas, in a 3 mm diameter NMR tube, the reaction did not occurred at all. We speculate that this difference in the rate and level of L-carnitine is perhaps either due to the differences in the rates of diffusion of oxygen across the sample or due to the surface area (Eppendorf tube > 5 mm diameter NMR tube > 3 mm diameter NMR tube) of the reaction containers.

In addition to maximum turnover, the reaction in Eppendorf tube also allowed us to easily quench the enzymatic reaction whenever required, hence presented a convenient approach to measure the L-carnitine/GBB concentration at a desired time. We found that 5 µL of 1M HCl can effectively quench (stop) the psBBOX catalysed GBB turnover into L-carnitine. In the presence of an inhibitory ligand, the L-carnitine signal was found to decrease in intensity as the inhibitory ligand concentration increase. For the IC_{50} measurement of the inhibitors tested in the study, the reaction was quenched at the 60^{th} second (linear rate region) (Figure 2.44).

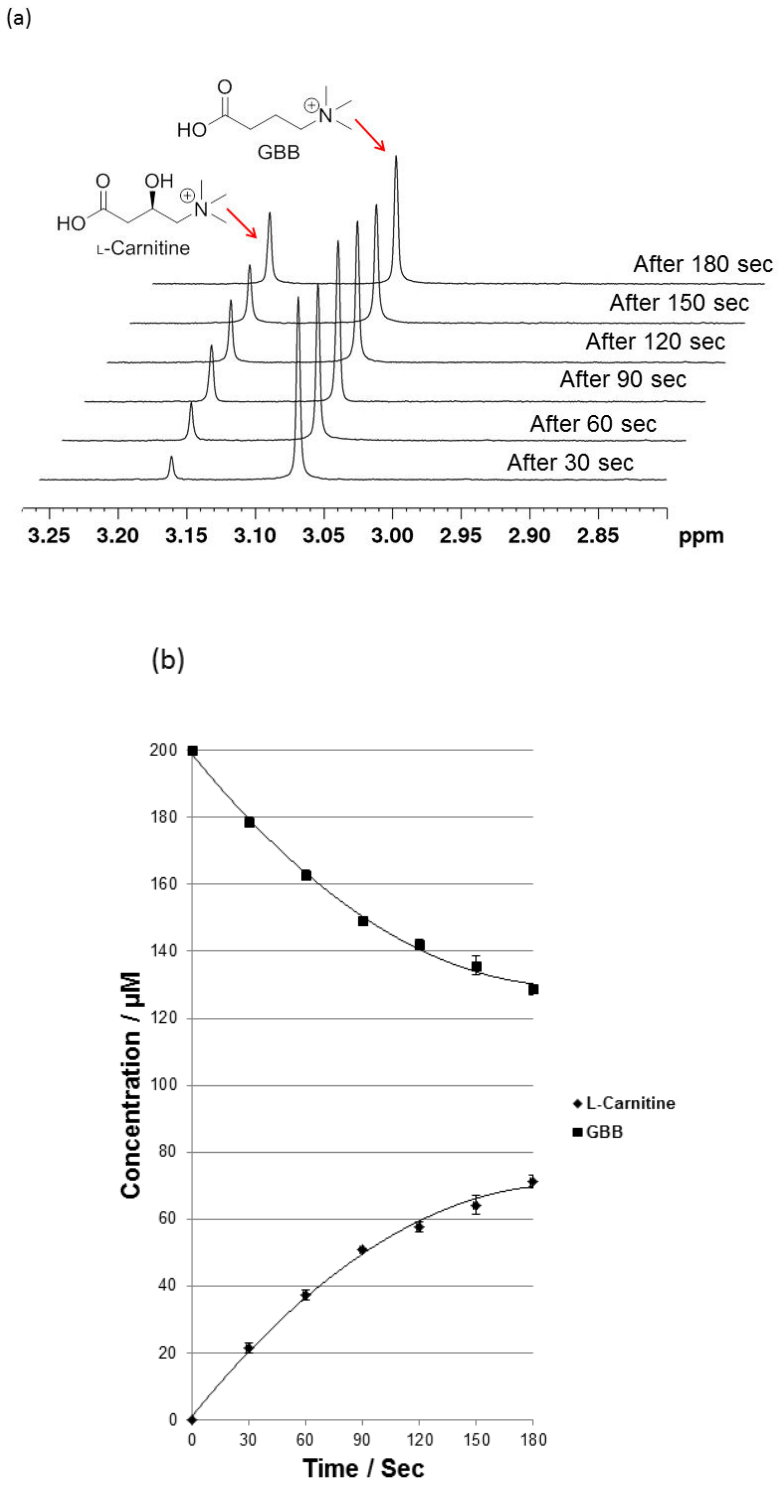


Figure 2.44 psBBOX catalysed GBB turnover into L-carnitine as observed by ^{1}H NMR.
(a) Time course of the psBBOX catalysed GBB turnover into L-Carnitine as monitored by ^{1}H NMR experiments (part of the spectra shown). (b) Graph of the time course data. The reaction was carried out in a 1.5 mL eppendorf tube and quenched by the addition of 5 µL of 1 M HCl. The assay mixture contained 200 µM GBB, 600 µM 2OG, 100 µM Fe^{II}, 140 nM psBBOX, 500 µM L-ascorbate, 80 mM KCl in 50 mM Tris-D11, pH 7.5 in D_2O. Error bars represents standard deviation from three separate measurements.

This technique offers one of the simplest approaches to measure the IC_{50} values of BBOX inhibitors. We applied this technique to measure the IC_{50} values of the isoquinoline derivatives tested in section 2.5.1. The inhibition potencies of these ligands are summarised in Table 2.3. An exemplary result from the assay is shown in Figure 2.45 for isoquinoline derivative containing the D-alanine side amino acid side chain (**2**), for which an IC_{50} of 0.20 µM was obtained. These ^{1}H NMR-based inhibition results were further validated by the reported fluoride ion release-based assay (performed by Robert K. Lesniak) (described in section 2.2.12). The results of this assay (Table 2.3 & Figure 2.46) also revealed that the ligand with glycine (**1**), D-Ala (**2**), L-Ala (**3**), L-Phe (**9**) and L-Trp (**11**) are the strongest psBBOX inhibitors in the series, with other values are also qualitatively paralleling those of the NMR-derived values.

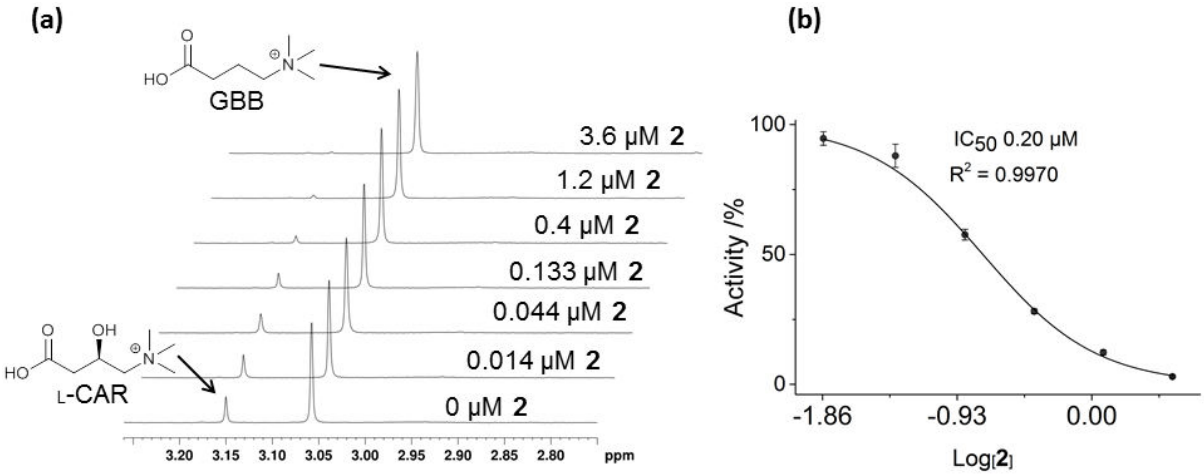
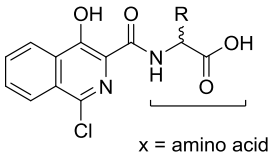


Figure 2.45 Measuremt of IC_{50} value of the isoquinoline-ligand (**2**) by for psBBOX using ^{1}H NMR.
(a) Bottom to top; ^{1}H NMR showing psBBOX catalysed GBB turnover into L-carnitine (CAR) in the presence of increasing amounts of (**2**). (b) Dose response curve of (a). Error bars represent standard deviations from three separate measurements. The assay mixture contained 200 µM GBB, 600 µM 2OG, 100 µM Fe^{II}, 500 µM L-ascorbate, 80 mM KCl and 0.140 µM BBOX in 50 mM Tris-D_{11}, pH 7.5 in D_2O.

The differences in IC_{50} values of ^{1}H NMR and fluoride ion release assays, particularly for weak inhibitors, are likely due to the different experimental conditions used in both the assays, notably the higher psBBOX concentration (1 µM) used in the fluoride ion release assay (0.140 µM in ^{1}H NMR assay). Also, the ^{1}H NMR assay was based on the inhibition of psBBOX whilst using the natural substrate GBB, whereas the fluoride assay is based on the inhibition of fluorinated substrate (3*S*)-3-fluoro-4-(trimethylammonio)butanoate (GBBF). Study by Rydzik *et al*[159] on hBBOX using fluoride ion release-based assay have shown that the L-configured ligands (**3**) and (**11**) are better inhibitors than their corresponding D-forms (**2**

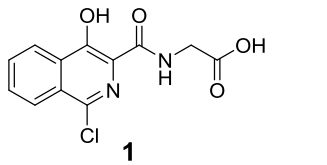
and **10**) (IC_{50} 6 µM vs 73 µM, 11 µM vs 33 µM). Similarly, (**1**) is a weak inhibitor (IC_{50} 31 µM) than (**3**).

Table 2.3 Comparison of the IC_{50} values measured by ^1H NMR assay (Figure 2.46) and fluoride ion release assay for the isoquinoline psBBOX inhibitors tested in the study.

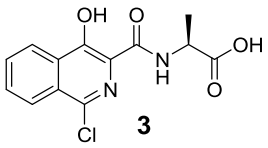


Inhibitor core skeleton

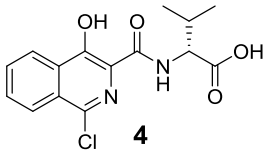
Compound no.	Amino acid (X)	IC_{50} (µM) by ^1H NMR assay	IC_{50} (µM) by Fluoride ion release assay
2	D-Ala	0.20 ± 0.02	0.9 ± 0.03
1	Gly	0.21 ± 0.01	1.2 ± 0.1
3	L-Ala	0.22 ± 0.01	0.7 ± 0.02
11	L-Trp	0.9 ± 0.03	26 ± 2
9	L-Phe	1 ± 0.03	13 ± 0.7
10	D-Trp	18 ± 1	> 1000
15	L-Glu	22 ± 2	147 ± 21
5	L-Val	23 ± 2	130 ± 14
7	L-Leu	26 ± 2	125 ± 6
8	D-Phe	38 ± 0.6	101 ± 6
12	D-Asp	38 ± 1	160 ± 61
13	L-Asp	76 ± 2	245 ± 62
14	D-Glu	86 ± 5	> 1000
6	D-Leu	94 ± 8	> 1000
4	D-Val	191 ± 6	> 1000



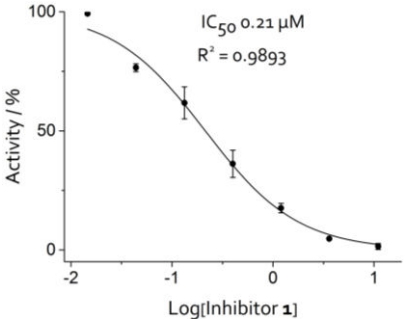
1

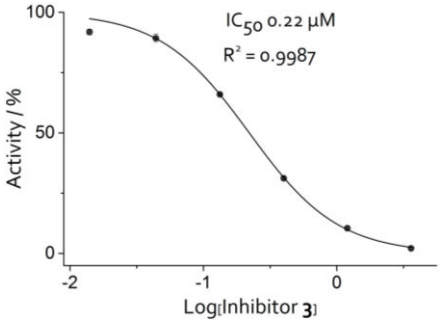


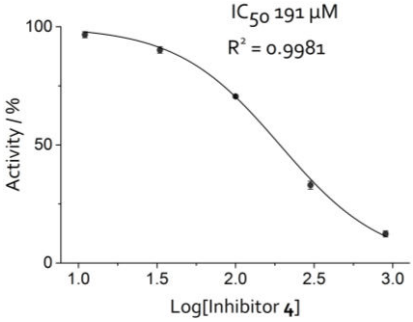
3

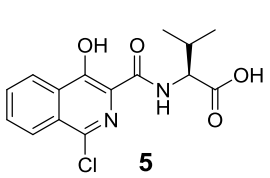


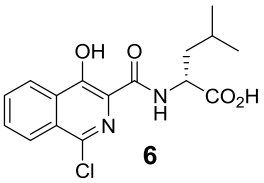
4

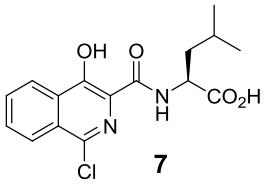


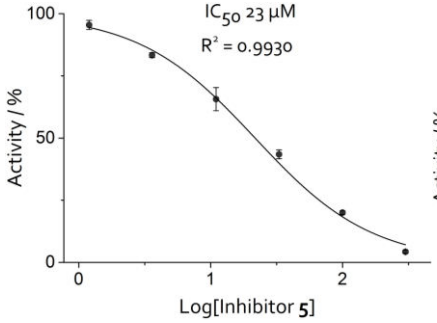


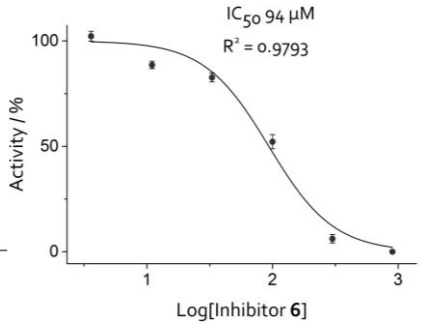


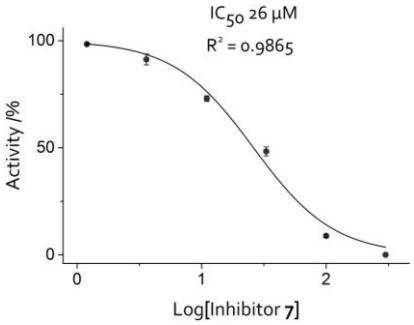

OH
O
N
H
OH
N
Cl
5

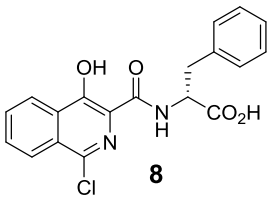

OH
O
N
H
CO2H
N
Cl
6

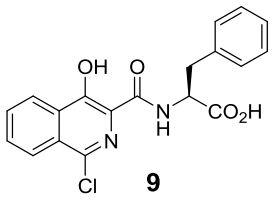

OH
O
N
H
CO2H
N
Cl
7

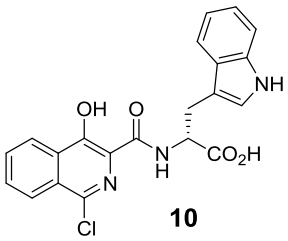

IC50 23 µM
R² = 0.9930
Activity / %
100
50
0
0
1
2
Log[Inhibitor 5]

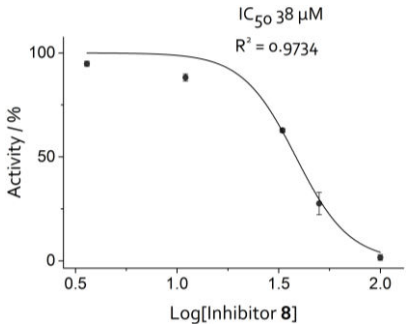

IC50 94 µM
R² = 0.9793
Activity / %
100
50
0
1
2
3
Log[Inhibitor 6]

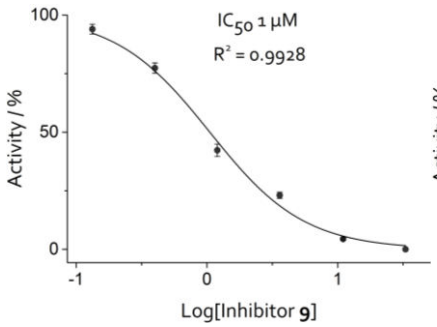

IC50 26 µM
R² = 0.9865
Activity /%
100
50
0
0
1
2
Log[Inhibitor 7]

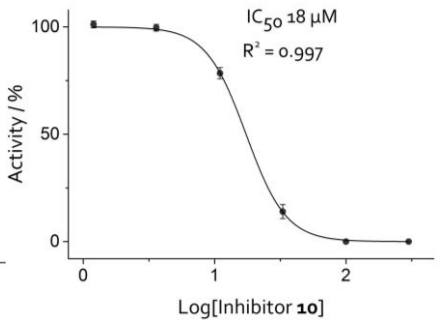

OH
O
N
H
CO2H
N
Cl
8

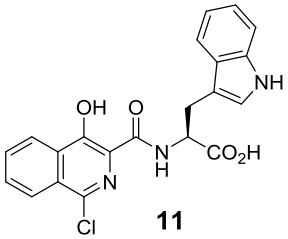

OH
O
N
H
CO2H
N
Cl
9

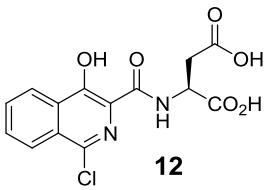

NH
OH
O
N
H
CO2H
N
Cl
10

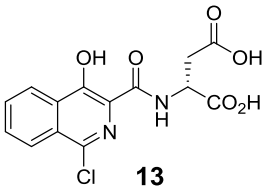

IC50 38 µM
R² = 0.9734
Activity / %
100
50
0
0.5
1.0
1.5
2.0
Log[Inhibitor 8]

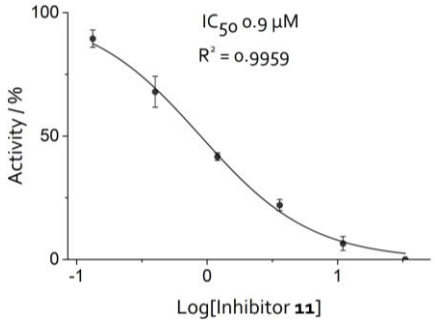

IC50 1 µM
R² = 0.9928
Activity / %
100
50
0
-1
0
1
Log[Inhibitor 9]

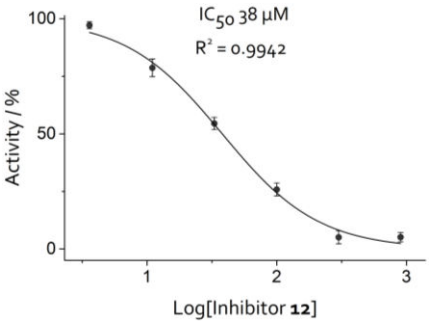

IC50 18 µM
R² = 0.997
Activity / %
100
50
0
0
1
2
Log[Inhibitor 10]

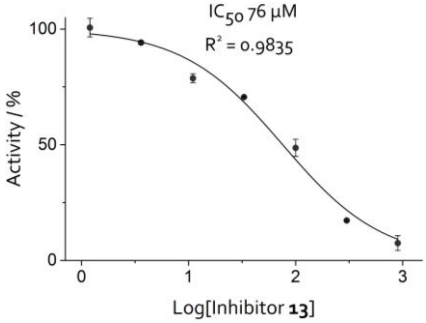

NH
OH
O
N
H
CO2H
N
Cl
11


O
OH
OH
O
N
H
CO2H
N
Cl
12


O
OH
OH
O
N
H
CO2H
N
Cl
13


IC50 0.9 µM
R² = 0.9959
Activity / %
100
50
0
-1
0
1
Log[Inhibitor 11]


IC50 38 µM
R² = 0.9942
Activity / %
100
50
0
1
2
3
Log[Inhibitor 12]


IC50 76 µM
R² = 0.9835
Activity / %
100
50
0
0
1
2
3
Log[Inhibitor 13]

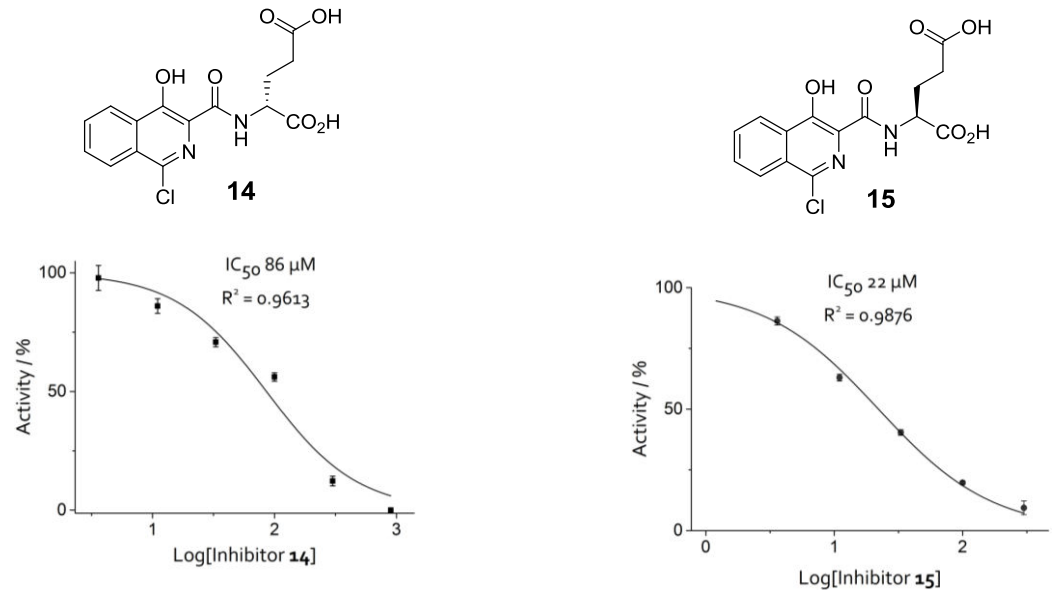
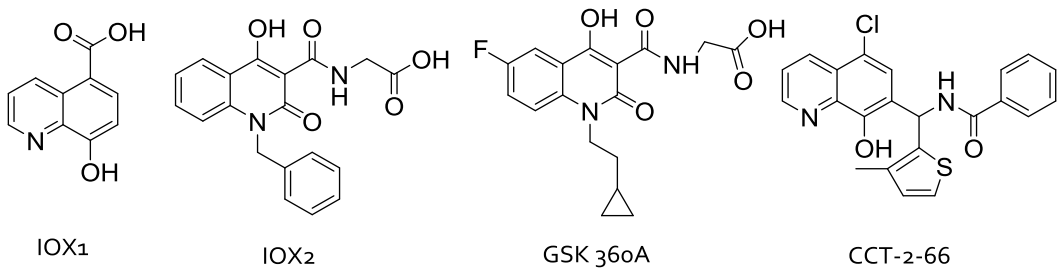


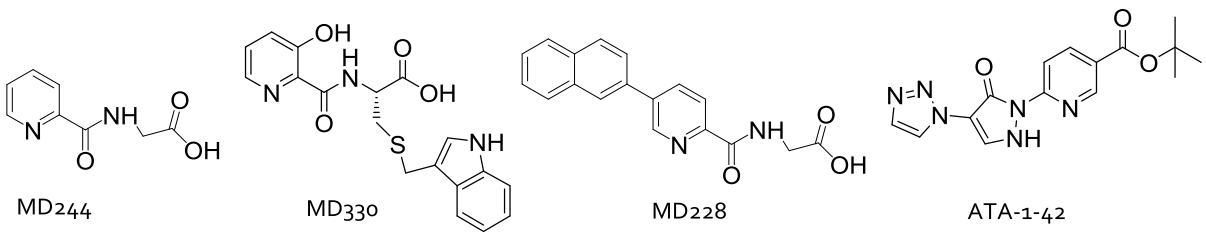
Figure 2.46 Dose response curves of the isoquinoline-based inhibitors of psBBOX tested in the SAR study conducted by ^{1}H NMR inhibition assay. The error bars represent standard deviation from three separate measurements.

2.6.1 Binding studies of selected quinoline, pyridine and pyrimidine based ligands with psBBOX by ^{1}H NMR dual-reporter displacement assay

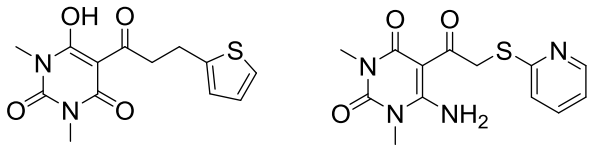
Many of the 2OG oxygenase inhibitors including hBBOX are also based on quinoline, pyridine and pyrimidine core-skeleton.[6,159] These inhibitors are primarily designed as 2OG competitors for active site metal chelation. Figure 2.47 shows the structure of some of selected quinoline, pyridine and pyrimidine based compounds that have been found to inhibit various 2OG oxygenases. As no crystal structure has reported for BBOX with these compounds, therefore we applied the ^{1}H NMR-based dual-reporter displacement assay to assess the binding nature of these compounds (provided by Dr Ivan K. H. Leung) to psBBOX. The results of these compounds are summarised in Figure 2.48. For the ligands that showed binding, we determined their IC_{50} values (Table 2.4 & Figure 2.49) by ^{1}H NMR-based assay as described in section 2.6.



Compounds with quinoline core skeleton



Compouns with pyridine core skeleton



Compounds with pyrimidine core skeleton

Figure 2.47 Structures of the compounds with different core skeleton tested for their binding with psBBOX by CPMG edited ^{1}H NMR-based dual-reporter displacement assay.

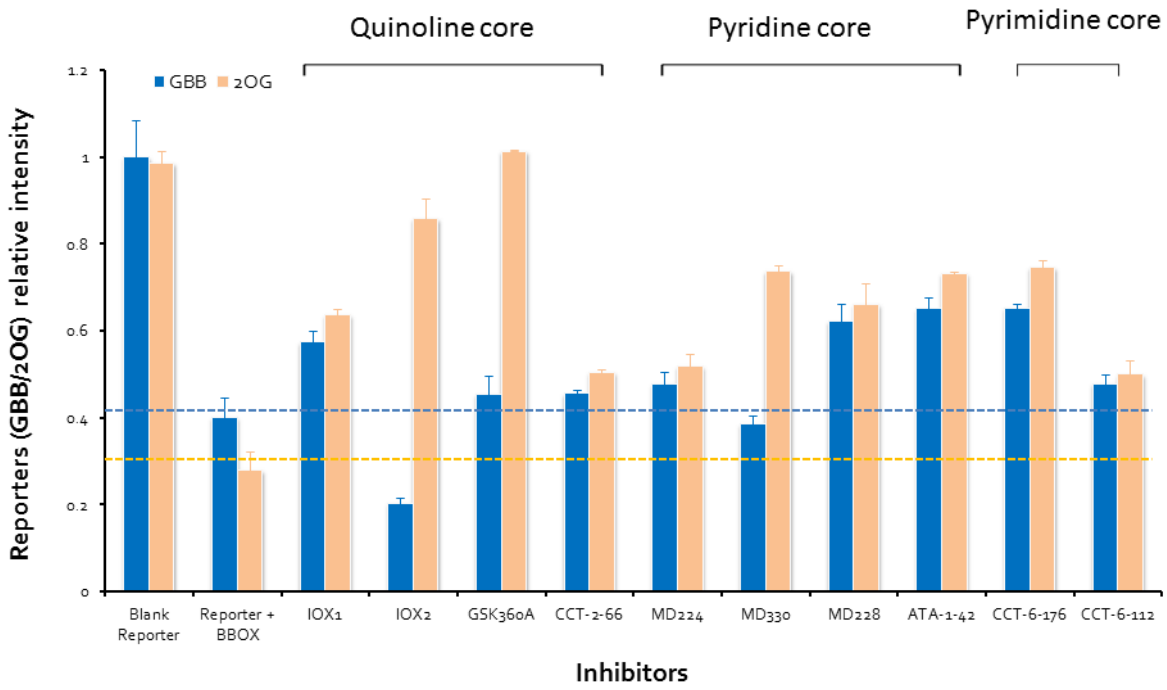
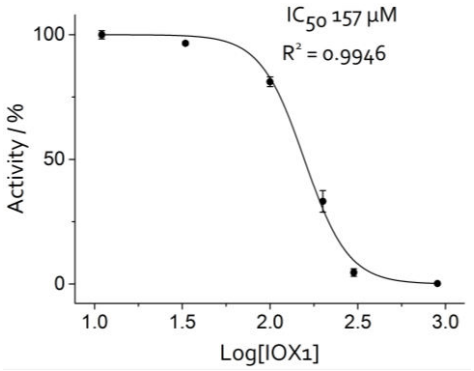
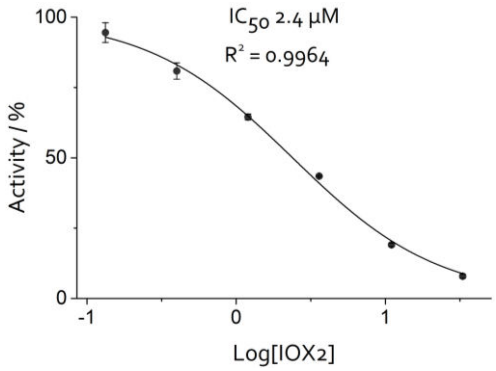


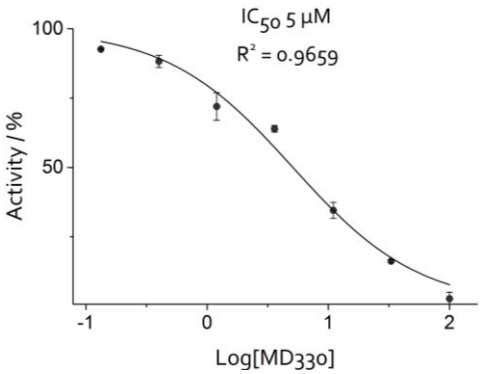
Figure 2.48 Investigating the binding of quinoline, pyridine and pyrimidine-based ligands with psBBOX by CPMG edited ^{1}H NMR-based dual-reporter displacement assay.
Chart showing the effect of addition of ligand of interest (x-axis) on the signal intensity (threshold lines) of the reporter ligands GBB and 2OG, whey they bind with psBBOX. After the addition of screening ligand of interest, the reporters intensity above the threshold line represents the level of their displacement, whereas, their intensity below this line represents further improvements in their binding. The final assay mixture contains 25 µM GBB, 300 µM 2OG, 150 µM Mn^{II}, 15 µM BBOX, 50 µM F_3CCOCH_3 (internal control), 80 mM KCl and 25 µM screening ligand in 50 mM Tris-D_{11} buffer, pH 7.5, in D_2O. The error bars represent standard deviation from three separate measurements.

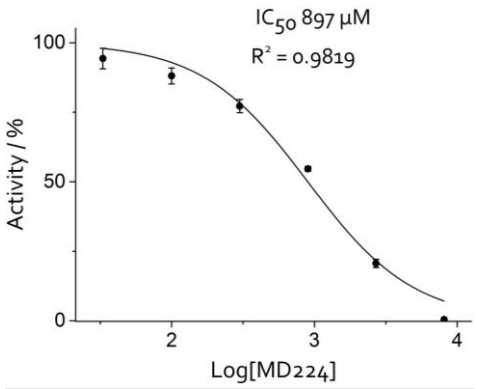
Table 2.4 IC_{50} values of psBBOX inhibitors determined by ^{1}H NMR assay

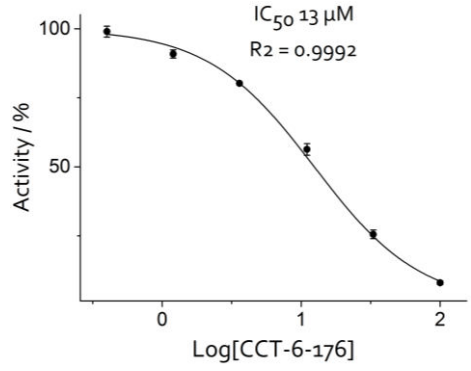
Ligand	IC_{50} (μM) by ^{1}H NMR assay
IOX1	157 ± 5
IOX2	2.4 ± 0.01
GSK360A	0.786 ± 0.02
CCT-2-66	not determined
MD224	897 ± 7 μM
MD330	5 ± 0.8
MD228	not determined
ATA-1-42	6 ± 0.6
CCT-6-176	13 ± 1
CCT-6-112	not determined

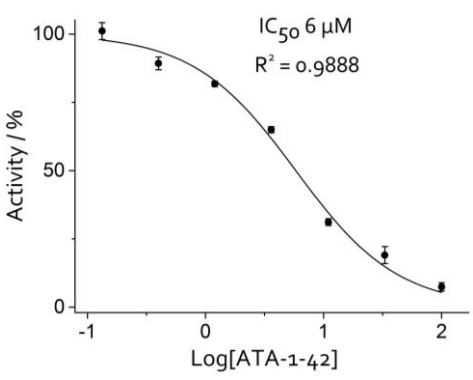












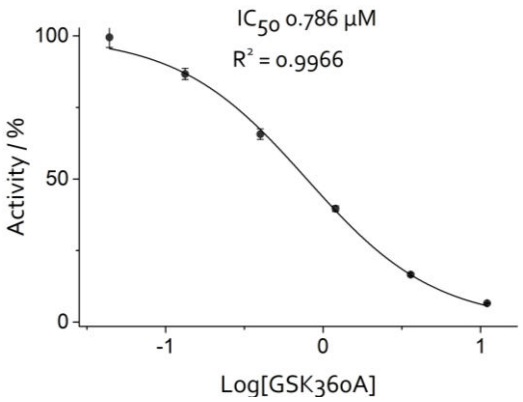


Figure 2.49 Dose response curves of the IC_{50} values of the inhibitors in Table 2.4. The error bars represent standard deviation from three separate measurements.

The above results (Figure 2.48 & Table 2.4) suggest that amongst the quinoline-based compounds, GSK360A appears to be the stronger psBBOX binder (hence a strong inhibitor with IC_{50} 0.786 µM) as demonstrated by high its high level of 2OG displacement. IOX_2, an inhibitor of HIF hydroxylases,[202] binds relatively weakly than GSK360A, as demonstrated by relatively lower level of 2OG displacement, with an IC_{50} 2.4 µM, also competing against only 2OG. The binding of IOX_2 also appears to improve the binding of GBB with psBBOX. IOX_1, which is a cell-active selective inhibitor of histone demethylases,[205] binds further weakly as demonstrated by its low level of 2OG and GBB displacement, with an IC_{50} of 157 µM. The data suggests that both GSK360A and IOX2 appears to bind in identical mode i.e both are competing with 2OG only, however, IOX_1 likely binds differently to GSK360A and IOX2, competing both 2OG and GBB, but perhaps in unfavourable positon as demonstrated by its lower IC_{50} 157 µM. CCT-2-66, an inhibitor of KDM4 appears to have no influence on the binding of both GBB and 2OG under our experimental conditions, likely a non-binder or very weaker binder of psBBOX (IC_{50} not determined).

Amongst the compounds containing the pyridine core skeleton, MD330 (nucleic acid demethylase inhibitor), MD228 and Ata-1-42 all appear to bind psBBOX with identical affinity as demonstrated by identical level of 2OG displacement by these compounds, hence also have similar inhibition potencies i.e IC_{50}; 5 µM (MD330), 6 µM (Ata-1-42), IC_{50} not determined for MD228. Although both MD228 and Ata-1-42 have different core skeleton, but they appear to bind identically; both are competing with both 2OG and GBB. MD330 binds differently to MD228 and Ata-1-42; competing with 2OG only. In this series the addition of MD224 has negligible influence on the binding of both the reporters and is a weakest inhibitor with IC_{50}; 897 µM.

In the pyrimidine series, the data suggests that CCT-6-176 (OGFOD1 inhibitor) binds psBBOX by competing with both 2OG and GBB (IC_{50} 13 μM), whereas, the CCT-6-112 (HIF hydroxylases inhibitor)[206] appears to have negligible effect on the binding of both GBB and 2OG and likely a very weak inhibitor of psBBOX (IC_{50} not determined).

The above results suggests that the ^{1}H NMR-based dual-reporter displacement assay (section 2.5) and ^{1}H NMR-based IC_{50} measurement assay (section 2.6) are useful techniques that readily reveal how inhibitors containing different core skeleton are likely to bind in the active site of BBOX, and what are their potencies relative to each other, respectively.

2.6.2 Binding of TCA cycle intermediates to BBOX

The acyl-coenzyme A (CoA) produced as a result of fatty acid β-oxidation in mitochondria enters the tricarboxylic acid (TCA) cycle, in which it ultimately breaks down and produces metabolic energy in the form of ATP (Figure 2.50). 2OG which is the co-substrate of BBOX, itself is also a TCA cycle intermediate. Due to this direct linkage between the fatty acid metabolism and β-oxidation, it is possible that other TCA cycle intermediates may also bind BBOX and perhaps may have a role in regulating cellular L-carnitine levels *via* psBBOX inhibition (2OG competitors).

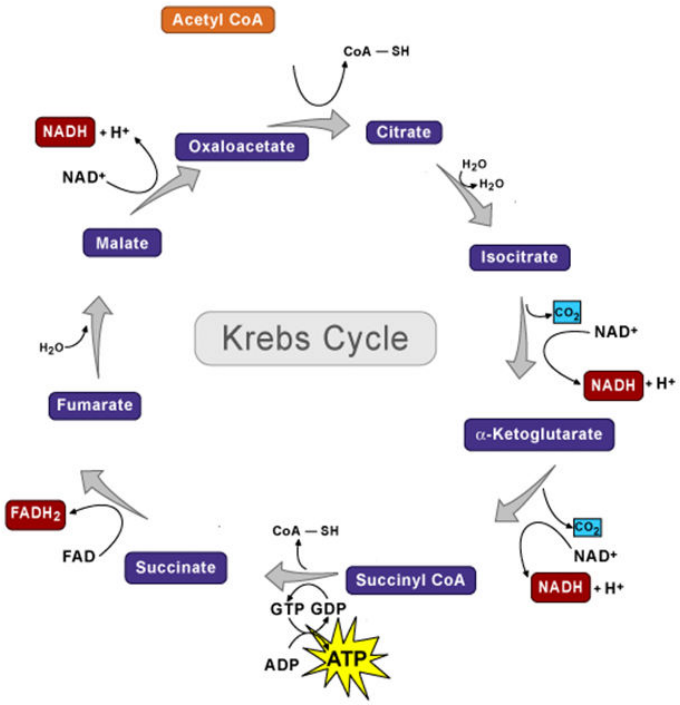


Figure 2.50 Schematic representation of the TCA cycle.

We therefore investigated the binding interaction of the TCA cycle intermediates (Figure 2.51) with psBBOX using the ^{1}H NMR-based dual-reporter displacement assay. The results are summarised in Figure 2.52. The data suggest that only fumaric acid and citric acid appear to bind to psBBOX, competing weakly with only 2OG (Figure 2.52a) without influencing the binding of GBB (Figure 2.52b). Previous studies have shown that fumarate is also a binder of PHD2, competing weakly with 2OG.[207] As fumarate does not have the α-ketoacid acid functionality typically involved in active site metal chelation, the only possible modes of its metal chelation may be either *via* the two carboxylate groups or the carboxylate and double bound.

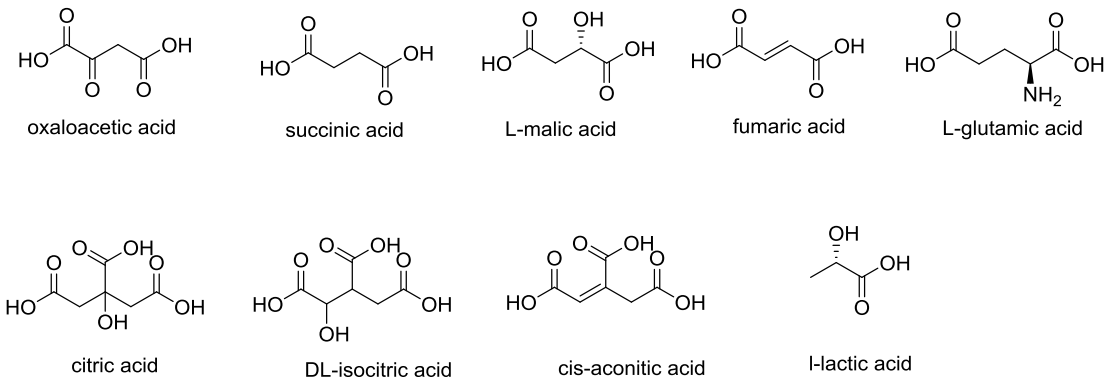
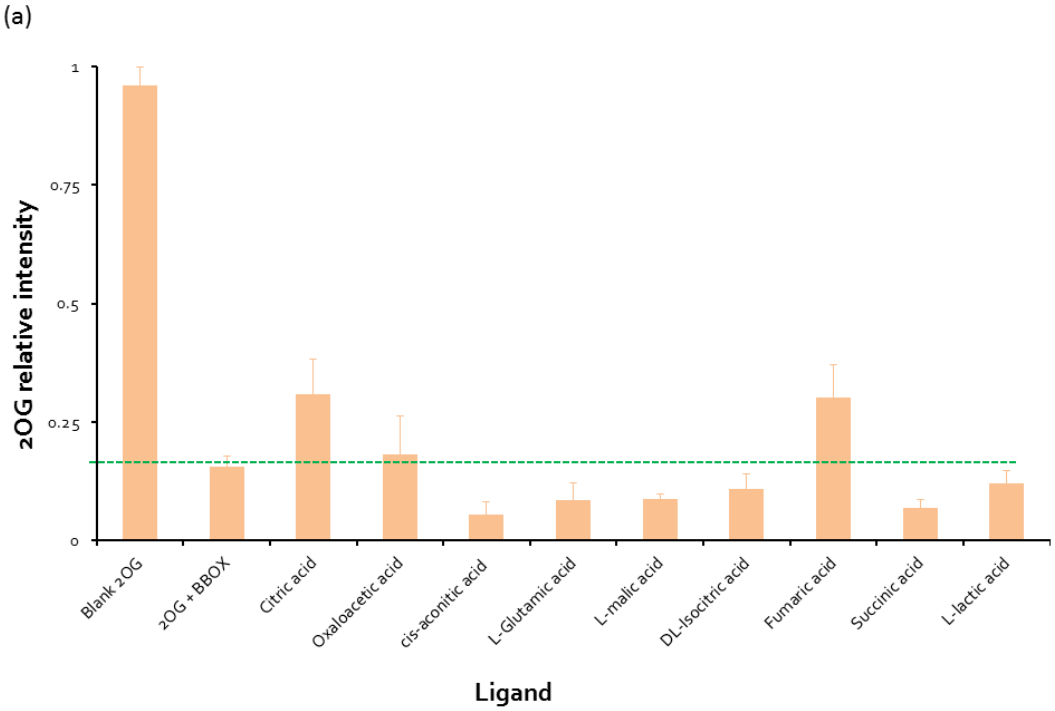


Figure 2.51 Structures of the TCA cycle intermediates.



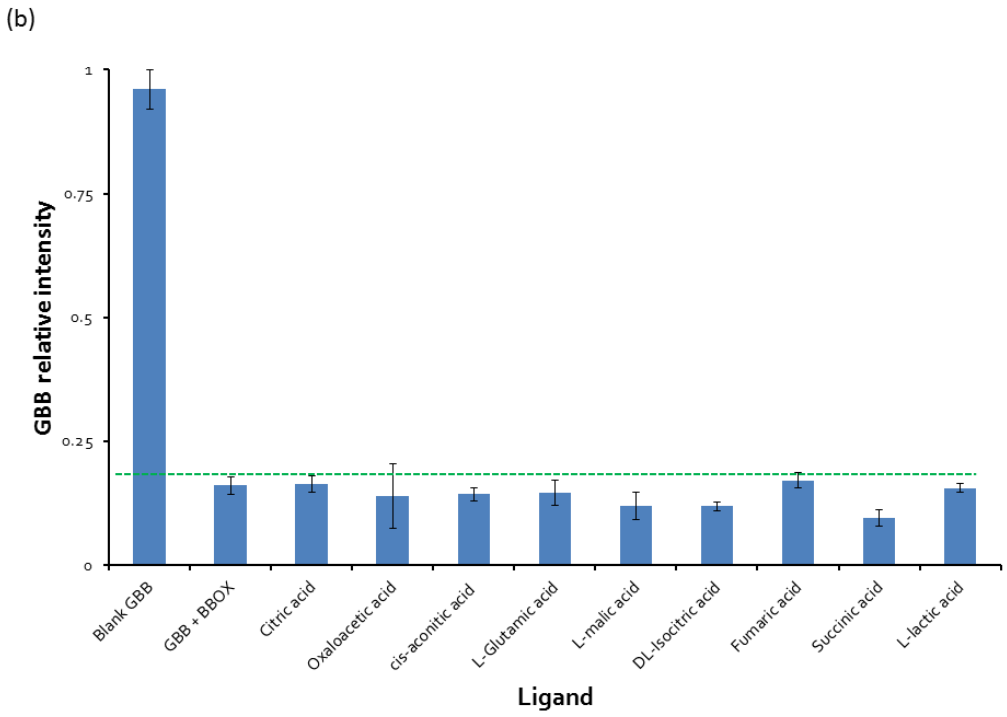


Figure 2.52 Investigating the binding of TCA cycle intermediates with psBBOX by CPMG edited ^{1}H NMR dual-reporter displacement assay.
Chart showing the effect of addition of TCA cycle intermediate on the signal intensity of 2OG (a) and GBB (b) In the ^{1}H NMR dual-reporter displacement assay for BBOX. The threshold lines represent the relative intensity of the reporters in the presence of BBOX. In the presence of competitive ligand, the reporter intensity above this line represents the level of their displacement, whereas their intensity below this line represents further improvements in their binding. The final assay mixture contains 25 μM GBB, 300 μM 2OG, 150 μM Mn^{II}, 15 μM BBOX, 50 μM F_3CCOCH_3 (internal control), 80 mM KCl and 100 μM screening ligand in 50 mM Tris-D11 buffer, pH 7.5, in D_2O. The error bars represent standard deviation from 3 separate measurements.

Citrate has been previously reported as a moderate inhibitor of psBBOX.[126] Its mode of metal chelation may perhaps be *via* the 1,4 decarboxylate groups, based on the structure of the molecule. It is possible that the binding of both fumarate and citrate to psBBOX may have potential role in regulating cellular carnitine levels. Hence further investigations along this line are proposed.

2.7 The role of cation-π interactions in psBBOX substrate recognition

Many protein systems contain an aromatic cage (or aromatic box) as part of the recognition module for specific binding of natural substrates. Aromatic cages typically consist of 2–4 aromatic amino acids (Trp, Tyr, Phe), and can be found either on the protein surface or buried inside the protein.[208-210] Pioneering work by Dougherty and co-workers has shown that aromatic cages recognise the positively-charged quaternary ammonium species *via* favourable cation–π interactions.[211-213] For example, cation–π interactions are involved in

associations of Cys-loop receptors and G-protein coupled receptors with various neurotransmitters, including acetylcholine, serotonin, dopamine, epinephrine, and histamine.[214] More recently, structural analyses of protein–peptide interactions on proteins involved in epigenetic gene regulation processes suggest that reader domain proteins that recognise dimethylarginine or trimethyllysine-containing histone tails interact *via* cation–π interactions.[210] Since their establishment and appearance in the literature in 1980s, cation–π interactions have been highlighted along with the hydrophobic effect, hydrogen bonding and ion pairing as the dominant non-covalent forces in protein–ligand associations.[215]

Analysis of the crystal structures of hBBOX reveals that the positively charged trimethylammonium group of GBB is located inside an aromatic cage formed by residues Tyr177, Tyr-194, Trp-181, Tyr-205 and Tyr-366 (Figure 2.53) The GBB carboxylate group also forms interactions with the side chains of Asn-292 (2.9 Å), Asn-191 (2.9 Å) and the backbone amide of Tyr-205 (2.9 Å) (Figure 2.53).

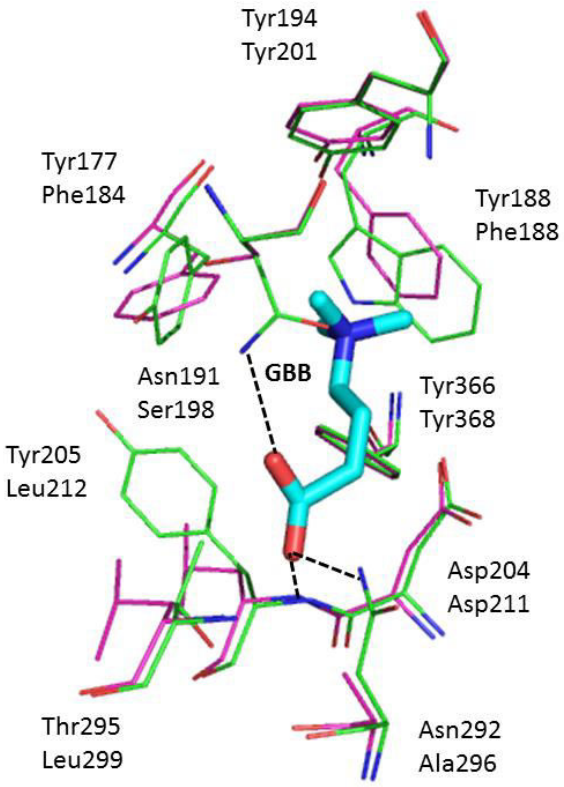


Figure 2.53 A view of the GBB binding site in the crystal structure of hBBOX (PDB ID: 3O2G) overlayed with a generated model of psBBOX.
hBBOX: Green sticks, upper residues numbers, psBBOX: pink sticks, lower residue numbers. GBB is shown as (cyan stick).

Previous studies on both recombinant human and psBBOX AK1 suggests that GBB analogues which lacks the carboxylate – OH group are not accepted as substrates by BBOX, suggesting that the interactions forming by the – OH group are crucial for the GBB binding. A radio-activity based assay has shown that a GBB analogue, in which the trimethylammonium group is replaced by a tertiary butyl group is a poor substrate of hBBOX.[165] However, in a recent study, it has also been shown that GBB analogue with trimethylammonium group replaced with an isopropyl group is not a hBBOX substrate.[163] Compared to the crystallographic studies, in solution physiological conditions it is still unknown whether it is the cation-π interactions between the GBB trimethylammonium group and BBOX aromatic cage residues or it is the hydrophobic interactions between the GBB trimethyl groups and aromatic cage residues that hold the GBB trimethylammonium group inside the aromatic cage in the BBOX active site.

Hence to investigate in more details, the role of cation-π interaction in BBOX substrate recognition, we replaced the positively-charged trimethylammonium group of GBB by three analogues functionalities: i) positively-charged trimethylphosphonium group; **I**) positively-charged trimethylarsonium group **II**; and uncharged tert-butyl group **III** (Figure 2.54) and tested their activity towards psBBOX for C-3 hydroxylation. Our aim was to comprehensively examine the binding and catalysis of these phospha, arsa and carba GBB analogues with BBOX using rigorous NMR and LC-MS analyses.[216]

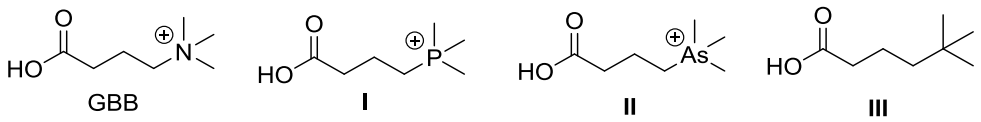


Figure 2.54 Structures of the phospha, arsa and carba analogues of GBB tested for C3 hydroxylation by psBBOX.

Both the phospha and arsa analogues possess a fixed positive charge, but are slightly larger than the GBB i.e atomic radius of P is 1.28 Å, As is 1.39 Å and N is 0.92 Å. The neutral carba analogue has virtually the same size (atomic radius of C is 0.91 Å) and shape as GBB, but lacks the presence of the positive charge.

Our analysis by ^{1}H NMR experiments revealed that the phospha (**I**) (Figure 2.55) and arsa (**II**) (Figure 2.56) GBB analogues, both of which possess positive charge on the central atom in the trimethylammonium moiety, are hydroxylated by psBBOX at C-3 position, whereas the carba analogue (**III**) which lacks positive charge is not hydroxylated by psBBOX (Figure

2.57), even with 10 folds (10 µM) higher psBBOX concentration as compared with phospha and carba ligands (1 µM psBBOX).

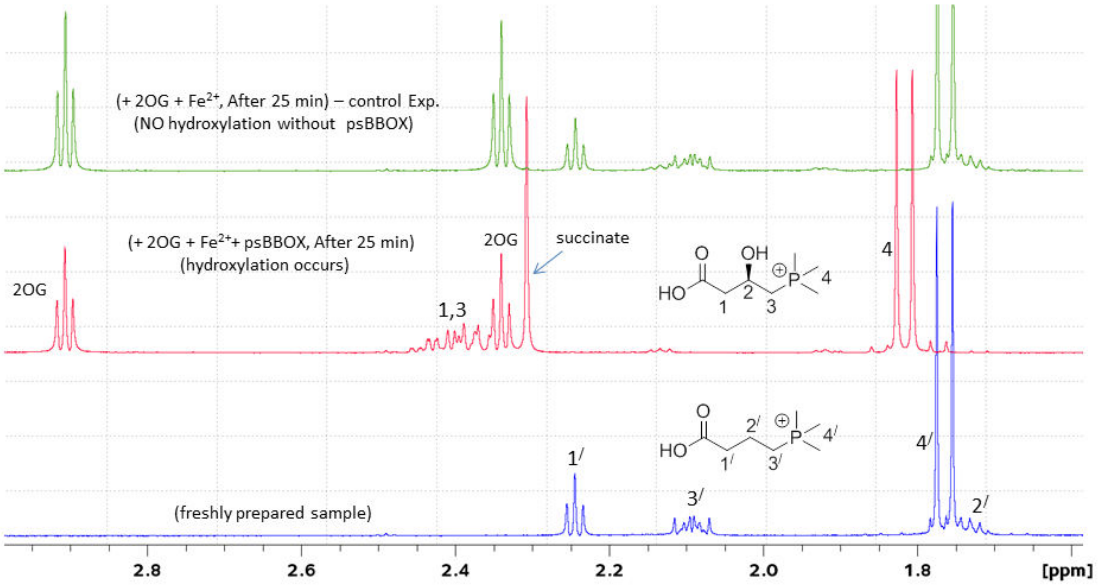


Figure 2.55 psBBOX catalysed C-3 hydroxylation of the phospha GBB analogue (**I**) as demonstrated by ^{1}H NMR.
Bottom spectrum shows the phospha ligand (**I**) on its own in buffer solution; middle spectrum shows the psBBOX catalysed formation of C-3 hydroxylated product in the presence of 2OG and Fe^{II}; top spectrum shows the control experiment without psBBOX. The assay mixture contains 200 µM phospha ligand (**I**), 600 µM 2OG, 100 µM Fe^{II} and 1 µM psBBOX in 50 mM Tris-D_{11}, pH 7.5 in D_2O.

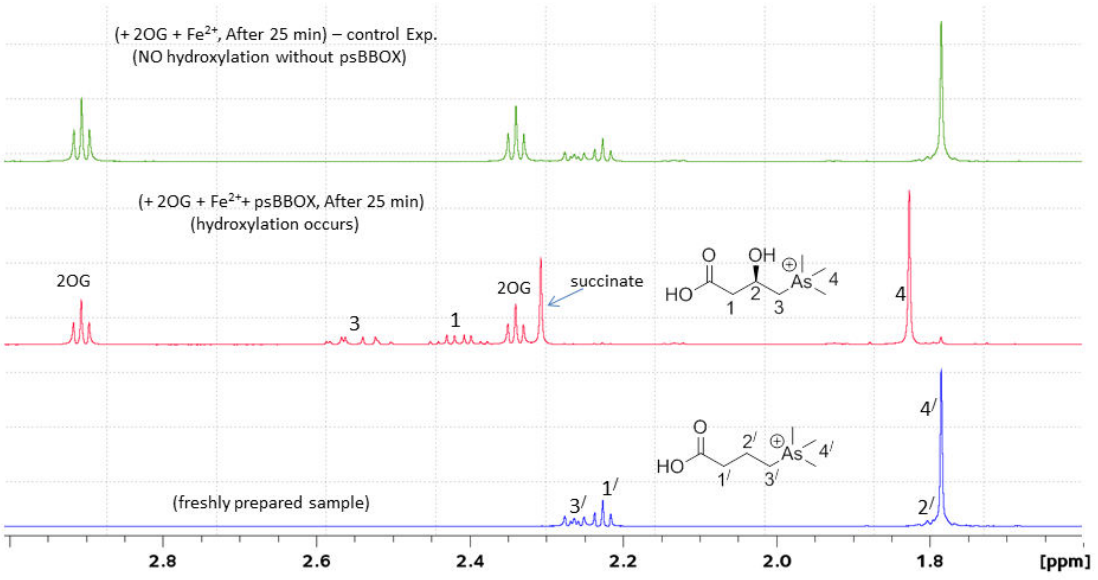


Figure 2.56 psBBOX catalysed C-3 hydroxylation of the arsa GBB analogue (**II**) as demonstrated by ^{1}H NMR.
Bottom spectrum shows the arsa ligand (**II**) on its own in buffer solution; middle spectrum shows the psBBOX catalysed formation of the hydroxylated product in the presence of 2OG and Fe^{II}; top spectrum shows the control experiment without psBBOX. The assay mixture contains 200 µM arsa ligand (**II**), 600 µM 2OG, 100 µM Fe^{II} and 1 µM psBBOX in 50 mM Tris-D_{11}, pH 7.5 in D_2O.

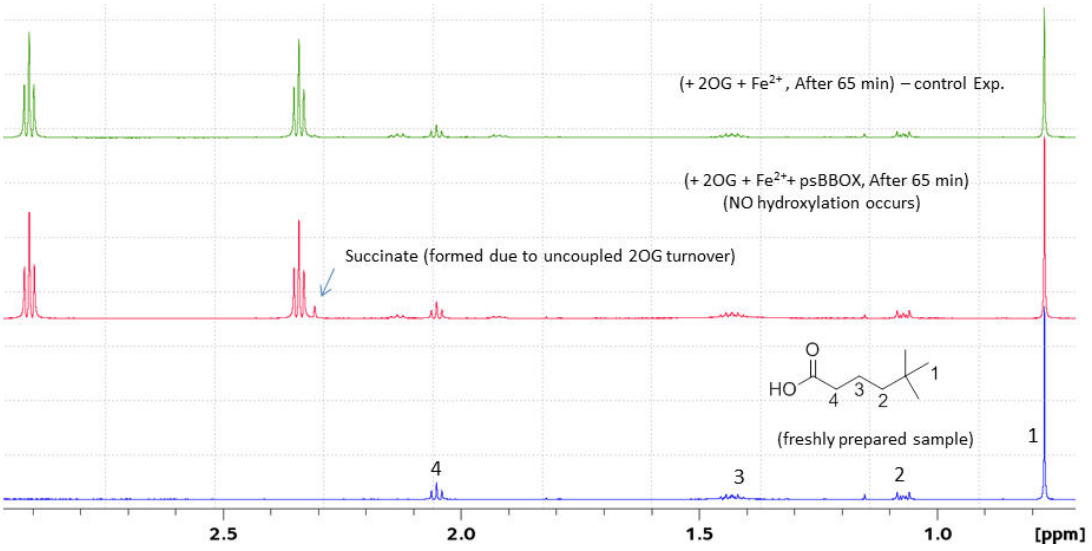


Figure 2.57 The carba GBB analogue (**III**) does not undergo hydroxylation by psBBOX as demonstrated by ^{1}H NMR. Bottom spectrum shows the carba ligand (**III**) on its own in buffer solution; middle spectrum shows the mixture of carba ligand (**III**), 2OG, Fe^{II} and psBBOX in solution; top spectrum shows the control experiment without psBBOX. The assay mixture contains 200 µM carba ligand (**III**), 600 µM 2OG, 100 µM Fe^{II} and 10 µM psBBOX in 50 mM Tris-D_{11}, pH 7.5 in D_2O.

Similarly, our investigation by LC-MS assays also demonstrated that the phospha (**I**) and arsa (**II**) GBB analogues undergo C-3 hydroxylation by psBBOX, whereas the carba analogue (**III**) does not (Figure 2.58) even with 10 folds high psBBOX as compared to the phospha (**I**) and arsa (**II**) ligands.

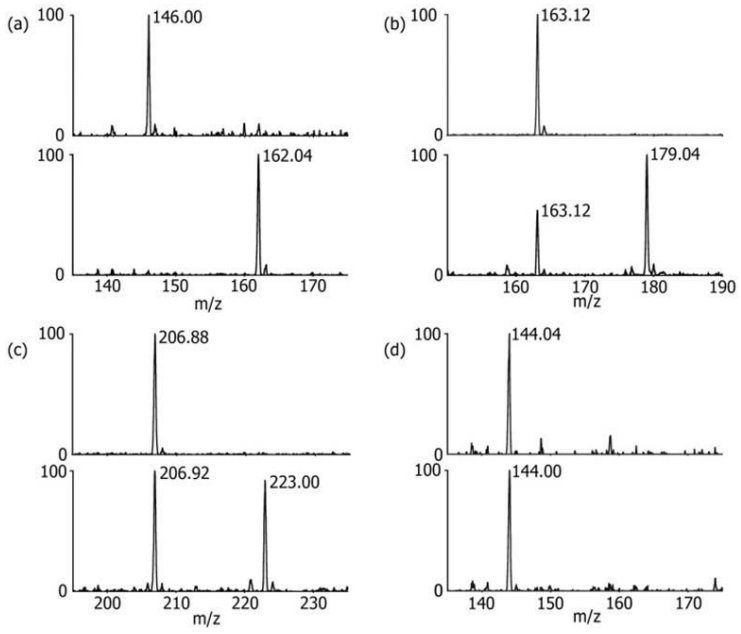


Figure 2.58 LCMS analysis of the psBBOX-catalysed hydroxylation of (a) GBB (b) phospha (**I**) (c) arsa (**II**) and (d) carba (**III**) ligands.
Top panel = starting substrate, bottom panel = psBBOX-catalysed reaction (a-c: 1 µM psBBOX, 5 minutes; d: 10 µM psBBOX, 3 hours).

The acceptance of phospha (**I**) and arsa (**II**) ligands by psBBOX for C-3 hydroxylation can be observed from the addition of 15.99 i.e atomic mass (Da) of oxygen to the molecular weights of phospha (**I**) (163.12 Da, Figure 2.58b) and arsa (**II**) (206.88 Da, Figure 2.58c) in the presence of psBBOX, whereas no such addition is observed in case of carba ligand (**III**) (144.04 Da, Figure 2.58d).

Control experiments were also carried out to demonstrate that in the absence of psBBOX, the incubation of the phospha (**I**) and arsa (**II**) ligands in their respective assay solutions do not undergo hydroxylation (Figure 2.59).

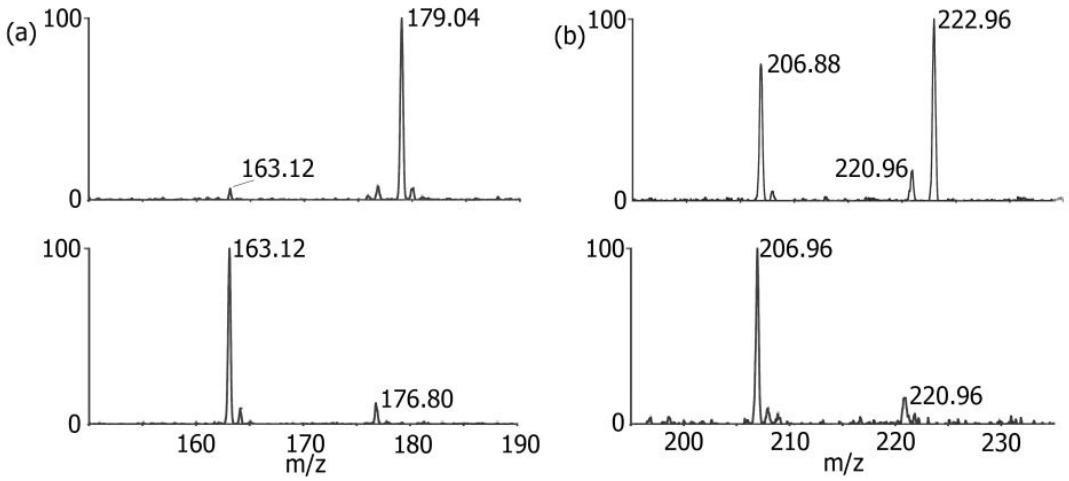


Figure 2.59 LCMS control experiments to prove the hydroxylation of phospha (**I**) and arsa (**II**) ligands occurs by psBBOX catalysis.
a) hydroxylation of the phospha ligand (**I**) in the presence of psBBOX (top) and in the absence of psBBOX (bottom); b) hydroxylation of arsa ligand (**II**) in the presence (top) and in the absence of psBBOX (bottom). A typical reaction was carried out with the following conditions psBBOX (2 µM), 2OG (1.5 mM), ascorbate (5 mM), $FeSO_4$ (50 µM) and substrate ligand (50 µM) in Tris buffer pH 7.5. Peak at 176.80 Da is from ascorbic acid, whereas, peak at 220.96 Da is from a contaminant in the LCMS instrument.

To demonstrate the requirement of molecular oxygen (as co-substrate) for psBBOX catalysed oxidation of phospha (**I**) and arsa ligands (**II**), we incubated their assay mixtures under (near) anaerobic conditions (under argon atmosphere), and only traces (< 5%) of hydroxylated products were observed (Figure 2.60).

In addition, the standard enzymatic reaction of phospha (**I**) and arsa (**II**) ligands in $H_2{}^{18}O$ (Tris-buffered saline, pH 7.5) afforded hydroxylated products that have the same molecular weight as products from the standard reaction in Tris-buffered $H_2{}^{16}O$ (Figure 2.61), demonstrating that the oxygen atom in the hydroxylated product is introduced by catalysis

and not from the solvent. Collectively, all these results confirm that phospha (**I**) and arsa (**I**) ligands undergo psBBOX-catalysed hydroxylation and that the oxygen atom in the hydroxylated products is derived from oxygen from the air, and not from oxygen from water, as also observed in the labelling studies on most other 2OG oxygenases.[22]

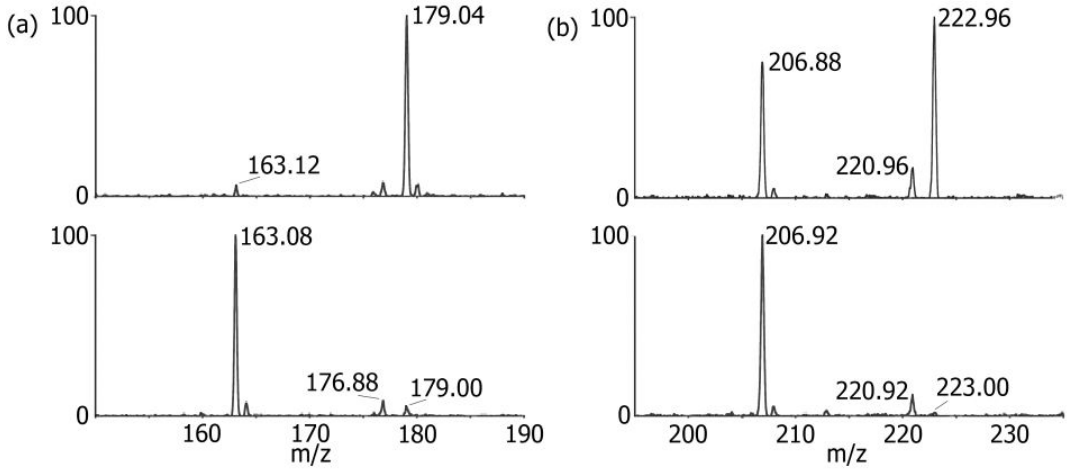


Figure 2.60 LCMS control experiment to prove that in the psBBOX catalysed hydroxylation of phospha (**I**) and arsa (**II**) ligands, the oxygen is required in the catalysis.
(a) psBBOX-catalysed hydroxylation of phospha ligand (**I**) under standard conditions (top) and under argon atmosphere (bottom); b) psBBOX-catalysed hydroxylation of arsa ligand (**II**) under standard conditions (top) and under argon atmosphere (bottom). A typical reaction with psBBOX (2 μM), 2OG (1.5 mM), ascorbate (5 mM), $FeSO_4$ (50 μM) and substrate ligand (50 μM) in Tris buffer, pH 7.5. Peaks at 176.88 Da and 220.96 Da derive from sodium ascorabate.

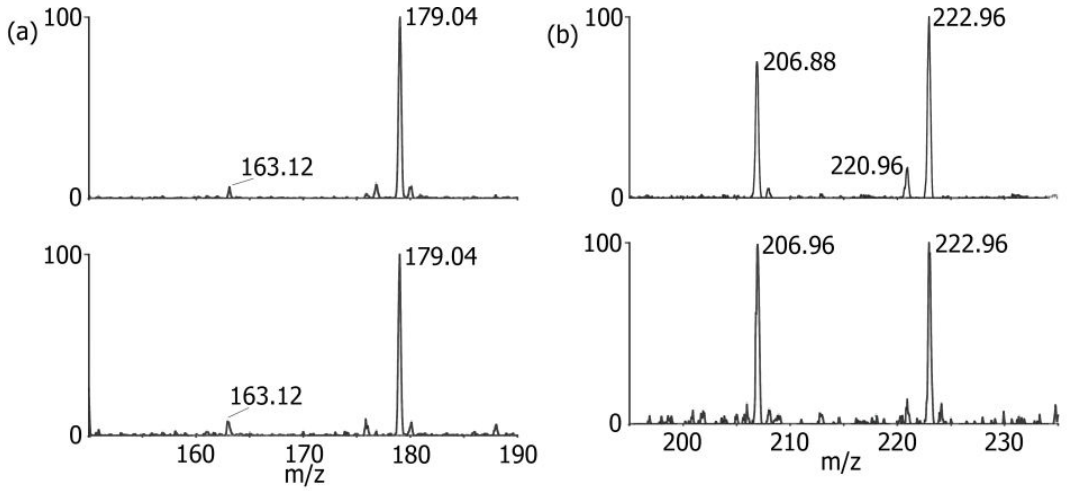


Figure 2.61 LCMS control experiment to prove that in the in the psBBOX catalysed hydroxylation of phospha (**I**) and arsa (**II**) ligands, the oxygen comes from O_2 and not from solvent.
a) psBBOX-catalysed hydroxylation of phospha ligand (**I**) under standard conditions in $H_2{}^{16}O$ (top) and in $H_2{}^{18}O$ (bottom); b) psBBOX-catalysed hydroxylation of arsa ligand (**II**) under standard conditions in $H_2{}^{16}O$ (top) and in $H_2{}^{18}O$ (bottom).

2.7.1 Characterisation of the phospha and arsa ligands C-3 hydroxylation products by 2D NMR techniques

After we proved that the phospha (**I**) and arsa (**II**) ligands are hydroxylated by psBBOX, we then used the 2D NMR techniques of COSY and HSQC to characterise these hydroxylation products. The C-3 hydroxylation product of the phospha ligand (**I**) was characterised by recording the following ^{1}H NMR spectrum (Figure 2.62), COSY spectrum (Figure 2.63), and HSQC spectrum (Figure 2.64) of the assay mixture.

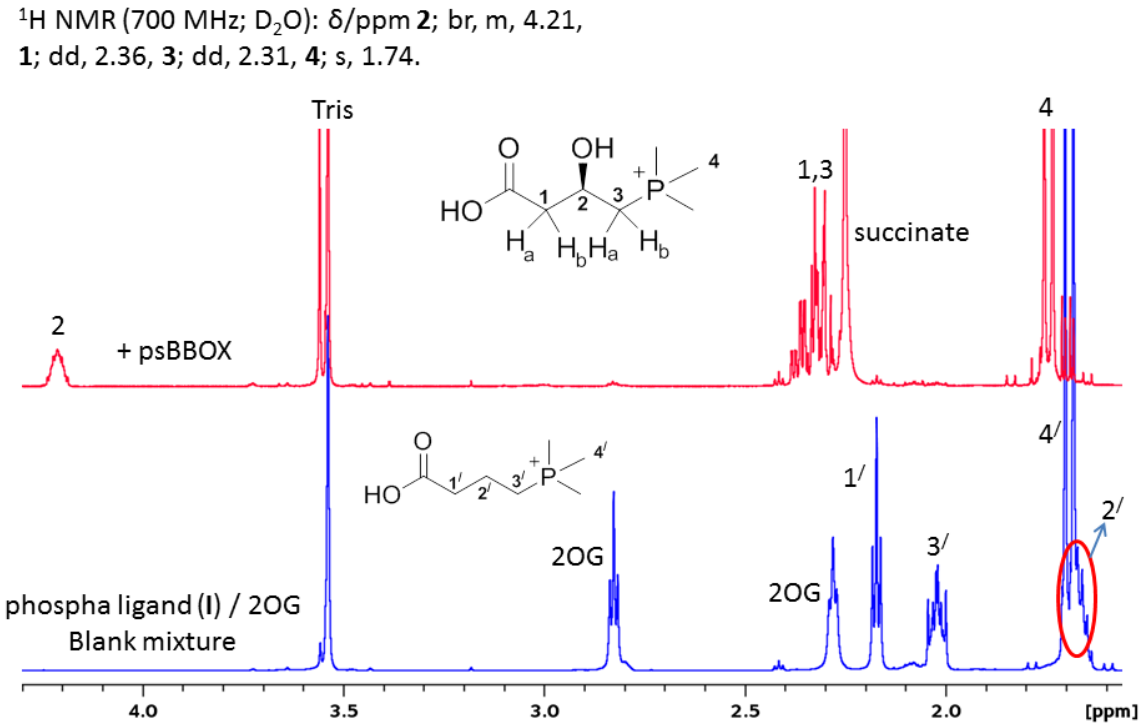


Figure 2.62 ^{1}H NMR analysis of the assay mixture of psBBOX catalysed hydroxylation of phospha ligand (**I**).
Bottom spectrum is the assay mixture without psBBOX and the top spectrum is the assay mixture incubated with psBBOX. The assay mixture contains 1 mM phospha substrate ligand (**I**), 1 mM 2OG, 80 mM KCl, 10 µM psBBOX and 100 µM Fe^{II} in 50 mM Tris-D11, pH 7.5 in D_2O and incubated for 25 min.

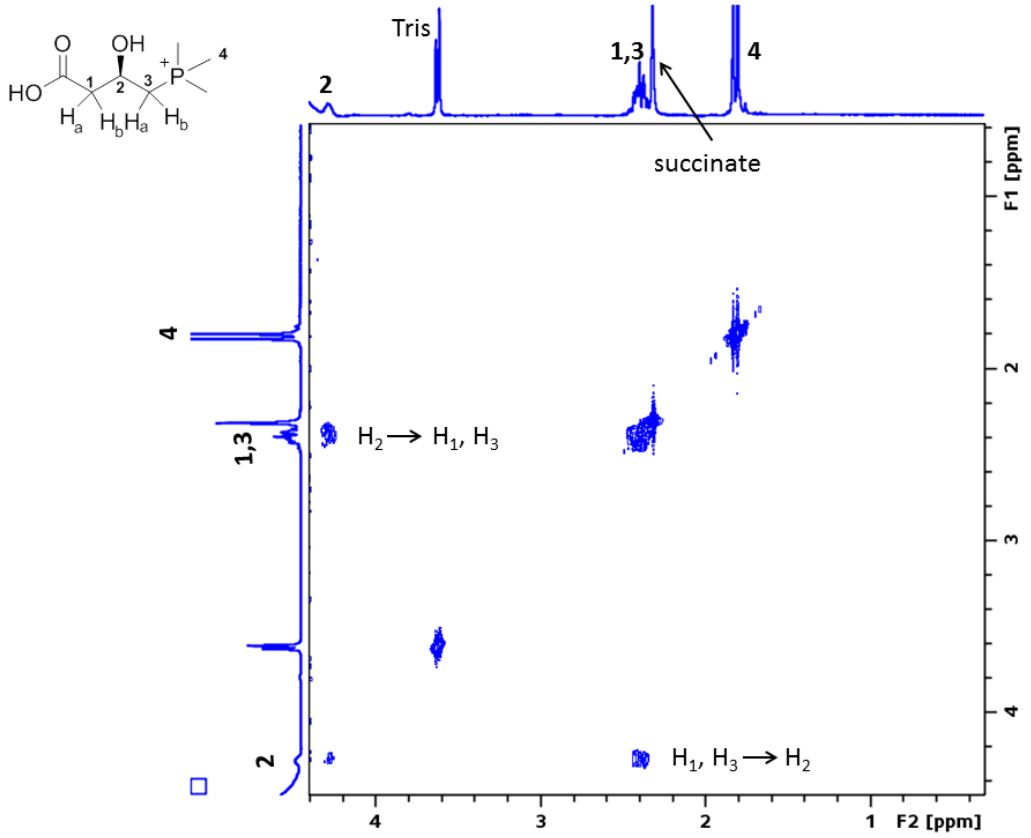
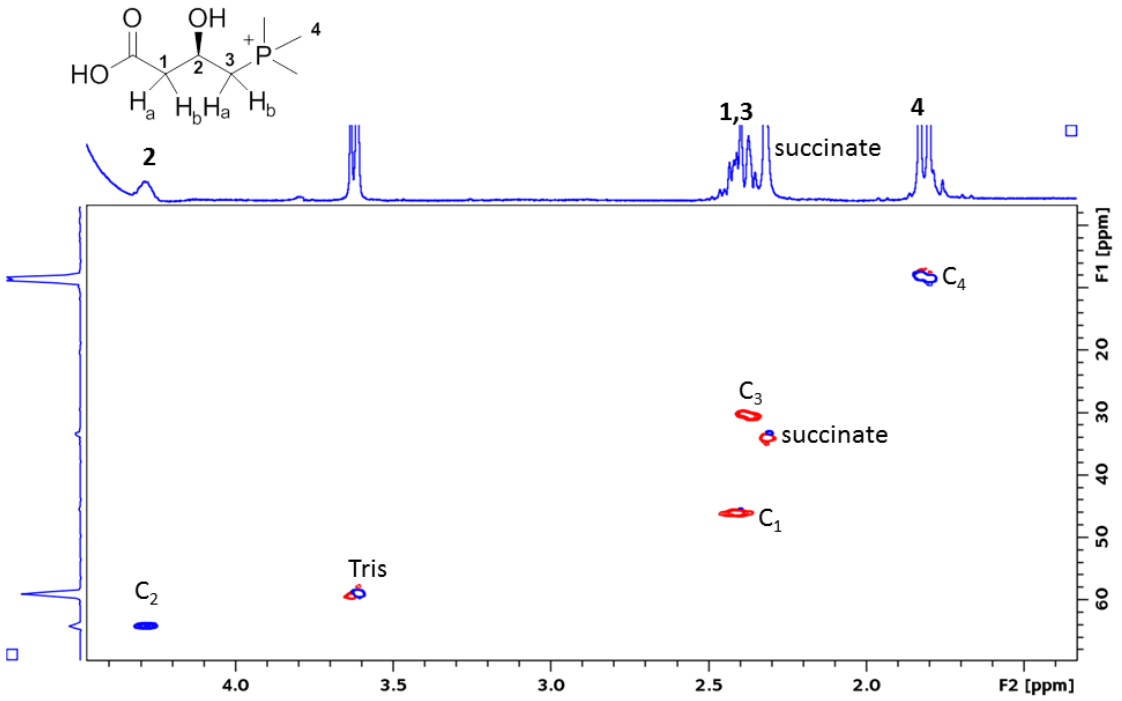


Figure 2.63 COSY NMR spectrum of the assay mixture of psBBOX catalysed hydroxylation of phospha ligand (**I**).



^{13}C NMR (500 MHz; D_2O): δ/ppm **2**; 64.1, **1**; 46.1, **3**; 30.4, **4**; 8.2.

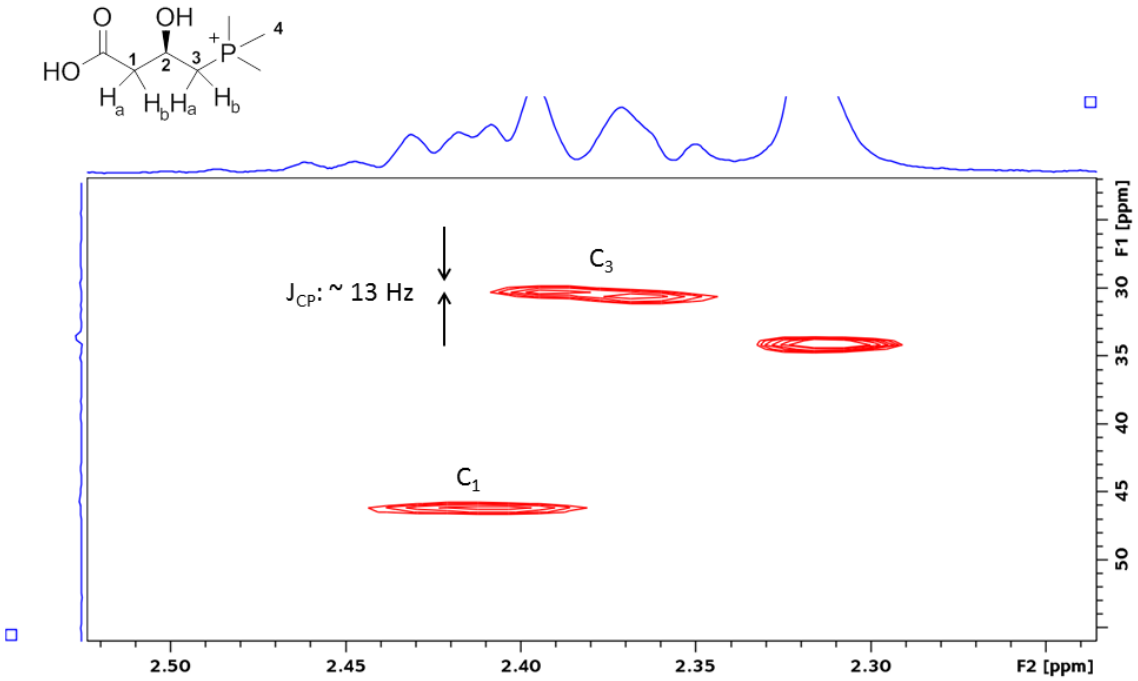


Figure 2.64 ^{1}H-^{13}C HSQC NMR spectrum of the assay mixture of psBBOX catalysed hydroxylation of phospha ligand (**I**). Bottom: A zoomed in view highlighting the presence of phosphorus coupling with C-3 in the hydroxylated product (structure shown at the top of the spectrum).

The assignments of this psBBOX catalysed formation of phospha C-3 hydroxylated product was also confirmed by the addition of synthetically synthesized (stereoselectively) (3*R*)-hydroxylated phospha product to the assay mixture of psBBOX catalysed hydroxylation of phospha ligand (**I**). It was shown that the addition of synthetic (3*R*)-hydroxylated phospha product resulted an increase (spiking) in the signals intensities of psBBOX catalysed phospha hydroxylated product in ^{1}H NMR spectrum (Figure 2.65).

The C-3 hydroxylation product of the arsa ligand (**II**) was characterised by recording the following ^{1}H NMR spectrum (Figure 2.66) COSY spectrum (Figure 2.67) and HSQC spectrum (Figure 2.68) of the assay mixture.

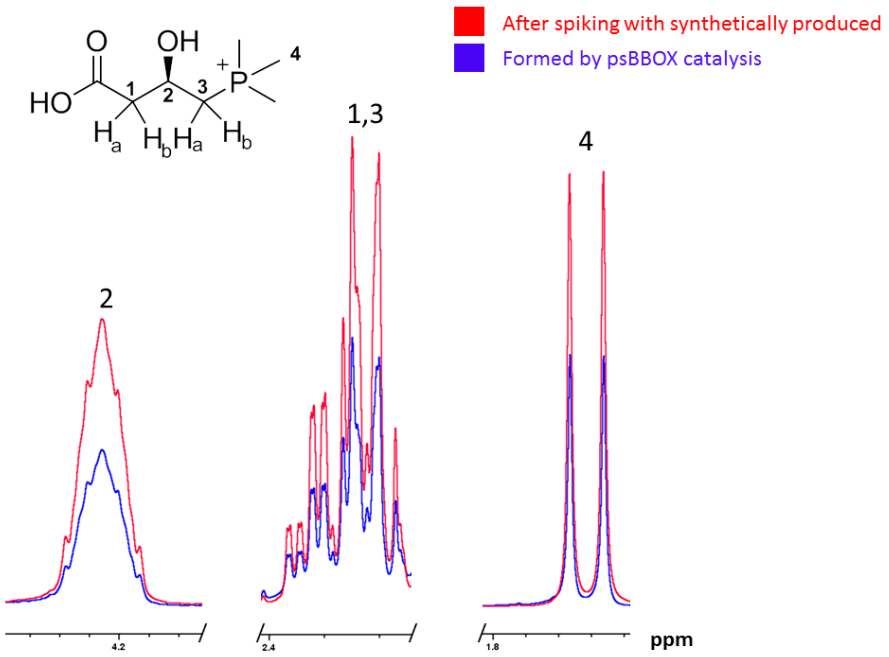


Figure 2.65 Conformation of NMR assignments of phospha hydroxylated product using spiking technique.
^{1}H NMR spectra (part shown for clarity) showing the spiking of the psBBOX catalysed phospha hydroxylated product (blue) with the same compound produced synthetically (red). The psBBOX catalysed hydroxylation of phospha ligand (**I**) was performed with an incubation of the assay mixture containing 1 mM phospha ligand (**I**), 1 mM 2OG, 80 mM KCl, 10 µM psBBOX and 100 µM Fe^{II} in 50 mM Tris-D_{11}, pH 7.5 in D_2O for 25 min. To this assay mixture 1 mM synthetically produced phospha hydroxylated product was added and a ^{1}H NMR spectra was recorded.

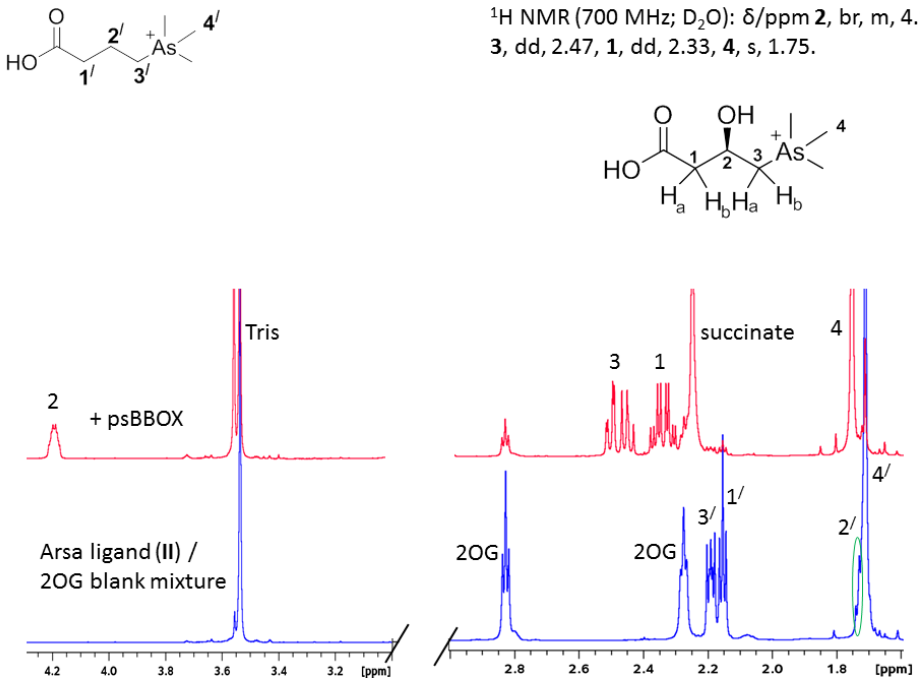


Figure 2.66 ^{1}H NMR spectrum of the assay mixture of psBBOX catalysed hydroxylation of arsa ligand (**II**).
Bottom spectrum is the assay mixture without psBBOX and the top spectrum is the assay mixture incubated with psBBOX. The assay mixture contains 1 mM arsa ligand (**II**), 1 mM 2OG, 80 mM KCl, 10 µM psBBOX and 100 µM Fe^{II} in 50 mM Tris-D_{11}, pH 7.5 in D_2O and incubated for 25 min.

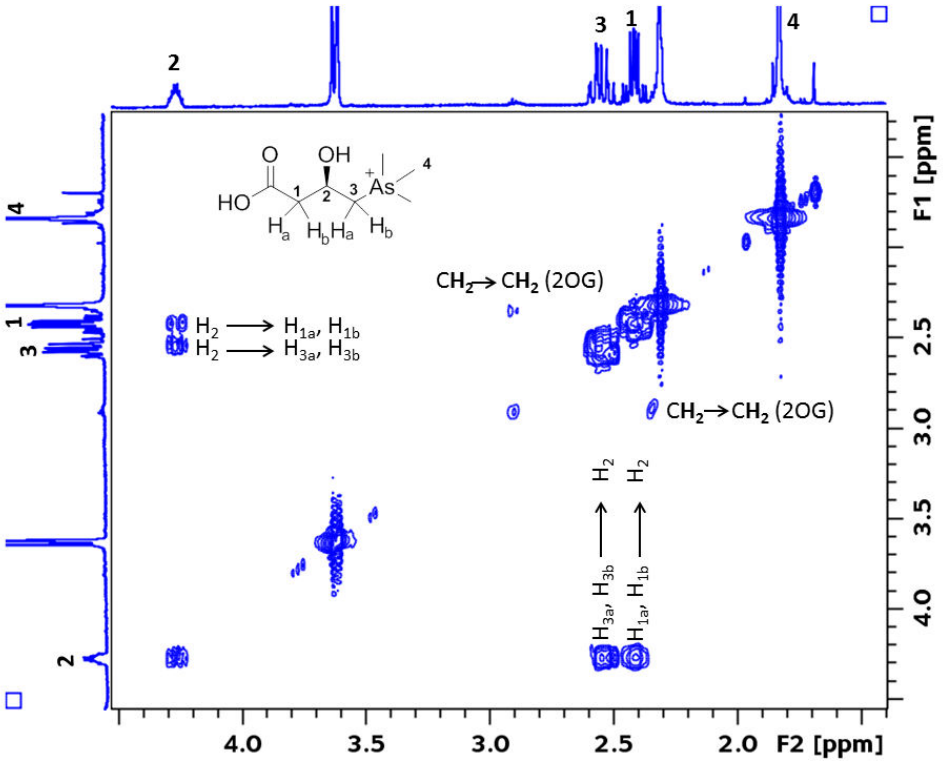


Figure 2.67 COSY NMR spectrum of the assay mixture of psBBOX catalysed hydroxylation of arsa ligand (**II**).

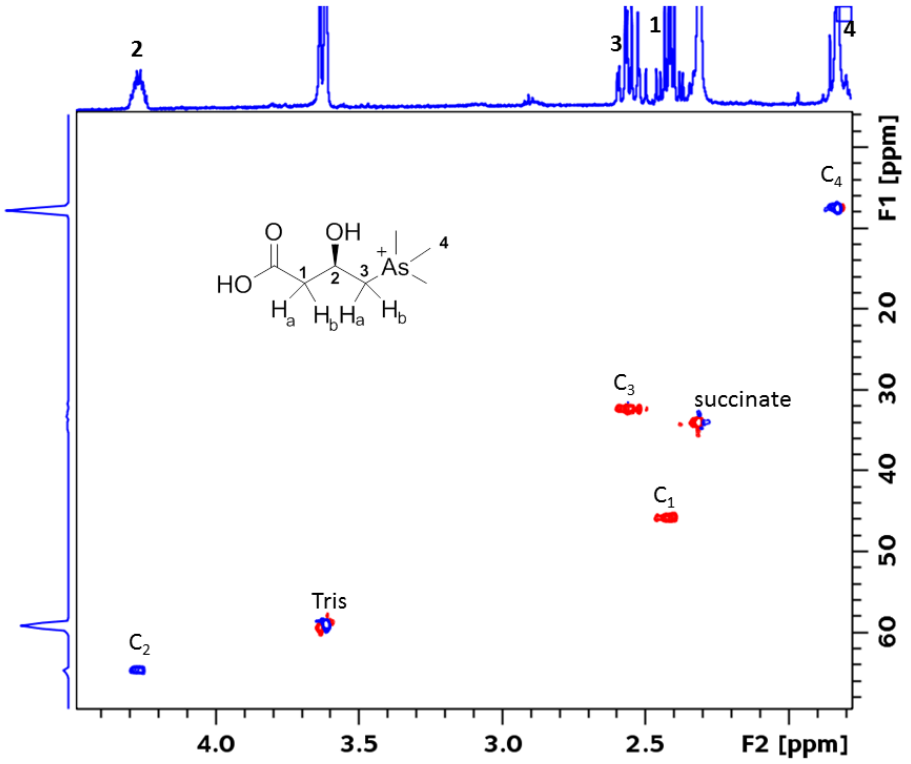


Figure 2.68 ^{1}H-^{13}C HSQC NMR spectrum of the assay mixture of psBBOX catalysed hydroxylation of arsa ligand (**II**).

Similarly, in this case also, synthetically produced (3*R*)-hydroxylated arsa product was added to the assay mixture of psBBOX catalysed hydroxylation of arsa ligand (**II**). It was shown that upon the addition of synthetic (3*R*)-hydroxylated arsa product, the signals intensities of the psBBOX catalysed arsa hydroxylated product were increased (spiking), (Figure 2.69) further confirming that the NMR assigments of the arsa hydroxylated product were correct.

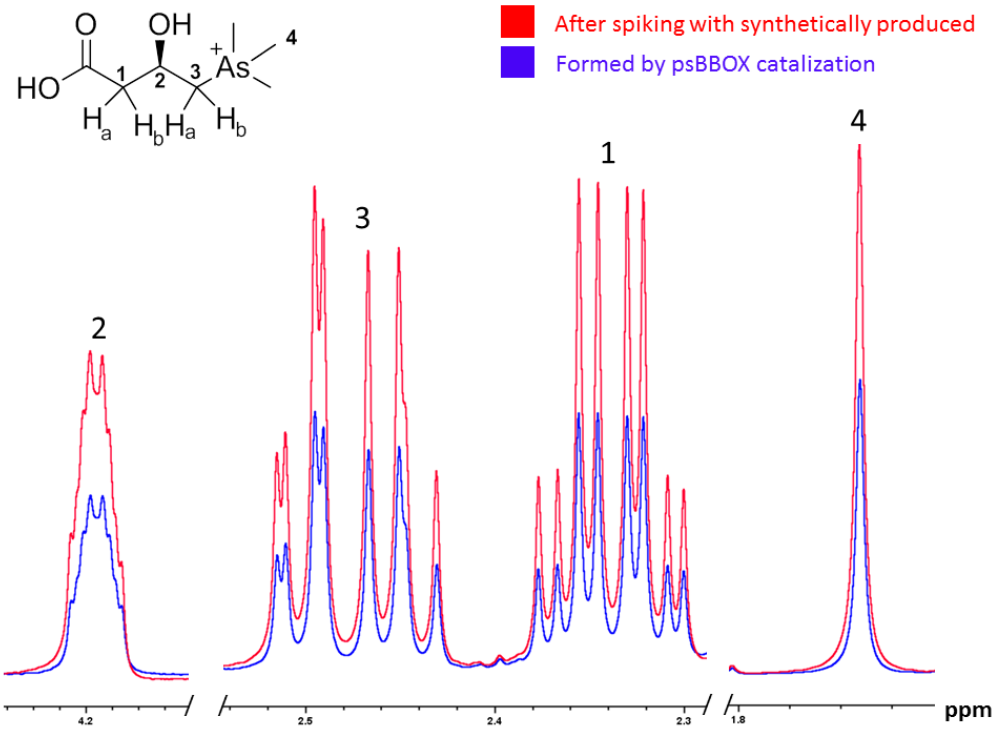


Figure 2.69 Conformation of NMR assignments of arsa hydroxylated product using spiking technique.
^{1}H NMR spectra (part shown for clarity) showing the spiking of the psBBOX catalysed arsa hydroxylated product (blue) with the same compound produced synthetically (red). The psBBOX catalysed hydroxylation of arsa ligand (**II**) was performed with an incubation of the assay mixture containing 1 mM arsa ligand (**II**), 1 mM 2OG, 80 mM KCl, 10 µM psBBOX and 100 µM Fe^{II} in 50 mM Tris-D_{11}, pH 7.5 in D_2O for 25 min. To this assay mixture 1 mM synthetically produced arsa hydroxylated product was added and a ^{1}H NMR spectra was recorded.

2.7.2 Comparison of the activities of GBB, phospha and arsa psBBOX substrates

Using LC-MS, the comparative enzyme kinetics analyses for the three positively-charged substrates (GBB, phospha (**I**) and arsa (**II**) ligands) was performed in order to quantitatively determine their ability to act as psBBOX substrates (Figure 2.70). In the presence of 1 µM of psBBOX, NMe_3^+ was almost instantly (within 5 minutes) fully hydroxylated. PMe_3^+ was also observed to be a very good substrate yielding 72% of the corresponding hydroxylated product, but less efficiently than NMe_3^+. Under the same conditions, $AsMe_3^+$ was found to be

the poorest of the three positively-charged substrates; only 39% of the hydroxylated product was formed.

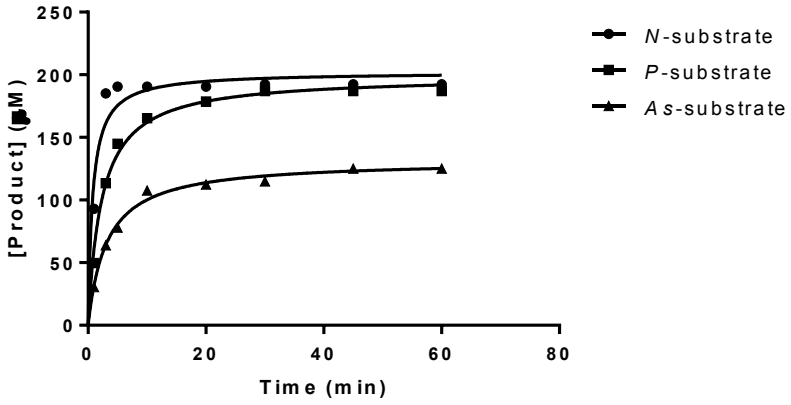
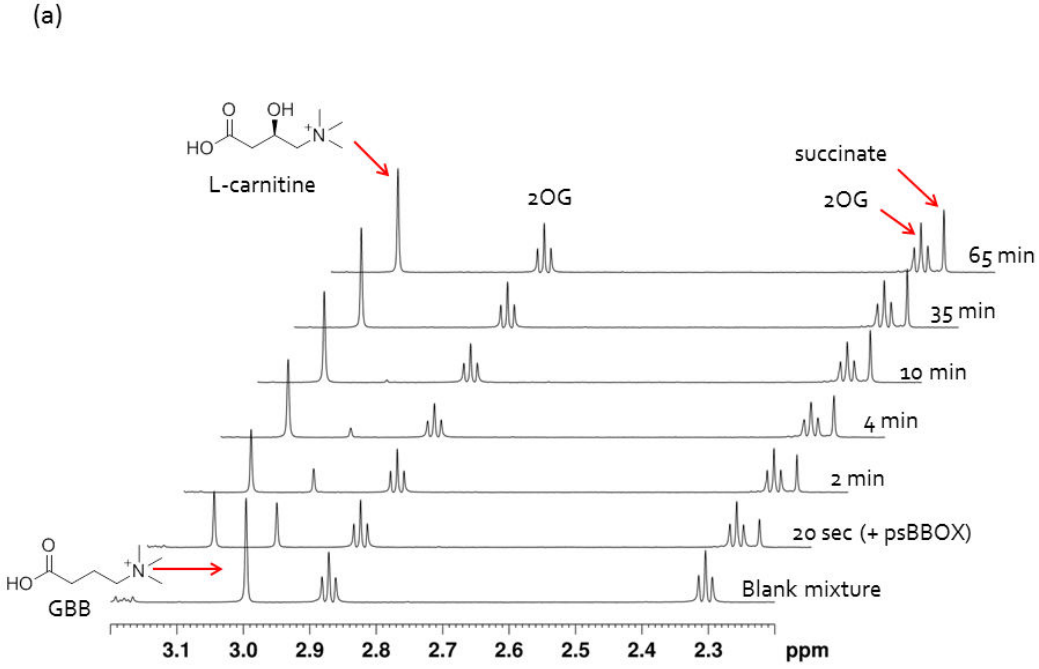
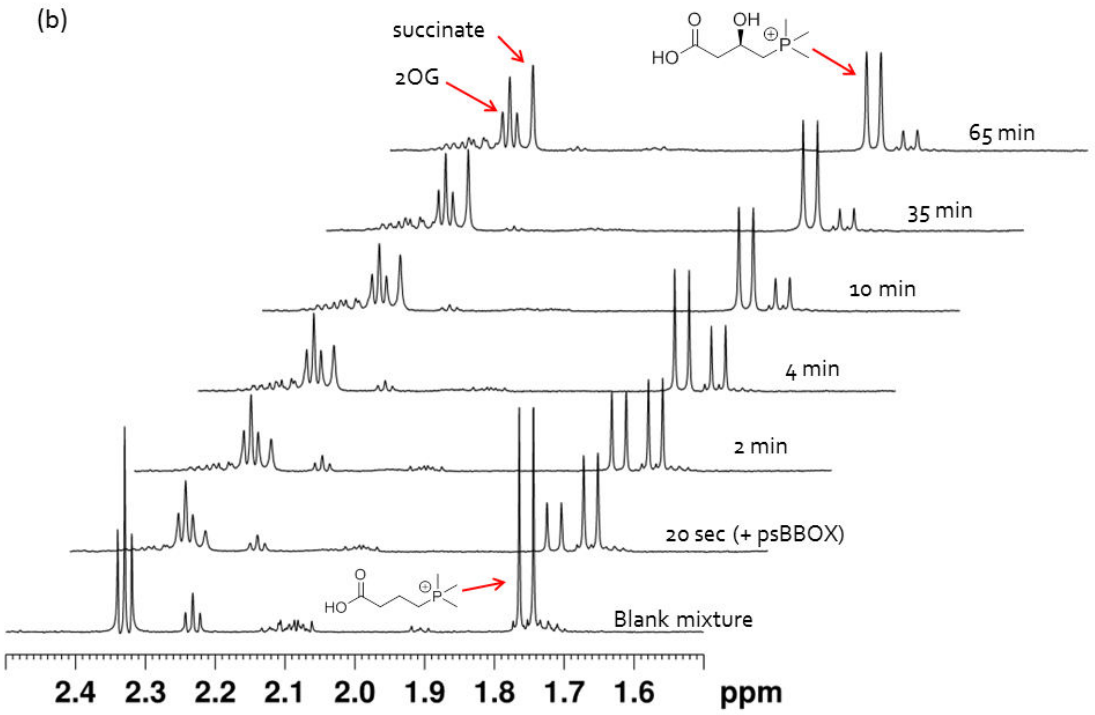
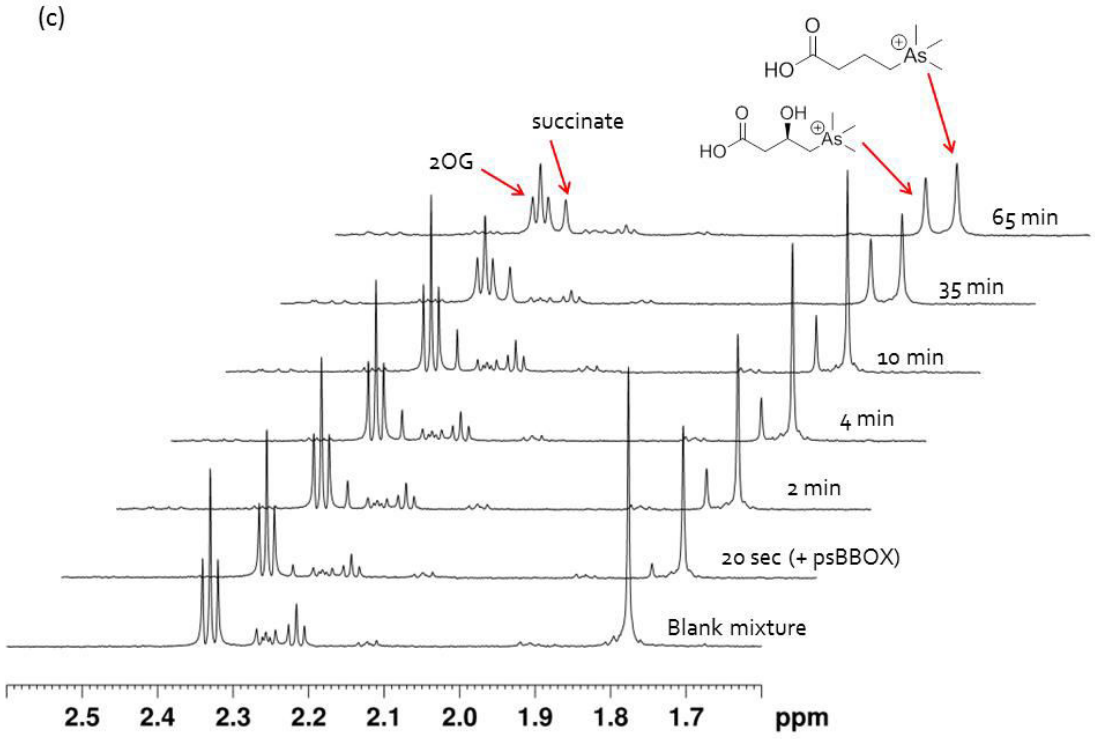


Figure 2.70 Comparison of the activities of GBB, phospha (**I**) and arsa (**II**) ligands using psBBOX kinetics studies performed on LCMS.
Time-course analyses of the psBBOX-catalysed hydroxylation of GBB (circle), *P*-substrate (**I**) (square) and *As*-substrate (triangle). Conditions used in these experiments are: psBBOX (1 μM), 2OG (1.5 mM), ascorbate (5 mM), $FeSO_4$ (50 μM) and substrate (50 μM) in a Tris (20 mM), NaCl (200 mM) buffer at pH 7.5.

A similar trend of activity (GBB > phospha > arsa) of these substrates was also observed by ^{1}H NMR in the presence of 575 nM psBBOX (Figure 2.71).




(b)
succinate
2OG
65 min
35 min
10 min
4 min
2 min
20 sec (+ psBBOX)
Blank mixture
2.4 2.3 2.2 2.1 2.0 1.9 1.8 1.7 1.6 ppm


(c)
succinate
2OG
65 min
35 min
10 min
4 min
2 min
20 sec (+ psBBOX)
Blank mixture
2.5 2.4 2.3 2.2 2.1 2.0 1.9 1.8 1.7 ppm

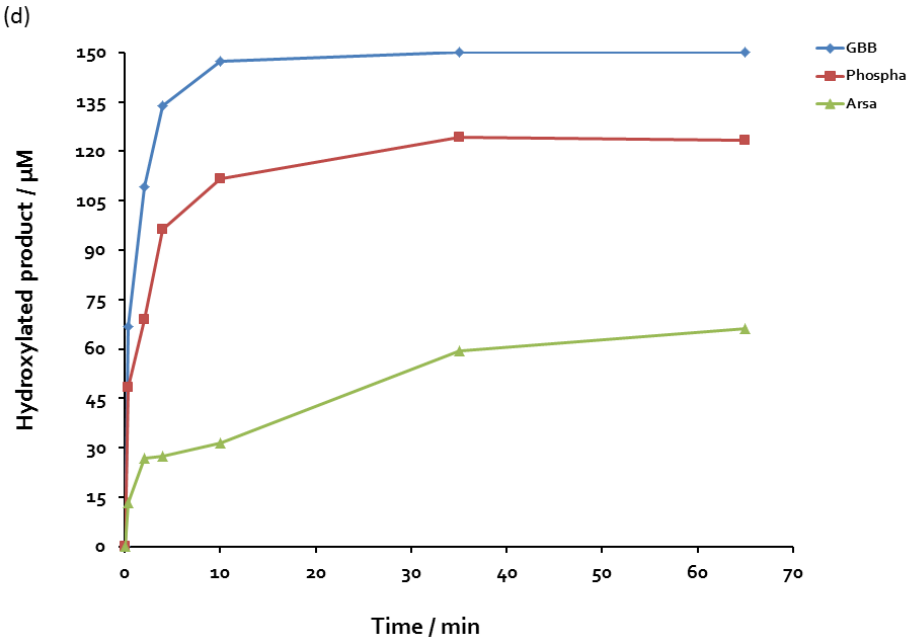


Figure 2.71 Comparing the activities of GBB, phospha (**I**) and arsa (**II**) ligands using psBBOX kinetics studies by ^{1}H NMR. psBBOX catalysed C-3 hydroxylation of (a): GBB, (b): phospha ligand (**I**) and (c): arsa ligand (**II**). The data shows that GBB is a better substrate than phospha (**II**), which in turn is a better substrate than arsa (**II**) ligand for psBBOX. (d): Plot showing the level of hydroxylation of all the three substrates as a function of time. The assay mixture contains 150 µM substrate ligand, 600 µM 2OG, 100 µM Fe^{II}, 575 nM psBBOX, 80 mM KCl in 50 mM Tris-D_{11}, pH 7.5 in D_2O. The reaction was performed in a 1.5 mL eppendorf tube and quenched by with the addition of 5 µL of 1M HCl.

LC-MS-based 'Michaelis-Menten' analyses were performed at 400 nM psBBOX with varying concentrations of substrates from 50 µM to 1.5 mM (Figure 2.72); reactions were quenched with acetonitrile after 1 minute at 23 °C. Maximal velocity (V_{max}) for the conversion of GBB into L-carnitine was significantly higher than V_{max} for the conversion of its phospha (**I**) and arsa (**II**) analogues into their corresponding hydroxylated products (2.81 vs. 1.65 vs. 1.04 µM s^{-1},Table 2.5); consequently, values of the rate constant (k_{cat}) showed that k_{cat} for GBB is about 1.7 and 2.7-fold higher than k_{cat} for its phospha and arsa analogues, respectively.

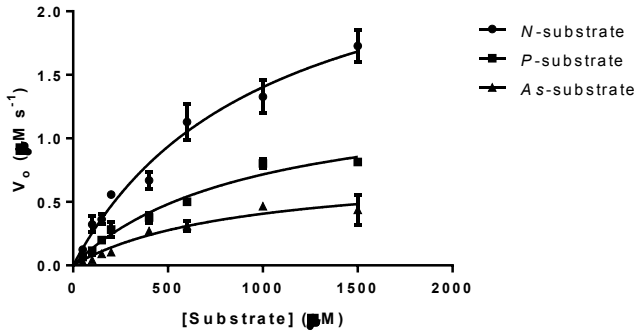


Figure 2.72 LC-MS-based 'Michaelis-Menten' analyses of GBB, phospha (**I**) and arsa ligands (**II**).
Conditions used in these experiments are: psBBOX (400 nM), 2OG (1.5 mM), ascorbate (5 mM), $FeSO_4$ (50 µM) and varying concentrations of substrate in a Tris (20 mM), NaCl (200 mM) buffer at pH 7.5. Reactions were quenched with acetonitrile (80% v/v final concentration) after 1 minute incubation at 23 °C. The error bars represents standard deviation from three separate measurements.

Table 2.5 Kinetic parameters for the conversion of GBB and its phospha (**I**) and arsa (**II**) analogues by psBBOX.

	GBB	Phospha	Arsa
V_{max} ($\mu M\ s^{-1}$)	2.81 ± 0.31	1.65 ± 0.16	1.04 ± 0.13
k_{cat} (s^{-1})	7.02 ± 0.77	4.13 ± 0.40	2.61 ± 0.31
K_M (μM)	1001 ± 205	1200 ± 202	1366 ± 275
k_{cat}/K_M ($s^{-1}\ M^{-1}$)	$7.01 \cdot 10^3$	$3.44 \cdot 10^3$	$1.91 \cdot 10^3$

We speculate that this observable decrease in V_{max} and k_{cat} values for phospha and arsa ligands relative to GBB, may be at least in part, due to slight alterations in the positioning of phospha and arsa substrates in the active site of psBBOX, which make the slower rate of the hydroxylation, as has been observed in substrate analogue studies with other 2OG oxygenases. It is possible that longer C-P and C-As bond lengths (1.87 Å and 1.98 Å, respectively) relative to C-N bond length (1.47 Å) forces the phospha and arsa substrates to adopt a slightly non-optimal binding modes, with the unactivated CH_2 groups further away from the highly oxidising Fe^{IV}=O enzymatic intermediate. Values of the K_M for all three positively-charged substrates (GBB 1001, phospha, 1200, arsa 1366 µM) were also observed to follow the order (GBB < phospha < arsa), implying that all three substrates are likely to be binding with psBBOX with affinities in close range in the order GBB > phospha > arsa. Thus, calculated values of enzyme efficiency (k_{cat}/K_M) clearly show that substrates follow the activity trend as: GBB > phospha > arsa.

2.7.3 Binding studies of phospha, arsa and carba ligands to psBBOX by ^{1}H NMR direct ligand-observation

Although both the qualitative ^{1}H NMR and quantitative LCMS-based kinetics studies proved that the positive charge in the trimethylcation XMe_3^+ moiety of the ligand is critical for psBBOX catalysis, the small difference in V_{max}, k_{cat} and K_M values and the complex mechanism of 2OG oxygenases, motivated us to study the binding interactions of phospha (**I**), arsa (**II**) and carba (**III**) ligands with psBBOX-Zn^{II}-2OG complex by CPMG ^{1}H NMR direct ligand observation by titrating BBOX. We obtained the binding dissociation constant (K_D) values of 7 µM and 17 µM for phospha ligand (**I**) (Figure 2.73) and arsa ligand (**II**) (Figure 2.74) respectively; the K_D of GBB was 5 µM (determined in section 2.4). These

results thus clearly demonstrate that GBB binds slightly better than the phospha ligand (**I**), which in turn binds better than the arsa ligand (**II**). These results are consistent with the trend in the V_{max}, k_{cat} and K_M values for these substrates as described above. Most importantly, the ^{1}H NMR-based binding studies also confirmed that due to the lack of positive charge on the central atom in the ligand tertiary butyl group, the carbon analogue (**III**) does not bind with psBBOX (Figure 2.75), hence consistent with the results of both ^{1}H NMR (Figure 2.57), and LC-MS (Figure 2.58d) based turnover assays showing that the carba analogue is not a psBBOX substrate.

Similar trends have also been reported in the inhibition of serine protease factor Xa, which has an aromatic cage in its S4 substrate residue binding pocket, which showed that the inhibitor containing a quaternary ammonium moiety has a similar potency as inhibitor containing analogues phosphonium group towards factor Xa, whereas, the carba analogue is a significantly less potent inhibitor (almost 60 fold less potent) .[217,218]

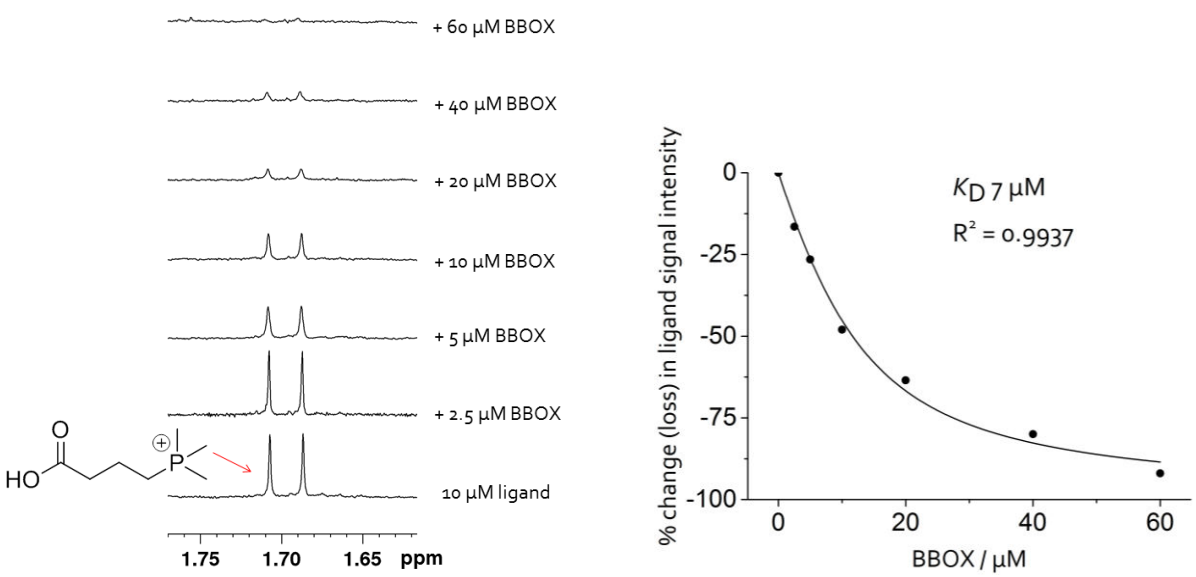


Figure 2.73 Measurements of the binding constant (K_D) of the phospha ligand (**I**) to psBBOX by CPMG edited ^{1}H NMR direct ligand observation.
The attenuation of the ligand Me_3 signal is measured as a function of BBOX concentration (left image). To the right is the fitting curve of the titration data points using Origin Pro. A K_D of 7 µM was obtained. The assay mixture contains 10 µM phospha ligand, 1 mM 2OG, 300 µM $ZnCl_2$ in 50 mM Tris-D_{11} buffer, pH 7.5 in 90% H_2O and 10% D_2O. The filter time used in CPMG experiments was 32 ms.

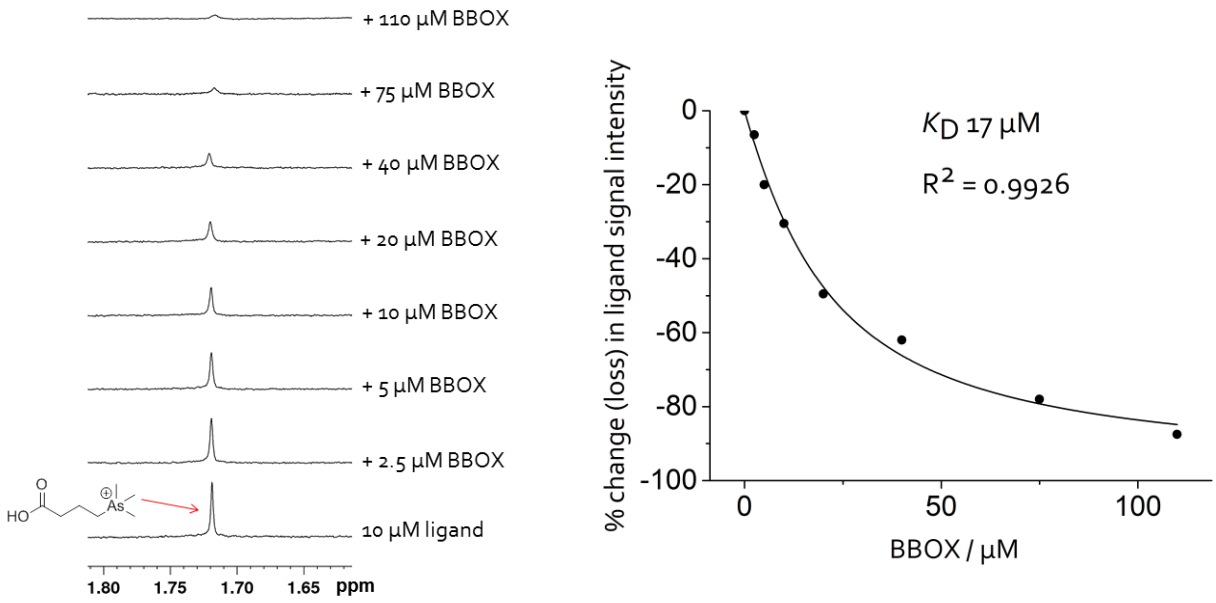


Figure 2.74 Measurements of the binding constant (K_D) of the arsa ligand (**II**) to psBBOX by CPMG edited ^{1}H NMR direct ligand observation.
The attenuation of the ligand Me_3 signal is measured as a function of BBOX concentration (left image). To the right is the fitting curve of the titration data points using Origin Pro. A K_D of 17 µM was obtained. The assay mixture contains 10 µM arsa ligand (**II**), 1 mM 2OG, 300 µM $ZnCl_2$ in 50 mM Tris-D_{11} buffer, pH 7.5 in 90% H_2O and 10% D_2O. The filter time used in CPMG experiments was 32 ms.

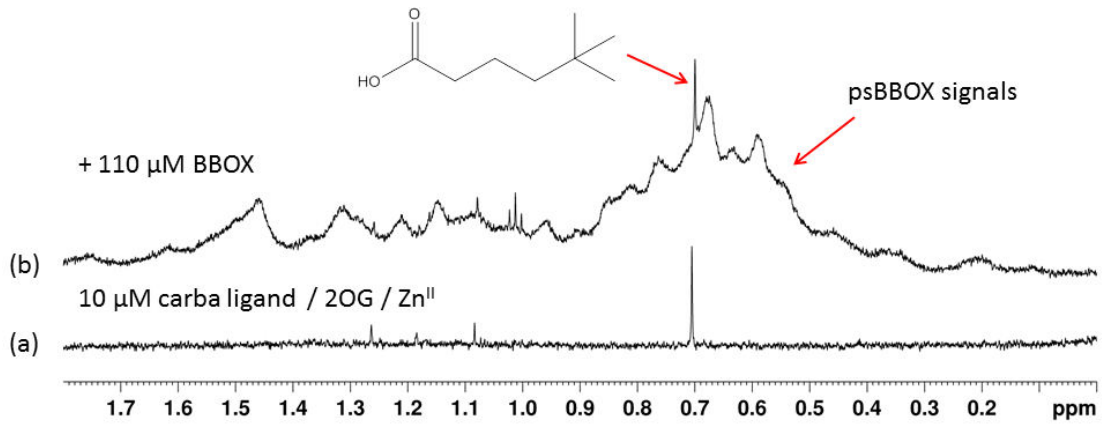


Figure 2.75 Monitoring the binding of carba ligand (**III**) with psBBOX by ^{1}H NMR direct ligand observation.
Monitoring the broadness/attenuation of ligand NMR signals (Me_3 signal at 0.7 ppm shown) in the presence of psBBOX (broad signals) as a probe of binding interaction. (a) Spectrum (part shown) of the carba ligand in the presence of co-factor Zn^{II} and co-substrate 2OG in the absence of psBBOX. (b) In the presence of psBBOX. From spectrum (b), it is evident that the ligand signals are not attenuated / broadened in the presence of psBBOX (with 1:11, ligand:BBOX concentration ratio), demonstrating that the carba ligand doesn't bind with psBBOX . The assay mixture contains 10 µM carbon analogue , 300 µM Zn^{II}, 1 mM 2OG, 110 µM psBBOX in 50 mM Tris-D_{11}, pH 7.5 in 90% H_2O: 10% D_2O.

2.7.4 Docking studies of phospha and arsa ligands

To investigate the binding mode of phospha (**I**) and arsa (**II**) ligands with BBOX, the crystal structure of hBBOX in complex with Zn^{II} and NOG (PDB ID: 3O2G) was used to make a homology model of psBBOX. All the three substrates i.e GBB and its phospha and arsa analogues were predicted to bind in the same way inside the psBBOX active site for stereoselective C-3 hydroxylation, with XMe_3^+ group located inside the aromatic cage (Figure 2.76).

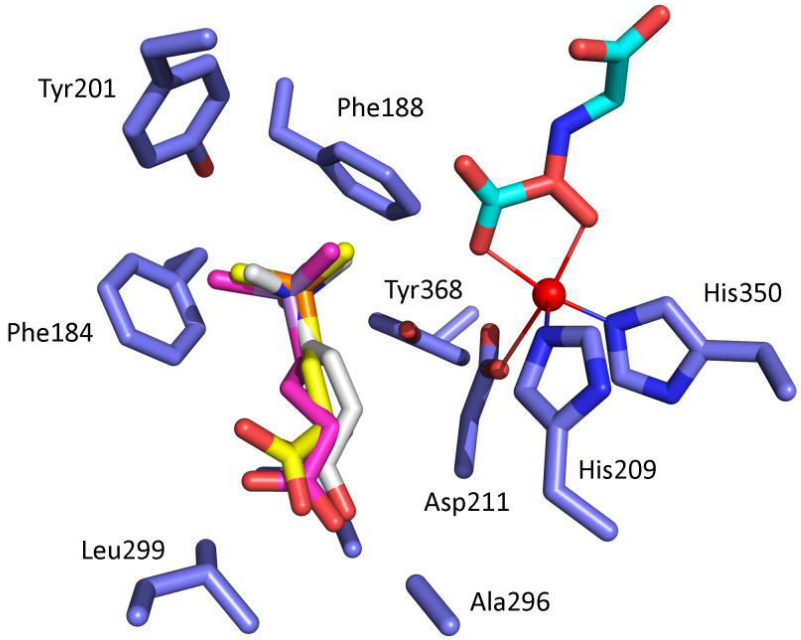


Figure 2.76 Docking studies of the binding of GBB, phospha (**I**) and arsa (**II**) ligands in psBBOX active site.
The model structure of psBBOX (purple) complexed with Zn^{II} (red), *N*-oxalylglycine (cyan) and substrates GBB (white), phospha (yellow) and arsa (magenta) ligands.

In addition, the CHELPG (Charges from Electrostatic Potentials using a Grid-based method) atomic charges of the central X atom and its surrounding 3 carbons and 9 hydrogen atoms for the docked conformations for GBB and its phospha, arsa analogues and the minimized energy conformation for the carba analogue (Table 2.6) were also calculated (performed by Hwanho Choi). Notably, the average partial charge of 9 hydrogen atoms of XMe_3^+ shows slightly incremental trend of +0.1506 (for arsa), +0.1596 (for phospha) and +0.1652 (for GBB). The calculated partial charge of hydrogen atoms in neutral tert-butyl group of carba analogue is +0.0884. These results are in agreement with trends in binding affinities as observed by NMR; the more positively-charged H atoms of XMe_3^+ results in stronger cation–π interactions with the aromatic cage of psBBOX.

Table 2.6 Calculated CHELPG charges (in e) of the X (As, P, N and C) and its surrounding atoms in the model system for the psBBOX-substrate complex.

Partial charge (CHELPG)	As	P	N (GBB)	C
X= As, P, N, C	1.083874	0.982078	0.346632	0.781339
C6	-0.62055	-0.62769	-0.44605	-0.52202
C7	-0.51859	-0.54974	-0.36109	-0.48489
C8	-0.55633	-0.49209	-0.33038	-0.46134
H15	0.124747	0.16353	0.168474	0.091036
H16	0.133143	0.148607	0.168991	0.091743
H17	0.146389	0.138545	0.185918	0.093877
H18	0.153665	0.136004	0.146934	0.093287
H19	0.127262	0.199521	0.157062	0.076114
H20	0.176385	0.184089	0.151968	0.099502
H21	0.13269	0.125163	0.258579	0.087137
H22	0.174563	0.157739	0.14222	0.070371
H23	0.187265	0.1833	0.106538	0.092407
H_average	0.150679	0.159611	0.165187	0.088386

2.7.5 Enantiopure (*R*) and (*S*) C-3 hydroxylated phospha and arsa ligands as psBBOX substrates

We also tested the synthesised enantiopure (3*R*)- and (3*S*)-hydroxylated phospha-and arsa ligands (Figure 2.77) as substrates of psBBOX using both LCMS and ^{1}H NMR-based turnover assays.

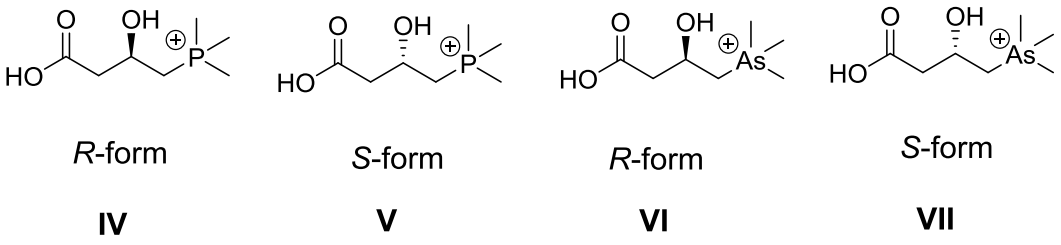
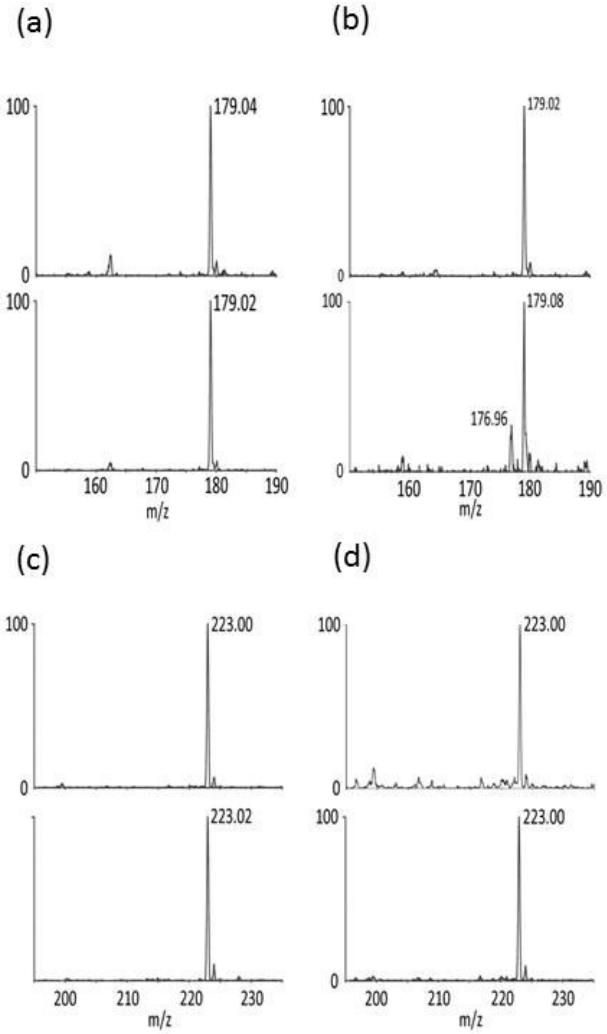


Figure 2.77 Synthesized enantiopure phospha and arsa analogues of carnitine tested as substrates of psBBOX.

Our LCMS based results showed that in the presence of 20 µM psBBOX, the (3*R*)-phospha hydroxylated ligand **IV** is not a substrate of psBBOX within the limits of detection (Figure 2.78a), whereas, the (3*S*)-phospha hydroxylated ligand **V** is a poor psBBOX substrate, forming a small amount of 3-keto product (mass 176.96 Da) (Figure 2.78b,e). We, however,

did not detect the formation of 3-keto product from (3*R*) **VI** (Figure 2.78c) and (3*S*) **VII** arsa hydroxylated ligands (Figure 2.78d).



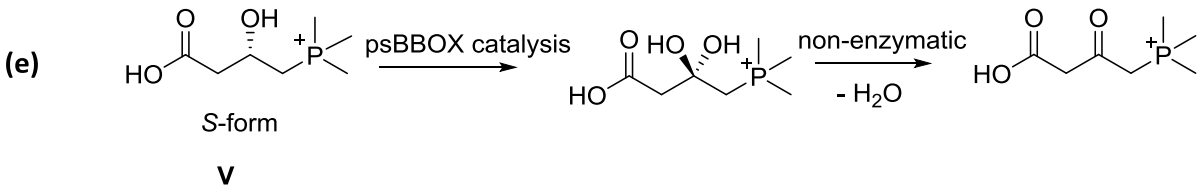


Figure 2.78 LCMS analysis of the psBBOX catalysed hydroxylation of ligands **IV** to **VII**.
Investigation of enantiopure (a) (3*R*) hydroxylated phospha **IV** and (b) (3*S*) hydroxylated phospha **V** (c) (3*R*) hydroxylated arsa **VI** and (d) (3*S*) hydroxylated arsa **VII** as psBBOX substrates using LCMS. Top panel = starting substrate, bottom panel = after the addition of psBBOX. A typical reaction was carried out with psBBOX (20 μM), 2OG (1.5 mM), ascorbate (5 mM), $FeSO_4$ (100 μM) and substrate (200 μM) in Tris buffer pH 7.5. (e) psBBOX catalysed hydroxylation of (3*S*) hydroxylated phospha **V**. The hydroxylated product is unstable, undergoes dehydration in the assay mixture and forms the 3-keto derivative.

However, our ^{1}H NMR results revealed that all the four hydroxylated ligands **IV–VII** are poor psBBOX substrates with low levels of turnover being observed in the presence of 10 μM psBBOX, likely to the corresponding C-3 ketone compounds or perhaps their acetyl derivatives (Figure 2.79), as these could not be fully characterised due to the low level of conversion.

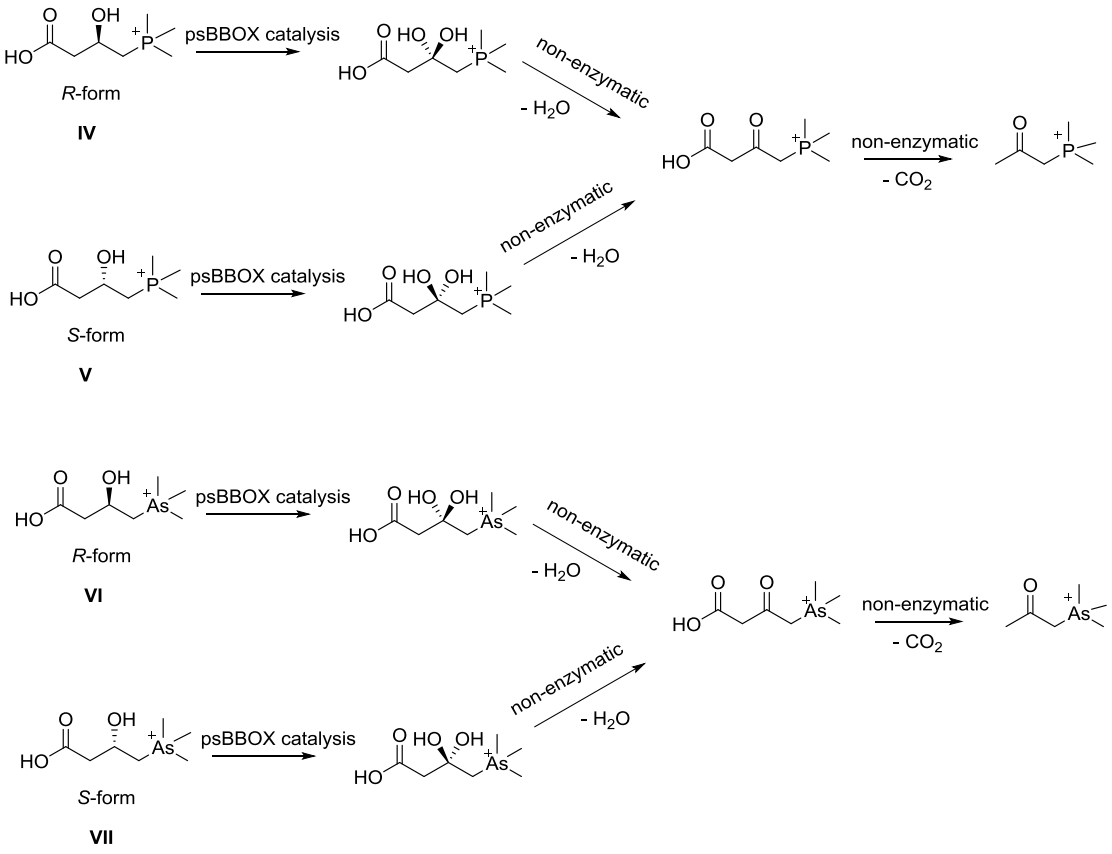


Figure 2.79 Reaction schemes of the psBBOX catalysed hydroxylation of ligands **IV** to **VII**.

Figure 2.80 shows the psBBOX catalysed hydroxylation reaction for (3*R*)-hydroxylated phospha ligand **IV**, Figure 2.81 for the (3*S*)-hydroxylated phospha ligand **V**, Figure 2.82 for the (3*R*)-hydroxylated arsa ligand **VI**, Figure 2.83 for the (3*S*)-hydroxylated arsa ligand **VII.** The observation of a product with the same chemical shift for the XMe_3^+ group from **IV**-**VII** supports the proposed formation of the corresponding 3-ketone product or perhaps their acetyl derivative in all four cases (Figure 2.79). As expected, in both the forms of phospha and arsa hydroxylated ligands, the (3*S*)-enantiomers (**V** and **VII**) were proved as relatively better substrates than their (3*R*)-enantiomers i.e **IV** and **VI** (turnover: 36% vs 7% for phospha **V** and **IV**, 22% vs 3% for arsa **VII** and **VI**). The relatively higher level of turnover for **V** compared with **VII** is consistent with the observations that the phospha GBB analogue is a

better psBBOX substrate than its arsa analogue as demonstrated by both of our ^{1}H NMR and LCMS based assays as stated previously. These results are consistent with previous results of hBBOX- and psBBOX-catalysed hydroxylation of D (*S*-form) and L (*R*-form) carnitine, for which as expected the level of conversion was higher for both *S*- and *R*- forms, as compared with the *S*- and *R*- forms of hydroxylated phospha and arsa analogues.[163,167]

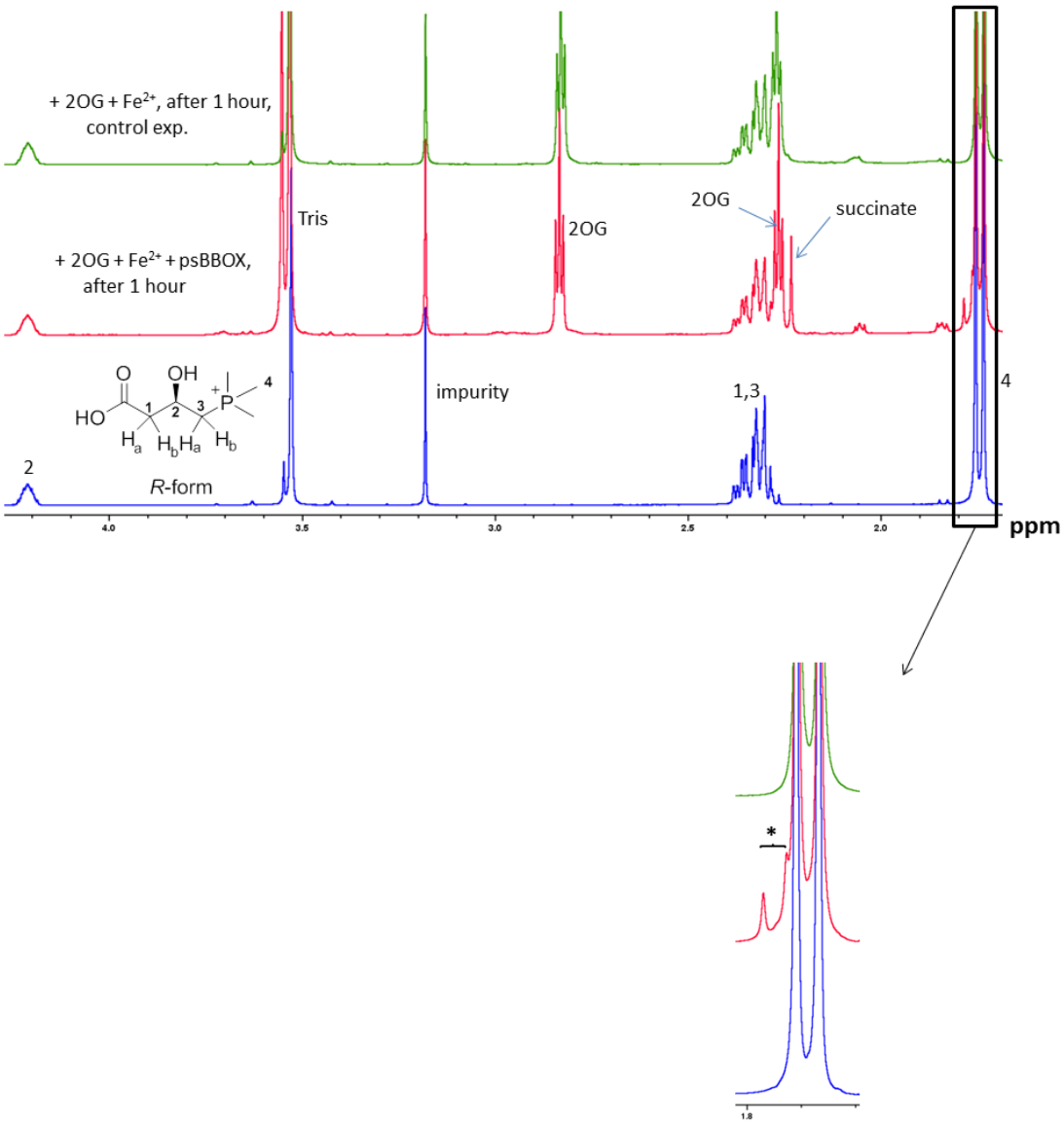


Figure 2.8 o NMR analysis of the psBBOX catalysed hydroxylation of (3*R*)-hydroxylated phospha ligand **IV**.
Hydroxylation of (3*R*)-hydroxylated phospha ligand **IV** (bottom, blue) in the presence (middle, red) and absence (top, green) of psBBOX. The appearance of the new P**Me$_3$** $^{+}$ peak (doublet, marked with *) indicates the formation of the product, likely the corresponding 3-keto derivative of **IV** (as observed by LC-MS analyses) or perhaps the acetyl derivative. The assay mixture contained 1 mM (3*R*)-hydroxylated phospha ligand **IV**, 1 mM 2OG, 80 mM KCl, 10 µM psBBOX and 100 µM Fe^{II} in 50 mM Tris-D_{11}, pH 7.5 in D_2O and incubated for 1 hour.

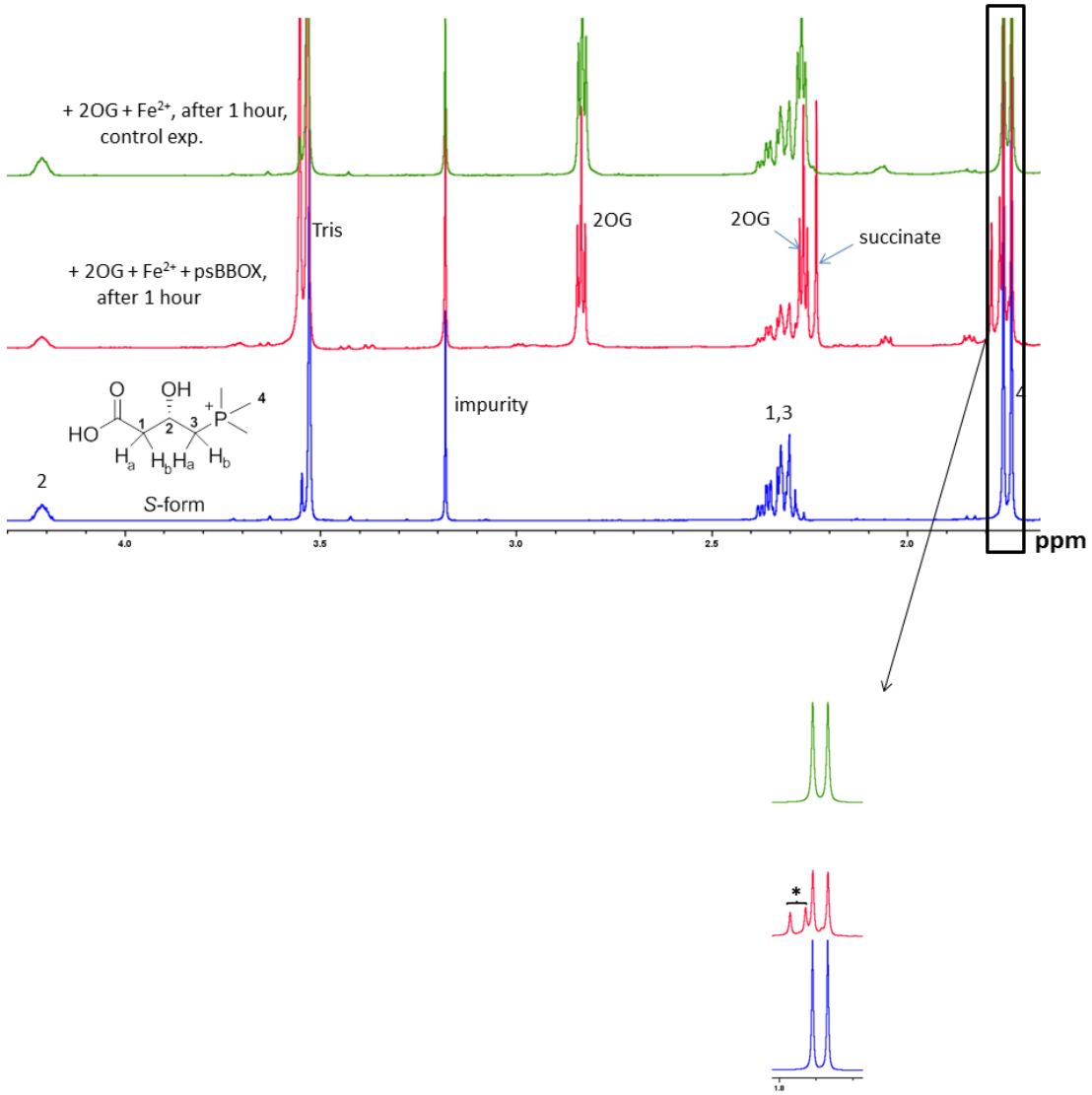


Figure 2.81 NMR analysis of the psBBOX catalysed hydroxylation of (3*S*)-hydroxylated phospha ligand **V**.
Hydroxylation of (3*S*)-hydroxylated phospha ligand **V** (bottom, blue) in the presence (middle, red) and absence (top, green) of psBBOX. The appearance of the new P**Me**$_3$$^+$ peak (doublet, marked with *) indicates the formation of the product, likely the corresponding 3-keto or acetyl derivative of **V**. The assay mixture contained 1 mM (3*S*)-hydroxylated phospha ligand **V**, 1 mM 2OG, 80 mM KCl, 10 µM psBBOX and 100 µM Fe^{II} in 50 mM Tris-D_{11}, pH 7.5 in D_2O and incubated for 1 hour.

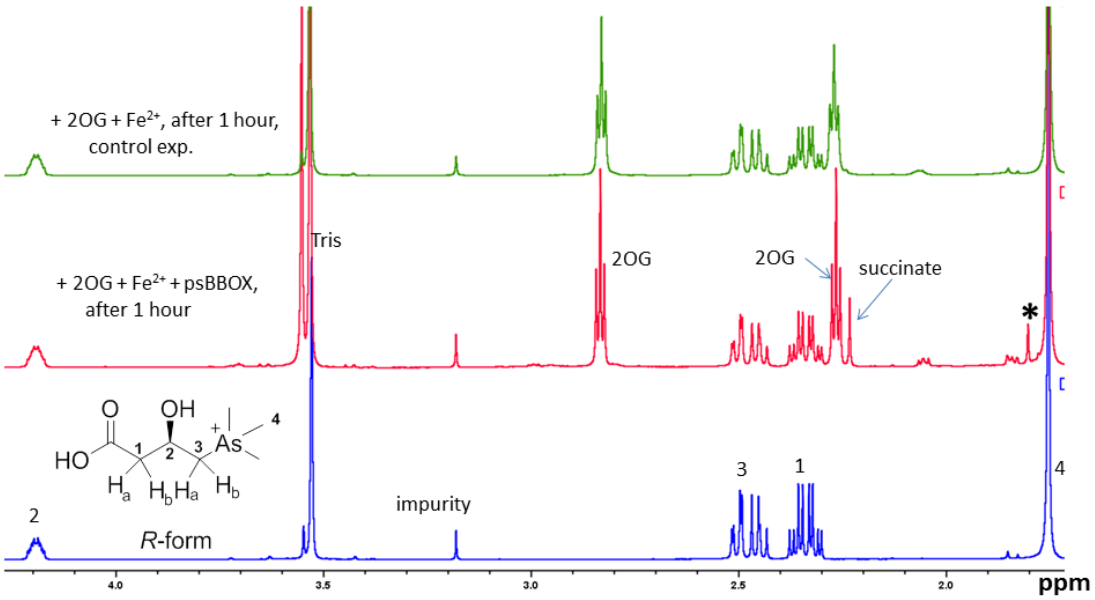


Figure 2.82 NMR analysis of the psBBOX catalysed hydroxylation of (3*R*)-hydroxylated arsa ligand **VI**.
Hydroxylation of (3*R*)-hydroxylated arsa ligand **VI** (bottom, blue) in the presence (middle, red) and absence (top, green) of psBBOX. The appearance of the new $AsMe_3$+ peak (marked with *) indicates the formation of the product, likely the corresponding 3-keto or acetyl derivative of **VI**. The assay mixture contained 1 mM (3*R*)-hydroxylated arsa ligand **VI**, 1 mM 2OG, 80 mM KCl, 10 µM psBBOX and 100 µM Fe^{II} in 50 mM Tris-D_{11}, pH 7.5 in D_2O and incubated for 1 hour.

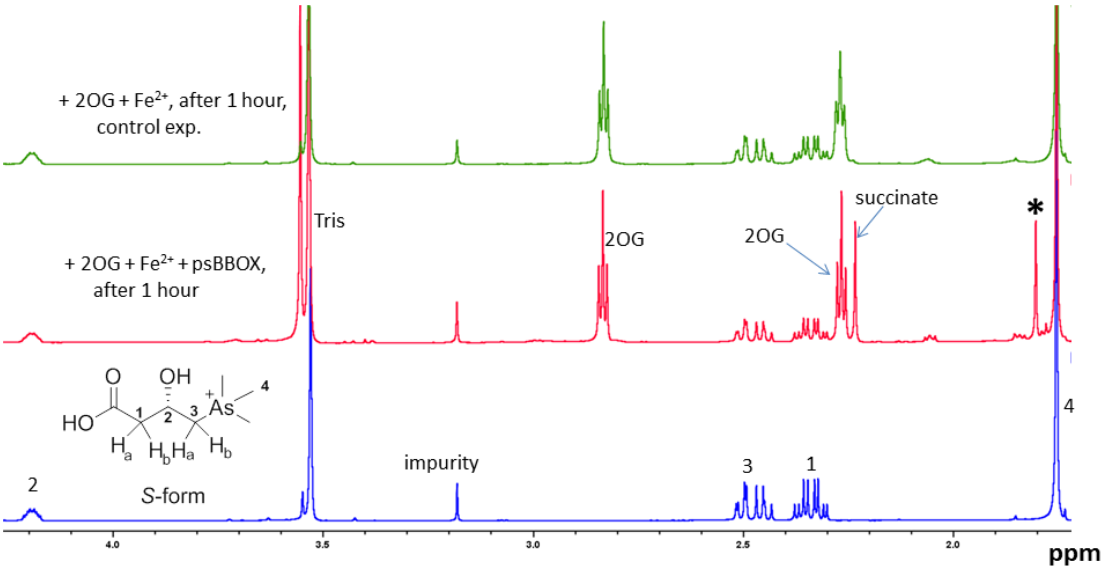


Figure 2.83 NMR analysis of the psBBOX catalysed hydroxylation of (3*S*)-hydroxylated arsa ligand **VII**.
Hydroxylation of (3*S*)-hydroxylated arsa **VII** (bottom, blue) in the presence (middle, red) and absence (top, green) of psBBOX. The appearance of the new $AsMe_3$+ peak (marked with *) indicates the formation of the product, likely the corresponding 3-keto or acetyl derivative of **VII**. The assay mixture contained 1 mM (3*S*)-hydroxylated arsa ligand **VII**, 1 mM 2OG, 80 mM KCl, 10 µM psBBOX and 100 µM Fe^{II} in 50 mM Tris-D_{11}, pH 7.5 in D_2O and incubated for 1 hour.

2.8 The complexity of binding of GBB to psBBOX

In exploring the psBBOX catalysis, some unexpected results were also observed. We noted that the BBOX natural substrate GBB cannot be fully out competed by a GBB analogue ligand (as observed by ^{1}H NMR-based dual-reporter displacement assay discussed in section 2.5), whereas a bulky ligand such as AMG-1957661 was able to fully displace GBB from its binding pocket. Figure 2.84 shows the results of the CPMG edited ^{1}H NMR reporter displacement assay for monitoring the competitive binding of the phospha GBB analogue to psBBOX. Part of the spectra showing only the GBB Me_3 signal (reporter ligand) at 3.05 ppm is shown. Figure 2.84a (II) shows the binding of 5 µM of GBB to 25 µM psBBOX in the presence of 700 µM 2OG and 250 µM Zn^{II}. The complete attenuation of GBB signal with the addition of psBBOX demonstrates that almost all the GBB molecules are bound to the protein. To this sample the phospha competitive ligand, which has almost similar activity as GBB as psBBOX substrate, was titrated in 100 µM (III), 300 µM (IV) and 600 µM (V) concentrations. As a result of this titration, only approximately half of the GBB molecules were displaced from the protein (Figure 2.84b).

In another experiment, we incubated the assay mixture of 1.5 mM phospha ligand/700 µM 2OG/250 µM Zn^{II}/25 µM psBBOX/buffer for 10 minutes in order to first fully saturate all the psBBOX molecules with the phospha ligands (in order to block all the GBB binding sites with the phospha ligand). To this sample, we then added 5 µM of GBB to see if GBB still binds. Unexpectedly, it was observed that the GBB signal attenuates by almost 50% (Figure 2.84aVI). To this sample, we then added the bulky ligand AMG (250 µM), and as it be seen that the addition of AMG-1957661 has resulted in almost complete displacement of GBB from its binding pocket (Figure 2.84aVII). The results of this experiment suggest that some of the GBB (possibly half of molecules) binds in a pocket, which is not accessed by the phospha ligand but accessed by the bulky ligand AMG-19557661.The displacement of GBB by AMG-1957661 is not surprising as its analogue has been previously shown to block the GBB binding pocket (in addition to 2OG binding pocket) in the crystal structure of hBBOX.[159]

The observation that why a GBB mimic ligand cannot fully displace GBB from its binding pocket is interesting. Since we have shown in section 2.7.2 that the phospha ligand has almost similar activity to psBBOX as GBB, therefore 600 µM of phospha ligand is more than enough to displace 5 µM GBB, however, it failed to do so. The partial displacement of GBB by the phospha ligand and the subsequent displacement of the remaining bound GBB by

AMG-1957661 may imply that the rest of the GBB molecules could perhaps be bound elsewhere, possibly in the other monomer (as BBOX is a dimer) and which is difficult to displace by another GBB mimic ligand. It is possible that GBB may be occupying both the binding sites in the dimer with equal population. Following the binding of GBB (and 2OG and active site metal), the loop above the active site (as shown in section 2.2.4, Figure 2.16) locks the substrate binding pocket. This makes the substrate difficult to be removed from its binding pocket by a substrate analogue competitive ligand. Perhaps, the substrate is released after when it is hydroxylated, with the loop above its binding pocket likely to open in order to release the product. In another scenario it is also possible that at a time only one monomer may be participating in catalysis, instead of both. The monomer not participating in the catalysis may have the substrate locked in a specific conformation and which cannot be displaced by another GBB mimic ligand. The displacement of GBB that occurs with the phospha ligand may be from the other monomer in which GBB is not locked. This partial displacement of GBB by another GBB mimic ligand appears to be a complex process. More insights perhaps may be obtained by further studies, including fluorine-labelling at both the catalytic sites of the BBOX dimer to probe protein conformational changes upon GBB binding, mass spectrometric studies to investigate the number of GBB molecules bound per BBOX dimer etc. This incomplete displacement of GBB by other GBB mimic ligands raises important questions as what would be the exact mechanism of binding of such ligands with BBOX.

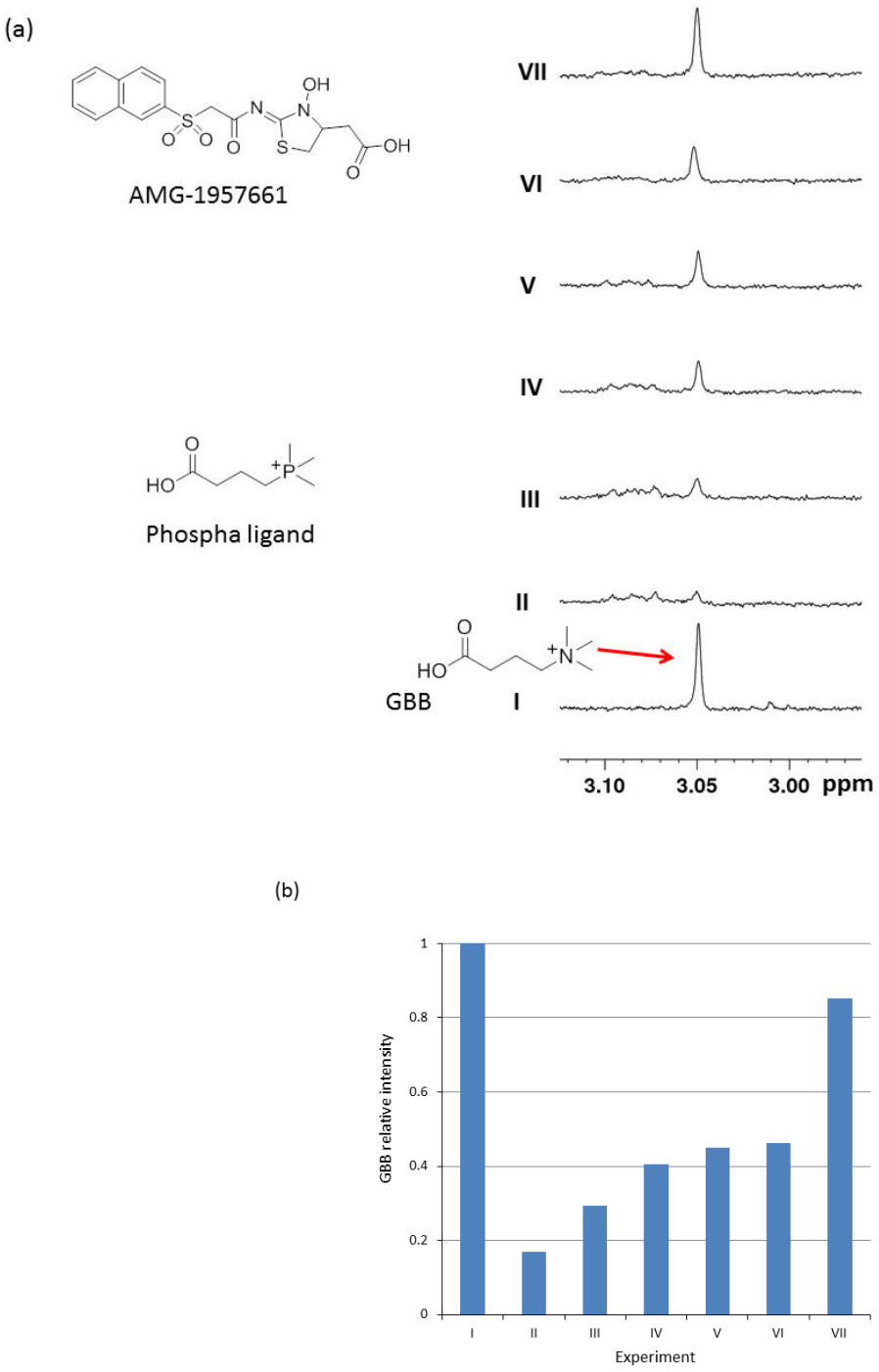


Figure 2.84 Partial displacement of GBB by its phospha analogue (**I**) as observed by CPMG edited ^{1}H NMR GBB displacement experiment.
(a) The titration of phospha ligand to GBB-psBBOX-2OG-Zn^{II} (only the GBB Me_3 signal at 3.05 ppm is shown). I: 5 µM GBB (+ 700 µM 2OG and 250 µM Zn^{II} in 50 mM Tris-D_{11}, pH 7.5 in D_2O.). II: after the addition of 25 µM psBBOX. III-V: titration of the phospha psBBOX substrate to II in 100 µM (III), 300 µM (IV) and 600 µM (V) concentration. VI: The 1.5 mM phospha ligand/700 µM 2OG/250 µM Zn^{II}/25 µM psBBOX/Tris buffer was first incubated for 10 minutes. After 10 min 5 µM GBB was added and a spectrum was recorded. VII: After the addition of 500 µM AMG-19557661 to VI. (b) Graphical representation of data of (a).

Similar results were also obtained for mildronate (THP), which is also a competitive substrate of BBOX.[155,163] The data showed that 1.5 mM of mildronate could not fully displace 5 µM of GBB form its binding pocket (Figure 2.85).

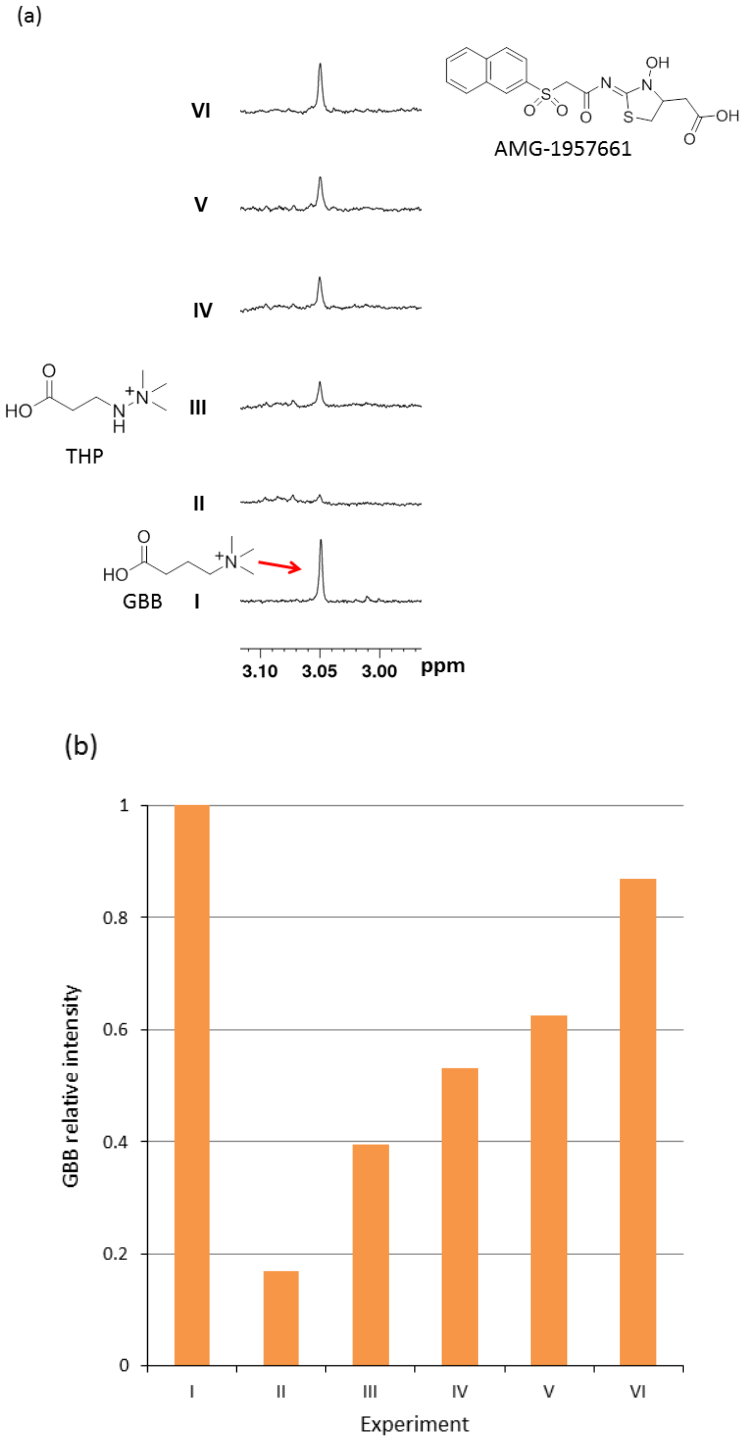


Figure 2.85 Partial displacement of GBB by THP as observed by CPMG edited ^{1}H NMR GBB displacement experiment. (a) Titration of mildronate to GBB-psBBOX-2OG-Zn^{II} in CPMG edited ^{1}H NMR experiments (only parts of spectra shown for clarity). I: Assay mixture of 5 µM GBB/700 µM 2OG/250 µM Zn^{II} in 50 mM Tris-D_{11}, pH 7.5 in D_2O. II: After the addition of 25 µM psBBOX. III-V: Addition of 500 µM THP, 1 mM THP and 1.5 mM THP respectively to II. Resonances of THP are not shown for clarity. VI: After the addition of 500 µM AMG-19557661 to V. (b) Graphical representation of data of (a).

Similarly, in case of GBBF (Figure 2.86III) and GBBNF (Figure 2.86IV), which are also substrates for both human and psBBOX,[166] of comparable activity to GBB; we however observed no displacement of GBB by these substrates.

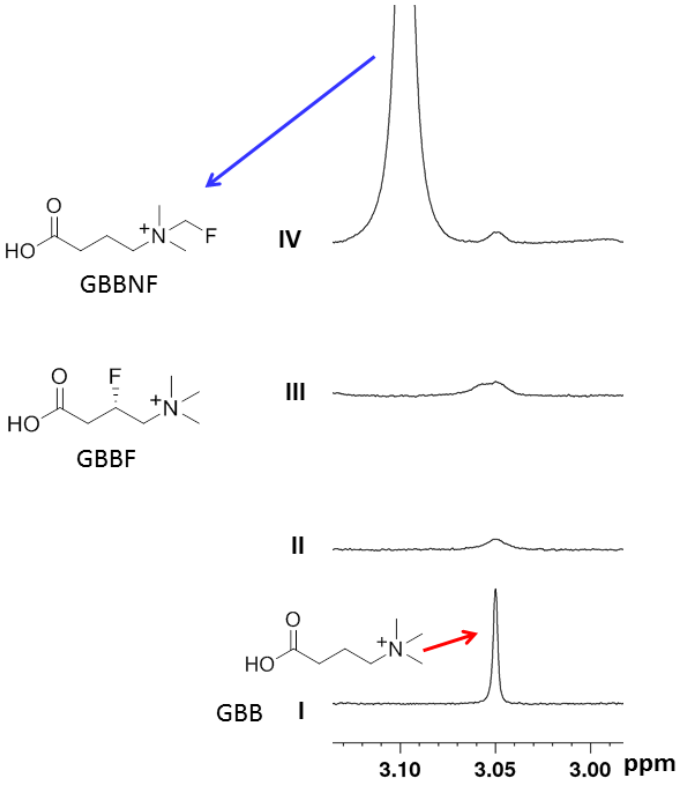


Figure 2.86 Partial displacement of GBB by GBBF and GBBNF as observed by CPMG edited ^{1}H NMR GBB displacement experiments.
I: Assay mixture of 25 µM GBB/300 µM 2OG/150 µM Mn^{II} in 50 mM Tris-D_{11}, pH 7.5 in D_2O. II: After the addition of 15 µM psBBOX. III: After the addition of 700 µM GBBF to II. IV: After the addition of 1 mM GBBNF to II.

2.9 Summary

In summary, the work described in this chapter demonstrates that the ^{1}H NMR direct ligand-observation is a useful technique for studying small molecules binding interaction with BBOX. In this assay, it was shown that the NMR sensitivity to ligand binding can be improved by the use of PRE effect using paramagnetic metal co-factor Mn^{II}. With the application of PRE effect, the consumption of protein in the assay was also significantly reduced. Based on this technique, using both GBB and 2OG simultaneously as reporter ligands, a NMR dual-reporter displacement assay was developed to assess the nature of an inhibitor binding to BBOX i.e whether the inhibitor binds by competing with GBB or 2OG or GBB. The applicability of this assay was demonstrated with a structure-activity relationship

(SAR) study using a series of isoquinoline-based ligands. We envisage that this assay will aid in the development of new potent BBOX inhibitors, including those selective for BBOX over other 2OG oxygenases.

Similarly, an efficient ^{1}H NMR-based assay was also developed to study BBOX inhibition by a ligand of interest. In this technique, the BBOX catalytic assay is performed in an eppendorf tube. The activity of the enzyme is then calculated from the concentrations of GBB and L-carnitine, which can be conveniently measured from the signal of their respective Me_3 group in a standard ^{1}H NMR spectrum. This assay was applied to measure the IC_{50} values of a series of inhibitors and the results obtained were in reasonable correlation with the values obtained with the reported fluoride ion release assay for BBOX.

In this work, it was also shown that the cation-π interaction between the positive charge of the substrate and the aromatic cage residues of the BBOX plays a critical role in BBOX substrate recognition. This was demonstrated by studying the binding interaction and turnover of three GBB analogues in which the nitrogen of the trimethylammonium group was replaced with phosphorous, arsenic and carbon atoms. The phospha and arsa analogues which contained the positive charge were proved to bind with the BBOX and undergo C-3 hydroxylation, whereas, the analogue with the neutral carbon atom did not show turnover with the BBOX, because it doesn't bind BBOX due to the lack of the positive charge.

Finally, it was also shown that mildronate (THP), which binds to hBBOX in the same site as GBB, however, cannot fully out compete GBB from its binding site. It is proposed that this unusual binding mode of GBB may have some potential biological relevance in terms of modulating the function of BBOX.

2.10 References

1. Hausinger, R. P.; Schofield, C. J. *2-Oxoglutarate-Dependent Oxygenases*. Royal Society of Chemistry: Cambridge, UK, **2015**.
2. Hausinger, R. P. Fe^{II}/alpha-ketoglutarate-dependent hydroxylases and related enzymes. *Crit. Rev. Biochem. Mol. Biol.* **2004**, 39, 21-68.
3. Costas, M.; Mehn, M. P.; Jensen, M. P.; Que, L., Jr. Dioxygen activation at mononuclear nonheme iron active sites: enzymes, models, and intermediates. *Chem. Rev.* **2004**, 104, 939-86.
4. Prescott, A. G.; Lloyd, M. D. The iron(II) and 2-oxoacid-dependent dioxygenases and their role in metabolism. *Nat. Prod. Rep.* **2000**, 17, 367-83.
5. Schofield, C. J.; Zhang, Z. Structural and mechanistic studies on 2-oxoglutarate-dependent oxygenases and related enzymes. *Curr. Opin. Struct. Biol.* **1999**, 9, 722-31.
6. Rose, N. R.; McDonough, M. A.; King, O. N.; Kawamura, A.; Schofield, C. J. Inhibition of 2-oxoglutarate dependent oxygenases. *Chem. Soc. Rev.* **2011**, 40, 4364-97.
7. Loenarz, C.; Schofield, C. J. Physiological and biochemical aspects of hydroxylations and demethylations catalyzed by human 2-oxoglutarate oxygenases. *Trends Biochem. Sci.* **2011**, 36, 7-18.
8. Loenarz, C.; Schofield, C. J. Expanding chemical biology of 2-oxoglutarate oxygenases. *Nat. Chem. Biol.* **2008**, 4, 152-6.
9. Kovaleva, E. G.; Lipscomb, J. D. Versatility of biological non-heme Fe(II) centers in oxygen activation reactions. *Nat. Chem. Biol.* **2008**, 4, 186-93.
10. Ozer, A.; Bruick, R. K. Non-heme dioxygenases: cellular sensors and regulators jelly rolled into one? *Nat. Chem. Biol.* **2007**, 3, 144-53.
11. Flashman, E.; Schofield, C. J. The most versatile of all reactive intermediates? *Nat. Chem. Biol.* **2007**, 3, 86-7.
12. Hewitson, K. S.; Granatino, N.; Welford, R. W.; McDonough, M. A.; Schofield, C. J. Oxidation by 2-oxoglutarate oxygenases: non-haem iron systems in catalysis and signalling. *Philos. Trans. A Math. Phys. Eng. Sci.* **2005**, 363, 807-28; discussion 1035-40.
13. Sono, M.; Roach, M. P.; Coulter, E. D.; Dawson, J. H. Heme-Containing Oxygenases. *Chem. Rev.* **1996**, 96, 2841-2888.

14. Hanauske-Abel, H. M.; Gunzler, V. A stereochemical concept for the catalytic mechanism of prolylhydroxylase: applicability to classification and design of inhibitors. *J. Theor. Biol.* **1982**, 94, 421-55.

15. Solomon, E. I.; Decker, A.; Lehnert, N. Non-heme iron enzymes: contrasts to heme catalysis. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, 100, 3589-94.

16. Myllyla, R.; Tuderman, L.; Kivirikko, K. I. Mechanism of the prolyl hydroxylase reaction. 2. Kinetic analysis of the reaction sequence. *Eur. J. Biochem.* **1977**, 80, 349-57.

17. Elkins, J. M.; Hewitson, K. S.; McNeill, L. A.; Seibel, J. F.; Schlemminger, I.; Pugh, C. W.; Ratcliffe, P. J.; Schofield, C. J. Structure of Factor-inhibiting Hypoxia-inducible Factor (HIF) Reveals Mechanism of Oxidative Modification of HIF-1α. *J. Biol. Chem.* **2003**, 278, 1802-1806.

18. Hoffart, L. M.; Barr, E. W.; Guyer, R. B.; Bollinger, J. M., Jr.; Krebs, C. Direct spectroscopic detection of a C-H-cleaving high-spin Fe(IV) complex in a prolyl-4-hydroxylase. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, 103, 14738-43.

19. Riggs-Gelasco, P. J.; Price, J. C.; Guyer, R. B.; Brehm, J. H.; Barr, E. W.; Bollinger, J. M., Jr.; Krebs, C. EXAFS spectroscopic evidence for an Fe=O unit in the Fe(IV) intermediate observed during oxygen activation by taurine:alpha-ketoglutarate dioxygenase. *J. Am. Chem. Soc.* **2004**, 126, 8108-9.

20. Proshlyakov, D. A.; Henshaw, T. F.; Monterosso, G. R.; Ryle, M. J.; Hausinger, R. P. Direct detection of oxygen intermediates in the non-heme Fe enzyme taurine/alpha-ketoglutarate dioxygenase. *J. Am. Chem. Soc.* **2004**, 126, 1022-3.

21. Price, J. C.; Barr, E. W.; Tirupati, B.; Bollinger, J. M., Jr.; Krebs, C. The first direct characterization of a high-valent iron intermediate in the reaction of an alpha-ketoglutarate-dependent dioxygenase: a high-spin FeIV complex in taurine/alpha-ketoglutarate dioxygenase (TauD) from Escherichia coli. *Biochemistry* **2003**, 42, 7497-508.

22. Welford, R. W.; Kirkpatrick, J. M.; McNeill, L. A.; Puri, M.; Oldham, N. J.; Schofield, C. J. Incorporation of oxygen into the succinate co-product of iron(II) and 2-oxoglutarate dependent oxygenases from bacteria, plants and humans. *FEBS Lett.* **2005**, 579, 5170-4.

23. Grzyska, P. K.; Appelman, E. H.; Hausinger, R. P.; Proshlyakov, D. A. Insight into the mechanism of an iron dioxygenase by resolution of steps following the Fe^{IV}=HO species. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, 107, 3982-7.

24. Myllyla, R.; Majamaa, K.; Gunzler, V.; Hanauske-Abel, H. M.; Kivirikko, K. I. Ascorbate is consumed stoichiometrically in the uncoupled reactions catalyzed by prolyl 4-hydroxylase and lysyl hydroxylase. *J. Biol. Chem.* **1984**, 259, 5403-5.
25. Ryle, M. J.; Liu, A.; Muthukumaran, R. B.; Ho, R. Y.; Koehntop, K. D.; McCracken, J.; Que, L., Jr.; Hausinger, R. P. O2- and alpha-ketoglutarate-dependent tyrosyl radical formation in TauD, an alpha-keto acid-dependent non-heme iron dioxygenase. *Biochemistry* **2003**, 42, 1854-62.
26. Barlow, J. N.; Zhang, Z.; John, P.; Baldwin, J. E.; Schofield, C. J. Inactivation of 1-aminocyclopropane-1-carboxylate oxidase involves oxidative modifications. *Biochemistry* **1997**, 36, 3563-9.
27. McCusker, K. P.; Klinman, J. P. Modular behavior of tauD provides insight into the origin of specificity in alpha-ketoglutarate-dependent nonheme iron oxygenases. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, 106, 19791-5.
28. Mantri, M.; Zhang, Z.; McDonough, M. A.; Schofield, C. J. Autocatalysed oxidative modifications to 2-oxoglutarate dependent oxygenases. *FEBS J* **2012**, 279, 1563-75.
29. Mantri, M.; Webby, C. J.; Loik, N. D.; Hamed, R. B.; Nielsen, M. L.; McDonough, M. A.; McCullagh, J. S. O.; Bottger, A.; Schofield, C. J.; Wolf, A. Self-hydroxylation of the splicing factor lysyl hydroxylase, JMJD6. *MedChemComm* **2012**, 3, 80-85.
30. Valegard, K.; Terwisscha van Scheltinga, A. C.; Dubus, A.; Ranghino, G.; Oster, L. M.; Hajdu, J.; Andersson, I. The structural basis of cephalosporin formation in a mononuclear ferrous enzyme. *Nat. Struct. Mol. Biol.* **2004**, 11, 95-101.
31. Tarhonskaya, H.; Szöllössi, A.; Leung, I. K. H.; Bush, J. T.; Henry, L.; Chowdhury, R.; Iqbal, A.; Claridge, T. D. W.; Schofield, C. J.; Flashman, E. Studies on Deacetoxycephalosporin C Synthase Support a Consensus Mechanism for 2-Oxoglutarate Dependent Oxygenases. *Biochemistry* **2014**, 53, 2483-2493.
32. Mecinovic, J.; Chowdhury, R.; Lienard, B. M.; Flashman, E.; Buck, M. R.; Oldham, N. J.; Schofield, C. J. ESI-MS studies on prolyl hydroxylase domain 2 reveal a new metal binding site. *ChemMedChem* **2008**, 3, 569-72.
33. Sekirnik, R.; Rose, N. R.; Mecinovic, J.; Schofield, C. J. 2-Oxoglutarate oxygenases are inhibited by a range of transition metals. *Metallomics* **2010**, 2, 397-9.
34. Kalliri, E.; Grzyska, P. K.; Hausinger, R. P. Kinetic and spectroscopic investigation of CoII, NiII, and N-oxalylglycine inhibition of the FeII/alpha-ketoglutarate dioxygenase, TauD. *Biochem. Biophys. Res. Commun.* **2005**, 338, 191-7.

35. Hirsila, M.; Koivunen, P.; Xu, L.; Seeley, T.; Kivirikko, K. I.; Myllyharju, J. Effect of desferrioxamine and metals on the hydroxylases in the oxygen sensing pathway. *FASEB J.* **2005**, 19, 1308-10.

36. Flashman, E.; Davies, S. L.; Yeoh, K. K.; Schofield, C. J. Investigating the dependence of the hypoxia-inducible factor hydroxylases (factor inhibiting HIF and prolyl hydroxylase domain 2) on ascorbate and other reducing agents. *Biochem. J.* **2010**, 427, 135-42.

37. Cunliffe, C. J.; Franklin, T. J.; Hales, N. J.; Hill, G. B. Novel inhibitors of prolyl 4-hydroxylase. 3. Inhibition by the substrate analogue N-oxaloglycine and its derivatives. *J. Med. Chem.* **1992**, 35, 2652-8.

38. Wehbie, R. S.; Punekar, N. S.; Lardy, H. A. Rat liver .gamma.-butyrobetaine hydroxylase catalyzed reaction: influence of potassium, substrates, and substrate analogs on hydroxylation and decarboxylation. *Biochemistry* **1988**, 27, 2222-2228.

39. Majamaa, K.; Hanauske-Abel, H. M.; Gunzler, V.; Kivirikko, K. I. The 2-oxoglutarate binding site of prolyl 4-hydroxylase. Identification of distinct subsites and evidence for 2-oxoglutarate decarboxylation in a ligand reaction at the enzyme-bound ferrous ion. *Eur. J. Biochem.* **1984**, 138, 239-45.

40. Tarhonskaya, H.; Rydzik, A. M.; Leung, I. K. H.; Loik, N. D.; Chan, M. C.; Kawamura, A.; McCullagh, J. S. O.; Claridge, T. D. W.; Flashman, E.; Schofield, C. J. Non-enzymatic chemistry enables 2-hydroxyglutarate-mediated activation of 2-oxoglutarate oxygenases. *Nature Communications* **2014**, 5, 3423.

41. McDonough, M. A.; Loenarz, C.; Chowdhury, R.; Clifton, I. J.; Schofield, C. J. Structural studies on human 2-oxoglutarate dependent oxygenases. *Curr. Opin. Struct. Biol.* **2010**, 20, 659-72.

42. Clifton, I. J.; McDonough, M. A.; Ehrismann, D.; Kershaw, N. J.; Granatino, N.; Schofield, C. J. Structural studies on 2-oxoglutarate oxygenases and related double-stranded beta-helix fold proteins. *J. Inorg. Biochem.* **2006**, 100, 644-69.

43. Stirk, H. J.; Woolfson, D. N.; Hutchinson, E. G.; Thornton, J. M. Depicting topology and handedness in jellyroll structures. *FEBS Lett.* **1992**, 308, 1-3.

44. Hegg, E. L.; Que, L., Jr. The 2-His-1-carboxylate facial triad--an emerging structural motif in mononuclear non-heme iron(II) enzymes. *Eur. J. Biochem.* **1997**, 250, 625-9.

45. Hewitson, K. S.; Holmes, S. L.; Ehrismann, D.; Hardy, A. P.; Chowdhury, R.; Schofield, C. J.; McDonough, M. A. Evidence that two enzyme-derived histidine

ligands are sufficient for iron binding and catalysis by factor inhibiting HIF (FIH). *J. Biol. Chem.* **2008**, 283, 25971-8.

46. Aik, W.; McDonough, M. A.; Thalhammer, A.; Chowdhury, R.; Schofield, C. J. Role of the jelly-roll fold in substrate binding by 2-oxoglutarate oxygenases. *Curr. Opin. Struct. Biol.* **2012**, 22, 691-700.
47. Nagel, S.; Talbot, N. P.; Mecinovic, J.; Smith, T. G.; Buchan, A. M.; Schofield, C. J. Therapeutic manipulation of the HIF hydroxylases. *Antioxid Redox Signal* **2010**, 12, 481-501.
48. Fraisl, P.; Aragones, J.; Carmeliet, P. Inhibition of oxygen sensors as a therapeutic strategy for ischaemic and inflammatory disease. *Nat Rev Drug Discov* **2009**, 8, 139-52.
49. Hewitson, K. S.; Schofield, C. J. The HIF pathway as a therapeutic target. *Drug Discov Today* **2004**, 9, 704-11.
50. Semenza, G. L. Targeting HIF-1 for cancer therapy. *Nat. Rev. Cancer* **2003**, 3, 721-32.
51. Sesti, C.; Simkhovich, B. Z.; Kalvinsh, I.; Kloner, R. A. Mildronate, a novel fatty acid oxidation inhibitor and antianginal agent, reduces myocardial infarct size without affecting hemodynamics. *J. Cardiovasc. Pharmacol.* **2006**, 47, 493-9.
52. Rademacher, W. Growth Retardants: Effects on Gibberellin Biosynthesis and Other Metabolic Pathways. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **2000**, 51, 501-531.
53. Yeoh, K. K.; Chan, M. C.; Thalhammer, A.; Demetriades, M.; Chowdhury, R.; Tian, Y. M.; Stolze, I.; McNeill, L. A.; Lee, M. K.; Woon, E. C.; Mackeen, M. M.; Kawamura, A.; Ratcliffe, P. J.; Mecinovic, J.; Schofield, C. J. Dual-action inhibitors of HIF prolyl hydroxylases that induce binding of a second iron ion. *Org. Biomol. Chem.* **2013**, 11, 732-45.
54. Tegley, C. M.; Viswanadhan, V. N.; Biswas, K.; Frohn, M. J.; Peterkin, T. A.; Chang, C.; Burli, R. W.; Dao, J. H.; Veith, H.; Rogers, N.; Yoder, S. C.; Biddlecome, G.; Tagari, P.; Allen, J. R.; Hungate, R. W. Discovery of novel hydroxy-thiazoles as HIF-alpha prolyl hydroxylase inhibitors: SAR, synthesis, and modeling evaluation. *Bioorg. Med. Chem. Lett.* **2008**, 18, 3925-8.
55. McDonough, M. A.; McNeill, L. A.; Tilliet, M.; Papamicael, C. A.; Chen, Q. Y.; Banerji, B.; Hewitson, K. S.; Schofield, C. J. Selective inhibition of factor inhibiting hypoxia-inducible factor. *J. Am. Chem. Soc.* **2005**, 127, 7680-1.

56. Chang, K. H.; King, O. N.; Tumber, A.; Woon, E. C.; Heightman, T. D.; McDonough, M. A.; Schofield, C. J.; Rose, N. R. Inhibition of histone demethylases by 4-carboxy-2,2'-bipyridyl compounds. *ChemMedChem* **2011**, 6, 759-64.

57. Hales, N. J.; Beattie, J. F. Novel inhibitors of prolyl 4-hydroxylase. 5. The intriguing structure-activity relationships seen with 2,2'-bipyridine and its 5,5'-dicarboxylic acid derivatives. *J. Med. Chem.* **1993**, 36, 3853-8.

58. Warshakoon, N. C.; Wu, S.; Boyer, A.; Kawamoto, R.; Sheville, J.; Renock, S.; Xu, K.; Pokross, M.; Evdokimov, A. G.; Walter, R.; Mekel, M. A novel series of imidazo[1,2-a]pyridine derivatives as HIF-1alpha prolyl hydroxylase inhibitors. *Bioorg. Med. Chem. Lett.* **2006**, 16, 5598-601.

59. Dao, J. H.; Kurzeja, R. J.; Morachis, J. M.; Veith, H.; Lewis, J.; Yu, V.; Tegley, C. M.; Tagari, P. Kinetic characterization and identification of a novel inhibitor of hypoxia-inducible factor prolyl hydroxylase 2 using a time-resolved fluorescence resonance energy transfer-based assay technology. *Anal. Biochem.* **2009**, 384, 213-23.

60. Stubbs, C. J.; Loenarz, C.; Mecinovic, J.; Yeoh, K. K.; Hindley, N.; Lienard, B. M.; Sobott, F.; Schofield, C. J.; Flashman, E. Application of a proteolysis/mass spectrometry method for investigating the effects of inhibitors on hydroxylase structure. *J. Med. Chem.* **2009**, 52, 2799-805.

61. Kwon, H. S.; Choi, Y. K.; Kim, J. W.; Park, Y. K.; Yang, E. G.; Ahn, D. R. Inhibition of a prolyl hydroxylase domain (PHD) by substrate analog peptides. *Bioorg. Med. Chem. Lett.* **2011**, 21, 4325-8.

62. Simkhovich, B. Z.; Shutenko, Z. V.; Meirena, D. V.; Khagi, K. B.; Mezapuke, R. J.; Molodchina, T. N.; Kalvins, I. J.; Lukevics, E. 3-(2,2,2-Trimethylhydrazinium)propionate (THP)--a novel gamma-butyrobetaine hydroxylase inhibitor with cardioprotective properties. *Biochem. Pharmacol.* **1988**, 37, 195-202.

63. Liepinsh, E.; Makrecka-Kuka, M.; Kuka, J.; Vilskersts, R.; Makarova, E.; Cirule, H.; Loza, E.; Lola, D.; Grinberga, S.; Pugovics, O.; Kalvins, I.; Dambrova, M. Inhibition of L-carnitine biosynthesis and transport by methyl-gamma-butyrobetaine decreases fatty acid oxidation and protects against myocardial infarction. *Br. J. Pharmacol.* **2015**, 172, 1319-32.

64. Jayatunga, M. K. P.; Thompson, S.; McKee, T. C.; Chan, M. C.; Reece, K. M.; Hardy, A. P.; Sekirnik, R.; Seden, P. T.; Cook, K. M.; McMahon, J. B.; Figg, W. D.; Schofield, C. J.; Hamilton, A. D. Inhibition of the HIF1α-p300 interaction by

quinone- and indandione-mediated ejection of structural Zn(II). *Eur. J. Med. Chem.* **2015**, 94, 509-516.

65. Sekirnik, R.; Rose, N. R.; Thalhammer, A.; Seden, P. T.; Mecinovic, J.; Schofield, C. J. Inhibition of the histone lysine demethylase JMJD2A by ejection of structural Zn(II). *Chem. Commun. (Camb.)* **2009**, 6376-8.
66. Scozzafava, A.; Casini, A.; Supuran, C. T. Targeting cysteine residues of biomolecules: new approaches for the design of antiviral and anticancer drugs. *Curr. Med. Chem.* **2002**, 9, 1167-85.
67. Rydzik, A. M.; Brem, J.; Struwe, W. B.; Kochan, G. T.; Benesch, J. L.; Schofield, C. J. Ejection of structural zinc leads to inhibition of gamma-butyrobetaine hydroxylase. *Bioorg. Med. Chem. Lett.* **2014**, 24, 4954-7.
68. Chowdhury, R.; Flashman, E.; Mecinovic, J.; Kramer, H. B.; Kessler, B. M.; Frapart, Y. M.; Boucher, J. L.; Clifton, I. J.; McDonough, M. A.; Schofield, C. J. Studies on the reaction of nitric oxide with the hypoxia-inducible factor prolyl hydroxylase domain 2 (EGLN1). *J. Mol. Biol.* **2011**, 410, 268-79.
69. Metzen, E.; Zhou, J.; Jelkmann, W.; Fandrey, J.; Brune, B. Nitric oxide impairs normoxic degradation of HIF-1alpha by inhibition of prolyl hydroxylases. *Mol. Biol. Cell* **2003**, 14, 3470-81.
70. Tuderman, L.; Myllyla, R.; Kivirikko, K. I. Mechanism of the prolyl hydroxylase reaction. 1. Role of co-substrates. *Eur. J. Biochem.* **1977**, 80, 341-8.
71. Begley, T. J.; Samson, L. D. AlkB mystery solved: oxidative demethylation of N1-methyladenine and N3-methylcytosine adducts by a direct reversal mechanism. *Trends Biochem. Sci.* **2003**, 28, 2-5.
72. Steiber, A.; Kerner, J.; Hoppel, C. L. Carnitine: a nutritional, biosynthetic, and functional perspective. *Mol. Aspects Med.* **2004**, 25, 455-73.
73. Fraenkel, G.; Friedman, S. Carnitine. *Vitam. Horm.* **1957**, 15, 73-118.
74. Czeczot, H.; Scibior, D. Role of L-carnitine in metabolism, nutrition and therapy. *Postepy Hig Med Dosw (Online)* **2005**, 59, 9-19.
75. Bremer, J. Carnitine--metabolism and functions. *Physiol. Rev.* **1983**, 63, 1420-80.
76. Rebouche, C. J.; Seim, H. Carnitine metabolism and its regulation in microorganisms and mammals. *Annu. Rev. Nutr.* **1998**, 18, 39-61.
77. Rebouche, C. J. Carnitine function and requirements during the life cycle. *FASEB J.* **1992**, 6, 3379-86.

78. Panter, R. A.; Mudd, J. B. Carnitine levels in some higher plants. *FEBS Lett.* **1969**, 5, 169-170.
79. Kleber, H. P. Bacterial carnitine metabolism. *FEMS Microbiol. Lett.* **1997**, 147, 1-9.
80. Vaz, F. M.; Wanders, R. J. Carnitine biosynthesis in mammals. *Biochem. J.* **2002**, 361, 417-29.
81. Ramsay, R. R.; Gandour, R. D.; van der Leij, F. R. Molecular enzymology of carnitine transfer and transport. *Biochim. Biophys. Acta* **2001**, 1546, 21-43.
82. Kerner, J.; Hoppel, C. Fatty acid import into mitochondria. *Biochim. Biophys. Acta* **2000**, 1486, 1-17.
83. McGarry, J. D.; Brown, N. F. The mitochondrial carnitine palmitoyltransferase system. From concept to molecular analysis. *Eur. J. Biochem.* **1997**, 244, 1-14.
84. Lowenstein, J. M. *Methods in Enzymology*. Academic Press: Boston, 1669; Vol. 13.
85. Schulz, H. Beta oxidation of fatty acids. *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism* **1991**, 1081, 109-120.
86. Opie, L. H. Metabolism of the heart in health and disease. Part I. *Am. Heart J.* **1968**, 76, 685-698.
87. Bing, R. J. Cardiac Metabolism. *Physiol. Rev.* **1965**, 45, 171-213.
88. Jakobs, B. S.; Wanders, R. J. Fatty acid beta-oxidation in peroxisomes and mitochondria: the first, unequivocal evidence for the involvement of carnitine in shuttling propionyl-CoA from peroxisomes to mitochondria. *Biochem. Biophys. Res. Commun.* **1995**, 213, 1035-41.
89. Hoppel, C. L. Carnitine and carnitine palmitoyltransferase in fatty acid oxidation and ketosis. *Fed. Proc.* **1982**, 41, 2853-7.
90. Fritz, I. B. Carnitine and its role in fatty acids metabolism. *Adv. Lipid Res.* **1963**, 1, 285-334.
91. Carter, A. L.; Abney, T. O.; Lapp, D. F. Biosynthesis and metabolism of carnitine. *J. Child Neurol.* **1995**, 10 Suppl 2, S3-7.
92. Duran, M.; Loof, N. E.; Ketting, D.; Dorland, L. Secondary carnitine deficiency. *J. Clin. Chem. Clin. Biochem.* **1990**, 28, 359-63.
93. Bremer, J. Carnitine in intermediary metabolism. Reversible acetylation of carnitine by mitochondria. *J. Biol. Chem.* **1962**, 237, 2228-31.
94. Friedman, S.; Fraenkel, G. Reversible enzymatic acetylation of carnitine. *Arch. Biochem. Biophys.* **1955**, 59, 491-501.

95. Liedtke, A. J.; Nellis, S. H.; Whitesell, L. F.; Mahar, C. Q. Metabolic and mechanical effects using L- and D-carnitine in working swine hearts. *Am. J. Physiol. Heart Circ. Physiol.* **1982**, 243, H691-H697.

96. Lennon, D. L.; Shrago, E. R.; Madden, M.; Nagle, F. J.; Hanson, P. Dietary carnitine intake related to skeletal muscle and plasma carnitine concentrations in adult men and women. *Am. J. Clin. Nutr.* **1986**, 43, 234-8.

97. Paulson, D. J. Carnitine deficiency-induced cardiomyopathy. *Mol. Cell. Biochem.* **1998**, 180, 33-41.

98. Rippa, S.; Zhao, Y.; Merlier, F.; Charrier, A.; Perrin, Y. The carnitine biosynthetic pathway in Arabidopsis thaliana shares similar features with the pathway of mammals and fungi. *Plant Physiol. Biochem.* **2012**, 60, 109-14.

99. Strijbis, K.; Vaz, F. M.; Distel, B. Enzymology of the carnitine biosynthesis pathway. *IUBMB Life* **2010**, 62, 357-62.

100. Lindstedt, G.; Lindstedt, S.; Midtvedt, T.; Tofft, M. The Formation and Degradation of Carnitine in Pseudomonas. *Biochemistry* **1967**, 6, 1262-1270.

101. Dunn, W. A.; Rettura, G.; Seifter, E.; Englard, S. Carnitine biosynthesis from gamma-butyrobetaine and from exogenous protein-bound 6-N-trimethyl-L-lysine by the perfused guinea pig liver. Effect of ascorbate deficiency on the in situ activity of gamma-butyrobetaine hydroxylase. *J. Biol. Chem.* **1984**, 259, 10764-70.

102. Hulse, J. D.; Ellis, S. R.; Henderson, L. M. Carnitine biosynthesis. beta-Hydroxylation of trimethyllysine by an alpha-ketoglutarate-dependent mitochondrial dioxygenase. *J. Biol. Chem.* **1978**, 253, 1654-9.

103. LaBadie, J.; Dunn, W. A.; Aronson, N. N., Jr. Hepatic synthesis of carnitine from protein-bound trimethyl-lysine. Lysosomal digestion of methyl-lysine-labelled asialo-fetuin. *Biochem. J.* **1976**, 160, 85-95.

104. Tanphaichitr, V.; Broquist, H. P. Role of lysine and -N-trimethyllysine in carnitine biosynthesis. II. Studies in the rat. *J. Biol. Chem.* **1973**, 248, 2176-81.

105. Horne, D. W.; Broquist, H. P. Role of lysine and -N-trimethyllysine in carnitine biosynthesis. I. Studies in Neurospora crassa. *J. Biol. Chem.* **1973**, 248, 2170-5.

106. Tanphaichitr, V.; Horne, D. W.; Broquist, H. P. Lysine, a precursor of carnitine in the rat. *J. Biol. Chem.* **1971**, 246, 6364-6.

107. Paik, W. K.; Kim, S. Protein methylation. *Science* **1971**, 174, 114-9.

108. Morse, R. K.; Vergnes, J. P.; Malloy, J.; McManus, I. R. Sites of biological methylation of proteins in cultured chick muscle cells. *Biochemistry* **1975**, 14, 4316-25.

109. Huszar, G. Tissue-specific biosynthesis of epsilon-N-monomethyllysine and epsilon-N-trimethyllysine in skeletal and cardiac muscle myosin: a model for the cell-free study of post-translational amino acid modifications in proteins. *J. Mol. Biol.* **1975**, 94, 311-26.

110. Vaz, F. M.; Ofman, R.; Westinga, K.; Back, J. W.; Wanders, R. J. Molecular and Biochemical Characterization of Rat N^{ε}-Trimethyllysine Hydroxylase, the First Enzyme of Carnitine Biosynthesis. *J. Biol. Chem.* **2001**, 276, 33512-7.

111. Kazaks, A.; Makrecka-Kuka, M.; Kuka, J.; Voronkova, T.; Akopjana, I.; Grinberga, S.; Pugovics, O.; Tars, K. Expression and purification of active, stabilized trimethyllysine hydroxylase. *Protein Expr. Purif.* **2014**, 104c, 1-6.

112. McNeil, J. B.; Flynn, J.; Tsao, N.; Monschau, N.; Stahmann, K.; Haynes, R. H.; McIntosh, E. M.; Pearlman, R. E. Glycine metabolism in Candida albicans: characterization of the serine hydroxymethyltransferase (SHM1, SHM2) and threonine aldolase (GLY1) genes. *Yeast* **2000**, 16, 167-75.

113. Schirch, L.; Peterson, D. Purification and properties of mitochondrial serine hydroxymethyltransferase. *J. Biol. Chem.* **1980**, 255, 7801-6.

114. Hassan, M.; Okada, M.; Ichiyanagi, T.; Mori, N. 4-N-trimethylaminobutyraldehyde dehydrogenase: purification and characterization of an enzyme from Pseudomonas sp. 13CM. *Biosci. Biotechnol. Biochem.* **2008**, 72, 155-62.

115. Vaz, F. M.; Fouchier, S. W.; Ofman, R.; Sommer, M.; Wanders, R. J. Molecular and biochemical characterization of rat gamma-trimethylaminobutyraldehyde dehydrogenase and evidence for the involvement of human aldehyde dehydrogenase 9 in carnitine biosynthesis. *J. Biol. Chem.* **2000**, 275, 7390-4.

116. Hulse, J. D.; Henderson, L. M. Carnitine biosynthesis. Purification of 4-N'-trimethylaminobutyraldehyde dehydrogenase from beef liver. *J. Biol. Chem.* **1980**, 255, 1146-51.

117. Rebouche, C. J. Sites and regulation of carnitine biosynthesis in mammals. *Fed. Proc.* **1982**, 41, 2848-52.

118. Lin, S. W.; Chen, J. C.; Hsu, L. C.; Hsieh, C. L.; Yoshida, A. Human gamma-aminobutyraldehyde dehydrogenase (ALDH9): cDNA sequence, genomic

organization, polymorphism, chromosomal localization, and tissue expression. *Genomics* **1996**, 34, 376-80.

119. Kikonyogo, A.; Pietruszko, R. Aldehyde dehydrogenase from adult human brain that dehydrogenates gamma-aminobutyraldehyde: purification, characterization, cloning and distribution. *Biochem. J.* **1996**, 316 (Pt 1), 317-24.
120. Abe, T.; Takada, K.; Ohkawa, K.; Matsuda, M. Purification and characterization of a rat brain aldehyde dehydrogenase able to metabolize gamma-aminobutyraldehyde to gamma-aminobutyric acid. *Biochem. J.* **1990**, 269, 25-9.
121. Lindstedt, G.; Lindstedt, S. On the biosynthesis and degradation of carnitine. *Biochem. Biophys. Res. Commun.* **1961**, 6, 319-323.
122. Englard, S.; Blanchard, J. S.; Midelfort, C. F. Gamma-butyrobetaine hydroxylase: stereochemical course of the hydroxylation reaction. *Biochemistry* **1985**, 24, 1110-6.
123. Rigault, C.; Le Borgne, F.; Demarquoy, J. Genomic structure, alternative maturation and tissue expression of the human BBOX1 gene. *Biochim. Biophys. Acta* **2006**, 1761, 1469-81.
124. Vaz, F. M.; van Gool, S.; Ofman, R.; Ijlst, L.; Wanders, R. J. Carnitine biosynthesis: identification of the cDNA encoding human gamma-butyrobetaine hydroxylase. *Biochem. Biophys. Res. Commun.* **1998**, 250, 506-10.
125. Lindstedt, G.; Lindstedt, S.; Nordin, I. γ-Butyrobetaine hydroxylase in human kidney. *Scand. J. Clin. Lab. Invest.* **1982**, 42, 477-485.
126. Lindstedt, G.; Lindstedt, S. Cofactor requirements of gamma-butyrobetaine hydroxylase from rat liver. *J. Biol. Chem.* **1970**, 245, 4178-86.
127. Lindstedt, G.; Lindstedt, S.; Olander, B.; Tofft, M. Alpha-ketoglutarate and hydroxylation of gamma-butyrobetaine. *Biochim. Biophys. Acta* **1968**, 158, 503-5.
128. Galland, S.; Le Borgne, F.; Guyonnet, D.; Clouet, P.; Demarquoy, J. Purification and characterization of the rat liver gamma-butyrobetaine hydroxylase. *Mol. Cell. Biochem.* **1998**, 178, 163-8.
129. Paul, H. S.; Sekas, G.; Adibi, S. A. Carnitine biosynthesis in hepatic peroxisomes. Demonstration of gamma-butyrobetaine hydroxylase activity. *Eur. J. Biochem.* **1992**, 203, 599-605.
130. Lindstedt, G. Hydroxylation of gamma-butyrobetaine to carnitine in rat liver. *Biochemistry* **1967**, 6, 1271-82.
131. Kondo, A.; Blanchard, J. S.; Englard, S. Purification and properties of calf liver gamma-butyrobetaine hydroxylase. *Arch. Biochem. Biophys.* **1981**, 212, 338-46.

132. Lindstedt, G.; Lindstedt, S.; Tofft, M. Gamma-butyrobetaine hydroxylase from Pseudomonas sp AK 1. *Biochemistry* **1970**, 9, 4336-42.
133. Nechaeva, G. I.; Potashov, D. A.; Zheltikova, E. N. [[Meldonium in treatment of patients with myocardial infarction]. *Kardiologiia* **2014**, 54, 93-6.
134. Liepinsh, E.; Vilskersts, R.; Loca, D.; Kirjanova, O.; Pugovichs, O.; Kalvinsh, I.; Dambrova, M. Mildronate, an inhibitor of carnitine biosynthesis, induces an increase in gamma-butyrobetaine contents and cardioprotection in isolated rat heart infarction. *J. Cardiovasc. Pharmacol.* **2006**, 48, 314-9.
135. Dambrova, M.; Liepinsh, E.; Kalvinsh, I. Mildronate: cardioprotective action through carnitine-lowering effect. *Trends Cardiovasc. Med.* **2002**, 12, 275-9.
136. Hayashi, Y.; Kirimoto, T.; Asaka, N.; Nakano, M.; Tajima, K.; Miyake, H.; Matsuura, N. Beneficial effects of MET-88, a gamma-butyrobetaine hydroxylase inhibitor in rats with heart failure following myocardial infarction. *Eur. J. Pharmacol.* **2000**, 395, 217-24.
137. Lavie, C. J.; Milani, R. V.; Ventura, H. O. Obesity and cardiovascular disease: risk factor, paradox, and impact of weight loss. *J. Am. Coll. Cardiol.* **2009**, 53, 1925-32.
138. Liepinsh, E.; Skapare, E.; Kuka, J.; Makrecka, M.; Cirule, H.; Vavers, E.; Sevostjanovs, E.; Grinberga, S.; Pugovics, O.; Dambrova, M. Activated peroxisomal fatty acid metabolism improves cardiac recovery in ischemia-reperfusion. *Naunyn Schmiedebergs Arch. Pharmacol.* **2013**, 386, 541-50.
139. Wang, W.; Lopaschuk, G. D. Metabolic therapy for the treatment of ischemic heart disease: reality and expectations. *Expert Rev. Cardiovasc. Ther.* **2007**, 5, 1123-34.
140. Ussher, J. R.; Wang, W.; Gandhi, M.; Keung, W.; Samokhvalov, V.; Oka, T.; Wagg, C. S.; Jaswal, J. S.; Harris, R. A.; Clanachan, A. S.; Dyck, J. R.; Lopaschuk, G. D. Stimulation of glucose oxidation protects against acute myocardial infarction and reperfusion injury. *Cardiovasc. Res.* **2012**, 94, 359-69.
141. Kuka, J.; Makrecka, M.; Grinberga, S.; Pugovics, O.; Liepinsh, E.; Dambrova, M. A short-term high-dose administration of sodium pivalate impairs pyruvate metabolism without affecting cardiac function. *Cardiovasc. Toxicol.* **2012**, 12, 298-303.
142. Korb, H.; Hoeft, A.; Hunneman, D. H.; Schraeder, R.; Wolpers, H. G.; Wober, W.; Hellige, G. Changes in myocardial substrate utilisation and protection of ischemic stressed myocardium by oxfenicine [(S)-4-hydroxyphenylglycine]. *Naunyn-Schmiedeberg's Archives of Pharmacology* **1984**, 327, 70-74.

143. Blackburn, K. J.; Burges, R. A.; Gardiner, D. G.; Higgins, A. J.; Morville, M.; Page, M. G. Protection against experimental myocardial ischaemia by L-4-hydroxy-phenylglycine, a new agent which alters myocardial metabolic balance in favour of carbohydrate utilisation [proceedings]. *Br. J. Pharmacol.* **1979**, 66, 443P-444P.

144. Atkinson, L.; Bergman, G.; Jackson, N.; Jewitt, D. E.; Metcalfe, J. M. Potential protective value of enhanced myocardial carbohydrate utilisation with reduced free fatty acid uptake after UK 25842 in patients with coronary artery disease [proceedings]. *Br. J. Pharmacol.* **1979**, 66, 444P-444P.

145. Wolf, H. P.; Eistetter, K.; Ludwig, G. Phenylalkyloxirane carboxylic acids, a new class of hypoglycaemic substances: hypoglycaemic and hypoketonaemic effects of sodium 2-[5-(4-chlorophenyl)-pentyl]-oxirane-2-carboxylate (B 807-27) in fasted animals. *Diabetologia* **1982**, 22, 456-63.

146. Eistetter, K.; Wolf, H. P. Synthesis and hypoglycemic activity of phenylalkyloxiranecarboxylic acid derivatives. *J. Med. Chem.* **1982**, 25, 109-13.

147. Liepinsh, E.; Skapare, E.; Svalbe, B.; Makrecka, M.; Cirule, H.; Dambrova, M. Anti-diabetic effects of mildronate alone or in combination with metformin in obese Zucker rats. *Eur. J. Pharmacol.* **2011**, 658, 277-83.

148. Pupure, J.; Isajevs, S.; Skapare, E.; Rumaks, J.; Svirskis, S.; Svirina, D.; Kalvinsh, I.; Klusa, V. Neuroprotective properties of mildronate, a mitochondria-targeted small molecule. *Neurosci. Lett.* **2010**, 470, 100-105.

149. Sjakste, N.; Gutcaits, A.; Kalvinsh, I. Mildronate: An Antiischemic Drug for Neurological Indications. *CNS Drug Reviews* **2005**, 11, 151-168.

150. Klusa, V.; Muceniece, R.; Isajevs, S.; Isajeva, D.; Beitnere, U.; Mandrika, I.; Pupure, J.; Rumaks, J.; Jansone, B.; Kalvinsh, I.; Vinters, H. V. Mildronate enhances learning/memory and changes hippocampal protein expression in trained rats. *Pharmacol. Biochem. Behav.* **2013**, 106, 68-76.

151. Jaudzems, K.; Kuka, J.; Gutsaits, A.; Zinovjevs, K.; Kalvinsh, I.; Liepinsh, E.; Liepinsh, E.; Dambrova, M. Inhibition of carnitine acetyltransferase by mildronate, a regulator of energy metabolism. *J. Enzyme Inhib. Med. Chem.* **2009**, 24, 1269-75.

152. Oppedisano, F.; Fanello, D.; Calvani, M.; Indiveri, C. Interaction of mildronate with the mitochondrial carnitine/acylcarnitine transport protein. *J. Biochem. Mol. Toxicol.* **2008**, 22, 8-14.

153. Georges, B.; Le Borgne, F.; Galland, S.; Isoir, M.; Ecosse, D.; Grand-Jean, F.; Demarquoy, J. Carnitine transport into muscular cells. Inhibition of transport and cell growth by mildronate. *Biochem. Pharmacol.* **2000**, 59, 1357-63.

154. Spaniol, M.; Brooks, H.; Auer, L.; Zimmermann, A.; Solioz, M.; Stieger, B.; Krahenbuhl, S. Development and characterization of an animal model of carnitine deficiency. *Eur. J. Biochem.* **2001**, 268, 1876-87.

155. Henry, L.; Leung, I. K. H.; Claridge, T. D. W.; Schofield, C. J. gamma-Butyrobetaine hydroxylase catalyses a Stevens type rearrangement. *Bioorg. Med. Chem. Lett.* **2012**, 22, 4975-8.

156. Leung, I. K.; Krojer, T. J.; Kochan, G. T.; Henry, L.; von Delft, F.; Claridge, T. D.; Oppermann, U.; McDonough, M. A.; Schofield, C. J. Structural and mechanistic studies on gamma-butyrobetaine hydroxylase. *Chem. Biol.* **2010**, 17, 1316-24.

157. Tars, K.; Leitans, J.; Kazaks, A.; Zelencova, D.; Liepinsh, E.; Kuka, J.; Makrecka, M.; Lola, D.; Andrianovs, V.; Gustina, D.; Grinberga, S.; Liepinsh, E.; Kalvinsh, I.; Dambrova, M.; Loza, E.; Pugovics, O. Targeting carnitine biosynthesis: discovery of new inhibitors against gamma-butyrobetaine hydroxylase. *J. Med. Chem.* **2014**, 57, 2213-36.

158. Rydzik, A. M.; Leung, I. K.; Kochan, G. T.; McDonough, M. A.; Claridge, T. D.; Schofield, C. J. Oxygenase-catalyzed desymmetrization of N,N-dialkyl-piperidine-4-carboxylic acids. *Angew. Chem., Int. Ed.* **2014**, 53, 10925-7.

159. Rydzik, A. M.; Chowdhury, R.; Kochan, G. T.; Williams, S. T.; McDonough, M. A.; Kawamura, A.; Schofield, C. J. Modulating carnitine levels by targeting its biosynthesis - selective inhibition of [gamma]-butyrobetaine hydroxylase. *Chemical Science* **2014**, 5, 1765-1771.

160. Tars, K.; Rumnieks, J.; Zeltins, A.; Kazaks, A.; Kotelovica, S.; Leonciks, A.; Sharipo, J.; Viksna, A.; Kuka, J.; Liepinsh, E.; Dambrova, M. Crystal structure of human gamma-butyrobetaine hydroxylase. *Biochem. Biophys. Res. Commun.* **2010**, 398, 634-9.

161. Krissinel, E.; Henrick, K. Inference of Macromolecular Assemblies from Crystalline State. *J. Mol. Biol.* **2007**, 372, 774-797.

162. Lindstedt, S.; Nordin, I. Multiple forms of gamma-butyrobetaine hydroxylase (EC 1.14.11.1). *Biochem. J.* **1984**, 223, 119-27.

163. Leung, I. K. H.; Krojer, T. J.; Kochan, G. T.; Henry, L.; von Delft, F.; Claridge, T. D. W.; Oppermann, U.; McDonough, M. A.; Schofield, C. J. Structural and mechanistic studies on gamma-butyrobetaine hydroxylase. *Chem. Biol.* **2010**, 17, 1316-24.

164. Zhang, Z.; Ren, J.; Harlos, K.; McKinnon, C. H.; Clifton, I. J.; Schofield, C. J. Crystal structure of a clavaminate synthase-Fe(II)-2-oxoglutarate-substrate-NO complex: evidence for metal centered rearrangements. *FEBS Lett.* **2002**, 517, 7-12.

165. Ziering, D. L.; Pascal, R. A. Mechanism-based inhibition of bacterial .gamma.-butyrobetaine hydroxylase. *J. Am. Chem. Soc.* **1990**, 112, 834-838.

166. Rydzik, A. M.; Leung, I. K. H.; Kochan, G. T.; Loik, N. D.; Henry, L.; McDonough, M. A.; Claridge, T. D. W.; Schofield, C. J. Comparison of the substrate selectivity and biochemical properties of human and bacterial gamma-butyrobetaine hydroxylase. *Org. Biomol. Chem.* **2014**, 12, 6354-8.

167. Rydzik, A. M.; Leung, I. K.; Kochan, G. T.; Loik, N. D.; Henry, L.; McDonough, M. A.; Claridge, T. D.; Schofield, C. J. Comparison of the substrate selectivity and biochemical properties of human and bacterial gamma-butyrobetaine hydroxylase. *Org. Biomol. Chem.* **2014**, 12, 6354-8.

168. Rydzik, A. M.; Leung, I. K. H.; Kochan, G. T.; Thalhammer, A.; Oppermann, U.; Claridge, T. D. W.; Schofield, C. J. Development and application of a fluoride-detection-based fluorescence assay for gamma-butyrobetaine hydroxylase. *Chembiochem* **2012**, 13, 1559-63.

169. Suelter, C. H. Enzymes activated by monovalent cations. *Science* **1970**, 168, 789-95.

170. Nakashima, K.; Tuboi, S. Size-dependent allosteric effects of monovalent cations on rabbit liver fructose-1,6-bisphosphatase. *J. Biol. Chem.* **1976**, 251, 4315-21.

171. Evans, H. J.; Sorger, G. J. Role of Mineral Elements with Emphasis on the Univalent Cations. *Annual Review of Plant Physiology* **1966**, 17, 47-76.

172. Melchior, J. B. The role of metal ions in the pyruvic kinase reaction. *Biochemistry* **1965**, 4, 1518-25.

173. Lowenstein, J. M. The stimulation of transphosphorylation by alkali-metal ions. *Biochem. J.* **1960**, 75, 269-74.

174. Nowak, T.; Mildvan, A. S. Nuclear magnetic resonance studies of the function of potassium in the mechanism of pyruvate kinase. *Biochemistry* **1972**, 11, 2819-28.

175. Hewitson, K. S.; Granatino, N.; Welford, R. W. D.; McDonough, M. A.; Schofield, C. J. Oxidation by 2-oxoglutarate oxygenases: non-haem iron systems in catalysis and signalling. *Philos. Trans. A Math. Phys. Eng. Sci.* **2005**, 363, 807-828.

176. Myllyla, R.; Kuutti-Savolainen, E. R.; Kivirikko, K. I. The role of ascorbate in the prolyl hydroxylase reaction. *Biochem. Biophys. Res. Commun.* **1978**, 83, 441-8.

177. Jung, H.; Jung, K.; Kleber, H. P. L-carnitine metabolization and osmotic stress response in Escherichia coli. *J. Basic Microbiol.* **1990**, 30, 409-13.

178. Emaus, R. K.; Bieber, L. L. A biosynthetic role for carnitine in the yeast Torulopsis bovina. *J. Biol. Chem.* **1983**, 258, 13160-5.

179. Lindstedt, G.; Lindstedt, S.; Midtvedt, T.; Tofft, M. The Formation and Degradation of Carnitine in Pseudomonas*. *Biochemistry* **1967**, 6, 1262-1270.

180. Ruetschi, U.; Nordin, I.; Odelhog, B.; Jornvall, H.; Lindstedt, S. gamma-Butyrobetaine hydroxylase. Structural characterization of the Pseudomonas enzyme. *Eur. J. Biochem.* **1993**, 213, 1075-80.

181. Lindstedt, G.; Lindstedt, S.; Nordin, I. Purification and properties of gamma-butyrobetaine hydroxylase from Pseudomonas sp AK 1. *Biochemistry* **1977**, 16, 2181-8.

182. Parvin, R.; Pande, S. V. Microdetermination of (−)carnitine and carnitine acetyltransferase activity. *Anal. Biochem.* **1977**, 79, 190-201.

183. Böhmer, T.; Rydning, A.; Solberg, H. E. Carnitine levels in human serum in health and disease. *Clin. Chim. Acta* **1974**, 57, 55-61.

184. Cederblad, G.; Lindstedt, S. A method for the determination of carnitine in the picomole range. *Clin. Chim. Acta* **1972**, 37, 235-243.

185. Pearson, D.; Chase, J.; Tubbs, P.; Lowenstein, J. Methods in Enzymology. In Academic Press, New York and London: **1969**.

186. Healy, J.; Ekkerman, S.; Pliotas, C.; Richard, M.; Bartlett, W.; Grayer, S. C.; Morris, G. M.; Miller, S.; Booth, I. R.; Conway, S. J.; Rasmussen, T. Understanding the structural requirements for activators of the Kef bacterial potassium efflux system. *Biochemistry* **2014**, 53, 1982-92.

187. Cametti, M.; Rissanen, K. Recognition and sensing of fluoride anion. *Chemical Communications* **2009**, 2809-2829.

188. Leung, I. K. H. NMR studies of 2-oxoglutarate oxygenases. PhD Thesis, University of Oxford, UK, **2012**.

189. Poppe, L.; Tegley, C. M.; Li, V.; Lewis, J.; Zondlo, J.; Yang, E.; Kurzeja, R. J.; Syed, R. Different modes of inhibitor binding to prolyl hydroxylase by combined use of X-ray crystallography and NMR spectroscopy of paramagnetic complexes. *J. Am. Chem. Soc.* **2009**, 131, 16654-5.

190. Bertini, I.; Luchinat, C.; Parigi, G.; Pierattelli, R. NMR spectroscopy of paramagnetic metalloproteins. *Chembiochem* **2005**, 6, 1536-49.
191. Jahnke, W. Spin labels as a tool to identify and characterize protein-ligand interactions by NMR spectroscopy. *Chembiochem* **2002**, 3, 167-73.
192. Jahnke, W.; Rudisser, S.; Zurini, M. Spin label enhanced NMR screening. *J. Am. Chem. Soc.* **2001**, 123, 3149-50.
193. Bertini, I.; Luchinat, C.; Parigi, G. *Solution NMR of paramagnetic molecules : applications to metallobiomolecules and models*. Elsevier Science Ltd.: Amsterdam, **2001**.
194. Kosen, P. A. Spin labeling of proteins. *Methods Enzymol.* **1989**, 177, 86-121.
195. Burton, D. R.; Forsen, S.; Karlstrom, G.; Dwek, R. A. Proton relaxation enhancement (PRE) in biochemistry: A critical survey. *Prog. Nucl. Magn. Reson. Spectrosc.* **1979**, 13, 1-45.
196. Mohr, P. J.; Taylor, B. N.; Newell, D. B. CODATA recommended values of the fundamental physical constants: 2006*. *Reviews of Modern Physics* **2008**, 80, 633-730.
197. Bertini, I.; Fragai, M.; Lee, Y. M.; Luchinat, C.; Terni, B. Paramagnetic metal ions in ligand screening: the Co(II) matrix metalloproteinase 12. *Angew. Chem., Int. Ed.* **2004**, 43, 2254-6.
198. Leung, I. K. H.; Flashman, E.; Yeoh, K. K.; Schofield, C. J.; Claridge, T. D. W. Using NMR solvent water relaxation to investigate metalloenzyme-ligand binding interactions. *J. Med. Chem.* **2010**, 53, 867-75.
199. Bertini, I.; Fragai, M.; Luchinat, C.; Talluri, E. Water-based ligand screening for paramagnetic metalloproteins. *Angew. Chem., Int. Ed.* **2008**, 47, 4533-7.
200. Krugh, T. R. *Spin-label-induced nuclear magnetic resonance relaxation studies of enzymes*. Academic Press: New York, **1976**; Vol. 1, p 339-372.
201. Aik, W.; Scotti, J. S.; Choi, H.; Gong, L.; Demetriades, M.; Schofield, C. J.; McDonough, M. A. Structure of human RNA N(6)-methyladenine demethylase ALKBH5 provides insights into its mechanisms of nucleic acid recognition and demethylation. *Nucleic Acids Res.* **2014**, 42, 4741-54.
202. Chowdhury, R.; Candela-Lena, J. I.; Chan, M. C.; Greenald, D. J.; Yeoh, K. K.; Tian, Y. M.; McDonough, M. A.; Tumber, A.; Rose, N. R.; Conejo-Garcia, A.; Demetriades, M.; Mathavan, S.; Kawamura, A.; Lee, M. K.; van Eeden, F.; Pugh, C.

W.; Ratcliffe, P. J.; Schofield, C. J. Selective small molecule probes for the hypoxia inducible factor (HIF) prolyl hydroxylases. *ACS Chem. Biol.* **2013**, 8, 1488-96.

203. McDonough, M. A.; Li, V.; Flashman, E.; Chowdhury, R.; Mohr, C.; Lienard, B. M.; Zondlo, J.; Oldham, N. J.; Clifton, I. J.; Lewis, J.; McNeill, L. A.; Kurzeja, R. J.; Hewitson, K. S.; Yang, E.; Jordan, S.; Syed, R. S.; Schofield, C. J. Cellular oxygen sensing: Crystal structure of hypoxia-inducible factor prolyl hydroxylase (PHD2). *Proc. Natl. Acad. Sci. U.S.A.* **2006**, 103, 9814-9.
204. Aik, W.; Demetriades, M.; Hamdan, M. K.; Bagg, E. A.; Yeoh, K. K.; Lejeune, C.; Zhang, Z.; McDonough, M. A.; Schofield, C. J. Structural basis for inhibition of the fat mass and obesity associated protein (FTO). *J. Med. Chem.* **2013**, 56, 3680-8.
205. King, O. N.; Li, X. S.; Sakurai, M.; Kawamura, A.; Rose, N. R.; Ng, S. S.; Quinn, A. M.; Rai, G.; Mott, B. T.; Beswick, P.; Klose, R. J.; Oppermann, U.; Jadhav, A.; Heightman, T. D.; Maloney, D. J.; Schofield, C. J.; Simeonov, A. Quantitative high-throughput screening identifies 8-hydroxyquinolines as cell-active histone demethylase inhibitors. *PLoS One* **2010**, 5, e15535.
206. Nangaku, M.; Izuhara, Y.; Takizawa, S.; Yamashita, T.; Fujii-Kuriyama, Y.; Ohneda, O.; Yamamoto, M.; van Ypersele de Strihou, C.; Hirayama, N.; Miyata, T. A novel class of prolyl hydroxylase inhibitors induces angiogenesis and exerts organ protection against ischemia. *Arterioscler. Thromb. Vasc. Biol.* **2007**, 27, 2548-54.
207. Leung, I. K. H.; Demetriades, M.; Hardy, A. P.; Lejeune, C.; Smart, T. J.; Szollossi, A.; Kawamura, A.; Schofield, C. J.; Claridge, T. D. W. Reporter ligand NMR screening method for 2-oxoglutarate oxygenase inhibitors. *J. Med. Chem.* **2013**, 56, 547-55.
208. Nagy, G. N.; Marton, L.; Contet, A.; Ozohanics, O.; Ardelean, L. M.; Revesz, A.; Vekey, K.; Irimie, F. D.; Vial, H.; Cerdan, R.; Vertessy, B. G. Composite aromatic boxes for enzymatic transformations of quaternary ammonium substrates. *Angew. Chem., Int. Ed.* **2014**, 53, 13471-6.
209. Salonen, L. M.; Ellermann, M.; Diederich, F. Aromatic rings in chemical and biological recognition: energetics and structures. *Angew. Chem., Int. Ed.* **2011**, 50, 4808-42.
210. Taverna, S. D.; Li, H.; Ruthenburg, A. J.; Allis, C. D.; Patel, D. J. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat. Struct. Mol. Biol.* **2007**, 14, 1025-40.

211. Gallivan, J. P.; Dougherty, D. A. Cation-pi interactions in structural biology. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, 96, 9459-64.

212. Ma, J. C.; Dougherty, D. A. The Cationminus signpi Interaction. *Chem. Rev.* **1997**, 97, 1303-1324.

213. Dougherty, D. A.; Stauffer, D. A. Acetylcholine binding by a synthetic receptor: implications for biological recognition. *Science* **1990**, 250, 1558-60.

214. Lester, H. A.; Dibas, M. I.; Dahan, D. S.; Leite, J. F.; Dougherty, D. A. Cys-loop receptors: new twists and turns. *Trends Neurosci.* **2004**, 27, 329-36.

215. Dougherty, D. A. The cation-pi interaction. *Acc. Chem. Res.* **2013**, 46, 885-93.

216. Kamps, J. J. A. G.; Khan, A.; Choi, H.; Lesniak, R. K.; Brem, J.; Rydzik, A. M.; McDonough, M. A.; Schofield, C. J.; Claridge, T. D. W.; Mecinović, J. Cation–π Interactions Contribute to Substrate Recognition in γ-Butyrobetaine Hydroxylase Catalysis. *Chem. Eur. J.* **2016**, 1270-1276.

217. Salonen, L. M.; Holland, M. C.; Kaib, P. S.; Haap, W.; Benz, J.; Mary, J. L.; Kuster, O.; Schweizer, W. B.; Banner, D. W.; Diederich, F. Molecular recognition at the active site of factor Xa: cation-pi interactions, stacking on planar peptide surfaces, and replacement of structural water. *Chemistry (Easton)* **2012**, 18, 213-22.

218. Scharer, K.; Morgenthaler, M.; Paulini, R.; Obst-Sander, U.; Banner, D. W.; Schlatter, D.; Benz, J.; Stihle, M.; Diederich, F. Quantification of cation-pi interactions in protein-ligand complexes: crystal-structure analysis of Factor Xa bound to a quaternary ammonium ion ligand. *Angew. Chem., Int. Ed.* **2005**, 44, 4400-4.

3 Non-enzymatic 2-oxoglutarate decarboxylation by ascorbate degradation

3.1 L-Ascorbic acid as co-factor/activity stimulator for 2OG oxygenases

A large number of 2OG and Fe^{II} dependent oxygenases in both plants and animals require L-ascorbic acid or simply ascorbate (Figure 3.1) for maximum enzymatic activities both in *in vitro* and *in vivo* environments (Table 3.1).[1-4] The stimulatory effect of L-ascorbic acid has also been observed for 2OG oxygenases originating from microorganisms that do no produce L-ascorbate.[5]

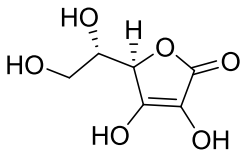


Figure 3.1 Structure of L-ascorbate

It is proposed that ascorbate serves as reducing agent to these enzymes, by maintaining a pool of Fe^{II} (by reducing Fe^{III} to Fe^{II}) available to the enzyme for catalysis, and hence promotes maximum catalytic activity. Studies have shown that in the 2OG oxygenase catalysed hydroxylation reactions, the rate of hydroxylation is similar in the linear region in both in the presence and absence of L-ascorbate; however in the absence of L-ascorbate after certain time the rate of hydroxylation then suddenly and markedly deviates from linearity to become negligible.[1-4] This significant reduction in the rate of hydroxylation by 2OG oxygenases in the absence of L-ascorbate is generally believed to be due to the inactivation of enzyme in the absence of Fe^{II} and thus L-ascorbate is used as reducing agent to reduce Fe^{III} back to the catalytic Fe^{II} and promote maximum enzyme activity. Since, no such active mechanism is known that transports iron into the active sites of 2OG oxygenases, it is therefore unknown if L-ascorbate promotes the catalytic activity by reducing the enzyme bound Fe^{III} to Fe^{II}, or it reduces the free Fe^{III} in solution to Fe^{II}. To date, no crystal structure of a 2OG oxygenase in complex with L-ascorbic acid has been reported that could suggest enzyme activation by ascorbate induced conformational changes. If the role of L-ascorbate was only to maintain a pool of Fe^{II}, then a strong reducing agent such as dithiothreitol (DTT) (Figure 3.2) which has more negative standard reduction potential as compared with L-ascorbate; -0.33 V[6] vs +0.28 V[7], would be more capable of reducing Fe^{III} into Fe^{II} in solution and hence would result in

significant activation of 2OG oxygenases activity compared with L-ascorbate. However, except one example (AlkB, a bacterial 2OG oxygenase)[8] no such study has so far demonstrated that any other reducing agent can replace L-ascorbate for promoting maximum 2OG oxygenase activity. The stimulation of 2OG oxygenases activity in the presence of L-ascorbate appears to be a complex process, the precise mechanism of which still remains unknown.

Table 3.1 List of the important human and plant 2OG oxygenases that require L-ascorbic acid as reducing agent/co-factor/activity stimulator (Reproduced from Reference [9])

Enzyme	Physiological Role	Enzymatic Activity	Metal Ion Centre
Deacetoxycephalosporin C synthetase	Antiobiotic metabolism (fungi)	Penicillin N to deacetylcephalosporin	Iron
γ-Butyrobetaine-2-oxoglutarate-4-dioxygenase E.C. 1.14.11.1	Carnitine biosynthesis	Hydroxylation of butyrobetaine to carnitine	Iron
Gibberellin-3-β-dioxygenase E.C. 1.14.11.15	Gibberin (plant hormone) biosynthesis	C_{20} oxidation decarboxylation and activation of gibberellins	Iron
Lysine hydroxylase E.C. 1.14.11.4	Collagen biosynthesis (animals) & Extensin biosynthesis (plants)	Hydroxylation of Lysine	Iron
Procollagen proline 2-oxoglutarate-3-dioxygenase E.C. 1.14.11.7	Procollagen biosynthesis (animals) Extensin biosynthesis (plants)	Hydroxylation of proline (3-hydroxylating)	Iron
Proline hydroxylase E.C. 1.14.11.2	Procollagen synthesis (animals)	Hydroxylation of proline (4-hydroxylating)	Iron
Pyrimidine deoxynucleoside 2'-dioxygenase E.C. 1.14.11.3	Pyrimidine metabolism (fungi)	Deoxyuridine to uridine	Iron
Thymine dioxygenase E.C. 1.14.11.6	Pyrimidine metabolism (fungi)	7-Hydroxylation of thymine	Iron
Trimethyllysine 2-oxoglutarate dioxygenase E.C. 1.14.11.8	Carnitine biosynthesis	Hydroxylation of trimethyl lysine	Iron

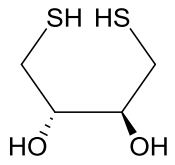


Figure 3.2 Structure of Dithiothreitol (DTT)

The stimulatory effect of some of 2OG oxygenases by L-ascorbic acid in the assay solution is discussed below,

3.1.1 Stimulatory effect of L-ascorbate on collagen hydroxylases

In animals after collagen (the major structural protein in the extracellular matrix of connective tissues) is biosynthesised, its post-translational modification involves the hydroxylation of some proline and lysine residues on the growing peptide chain.[10] Hydroxyproline and hydroxylysine are essentially required for the formation of a stable extracellular matrix and cross-linkage in the fibres. Its absence results in structurally unstable collagen. Hydroxylysine is also necessary for formation of the intermolecular crosslinks in collagen. These hydroxylation reactions are catalysed by prolyl (prolyl 4-hydroxylases, prolyl 3-hydroxylases, EC 1.14.11.2) and lysyl hydroxylases (EC 1.14.11.4) respectively, both of which belong to the Fe^{II} and 2OG dependent oxygenases super family of enzymes.[11-14] In the *in vitro* assays, it has been observed that these enzymes require L-ascorbic acid for maximum catalytic activity.[13-16] The formation of under hydroxylated collagen due to the lack of L-ascorbic acid has also been demonstrated in cell culture studies using ascorbate-deficient growth medium.[17-21] The direct involvement of ascorbic acid in collagen biosynthesis is well established and perhaps it is the most established biochemical role of this vitamin. In animals, the conditions of defective connective tissues (a symptom of scurvy), the biochemical basis of which have been known to be the impaired collagen biosynthesis, have been traced back to the dietary vitamin C deficiency. Collagens biosynthesis that may occur in the absence of L-ascorbic acid usually results abnormal fibres, causing in skin lesions, blood-vessel fragility, etc. Based on this biochemical link of L-ascorbic acid with collagen biosynthesis *in vivo*, it is possible that this link may be the activity dependency of these prolyl and lysyl hydroxylases (and possibly other enzymes) on L-ascorbic acid. L-ascorbate analogues/other reducing agents, including, its stereoisomer D-ascorbic acid (Figure 3.3), various reduced pteridines, a number of thiol compounds including DTT, cysteine, NADH2 and NADPH2 have been tested as alternative reducing agents, however ascorbic acid appears to be the most effective reductant for these enzymes.[22-25] The observation that some of these alternative reducing agents did increased the catalytic activity of these enzymes though to significantly low level as compared to L-ascorbate, the possibility existed that reductants other than vitamin C may perform this function *in vivo* and that impaired collagen synthesis in scurvy might imply a role for ascorbic acid in collagen metabolism other than one in hydroxylation. There appears to be no evidence, however, for the other reducing agents to replace L-ascorbic acid in the hydroxylation mechanism *in vivo*.

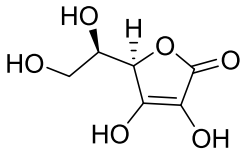


Figure 3.3 Structure of D-ascorbate

Several suggestions have been made regarding the possible role of ascorbate in the prolyl hydroxylase reactions,[13-16] but none of these have so far been proven. The exact mode of action of ascorbate in these hydroxylations has yet to be established.

If L-ascorbic acid was to act as a reducing agent by maintaining a pool of Fe^{II} to the enzyme, then in that case ascorbic acid should be consumed in stoichiometric proportions i.e one molecule of ascorbic acid should reduce two atoms of Fe^{III} to Fe^{II}. However, no such studies have proved this yet.[3,26-28] Both prolyl and lysyl hydroxylases can catalyse their reaction in complete absence of ascorbate at an essentially maximal rate for 5 to 15 s, corresponding to 15 to 30 reaction cycles, however after this time the rate of reaction dramatically decrease.[27,29,30] In a study by Raili *et al.*[30] in monitoring prolyl hydroxylase reaction, L-ascorbate was added to the reaction mixture after 15 s and it was observed that the rate of hydroxylation was suddenly restored with L-ascorbate addition to similar level to that in the presence of L-ascorbate, suggesting that the enzyme inactivation in the absence of L-ascorbate was not at least irreversible. To investigate if the enzyme activation was due to increase in concentration of free Fe^{II} species in solution (produced by reduction of Fe^{III} into Fe^{II} by L-ascorbic acid), they studied the oxidation of Fe^{II} into Fe^{III} in the absence of both enzyme and L-ascorbate, and it was found that only 1-2% of free Fe^{II} is oxidised into Fe^{III} per min. This suggests that in the absence of L-ascorbate, the enzyme inactivation is not solely and primarily due to the decrease in the concentration of free Fe^{II} in solution, because L-ascorbate was clearly required after 15 s, when only 0.25-0.5% of free Fe^{II} had become oxidised and about 99.5% or more still remain as Fe^{II}. This leaves the possibility that L-ascorbate may perhaps be reducing the enzyme-bound Fe^{III} to Fe^{II} in a specific manner.

Both prolyl and lysyl hydroxylases have also been reported to catalyse the uncoupled 2OG decarboxylation in the absence of the peptide substrate.[14,27,28,31] Some studies have shown that the rate of this uncoupled 2OG turnover increases in the presence of certain substrate competitive inhibitors with ascorbate.[31,32] It has been shown that unlike the complete reaction, this uncoupled 2OG decarboxylation involves stoichiometric L-ascorbate

consumption.[33,34] The mechanism of this 1:1 stoichiometry of uncoupled 2OG decarboxylation and L-ascorbate oxidation still remains unknown.

3.1.2 Stimulatory effect of L-ascorbate on HIF hydroxylases

HIF (hypoxia inducible factor) plays an important role in animal's response to low oxygen level (hypoxia).[35,36] In human cells, the levels of HIF-α domains are regulated in the presence of oxygen by four enzymes, called HIF hydroxylases.[37-40] These enzymes belong to the 2OG oxygenases super-family of enzymes. In normoxia, the prolyl hydroxylase domains 1-3 (PHDs 1-3) catalyse hydroxylation of two prolyl residues in HIF-α (in HIF-1α, Pro^{402}, Pro^{564}) in the N- and C-terminal ODDs (oxygen-dependent degradation domains) (NODD) and (CODD) respectively preparing them for subsequent ubiquitination and degradation by the proteasome.[41,42]

Studies have shown that PHD2 activity is significantly increased in the presence of L-ascorbic acid in isolated proteins,[1,41,43] insect cell extracts[44] and in cultured mammalian cells.[45]

FIH (factor inhibiting HIF) catalyses asparaginyl hydroxylation (Asn^{803}) in the CAD (C-terminal activation domain) of HIF-1α. This hydroxylation is a modification which is believed to stop the interaction of HIF-1α with the co-transcriptional activator, p300.[46] Studies have also shown that the catalytic activity of FIH also significantly increases in the presence of L-ascorbic acid.[1] FIH is also a Fe^{II} and 2OG dependent oxygenase that plays a key role in adaptation to low oxygen availability.

The role of L-ascorbic acid on the activities of both these enzymes (PHD2, FIH) has been proposed to be *via* the completion of uncoupled turnover cycles or *via* maintaining a pool of catalytic Fe^{II}.[45,47]

Work by Flashman *et al.*[1] have shown that neither DTT nor glutathione (GSH) (at 4 mM) was able to fully replace L-ascorbic acid in promoting PHD2 hydroxylating activity towards either CODD or NODD peptide fragments substrates of HIF-1α protein. They showed that ascorbate (at 4 mM) increased the level of CODD hydroxylation from 15% (in the absence of ascorbate) to 64%, whereas, the level of NODD hydroxylation was increased from 2% to 63%. DTT partially stimulated PHD2 hydroxylation of CODD to 48% compared with 64% of L-ascorbic acid. In case of NODD, the stimulatory effect by DTT was however lower than CODD, to 26% compared with 63% of L-ascorbic acid. Similarly, glutathione was also found to promote PHD2-catalysed CODD hydroxylation, although to a lesser extent compared with

DTT, (27%, as compared with 15% without L-ascorbic acid, and 17% with oxidized glutathione). Glutathione however, did not stimulate the NODD hydroxylation at all. Their study has shown that the dependence of FIH on L-ascorbic acid was lower than the dependence of PHD2. The level of hydroxylation increases only from 19% to 36% by the addition of L-ascorbate. In contrast to PHD2, replacement of L-ascorbate with either DTT or glutathione do not result in an increase in FIH-catalysed HIF hydroxylation relative to levels seen without L-ascorbate (22% and 12% respectively, compared with 19% without L-ascorbate). These results suggest that L-ascorbic acid possibly interacts with both these different enzymes *via* different modes/mechanism. It is therefore possible that specific reducing agents can be employed to selectively activate certain HIF hydroxylases, or indeed 2OG oxygenases.

Their study[1] of L-ascorbic acid analogues (Figure 3.4) have shown that both D-isoascorbate and iso-propylidine-L-ascorbate (IPPA) fully replace L-ascorbate in stimulating both PHD2 and FIH catalytic activities. These results suggest that at least a highly specific binding pocket for L-ascorbic acid in both these different enzymes is highly unlikely. To investigate the role of 2 and 3 -OH groups (ene-diol group of L-ascorbate) in enzyme stimulation, L-gulonic-γ–lactone, which has a *cis* 1,2-diol group but lacks the double bond, a sulfate analogue and dehydro-L-ascorbic acid had been tested but none of them had shown to stimulate the activities of either PHD2 or FIH,[1] suggesting that ene-diol part of the L-ascorbate appears to play a critical role in stimulating PHD2 and FIH catalytic activity. Similarly, 3,4-dihydroxyphenylacetic acid contains an ene-diol group within a planar ring, but lacks the lactone functionality and side chain, also do not stimulate the activity of both PHD2 and FIH. The exact mechanism on how L-ascorbate preferentially activates both PHD2 and FIH remains still unknown.

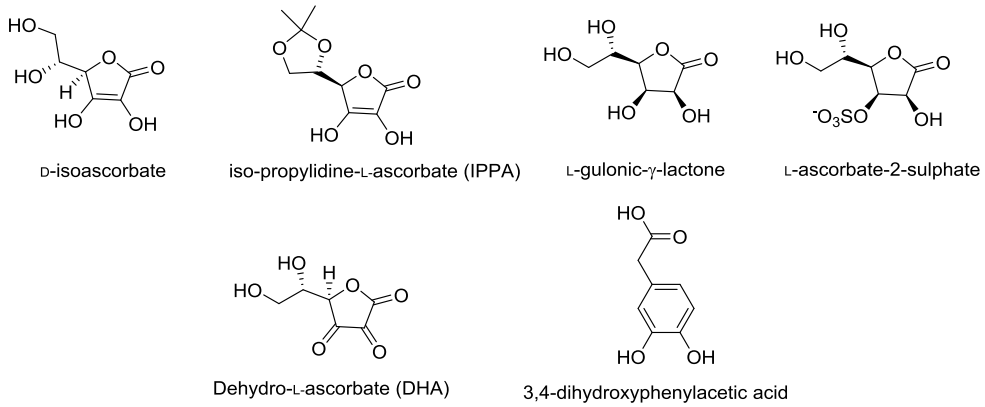
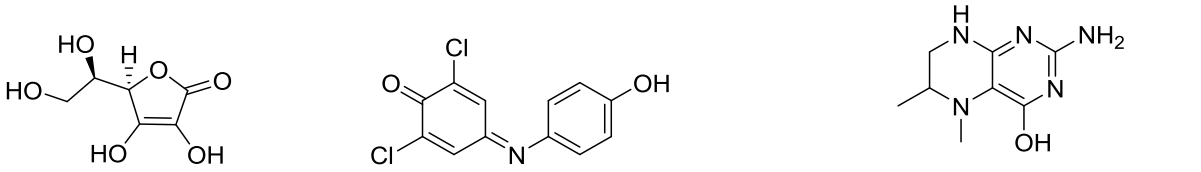


Figure 3.4 Structure of L-ascorbate analogues that had been tested for their stimulatory effect of PHD2 and FIH enzymes.

3.1.3 Stimulatory effect of L-ascorbate on 2OG oxygenases involved in L-carnitine biosynthesis pathway

L-ascorbic acid is also an essential co-factor for two human 2OG and Fe^{II} dependent oxygenases, N^{ϵ}-trimethyllysine hydroxylase (TMLH) and γ-butyrobetaine hydroxylase (BBOX) involved in the L-carnitine biosynthetic pathways in animals.[48-54]

Studies by Holme *et al.*[55] suggest that during the *in vitro* hydroxylation of GBB by human kidney BBOX, significant level of uncoupled 2OG decarboxylation occurred in the presence of L-ascorbate. Similarly, when human BBOX was incubated with 2OG in the absence of GBB, but in the presence of added D-carnitine, the level of uncoupling almost become doubled, from 20% in the first experiment to 36% in the last. This un-coupled 2OG decarboxylation was also observed when BBOX from *pseudomonas* was used, but the degree of uncoupling was considerably less. To these observations they had no obvious explanation. In another study on BBOX, isoascorbate (Figure 3.5) has been shown to be an effective stimulator, while, other reductants including 2,6-dichlorophenolindolphenols (DCPIP) and 2-amino-5,6-dimethyl-4-hydroxy-5,6,7,8-tetrahydropteredine (Figure 3.5) have been shown as less effective stimulator.[54]



Isoascorbate | 2,6-dichlorophenolindolphenols (DCPIP) | 2-amino-5,6-dimethyl-4-hydroxy-5,6,7,8-tetrahydropteredine

Figure 3.5 Strcuture reported reducing agents that had been tested for their stimulatory effect BBOX.

For N^{ϵ}-trimethyllysine hydroxylase, other reductants including NADH, NADP and DTT simply cannot replace L-ascorbate for maximum enzymatic activity.[50,51] Although the exact mechanism by which L-ascorbate stimulates the activities of both these enzymes remains unknown, but it is believed that L-ascorbate reduces the Fe^{III} into catalytic Fe^{II}, thus maintains a pool of Fe^{II} to the enzyme and hence promote maximum catalytic activity. The present knowledge concerning the *in vivo* functions of L-ascorbate in carnitine biosynthesis suggests that ascorbate deficiency in guinea pigs results in a variable decrease in several tissues carnitine levels.[56] The reduced activity of ascorbate-deficient guinea pigs liver BBOX can be restored rapidly by injection of ascorbate. Similarly, *in vivo*, ascorbate deficiency in guinea

pigs appears to diminish N^{ε}-trimethyllysine hydroxylase activity in kidney but not in liver. It is possible that the low carnitine level may be due the dependence of activity of both BBOX and TMLH on L-ascorbic acid.

3.1.4 Stimulatory effect of L-ascorbate on TET 2OG oxygenases

The Ten-eleven translocation (Tet) family of proteins was discovered in 2009.[57] These enzymes belong to the super family of 2OG and Fe^{II} dependent oxygenases.[57-60] Tet enzymes catalyse the oxidation of 5-methylcytosine (5mC) in the DNA strands (Figure 3.6).[57,59] Studies on these proteins in the past several years have established that 5-hydroxymethylcytosine (5hmC) is an intermediate in DNA demethylation process and that Tet proteins have important biological roles in epigenetic reprogramming in early embryos and primordial germ cells dioxygenases for the oxidation of 5-methylcytosine (5mC).[61] 5hmC undergoes further oxidation by Tet proteins to form 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC), which can be excised by thymine DNA glycosylase (TDG) followed by the reintroduction of unmethylated cytosine *via* the base-excision repair (BER) pathway.[59,60]

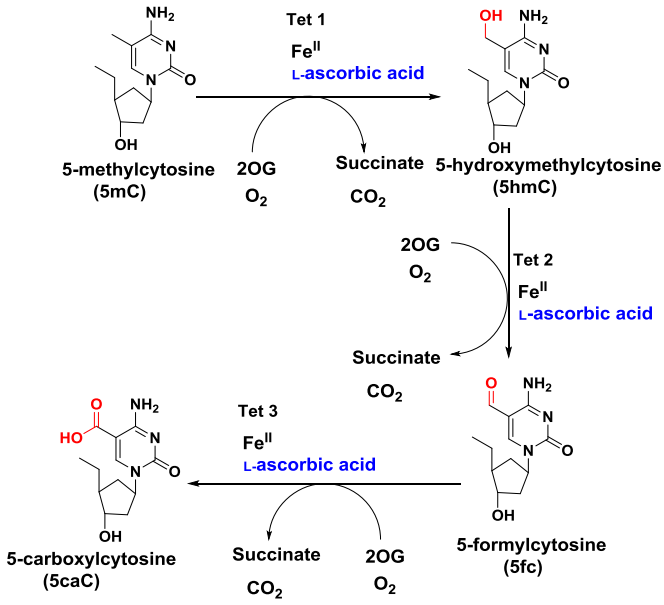


Figure 3.6 Reactions catalysed by Tet 2OG oxygenases in DNA strands.

It has been found that the activity of Tet 2OG oxygenases both in *in vitro* and *in vivo* assays is specifically and highly dependent on the presence of L-ascorbic acid in the assay mixture.[61-64] Studies by Ruichuan *et al.* have shown that in a dose-dependent manner, L-ascorbate (50-500 μM) increases the level of formation of 5hmC and 5fC by 4.0−7.0-fold and

by 4.6–8.9-fold respectively.[61] Their reaction time courses showed that in the presence of L-ascorbate the reaction rates of formation of 5hmC and 5fC increased by 8.2- and 6.2-fold. D-ascorbic acid was also as good as ascorbate in promoting the 5fc formation (3.8 fold increase). Other compounds including ascorbate analogues L-ascorbate-2-phosphate and L-ascorbate-2-sulfate and strong reducing agents including spermidine, vitamin B1, vitamin E, glutathione, NADPH, and L-cysteine cannot replace L-ascorbic acid for Tet catalysed oxidation reactions. They showed that L-ascorbic acid binds to the C-terminal catalytic domain of Tet2 enzyme as evidenced by changes in intrinsic fluorescence. They proposed that the binding of L-ascorbic acid to Tet2 may be either inducing structural changes in the enzyme to catalytically active form or it may be acting as a site specific reducing agent by recycling the Fe^{III} to catalytic Fe^{II}.

The regulation of DNA 5-hmC activity by L-ascorbic acid is a novel biochemical function of this vitamin, suggesting its role in modulating the epigenetic control of genome activity. The impaired generation of 5-hmC has also been recently associated with cancer.[65,66] This area of research may help to develop novel therapies for these conditions by rescuing the impaired generation of 5-hmC using ascorbate treatments in the future.

3.1.5 Stimulatory effect of L-ascorbate on Anthocyanidin synthase (ANS)

Anthocyanidin synthase (ANS), is a plant Fe^{II} and 2OG dependent oxygenase[67,68] that catalyses the second last step in the biosynthesis of the anthocyanin class of flavonoids (Figure 3.7). Flavonoids are common colorants in plants that are widely distributed throughout plant body and have long-established biomedicinal properties.[69,70] Flavonoids produce yellow or red/blue pigmentation in petals designed to attract pollinator animals. In higher plants, flavonoids are involved in UV filtration, symbiotic nitrogen fixation and floral pigmentation.

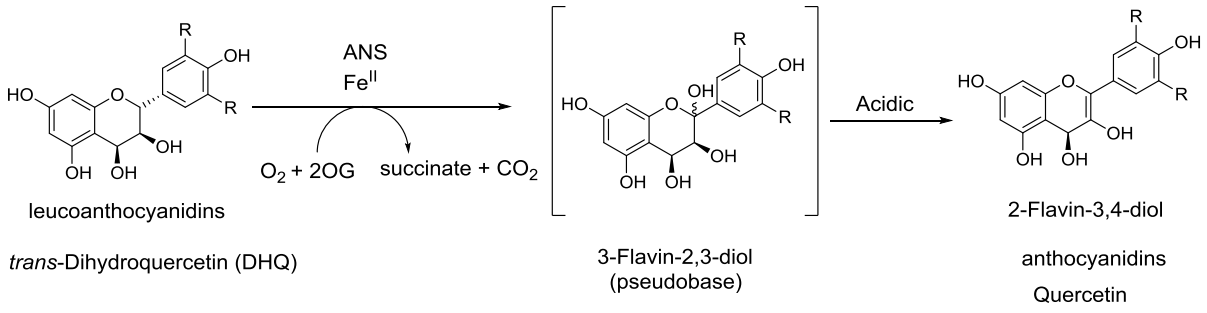


Figure 3.7 Reaction catalysed by ANS.

In *in vitro* studies have shown that ANS requires an unusually high concentration of ascorbate (40 mM) for optimal turnover.[67,68,71] In the absence of ascorbate, ANS has only 5% of full activity in standard *in vitro* assay conditions, suggesting a key role for L-ascorbic acid in catalysis.[71] Based on the presence of MES or 2-(*N*-morpholino)ethanesulfonic acid (Figure 3.8) buffer molecule in the active site of ANS in the holo-form crystal structure, it has been proposed that the MES binding site could be the L-ascorbic acid binding site, with its morpholino oxygen atom making interaction with the C-3 hydroxy group (2.71 Å) of quercetin and during crystallization MES may have displaced L-ascorbic acid.[71] It is possible that one of the enol hydroxyl groups of ascorbate can act as an acid or base through its interaction with the C-4 hydroxyl of dihydroquercetin (DHQ-1) or the C-3/C-4 hydroxyls of the leucoanthocyanidins as observed in the proposed role of ascorbate in the catalytic mechanism of myrosinase.[72] In the apo-crystal structure of ANS, the MES molecule was not found in the active site, but was found at a remote site away from the active site, suggesting that the binding of the MES in the active site was dependent on substrate binding.

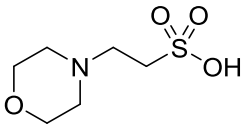


Figure 3.8 Structure of 2-(*N*-morpholino)ethanesulfonic acid or MES.

3.1.6 Stimulatory effect of L-ascorbate on AlkB 2OG oxygenase

AlkB is one of four proteins (AlkA, AidB, and Ada) involved in the adaptive response to DNA alkylation damage in *Escherichia coli*. [73] AlkB is highly conserved from bacteria to humans.[74] AlkB belongs to the Fe^{II} and 2OG-dependent oxygenase super-family of enzymes[75,76] which repairs the methylated DNA by direct demethylation of 1-methyladenine and 3-methylcytosine lesions (Figure 3.9). The demethylation proceeds *via* hydroxylation (Figure 3.9).

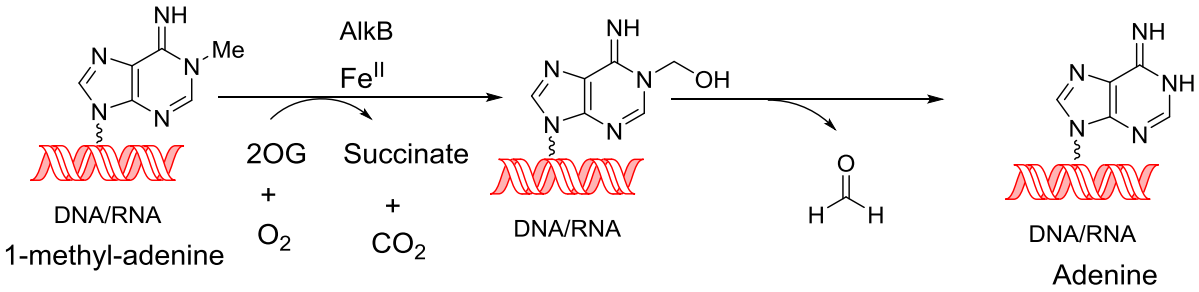


Figure 3.9 Oxidative demethylation of DNA/RNA 1-methyl-adenine by AlkB.

Welford *et al.* studied the catalysis of AlkB *in vitro* using 1-[^{14}C]2OG and measuring $^{14}CO_2$ released from succinate formation.[8] They found that in the presence of L-ascorbic acid and the nucleoside substrates analogues, including 1-methyladenosine, 1-methyl-2′-deoxyadenosine, 3-methylcytidine, and 3-methyl-2′-deoxycytidine (Figure 3.10) in the assay mixture, there is a stimulation of uncoupled 2OG turnover by AlkB. In these reactions, the prime substrate analogue completely fails to hydroxylate. In the absence of substrate/substrate analogues, only a low level of uncoupled 2OG turnover occurs, whereas, in case of 1-methyladenosine and 1-methyl-2′-deoxyadenosine, the uncoupled 2OG turnover was significantly increased, up to 7 fold, whereas, in case of 3-methylcytidine and 3-methyl-2′-deoxycytidine, the increase in the uncoupled 2OG turnover was 2 and 3 folds respectively.

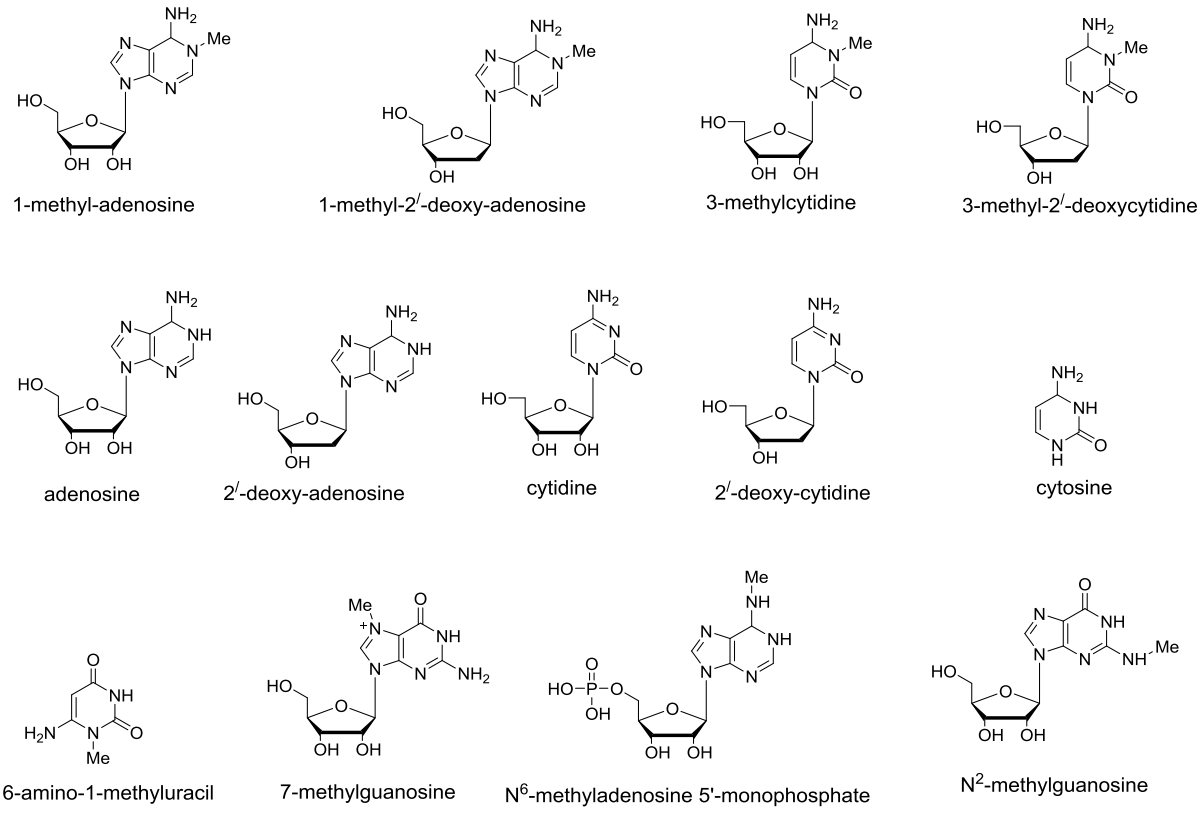


Figure 3.10 Structure of the AlkB substrate analogues that stimulate the uncoupled 2OG decarboxylation in the presence of L-ascorbic acid.

This uncoupled 2OG turnover by AlkB was found to be highly dependent on ascorbate both in the presence and absence of 1-methyladenosine, but more so in the former case. This uncoupled 2OG turnover increased by using upto 500 µM ascorbate. L-Ascorbate concentration above 500 µM (4 mM) made no influence on the uncoupled 2OG turnover. They also

observed that the amount of ascorbate required for optimal uncoupled 2OG turnover was far in excess of stoichiometry to 2OG.

In the presence of other nucleoside substrate analogues such as adenosine, 2′-deoxyadenosine and adenine (in the absence of the 1-methyl group), the rate of uncoupled 2OG turnover was similar to the rate observed in the absence of substrate (significantly low). The observation that in the presence of 1-methyladenine (AlkB substrate), 1-methyladenosine and 1-methyl-2′-deoxyadenosine, there was a significant increase in the rate of uncoupled 2OG turnover, indicate that the 1-methyl group is an absolute requirement for stimulation of the uncoupled 2OG turnover. In case of 1-methyladenosine and 1-methyl-2′-deoxyadenosine, the uncoupled 2OG conversion was higher compared to substrate 1-methyladenine, suggesting that the presence of a sugar moiety improves substrate recognition by AlkB, perhaps by direct contacts between the enzyme and the sugar giving rise to tighter enzyme/substrate analogue binding. Cytidine, 2′-deoxycytidine, cytosine, 6-amino-1-methyluracil, 7-methylguanosine, N^6-methyladenosine-5′-monophosphate and N^2-methylguanosine (Figure 3.10) all resulted in only a negligible increase in the level of uncoupled 2OG conversion, indicating that AlkB activity is stimulated by binding the methyl group of 3-methylcytidine nucleosides.

Crystallographic, spectroscopic, and solution studies of other 2OG oxygenases have indicated that binding of the prime substrate to the enzyme complex is required to initiate the binding of molecular oxygen and subsequent coupled catalysis.[77-79] It is therefore possible that 1-methyladenosine and 3-methylcytidine nucleosides may be binding in the AlkB active site, followed by the binding of dioxygen, but as the reactive oxidizing intermediate (Fe^{IV}=O) is positioned incorrectly to make the requisite connections with the DNA oligomer to effect hydroxylation, instead only uncoupled 2OG decarboxylation occurs, but how L-ascorbate promotes this completely uncoupled 2OG turnover, still remains unknown.

Other reducing agents including D-isoascorbate and DTT have also been able to fully replace L-ascorbate in this uncoupled 2OG decarboxylation, with the latter giving approximately 90% of the activity of L-ascorbate. β-mercaptoethanol results in only a moderate stimulation in the uncoupled 2OG turnover. Tris(2-carboxyethyl)phosphine (TCEP), dithionite and 4-nitrocatechol do not stimulate AlkB catalysed uncoupled 2OG turnover (Figure 3.11).

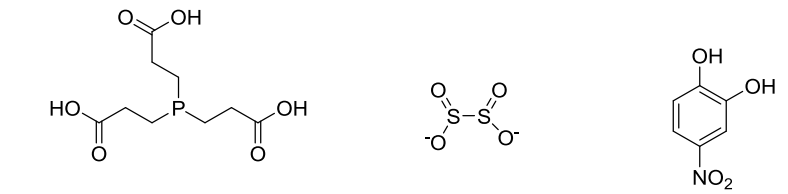


Figure 3.11 Structure of the reducing agents that do not stimulate AlkB catalysed uncoupled 2OG decarboxylation.

The exact mechanism of how the presence of certain nucleoside substrate analogues in the presence L-ascorbic acid or other reducing agents (DTT) results in significant level of uncoupled 2OG decarboxylation remains unknown.

3.1.7 Stimulatory effect of L-ascorbate on TfdA 2OG oxygenase

2,4-Dichlorophenoxyacetic acid (2,4-D) is a broad leaf herbicide that undergoes rapid degradation in the environment.[80] The first step in catabolism of 2,4-D is its oxidation to 2,4-dichlorophenol and glyoxylate (Figure 3.12). This reaction is catalysed by 2,4-D dioxygenase[81,82] commonly called, TfdA (from the *TfdA* gene)[83] in *Ralstonia eutropha* (formerly *Alcaligenes eutrophus*) JMP134. TfdA belongs to the 2OG oxygenases family of enzymes. In *in vitro* studies have shown that catalytic activity of TfdA towards the thiophenoxyacetic acid (TPAA, Figure 3.12) substrate (but not towards 2,4-D) specifically and essentially depends on the presence of L-ascorbic acid in the assay mixture.[5,82] The mechanism by which L-ascorbate stimulates the activity of TfdA is still unknown.

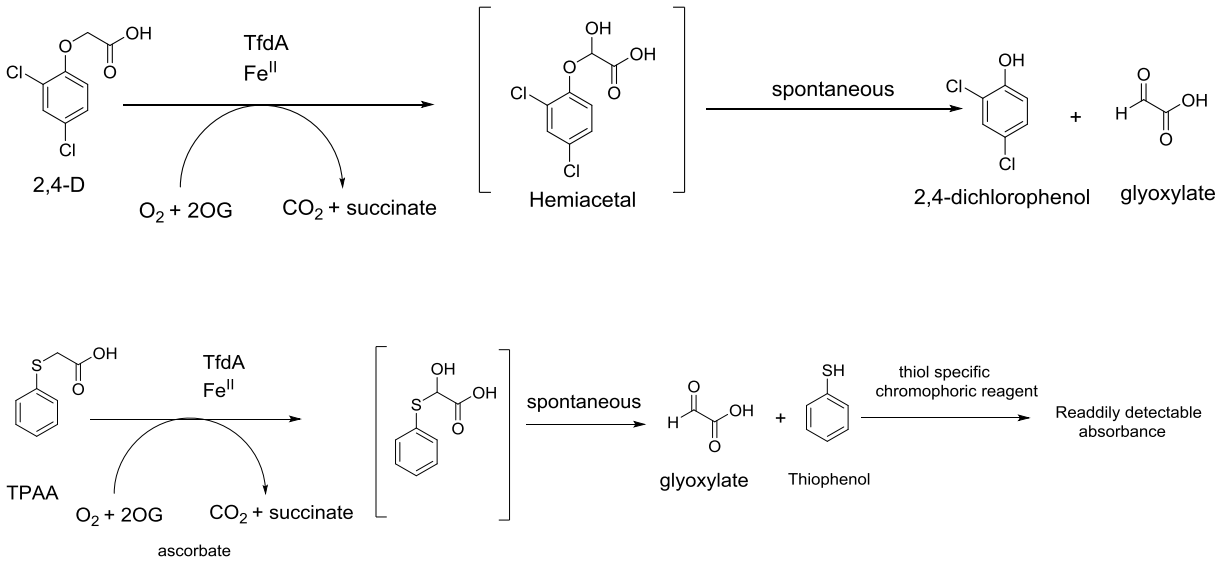


Figure 3.12 Reaction catalysed by TfdA.

3.1.8 Stimulatory effect of L-ascorbate on JMJD2E 2OG oxygenase

The 2OG oxygenase histone demethylase (HDM) subfamily of enzymes catalyse the demethylation of N^{ϵ}-methylated lysine residues in the N-terminal tails of histones proteins. The demethylation proceeds *via* the enzyme catalysed hydroxylation. Methylated histone tail lysines are involved in the formation of different chromatin states, and have roles in both gene silencing and activation.[84] Several different HDM subfamilies have been linked with many diseases, with the JMJD2 HDMs being associated to prostate and oesophageal cancers.[85]

In vitro studies have shown that L-ascorbate significantly stimulate the activity of JMJD2E towards K9me2/me3/K9Me1 peptides.[86] In these reactions uncoupled 2OG decarboxylation also occur over time.

3.2 Stimulation of non-2OG oxygenases enzymes by L-ascorbic acid

In addition to 2OG oxygenases, the activities of a large number of biologically important non-2OG oxygenases enzymes are also highly dependent on the presence of L-ascorbate both in *in vitro* assays and in cells. Most of these enzymes are mono or dioxygenases that contain a metal co-factor such as Cu^{II} in their active site and require molecular oxygen, and a reductant such as L-ascorbate for catalysis. Although, the exact role of L-ascorbic acid in the catalysis of these enzymes is still not well-established, but, it is generally believed that it acts as a reducing agent in the assays of these enzymes, which keep the enzyme catalytic metal in the reduced (catalytic) form and thus promote maximum enzyme activity. The role of L-ascorbate in the catalysis of some of these enzymes is discussed below,

Table 3.2 List of biologically important non-2OG dependent enzymes whose catalysis require L-ascorbic acid (Reproduced from Reference [9]).

Enzyme	Physiological Role	Enzymatic Activity	Metal Ion Centre
1-Aminocyclopropane-1-carboxylate Oxidase	Ethylene (plant hormone) biosynthesis	Oxidation of 1-Aminocyclopropane to ethylene and cyanoformic acid	Iron
Cholesterol 7-alpha monooxygenase E.C. 1.14.13.17	Cholesterol catabolism; bile acid synthesis (animals)	Hydroxylation of Cholesterol	-
Catechol-O-methyl transferase E.C. 2.1.1.6	Adrenaline (epinephrine) inactivation (animals)	Increased levels of adrenaline (epinephrine)	-
Dopamine-β- monooxygenase E.C. 1.14.17.1	Noradrenaline (norepinephrine) synthesis	β-hydroxylation of dopamine	Copper

4-Hydroxylphenylpyruvate dioxygenase E.C. 1.13.11.27	Tyrosine metabolism	Decarboxylation and hydroxylation of 4-hydroxyphenyl pyruvic acid to homogenistic acid	Iron
Mitochondrial glycerol-3-phosphate dehydrogenase E.C. 1.1.99.5	NAD(P) H and ATP production; aid in insulin release	Dehydrogenation of triose phosphate	Iron
Peptidyl glycine α-amidating monooxygenase E.C. 1.14.17.3	Peptide amidation in peptide hormone metabolism	C-terminal glycine amidation	Copper
Thioglucoside glucohydrolase E.C. 3.2.3.1	Catabolism of glucosinolates (plants)	Hydrolysis of S-glucosides	-
Violaxanthin de-epoxidase	Zeaxanthin biosynthesis and the xanthophylls cycle (plants)	De-epoxidation of violaxanthin and antheroxanthin	-

3.2.1 Stimulatory effect of L-ascorbate on dopamine-β-monooxygenase (DβM)

L-Ascorbic acid acts as a biological electron donor to dopamine-β-monooxygenase (DβM), a Cu^{II} containing aromatic amino acid oxygenase, that catalyse the hydroxylation of dopamine to norepinephrine (a neurotransmitter) (Figure 3.13).[87-90]

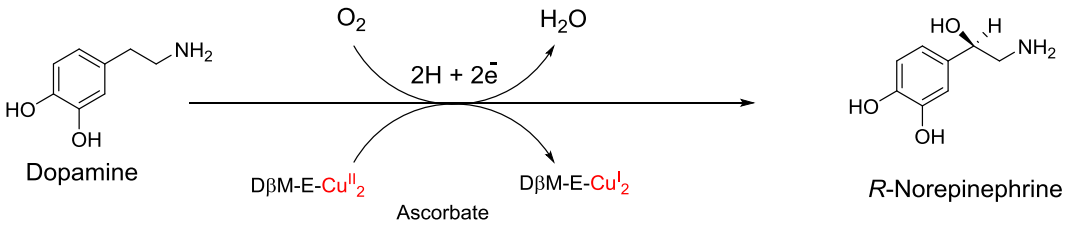


Figure 3.13 L-ascorbic acid (exogenous electron donor) in DβM catalysis

Studies by Kandatege *et. al*[91] have shown that L-ascorbic acid analogues such as 6-deoxy-L-ascorbic acid, 6-bromo-6-deoxy-L-ascorbic acid and 5,6-isopropylidene-L-ascorbic acid (Figure 3.14) also display reducing activity to DβM, similar as L-ascorbic acid, suggesting that the C-6 -OH group of L-ascorbic acid has no significant influence on the electron donor efficiency of the molecule. However, modification of the 3-OH group of L-ascorbic acid either with an alkyl group or an acetyl groups (Figure 3.14) results in complete loss of reducing activity of the molecule to the enzyme, demonstrating a crucial role of the 3-OH group of L-ascorbic acid for its activation of DβM. Analogue, 6-amino-6-deoxy-L-ascorbic acid (Figure 3.14) however was proved as not influencing DβM activity, although it is not expecting that 6-amino group can significantly affect the molecule reduction potential. The other possible explanation of this compound could be attributed to its altered structural features. Since at physiological conditions, the molecule should exists as a zwitterion (6-amino group positively charged and 3-OH group negatively charged), therefore the

electrostatic interactions of the two charged groups of the molecule may alter the conformation/mode in such a way that it could not interact with the enzyme efficiently.

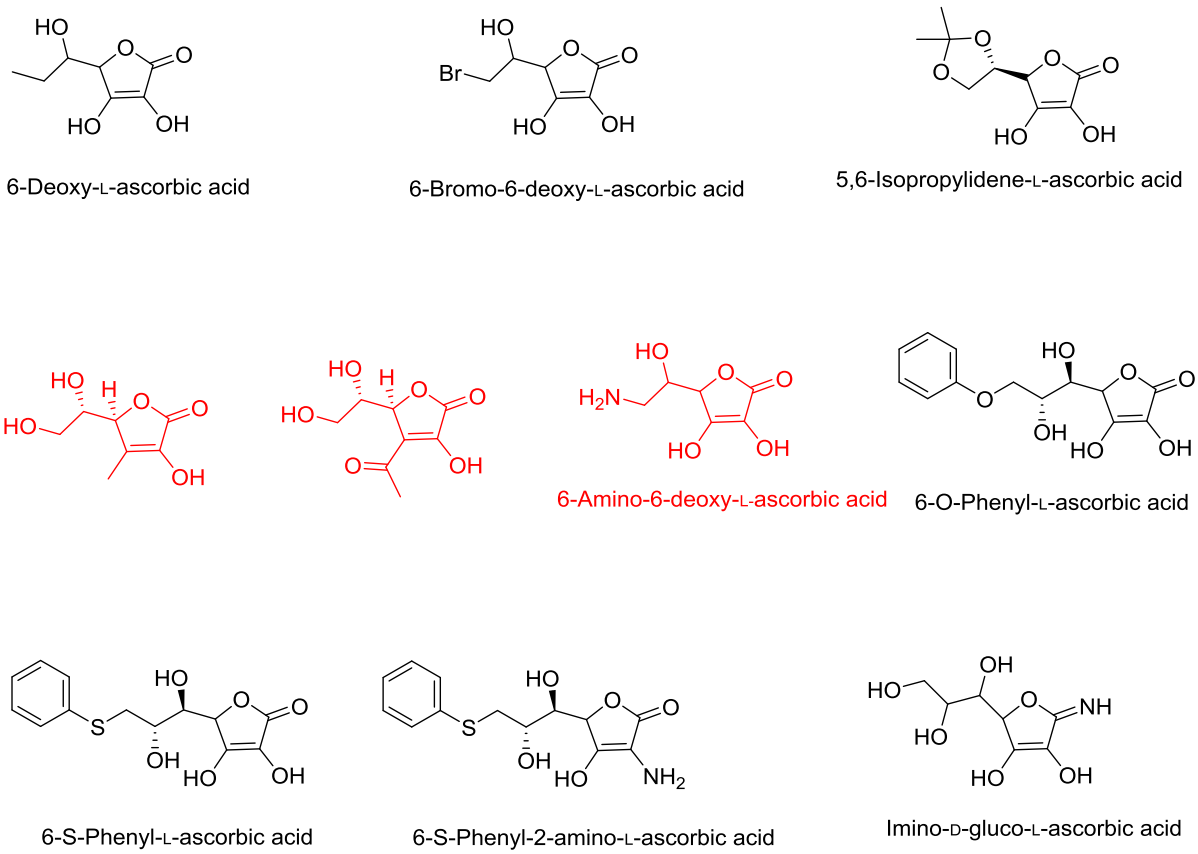


Figure 3.14 L-Ascorbic acid analogues that had been tested for their stimulation of DβM. Compounds coloured red do not activate DβM.

Clearly, it appears that the 5-OH and/or 6-OH groups of the L-ascorbate molecule are not essential for the efficient interactions with the enzyme; positively charged groups at position 6 are however, intolerable. L-Ascorbate analogues with bulky hydrophobic group at position six, such as 6-O-phenyl-L-ascorbic acid and 6-S-phenyl-L-ascorbic acid (Figure 3.14) have been proved as excellent reducing agents, stronger than L-ascorbic acid. The high affinity of these molecules towards the enzyme could simply be due to the hydrophobic interactions of the phenyl group of the molecule with some hydrophobic amino acids residues in the reduction site of the enzyme. However, in contrast, the 6-S-phenyl-2-amino-L-ascorbic acid (Figure 3.14) was shown as a competitive inhibitor of DβM similar to imino-D-glucoascorbic acid with respect to L-ascorbic acid. The unexpected and unparalleled behaviour of these two series of compounds suggest that they may be interacting with the enzyme in possibly two different modes. One possible important difference between these two series of compounds is the pKa values of the 3-OH group. In case of L-ascorbic acid and its 6-OH substituted

derivatives the pKa of the 3-OH group is 4.2, and therefore under physiological conditions it should be mostly negatively charged and the negatively charged species may interact with the enzymes preferentially. The pKa of the 3-OH group of 2-aminoascorbic acid and is 6-OH substituted derivatives is about 6.8 and hence under physiological conditions it should exist mostly uncharged and the uncharged species may interact with the enzyme by an altered mode. Imino-D-glucoascorbic acid (Figure 3.14), also a strong electron donor, has been shown as a potent competitive inhibitor of DβM with respect to L-ascorbic acid. The inability of the iminoglucoascorbic acid to act as a reductant for DβM must be due to the altered reduction potential of the molecule due to the replacement of the carbonyl group by an imino group. The strong inhibition of DβM activity by iminoglucoascorbic acid demonstrates that iminoglucoascorbic acid must be interacting exclusively with the oxidised form of the enzyme, in a binding mode that mimics the binding mode of L-ascorbic acid.

3.2.2 Stimulatory effect of L-ascorbate on ACCO

L-Ascorbic acid is crucially required for ethylene biosynthesis in plants. Ethylene serves as a hormone in plants.[92] Ethylene is involved in stimulating and regulating of the ripening of fruit, the opening of flowers, and the abscission (shedding) of leaves. The final step of ethylene biosynthesis involves the conversion of 1-aminocyclopropane-1-carboxylic acid (ACC) into ethylene (Figure 3.15).[93,94] This reaction is catalysed by enzyme ACC-oxidase (ACCO), which is a non-heme Fe-containing enzyme formerly known as the ethylene forming enzyme (EFE). In this reaction ACC undergoes two electrons oxidation into ethylene in the presence of dioxygen and two electrons (Figure 3.15). In addition to ethylene, CO_2 and cyanide are also formed.

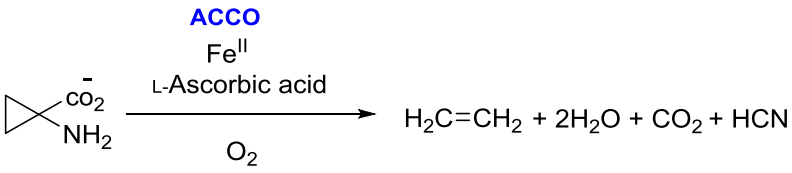


Figure 3.15 Reaction catalysed by ACCO.

In *in vitro* studies have shown that the native ACCO shows significantly low activity and in the presence of L-ascorbic acid, the enzymatic activity is significantly stimulated. Structurally, ACCO is very similar to the Fe^{II} and 2OG dependent oxygenases family of enzymes; however ACCO does not require 2OG for catalysis. Other reducing agents, including saccharic acid 1,4-lactone (a redox-inactive structural analogue of ascorbic acid)

DTT, β-mercaptoethanol NADH, NADPH and reduced cytochrome c are also able to stimulate ACCO activity, however only slightly as compared with L-ascorbic acid.[95] Dehydro-L-ascorbate, neither stimulate nor inhibit ACCO activity.[95,96] Studies by Zhang *et al.* have shown that 5,6-O-isopropylidine-L-ascorbate and D-ascorbate can also fully replace L-ascorbate for stimulating ACCO activity (5,6-O-isopropylidine L-ascorbate > D-ascorbate > L-ascorbate), whereas, L-galactono-γ-lactone is completely inactive towards ACCO.[96]

In addition, L-ascorbic acid is also crucially required by other biologically important enzymes including cholesterol 7-α-hydroxylase (cholesterol metabolism into bile acids),[97,98] aromatic amino acids hydroxylases such as L-tryptophan hydroxylase (TPH) (biosynthesis of serotonin or 5-hydroxytryptamine, a neurotransmitter)[9,99], phenylalanine hydroxylase,[100] L-tyrosine hydroxylase[9,99] (L-tyrosine metabolism into catecholamines) and 4-hydroxyphenyl pyruvate dioxygenase (HPPD)[101-103] (tyrosine catabolism).

3.3 Role of L-ascorbic acid as an antioxidant/free radical scavenger

In addition to its role as enzymes stimulator/co-factor, ascorbic acid also functions as an effective antioxidant in biological systems. Oxygen is the key to the survival of all life forms. During cellular oxygen metabolic processes in the mitochondria of animals and chloroplast of plants, where cellular energy is produced by the reaction of an oxygen molecule with 4 electrons and 4 protons resulting in the formation of water (Figure 3.16).

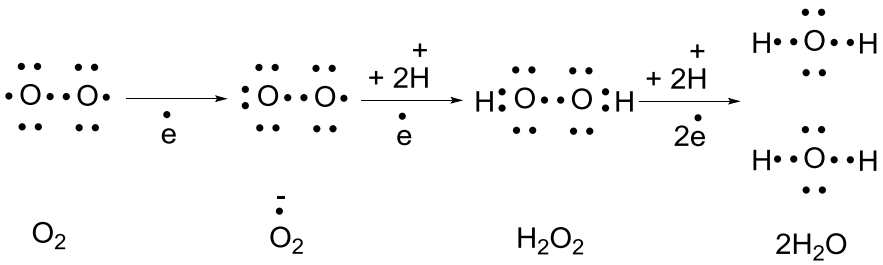


Figure 3.16 Four-electron reduction process of oxygen into water.

In these metabolic processes, reactive oxygen species (ROS), including peroxides, superoxide, hydroxyl radical, and singlet oxygen are also formed as natural by-products. Inside cells, these ROS are also generated from exogenous sources such as ionizing radiation. These ROS initiate free radical chain reactions resulting in serious damage to tissues.[104] These conditions are usually described as oxidative stress in chemical biology.[105] It is therefore highly essential that an active defensive mechanism be in place to control these toxic ROS from participating in harmful biological reactions inside the cells. L-Ascorbic acid

interacts with ROS and their derivatives, both enzymatically and non-enzymatically and neutralize their toxic effects (Figure 3.17).[104] These ROS, in particular, hydrogen peroxide reacts with transition metal ions in the cells, especially Fe^{II} and generate hydroxyl radical (Fenton reaction), a highly reactive radical, which immediately removes electrons from any molecule coming in its path, converting the molecule into a free radical, thus propagating a chain reaction. The hydroxyl radicals ($HO^{\cdot}$) react easily with cellular components such as proteins, DNA, lipids and membrane lipids (Figure 3.17).[104] Membrane lipids contain allylic hydrocarbon chains that can easily react with hydroxyl radicals. The resulting carbon-centred radicals rapidly react with oxygen and form alkyl peroxy radicals ($LOO^{\cdot}$), which then abstract a hydrogen atom from lipids and produce LOOH. Due to its stability, LOOH flows and reacts with metal ions and generate free radicals causing damage to membranes and other cellular bodies. Although, L-ascorbic acid cannot directly detoxify the lipophilic radicals, it acts as an available agent to the tocopherol for the reduction of lipid peroxide radicals.[104] Tocopherols (vitamin E) are also biological antioxidants, scavenging peroxyl radicals, by reacting with them and forming a tocopheryl radical. At the aqueous-lipid interface, L-ascorbic acid donates its hydrogen to the membrane-bound oxidized tocopherol radical and thus regenerate the active reduced tocopherol for continued peroxyl radicals scavenging.

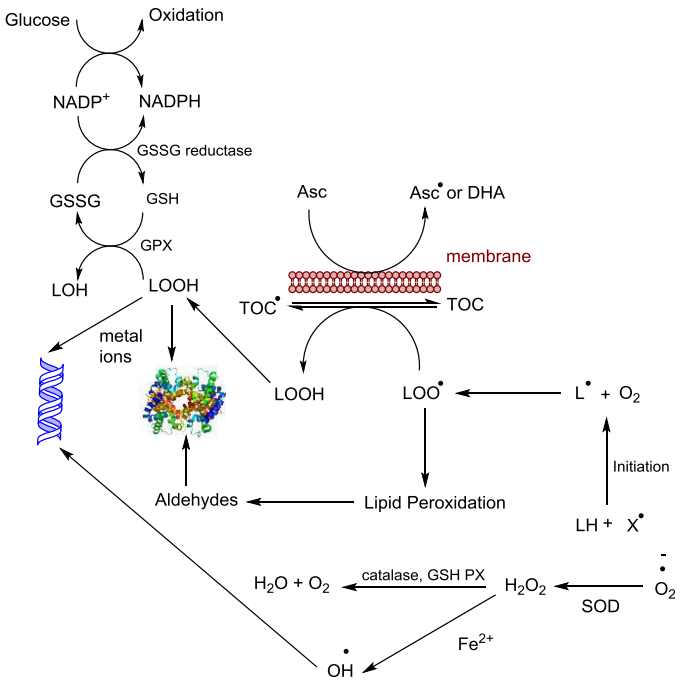


Figure 3.17 Schematic representation of the radical reactions in the cells and antioxidant activities.
Radical species are represented by X, tocopherol by TOC and glutathione peroxidase by GPX. Radical chain reactions lead to oxidative stress. Thus, cell antioxidants e.g. L-ascorbic acid, antioxidant enzymes and glucose supplies reducing power to fight oxidative stress. [104]

Compared with other antioxidants in the cells such as uric acid, carotenoids, flavonoids, α-tocopherol, etc. the mode of action of L-ascorbic acid as an anti-oxidant is more beneficial because it terminates the radical propagations reactions, while itself is converted into non-toxic oxidised forms e.g semidehydro-L-ascorbic acid radical and dehydro-L-ascorbic acid. Semidehydro-L-ascorbic acid radical disproportionates back to L-ascorbic acid and dehydro-L-ascorbic acid (Figure 3.18). This non-enzymatic antioxidant activity of L-ascorbic acid is very useful in maintaining the balance of keeping a wide range of important biological substrates in their reduced and active forms. For example, L-ascorbic acid maintain a pool of the reduced form of folic acid (vitamin B_9). Folic acid is required in many of the cellular one-carbon migration reactions, which are frequently involved in the formation of many of bioactive molecules.[104]

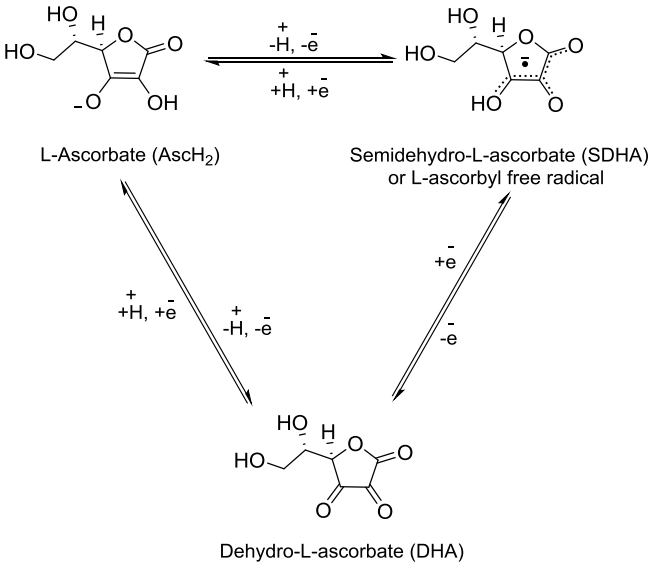


Figure 3.18 L-ascorbate redox system

3.4 Aim of the project

We are interested in exploring the underlying chemistry of the catalysis of 2OG oxygenases, which employ 2-oxoglutarate (2OG) as a co-substrate and Fe^{II} as a co-factor for catalysis.

As described above, many of these 2OG oxygenases require L-ascorbic acid for maximum enzymatic activities in the turnover assays. The aim of the work described in this chapter was in general to apply ^1H NMR spectroscopy to gain a better understanding of the complex mode of action of ascorbate in the *in vitro* assays. We aimed to investigate the role of L-ascorbate in the following contexts;

The 2OG oxygenases rate of reaction in the linear region remains same both in the absence and presence of ascorbate, but deviates suddenly in the absence of ascorbate after 10-15 s. However, no such study has so far proved that this sudden enzyme inactivation is due to the unavailability of free Fe^{II} species in solution and that ascorbate is required (as reductant) to convert Fe^{III} to Fe^{II} to activate the enzyme. This leaves the possibility that ascorbate may perhaps be reducing the enzyme-bond Fe^{III} to catalytic Fe^{II}, but again, no such evidence has so far emerged to prove that ascorbate specifically binds in the active site of a 2OG oxygenase. So what is the mechanism of the stimulation of a 2OG oxygenase activity in the presence of ascorbate in the *in vitro* assays?

Other important questions are; why in the presence or absence of a substrate, incubation of a 2OG oxygenase assay solutions with ascorbate over longer time results an increase in the level of uncoupled 2OG turnover? In addition, why the presence of certain compounds results a significant increase in uncoupled 2OG decarboxylation in the presence of ascorbic acid in the assay mixture as observed in the case of AlkB.

According to the reported literature, there is a considerable variation in the scope of ascorbate in the catalysis of 2OG oxygenases, with some enzymes showing a high dependency on ascorbate (specifically) for maximum catalytic activity, while some do not; some assays result in high uncoupled 2OG turnover with ascorbate over time, while some result in low uncoupled 2OG turnover. Hence, to understand this complex role of ascorbate, it is also highly important to assess the influence of typical assay conditions of 2OG oxygenases on ascorbate in the assay mixture. It is possible that the stimulation of 2OG oxygenases (high turnover) over time by incubation with ascorbate at least in parts could be through an indirect mechanism, perhaps through some chemistry between ascorbate and the components of the assay mixture.

3.5 Ascorbic acid degradation leads to non-enzymatic 2OG decarboxylation

In the presence of a known substrate, the activity (turnover reaction) of a 2OG oxygenase is generally followed either from the decrease in the level of the substrate or from the decarboxylation of 2OG into succinate, whereas, in case when the enzyme substrate is not known, the activity of the enzyme can be only assessed *via* 2OG turnover into succinate. These assays are performed in the presence of L-ascorbate as reducing agent to keep the iron predominantly in the reduced form (Fe^{II}) to promote maximum enzyme activities. For an accurate interpretation of the enzyme kinetics results, it is highly important to carry out necessary control experiments to eliminate any artifacts in the results that could arise from the assay conditions instead of enzyme catalysis.

In working towards understanding the chemistry of the complex role of L-ascorbate in the catalysis of 2OG oxygenases, using ^1H NMR spectroscopy, we observed that in the absence of a 2OG oxygenase, if a mixture of L-ascorbate (500 µM) and 2OG (200 µM) is incubated in 50 mM Tris-D_{11} buffer (one of the most commonly used buffers for 2OG oxygenases assays) solution at pH 7.5 in H_2O in an eppendorf tube at room temperature (298K), then over time, L-ascorbate in the mixture degrades and at the same time 2OG is decarboxylated forming succinate (~ 88 µM) (Figure 3.19). Due to the low amounts of the L-ascorbate degradation products, we didn't aim to investigate their identity.

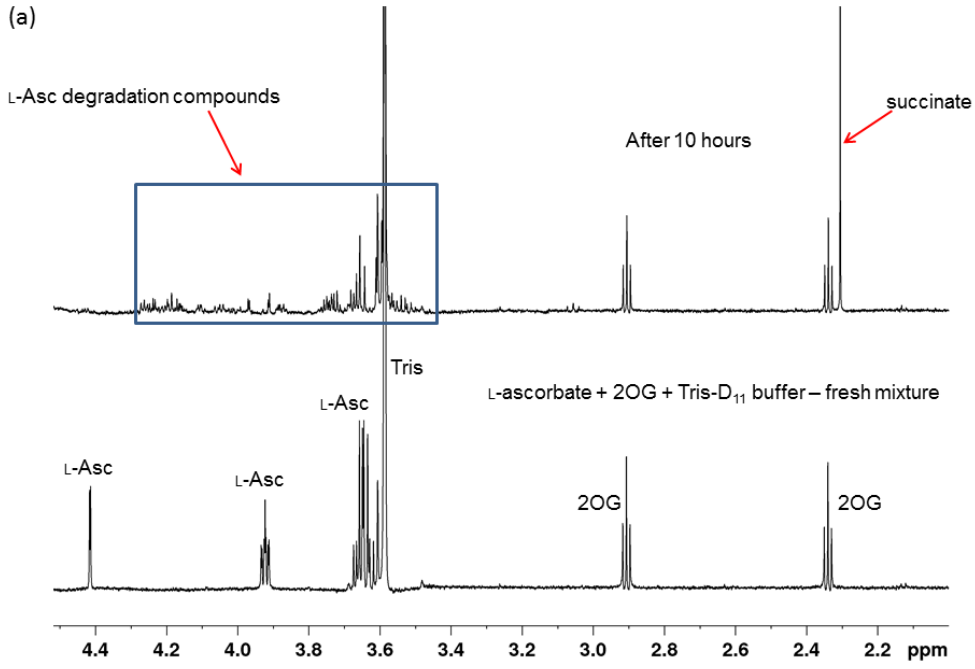
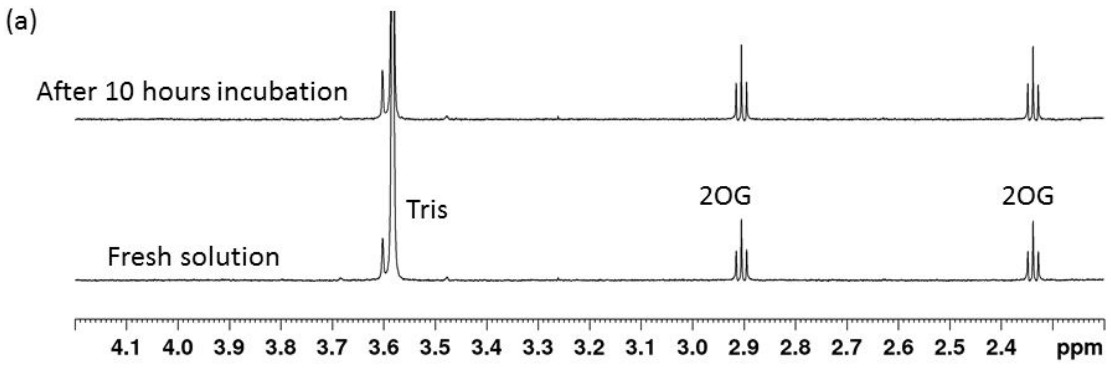


Figure 3.19 Non-enzymatic 2OG decarboxylation caused by ascorbic acid degradation in buffer.
(a) Overlay of a ^{1}H NMR spectrum of a fresh mixture of ascorbate and 2OG in buffer (bottom) with a mixture incubated for 10 hours at room temperature on a bench in the laboratory. (Top): During 10 hours incubation, ascorbate has been degraded and 2OG is turned over into succinate. The assay mixture contained 500 µM L-ascorbate (L-Asc), 200 µM 2OG in 50 mM Tris-D_{11}, pH 7.5 in H_2O. After 10 hours incubation, all (500 µM) of the L-ascorbate was found degraded (signals in blue rectangle) resulting in the formation of 88 µM of succinate. (b) Schematic representation of (a).

To eliminate any possible role of bacteria etc. which could potentially come from buffer, the stock solutions of L-ascorbate and 2OG or the eppendorf tube, the stock solutions of L-ascorbate, 2OG and buffer were separately made sterile and then the assay mixture was incubated in a sterile falcon tube. In this case also, it was found that L-ascorbate (500 µM) degrades and at the same time 2OG decarboxylates into succinate (~ 90 µM) demonstrating that the 2OG decarboxylation by L-ascorbate degradation is purely a non-enzymatic process.

In control experiments, 2OG alone was incubated in Tris-D_{11} buffer and no succinate was formed, implying that the 2OG decarboxylation directly correlates with L-ascorbate degradation (Figure 3.20a). Similarly, by incubating L-ascorbate alone in Tris buffer, it was found degraded, suggesting that its degradation is an automatic process, and that 2OG has no role in the L-ascorbate degradation (Figure 3.20b).



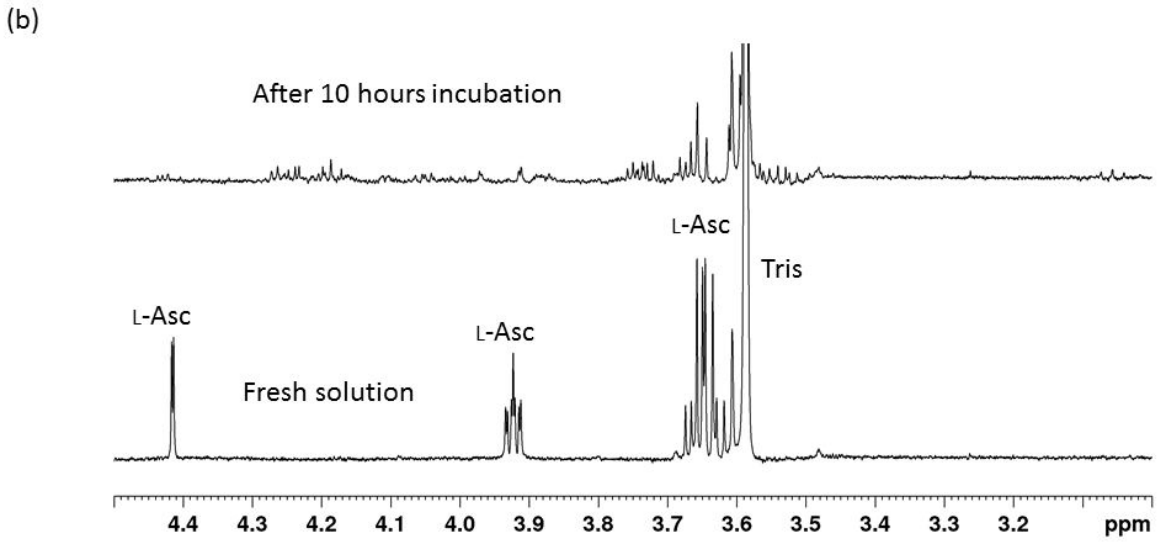


Figure 3.20 Control experiments for 2OG decarboxylation by L-ascorbate degradation.
(a) Overlay of ^{1}H NMR spectra of a solution of 2OG in Tris-D_{11} buffer. Bottom: freshly prepared solution. Top: after incubation for 10 hours. (b) Overlay of ^{1}H NMR spectra of a solution of L-ascorbate in Tris-D_{11} buffer. Bottom: freshly prepared solution. Top: After incubation for 10 hours. The concentrations used are 2OG 200 µM and L-ascorbate 500 µM.

To investigate how the degradation of L-ascorbate results in simultaneous 2OG decarboxylation, we monitored the L-ascorbate/2OG/Tris-D_{11} buffer mixture incubated in an eppendorf tube after every one hour over a period of ten hours. We observed that over time, L-ascorbate slowly degrades into dehydro-L-ascorbate (DHA), which due to its instability, degrades further into other compounds. During this L-ascorbate degradation process, 2OG is also decarboxylated forming succinate, demonstrating a direct link between the two processes (Figure 3.21). The peaks of DHA were identified from the spectrum of a reference sample. It was also observed that the rate of L-ascorbate degradation is much faster than 2OG turnover.

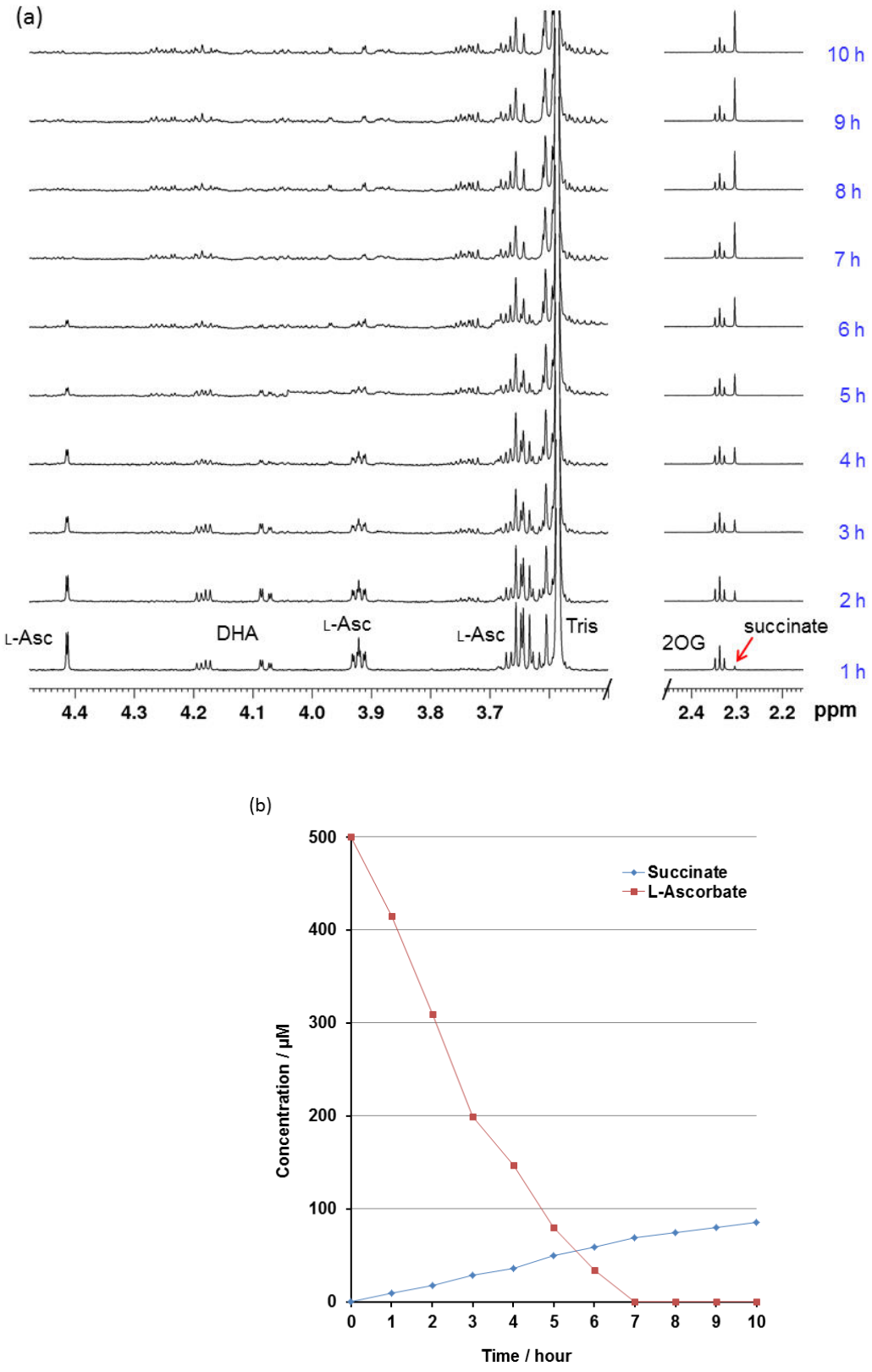


Figure 3.21 Time course of ascorbate degradation and concomitant 2OG turnover into succinate.
(a) Time course analysis by ^{1}H NMR, showing the degradation of L-ascorbic acid and the simultaneous decarboxylation of 2OG (parts of the spectra shown). (b) Graph showing the degradation of L-ascorbate and the simultaneous succinate formation. The assay mixture contained 500 µM L-ascorbate, 200 µM 2OG in 50 mM Tris-D_{11}, pH 7.5 in H_2O.

From our knowledge of the reported literature, we propose that this degradation of L-ascorbate (Figure 3.22)[106-110] in buffer is the oxidative degradation of L-ascorbate by dissolved oxygen in the assay mixture. The oxidative degradation of L-ascorbic acid is known as a complex process and involves a number of oxidation/reduction and inter-molecular rearrangement reactions.[109] At physiological pH ascorbate exists predominantly in the monoanion form (Figure 3.22). The anion loses an electron from its ene-diol group to form semi-dehydro-L-ascorbate radical. This reversible reaction is believed to be primarily responsible for the antioxidant properties of ascorbic acid to scavenge toxic free radicals produced in biological systems. Further oxidation of semi-dehydro-L-ascorbate radical forms the uncharged molecule DHA. DHA is a highly unstable compound in aqueous solutions at neutral pH, undergoes hydrolysis in two ways. One is by the formation of a bicyclic hemiketal (BH) *via* intramolecular ring closure. The other is by the formation of diketogulonic acid (DKG) *via* ring opening. Further oxidation of DKG by H_2O_2 produces 4,5,5,6-tetrahydroxy-2,3-diketohexanoic acid (THDH) (first route) and trihydroxybutanoic acid (threonic acid) along with oxalic acid (second route). In the third route, the hydrolysis of DKG results in the formation of two epimers of pentanoic acids (L-xylonic and L-lyxonic acids).

We also tested this non-enzymatic 2OG turnover in the presence of DHA in the assay mixture, and it was found that DHA also degrades and simultaneously results in 2OG decarboxylation (Figure 3.23). With DHA, however, the level of 2OG turnover was much lower as compared with L-ascorbate i.e 15% vs 45% under the same incubation conditions. Since, DHA can reduce back to L-ascorbate in the presence of H^+ under physiological conditions;[111] we speculate that the 2OG turnover occurs during the transformation of L-ascorbate into DHA.

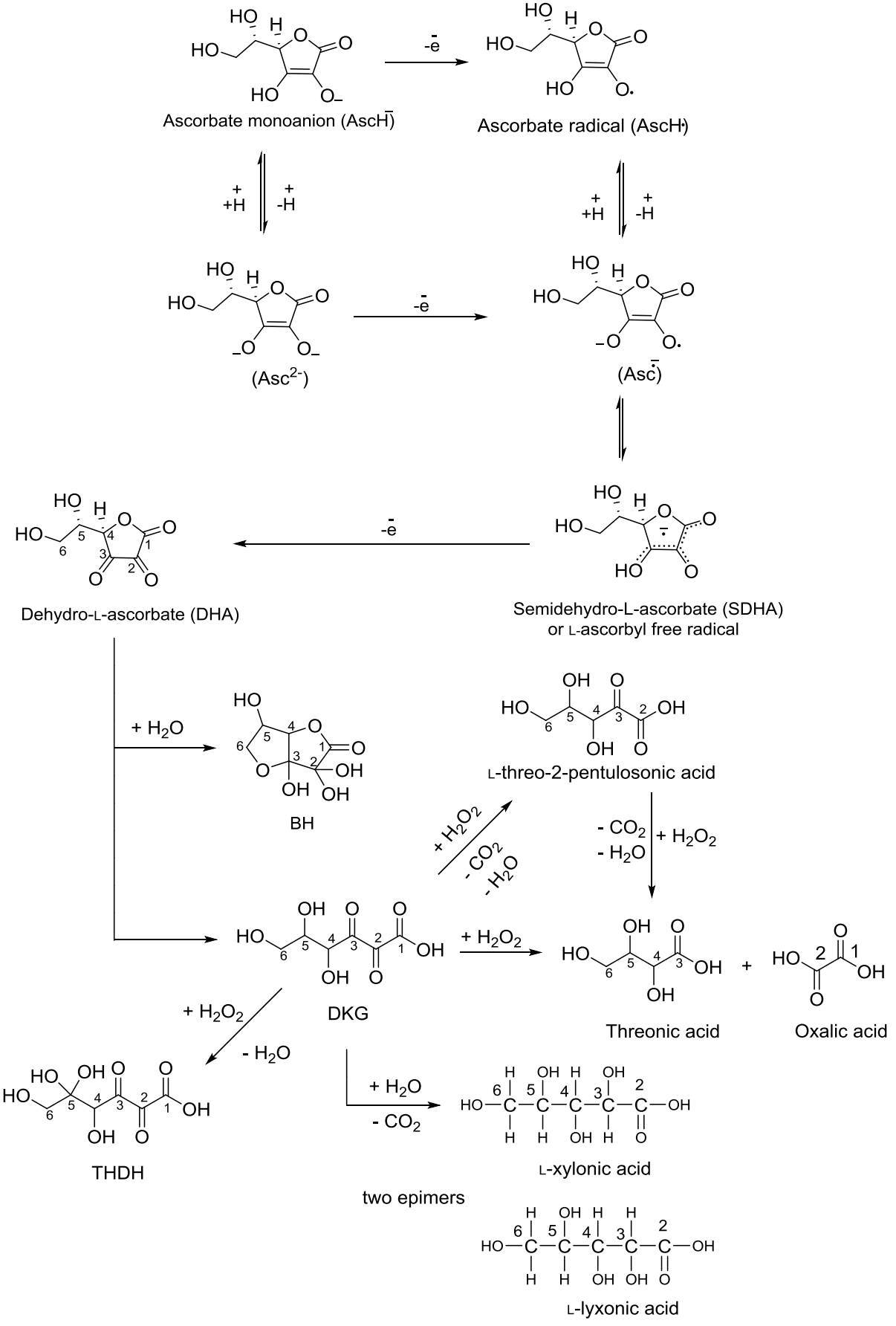


Figure 3.22 Degradation of L-ascorbate.

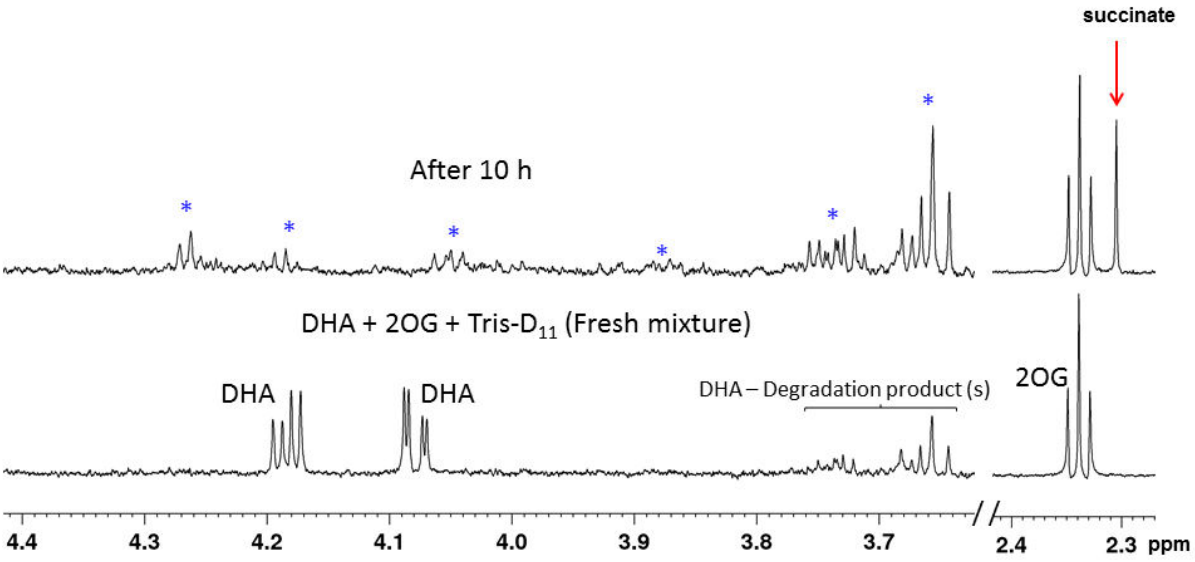


Figure 3.23 Non-enzymatic 2OG decarboxylation by DHA degradation as demonstrated by ^{1}H NMR.
The assay mixture contained 500 µM DHA, 200 µM 2OG in 50 mM Tris-D_{11}, pH 7.5 in H_2O. After 10 hours incubation, 31 µM succinate was formed. The blue asterisks represent the DHA degradation products.

3.5.1 L-Ascorbate oxidative degradation generates H_2O_2

L-Ascorbate undergoes auto-oxidation in the presence of molecular oxygen to produce dehydro-L-ascorbate (DHA) and hydrogen peroxide (H_2O_2)[112-119] in a 1:1 stoichiometry.[115] According to the reported literature, in this reaction H_2O_2 is believed to be formed *via* the highly reactive and toxic intermediate superoxide anion ($O_2^{\cdot-}$) (the product of the one-electron reduction of dioxygen O_2).[120] In contrast to its role as an antioxidant, such (H_2O_2 production) pro-oxidant properties have also been ascribed to ascorbic acid under certain conditions.[121-123] Ascorbic acid can cause strand breakage in the DNA in the presence of oxygen and can initiate cell death in tissue culture, through the generation of toxic reactive oxygen species such as H_2O_2, superoxide anion ($O_2^{\cdot-}$) and hydroxyl radical ($^{\cdot}OH$). Because of its property to produce H_2O_2 in physiological conditions, pharmacological ascorbate concentration has been also proposed as a prodrug for the delivery of H_2O_2 to tumors.[113,114,116,124,125]

We therefore propose that it is this H_2O_2 that generates from the auto-oxidation of L-ascorbate (Figure 3.24a) and that carrys out the 2OG decarboxylation (Figure 3.24b). Similar to L-ascorbate, many polyphenolic compounds (containing the 1,2-diol-ene moiety) have also been reported to generate H_2O_2 by auto-oxidation.[126] This H_2O_2 and/or other oxidation products can contribute to, or account for, many of the observed effects of ascorbate and

polyphenols on cells in culture. Similar to ascorbate, these polyphenols are also well-established anti-oxidants that scavenge (detoxify) the free radicals.[127]

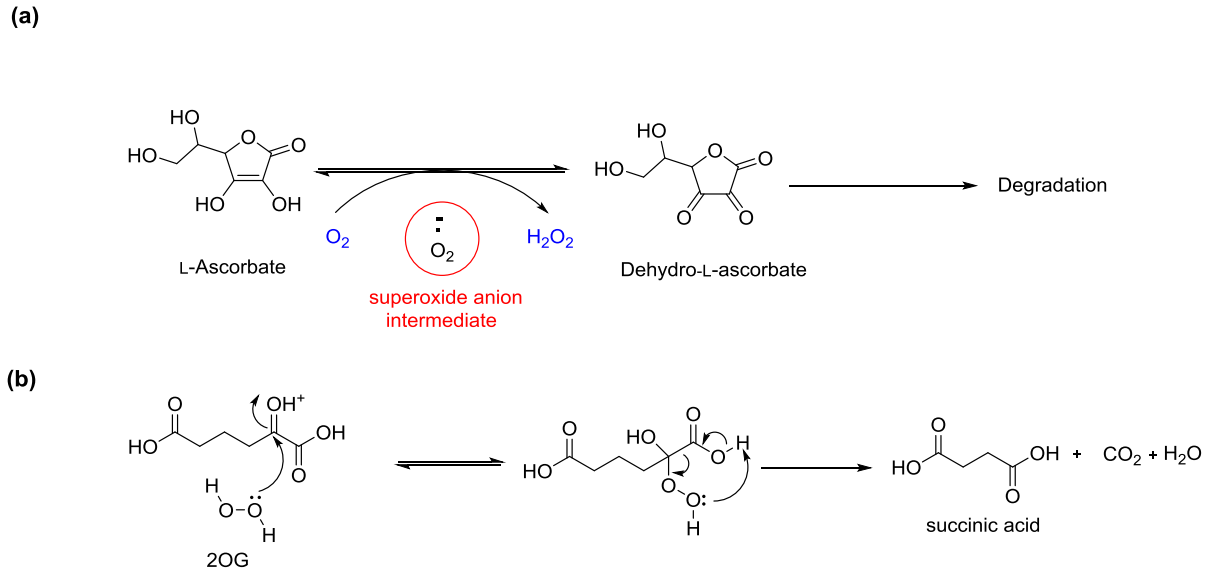


Figure 3.24 Proposed mechanism of 2OG decarboxylation in the presence of L-ascorbate auto-oxidative degradation.

α-Ketoacids have also been found to react (scavenge) with H_2O_2 and undergo decarboxylation (Figure 3.24b).[128-130] This reaction was first described by Holleman[131] in 1904.[132] Mechanistically, in this nucleophilic addition-elimination reaction, the α-ketoacid and H_2O_2 exist in equilibrium with a tetrahedral intermediate. Irreversible decarboxylation of the intermediate forms the corresponding carboxylic acid. The reaction of α-ketoacids with H_2O_2 has also been previously used to measure H_2O_2 induced toxicity in cells.[128-130,133-135] α-Ketoacids have been implicated as strong biological anti-oxidants reacting with H_2O_2 and concomitantly leading to its decomposition. There is renewal of interests in H_2O_2–mediated α-ketoacids oxidation due to the consideration regarding these chemicals to be useful from therapeutic point of view in several disease states from cancer to aging.[135-138] H_2O_2 has also been proposed to be involved in many brain pathologies such as neurodegenerative diseases and in acute diseases such as ischemia or trauma.[139-142] It has been found that pyruvic acid and other related α-ketoacids protects neurons from the toxic effect of H_2O_2 *via* this non-enzymatic detoxification (decarboxylation) process.[143] The protective role of pyruvate in H_2O_2-induced renal cellular injuries is also well-documented.[144]

To test our hypothesis that H_2O_2 (generated during L-ascorbate oxidation into dehydro-L-ascorbate by dissolved oxygen) is the real specie that results in 2OG decarboxylation, we monitored the 2OG (200 μM)/H_2O_2 (500 μM)/Tris-D_{11} buffer incubation mixture after every

one hour over a period of 10 hours at room temperature (298K, as were all subsequent experiments in this chapter) (Figure 3.25).

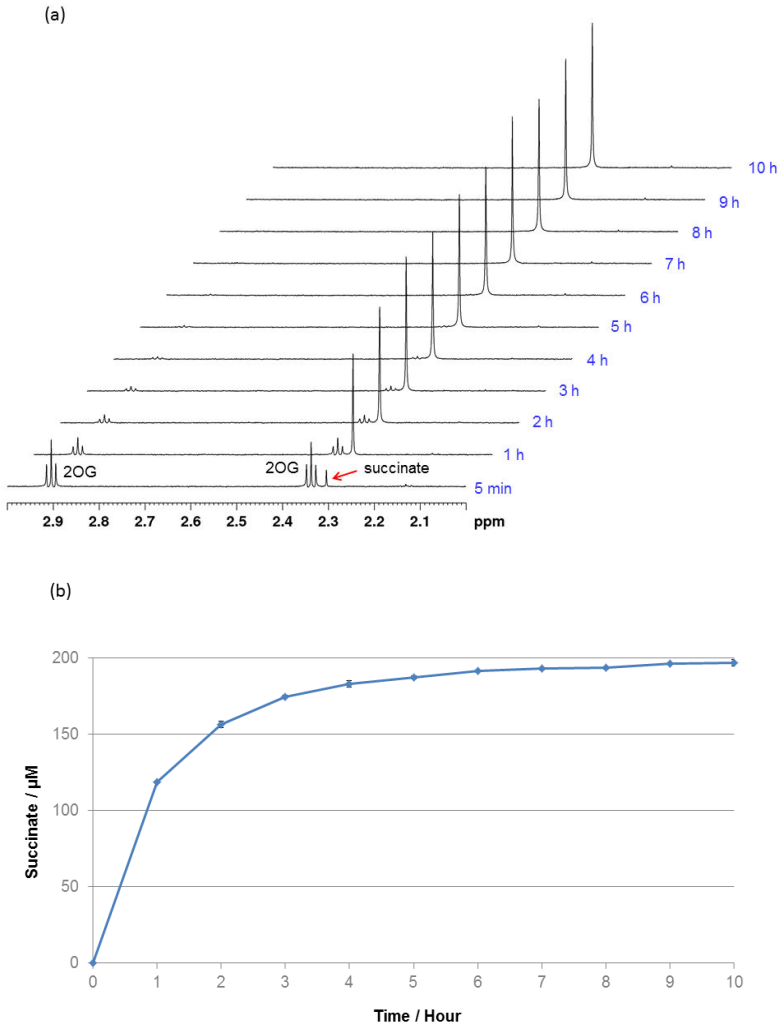
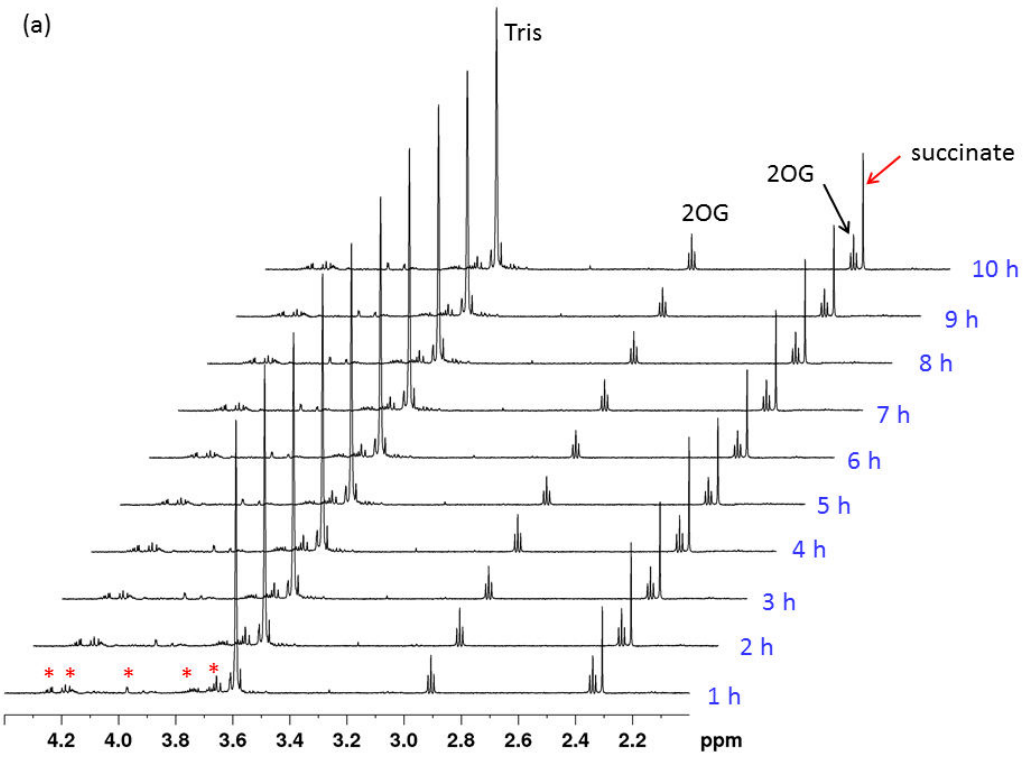


Figure 3.25 2OG decarboxylation by H_2O_2.
(a) ^{1}H NMR time course of 2OG decarboxylation by H_2O_2. (b) Plot of succinate formation vs time. The assay mixture contains 500 µM H_2O_2, 200 µM 2OG in 50 mM Tris-D_{11} pH 7.5, in H_2O.

It was observed that within three hours approximately 90% of the 2OG in the assay mixture was decarboxylated forming succinate, implying that the reaction of H_2O_2 with 2OG itself is a fast process. These results also demonstrate that in the L-ascorbate (500 µM)/2OG (200 µM)/Tris-D_{11} buffer incubation mixture (Figure 3.19), out of the 500 µM of H_2O_2 that generates from 500 µM L-ascorbate oxidation, only 18% (88 µM) of H_2O_2 reacts with 2OG (200 µM) forming 88 µM succinate, while the remaining 82% (412 µM) of H_2O_2 reacts somewhere else in the reaction mixture. We thus speculate that the slow 2OG decarboxylation in the L-ascorbate/2OG/Tris-D_{11} buffer incubation mixture (Figure 3.21) is first due to the slow L-ascorbate oxidation into DHA and second it is only a fraction of the

H_2O_2 left in the mixture that reacts with 2OG as most of H_2O_2 is utilized elsewhere in the assay mixture.

As H_2O_2 is known to result in oxidative degradation of both L-ascorbic acid (antioxidant property of ascorbate) and dehydro-L-ascorbic acid, forming the same degradation products from both compounds,[145,146] while itself reduces to water, it is possible that most of the H_2O_2 that generates from ascorbate oxidation to dehydro-L-ascorbate, may in turn be consuming in the degradation of both these compounds (and their degradation products). Hence to investigate whether most of the H_2O_2 (82%) that generates from L-ascorbate oxidation into dehydro-L-ascorbic acid, in the L-ascorbate (500 µM)/2OG (200 µM)/Tris-D_{11} buffer incubation mixture, could be involved in the oxidation of ascorbate again (an H_2O_2 recycling process), we monitored the H_2O_2 (500 µM)/2OG (200 µM)/L-ascorbate (500 µM)/Tris-D_{11} buffer incubation mixture after every one hour for a period of 10 hours (Figure 3.26). In this case if L-ascorbate can react with H_2O_2, it will slow down the rate and decrease the level of 2OG decarboxylation relative to the rate and level of 2OG decarboxylation in the presence of only H_2O_2, because most of H_2O_2 would be consumed in reaction with L-ascorbate.



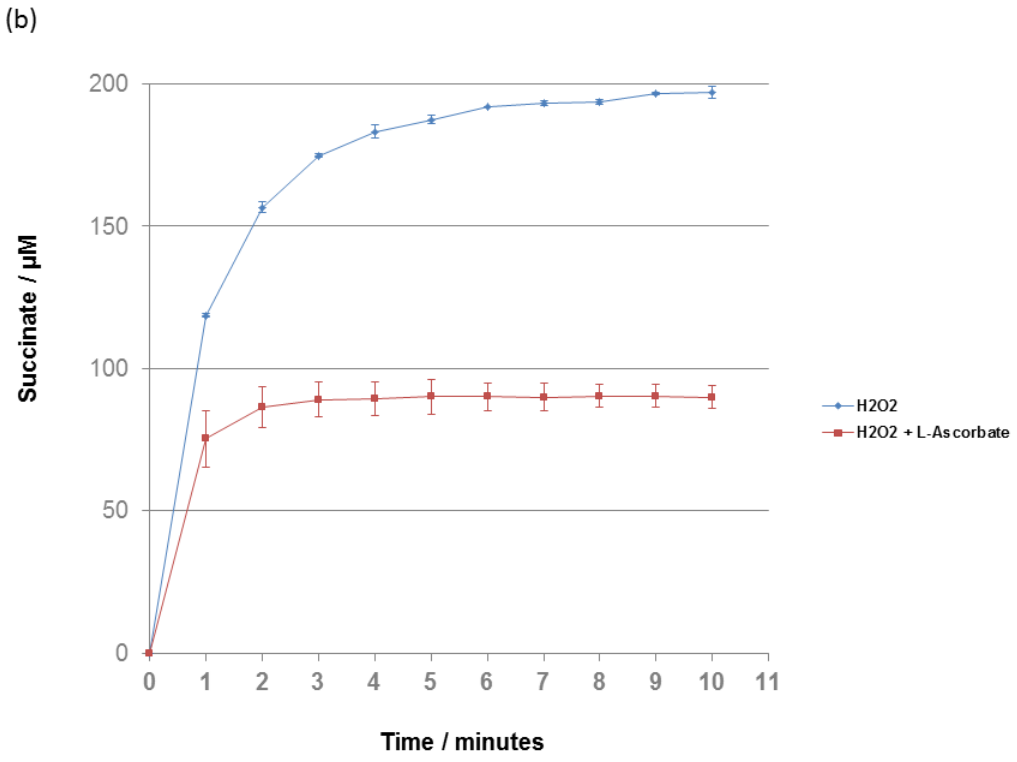


Figure 3.26 2OG decarboxylation by H_2O_2 in the presence of L-ascorbate.
(a) Overlay of ^{1}H NMR spectra (part shown) showing time course analysis of mixture of H_2O_2/2OG/L-ascorbate/Tris-D11 buffer. The red asterisks represent L-ascorbate degradation products. The assay mixture contains 500 µM H_2O_2, 500 µM ascorbate, 200 µM 2OG in 50 mM Tris-D_{11}, pH 7.5 in H_2O. (b) Plots comparing the rate and level of 2OG decarboxylation in the presence of only H_2O_2 (blue curve) and in the presence of both H_2O_2 and L-ascorbate (brown curve, from (a). The error bars represents standard deviation from three separate measurements.

Interestingly, in this mixture it was observed that both the rate and level of 2OG decarboxylation was significantly reduced as compared to 2OG/H_2O_2/Tris-D_{11} mixture, forming only 90 µM of succinate (45% 2OG turnover), (Figure 3.26) which is similar to the amount of succinate formed in the L-ascorbate/2OG /Tris-D_{11} mixture (Figure 3.19). In this case it was also noted that at 1 hour, all of the L-ascorbate in the assay mixture was found degraded, which is significantly faster compared to the L-ascorbate auto-degradation in L-ascorbate (500 µM)/2OG (200 µM)/Tris-D_{11} buffer incubation mixture (Figure 3.21). These results suggest that H_2O_2 speeds up L-ascorbate degradation. Separately, in the absence of 2OG, we also investigated the L-ascorbate degradation in the presence of different amounts of H_2O_2 (Figure 3.27) and it was observed that H_2O_2 speeds up L-ascorbate degradation in a concentration dependent manner.

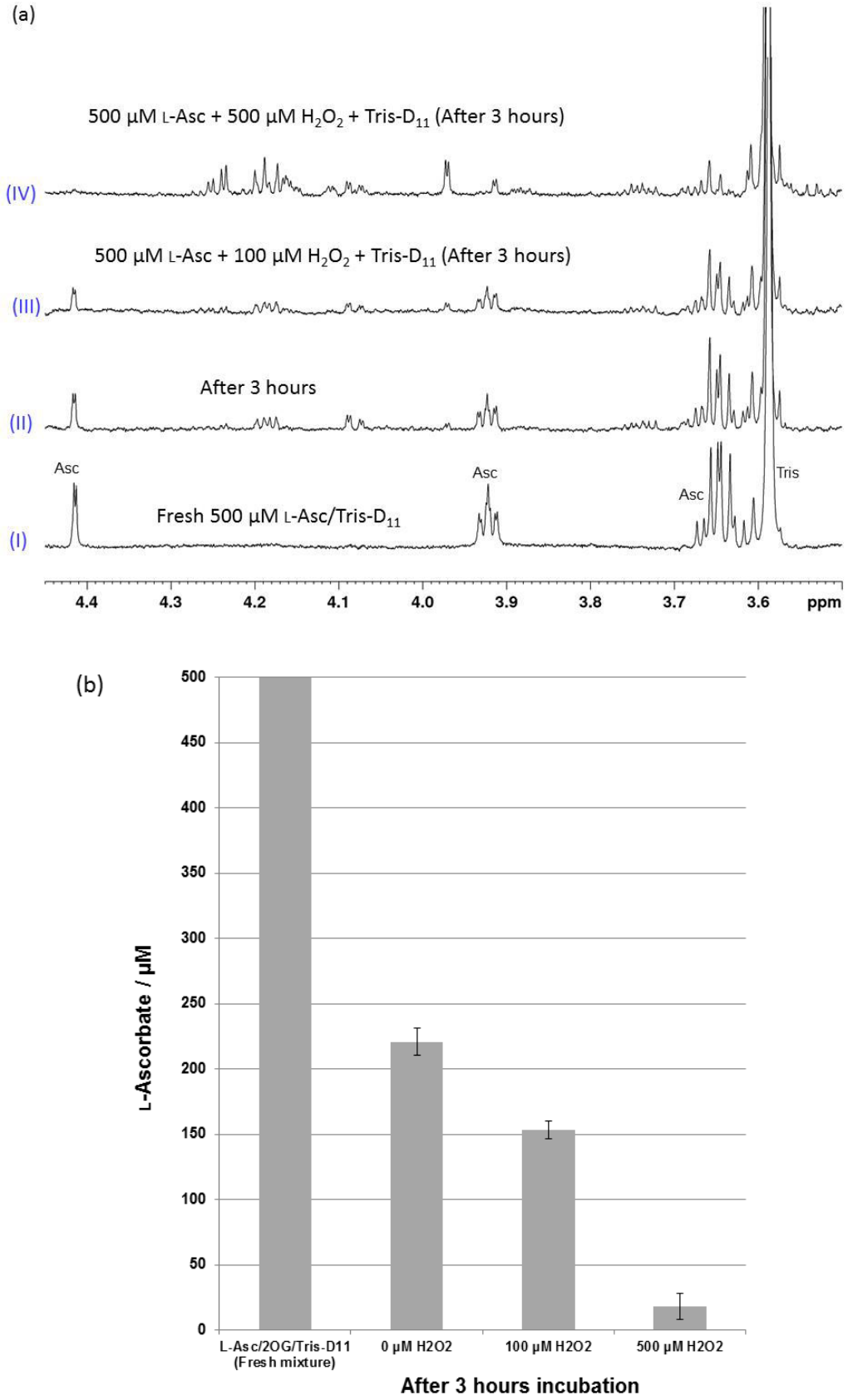


Figure 3.27 H_2O_2 speeds up L-ascorbate degradation.
(a) ^{1}H NMR spectra comparing L-ascorbate degradation in (I,II) absence of H_2O_2 (III, IV) in the presence of low (100 µM) to high amount (500 µM) of H_2O_2. The solutions were incubated for 3 hours in a 1.5 mL eppendorf tube. (b) Chart showing the level of ascorbate degradation from the ^{1}H NMR spectra in (a). The error bars represent standard deviation from 3 separate measurements.

Hence based on the results of Figure 3.26 and Figure 3.27, we propose that most (82%) of the H_2O_2 that generates from L-ascorbate oxidation into dehydro-L-ascorbate in the L-ascorbate (500 µM)/2OG (200 µM)/Tris-D_{11} buffer incubation mixture (Figure 3.19 & Figure 3.21), is recycled in the oxidative-degradation of L-ascorbate and its degradation products.

In another test for the parallel reaction of L-ascorbate with H_2O_2 in the L-ascorbate/2OG/Tris-D_{11} buffer incubation mixture, we observed that in the presence of catalase (which breaks H_2O_2 into H_2O and O_2), L-ascorbate does not appear to oxidise/degrade and hence no 2OG decarboxylation occurs (Figure 3.28a). This suggest that the process of oxidation of L-ascorbate to dehydro-L-ascorbate is primarily driven by hydrogen peroxide in agreement with the reported literature[146]. In control experiment it was shown that catalase inhibits the H_2O_2 mediated 2OG decarboxylation (Figure 3.28b,c).

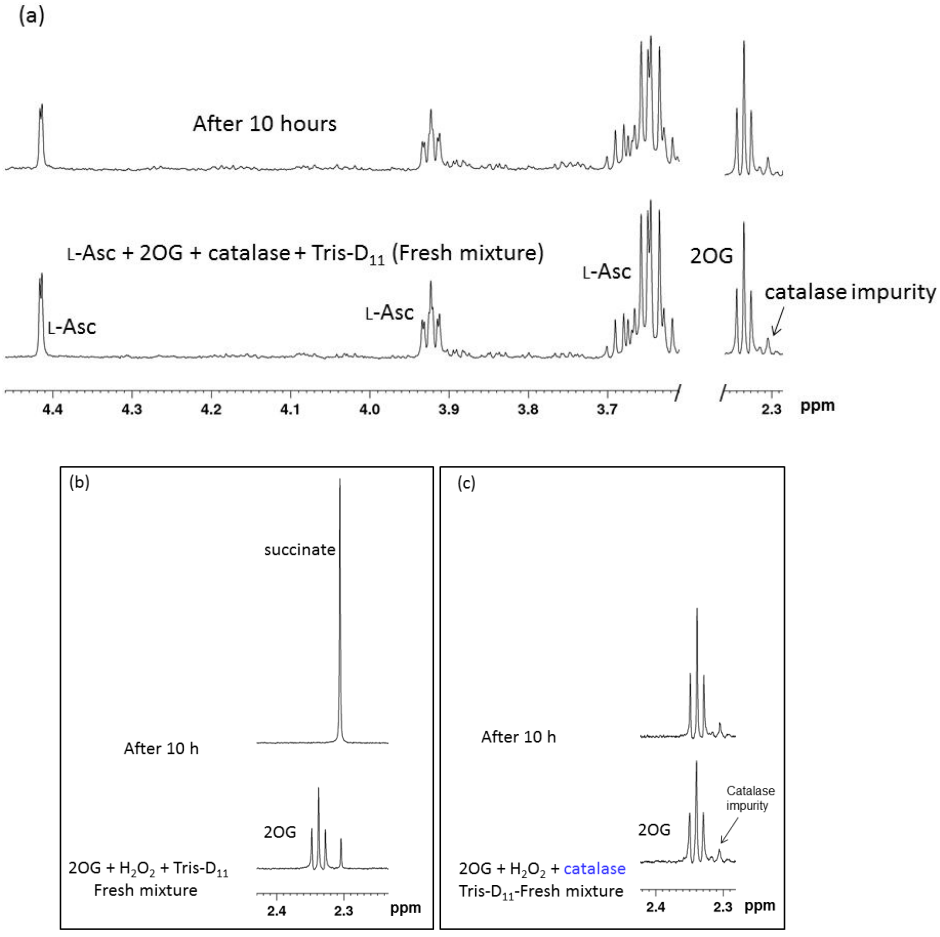


Figure 3.28 Investigating 2OG decarboxylation by L-ascorbate degradation in the presence of catalase.
(a) Overlay of ^{1}H NMR spectra (parts shown) of a freshly prepared mixture of L-ascorbate/2OG/Tris-D_{11}/catalase (bottom spectrum) and mixture after 10 hours incubation (top spectrum). (b) Control experiments; 2OG decarboxylation by H_2O_2 in the absence of catalase. (c) In the presence of catalase. The concentrations used are; 500 µM L-ascorbate, 200 µM 2OG, 500 µM H_2O_2 , 3 µL catalase (45 mg/ml; 3×45= 135 µg, batch activity 12852 units /mg; hence activity = 12852×135/1000 = 1735 units) in 50 mM Tris-D_{11}, pH 7.5 in H_2O.

To prove that oxygen is required for the generation of H_2O_2 from L-ascorbate oxidation in the assay, we incubated the L-ascorbate/2OG/Tris-D_{11} buffer mixture in anaerobic conditions for 10 hours. It was observed that under anaerobic conditions, L-ascorbate does not appear to oxidise and hence does not generate H_2O_2 and in turn no 2OG decarboxylation occurs (Figure 3.29).

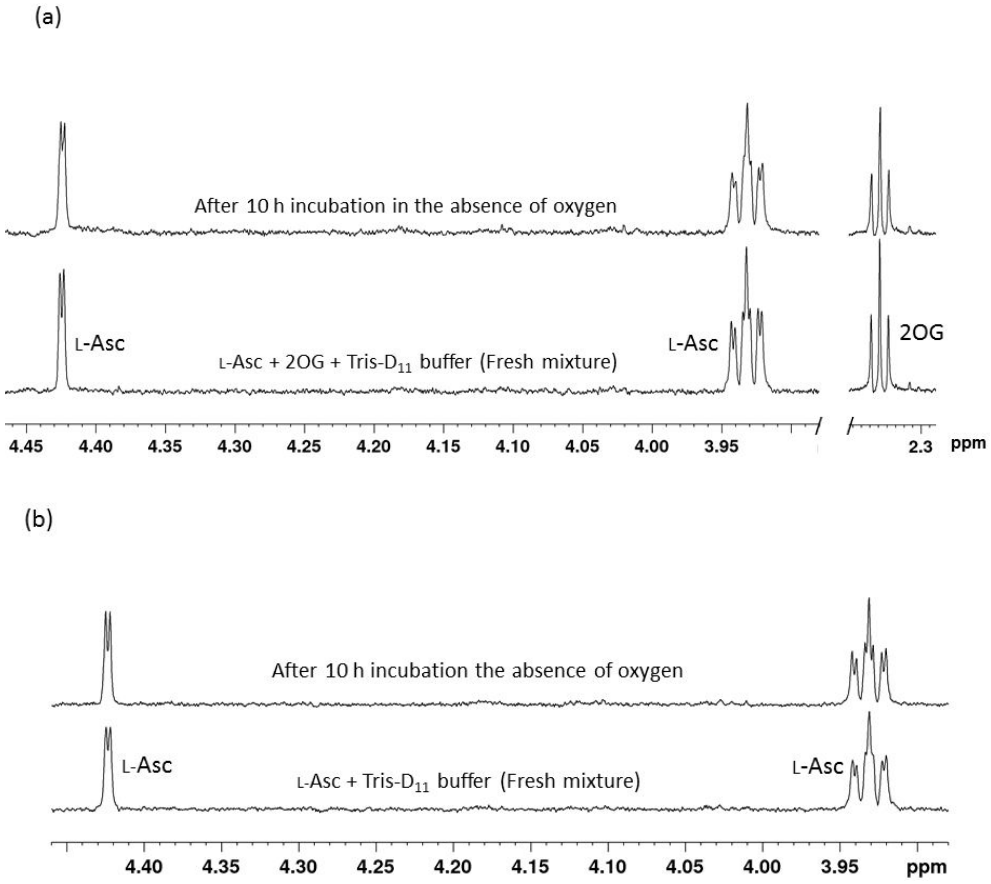


Figure 3.29 Testing 2OG decarboxylation by L-ascorbate degradation in anaerobic conditions.
(a) Overlay of ^{1}H NMR spectra (parts shown) showing a fresh mixture of L-ascorbate and 2OG in buffer (bottom) with that incubated for 10 hours (top), in anaerobic conditions. In the absence of oxygen, ascorbate does not oxidise, hence does not generate H_2O_2, and in turn no 2OG decarboxylation occurs. (b) Control experiment of (a); assay mixture in the absence of 2OG. The final assay mixture contains 500 µM L-ascorbate and 200 µM 2OG in 50 mM Tris-D_{11}, pH 7.5 in H_2O.

Similarly, in a second test for the requirement of oxygen for the generation of H_2O_2 from L-ascorbate oxidative degradation, the L-ascorbate/2OG/Tris-D_{11} buffer mixture was incubated in a 3 and 5 mm (diameter) NMR tubes in the NMR spectrometer for 10 hours. Under these conditions, the assay mixture was now left with only a limited amount of oxygen i.e only the dissolved oxygen in the solution mixture, as the flow of oxygen from outside to inside of the NMR probe (and then through the assay mixture) was almost prevented. It was observed that both the rate and level of 2OG decarboxylation is reduced due to slow L-ascorbate breakdown forming only 25 µM (3 mm tube) and 50 µM (5 mm tube) of succinate compared to ~ 90 µM

in the eppendorf tube in the presence of oxygen under our standard incubation time of 10 hours. These results thus clearly demonstrate that oxygen is the main factor responsible for L-ascorbate degradation in agreement with the reported literature work.

To test if the dark in the NMR probe or glass of the NMR tube could account for the reduced level of ascorbate oxidation/2OG decarboxylation, the L-ascorbate/2OG/Tris-D_{11} buffer mixture was incubated in plastic eppendorf in the dark as well as a mixture left in an identical size glass vial in light. After 10 hours of incubation (our experimental incubation time), in both cases, identical levels of 2OG decarboxylation were observed (Figure 3.30). These results thus clearly demonstrate that the low level of 2OG decarboxylation in the presence of L-ascorbate in the NMR tubes incubated in the NMR probe was primarily due to limited supply of oxygen to the assay mixture and neither due to the influence of either dark in the NMR probe nor due to the glass of NMR tubes.

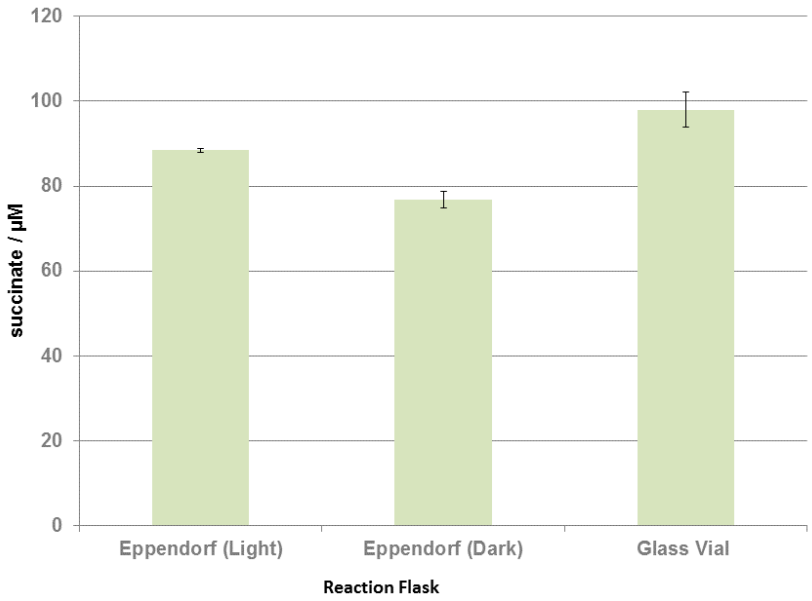


Figure 3.30 Monitoring 2OG decarboxylation by L-ascorbate degradation by incubation the reaction mixture in a plastic eppendorf tube in light, dark and in a glass vial in light.

3.5.2 Effect of different L-ascorbate concentration on 2OG decarboxylation

2OG oxygenases show various degrees of dependencies on the presence of L-ascorbic acid in the assay mixture for maximum catalytic activity. In most cases, a concentration of 0.5-1 mM of L-ascorbate is believed to be sufficient for maximum catalytic activity. However, in some cases such as in ANS, a significantly higher ascorbate concentration (40 mM) have been employed. Hence, it is important to assess as what ascorbate concentration could be

producing maximum artifacts of the ascorbate-induced 2OG turnover in the level of the uncoupled 2OG turnover in the presence of a 2OG oxygenase. We therefore performed our assay in the presence of low (50 µM) to high concentration (20 mM) of ascorbate. Our results (Figure 3.31) showed that maximum 2OG turnover occurs with 0.5-1 mM ascorbic acid concentration, suggesting that at this concentration of L-ascorbate there is a high probability of artifacts of ascorbate-induced uncoupled 2OG turnover in the enzymatic assays. The data further suggest that under our assay conditions, L-ascorbate concentration at and above 1 mM begins to alter the solution environment, possibly the pH (as indicated by changes in the chemical shift of the signals in ^{1}H NMR spectrum) to become unfavourable to promote L-ascorbate oxidative degradation and hence concomitant 2OG decarboxylation.

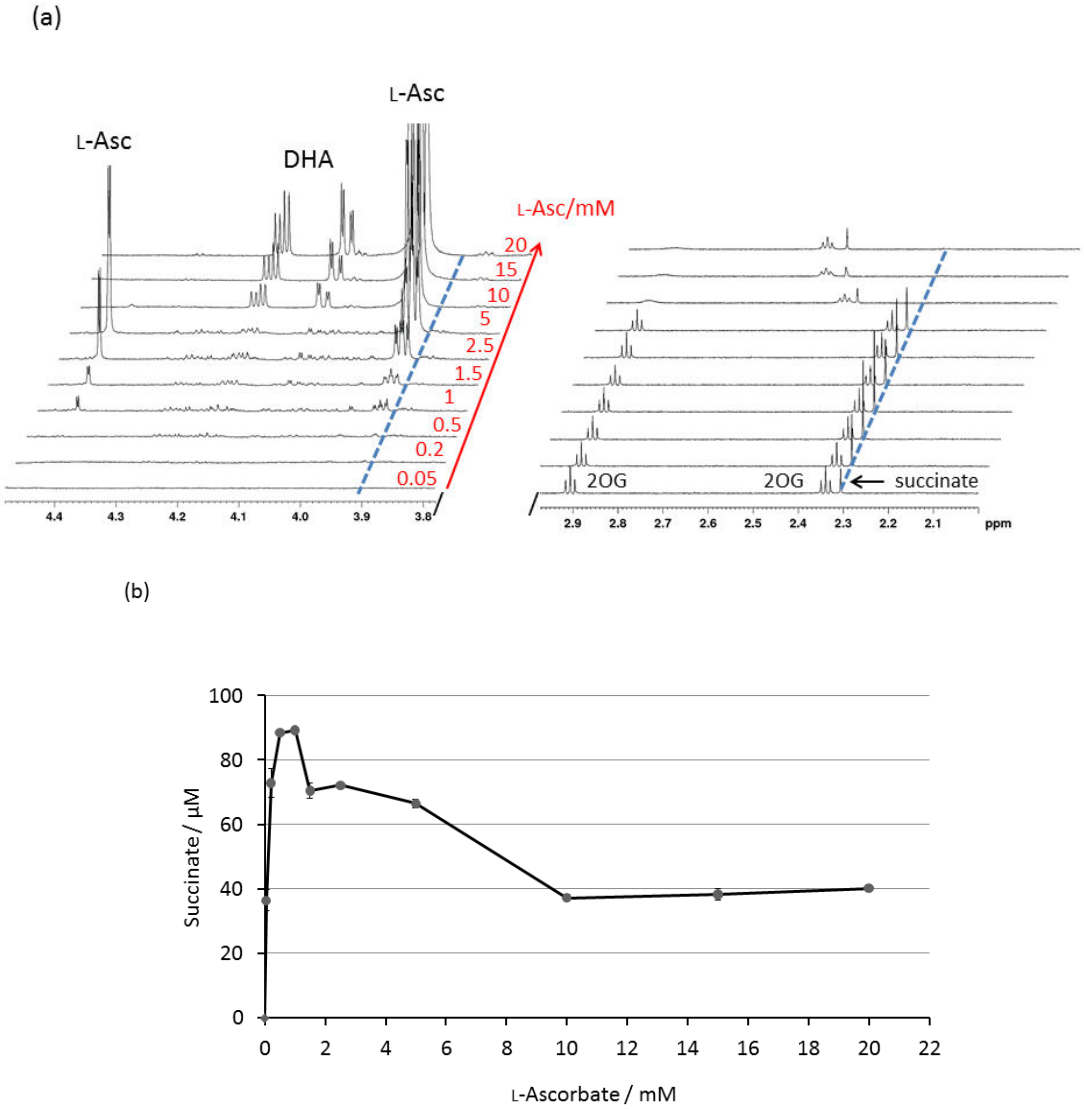


Figure 3.31 2OG decarboxylation in the presence of different L-ascorbate concentration.
(a) ^{1}H NMR spectra recorded after 10 hours incubation of a mixture of L-ascorbate, 2OG (200 µM) and 50 mM Tris-D11, pH 7.5, in H_2O. The concentration of L-ascorbate used are, 0.05, 0.2, 0.5, 1, 1.5, 2.5, 5, 10, 15 and 20 mM from bottom to top. The blue line is drawn to aid visualization of the changes in the chemical shift of the signals (b) Plot of L-ascorbate concentration vs succinate concentration.

3.5.3 Effect of different buffers on ascorbate-induced 2OG decarboxylation

The activity of a protein is generally dependent upon its incubation buffer and hence may exhibit different level of activity in different buffer systems. Most enzymatic assays employ buffer concentration ranges from 50-200 mM. At such high concentrations, the interaction of buffer with protein in some form (proton transfer, salt bridge formation or Schiff's base formation) is highly likely, and perhaps these interactions stabilise the proteins. Many buffer systems have been found to bind the enzymes and possibly play a role in the catalysis. For example, in Anthocyanidin synthase (ANS), the in *vitro* turnover activity of which is highly dependent on L-ascorbate, a 2-(*N*-morpholino)ethanesulfonic acid or MES buffer molecule was found bound in the crystal structure. The binding site has been proposed to serve as the binding site of L-ascorbic acid (MES may have displaced L-ascorbate during crystallization).[71] In the context of this study, it is possibly that a buffer system may interact with ascorbate and hence may either promote or demote the level of artifacts of ascorbate-induced 2OG turnover in the presence of a 2OG oxygenase. We therefore investigated the effect of different buffers on the ascorbate-induced 2OG decarboxylation assay. The different buffer systems tested include, sodium phosphate, *N*-(tri(hydroxymethyl)methyl)glycine (Tricine), *N*-[tris(hydroxymethyl)methyl]-2-aminoethanesulfonic (TES), 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), 1,4-piperazinediethanesulfonic acid (PIPES) and 4-morpholinepropanesulfonic acid (MOPS). The buffers were tested at 50 mM concentration at pH 7.5 in H_2O. ^1H NMR spectra were recorded after 10 hours incubation of the mixture of L-ascorbate/2OG/buffer (Figure 3.32a). It was observed that in case of sodium phosphate buffer, L-ascorbate remains stable and does not undergo oxidative degradation and hence no 2OG decarboxylation occurs. In the case of Tricine, also most of the L-ascorbate in the mixture remained undegraded, and hence only a very low amount of succinate was formed. In contrast, in case of TES, HEPES, PIPES and MOPS, the entire L-ascorbate in the mixture was found degraded after 10 hours incubation, but different amount of succinate was formed in each case (Figure 3.32). The level of succinate formed was in the order Tris-D_{11} > HEPES > PIPES > MOPS > TES. These results suggest that apart from their buffering capacity, these buffers also display chemistry towards the components of the assay mixture (ascorbate).

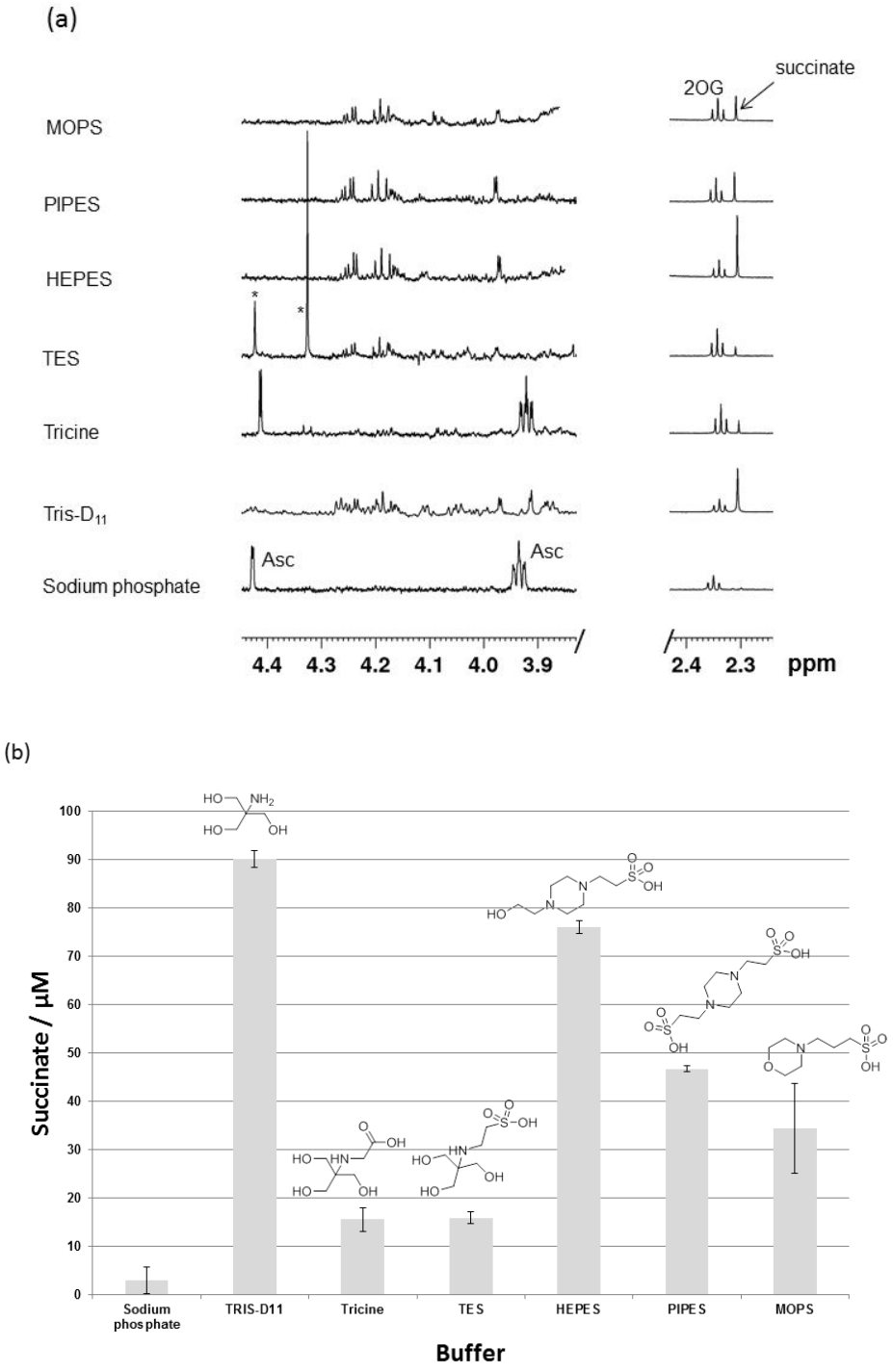


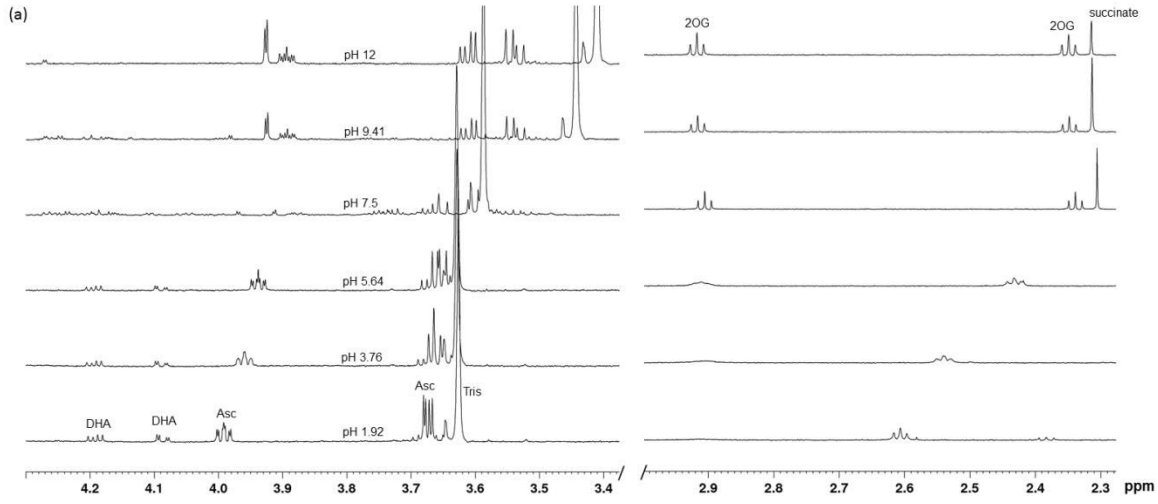
Figure 3.32. The effect of different buffer system on L-ascorbate induced 2OG decarboxylation.
(a) ^{1}H NMR spectra (part shown) recorded after 10 hours incubation of a mixture of L-Asc/2OG/buffer. (b) Chart showing the amount of succinate formed in case of each buffer system. The assay mixture contains 500 µM ascorbate, 200 µM 2OG in 50 mM buffer at pH 7.5 in H_2O.

According to these results, sodium phosphate and Tricine clearly do not appear to promote ascorbate oxidative degradation and hence 2OG turnover. This may possibly be due to the reaction of these buffers with either oxygen or H_2O_2, the only two species that can cause ascorbate oxidative degradation in the assay mixture. Although, TES does appear to promote

ascorbate degradation, however it does not result in 2OG turnover, perhaps by its (buffer) reaction with H_2O_2. Hence, in comparison to the rest of the buffers tested, the use of sodium phosphate, Tricine or TES buffer with ascorbate in the assay mixture may provide an accurate measurement of the level of uncoupled 2OG turnover in the presence of a 2OG oxygenase.

3.5.4 Effect of buffer pH on ascorbate induced 2OG decarboxylation

In order to investigate what pH of the assay mixture promotes the ascorbate-induced 2OG turnover, we incubated the ascorbate/2OG/Tris-D_{11} buffer mixture with different pH of the buffer. The buffer was tested at pH 1.92, 3.76, 5.64, 7.5, 9.41 and 12. The results are shown in Figure 3.33. From these results it appear that in the acidic pH range, i.e pH 1.92, 3.76, 5.64, L-ascorbate does not appear to undergoes oxidative degradation and hence does not cause 2OG decarboxylation, whereas, in the neutral and alkaline pH range, i.e 7.5, 9.41, 12, L-ascorbate undergoes oxidative degradation and results in 2OG decarboxylation, with maximum turnover occurs between 7.5-9. As enzymatic (2OG oxygenase) assays are performed at physiological pH (7-7.5), therefore at this pH, there is a high probability of artifacts of ascorbate-induced uncoupled 2OG turnover in the enzymatic assays.



(b)

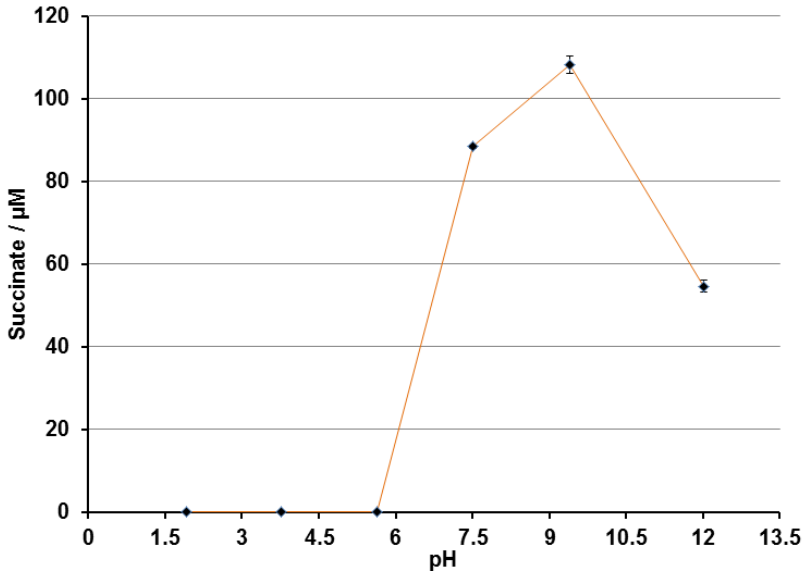


Figure 3.33 The effect of different pH of Tris-D_{11} on the ascorbate induced 2OG decarboxylation.
(a) ^{1}H NMR recorded after 10 hours incubation of L-ascorbate/2OG/Tris-D_{11} of different pH. (b) Plot of level of succinate formation at different pH of Tris-D_{11}. The assay mixture contains 500 µM ascorbate, 200 µM 2OG, in 50 mM buffer in H_2O. The error bars represents standard deviation from three separate measurements.

3.5.5 Effect of different metals on L-ascorbate induced 2OG decarboxylation

2OG oxygenases are metalloenzymes that requires Fe^{II} for catalysis. Other metals such as Zn^{II}, Ni^{II}, Mn^{II} and Cu^{II} etc. are routinely employed in the binding assays of 2OG oxygenases instead of catalytic Fe^{II} to prevent the catalysis. In general, also, most of the reagents used in the assays or chemical reactions are used as supplied by the manufacturer. These reagents are sometimes contaminated with traces of metals. Even in trace amounts, these metals are potential sources of un-wanted chemical reactions and hence their presence may lead to ambiguity in the actual experimental results. We therefore decided to explore the effect of addition of different metals such as Na^{I}, K^{I}, Mg^{II}, Ca^{II}, Fe^{II}, Zn^{II}, Cu^{II}, Ni^{II}, Co^{II} and Mn^{II} on the L-ascorbate induced 2OG decarboxylation (Figure 3.34). It was observed that in all these samples the entire concentration of L-ascorbate in the assay mixture was found degraded after our standard incubation period of 10 hours, however, different amounts of succinate was formed in the presence of different metals. In cases of both Fe^{II} and Cu^{II}, significantly lower amounts of succinate were formed as compared to the assay mixture without a metal, whereas in case of Zn^{II}, relatively higher amount of succinate was formed in comparison to the assay mixture without a metal.

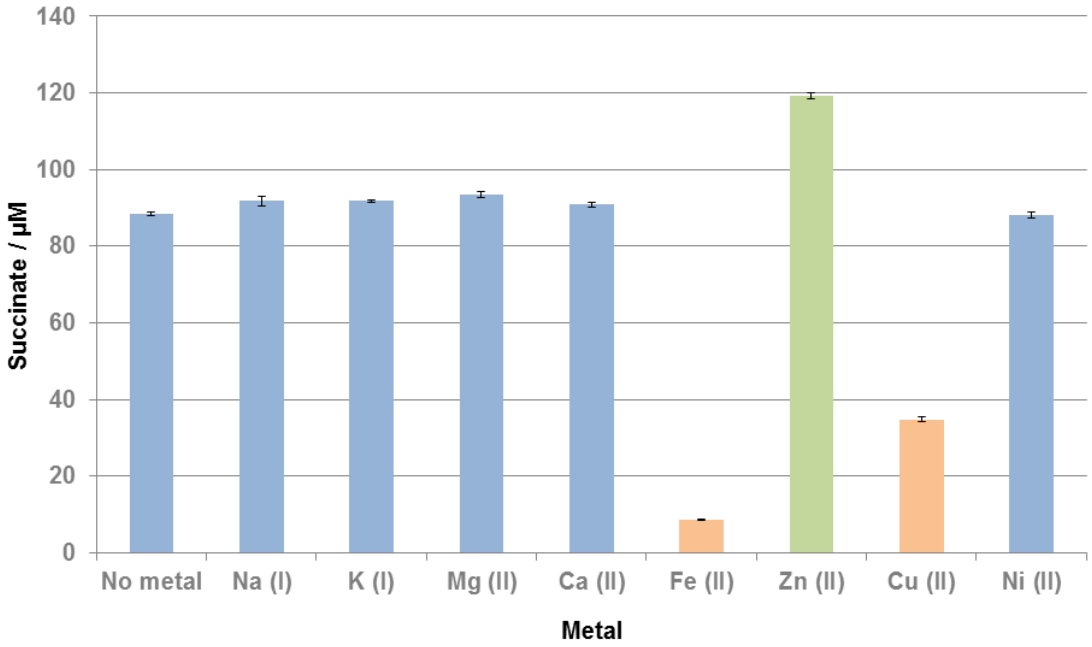


Figure 3.34 2OG decarboxylation by L-ascorbic acid degradation in the presence of different metals.
Chart showing the amount of succinate formed by incubation of 500 µM L-ascorbic/200 µM 2OG/100 µM metal/Tris-D_{11} buffer mixture for 10 hours. The error bars represents standard deviation from three separate measurements.

Based on our proposal that it is H_2O_2 generated from L-ascorbate oxidative degradation that causes 2OG decarboxylation; that is the amount of succinate formation is directly proportional to the amount of H_2O_2 in the assay mixture. Therefore in order to evaluate the mechanism by which Fe^{II} decreases the level of 2OG turnover, we studied the 2OG decarboxylation by H_2O_2 in the presence of Fe^{II} (Figure 3.35) It was revealed that Fe^{II} reduced the rate and level of succinate formation most likely by breaking down H_2O_2 into hydroxyl radical and hydroxyl anion possibly *via* Fenton chemistry (Figure 3.36). In our experiments, the copper metal was used as $CuCl_2$. Based on the reported literature, we speculate that the mechanism of breakdown of H_2O_2 by Cu^{II} (Figure 3.36) perhaps is either *via* Fenton chemistry,[147] or through its reaction with superoxide anion radical[148] that generates from the reaction of ascorbic acid and oxygen, which then readily forms H_2O_2. The reaction of Cu^{II} with superoxide anion possibly forms Cu^{I} and O_2, thus prevent the superoxide anion radical from forming H_2O_2.

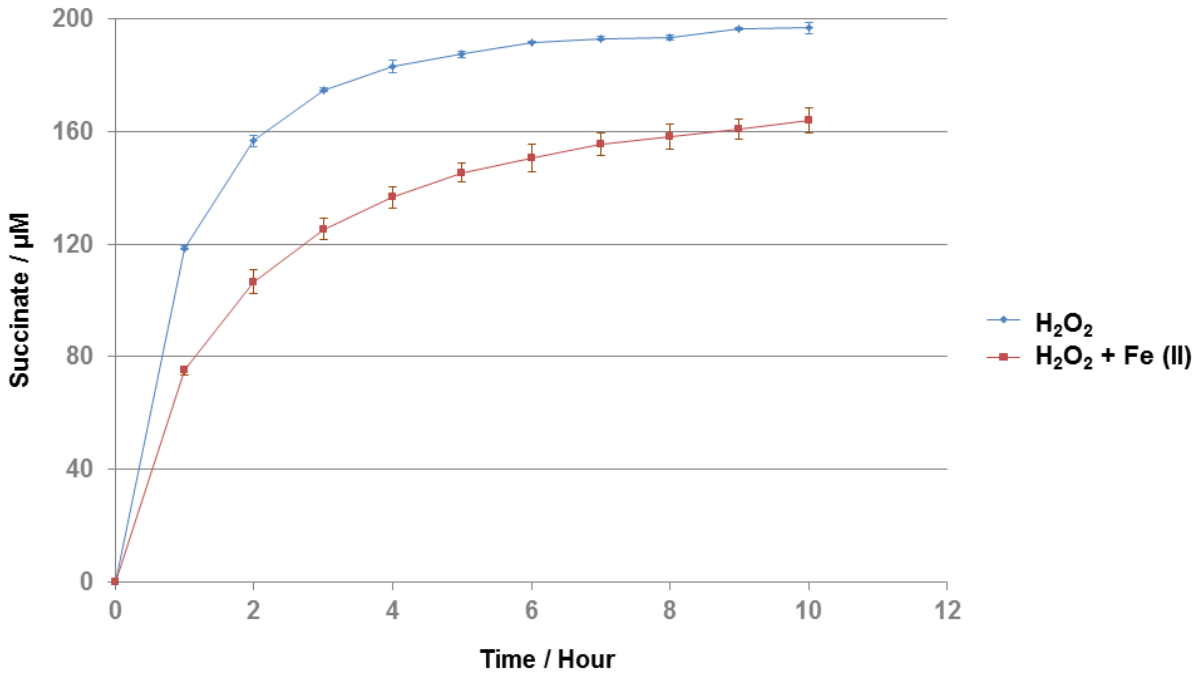


Figure 3.35 2OG decarboxylation by H_2O_2 in the absence and presence of Fe (II).
^{1}H NMR time course profiles of 2OG decarboxylation by H_2O_2 in the absence (blue curve) and presence (brown curve) of Fe (II). The assay mixture contains 200 µM 2OG/500 µM H_2O_2/100 µM Fe (II) /50 mM Tris-D11, pH 7.5, H_2O.

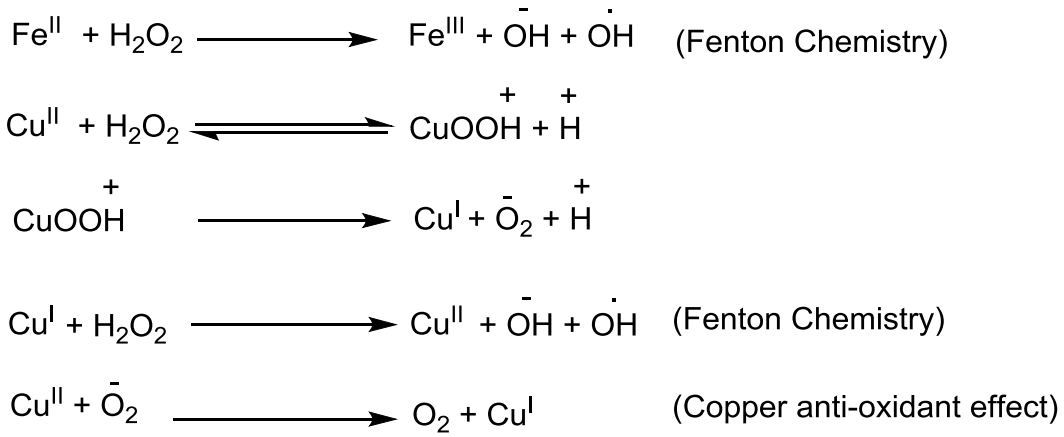


Figure 3.36 Breakdown of H_2O_2 *via* Fenton chemistry

To investigate the concentration of Zn^{II} that may yield high level of 2OG turnover; the 500 µM ascorbate/200 µM 2OG/Tris-D_{11} assay mixture was incubated with 25 µM, 100 µM, 500 µM, 2.5 mM, and 5 mM of Zn^{II}. The results (Figure 3.37, blue curve) obtained suggest that maximum 2OG turnover occurs with 500 µM of Zn^{II} (succinate; 136 µM vs 88 µM without Zn^{II}). The concentration of Zn^{II} above 500 µM however did not further increase the level of 2OG turnover. These results suggest that the mode of action of Zn^{II} in this process may be *via* some type of its interaction with ascorbate, possibly in a 1:1 ratio. Since earlier in section

3.5.2, we noted that the concentration of L-ascorbate in the (ascorbate/200 µM 2OG/Tris) mixture at and above 1 mM does not fully degrade during our standard incubation time of 10 hours, and result in identical level of 2OG turnover (73 µM succinate) as with 500 µM of ascorbate (88 µM succinate), we envisaged that the higher 2OG turnover in the assay with Zn^{II} compared to without Zn^{II}, may perhaps be due to the rapid ascorbate degradation with Zn^{II} than without Zn^{II}. To test this catalytic role of Zn^{II}, the 1.5 mM ascorbate/200 µM 2OG/Tris assay mixture was incubated for 10 hours with 0.5 µM, 2.5 mM and 5 mM of Zn^{II}. Interestingly, it was observed that with 2.5 mM of Zn^{II}, the entire 1.5 mM ascorbate in the assay mixture was found degraded after 10 hours incubation (Figure 3.38) and simultaneously the level of 2OG turnover was also significantly increased from 73 µM succinate (with 1.5 mM ascorbate) to ~ 164 µM succinate (Figure 3.37, brown curve). These results thus clearly demonstrate that Zn^{II} speeds up ascorbate oxidative degradation and thus increases the level of 2OG turnover.

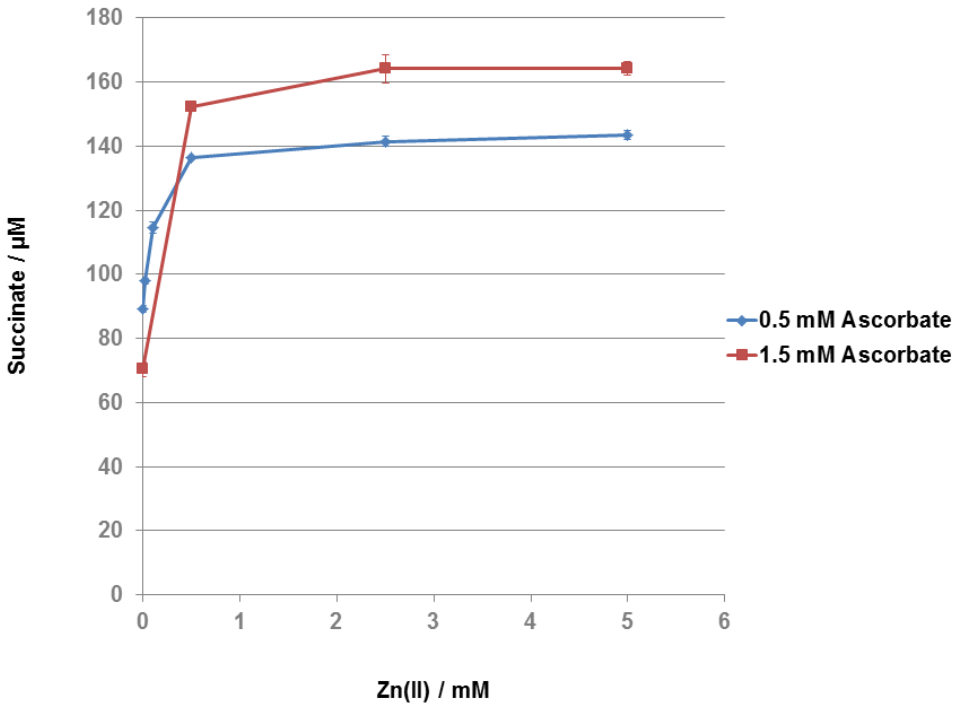


Figure 3.37 2OG decarboxylation by L-ascorbate degradation in the presence of different concentration of Zn (II).
The concentrations of Zn (II) used are 25 µM, 100 µM, 500 µM, 2.5 mM, and 5 mM. The error bars represent standard deviations from three separate measurements.

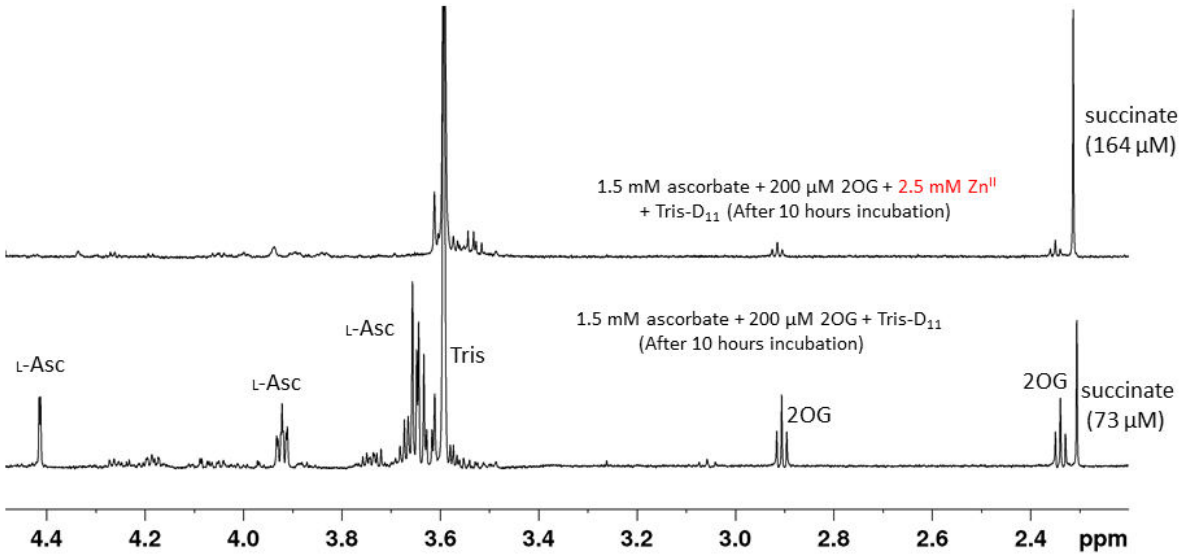


Figure 3.38 Zn(II) speeds up L-ascorbate oxidative degradation.
Bottom is the ^{1}H NMR spectrum of the assay mixture of 1.5 mM ascorbate, 200 µM 2OG in 50 mM, pH 7.5 in H_2O after 10 hours incubation. Top is the ^{1}H NMR spectrum of the corresponding mixture in the presence of 2.5 mM Zn(II). Note that in the top spectrum all of the L-ascorbate is degraded and hence more succinate is produced as compared to the spectrum below.

The direct reaction of dioxygen with biomolecules e.g ascorbate is spin forbidden, because ground state dioxygen is a biradical. Hence these reactions proceed very slowly. Transition metals can exist in several spin-states and thus they are able to relieve the spin restriction of dioxygen. These metals allow the simultaneous binding of a biomolecule and di-oxygen, thereby providing a bridge between O_2 and biomolecule and hence speed up the oxidation (Figure 3.39).

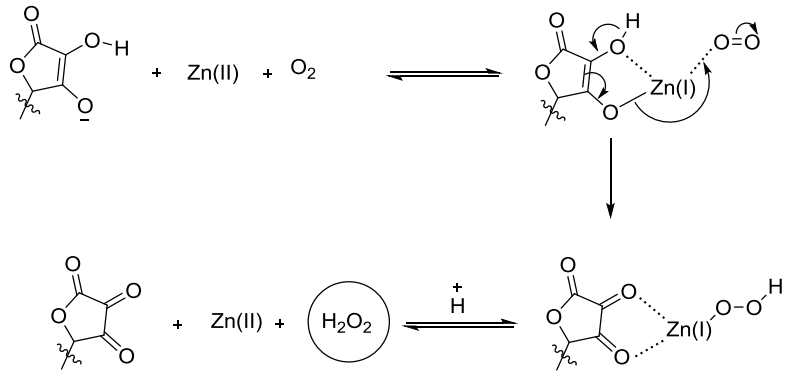
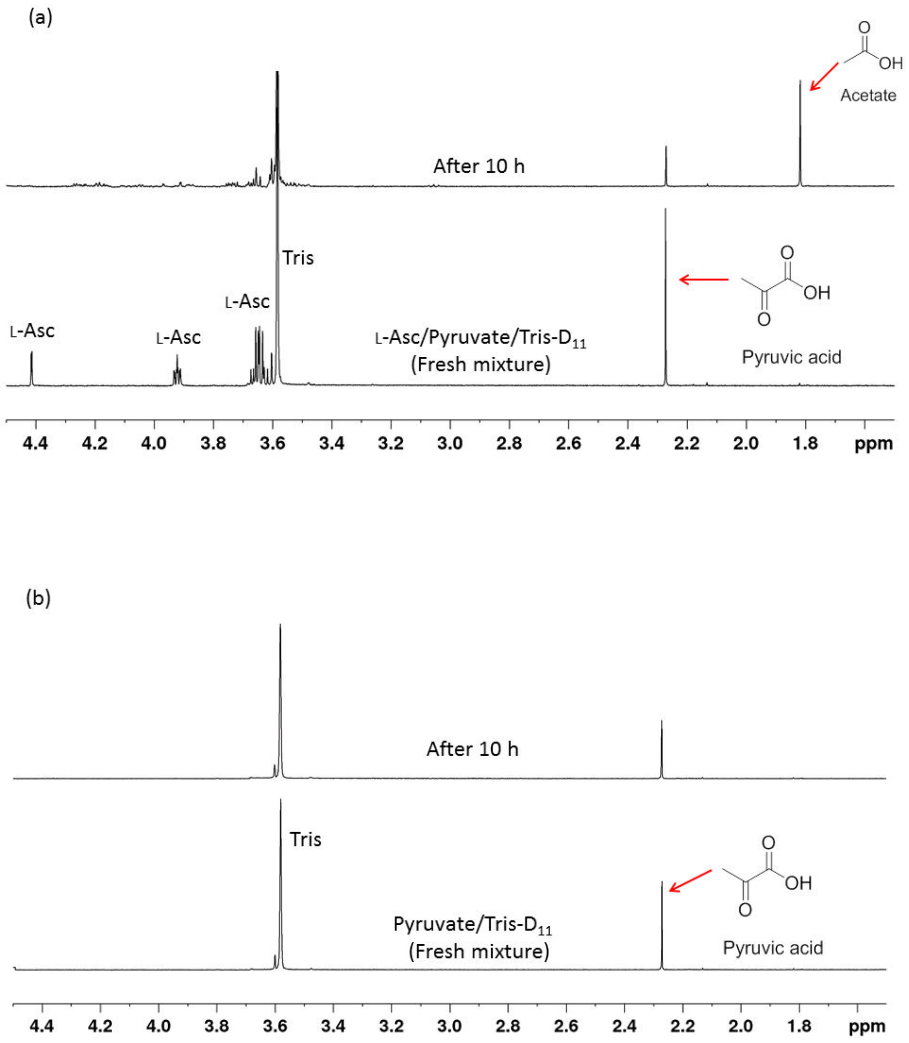


Figure 3.39 Proposed mechanism for the rapid oxidative degradation of ascorbate with Zn(II).

3.6 Decarboxylation of other α-ketoacids by L-ascorbate degradation

To explore if other biologically important α-ketoacids could also undergo simultaneous decarboxylation with L-ascorbate degradation, we tested pyruvic acid (Figure 3.40), oxaloacetic acid (Figure 3.41) and 4-hydroxyphenyl pyruvic acid (HPP) (Figure 3.42). It was observed that all these α-ketoacids undergoe simultaneous decarboxylation with L-ascorbate degradation. Control experiments of these α-ketoacids with H_2O_2 produced identical decarboxylation compounds as observed with ascorbate degradation. Also, it was shown that similar to 2OG, these α-ketoacids also remain stable in buffer alone under our standard incubation period of 10 hours. These observations further support our proposal that H_2O_2 is the specie produced from ascorbate oxidation/degradation and is the actual decarboxylating agent towards 2OG.



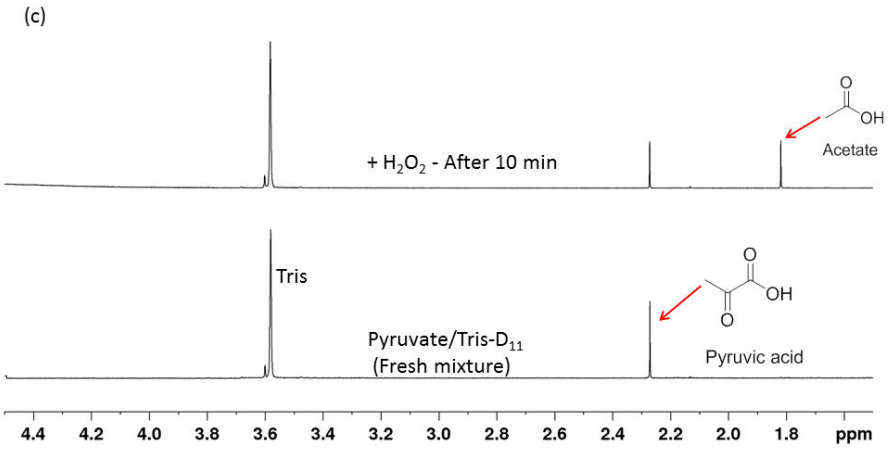
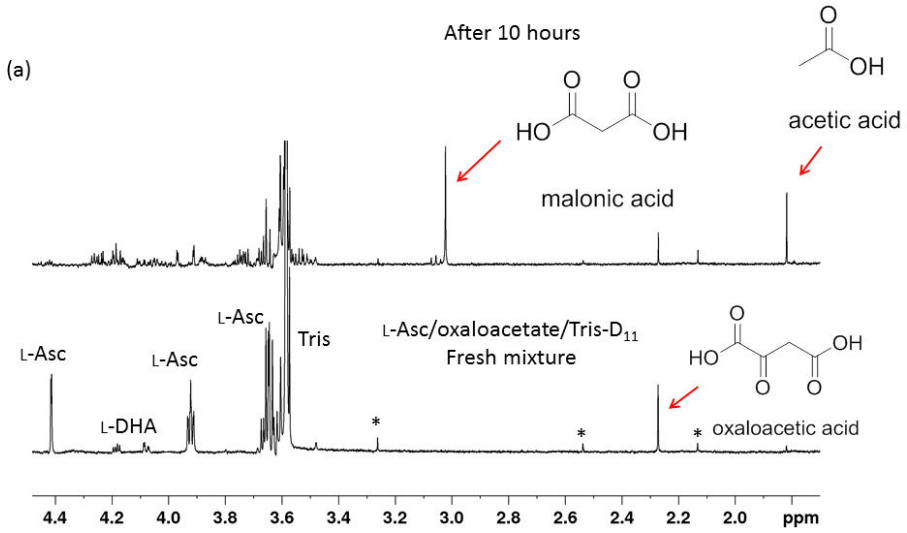
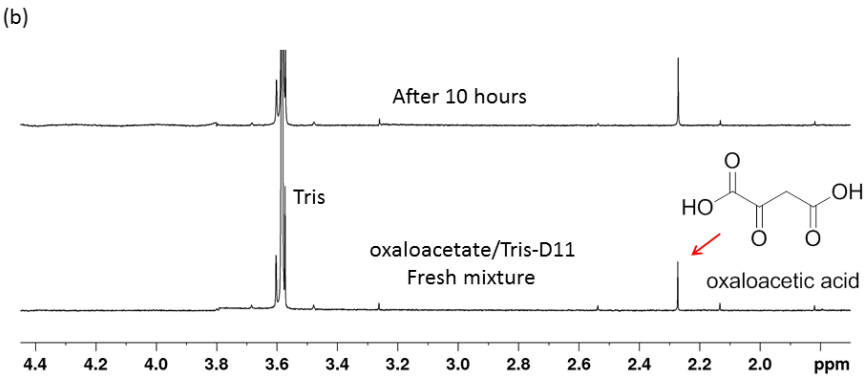


Figure 3.40 Pyruvic acid decarboxylation by L-ascorbic acid degradation.
(a) ^{1}H NMR spectra (part shown) showing a fresh mixture of L-ascorbate/pyruvic acid/Tris-D_{11} (bottom) and mixture after 10 hours incubation (top). (b) Control experiment to show that pyruvic acid remains stable after incubation with Tris-D_{11} for 10 hours. (c) Control experiment to show that pyruvic acid undergoes decarboxylation by reaction with H_2O_2. The concentration used are 500 µM L-ascorbate, 200 µM pyruvic acid, 500 µM H_2O_2 in 50 mM Tris-D_{11}, pH 7.5 in H_2O.





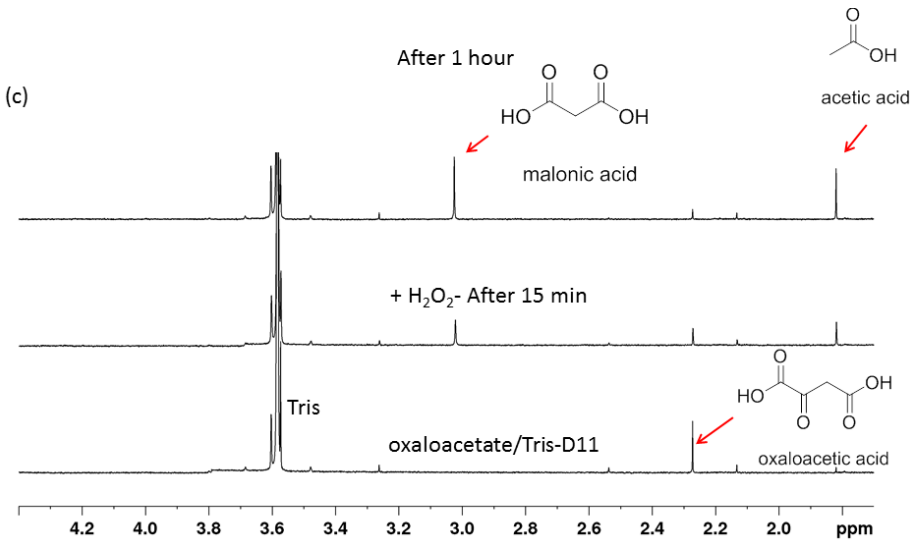
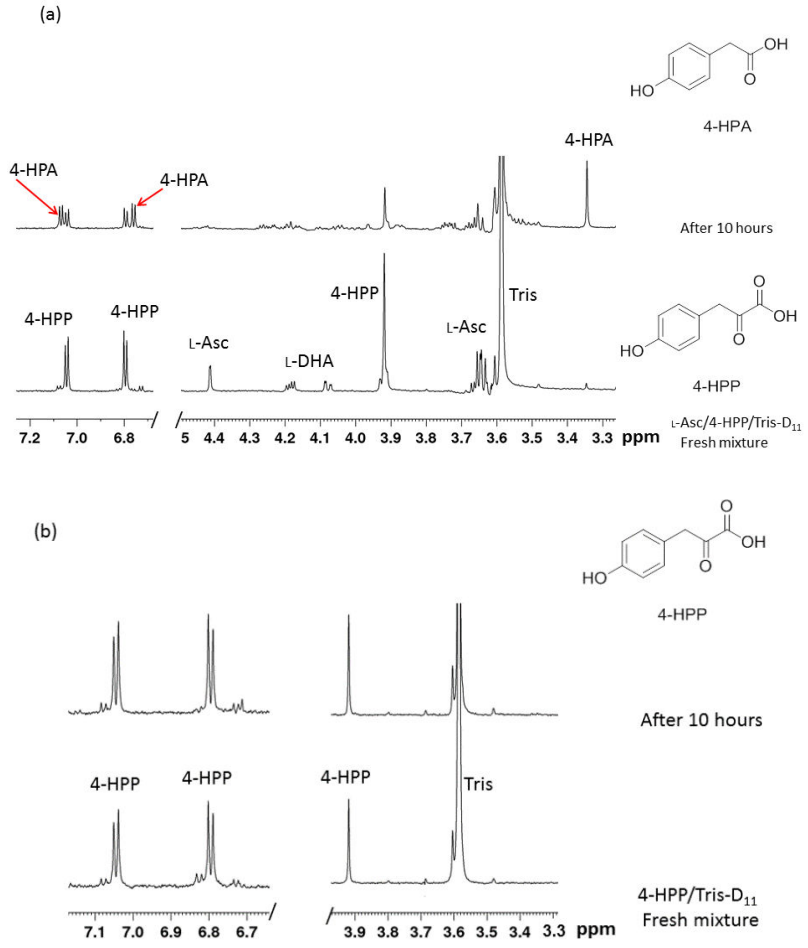


Figure 3.41 Oxaloacetic acid decarboxylation by L-ascorbic acid degradation.
(a) ^{1}H NMR spectra (part shown) showing a fresh mixture of L-ascorbate/oxaloacetic acid/Tris-D_{11} (bottom) and mixture after 10 hours incubation (top). The asterisk represent signals of the impurities in the assay solution (b) Control experiment to show that oxaloacetic acid remains stable after incubation with Tris-D_{11} for 10 hours. (c) Control experiment to show that oxaloacetic acid undergoes decarboxylation by reaction with H_2O_2. The concentration used are 500 µM L-ascorbate, 200 µM oxaloacetic acid, 500 µM H_2O_2 in 50 mM Tris-D_{11}, pH 7.5 in H_2O.



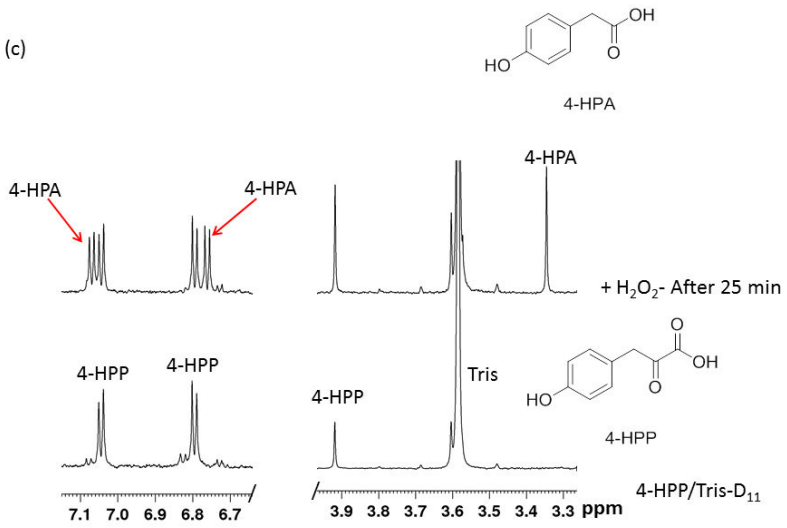


Figure 3.42 4-Hydroxyphenyl pyruvic acid (HPP) decarboxylation by L-ascorbic acid degradation.
(a) ^{1}H NMR spectra (part shown) showing a fresh mixture of L-ascorbate/4-Hydroxyphenyl pyruvic acid /Tris-D_{11} (bottom) and mixture after 10 hours incubation (top). (b) Control experiment to show that 4-Hydroxyphenyl pyruvic acid remains stable after incubation with Tris-D_{11} for 10 hours. (c) Control experiment to show that 4-Hydroxyphenyl pyruvic acid undergoes decarboxylation by reaction with H_2O_2. The concentration used are 500 µM L-ascorbate, 200 µM 4-hydroxyphenylpyruvic acid, 500 µM H_2O_2 in 50 mM Tris-D_{11}, pH 7.5 in H_2O.

In addition, it was also observed that Fe^{II} significantly reduces the level of pyruvate decarboxylation as observed in case of 2OG (Figure 3.43). Similarly, in the presence of Zn^{II}, the increase in the level pyruvate decarboxylation was relatively less as compared to 2OG (Figure 3.43). Similar to 2OG, catalase was also found to prevent the L-ascorbate oxidative degradation and concomitant pyruvate decarboxylation (Figure 3.43).

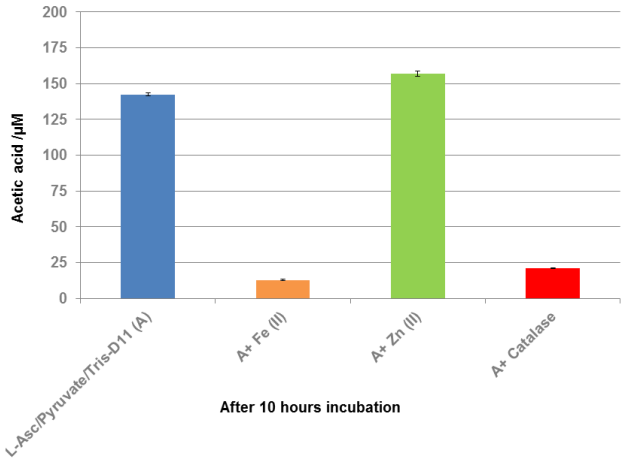


Figure 3.43 Monitoring the effect of Fe^{II}, Zn^{II} and catalase on pyruvic acid decarboxylation by L-ascorbate degradation.
Chart showing the level of acetate acid formed after 10 hours incubation of (blue): mixture of L-ascorbate/pyruvic acid/Tris-D_{11} buffer. (orange): mixture of L-ascorbate/pyruvic acid/Fe^{II}/Tris-D_{11} buffer: (green) mixture of L-ascorbate/pyruvic acid/Zn^{II}/Tris-D_{11} buffer. (Red): mixture of L-ascorbate/pyruvic acid/catalase/Tris-D_{11} buffer. The concentration used are 500 µM L-ascorbate, 200 µM pyruvic acid, 100 µM metal in 50 mM Tris-D_{11}, pH 7.5 in H_2O.

Other hydroxy acids such as citric acid (Figure 3.44), DL-isocitric acid (Figure 3.45) and maleic acid (Figure 3.46) and 1,3 dicarbonyl compounds including benzoylacetone, methylacetoacetate and 2,4 pentanedione were also tested, however they were found to be stable with the degradation of L-ascorbic acid, suggesting the requirement of an α-keto acidic functionality for decarboxylation by H_2O_2.

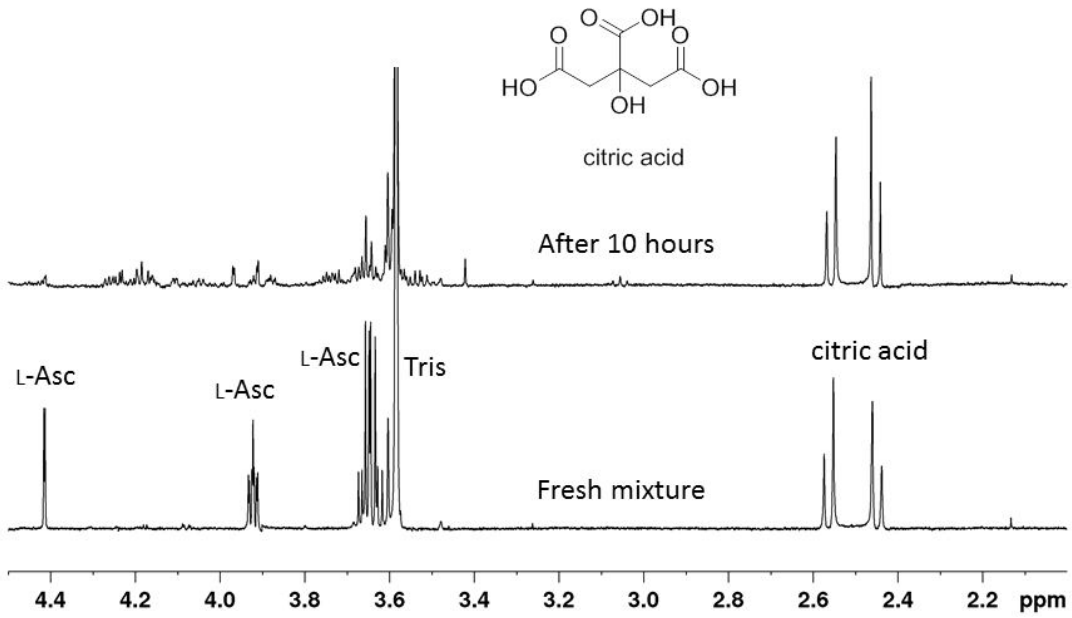


Figure 3.44 Investigating the effect of L-ascorbate degradation on citric acid.
The assay mixture contains 500 µM L-ascorbate, 200 µM citric acid, in 50 mM Tris-D_{11}, pH 7.5 in H_2O.

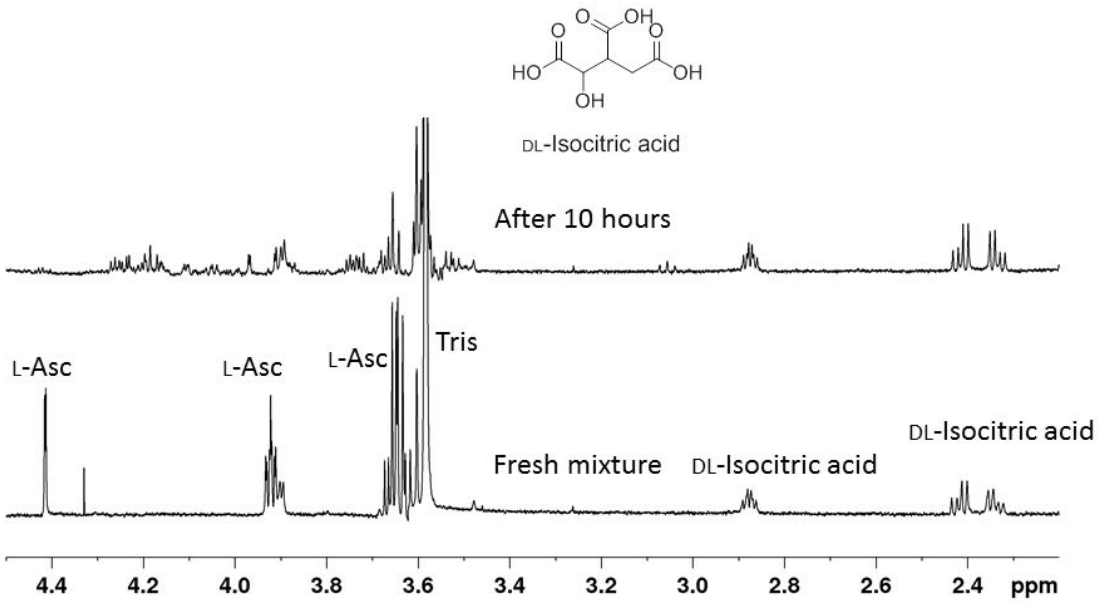


Figure 3.45 Investigating the effect of L-ascorbate degradation on DL-isocitric acid.
The assay mixture contains 500 µM L-ascorbate, 200 µM DL-isocitric acid, in 50 mM Tris-D_{11}, pH 7.5 in H_2O.

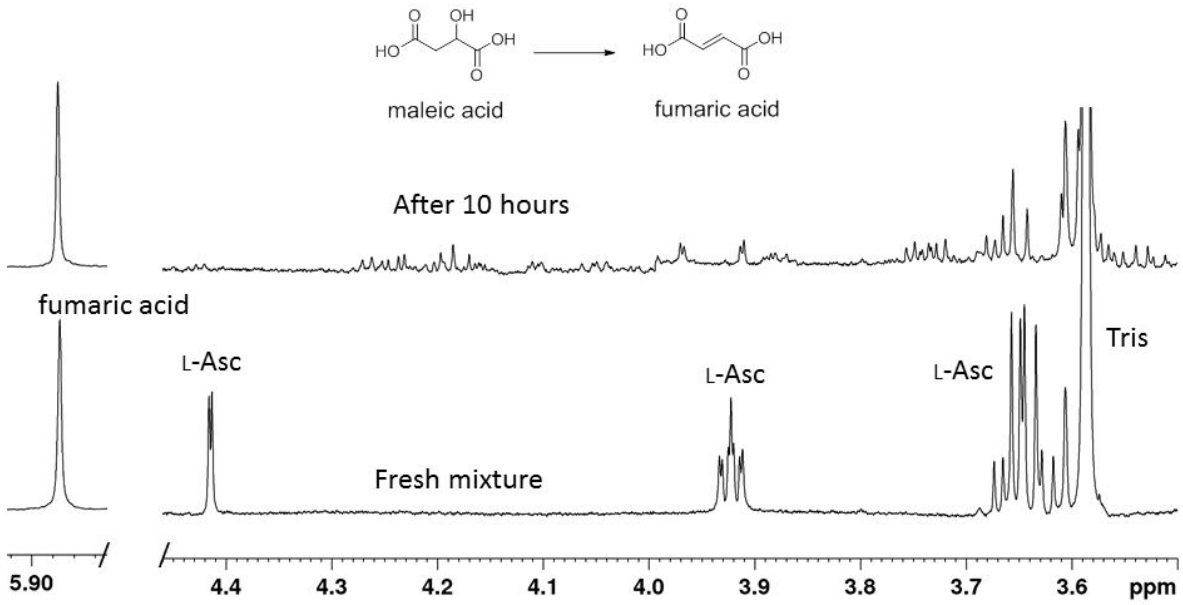


Figure 3.46 Investigating the effect of L-ascorbate degradation on maleic acid.
The assay mixture contains 500 µM L-ascorbate, 200 µM maleic acid, in 50 mM Tris-D_{11}, pH 7.5 in H_2O.

3.7 2OG decarboxylation in the presence of other reducing agents

Finally, other biological reducing agents/anti-oxidants were explored for their H_2O_2 generating ability from their oxidation by oxygen and hence causing non-enzymatic 2OG decarboxylation (Figure 3.47). The results revealed that similar to L-ascorbic acid, the level of 2OG decarboxylation is directly related on the level of oxidation of these compounds by oxygen. The higher the level of oxidation, the higher 2OG turnover and vice versa. It was observed that under our standard experimental incubation period of 10 hours, baicalein (Figure 3.48) and DTT (Figure 3.49) were rapidly oxidised by oxygen and hence produced more succinate, 115 µM and 72 µM respectively. Whereas, propyl gallate (Figure 3.50), 3,4-dihydroxybenzoic acid or protocatechuic acid (PCA) (Figure 3.51), catechol (Figure 3.52) and glutathione (GSH) (Figure 3.53), due to their significantly slower oxidation (by oxygen), produced 24 µM, 5 µM, 22 µM and 6 µM succinate respectively. In the case of glutathione, Zn^{II} was also added, in case it will speed up its oxidation and hence increase the level of succinate formation, however, no such catalytic role for Zn^{II} was observed (Figure 3.54). It is possible that Zn^{II} may be not interacting with glutathione in the same way as with ascorbate to promote its oxidation. In case of all these reducing agents, unlike dehydro-L-ascorbic acid, the oxidised forms of all these reducing agents were stable under our assay incubation conditions. In the case of Tris(2-carboxyethyl)phosphine (TCEP), we did not detect the

formation of its oxidised form in the spectrum, and hence no 2OG decarboxylation was detected (Figure 3.55).

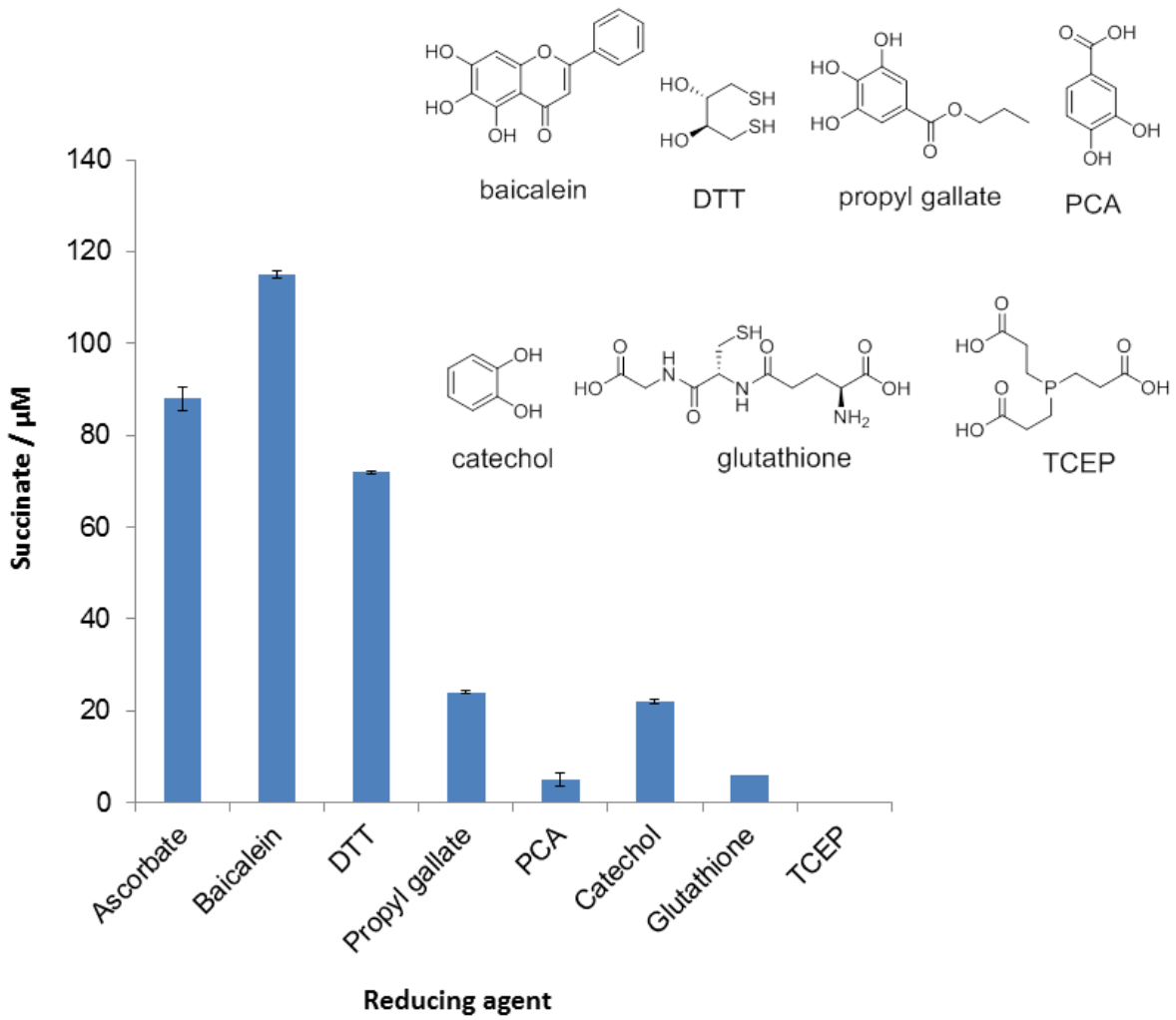


Figure 3.47 Chart showing the Level of 2OG decarboxylation by reducing agents other than ascorbate.
Under the same condition with ascorbic acid, 88 µM of succinate was formed. The concentrations used are 200 µM 2OG and 500 µM reducing agent in 50 mM Tris-D_{11}, pH 7.5 in H_2O. The assay mixture was incubated for 10 hours in an eppendorf tube. The error bars represent standard deviation from three separate measurements.

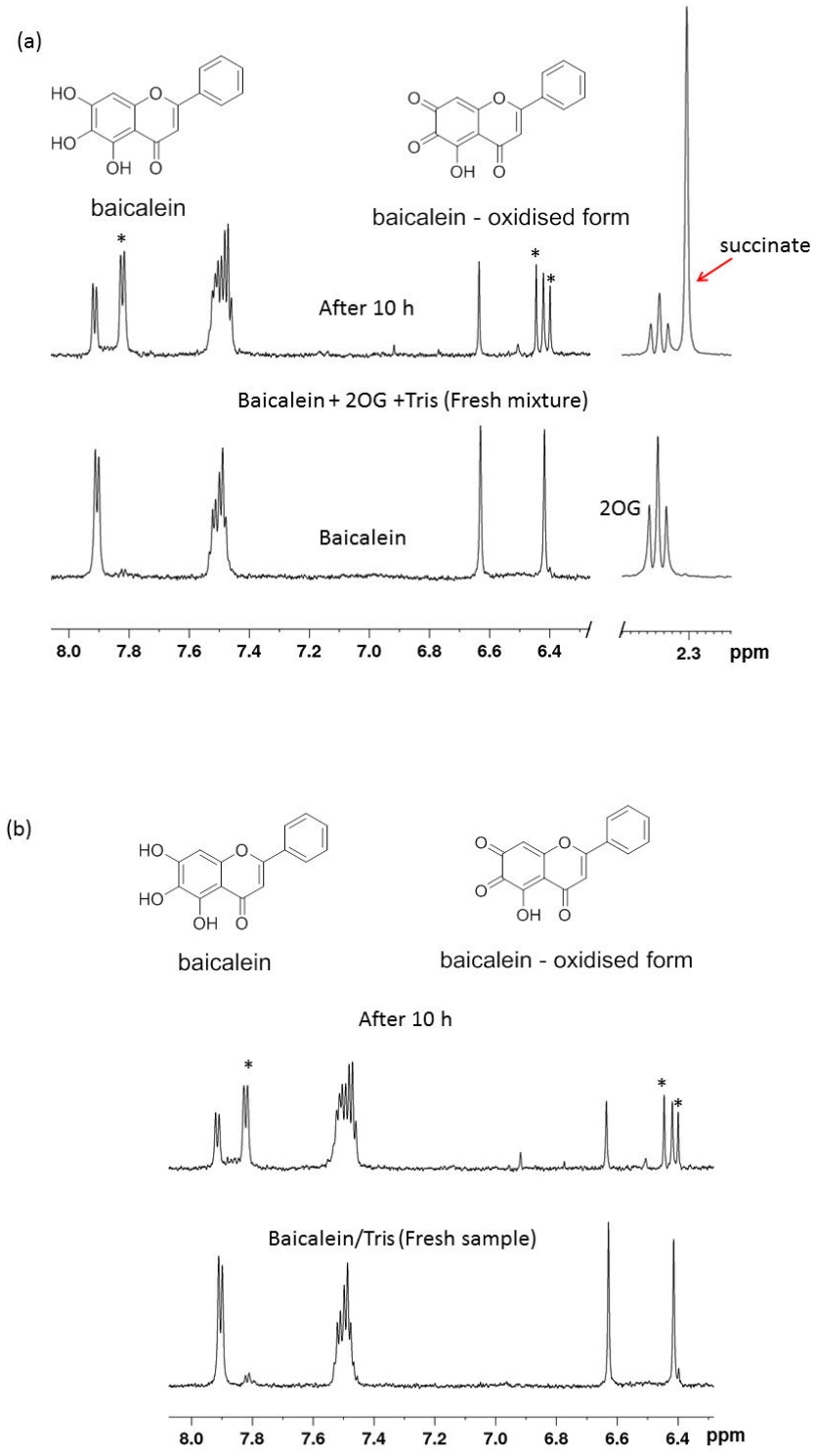


Figure 3.48 2OG decarboxylation by baicalein auto-oxidation.
(a) ^{1}H NMR spectra (part shown) showing a fresh mixture of baicalein/2OG/Tris-D_{11} (bottom) and mixture after 10 hours incubation (top). (b) Control experiment to show that baicalein undergoes auto-oxidation in the presence of oxygen. The asterisks are believed to represent the oxidised form of baicalein. The concentration used are 500 µM baicalein, 200 µM 2OG in 50 mM Tris-D_{11}, pH 7.5 in H_2O.

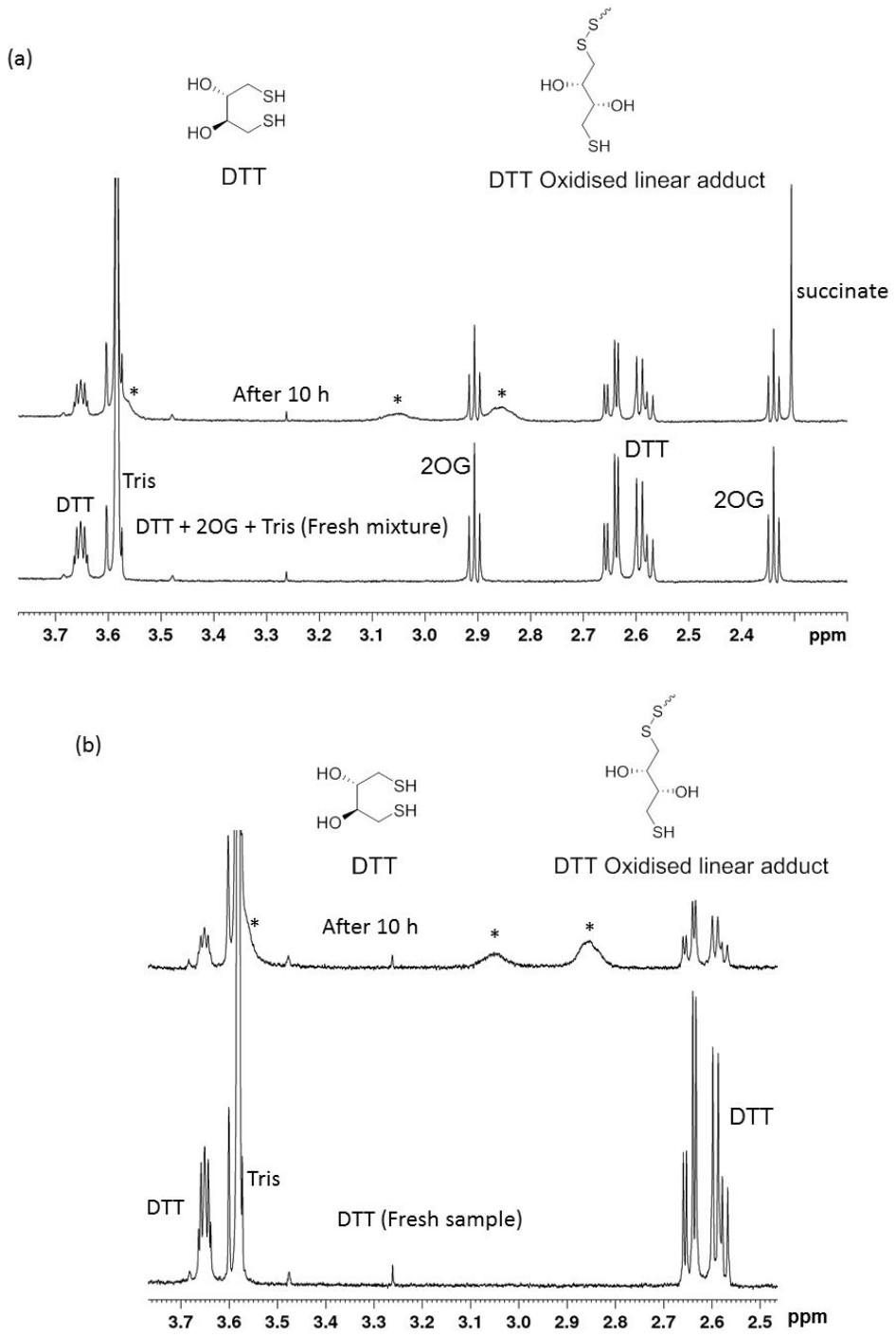


Figure 3.49 2OG decarboxylation by DTT auto-oxidation.
(a) ^{1}H NMR spectra (part shown) showing a fresh mixture of DTT/2OG/Tris-D_{11} (bottom) and mixture after 10 hours incubation (top). (b) Control experiment to show that DTT undergoes auto-oxidation in the presence of oxygen. The broad resonances marked with asterisks are believed to be of the oxidised form of DTT, possibly the linear adduct polymeric form. The concentration used are 500 μM DTT, 200 μM 2OG in 50 mM Tris-D_{11}, pH 7.5 in H_2O.

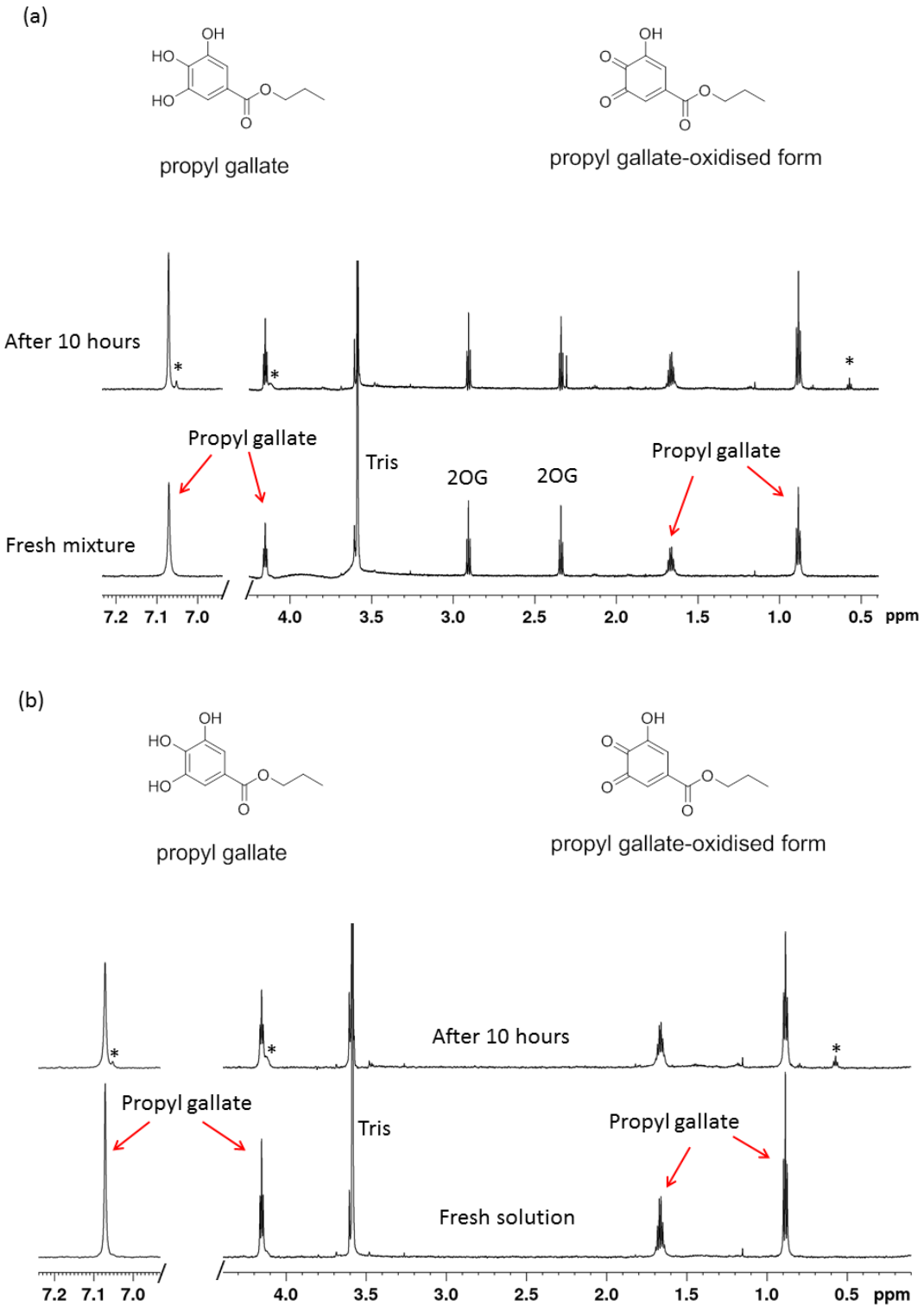


Figure 3.50 2OG decarboxylation by propyl gallate auto-oxidation.
(a) ^{1}H NMR spectra (part shown) showing a fresh mixture of propyl gallate/2OG/Tris-D_{11} (bottom) and mixture after 10 hours incubation (top). (b) Control experiment to show that propyl gallate undergoes auto-oxidation in the presence of oxygen. Signals marked with the asterisk are proposed to denote the oxidised form of propyl gallate. The concentration used are 500 µM propyl gallate, 200 µM 2OG in 50 mM Tris-D_{11}, pH 7.5 in H_2O.

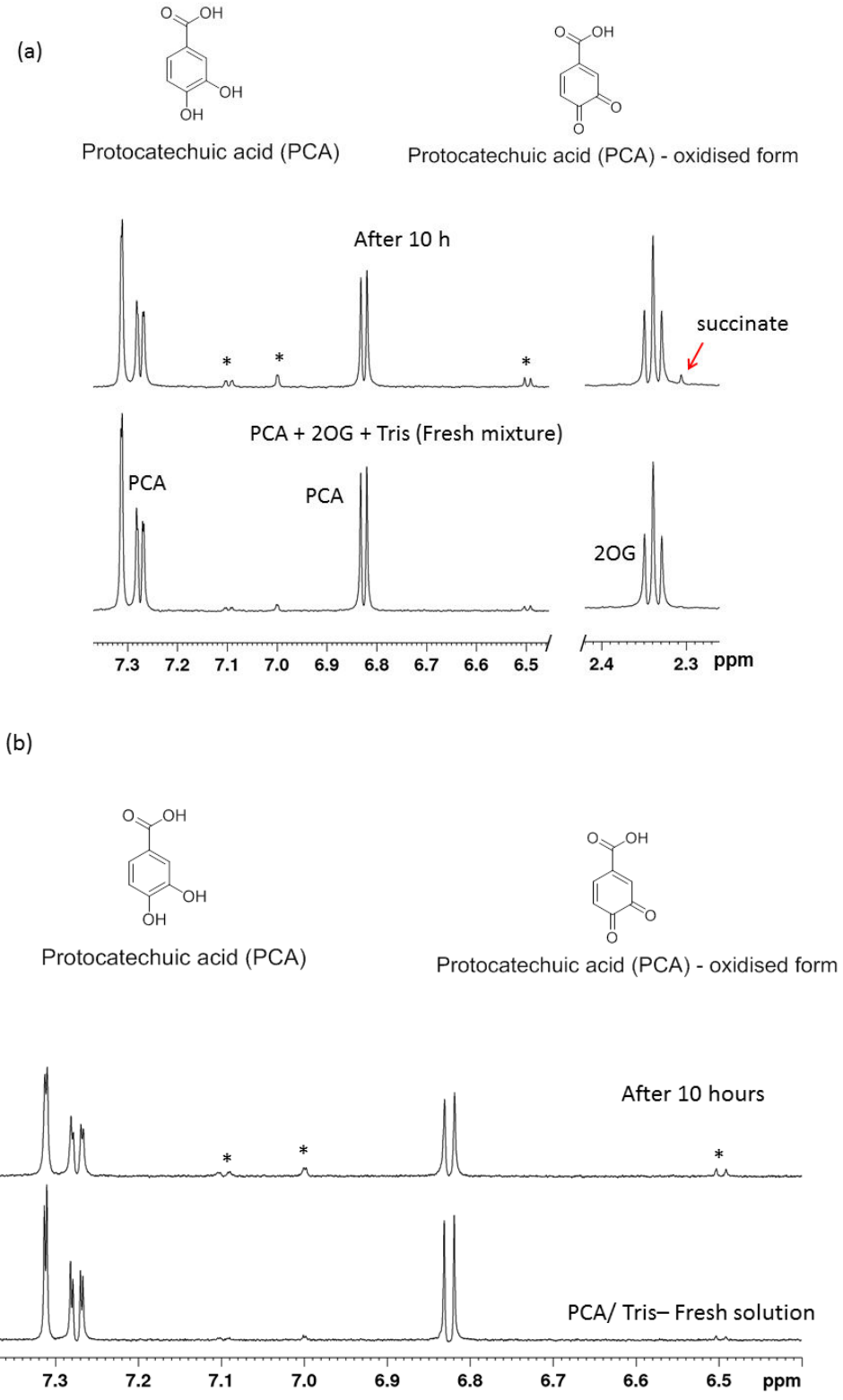


Figure 3.51 2OG decarboxylation by protocatechuic acid (PCA) auto-oxidation.
(a) ^{1}H NMR spectra (part shown) showing a fresh mixture of protocatechuic acid /2OG/Tris-D_{11} (bottom) and mixture after 10 hours incubation (top). (b) Control experiment to show that protocatechuic acid undergoes auto-oxidation in the presence of oxygen. The asterisks possibly represent the signals of the oxidised form of PCA. The concentration used are 500 µM protocatechuic acid, 200 µM 2OG in 50 mM Tris-D11, pH 7.5 in H2O.

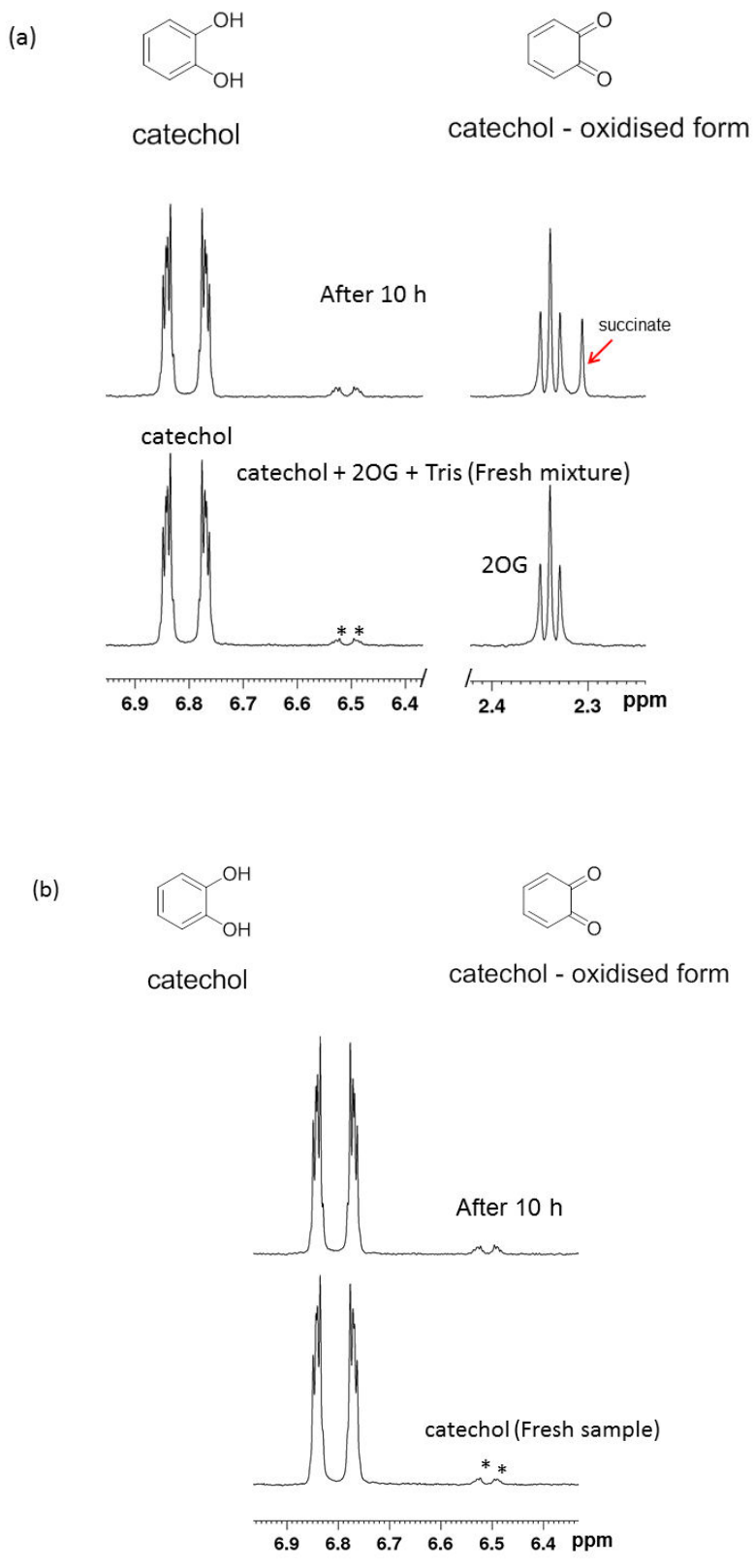


Figure 3.52 2OG decarboxylation by catechol auto-oxidation.
(a) ^{1}H NMR spectra (part shown) showing a fresh mixture of catechol/2OG/Tris-D11 (bottom) and mixture after 10 hours incubation (top). (b) Control experiment to show that catechol undergoes auto-oxidation in the presence of oxygen. The asterisks possibly represent the oxidised form of catechol. The concentration used are 500 µM catechol, 200 µM 2OG in 50 mM Tris-D11, pH 7.5 in H2O.

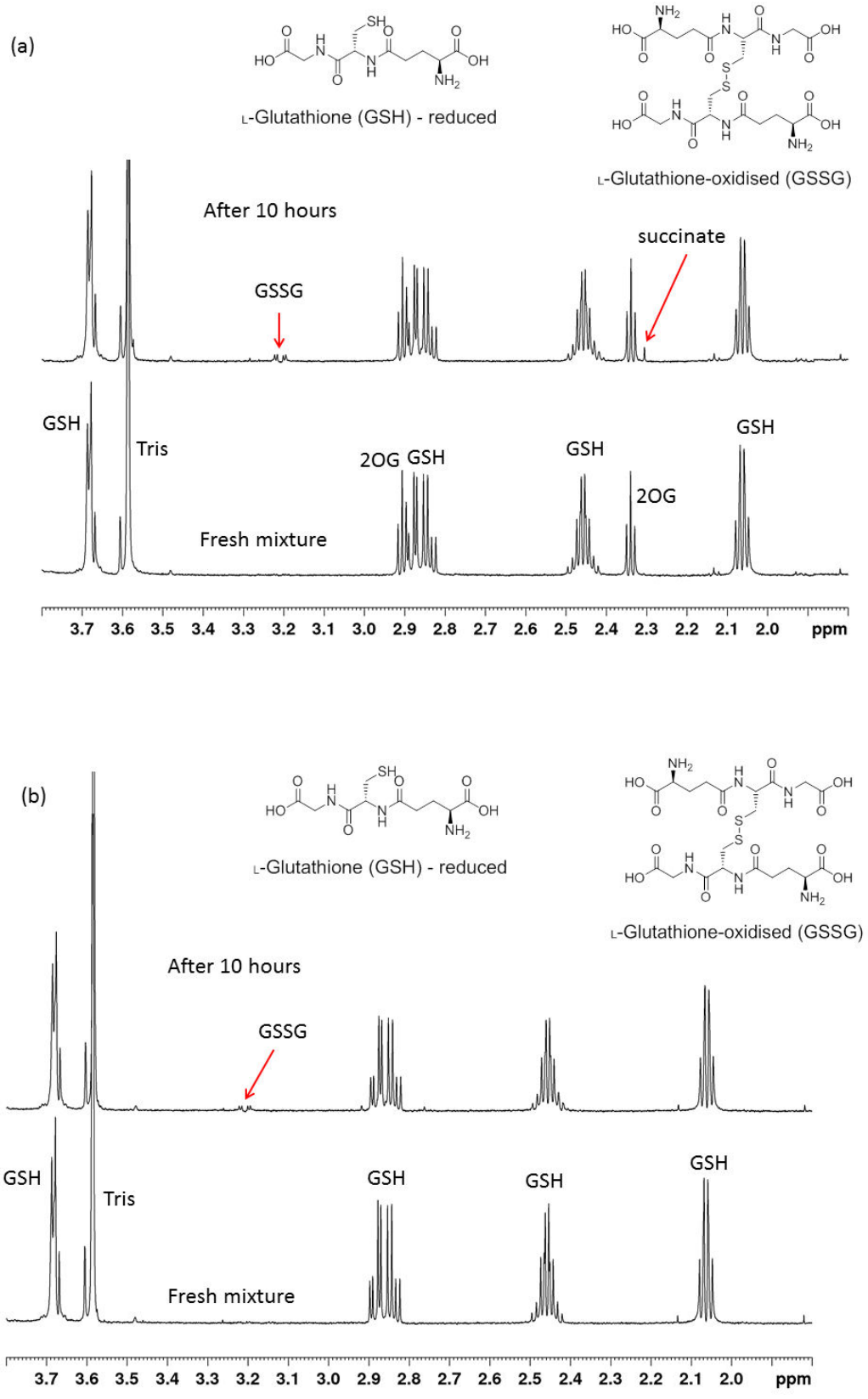


Figure 3.53 2OG decarboxylation by glutathione auto-oxidation.
(a) ^{1}H NMR spectra (part shown) showing a fresh mixture of glutathione/2OG/Tris-D11 (bottom) and mixture after 10 hours incubation (top). (b) Control experiment to show that glutathione undergoes auto-oxidation in the presence of oxygen. The concentration used are 500 µM glutathione, 200 µM 2OG in 50 mM Tris-D11, pH 7.5 in H2O.

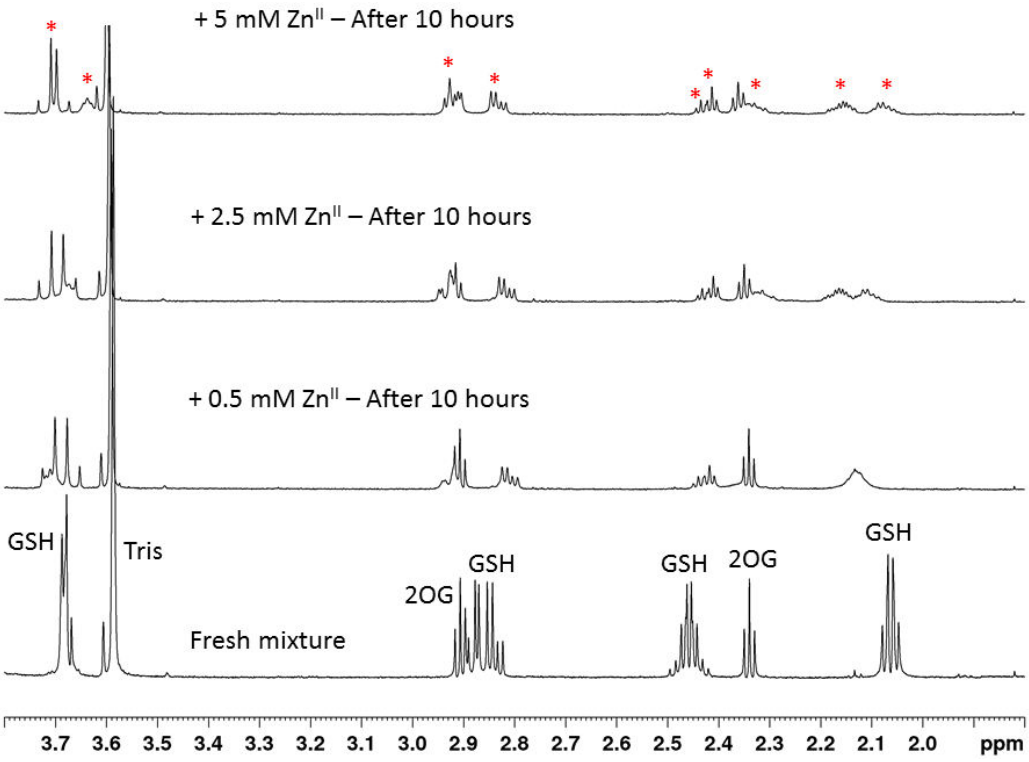
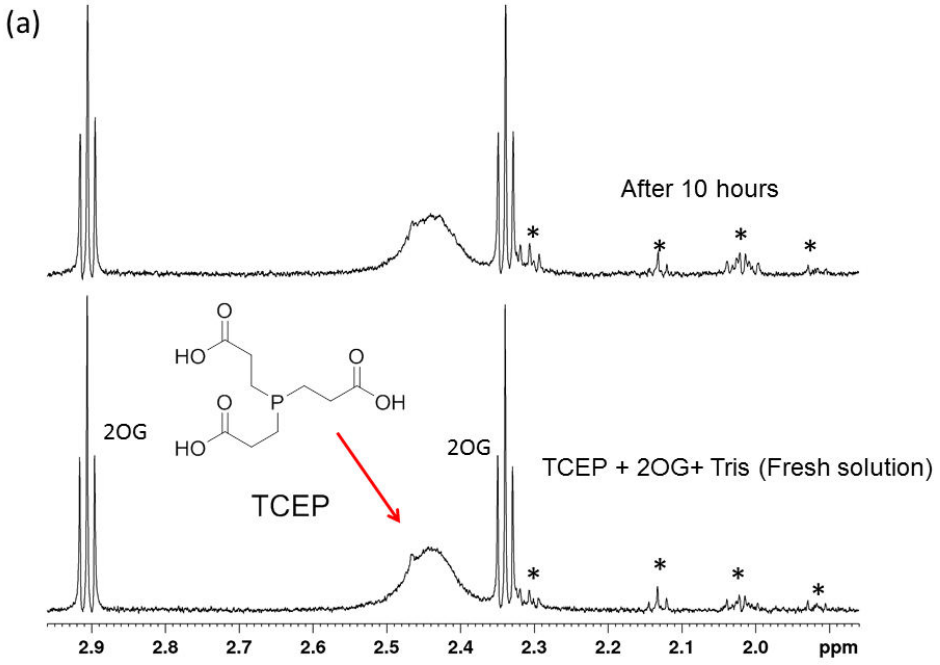


Figure 3.54 Monitoring if Zn^{II} can speed up the auto-oxidation of glutathione.
The glutathione concentration used was 500 µM. The red asterisks possibly represent the signals of the Zn^{II} complexation product with GSH.



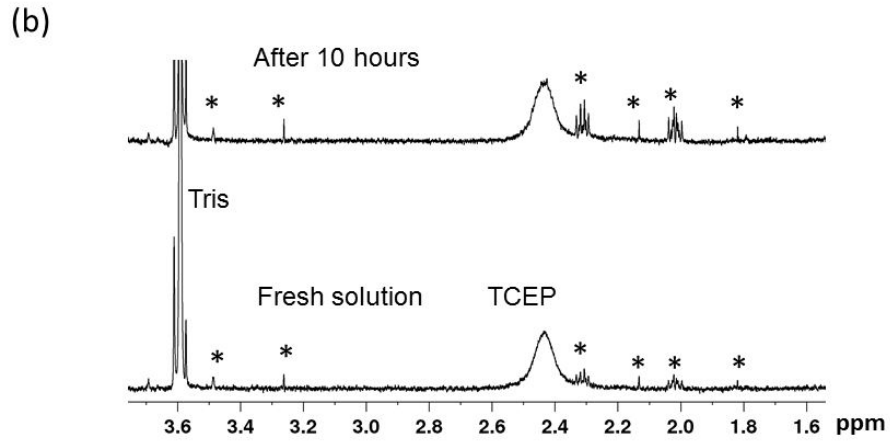


Figure 3.55 Investigating 2OG decarboxylation by TCEP auto-oxidation.
(b) ^{1}H NMR spectra (part shown) showing a fresh mixture of TCEP/2OG/Tris-D11 (bottom) and mixture after 10 hours incubation (top). (b) Control experiment to show that TCEP does not undergoes auto-oxidation in the presence of oxygen. The asterisks are the signals from impurities in the buffer solution. The concentration used are 500 μM TCEP, 200 μM 2OG in 50 mM Tris-D11, pH 7.5 in H2O.

3.8 Summary

The work described in this chapter demonstrates that in the assay solution of 2OG oxygenases, in the absence of both enzyme and substrate, ascorbate is slowly degraded by the dissolved oxygen in the mixture and produces H_2O_2. We speculate that in this process ascorbate donates a single electron to the dissolved molecular oxygen in the mixture converting it into highly reactive superoxide anion, which then yields H_2O_2. Although, we have not directly detected the formation of H_2O_2 in the mixture, however, the results obtained strongly suggest that H_2O_2 is formed from ascorbate oxidation (degradation). This H_2O_2 then reacts with the 2OG oxygenase co-substrate 2OG and results in its decarboxylation forming succinate. This ascorbate-induced 2OG turnover is thus proposed as a potential source of artifacts in the level of uncoupled 2OG turnover in the turnover assays of 2OG oxygenases both in the absence and presence of enzyme substrate.

The results also suggest that, as reducing agents, compounds containing 1,2 ene-diol functionality such as ascorbate, baicalein etc. behaves similarly as compounds without 1,2 ene-diol functionality (DTT) i.e both types of compounds donate electron to O_2 to form superoxide to yield H_2O_2. They were both proved to result in 2OG decarboxylation into succinate under our assay conditions. Hence, if the activation of the 2OG oxygenase activity by ascorbate is due to maintaining a pool of Fe^{II} to the enzyme by converting Fe^{III} to Fe^{II} in solution (as electron donor or reductant), then, in that case other reducing agent agents such

as DTT should also activate the enzyme to a similar level as ascorbate, but in most 2OG oxygenases this is not the case. For example, in case of PHD2, maximum enzymatic activity is obtained only with ascorbate and not with DTT, whereas, in case of AlkB, DTT can replace 80% of the enzyme activity with ascorbate. This suggests that these reducing agents perhaps interact with the 2OG oxygenases in a specific manner; not through a common mechanism (by reducing the free Fe^{III} to Fe^{II} in solution). Our results thus support the theory that these compounds may be activating the 2OG oxygenase enzymes by reducing the enzyme bound Fe^{III} to Fe^{II} as also supported by the structural-activity relation studies of various ascorbate analogues in both 2OG oxygenases (PHD2) and non-2OG oxygenase enzymes (dopamine-β-monooxygenase or DβM).

3.9 References

1. Flashman, E.; Davies, S. L.; Yeoh, K. K.; Schofield, C. J. Investigating the dependence of the hypoxia-inducible factor hydroxylases (factor inhibiting HIF and prolyl hydroxylase domain 2) on ascorbate and other reducing agents. *Biochem. J.* **2010**, 427, 135-42.
2. Rebouche, C. J. Ascorbic acid and carnitine biosynthesis. *Am. J. Clin. Nutr.* **1991**, 54, 1147-1152.
3. Myllylä, R.; Kuutti-Savolainen, E.-R.; Kivirikko, K. I. The role of ascorbate in the prolyl hydroxylase reaction. *Biochem. Biophys. Res. Commun.* **1978**, 83, 441-448.
4. Kivirikko, K. I.; Laitinen, O.; Prockop, D. J. Modifications of a specific assay for hydroxyproline in urine. *Anal. Biochem.* **1967**, 19, 249-55.
5. Saari, R. E.; Hausinger, R. P. Ascorbic acid-dependent turnover and reactivation of 2, 4-dichlorophenoxyacetic acid/α-ketoglutarate dioxygenase using thiophenoxyacetic acid. *Biochemistry* **1998**, 37, 3035-3042.
6. Inaba, K.; Ito, K. Paradoxical redox properties of DsbB and DsbA in the protein disulfide-introducing reaction cascade. *EMBO J.* **2002**, 21, 2646-54.
7. Barycki, J. J.; Asard, H.; Stone, J. M.; Wilson, M. A.; Banerjee, R.; Becker, D. F. Antioxidant Molecules and Redox Cofactors. In *Redox Biochemistry*, John Wiley & Sons, Inc.: **2007**; pp 11-47.
8. Welford, R. W.; Schlemminger, I.; McNeill, L. A.; Hewitson, K. S.; Schofield, C. J. The selectivity and inhibition of AlkB. *J. Biol. Chem.* **2003**, 278, 10157-61.
9. Davey, M. W.; Montagu, M. V.; Inzé, D.; Sanmartin, M.; Kanellis, A.; Smirnoff, N.; Benzie, I. J. J.; Strain, J. J.; Favell, D.; Fletcher, J. Plant L-ascorbic acid: chemistry, function, metabolism, bioavailability and effects of processing. *J. Sci. Food Agric.* **2000**, 80, 825-860.
10. Kivirikko, K. I.; Myllylä, R. Posttranslational enzymes in the biosynthesis of collagen: intracellular enzymes. *Methods Enzymol.* **1981**, 82, 245-304.
11. Kivirikko, K. I.; Myllyla, R. Posttranslational enzymes in the biosynthesis of collagen: intracellular enzymes. *Methods Enzymol.* **1982**, 82 Pt A, 245-304.
12. Kivirikko, K.; Myllylä, R.; Freedman, R.; Hawkins, H. *The Enzymology of Post-translational Modification of Proteins*. Academic Pres: New York, **1980**.
13. Prockop, D.; Berg, R.; Kivirikko, K.; Uitto, J. *Biochemistry of collagen.* Ramachandran, G. N & Reddi, A. H: Plenum, New York, USA, **1976**; p 163-273.

14. Cardinale, G. J.; Udenfriend, S. Prolyl hydroxylase. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1974**, 41, 245-300.
15. Kivirikko, K. I.; Risteli, L. Biosynthesis of collagen and its alterations in pathological states. *Med. Biol.* **1976**, 54, 159-86.
16. Barnes, M. J. Function of ascorbic acid in collagen metabolism. *Ann. N. Y. Acad. Sci.* **1975**, 258, 264-77.
17. Silver, I.; Kulonen, E.; Pikkarainen, J. Biology of Fibroblast. In Academic Press, London and New York: **1973**.
18. Ramaley, P.; Jimenez, S.; Rosenbloom, J. Conformation of underhydroxylated collagen synthesized by 3T6 fibroblasts in culture. *FEBS Lett.* **1973**, 33, 187-191.
19. Peterkofsky, B. The effect of ascorbic acid on collagen polypeptide synthesis and proline hydroxylation during the growth of cultured fibroblasts. *Arch. Biochem. Biophys.* **1972**, 152, 318-328.
20. Peterkofsky, B. Regulation of collagen secretion by ascorbic acid in 3T3 and chick embryo fibroblasts. *Biochem. Biophys. Res. Commun.* **1972**, 49, 1343-1350.
21. Bates, C.; Prynne, C.; Levene, C. The synthesis of underhydroxylated collagen by 3T6 mouse fibroblasts in culture. *Biochimica et Biophysica Acta (BBA)-Protein Structure* **1972**, 263, 397-405.
22. Levene, C. I.; Aleo, J. J.; Prynne, C. J.; Bates, C. J. The activation of protocollagen proline hydroxylase by ascorbic acid in cultured 3T6 fibroblasts. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1974**, 338, 29-36.
23. Rhoads, R. E.; Udenfriend, S. Purification and properties of collagen proline hydroxylase from newborn rat skin. *Arch. Biochem. Biophys.* **1970**, 139, 329-39.
24. Kutnink, M. A.; Tolbert, B. M.; Richmond, V. L.; Baker, E. M. Efficacy of the ascorbic acid stereoisomers in proline hydroxylation in vitro. *Proc. Soc. Exp. Biol. Med.* **1969**, 132, 440-2.
25. Hutton Jr, J. J.; Tappel, A. L.; Udenfriend, S. Cofactor and substrate requirements of collagen proline hydroxylase. *Arch. Biochem. Biophys.* **1967**, 118, 231-240.
26. Nietfeld, J. J.; Kemp, A. The function of ascorbate with respect to prolyl 4-hydroxylase activity. *Biochim. Biophys. Acta* **1981**, 657, 159-67.
27. Puistola, U.; Turpeenniemi-Hujanen, T. M.; Myllyla, R.; Kivirikko, K. I. Studies on the lysyl hydroxylase reaction. I. Initial velocity kinetics and related aspects. *Biochim. Biophys. Acta* **1980**, 611, 40-50.

28. Tuderman, L.; Myllyla, R.; Kivirikko, K. I. Mechanism of the prolyl hydroxylase reaction. 1. Role of co-substrates. *Eur. J. Biochem.* **1977**, 80, 341-8.
29. de Jong, L.; Albracht, S. P.; Kemp, A. Prolyl 4-hydroxylase activity in relation to the oxidation state of enzyme-bound iron. The role of ascorbate in peptidyl proline hydroxylation. *Biochim. Biophys. Acta* **1982**, 704, 326-32.
30. Myllyla, R.; Kuutti-Savolainen, E. R.; Kivirikko, K. I. The role of ascorbate in the prolyl hydroxylase reaction. *Biochem. Biophys. Res. Commun.* **1978**, 83, 441-8.
31. Counts, D. F.; Cardinale, G. J.; Udenfriend, S. Prolyl hydroxylase half reaction: peptidyl prolyl-independent decarboxylation of alpha-ketoglutarate. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, 75, 2145-9.
32. Rao, N. V.; Adams, E. Partial reaction of prolyl hydroxylase. (Gly-PRO-Ala)n stimulates alpha-ketoglutarate decarboxylation without prolyl hydroxylation. *J. Biol. Chem.* **1978**, 253, 6327-30.
33. Myllyla, R.; Majamaa, K.; Gunzler, V.; Hanauske-Abel, H. M.; Kivirikko, K. I. Ascorbate is consumed stoichiometrically in the uncoupled reactions catalyzed by prolyl 4-hydroxylase and lysyl hydroxylase. *J. Biol. Chem.* **1984**, 259, 5403-5.
34. De Jong, L.; Kemp, A. Stoicheiometry and kinetics of the prolyl 4-hydroxylase partial reaction. *Biochim. Biophys. Acta* **1984**, 787, 105-11.
35. Wang, G. L.; Jiang, B.-H.; Rue, E. A.; Semenza, G. L. Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O_2 tension. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, 92, 5510-5514.
36. Semenza, G. L.; Wang, G. L. A nuclear factor induced by hypoxia via de novo protein synthesis binds to the human erythropoietin gene enhancer at a site required for transcriptional activation. *Mol. Cell. Biol.* **1992**, 12, 5447-5454.
37. Hewitson, K. S.; McNeill, L. A.; Riordan, M. V.; Tian, Y.-M.; Bullock, A. N.; Welford, R. W.; Elkins, J. M.; Oldham, N. J.; Bhattacharya, S.; Gleadle, J. M. Hypoxia-inducible factor (HIF) asparagine hydroxylase is identical to factor inhibiting HIF (FIH) and is related to the cupin structural family. *J. Biol. Chem.* **2002**, 277, 26351-26355.
38. Taylor, M. S. Characterization and comparative analysis of the EGLN gene family. *Gene* **2001**, 275, 125-132.
39. Epstein, A. C.; Gleadle, J. M.; McNeill, L. A.; Hewitson, K. S.; O'Rourke, J.; Mole, D. R.; Mukherji, M.; Metzen, E.; Wilson, M. I.; Dhanda, A. C. elegans EGL-9 and

mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation. *Cell* **2001**, 107, 43-54.

40. Bruick, R. K.; McKnight, S. L. A conserved family of prolyl-4-hydroxylases that modify HIF. *Science* **2001**, 294, 1337-1340.
41. Jaakkola, P.; Mole, D. R.; Tian, Y.-M.; Wilson, M. I.; Gielbert, J.; Gaskell, S. J.; von Kriegsheim, A.; Hebestreit, H. F.; Mukherji, M.; Schofield, C. J. Targeting of HIF-α to the von Hippel-Lindau ubiquitylation complex by O_2-regulated prolyl hydroxylation. *Science* **2001**, 292, 468-472.
42. Ivan, M.; Kondo, K.; Yang, H.; Kim, W.; Valiando, J.; Ohh, M.; Salic, A.; Asara, J. M.; Lane, W. S.; Kaelin Jr, W. G. HIFα targeted for VHL-mediated destruction by proline hydroxylation: implications for O_2 sensing. *Science* **2001**, 292, 464-468.
43. McNeill, L. A.; Flashman, E.; Buck, M. R.; Hewitson, K. S.; Clifton, I. J.; Jeschke, G.; Claridge, T. D.; Ehrismann, D.; Oldham, N. J.; Schofield, C. J. Hypoxia-inducible factor prolyl hydroxylase 2 has a high affinity for ferrous iron and 2-oxoglutarate. *Molecular BioSystems* **2005**, 1, 321-324.
44. Hirsilä, M.; Koivunen, P.; Günzler, V.; Kivirikko, K. I.; Myllyharju, J. Characterization of the human prolyl 4-hydroxylases that modify the hypoxia-inducible factor. *J. Biol. Chem.* **2003**, 278, 30772-30780.
45. Knowles, H. J.; Raval, R. R.; Harris, A. L.; Ratcliffe, P. J. Effect of Ascorbate on the Activity of Hypoxia-inducible Factor in Cancer Cells. *Cancer Res.* **2003**, 63, 1764-1768.
46. Lando, D.; Peet, D. J.; Whelan, D. A.; Gorman, J. J.; Whitelaw, M. L. Asparagine hydroxylation of the HIF transactivation domain: a hypoxic switch. *Science* **2002**, 295, 858-861.
47. Welford, R. Studies on the DNA repair enzyme AlkB and the ascorbate dependence of the, 2-oxoglutarate oxygenases. PhD Thesis, University of Oxford, UK **2004**.
48. Stein, R.; Englard, S. Properties of rat 6-N-trimethyl-L-lysine hydroxylases: similarities among the kidney, liver, heart, and skeletal muscle activities. *Arch. Biochem. Biophys.* **1982**, 217, 324-31.
49. Stein, R.; Englard, S. The use of a tritium release assay to measure 6-N-trimethyl-L-lysine hydroxylase activity: synthesis of 6-N-[3-3H]trimethyl-DL-lysine. *Anal. Biochem.* **1981**, 116, 230-6.
50. Sachan, D. S.; Hoppel, C. L. Carnitine biosynthesis. Hydroxylation of N^6-trimethyl-lysine to 3-hydroxy-N^6-trimethyl-lysine. *Biochem. J.* **1980**, 188, 529-34.

51. Hulse, J. D.; Ellis, S. R.; Henderson, L. M. Carnitine biosynthesis. beta-Hydroxylation of trimethyllysine by an alpha-ketoglutarate-dependent mitochondrial dioxygenase. *J. Biol. Chem.* **1978**, 253, 1654-9.

52. Englard, S.; Horwitz, L. J.; Mills, J. T. A simplified method for the measurement of gamma-butyrobetaine hydroxylase activity. *J. Lipid Res.* **1978**, 19, 1057-63.

53. Englard, S.; Carnicero, H. H. gamma-Butyrobetaine hydroxylation to carnitine in mammalian kidney. *Arch. Biochem. Biophys.* **1978**, 190, 361-4.

54. Lindstedt, G.; Lindstedt, S. Cofactor requirements of gamma-butyrobetaine hydroxylase from rat liver. *J. Biol. Chem.* **1970**, 245, 4178-86.

55. Holme, E.; Lindstedt, S.; Nordin, I. Uncoupling in the γ-butyrobetaine hydroxylase reaction by D and L carnitine. *Biochem. Biophys. Res. Commun.* **1982**, 107, 518-524.

56. Nelson, P. J.; Pruitt, R. E.; Henderson, L. L.; Jenness, R.; Henderson, L. M. Effect of ascorbic acid deficiency on the in vivo synthesis of carnitine. *Biochimica et Biophysica Acta (BBA)-General Subjects* **1981**, 672, 123-127.

57. Tahiliani, M.; Koh, K. P.; Shen, Y.; Pastor, W. A.; Bandukwala, H.; Brudno, Y.; Agarwal, S.; Iyer, L. M.; Liu, D. R.; Aravind, L. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **2009**, 324, 930-935.

58. Wu, H.; Zhang, Y. Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation. *Genes Dev.* **2011**, 25, 2436-2452.

59. Ito, S.; Shen, L.; Dai, Q.; Wu, S. C.; Collins, L. B.; Swenberg, J. A.; He, C.; Zhang, Y. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* **2011**, 333, 1300-1303.

60. He, Y.-F.; Li, B.-Z.; Li, Z.; Liu, P.; Wang, Y.; Tang, Q.; Ding, J.; Jia, Y.; Chen, Z.; Li, L. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* **2011**, 333, 1303-1307.

61. Yin, R.; Mao, S.-Q.; Zhao, B.; Chong, Z.; Yang, Y.; Zhao, C.; Zhang, D.; Huang, H.; Gao, J.; Li, Z. Ascorbic acid enhances Tet-mediated 5-methylcytosine oxidation and promotes DNA demethylation in mammals. *J. Am. Chem. Soc.* **2013**, 135, 10396-10403.

62. Minor, E. A.; Court, B. L.; Young, J. I.; Wang, G. Ascorbate Induces Ten-Eleven Translocation (Tet) Methylcytosine Dioxygenase-mediated Generation of 5-Hydroxymethylcytosine. *J. Biol. Chem.* **2013**, 288, 13669-13674.

63. Dickson, K. M.; Gustafson, C. B.; Young, J. I.; Zuchner, S.; Wang, G. Ascorbate-induced generation of 5-hydroxymethylcytosine is unaffected by varying levels of iron and 2-oxoglutarate. *Biochem. Biophys. Res. Commun.* **2013**, 439, 522-7.

64. Blaschke, K.; Ebata, K. T.; Karimi, M. M.; Zepeda-Martinez, J. A.; Goyal, P.; Mahapatra, S.; Tam, A.; Laird, D. J.; Hirst, M.; Rao, A.; Lorincz, M. C.; Ramalho-Santos, M. Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. *Nature* **2013**, 500, 222-6.

65. Lian, C. G.; Xu, Y.; Ceol, C.; Wu, F.; Larson, A.; Dresser, K.; Xu, W.; Tan, L.; Hu, Y.; Zhan, Q. Loss of 5-hydroxymethylcytosine is an epigenetic hallmark of melanoma. *Cell* **2012**, 150, 1135-1146.

66. Ko, M.; Huang, Y.; Jankowska, A. M.; Pape, U. J.; Tahiliani, M.; Bandukwala, H. S.; An, J.; Lamperti, E. D.; Koh, K. P.; Ganetzky, R. Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature* **2010**, 468, 839-843.

67. Turnbull, J. J.; Sobey, W. J.; Aplin, R. T.; Hassan, A.; Firmin, J. L.; Schofield, C. J.; Prescott, A. G. Are anthocyanidins the immediate products of anthocyanidin synthase? *Chem. Commun.* **2000**, 2473-2474.

68. Saito, K.; Kobayashi, M.; Gong, Z.; Tanaka, Y.; Yamazaki, M. Direct evidence for anthocyanidin synthase as a 2-oxoglutarate-dependent oxygenase: molecular cloning and functional expression of cDNA from a red forma ofPerilla frutescens. *The Plant Journal* **1999**, 17, 181-189.

69. Harborne, J. B.; Williams, C. A. Advances in flavonoid research since 1992. *Phytochemistry* **2000**, 55, 481-504.

70. Bohm, B. A. *Introduction to flavonoids*. Harwood Academic Publishers: Amsterdam, Netherlands, **1998**.

71. Wilmouth, R. C.; Turnbull, J. J.; Welford, R. W.; Clifton, I. J.; Prescott, A. G.; Schofield, C. J. Structure and mechanism of anthocyanidin synthase from Arabidopsis thaliana. *Structure* **2002**, 10, 93-103.

72. Burmeister, W. P.; Cottaz, S.; Rollin, P.; Vasella, A.; Henrissat, B. High resolution X-ray crystallography shows that ascorbate is a cofactor for myrosinase and substitutes for the function of the catalytic base. *J. Biol. Chem.* **2000**, 275, 39385-93.

73. Lindahl, T.; Sedgwick, B.; Sekiguchi, M.; Nakabeppu, Y. Regulation and expression of the adaptive response to alkylating agents. *Annu. Rev. Biochem.* **1988**, 57, 133-157.

74. Wei, Y. F.; Carter, K. C.; Wang, R. P.; Shell, B. K. Molecular cloning and functional analysis of a human cDNA encoding an Escherichia coli AlkB homolog, a protein involved in DNA alkylation damage repair. *Nucleic Acids Res.* **1996**, 24, 931-37.

75. Friedberg, E. C.; Walker, G. C.; Siede, W.; Wood, R. D. *DNA repair and mutagenesis*. American Society for Microbiology Press: Washington, D.C, **2005**.

76. Trewick, S. C.; Henshaw, T. F.; Hausinger, R. P.; Lindahl, T.; Sedgwick, B. Oxidative demethylation by *Escherichia coli* AlkB directly reverts DNA base damage. *Nature* **2002**, 419, 174-178.

77. Zhang, Z.; Ren, J.; Harlos, K.; McKinnon, C. H.; Clifton, I. J.; Schofield, C. J. Crystal structure of a clavaminate synthase-Fe(II)-2-oxoglutarate-substrate-NO complex: evidence for metal centered rearrangements. *FEBS Lett.* **2002**, 517, 7-12.

78. Solomon, E. I.; Brunold, T. C.; Davis, M. I.; Kemsley, J. N.; Lee, S. K.; Lehnert, N.; Neese, F.; Skulan, A. J.; Yang, Y. S.; Zhou, J. Geometric and electronic structure/function correlations in non-heme iron enzymes. *Chem. Rev.* **2000**, 100, 235-350.

79. Holme, E. A kinetic study of thymine 7-hydroxylase from neurospora crassa. *Biochemistry* **1975**, 14, 4999-5003.

80. Yabuuchi, E.; Kosako, Y.; Yano, I.; Hotta, H.; Nishiuchi, Y. Transfer of two Burkholderia and an Alcaligenes species to Ralstonia gen. Nov.: Proposal of Ralstonia pickettii (Ralston, Palleroni and Doudoroff 1973) comb. Nov., Ralstonia solanacearum (Smith 1896) comb. Nov. and Ralstonia eutropha (Davis 1969) comb. Nov. *Microbiol. Immunol.* **1995**, 39, 897-904.

81. Fukumori, F.; Hausinger, R. P. Alcaligenes eutrophus JMP134" 2, 4-dichlorophenoxyacetate monooxygenase" is an alpha-ketoglutarate-dependent dioxygenase. *J. Bacteriol.* **1993**, 175, 2083-2086.

82. Fukumori, F.; Hausinger, R. Purification and characterization of 2, 4-dichlorophenoxyacetate/alpha-ketoglutarate dioxygenase. *J. Biol. Chem.* **1993**, 268, 24311-24317.

83. Streber, W. R.; Timmis, K. N.; Zenk, M. H. Analysis, cloning, and high-level expression of 2, 4-dichlorophenoxyacetate monooxygenase gene tfdA of Alcaligenes eutrophus JMP134. *J. Bacteriol.* **1987**, 169, 2950-2955.

84. Klose, R. J.; Zhang, Y. Regulation of histone methylation by demethylimination and demethylation. *Nat. Rev. Mol. Cell Biol.* **2007**, 8, 307-18.

85. Agger, K.; Christensen, J.; Cloos, P. A.; Helin, K. The emerging functions of histone demethylases. *Curr. Opin. Genet. Dev.* **2008**, 18, 159-68.

86. Hopkinson, R. J.; Hamed, R. B.; Rose, N. R.; Claridge, T. D.; Schofield, C. J. Monitoring the activity of 2-oxoglutarate dependent histone demethylases by NMR spectroscopy: direct observation of formaldehyde. *Chembiochem* **2010**, 11, 506-10.

87. Kaufman, S.; Friedman, S. Dopamine-beta-Hydroxylase. *Pharmacol. Rev.* **1965**, 17, 71-100.

88. Friedman, S.; Kaufman, S. 3,4-dihydroxyphenylethylamine beta-hydroxylase. Physical properties, copper content, and role of copper in the catalytic acttivity. *J. Biol. Chem.* **1965**, 240, 4763-73.

89. Levin, E. Y.; Kaufman, S. Studies on the enzyme catalyzing the conversion of 3, 4-dihydroxyphenylethylamine to norepinephrine. *J. Biol. Chem.* **1961**, 236, 2043-2049.

90. Levin, E. Y.; Levenberg, B.; Kaufman, S. The enzymatic conversion of 3, 4-dihydroxyphenylethylamine to norepinephrine. *J. Biol. Chem.* **1960**, 235, 2080-2086.

91. Wimalasena, K.; Dharmasena, S.; Wimalasena, D. S. Ascorbate based novel high affinity alternate reductants and competitive inhibitors of dopamine beta-monooxygenase. *Biochem. Biophys. Res. Commun.* **1994**, 200, 113-9.

92. Lin, Z.; Zhong, S.; Grierson, D. Recent advances in ethylene research. *J. Expt. Bot* **2009**, 60, 3311-3336.

93. Macháčková, I. Pech, JC, Latché, A., Balagué, C.(ed.): Cellular and molecular aspects of the plant hormone ethylene.(Current plant science and biotechnology in agriculture, Vol. 16). *Biol. Plant.* **1994**, 36, 14-14.

94. Adams, D. O.; Yang, S. F. Ethylene biosynthesis: Identification of 1-aminocyclopropane-1-carboxylic acid as an intermediate in the conversion of methionine to ethylene. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, 76, 170-4.

95. Rocklin, A. M.; Kato, K.; Liu, H. W.; Que, L., Jr.; Lipscomb, J. D. Mechanistic studies of 1-aminocyclopropane-1-carboxylic acid oxidase: single turnover reaction. *J. Biol. Inorg. Chem.* **2004**, 9, 171-82.

96. Zhang, Z.; Schofield, C. J.; Baldwin, J. E.; Thomas, P.; John, P. Expression, purification and characterization of 1-aminocyclopropane-1-carboxylate oxidase from tomato in Escherichia coli. *Biochem. J.* **1995**, 307 (Pt 1), 77-85.

97. Holloway, D. E.; Rivers, J. M. Influence of chronic ascorbic acid deficiency and excessive ascorbic acid intake on bile acid metabolism and bile composition in the guinea pig. *The Journal of nutrition* **1981**, 111, 412-424.

98. Ginter, E. Ascorbic acid in cholesterol and bile acid metabolism. *Ann. N. Y. Acad. Sci.* **1975**, 258, 410-421.

99. Davies, M. B.; Austin, J.; Partridge, D. A. *Vitamin C: its chemistry and biochemistry*. Royal Society of Chemistry: London, **1991**.

100. Fisher, D. B.; Kirkwood, R.; Kaufman, S. Rat liver phenylalanine hydroxylase, an iron enzyme. *J. Biol. Chem.* **1972**, 247, 5161-5167.

101. La Du, B. N., Jr.; Greenberg, D. M. Ascorbic acid and the oxidation of tyrosine. *Science* **1953**, 117, 111-2.

102. Sealock, R. R.; Goodland, R. L. Ascorbic acid, a coenzyme in tyrosine oxidation. *Science* **1951**, 114, 645-6.

103. Knox, W. E.; Le, M.-K. M. The oxidation in liver of l-tyrosine to acetoacetate through p-hydroxyphenylpyruvate and homogentisic acid. *Biochem. J.* **1951**, 49, 686-93.

104. Kojo, S. Vitamin C: basic metabolism and its function as an index of oxidative stress. *Curr. Med. Chem.* **2004**, 11, 1041-64.

105. Betteridge, D. J. What is oxidative stress? *Metabolism* **2000**, 49, 3-8.

106. Yuan, J.-P.; Chen, F. Degradation of ascorbic acid in aqueous solution. *J. Agric. Food Chem.* **1998**, 46, 5078-5082.

107. Tsao, C.; Packer, L.; Fuchs, J. *An overview of ascorbic acid chemistry and biochemistry*. Marcel Dekker: New York, NY, USA, **1997**; p 25-58.

108. Solomon, O.; Svanberg, U.; Sahlström, A. Effect of oxygen and fluorescent light on the quality of orange juice during storage at 8°C. *Food Chem.* **1995**, 53, 363-368.

109. Kimoto, E.; Tanaka, H.; Ohmoto, T.; Choami, M. Analysis of the transformation products of dehydro-L-ascorbic acid by ion-pairing high-performance liquid chromatography. *Anal. Biochem.* **1993**, 214, 38-44.

110. Seib, P. A.; Tolbert, B. M. *In ascorbic acid: chemistry, metabolism and uses*. American Chemical Society: Washington, D.C., **1982**; p 101-123.

111. Deutsch, J. C. Ascorbic acid and dehydroascorbic acid interconversion without net oxidation or reduction. *Anal. Biochem.* **1997**, 247, 58-62.

112. Long, L. H.; Halliwell, B. Artefacts in cell culture: α-Ketoglutarate can scavenge hydrogen peroxide generated by ascorbate and epigallocatechin gallate in cell culture media. *Biochem. Biophys. Res. Commun.* **2011**, 406, 20-24.

113. Du, J.; Martin, S. M.; Levine, M.; Wagner, B. A.; Buettner, G. R.; Wang, S.-h.; Taghiyev, A. F.; Du, C.; Knudson, C. M.; Cullen, J. J. Mechanisms of ascorbate-induced cytotoxicity in pancreatic cancer. *Clin. Cancer Res.* **2010**, 16, 509-520.

114. Chen, Q.; Espey, M. G.; Sun, A. Y.; Lee, J. H.; Krishna, M. C.; Shacter, E.; Choyke, P. L.; Pooput, C.; Kirk, K. L.; Buettner, G. R.; Levine, M. Ascorbate in pharmacologic concentrations selectively generates ascorbate radical and hydrogen peroxide in extracellular fluid in vivo. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, 104, 8749-54.

115. Kramarenko, G. G.; Hummel, S. G.; Martin, S. M.; Buettner, G. R. Ascorbate reacts with singlet oxygen to produce hydrogen peroxide. *Photochem. Photobiol.* **2006**, 82, 1634-1637.

116. Chen, Q.; Espey, M. G.; Krishna, M. C.; Mitchell, J. B.; Corpe, C. P.; Buettner, G. R.; Shacter, E.; Levine, M. Pharmacologic ascorbic acid concentrations selectively kill cancer cells: action as a pro-drug to deliver hydrogen peroxide to tissues. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, 102, 13604-13609.

117. Clément, M.-V.; Ramalingam, J.; Long, L. H.; Halliwell, B. The in vitro cytotoxicity of ascorbate depends on the culture medium used to perform the assay and involves hydrogen peroxide. *Antioxidants and redox signaling* **2001**, 3, 157-163.

118. Sestili, P.; Brandi, G.; Brambilla, L.; Cattabeni, F.; Cantoni, O. Hydrogen peroxide mediates the killing of U937 tumor cells elicited by pharmacologically attainable concentrations of ascorbic acid: cell death prevention by extracellular catalase or catalase from cocultured erythrocytes or fibroblasts. *J. Pharmacol. Exp. Ther.* **1996**, 277, 1719-1725.

119. Calcutt, G. The formation of hydrogen peroxide during the autoxidation of ascorbic acid. *Cell. Mol. Life Sci.* **1951**, 7, 26-26.

120. Scarpa, M.; Stevanato, R.; Viglino, P.; Rigo, A. Superoxide ion as active intermediate in the autoxidation of ascorbate by molecular oxygen. Effect of superoxide dismutase. *J. Biol. Chem.* **1983**, 258, 6695-6697.

121. Frei, B.; Lawson, S. Vitamin C and cancer revisited. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, 105, 11037-11038.

122. Halliwell, B. Vitamin C: poison, prophylactic or panacea? *Trends Biochem. Sci.* **1999**, 24, 255-259.

123. Buettner, G. R.; Jurkiewicz, B. A. Catalytic metals, ascorbate and free radicals: combinations to avoid. *Radiat. Res.* **1996**, 145, 532-541.

124. Du, J.; Cullen, J. J.; Buettner, G. R. Ascorbic acid: chemistry, biology and the treatment of cancer. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer* **2012**, 1826, 443-457.

125. Chen, Q.; Espey, M. G.; Sun, A. Y.; Pooput, C.; Kirk, K. L.; Krishna, M. C.; Khosh, D. B.; Drisko, J.; Levine, M. Pharmacologic doses of ascorbate act as a prooxidant and decrease growth of aggressive tumor xenografts in mice. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, 105, 11105-11109.

126. Long, L. H.; Hoi, A.; Halliwell, B. Instability of, and generation of hydrogen peroxide by, phenolic compounds in cell culture media. *Arch. Biochem. Biophys.* **2010**, 501, 162-9.

127. Perron, N. R.; Brumaghim, J. L. A review of the antioxidant mechanisms of polyphenol compounds related to iron binding. *Cell Biochem. Biophys.* **2009**, 53, 75-100.

128. Nath, K. A.; Ngo, E. O.; Hebbel, R.; Croatt, A.; Zhou, B.; Nutter, L. alpha-Ketoacids scavenge H2O2 in vitro and in vivo and reduce menadione-induced DNA injury and cytotoxicity. *Am. J. Physiol. Cell Physiol.* **1995**, 268, C227-C236.

129. O'Donnell-Tormey, J.; Nathan, C.; Lanks, K.; DeBoer, C.; De La Harpe, J. Secretion of pyruvate. An antioxidant defense of mammalian cells. *J. Exp. Med* **1987**, 165, 500-514.

130. Andrae, U.; Singh, J.; Ziegler-Skylakakis, K. Pyruvate and related α-ketoacids protect mammalian cells in culture against hydrogen peroxide-induced cytotoxicity. *Toxicol. Lett.* **1985**, 28, 93-98.

131. Holleman, M. Note on the action of oxygenated water on alpha-keto acids and 1, 2, dicetones. *Recl Tran Chi Pays-bas Belq* **1904**, 23, 169-172.

132. Siegel, B.; Lanphear, J. Kinetics and mechanisms for the acid-catalyzed oxidative decarboxylation of benzoylformic acid. *J. Org. Chem.* **1979**, 44, 942-946.

133. Perera, A.; Parkes, H.; Herz, H.; Haycock, P.; Blake, D.; Grootveld, M. High resolution ^{1}H NMR investigations of the reactivities of a-keto acid anions with hydrogen peroxide. *Free Radic. Res.* **1997**, 26, 145-157.

134. Ramachandran, S.; Morris, S. M.; Devamanoharan, P.; Henein, M.; Varma, S. D. Radio-isotopic determination of hydrogen peroxide in aqueous humor and urine. *Exp. Eye Res.* **1991**, 53, 503-506.

135. Varma, S. D.; Devamanoharan, P. Excretion of hydrogen peroxide in human urine. *Free Radic. Res.* **1990**, 8, 73-78.

136. Nakamichi, N.; Kambe, Y.; Oikawa, H.; Ogura, M.; Takano, K.; Tamaki, K.; Inoue, M.; Hinoi, E.; Yoneda, Y. Protection by exogenous pyruvate through a mechanism

related to monocarboxylate transporters against cell death induced by hydrogen peroxide in cultured rat cortical neurons. *J. Neurochem.* **2005**, 93, 84-93.

137. Lee, Y.-J.; Kang, I.-J.; Bünger, R.; Kang, Y.-H. Enhanced survival effect of pyruvate correlates MAPK and NF-κB activation in hydrogen peroxide-treated human endothelial cells. *J. Appl. Physiol.* **2004**, 96, 793-801.
138. Mazzio, E. A.; Soliman, K. F. Cytoprotection of pyruvic acid and reduced β-nicotinamide adenine dinucleotide against hydrogen peroxide toxicity in neuroblastoma cells. *Neurochem. Res.* **2003**, 28, 733-741.
139. Simonian, N.; Coyle, J. Oxidative stress in neurodegenerative diseases. *Annu. Rev. Pharmacol. Toxicol.* **1996**, 36, 83-106.
140. Olanow, C. A radical hypothesis for neurodegeneration. *Trends Neurosci.* **1993**, 16, 439-444.
141. Traystman, R. J.; Kirsch, J. R.; Koehler, R. C. Oxygen radical mechanisms of brain injury following ischemia and reperfusion. *J. Appl. Physiol.* **1991**, 71, 1185-1195.
142. Siesjö, B.; Agardh, C.-D.; Bengtsson, F. Free radicals and brain damage. *Cerebrovasc. Brain Metab. Rev.* **1988**, 1, 165-211.
143. Desagher, S.; Glowinski, J.; Prémont, J. Pyruvate protects neurons against hydrogen peroxide-induced toxicity. *J. neurosci* **1997**, 17, 9060-9067.
144. Salahudeen, A. K.; Clark, E. C.; Nath, K. Hydrogen peroxide-induced renal injury. A protective role for pyruvate in vitro and in vivo. *J. Clin. Invest.* **1991**, 88, 1886.
145. Nappi, A. J.; Vass, E. Hydroxyl radical production by ascorbate and hydrogen peroxide. *Neurotox. Res.* **2000**, 2, 343-355.
146. Deutsch, J. C. Ascorbic acid oxidation by hydrogen peroxide. *Anal. Biochem.* **1998**, 255, 1-7.
147. Pham, A. N.; Xing, G.; Miller, C. J.; Waite, T. D. Fenton-like copper redox chemistry revisited: Hydrogen peroxide and superoxide mediation of copper-catalyzed oxidant production. *J. Catal* **2013**, 301, 54-64.
148. Perez-Benito, J. F. Reaction pathways in the decomposition of hydrogen peroxide catalyzed by copper (II). *J. Inorg. Biochem.* **2004**, 98, 430-438.

4 NMR studies on bacterial potassium ion efflux system (Kef)[†]

4.1 Bacterial potassium ion efflux system (Kef)

Cellular homeostasis is the key for the survival of all life forms. In many gram-negative pathogenic bacteria, cellular homeostasis is maintained by controlling the permeability of cytoplasmic membrane to ions and solutes, a function which is performed by transmembrane regulatory proteins (domains) i.e cytosolic domains (RCK; regulator of conductance of K^+ or KTN; K^+ transport and nucleotide binding domains). These proteins function as gateways or `loading docks′ to permit or deny the transport of specific substances across the cytoplasmic membrane, to get into the cell, or out of the cell. Potassium ions (K^+) are central to a bacterium life. In bacteria, cytoplasmic K^+ is required for a diverse range of processes such as the activation of cytoplasmic enzymes, the maintenance of cell turgor, the homeostasis of cytoplasmic pH etc. hence critical for the survival of bacterial life.[1,2] In most species of bacteria, the concentration of cytoplasmic K^+ ranges from 300-500 mM, that is kept relatively constant by a controlled active K^+-uptake (Trk/Ktr; superfamiliy of membrane proteins) and K^+-efflux systems (Kef proteins).[3-5] The regulation of this K^+-transport is thus a major determinant of the bacterial growth and survival.[1,2] In most species of bacteria two or more independent saturable K^+-transport systems are usually present.

[†] *The work described in this chapter was designed and conducted in collaboration with Dr. Samuel C. Grayer, Dr. Anthony K. N. Chan and Silvia Ekkerman.*

S.C.G synthesized the glutathione-S-conjugates and performed the florescence binding assay. A.K.N.C produced the sdKef-QCTD and performed the AMP HPLC analysis. S.E performed the structural and mutation studies and the K^+ efflux assay.

Parts of the work from this chapter have been drafted into the following manuscripts:

1. Amjad Khan, Samuel C. Grayer, Anthony K. N. Chan, Alexander Axer, Jessica Healy, Stuart J. Conway and Timothy D. W. Claridge., An efficient ^{19}F NMR reporter displacement assay for ligand binding to the bacterial K^+ Efflux (Kef) system. (Manuscript prepared and is currently being finalised for publication).

2. Sylvia Ekkerman, Wendy Bartlett, Tim Rasmussen, Samantha Miller, Ian R. Booth, Christos Pliotas, Phedra, Marius, and James H Naismith, Samuel C. Grayer, Anthony K. N. Chan, Jessica Healy, Amjad Khan, Wilian A. Cortopassi, Robert S Paton, Tim D. W. Claridge and Stuart J. Conway; The role of high affinity binding of AMP to the KTN domain from Shewanella denitrificans Kef channel. (Manuscript currently under preparation).

The K^+ efflux system (Figure 4.1) is a ligand-gated ion channel system involved the KTN domains/proteins. The most extensively studied KTN-regulated system is KefC from *Escherichia coli*, a transporter with homologues KefB and KefC in most of the Gram negative pathogens.[6]

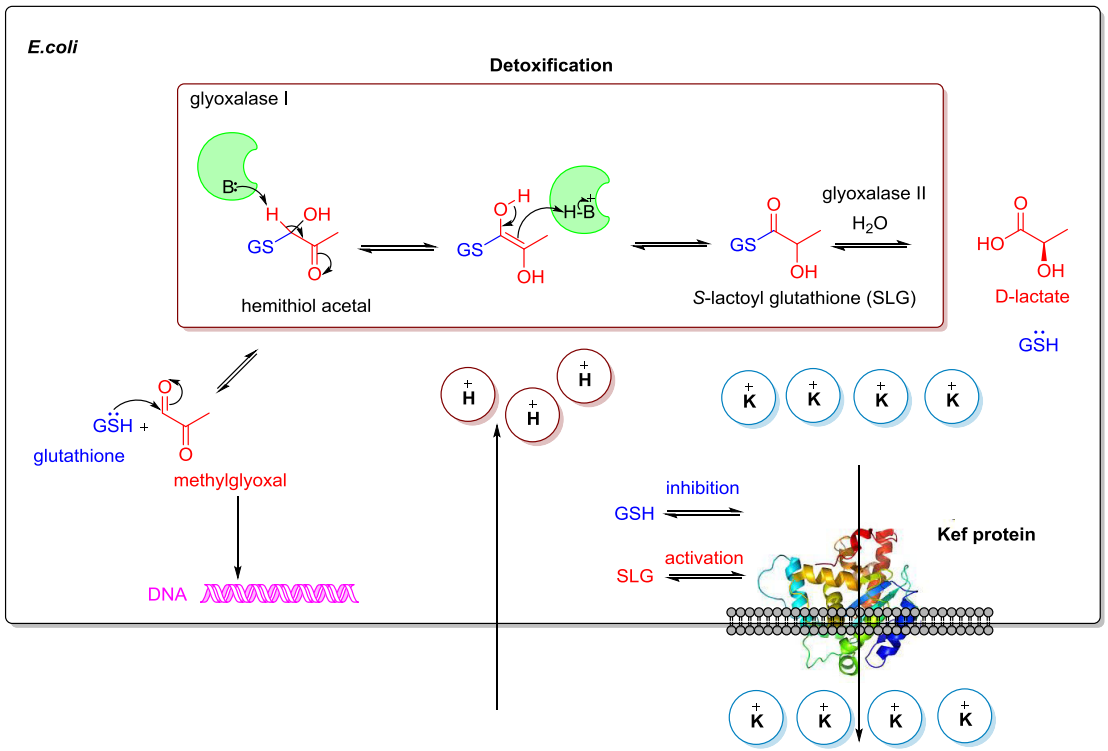


Figure 4.1 Illustration of bacterial K^+ efflux system.
Toxic electrophilic species (X) e.g methylglyoxal occurs widely in bacterial cells (left bottom). They react with bacterial DNA nucleophiles and causing them serious damage. In a parallel detoxification process, cellular glutathione (GSH) also reacts with these toxic electrophiles and form less toxic adducts (GS-X). Glyoxalase enzymes carry out the detoxification of methylglyoxal and the other reactive aldehydes that are produced as by-products of cellular metabolism (schemes in box). This detoxification involves the sequential action of two thiol-dependent enzymes; glyoxalase I, which catalyses the isomerisation of the spontaneously formed hemithioacetal adduct between GSH and 2-oxoaldehydes (such as methylglyoxal) into *S*-2-hydroxyacylglutathione. Glyoxalase II hydrolyses these thiolesters and in the case of methylglyoxal catabolism, produces D-lactate and GSH from *S*-D-lactoyl-glutathione. GSH also binds to the K^+ efflux protein and its binding results in closing (inhibition) of the K^+ channels (right bottom). Upon exposure to toxic electrophiles, the binding of GS-X to K^+ efflux protein results in opening of the K^+ channels allowing K^+ efflux and concomitant H^+ influx and thus protects bacterial *via* acidification of cellular matrix.

Toxic electrophilic species such as methylglyoxal exist widely inside bacterial cellular matrix (Figure 4.1). These electrophiles cause damage to bacterial cells through their reaction with the nucleophiles of bacterial DNA bases and functionally important side chains especially cysteine.[7,8] The cytoplasmic glutathione, which is abundant in cells (at concentration ~ 10-20 mM) acts as a scavenger of these harmful electrophiles (X), by reacting with them either chemically or enzymatically (glutathione-S-transferase),[9] forming glutathione-S-X conjugates (GS-X) and thus protects bacteria against the detrimental effects of these electrophiles

(Figure 4.1). In addition to its protective role, the cytoplasmic glutathione also binds the Kef protein (KTN domain) and causes the closing of the K^+ ion channels, resulting in inactivation of the K^+ efflux system of the cells (Figure 4.1).[10] Interestingly, however, the binding of glutathione-S-conjugates (GS-X) such as *S*-lactoyal glutathione (SLG) causes the opening of the K^+ ion channels[6] and thus helps in regulating the K^+ concentrations in the cells (Figure 4.1). It is believed that the additional steric bulk on X in GS-X induces conformational changes in the Kef protein which triggers the opening of the K^+ channels.[6] These glutathione-S-conjugates or adducts formed, subsequently undergo metabolism to less harmful species in enzyme-mediated detoxification reactions that also regenerate the cytoplasmic GSH (Figure 4.1).[11,12] Despite this active detoxification system of toxic cellular species, their potentially lethal direct reaction with DNA and other important macromolecules in the cells always occur, causing damage to them. During electrophilic exposure, bacterial survival has been found to increase significantly through acidification of the cellular environment, presumably due to the decrease in nucleophile reactivity caused by protonation or partial protonation at lower pH values.[7,8] Therefore the opening of the K^+ channels results in the cellular K^+ efflux and concomitant H^+ influx (Figure 4.1), causing acidification (lowering of the pH) of the cytoplasm and thus protects bacteria against toxic electrophiles. Hence Kef activation is part of a bacterial combined protection system in which toxic electrophiles are broken down into non-toxic species and simultaneously neutralizing their effect by modulation of the cellular pH.

4.1.1 Crystal structure of Kef proteins

More than 17 crystal structures have been reported in the literature for the KTN domains.[10,13-19] KTN domains exist in different multimeric forms. For example in Kef efflux system, these are dimeric, while in potassium uptake channels (CgIK, Kch, MthK), these domains are octomeric (dimers of tetramers). In *Escherichia coli*, the Kef system consists of KefC and KefB. The *E. coli* KefC represents the most extensively and well-characterised K^+ efflux system.[8,20-23] The KefC consists of a membrane domain attached to the carboxy-terminal KTN domain through a flexible linker. Some Kef proteins require complexation with another ancillary protein to form stabilized conformation such as *E.coli* KefC (Figure 4.2, PDB ID: 3L9W) and KefB require binding with KefF and KefG, respectively for full activity,[19] while some Kef proteins such as Kef-QCTD (Q-linker C-terminal domain) from *Shewanella denitrificans* (hereafter denoted as *sd*) have been found to be not dependent on such ancillary proteins for stability(activity).[24] KefF is a small flavoprotein, which binds directly to the

KTN domain.[10,19] The *E. coli* KefC C-terminal domain (CTD) is a 380 residues dimeric protein, with each monomer complexed to the ancillary protein, the KefF (Figure 4.2). The GSH/GS-X binding site is located at the interface of the two subunits (Figure 4.2, top). A comparison of *E. coli* KefC C-terminal domain (CTD) crystal structures in complex with inhibitory ligand GSH (PDB ID: 3L9W, Figure 4.3a) and with synthetic activating ligand *N*-ethylsuccinimido-S-glutathione (ESG) (PDB ID: 3L9X, Figure 4.3b) reveals that both the GSH and ESG (GS-X) binds in the same pocket and in the same mode.[6]

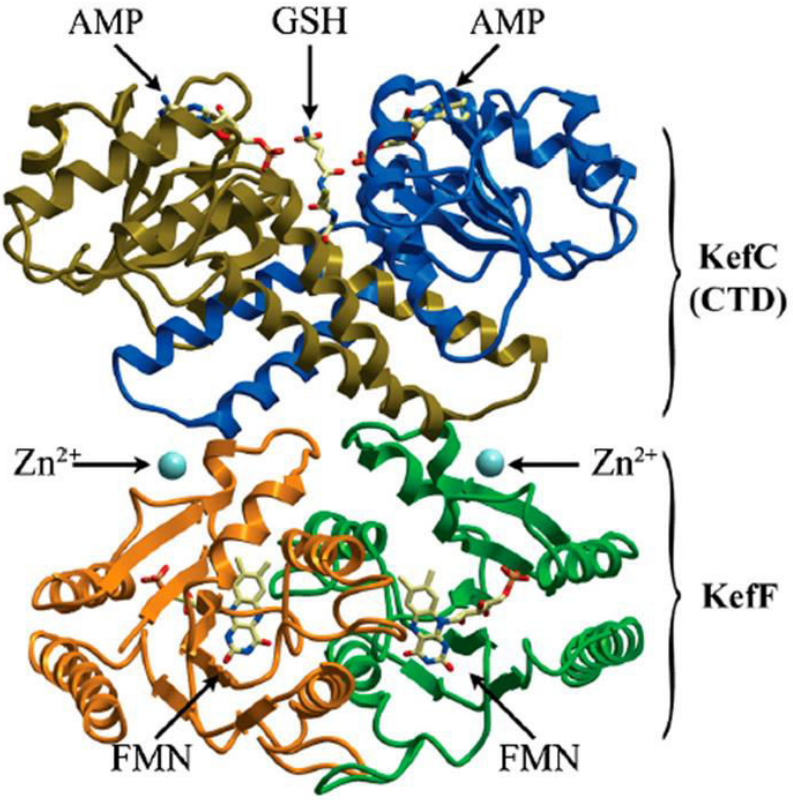
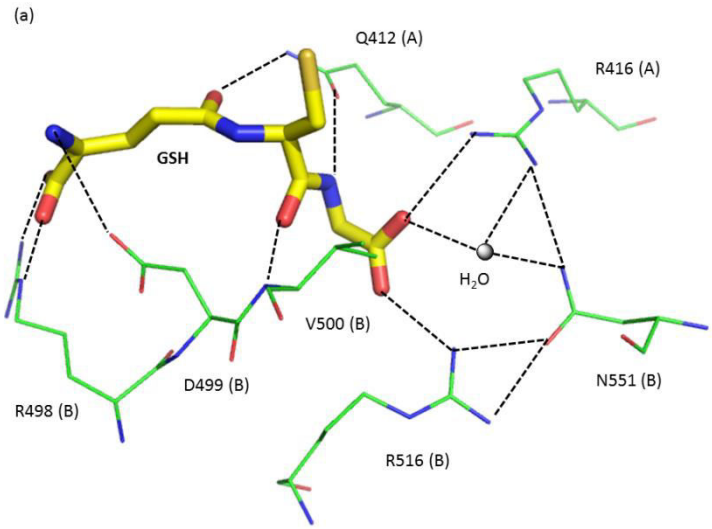
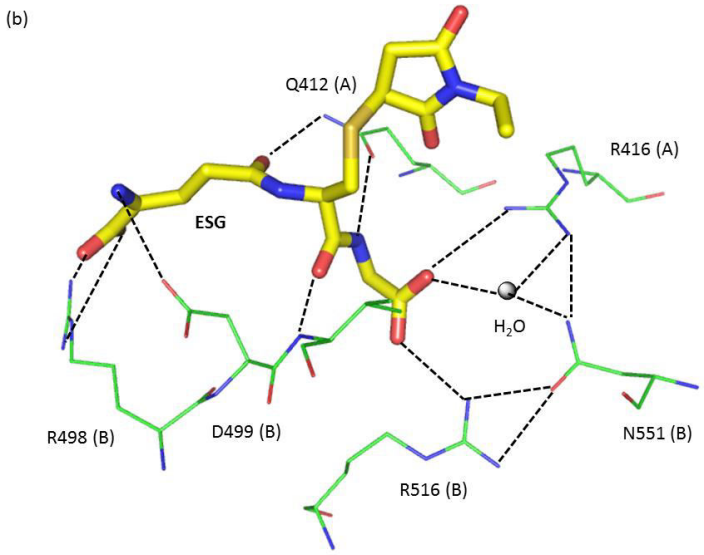


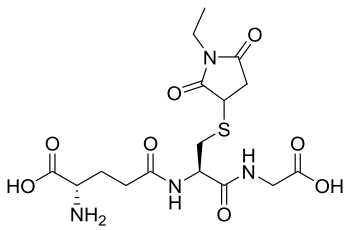
Figure 4.2 Crystal structure of the *E.coli* KefC-CTD-KefF dimer (PDB ID: 3L9W).
The two KefC-CTD monomers are blue and chartreuse. The GSH/GSX binding site is located at the interface of dimerization of the two KefC-CTD monomers. Each monomer contains an AMP molecule, which is believed to stabilize the dimeric form of the protein. Each KefC-CTD monomer is complexed with an ancillary protein KefF (a metalloenzyme) (green and golden). In the active site of KefF, flavin mononucleotide (FMN) and Zn^{II} are also shown. The complexation of KefC-CTD with KefF is believed to enhance the stability and activity of KefC-CTD. (Reproduced from reference [6])

Several key interactions from residues of both KTN domains can explain the binding of GSH/GS-X to the protein (Figure 4.3). Arg416 and Arg516 directly coordinate to the carboxylate of glycine, whereas, Asn551 indirectly forms interaction with the glycine carboxylate through a well-structured water molecule. Gln412 forms interactions with glycine –NH hydrogen and glutamate C-5 carbonyl oxygen. Arg498 forms interactions with

the glutamate C-1 carboxylate. Asp499 forms interaction with the glutamate (Glu) C-2 amino hydrogen. Val500 forms interaction with the cysteine carbonyl oxygen.







Structure of ESG

Figure 4.3 View from the active site of *E.coli* KefC-CTD crystal structures, in complex with (a) GSH (PDB ID: 3L9W) (b) ESG (PDB ID: 3L9X).
Key interactions are shown by dotted lines. A and B represents the two subunits of the dimer. Bottom: Structure of the *N*-ethylsuccinimido-S-glutathione (ESG).

Studies have shown that mutations of Q412 cause effects varying from reduced inhibition by GSH (Q412A) to complete insensitivity (Q412K) to both GSH and GS-X, suggesting it to a be crucial residue for the binding of GSH/GS-X.[6] It has been shown that Q412K mutation is functionally equivalent to a triple mutation affecting all three GSH glycine carboxylate group coordinating residues, R416, R516, and N551.[6] Similar results have also been reported for D499, where substitution of this residue with another hydrophilic group (D499S) does not alter the protein activity significantly, whereas substitution with either glycine or alanine results in a significant loss of Kef activation by ESG, suggesting it to be also a highly critical residue for the binding of ESG.[6] These mutations however do not affect the inactivation by GSH.

In the *E. coli* KefC-CTD crystal structure with GSH (PDB ID: 3L9W), the thiol group of GSH stabilizes the binding of Phe441 from α-helix 2 of one KTN subunit into a hydrophobic pocket formed by residues from α-helices 7 (M558) and 8 (T565, A569, Y572) from by the partnering KTN subunit (Figure 4.4A). In the crystal structure with activating ligand ESG (PDB ID: 3L9X) (Figure 4.4B), the attached group (electrophile) displaces this Phe441 (coloured red) to a different conformation (12 Å away). After displacement, Phe441 appears to form hydrophobic interactions with its own chain residues. It is believed that the conformational changes produced in the protein as a result of movement (displacement) of this Phe441 are likely triggering the opening of the K^+ channels by GSX in contrast to GSH. The binding of ESG also disrupts (breaks) the interaction between the α-helices 7 and 8 of one sub-unit with the α-helix 2 of the other sub-unit, making it completely disorder (not visible in the crystal structure). Mutation studies of Phe441 with either tyrosine or tryptophan or with a smaller hydrophobic residue i.e F441D have been shown to lower the level of KefC activation by ESG, without affecting the binding activity of ESG and GSH, suggesting that Phe441 acts a molecular sensor for Kef activation.[6]

The KTN domains also contain a highly characteristic Rossmann fold that hosts a nucleotide binding site (Figure 4.5). The binding of a nucleotide molecule has been proposed to stabilize the folded dimeric form of the KTN domains.

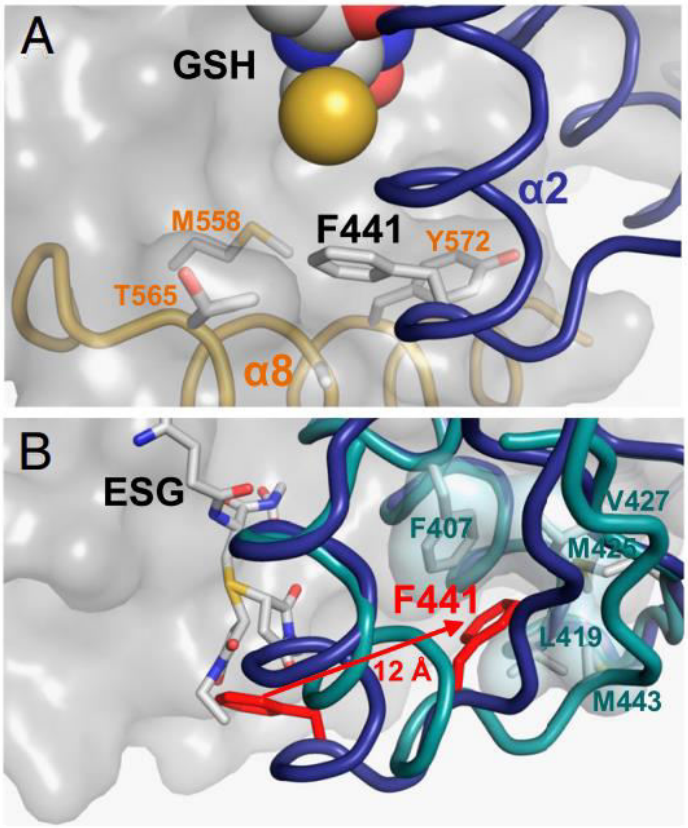


Figure 4.4 Conformational changes in *E.coli* KefC-CTD upon ESG binding.
Crystal structure with (A) GSH (PDB ID: 3L9W). (B) ESG (PDB ID: 3L9X). In the GSH bound structure (A), the GSH thiol group (yellow sphere) stabilizes the binding of F441 from α-helix 2 of one KTN subunit (blue) into a hydrophobic pocket formed by M558 (from α-helix 7), T565 and Y572 (from α-helix 8) from the partnering KTN subunit (gray surface). In the ESG (GS-X) bound structure (B), the attached chemical group of *N*-ethylsuccinimido-S-glutathione displaces crucial F441 (red) to an alternative conformation 12 Å away (shown by the arrow), where it forms new interactions with hydrophobic residues within its own chain. The binding of the ESG also disrupts the interactions between α-helices 7 and 8 of one sun-unit with α-helix 2 of the other sub-unit. (Reproduced from reference [6])

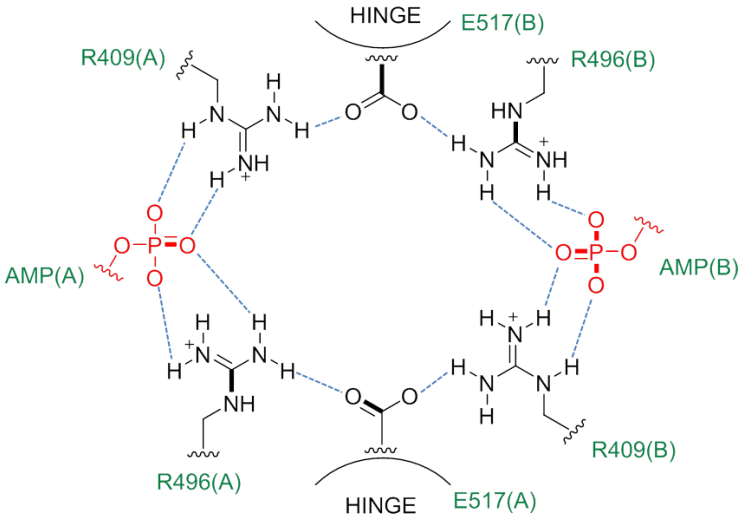


Figure 4.5 Interactions made by AMP in the *E.coli* KefC-CTD crystal structure in complex with GSH (PDB ID: 3L9W).
AMP is proposed to stabilize the dimeric form of KTN domains. A and B represents the two subunits of the dimer. The AMP phosphate group (red) appears to form a complex net of interactions including that with E517, the residue located centrally on the KTN hinge.

4.1.2 Structural requirements of GS-X for Kef activation

The fact that Gram-negative bacteria possess very large pools of GSH (~ 10-20 mM); this can potentially hinder the binding of the activating molecule (GS-X) to Kef protein (KTN domains). Therefore, the activating ligand must be a strong binder of Kef to compete strongly with such a large concentration of GSH.

The inhibitory ligand GSH has been reported to bind *sd*Kef-QCTD with an affinity (K_D) of 900 µM, whereas the activating ligand ESG binds much strongly with reported K_D values of 12 µM (fluorescence emission assay) and 23 µM (ITC).[24] This significant difference in the binding affinities of GSH and ESG to *sd*Kef-QCTD is attributed to the presence of large rigid hydrophobic ring of ESG. Interestingly, in *vivo* studies have shown that the formation of relatively low concentrations of ESG in the cytoplasm results in strong activation of KefC despite the presence of excess GSH.[6] The reported literature on the structure-activity relationship of glutathione-S-conjugates (GS-X) suggests that the presence of a bulky hydrophobic group in a rigid structure is the key for activation of the K^+ efflux system (*sd*Kef-QCTD) (Figure 4.6).[22-24]

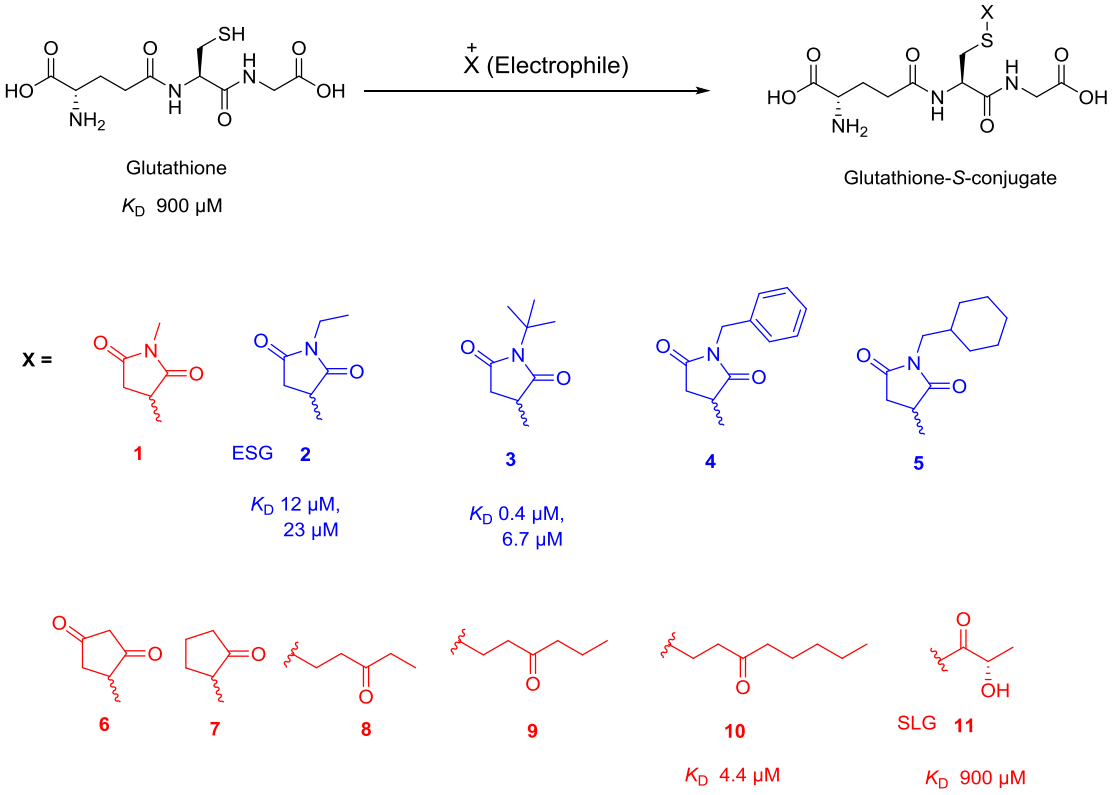


Figure 4.6 Structures of the reported glutathione-*S*-conjugates (GS-X) that had been tested for their effect in activating the *sd*KefC-QCTD K^+ efflux system.
GS-X with blue coloured electrophiles are strong activators of K^+ efflux system, whereas GS-X with red coloured electrophile shows less or no activation of *sd*Kef K^+ efflux system.[24]

In this series, compounds containing a rigid ring with a bulky substituent (coloured blue, **2**-**5**) have been shown to effectively activate the K^+ efflux system (with similar level of activation by all).[24] However, compound **1** with smaller hydrophobic substituent, compound **6**, which contains two keto groups, similar to the succinamide of compounds **2**-**5**, but lacks the substituted nitrogen atom, and compound **7** which possess only one keto group and also no substituted nitrogen atom, have been shown to result in little or no activation of K^+ efflux. Similarly the open structure compounds, **8**-**11** have also been shown to result in significantly less activation of potassium efflux supporting the requirements of a rigid bully hydrophobic group (X) for effective activation of potassium efflux systems.

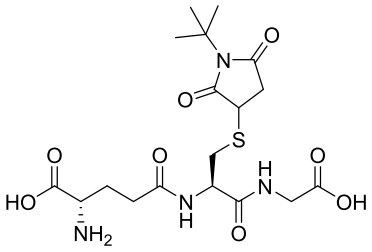
The observation that why glutathione-S-conjugates with small electrophile do not activate K^+ efflux system,[6] is because they can fit within the GSH/GS-X binding pocket without displacing the critical F441 residue. This may possibly occur either by slight repositioning of the F441 ring, or by the chemical group of the adduct occupying an alternate pocket adjacent to the thiol moiety of GSH. Such flexibility is not available for the more rigid ring structure like in the activating ligand ESG.

4.2 Aims of the project

In humans, to date, there are no reported homologues of KefB or KefC. Since the rate and extent of Kef activation during bacterial exposure to toxic electrophiles is a critical determinant of their survival, therefore, perturbation of Kef activity is potentially a novel target for the development of antibiotic drugs. Antibiotic resistance is an escalating problem, with multi-drug resistant pathogens becoming commonplace. To circumvent this resistance, it is essential that antibiotics demonstrating novel mechanisms of action are developed.

We are interested in developing non-peptidic activators of K^+ efflux systems. In order to develop such novel potential antibacterial drug molecules, an understanding of the binding interaction of GS-X with Kef KTN domains at atomic levels is highly critical. Although the crystallographic and genetic studies of *E.coli* KefC have identified residues that are likely to be involved in GSH/ESG binding, however, it is still unknown whether all the functional groups of the GS-X that are making interactions with the protein residues (as shown in the crystal structures, Figure 4.3) are required for binding or perhaps some groups are crucial for binding and some are not. The identification of these essential functional groups of the GS-X would be of significant importance in the development of non-peptidic antibacterial agents of Kef. Hence, one of the aims of this project was to identify the functional groups or atoms of

the GS-X that are crucial for binding with the Kef protein, using the glutathione-S-conjugate formed with the *N*-tertiary butylsuccinimide (hereafter termed as ^{t}BuSG for simplicity) (Figure 4.7), which is the most strongly binding (K_D 0.4 µM) activator of *sd*Kef-QCTD.[24]



3

^{t}BuSG

Figure 4.7 Structure of glutathione-S-conjugate with *N*-tertiary butylsuccinimide (^{t}BuSG)

As previously, biophysical techniques such as isothermal calorimetric titration (ITC), fluorescence-competition assays and differential scanning fluorimetric assays (thermal shift analysis) have been employed to study small molecule binding interactions with *sd*Kef-QCTD protein,[24] we therefore aimed to develop efficient solution-based NMR assays as another technique to readily detect and quantify small molecule binding events with Kef proteins.

To date it has also been unclear which nucleotide-based ligand binds in the Rossman fold of Kef KTN domains, where it comes from, and what is its exact role in the biology of these proteins. Therefore we also aimed to explore the identity of this nucleotide ligand and to evaluate its importance in the biological function of Kef proteins.

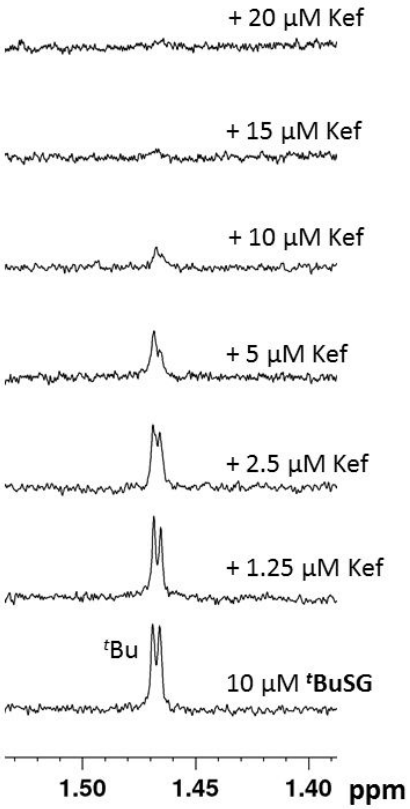
Since the instability of *E. coli* KefC-KefF dimeric protein has been previously shown to affect the assay results,[24] our investigations used the stable *sd*Kef-QCTD protein produced by Dr Anthony K. N. Chan), which also does not require the ancillary protein (KefF) for activity unlike *E. coli* KefC. As far as the biological activity of *sd*Kef-QCTD is concerned, both the *sd*Kef-QCTD and *E.coli* KefC-KefF behaves similarly, i.e Q419K mutation has been found to result in significant loss in GSH binding in *sd*Kef-QCTD as observed in *E. coli* KefC-CTD-KefF Q412K. In addition, similar to *E. coli* KefC-CTD-KefF F441, the *sd*Kef-QCTD F448 has also been shown to be crucially required for activation of the K^+ efflux.[24]

4.3 Identification of the critical functional groups of glutathione-S-*N*-tertiary butylsuccinimide (GS-X) required for binding with *sd*Kef-QCTD

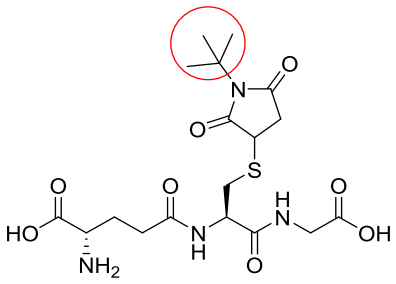
We first focused on identification of the glutathione-S-*N*-tertiary butylsuccinimide (^{t}BuSG) functional groups that form crucial interactions with *sd*Kef-QCTD. We first synthesized these novels dissection compounds of glutathione-S-*N*-tertiary butylsuccinimide (^{t}BuSG) (synthesis by Dr Samuel C. Grayer) and then we applied the CMPG edited ^{1}H NMR direct-ligand observation technique, monitoring the broadening and/or attenuation of ligand NMR resonances in the presence of protein (described in section 1.4.1) to study their binding interaction with *sd*Kef-QCTD. The results of this section are discussed below.

To test first the applicability of CPMG edited ^{1}H NMR technique in studying small molecules binding interaction with *sd*Kef-QCTD, we first applied this methodology to study the binding interaction of ^{t}BuSG to *sd*Kef-QCTD. Due to its high NMR sensitivity, we decided to follow the ligand tertiary butyl group Me_3 signal at 1.48 ppm for binding interaction with the protein. We observed that in the presence of *sd*KefC-QCTD, the ^{t}BuSG signals are only attenuated in intensity without broadening (Figure 4.8), suggesting its strong binding interaction with *sd*Kef-QCTD (NMR slow exchange system). By titration of *sd*Kef-QCTD to a fixed concentration of ^{t}BuSG (Figure 4.8a), and measuring the attenuation (loss) of ^{t}BuSG ^{t}Bu signal (circled red) as a function of protein concentration (Figure 4.8b), we obtained the ^{t}BuSG binding dissociation constant (K_D) value of 0.54 ± 0.03 μM (Figure 4.8c), which is in good agreement to its reported binding affinity value of 0.4 ± 0.2 μM determined by using a fluorescence-based competition assay.[24]

(a)



Kef/ µM	I_P	$\Delta I\,(\%) = (\frac{I_P - I_B}{I_B}) * 100$
0	1	0
1.25	0.8926	-10.74
2.5	0.7768	-22.3
5	0.5153	-48.47
10	0.2214	-77.86
15	0.079	-92.09
20	0.0501	-94.98



3

^{t}BuSG

(b)

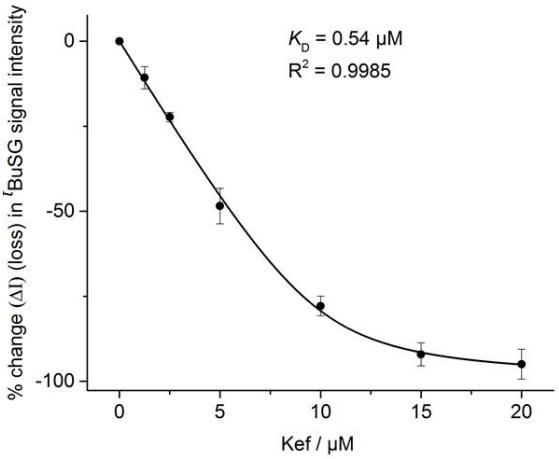


Figure 4.8 Demonstrating the binding interaction of ^{t}BuSG with *sd*Kef-QCTD by CPMG edited ^{1}H NMR direct ligand observation.

(a) NMR spectra (part shown) showing *sd*Kef-QCTD titration to fixed concentration of ^{t}BuSG; only the ligand Me_3 peak (in red circled in the structure) at 1.48 ppm is shown. The appearance of doublet for tertiary butyl group of the ^{t}BuSG is because the ligand solution is a racemic mixture. Table showing the relative loss (average response of the doublet) in the ^{t}BuSG Me_3 peak intensity as a function of protein concentration. The intensity of this signal in the absence of protein (I_B) was set to unity. I_P represents ^{t}BuSG Me_3 peak intensity in the presence of protein and ΔI (%) denotes the percent change (decrease) in the ligand signal intensity in the presence of protein. (b) Fitting of the titration data from the table according to the quadratic equation for K_D, using Origin Pro. A K_D of 0.54 ± 0.03 µM was obtained for ^{t}BuSG. The error bars represents standard deviation from three separate measurements. The concentration of ^{t}BuSG was 10 µM. The assay mixture was prepared in 50 mM sodium phosphate (NaH_2PO_4/Na_2HPO_4, supplemented with 150 mM NaCl) buffer, pH 7.4 in D_2O.The filter time used in the CPMG ^{1}H NMR experiments was 88 ms.

To prove that the binding interaction of ^{t}BuSG with *sd*Kef-QCTD was specifically in the GSH binding site, two types of control experiments were carried out; these are,

Firstly, a CPMG edited ^{1}H NMR reporter displacement experiment (as described in section 1.4.2) was performed, using ^{t}BuSG as reporter ligand and ESG (which is known to bind in the GSH binding site from its reported crystal structure, PDB ID: 3L9X) as competitive ligand (Figure 4.9). It was observed that the addition of ESG to the sample of ^{t}BuSG plus *sd*Kef-QCTD in buffer solution resulted in recovery of the ^{t}BuSG signals in the spectrum, because ^{t}BuSG is no longer bound to the protein i.e displaced by ESG binding, demonstrating that ^{t}BuSG binds to *sd*Kef-QCTD specifically in the ESG (GS-X) binding site.

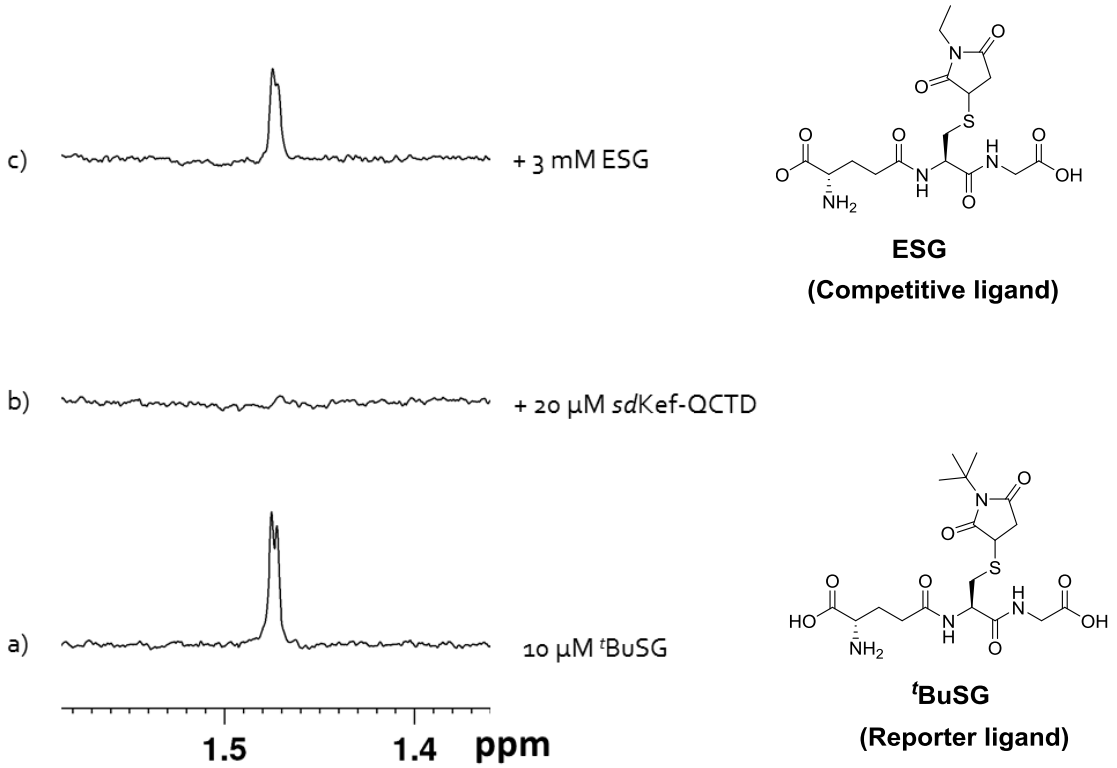


Figure 4.9 Demonstrating the site specific binding interaction of ^{t}BuSG with *sd*Kef-QCTD by NMR reporter displacement experiment in CPMG edited ^{1}H NMR.
(a) Blank ^{t}BuSG as reporter ligand, ^{t}Bu signal at 1.48 ppm is shown in the spectrum. (b) The ^{t}BuSG signals are attenuated in the presence of *sd*Kef-QCTD, demonstrating its binding interaction with the *sd*Kef-QCTD. (c) The addition of ESG as competitive ligand to the sample of (b) results in recovery of ^{t}BuSG signals, suggesting that ESG has displaced ^{t}BuSG from its binding site into solution. The signals of ESG are not shown in the spectrum.

As *sd*Kef-QCTD Q419 residue (Q412 in *E. coli* Kef-CTD) has been previously shown as a critical residue for GSH binding,[24] we therefore prepared Q419K mutant *sd*Kef-QCTD and investigated the binding interaction of ^{t}BuSG to this mutant protein. It was observed that in the presence of Q419K mutant *sd*Kef-QCTD, the ^{t}BuSG signals were neither attenuated nor

broadened, suggesting that ^{t}BuSG does not bind to Q419K mutant *sd*Kef-QCTD (Figure 4.10). This control experiment further confirmed that the binding of ^{t}BuSG with *sd*Kef-QCTD was specifically in the GSH binding site.

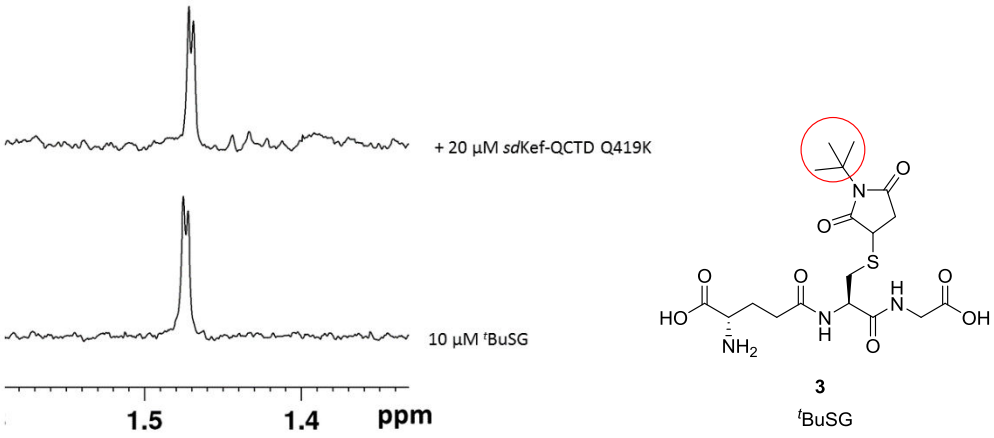


Figure 4.10 Experiment to prove that ^{t}BuSG doesn't bind mutant Q419K *sd*Kef-QCTD.
Monitoring the binding interaction of ^{t}BuSG (only the 9H ^{t}Bu signal shown in red circle in the structure at 1.48 ppm is shown in the spectra) with mutant Q419K *sd*Kef-QCTD by direct ligand observation in CPMG edited ^{1}H NMR experiment. Bottom spectrum is the ^{t}BuSG solution in buffer in the absence of Q419K *sd*Kef-QCTD and top spectrum after the addition of Q419K *sd*Kef-QCTD. In the presence of mutant Q419K *sd*Kef-QCTD, the ^{t}BuSG NMR resonances are neither broadened, nor attenuated, demonstrating that ^{t}BuSG doesn't bind to the mutant Q419K *sd*Kef-QCTD.

After establishing that the CPMG edited ^{1}H NMR direct ligand-observation is a suitable technique for studying ligand binding to *sd*Kef-QCTD, we then begin to assess the influence of each functional group of ^{t}BuSG on its binding interaction with *sd*Kef-QCTD. The results of this structure-activity relationship study are shown in Figure 4.11. The binding dissociation constant (K_D) values of all these ligands were determined by the CPMG ^{1}H NMR technique as shown in Figure 4.8. For the ligands shown as non-binders of *sd*Kef-QCTD, no broadening and/or attenuation of their NMR resonances were observed in the presence of protein. The spectra and binding curves of this structure-activity relationship study are shown in Appendix, Section 7.1

We began the dissection of ^{t}BuSG (binding affinity with *sd*Kef-QCTD 0.54 ± 0.03 µM) skeleton from the left side of the molecule, by first removing the carboxylate of glutamate (compound **A**/SG91D, Figure 4.11a; from left to right). It was observed that removing the glutamate carboxylate group of ^{t}BuSG does not lead to any significant loss in the ligand binding affinity to *sd*Kef-QCTD i.e K_D 1 µM, implying that the ^{t}BuSG glutamate carboxylate group does not seem to play a critical role in its binding interaction with *sd*Kef-QCTD, in agreement with the *sd*Kef-QCTD crystal structure i.e I505 interaction not possible unlike *E. coli* R498, Figure 4.11b).

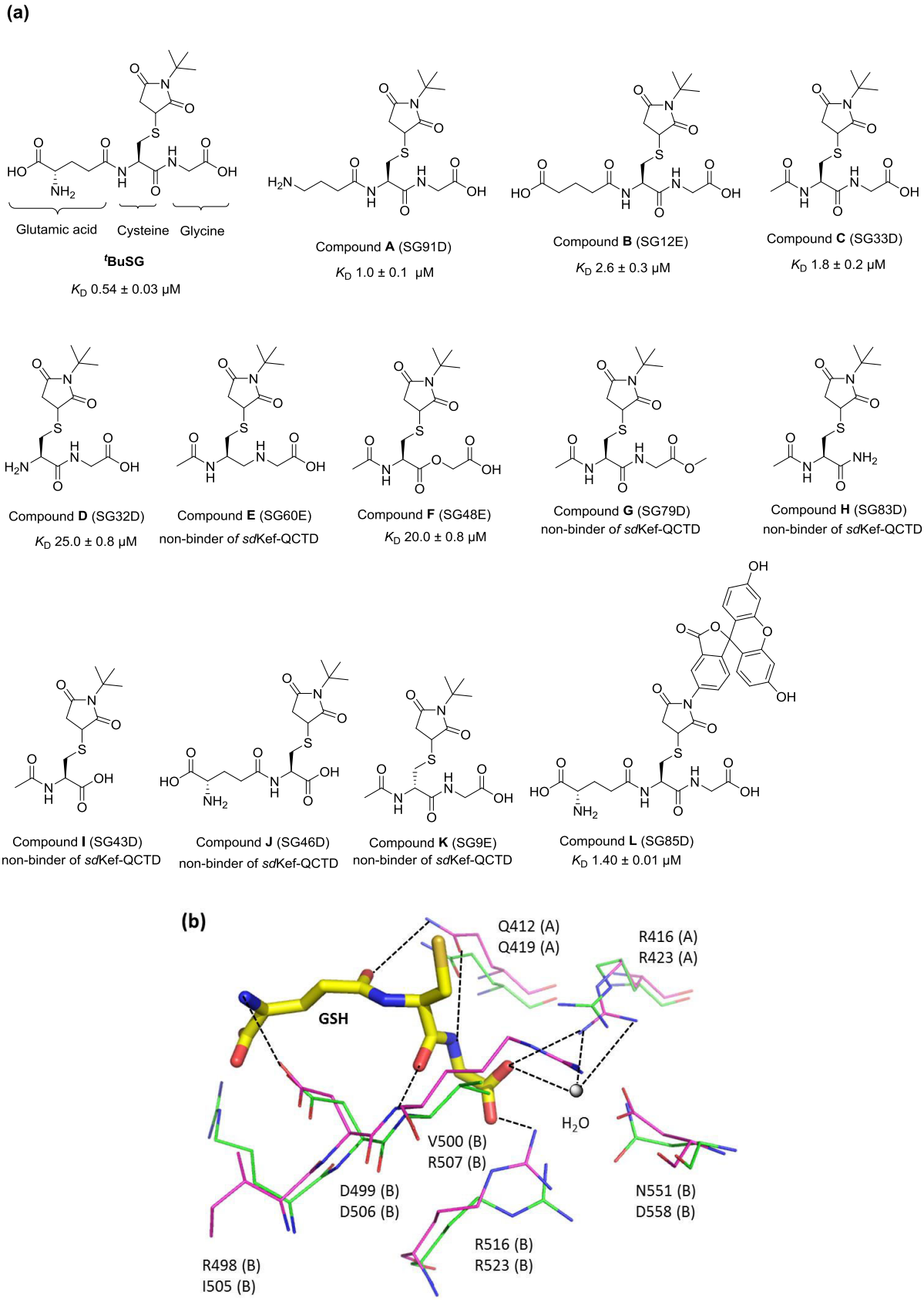


Figure 4.11 Assessing the influence of different functional groups of ^{t}BuSG on its binding interaction with *sd*Kef-QCTD. (a) Structure of the tBuSG dissection compounds (GS-X).(b) Overlay of the X-ray crystal structure of *E.coli* KefC (PDB ID: 3L9W) (green sticks, upper residue numbers) in complex with inhibitory ligand GSH (Yellow) with *sd*Kef (pink sticks, lower residue numbers). Interactions between the ligand (GSH) functional groups and the protein residues are highlighted by black dashed lines.

In the second compound (compound **B**/SG12E, Figure 4.11a) the amino group of glutamate was removed and it was found that the removal of glutamate amino group also does not result in any significant loss in the ligand binding affinity to *sd*Kef-QCTD i.e K_D 2.6 ± 0.3 µM, suggesting that the interaction of D506 carboxylate group with the ^{t}BuSG glutamate amino group do not play a critical role in the binding of ^{t}BuSG with *sd*Kef-QCTD.

In the third compound (compound **C**/SG33D, Figure 4.11a), both the carboxylate and amino groups together with the β-carbon of glutamate were removed and it was observed that this truncation of glutamate also do not cause any significant loss in the ligand binding affinity with *sd*Kef-QCTD i.e K_D 1.8 ± 0.2, implying that the presence of carboxylate and amino group together with the β-carbon of glutamate in ^{t}BuSG do not appear to play a major role in its binding interaction with *sd*Kef-QCTD.

In the fourth compound (compound **D**/SG32D, Figure 4.11a) the glutamate was completely removed and it was observed that removal of complete glutamate results in a significant loss in the ligand binding affinity with *sd*Kef-QCTD i.e K_D 25.0 ± 8 µM. This suggests that the interaction of Q419 with the ^{t}BuSG glutamate ε-carbonyl oxygen plays a significant role in the ligand binding interaction with *sd*Kef-QCTD, in agreement with the previous mutation studies of this residue in both *sd*Kef-QCTD[24] and *E. coli* KefC-KefF.[6]

In the fifth compound (compound **E**/SG60E, Figure 4.11a), the cysteine carbonyl oxygen of **C** was replaced with hydrogens and it was found that removal of this oxygen results in a complete loss of ligand binding with *sd*Kef-QCTD, demonstrating that the presence of cysteine carbonyl oxygen in ^{t}BuSG is necessarily required for its binding (interaction with R507 is critical, Figure 4.11b) with *sd*Kef-QCTD.

In the sixth compound (compound **F**/SG48E, Figure 4.11a), the -NH group of glycine of **C** was replaced with an oxygen and it was observed that the removal of this –NH group results in a significant loss in the binding affinity of **C** with *sd*Kef-QCTD i.e K_D 20.0 ± 0.8 µM due to the loss of crucial interaction with Q419 (Figure 4.11b). Hence, the presence of this –NH group in ^{t}BuSG also significantly contribute to its binding interaction with *sd*Kef-QCTD, but not necessarily required.

In the seventh compound (compound **G**/SG79D, Figure 4.11a), the glycine carboxylic acid hydrogen of **C** was replaced with a methyl group and it was found that this substitution results in complete loss of binding of **C** with *sd*Kef-QCTD. It is possible that the cavity

occupied by hydrogen is small and may not accommodate a methyl group or perhaps due to attachment of methyl group to the carboxylate oxygen, this oxygen may no longer able to form hydrogen bonding interaction with R523 (Figure 4.11b). Thus, for the binding of ^{t}BuSG with *sd*Kef-QCTD, this hydrogen also plays a critical role in binding of the GS-X ligand to the *sd*Kef-QCTD.

In the eighth compound (compound **H**/SG83D, Figure 4.11a), the glycine carboxylate together with its α-carbon is removed from **C** and it was revealed that this truncation of glycine results in a complete loss of binding of **C** with *sd*Kef-QCTD. Hence, the presence of glycine carboxylate and its α-carbon are necessarily required for the binding interaction of ^{t}BuSG with *sd*Kef-QCTD.

In the ninth compound (compound **I**/SG43D, Figure 4.11a), to investigate if the whole glycine is required for the binding of **C** with *sd*Kef-QCTD, we replaced the glycine with a –OH group. However, it resulted in complete loss of binding of **C** with *sd*Kef-QCTD, suggesting that the presence of complete glycine in ^{t}BuSG is necessarily required for its binding interaction with *sd*Kef-QCTD.

The provide further support to our results that glutamate carboxylate and amino group together with the **C** β-carbon has no significant role in the binding of ^{t}BuSG with *sd*Kef-QCTD, we extended the acetyl group of compound **I** to full glutamate group (compound **J**/SG46D, Figure 4.11a). In spite of this modification, compound **J** still failed to bind *sd*Kef-QCTD.

In investigating, the requirement of configuration (stereochemistry) at the cysteine stereocentre of **C**, we found that the compound possessing *R* configuration (**K**/SG9E, Figure 4.11a) fails to bind with *sd*Kef-QCTD, suggesting that *S* configuration of cysteine stereocentre ^{t}BuSG is essential for its binding with *sd*Kef-QCTD.

Finally, to explore the size of the hydrophobic group attached to the maleimide nitrogen of ^{t}BuSG, the size and rigidity of which in the glutathione-S-conjugates are critical for activation of K^{+} efflux due to the displacement of F441 (*E. coli,* F448 in *sd*Kef-QCTD), we replaced the tertiary butyl group of ^{t}BuSG with a fluorescein (compound **L**/SG85D, Figure 4.11a), one of the bulkiest hydrophobic groups. Interestingly, it was found that the compound still binds with *sd*Kef-QCTD with an affinity (K_D 1.40 ± 0.01 μM) in close range to that of

^{t}BuSG (K_D 0.54 ± 0.03 μM). This suggests that the hydrophobic group of ^{t}BuSG is likely exposed towards an open space on the protein.

The results of this structure-activity relationship (SAR) study conducted by CPMG edited ^{1}H NMR direct ligand-observation were then validated with a reported qualitative fluorescence-competition assay[24] (performed by Dr Samuel C. Grayer) described below (Figure 4.12).

In this technique the glutathione-S-conjugate with the dansyl chromophore (DNGSH, Figure 4.12a) is used as a fluorescence probe. The dansyl chromophore is known as highly sensitive to the nature of its local hydrophobic chemical environment. The small size of the dansyl group also makes it a good choice as a fluorescence probe as it is less likely to hinder ligand binding. This assay has been previously used for studying the binding interactions of both of natural and synthetic glutathione-S-conjugates to *sd*Kef-QCTD.[24]

In this fluorescence competition assay (Figure 4.12) when DNGSH binds to *sd*Kef-QCTD (both are incubated), its fluorescence quantum yield (intensity) increases and a significant shift in its λ_{max} is observed in the emission spectrum (Figure 4.12b). These changes in the fluorescence properties of DNGSH reflect changes in the hydrophobicity of the local chemical environment experienced by the probe. Under our experimental conditions, the λ_{max} of unbound probe (λ_L) is found at 572.5 nm and that of the bound probe (λ_{ML}) with high *sd*Kef-QCTD concentration could be found at ~ 530 nm. With *sd*Kef-QCTD Q419K, no changes in the λ_{max} of DNGSH occurred, suggesting that DNGSH binds *sd*Kef-QCTD in the GSH binding site. Upon displacement of DNGSH from its binding site by a competitive ligand, a decrease in its fluorescence intensity (Figure 4.12c) and a shift in its λ_{max} towards the free probe in solution (λ_L) are observed. The binding of the competitive ligand is denoted as a $DNGSH_{bound}$:$DNGSH_{ligand}$ ratio i.e [F_B:F_L], (B for bound and L for ligand) with a higher ratio represents a high affinity interaction of the competitive ligand and vice versa.

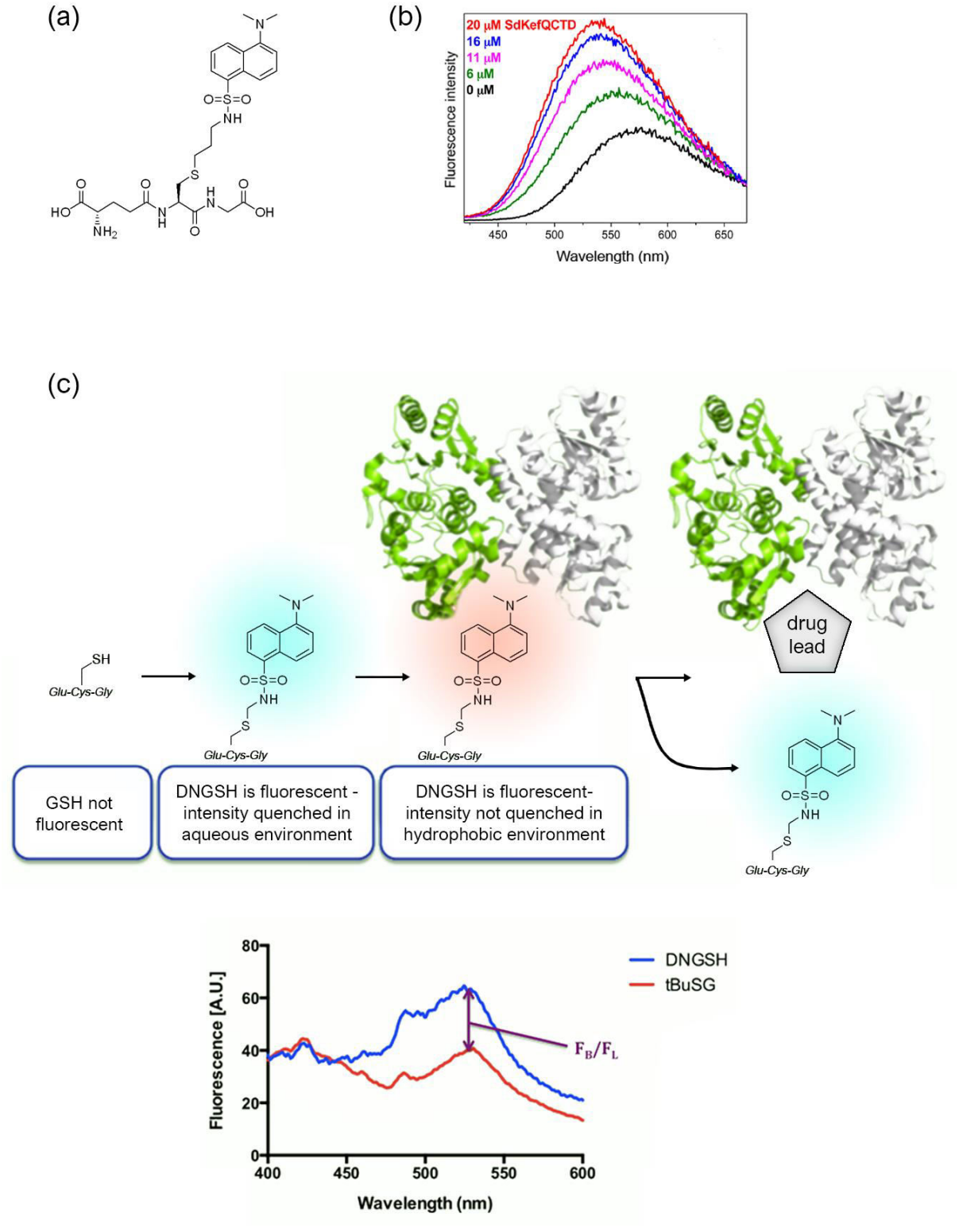


Figure 4.12 Illustration of the fluorescence competition assay for studying ligand binding to *sd*Kef-QCTD using DNGSH as a fluorescence reporter ligand.
(a) Structure of DNGSH (b) Titration of *sd*Kef-QCTD to 5 μM of DNGSH. (c) Overview of the fluorescence competition binding assay using DNGSH (5 μM) and *sd*Kef-QCTD (6 μM). The graph below shows DNGSH fluorescence signal intensity before and after the addition of competitive ligand ^{t}BuSG (1 mM).

This fluorescence competition assay was applied here to study the binding interaction of the ^{t}BuSG dissection compounds tested in the structure-activity relationship study (Figure 4.11) and the results are summarised below (Figure 4.13).

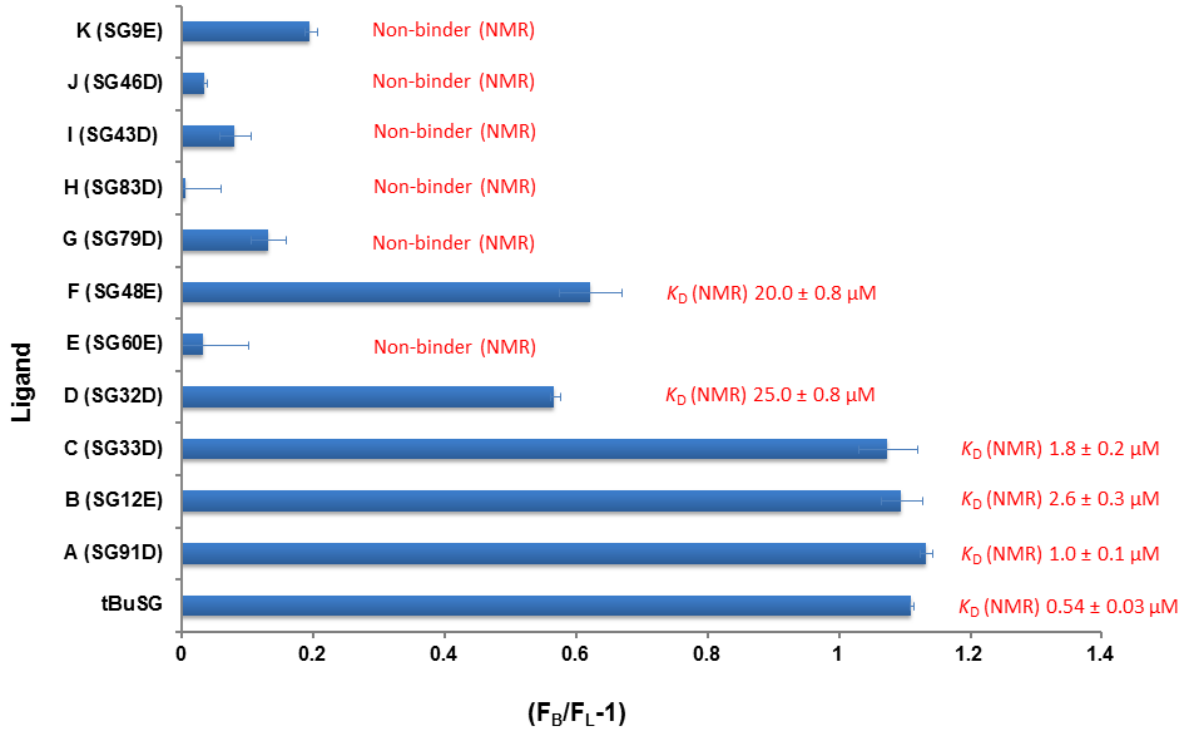


Figure 4.13 Results of the fluorescence competition binding assay for the compounds tested in SAR study.

These results demonstrate that ^{t}BuSG, compound **A** (SG91D), **B** (SG21E) and **C** (SG33D) have high values of F_B:F_L ratio, suggesting that they are the strong binders of *sd*Kef-QCTD in this series. The close F_B:F_L ratios of these four compounds suggest that the binding affinities of these ligands with *sd*Kef-QCTD are very close to each other, in agreement with the results of CPMG edited ^{1}H NMR assay for these ligands. Relative to these four compounds, compound **D** (SG32D) and **F** (SG48E) have lower F_B:F_L ratios, suggesting that they are relatively weak binders of *sd*Kef-QCTD in this series, also in agreement with the results of CPMG edited ^{1}H NMR for these ligands. Similarly, the significantly lower values of F_B:F_L ratios for compounds **E** (SG60E), **H** (SG83), **I** (SG43D), **J** (SG46D) and **K** (SG9E) suggests that these ligands are likely the non-binders or extremely poor binders of *sd*KefC-QCTD in the series.

Hence, based on our SAR results, in our journey towards the development of non-peptidic antibiotic of Kef systems, we replaced the glycine carboxylate of **C** (SG33D) with a non-peptidic moiety; the tetrazole ring (compound **M**/SG79E, Figure 4.14). Interestingly, the

compound was evaluated as a good binder of *sd*Kef-QCTD with an affinity of 7.6 ± 0.5 µM determined by CPMG edited ^{1}H NMR assay. The binding of this compound is thus a promising step towards development of novel non-peptidic drugs against Kef systems.

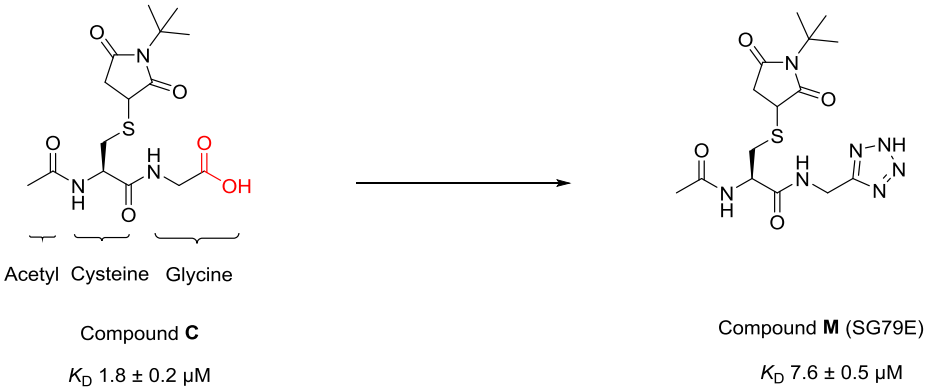


Figure 4.14 Structure of the non-peptidic GS-X synthesised as a results of ^{t}BuSG disection studies.
In the non-peptidic ligand **M** (SG79E) of *sd*Kef-QCTD, the carboxylate of glycine of **C** was replaced with a non-peptidic moiety; the tetrazole ring.

4.4 Development of an efficient ligand-based ^{19}F NMR binding assay for *sd*Kef-QCTD

The application of the CPMG edited ^{1}H NMR-based reporter displacement assays (discussed in section 1.4.2) in ligand screening can be sometimes limited due to the signal overlap of the reporter ligand with other molecules in the assay mixture. These overlap may arise from either small molecules impurities encountered in laboratory stock solutions, buffers, from contaminants protein purification e.g glycerol or the competitive ligands themselves.

Similarly, measuring a ligand binding dissociation constant (K_D) by NMR titration is a labour intensive technique and may require a large amount of materials and instrument time as compared with other biophysical techniques due to low intrinsic sensitivity of NMR. This usually limits the application of NMR techniques to only hit identification/validation and do not give an estimation of ligand binding affinity, which is an important parameter in ranking potential leads in early stages of drug discovery processes. Hence, there is a need for efficient NMR assays which can be used both for ligand screening and measuring their binding constant in a high throughput manner.

To overcome the problem of the reporter signals overlap in ^{1}H NMR-based reporter displacement assays, several nuclei of the reporter ligand such as phosphorus (^{31}P) [25], fluorine (^{19}F) [26], and carbon (^{13}C) [27,28] have been used. Amongst these, ^{31}P and ^{31}C reporter

displacement assays have received less attention for ligand screening due to their low intrinsic NMR sensitives.

On the other hand, ^{19}F nucleus has a comparable NMR sensitivity to that of a ^{1}H i.e 0.83 times that of a proton, and has a 100% natural abundance and therefore is attractive for ligand screening. The advantages of fluorine (^{19}F) NMR detection for ligand-based NMR screening applications as well as for ^{19}F magnetic resonance imaging (MRI) have been convincingly demonstrated in the past.[26,29-35] For example in pharma industries, the application of the FAXS (fluorine chemical shift anisotropy and exchange for screening) technique allows to screen mixture of small molecules against a target protein in a high throughput manner. Due to its large chemical shift anisotropy (CSA), the ^{19}F nucleus is very responsive (sensitive) to changes of molecular weight that accompany the binding events and has a large chemical shift range of ~ 300-400 ppm.

Typically, in ^{19}F NMR reporter displacement assays, a reporter ligand containing a CF_3 group is preferred, because of the advantage of the high sensitivity of the fluorine signal due to chemically similar three fluorine atoms.

In addition to removing the problem of signals overlap, the most striking feature of the use of a fluorinated reporter ligand is the possibility to observe two distinct NMR signals in the spectrum for a slow exchange ligand binding regime on NMR time scale; i.e a sharp signal for the free population of the reporter ligand and a broad signal for the bound population of the reporter ligand (as discussed in section 1.4.4).[36-39] By definition, in an NMR slow exchange system, the frequency of interconversion of the two forms of a system is lower than the difference of their frequencies on chemical shift scale. Hence the two forms appear as separate signals in the spectrum. The bound ligand signal is not usually possible to observe in ^{1}H NMR spectrum as it is usually buried under the protein resonances; instead the signal of the free reporter ligand is monitored for binding. In fast exchange reporter binding regime in both ^{1}H and ^{19}F NMR spectrum, it is the average single signal of the ligand between the bound and free form, which is being monitored for binding. The slow exchange reporter binding by ^{19}F NMR spectrum serves as an ideal system for reporter displacement assay for ligand screening in many aspects discussed as below,

The binding interaction of a small ligand (L) to a protein (P) at a single site can be represented as (described in section 1.2.2),

$$P + L \rightleftharpoons PL$$

The binding dissociation constant (K_D) of the ligand for this system is calculated as,

$$K_D = \frac{[P][L]}{[PL]} \qquad (1)$$

Where, [L], [P] and [PL] represent the concentrations of free ligand, free protein and protein-ligand complex at equilibrium respectively.

The values of [P] or [L] can be calculated from,

$$[P_T] = [P] + [PL]$$

Or, $$[P] = [P_T] - [PL] \qquad (2)$$

Similarly,

$$[L_T] = [L] + [PL]$$

Or, $$[L] = [L_T] - [PL] \qquad (3)$$

Where, $[P_T]$ represents the total concentration of protein and $[L_T]$ the total concentration of ligand present at equilibrium.

Since in the ^{19}F NMR slow exchange ligand binding system, the concentrations of free ligand [L] and protein-ligand complex [PL] are directly known from the spectrum, and by obtaining the concentration of free protein [P] from eq. 2, the binding dissociation constant (K_D) of the ligand can then be readily calculated from eq. 1, in contrast to measuring its K_D either by titrating protein and measuring the attenuation of the ligand peak intensity or by using other biophysical titration methods.

When the competitive ligand displaces the reporter ligand in such binding regime, the intensity of the signal of the bond population of the reporter ligand in the spectrum is decreased and that of the free population of the reporter ligand is increased. Again, from this spectrum, in the presence of competitive ligand, the apparent binding constant ($K_{D\ app}$) of the reporter ligand is directly calculated using eq. 1. From the knowledge of the actual and

apparent binding dissociation constant values of the reporter ligand, the binding constant of the competitive ligand ($K_{D\ comp}$) can be calculated using eq. 4 as described by Dalvit *et al.*[31]

$$K_{D\ comp} = \frac{[C]K_{D\ rep}}{K_{D\ app} - K_{D\ rep}} \qquad (4)$$

Where [C] represents the concentration of the competing ligand.

In general, both in slow and fast exchange reporter binding in ^{1}H NMR and fast exchange ^{19}F NMR experiments, the apparent K_D of the reporter ligand is measured by titration of the competitive ligand, while monitoring signal recovery of the reporter ligand.

Although there are some reported examples of the slow exchange reporter ligand binding system by ^{19}F NMR, however to our knowledge, in non-of these studies such system have been used for quantitative screening of the competitive ligands.[36-39]

The methodology and the application of this technique is ligand screening studies to *sd*Kef-QCTD is described here as below.

Two novel glutathione-S-conjugates termed here as SG59D and SG48D (Figure 4.15), one containing a single fluorine atom (SG49D), and the other containing three chemically similar fluorine atoms (SG48D), were synthesized (by Dr Samuel C. Grayer) according to the reported protocols in the literature.[24]

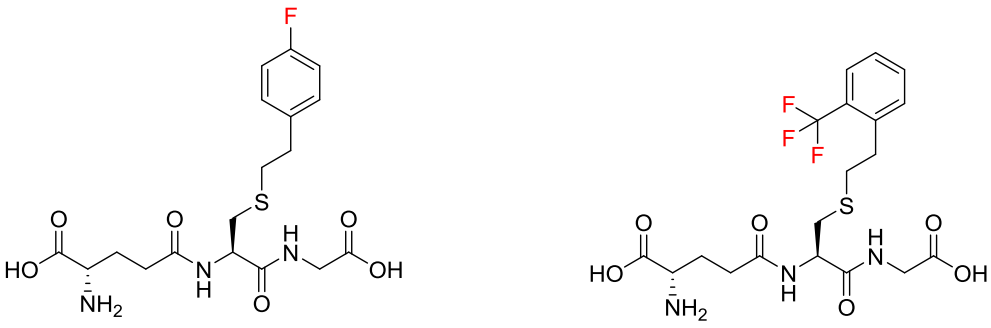


Figure 4.15 Structures of the two novel fluorinated glutathione-*S*-conjugates SG59D and SG48D synthesised for use as reporters ligand in ^{19}F NMR reporter displacement assay for ligand screening to *sd*Kef-QCTD.

By first investigating their stability under our assay conditions, it was observed that SG59D decomposes overtime, whereas, SG48D remains stable. Therefore, SG48D was selected as a reporter ligand.

We then investigated the binding interaction of SG48D with *sd*Kef-QCTD by standard ^{19}F NMR direct ligand observation (Figure 4.16). Interestingly, in the presence of *sd*Kef-QCTD, two distinct NMR resonances were observed in the spectrum, a sharp signal (F) at -59.07 ppm corresponds to the free population of SG48D and a broad signal (B) at -59.74 ppm corresponds to its bound populations, presenting a typical NMR slow exchange ligand binding system. This observation reflects the tight binding interaction of SG48D to *sd*KefC-QCTD.

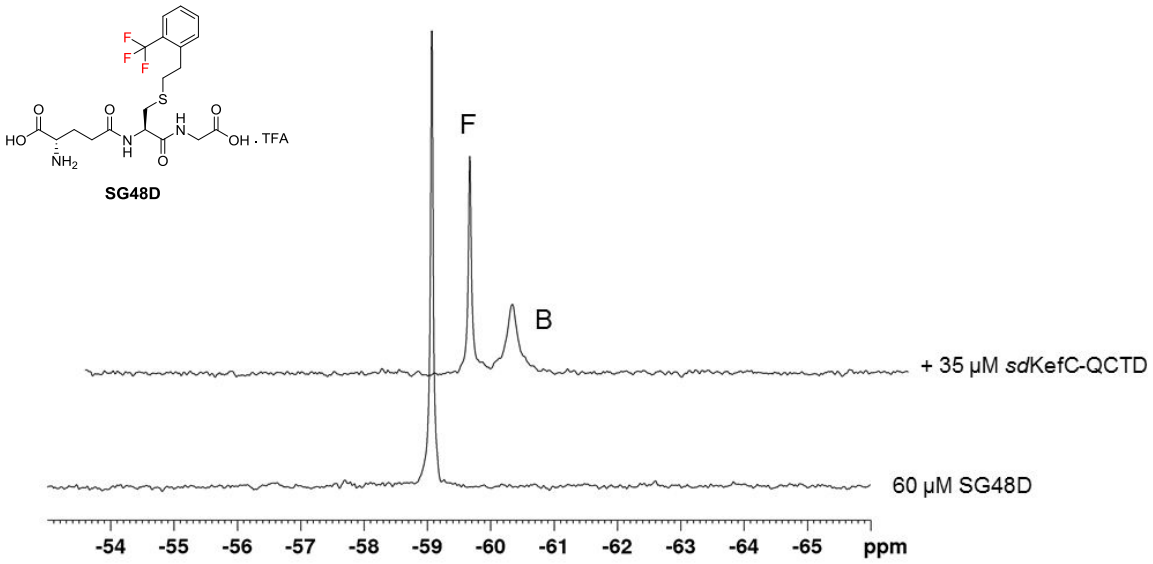


Figure 4.16 NMR slow exchange binding of SG48D with *sd*Kef-QCTD as demonstrated by direct ligand observation. The F_3 (coloured red in the structure) signal of SG48D is monitored in the spectrum. F represents the free population of SG48D and B the bound population of SG48D. (Note; the spectra are overlayed in a way so as to clearly see the signals in both the spectra. The chemical shift of F remains the same in both the spectra). The assay solution was made in 50 mM sodium phosphate buffer. The assay mixture contained 60 µM SG48D, 35 µM *sd*Kef-QCTD and 20 µM 1,1,1 trifluoroacetone (internal NMR control) in 50 mM sodium phosphate buffer, pH 7.4 in D_2O.

In this assay, we used the SG48D as a trifluoroacetic acid (TFA) salt. The TFA fluorine signal remained unaffected in the presence of *sd*Kef-QCTD (non-binder) (Figure 4.17). Similarly, in the assay mixture, 1,1,1 trifluoroacetone was used as an internal NMR control, and its signal also remained unaffected in the presence of *sd*Kef-QCTD (non-binder) (Figure 4.17).

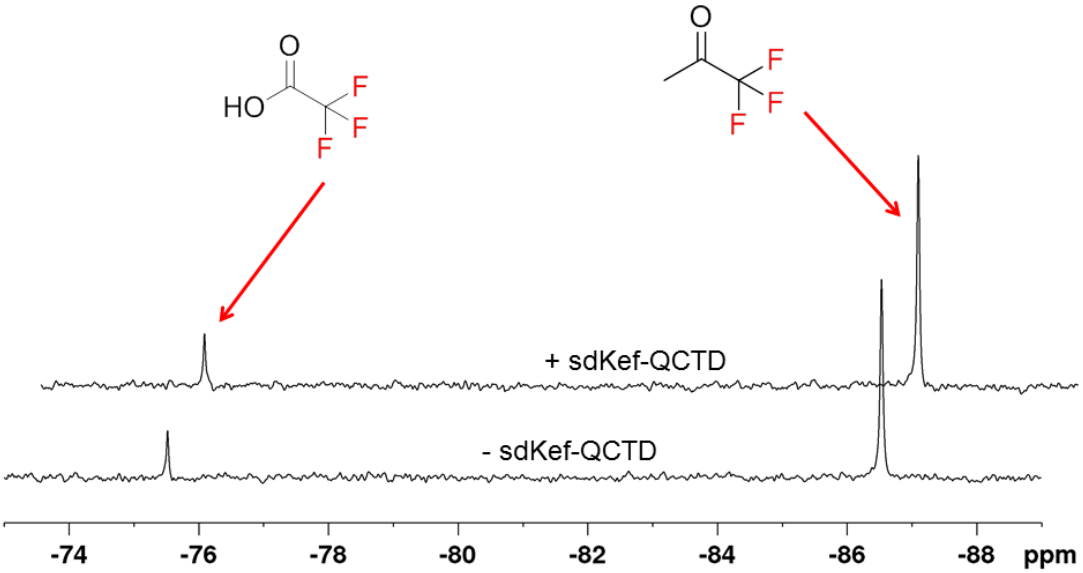


Figure 4.17 ^{19}F NMR spectra showing that the signals of the non-binders are not affected in the presence of *sd*KefC-QCTD. The concentration of 1,1,1 trifluoroacetone used in the assay was 20 µM. (note: The chemical shift of both these signals remain the same without and with protein).

To further confirm that the broad resonance in the spectrum (Figure 4.16) is due to the bound population of SG48D, we titrated *sd*Kef-QCTD to a fixed concentration of SG48D, and it was observed that the intensity of the broad resonance (B) increases as the *sd*Kef-QCTD concentration increases, whereas the intensity of the sharp resonance (F) decreases (Figure 4.18), implying that with increase in protein concentration, the population of the bound SG48 also increases, whereas that of the free SG48D decreases.

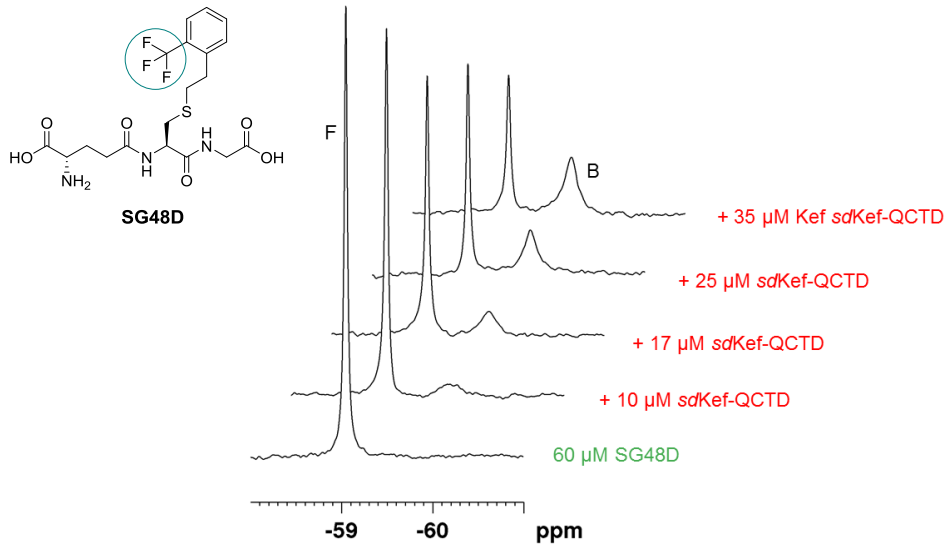


Figure 4.18 *sd*Kef-QCTD titration to SG48D showing the free and bound population signals of SG48D in ^{19}F NMR slow exchange system.
As *sd*Kef-QCTD concentration increases, the intensity of B (bound population of SG48D) increases and the intensity of F (free population of SG48D) decreases.

In a second experiment for proving the broad signal (Figure 4.16) as the signal of the bound population of SG48D, the thermal denaturation (heated at 100 °C for 5 minutes) of *sd*Kef-QCTD in the sample of SG48D + *sd*Kef-QCTD in buffer solution was carried out, and it was shown that denaturation of *sd*Kef-QCTD lead to the disappearance of the broad signal (B) and complete recovery of the sharp signal (F) in the spectrum (Figure 4.19).

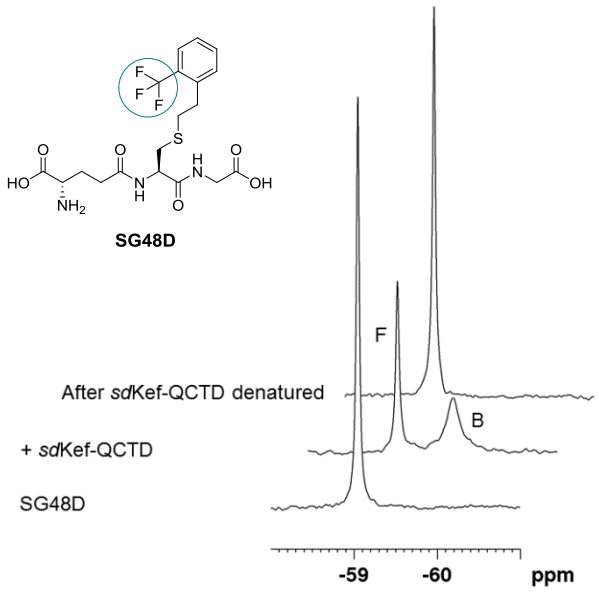


Figure 4.19 Control experiment for the slow exchange binding of SG48D with *sd*Kef-QCTD as observed by ^{19}F NMR. Thermal denaturation of *sd*Kef-QCTD in the sample of SG48D + *sd*Kef-QCTD results in the disappearance of signal B and increase in intensity of signal F, demonstrating that single B represents the bond population of SG48D in the mixture of SG48D + *sd*Kef-QCTD in buffer.

The binding interaction of SG48D to *sd*Kef-QCTD was also further validated by a reported differential scanning fluorimetric assay (thermal shift analysis performed by Dr Samuel C. Grayer)[24] which demonstrated that SG48D has a stabilizing effect when it binds to *sd*Kef-QCTD (Figure 4.20).

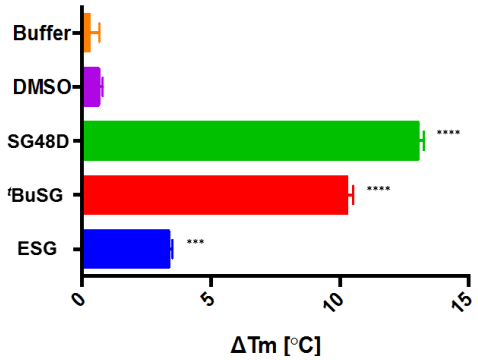


Figure 4.20 Differential scanning fluorimetry (DSF) analysis to determine whether SG48D has a stabilising effect on *sd*Kef-QCTD upon binding.
The graph shows changes in melting temperature in the presence of 1 mM ligand relative to that of *sd*Kef-QCTD + buffer or *sd*Kef-QCTD + DMSO (negative controls). ^{t}BuSG and ESG were run as positive control. The stabilizing effect of *sd*Kef-QCTD in the presence of SG48D suggests that SG48D binds to *sd*Kef-QCTD. Error bars indicate one standard deviation of uncertainty (n = 3). Significance of changes evaluated by a Student's *t*-test (where ****$p \leq 0.0001$, ***$p \leq 0.001$, **$p \leq 0.01$, *$p \leq 0.05$).

Finally, by the fluorescence competition assay (performed by Dr Samuel C. Grayer) as described in Figure 4.12, it was shown that SG48D binds specifically to *sd*Kef-QCTD in the GSH binding site (Figure 4.21).

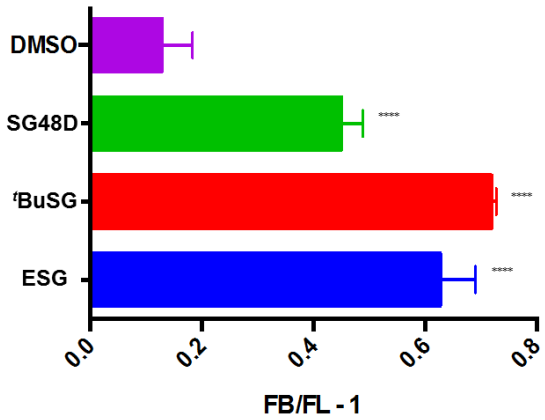


Figure 4.21 Fluorescence competition assay to determine whether SG48D binds to *sd*KefC-QCTD specifically in the GSH binding site.
The graph shows (FB/FL – 1) values of each compound (1 mM) relative to the positive controls ^{t}BuSG and ESG and negative control DMSO at 525.5 nm emission. The graph suggests that SG48D binds to sdKefC-QCTD with an affinity in close range to ^{t}BuSG and ESG . Error bars indicate one standard deviation of uncertainty (n = 3). Significance of changes evaluated by a Student's t-test (where ****p ≤ 0.0001, ***p ≤ 0.001, **p ≤ 0.01, *p ≤ 0.05).

From this ^{19}F NMR slow exchange binding of SG48D with *sd*Kef-QCTD (Figure 4.16), the binding dissociation constant (K_D) of SG48D was directly calculated according to eq. 1. The conditions chosen in the assay to measure the K_D of SG48D were 60 µM SG48D and 35 µM *sd*Kef-QCTD (Figure 4.16).

The basic expression for calculating a ligand K_D is shown below (eq. 1),

$$K_D = \frac{[P][L]}{[PL]} \qquad (1)$$

Or

$$K_D = \frac{[Kef][SG48\ D]}{[Kef.SG48D]}$$

As the values known from the spectrum (Figure 4.16) are,

[SG48D] = concentration of un-complxed SG48D at equilibrium = 27.07 µM

[Kef.SG48D] = concentration of Kef.SG48D complex at equilibrium = 32.93 µM

[Kef] = concentration of un-complxed Kef at equilibrium = 2.074 µM (from $[P]=[P_T]-[PL]$

Hence, by putting the values of [Kef], [SG48D] and [Kef.SG48D] in eq. 1, the K_D of SG48D was calculated as 1.71 µM. An average value of 1.36 ± 0.47 µM was obtained from three separate samples (Table 4.1).

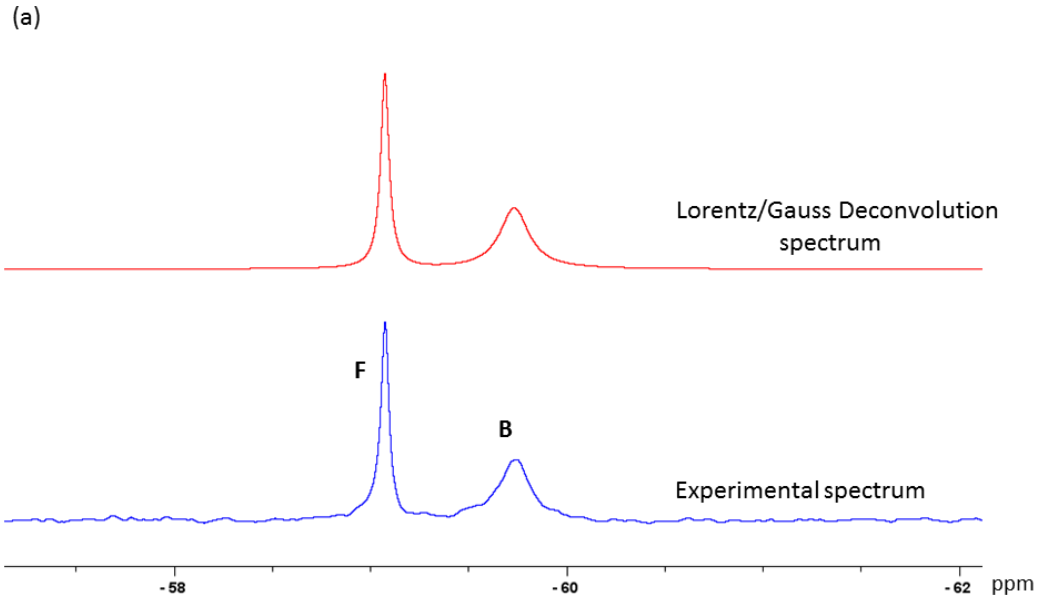
Table 4.1 Calculation of the ligand binding dissociation constant by ^{19}F NMR slow exchange ligand binding

$[P_T]$ = Protein concentration used in the assay = 35 µM

$[L_T]$ = Ligand concentration used in the assay = 60 µM

Sample No.	[L] / µM (from peak area of **F**)	[PL] / µM (from peak area of **B**)	[P] = $[P_T]$-[PL] / µM	$K_D = \frac{[P][L]}{[PL]}$ /µM
1	27.07	32.93	2.07	1.71
2	26.07	33.93	1.07	0.82
3	26.89	33.11	1.89	1.54
				Ave. K_D = 1.36 ± 0.47 µM

In the data analysis, instead by integration, the area of peaks F and B, were measured by Lorentzian line shape deconvolution (for theory and explanation see references [40,41]) (also called Line Fitting) method automatically through the Topspin (NMR data processing software), as this was observed to be more tolerant of variable noise levels and baseline distortion (Figure 4.22).



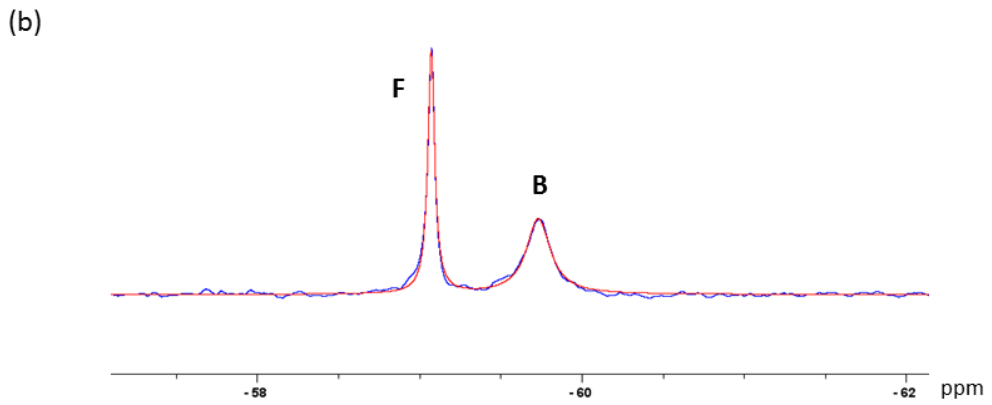
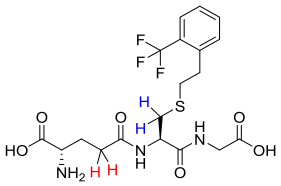


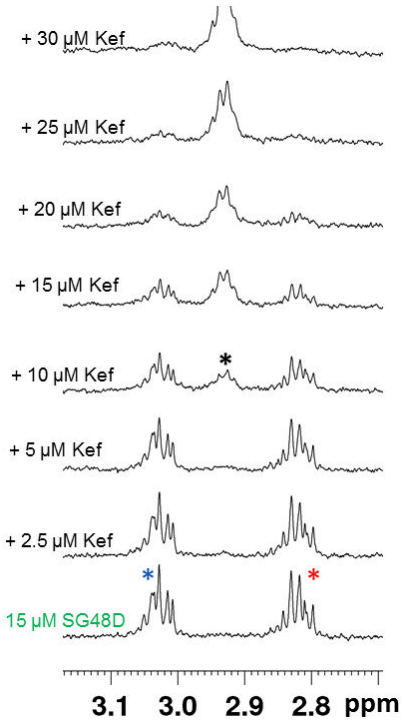
Figure 4.22 Representation of the Lorentz/Gauss Deconvolution applied to measure the peak areas of F and B in the slow exchange binding of SG48D with *sd*Kef-QCTD as observed by ^{19}F NMR.
(a) A comparison of the experimental spectrum vs the Lorenz/Gauss Deconvolution spectrum (b) Superimposition of the spectra in (a).

To confirm the validity of this K_D value of SG48D (1.36 ± 0.47 µM), we studied its binding interaction with *sd*Kef-QCTD by CPMG edited ^{1}H NMR direct ligand observation (Figure 4.23). We observed that in the presence of *sd*Kef-QCTD, the signals of SG48D are only attenuated without broadening (Figure 4.23a), demonstrating a strong binding interaction and a NMR slow exchange regime. By titration of *sd*Kef-QCTD, we measured the binding dissociation constant value of SG48D as 1.8 ± 0.3 µM (average of triplicates), which is in good agreement to its K_D value of 1.36 ± 0.47 µM determined, by its slow exchange binding observed in ^{19}F NMR spectrum. Hence, the NMR assays performed on both the nuclei i.e ^{19}F and ^{1}H of SG48D clearly reflects the same nature of its binding (tight binding) with *sd*Kef-QCTD. This thus confirms that the a slow exchange ligand binding observed by ^{19}F NMR provides the extra benefit of ligand binding calculation from single shot experiment.

After establishing the binding of SG48D with *sd*Kef-QCTD, we then assessed the binding interaction of the ^{t}BuSG to *sd*Kef-QCTD *via* NMR reporter displacement assay, using the bound population (signal B) of SG48D as a reporter system (Figure 4.24). It was observed that the addition of ^{t}BuSG to the assay mixture of SG48D + *sd*Kef-QCTD results a decrease in signal intensity of B and an increase in signal intensity of F (Figure 4.24). The decrease in intensity of B demonstrates that the population of bound SG48D has now decreased, because these protein molecules have now ^{t}BuSG bound instead of SG48D i.e SG48D is displaced by competitive ligand ^{t}BuSG. The increase in intensity of F demonstrates an increase in populations of free SG48D in solution, because of displacement from their binding pocket by ^{t}BuSG.



SG48D



Kef / µM	I_P	$\Delta I(\%) = (\frac{I_P - I_B}{I_B}) * 100$
0	1	0
2.5	0.88.31	-11.69
5	0.7199	-28
10	0.5114	-48.86
15	0.3368	-66.32
20	0.1731	-82.69
25	0.0907	-90.93
30	0.0343	-95.56

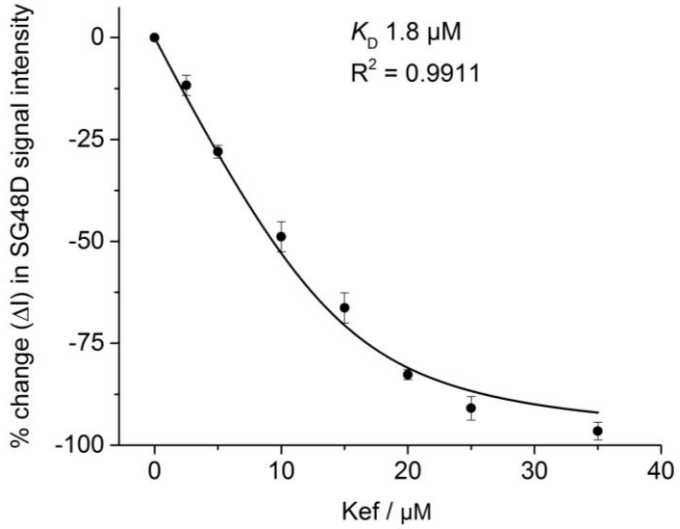


Figure 4.23 Measurement of the binding constant of SG48D with *sd*Kef-QCTD by CPMG edited ^{1}H NMR direct ligand observation.

Signals of SG48D blue and red protons are followed for binding. Left to right: to a fixed concentration of SG48D, *sd*Kef-QCTD was titrated and the SG48D (structure shown) C**H$_2$** protons (red/blue). *: impurity in buffer containing protein. Table showing the relative loss of the SG48D C**H2** signals peak intensity as a function of *sd*Kef-QCTD concentration. I_B represents SG48D signal intensity in the absence of protein; and is set to unit intensity. I_P represents its signal intensity in the presence of protein. Fitting of the data in (table) according to the quadratic K_D equation using OriginPro 9.0. The binding dissociation constant (K_D) of SG48D was calculated as 1.8 ± 0.3 µM. The error bars represents standard deviation from 3 separate measurements.

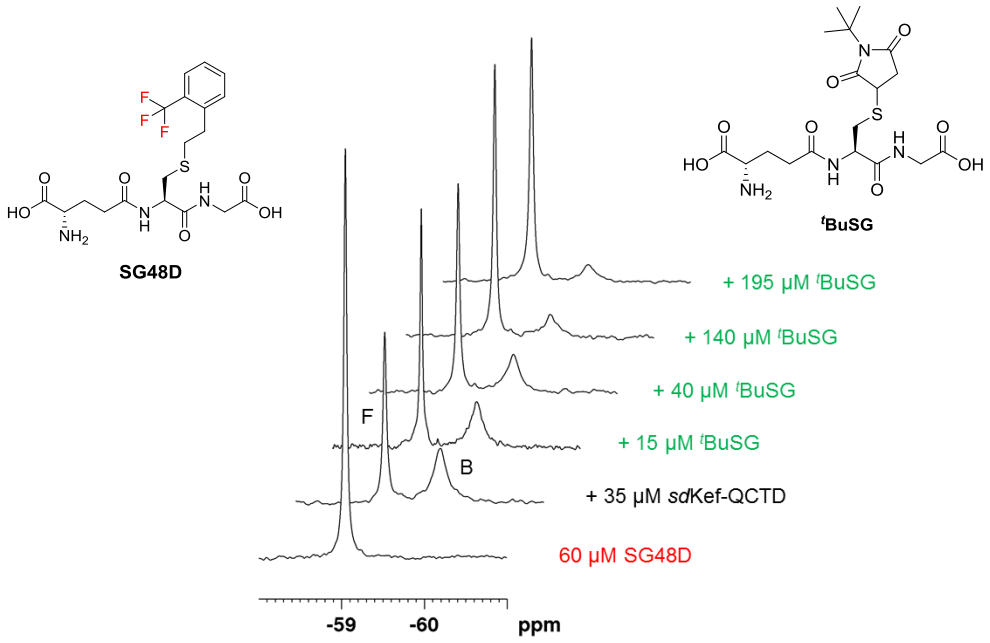


Figure 4.24 ^{19}F NMR competition experiment for ligand screening to *sd*Kef-QCTD using SG48D as a reporter.
F and B represents its free and bound (with *sd*Kef-QCTD) populations signals. After displacement of SG48D by ^{t}BuSG, the intensity of B decreases and the intensity of F increase. The displacement experiment is performed with low to high concentration of ^{t}BuSG.

Thus, after displacement of reporter ligand (SG48D), by the competitive ligand ^{t}BuSG, in conditions; 60 µM SG48D, 35 µM *sd*KefC-QCTD,

[SG48D] = concentration of free SG48D at equilibrium = 35.6 µM

[Kef.SG48D] = concentration of Kef.SG48D complex at equilibrium = 24.4 µM

[Kef] = concentration of Kef un-complexed with SG48D at equilibrium = 10.6 µM

After, putting these values in eq. 1, in the presence of competing ligand ^{t}BuSG, the apparent binding dissociation constant ($K_{D\ app}$) of SG48D was calculated as 15.46 µM (Table 4.2).

As now both the actual (1.36 µM) and apparent (15.46 µM) K_D values of SG48D are known and the concentration (C) of ^{t}BuSG used in the assay was 40 µM, the binding dissociation constant of the competitive ligand ($K_{D\ comp}$) ^{t}BuSG, was then calculated as 2.1 µM (average 2.07 ± 0.24 µM from three separate samples, Table 4.2) using eq. 2, which is close to its reported K_D values of 0.4 µM (emission spectroscopy) and 6.7 µM (isothermal calorimetric titration)[24] and 0.54 ± 0.03 by CPMG ^{1}H NMR assay.

Table 4.2 Calculation of the binding dissociation constant of competitive ligand ($K_{D\ comp}$) ^{t}BuSG with *sd*Kef-QCTD by ^{19}F NMR reporter (slow exchange binding) displacement assay.

$[P_T]$ = Protein concentration used in the assay = 35 µM

$[L_T]$ = Ligand concentration used in the assay = 60 µM

Sample No.	[L] / µM (from peak area of **F**)	[PL] / µM (from peak area of **B**)	[P] = $[P_T]$-[PL] / µM (protein not complexed with SG48D)	$K_{D\ app} = \frac{[P][L]}{[PL]}$ / (µM)	$K_{D\ comp} = \frac{[C]K_{D\ rep}}{K_{D\ app} - K_{D\ rep}}$ $K_{D\ rep}$ = 1.36 µM [C] = 40 µM
1	35.60	24.40	10.60	15.46	2.15
2	35.42	24.58	10.42	15.02	2.25
3	36.26	23.74	11.26	17.19	1.8
					Ave. K_D = 2.07 ± 0.24 / µM

The validity of this assay was further examined with other glutathione-S-conjugates as competitive ligands of *sd*Kef-QCTD (Figure 4.25) and the values of their binding constant obtained with this ^{19}F NMR reporter displacement assay were in close range to their reported literature values (Table 4.3).[24] The spectra and calculation of the K_D values obtained are shown in Appendix, Section 7.2.

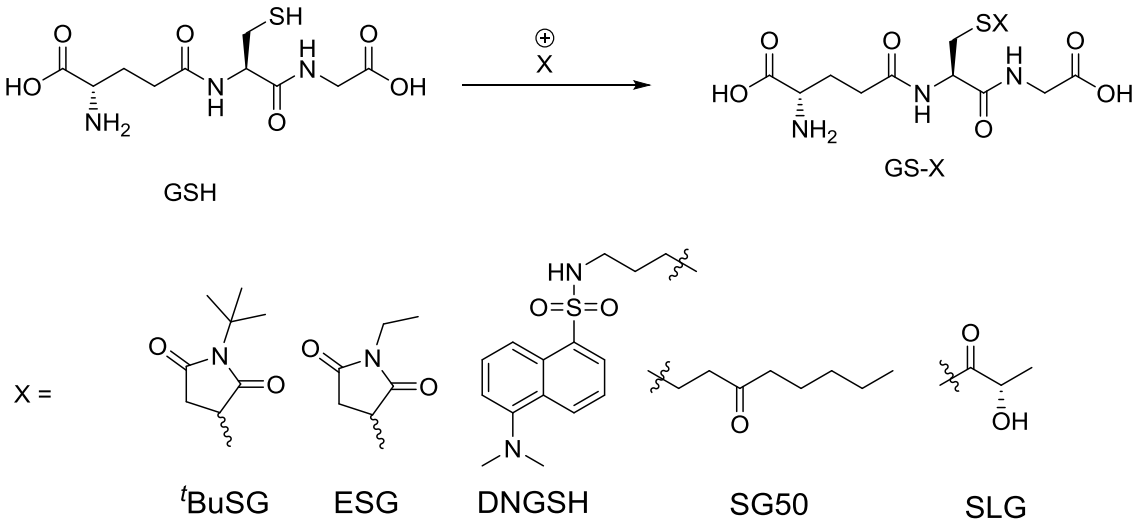


Figure 4.25 Structure of the ligands (GS-X) tested for their binding interaction with Kef by ^{19}F NMR reporter displacement assay.

Table 4.3 Comparisons of the K_D values of selected glutathione-*S*-conjugates to *sd*KefC-QCTD obtained by ^{19}F NMR reporter displacement assay with the reported literature values.[24] The ^{19}F NMR reporter displacement spectra and raw data for these ligands are shown in Appendix, section 7.2.

Ligand	K_D by ^{19}F NMR	K_D by other methods
GSH	881 ± 159 µM	900 ± 200 µM[c]
^{t}BuSG	2 ± 0.2 µM	6.7 ± 0.3 µM[b] 0.4 ± 0.2 µM[c] 0.54 ± 0.03 µM[a]
ESG	21 ± 1 µM	23 ± 4 µM[b] 12 ± 3 µM[c]
DNGSH	19 ± 1.5 µM	19 ± 6 µM[b] 6 ± 2 µM[c] 8 ± 2 µM[d]
SG50	23 ± 1.5 µM	4.4 ± 0.5 µM[c]
SLG	1176 ± 95 µM	900 ± 200 µM[c]

[a]: CPMG ^{1}H NMR, [b]: ITC, [c]: Fluorescence emission spectra, [d]: Fluorescence anisotropy

4.5 Probing conformational changes in *sd*Kef-QCTD using NMR slow exchange binding of a fluorinated ligand

Due to its high sensitivity of ^{19}F chemical shift to surrounding chemical environments, ^{19}F NMR spectroscopy is also employed for monitoring ligand binding events to proteins and the conformational/structural changes that occur in proteins as a result of ligand binding.[42-46] In this technique in the interested region of the protein, one or more residues are modified with a fluorine atom using various semi-synthetic and synthetic protocols (described in section 1.3.2). But this technique has two potential complications, which sometimes restrict its application in probing protein conformational changes due to ligand binding. (1) Expression

of site-specifically labelled protein is sometimes labour-intensive. (2) Upon ^{19}F-labelling, proteins sometimes lose stability (activity).

In contrast, in the ligand-based ^{19}F NMR slow exchange binding system, it is possible that the signal of the bound population of the ligand can be used to assess structural changes in the protein that may arise from the interaction of the protein with another ligand. From our knowledge of the reported literature, there are only a handful of examples of the ligand-based ^{19}F NMR slow exchange binding systems.[36-39] In these studies the signal of the bound populations of the ligand has been used as a probe to study the competitive binding of another ligand by reporter displacement assays. The ligand-based ^{19}F NMR slow exchange binding systems have received less attention for monitoring changes in the local environment of the protein due to binding events.

Herein, we employed the signal of the bound population of fluorinated ligand SG48D to probe structural changes in *sd*Kef-QCTD believed to be occurring when it interacts with the inhibitory and activating ligands. The binding of inhibitory ligand glutathione (GSH) to Kef has been proposed to close the K^+ ion channels, whereas, the binding of glutathione-S-conjugates (Kef activating ligands) such as ESG, is believed to open the K^+ ion channels.[6] GSH binds with *sd*Kef-QCTD weakly with a reported affinity of 900 ± 200 µM[24] and 881 ± 159 µM measured in this study by ^{19}F NMR. In monitoring the binding of GSH to *sd*Kef-QCTD by displacement of SG48D from its binding site, we observed that the addition of GSH to the solution of SG48D plus *sd*Kef-QCTD results in displacement of SG48D and simultaneously a change in the chemical shift of the residual SG48D-*sd*Kef-QCTD signal (Figure 4.26a,b). Titration of GSH showed that as more GSH binds with *sd*Kef-QCTD, the signal intensity of free SG48D (F) increases (without any change in its chemical shift) whereas, the intensity of signal of bound population of SG48D (B) decreases (due to displacement) with concomitant change in its chemical shift in NMR fast exchange regime indicating the weak binding of glutathione with *sd*Kef-QCTD. The GSH titrations made no influence on the chemical shift of the signal of the control molecule (1,1,1-trifluoroacetone: -86.53 ppm) (Figure 4.26c) in the assay mixture, implying that the changes in the chemical shift of the residual signal of the bond populations of SG48D due to GSH titration is not an artifact but perhaps is reflecting some structural changes in *sd*Kef-QCTD upon its interaction with GSH.

Figure 4.27 represents a possible explanation of these results. In the reported crystal structures of *E. coli* KefC-CTD in complex with GSH (PDB ID: 3L9W) and ESG (PDB ID: 3L9X), the protein is a homodimer with one molecule of GSH/ESG bound at the dimerization interface, but as each sub-unit has an AMP molecule and complexed to the ancillary protein Kef-F, it is also possible that there may be another GSH/ESG molecule bound at another site in the dimer, and which was not observed/refined in the crystal structures.

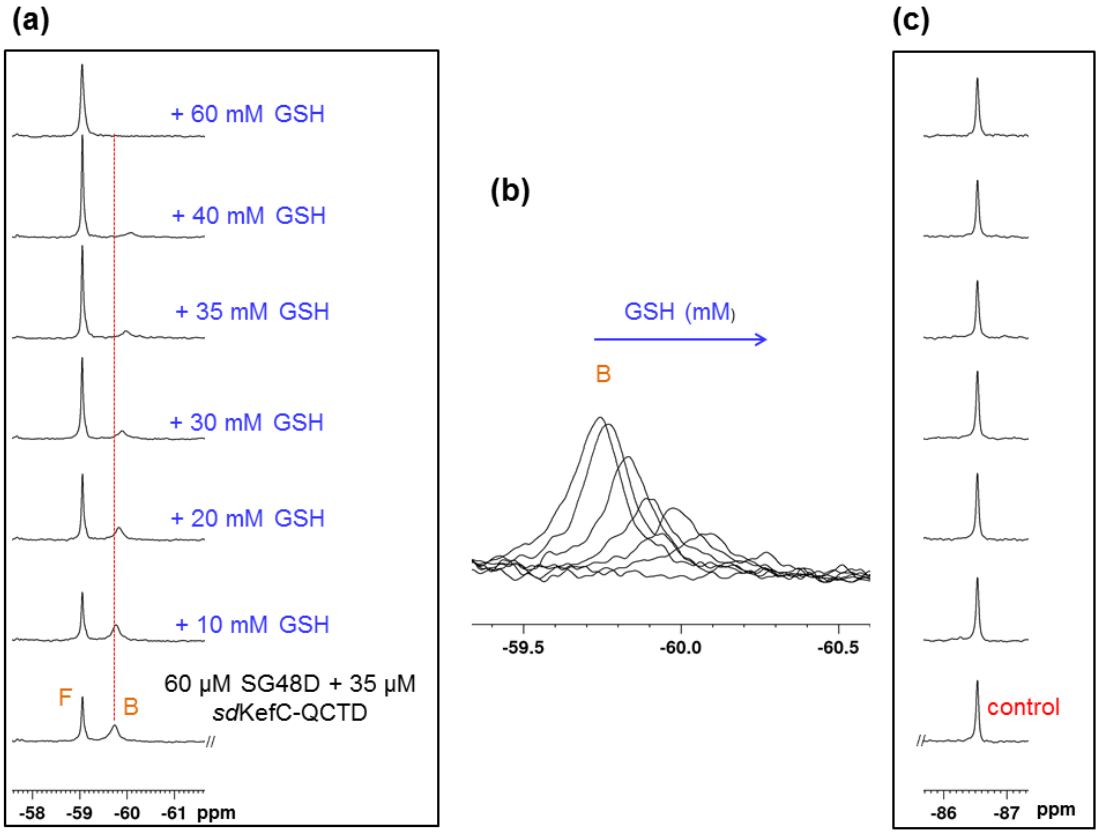


Figure 4.26 Monitoring changes in the local chemical environment of SG48D- *sd*Kef-QCTD complex due to GSH binding as demonstrated by ligand-based ^{19}F NMR assay.
(a) Binding of GSH to *sd*Kef-QCTD results in displacement of SG48D as indicated by decrease in bound population of SG48D (signal B) and increase in signal intensity of free SG48D (signal F) in solution. The GSH binding also appears to alter the local environment of (B) as demonstrated by changes in the chemical shift of the residual bound SG48D signal. (b) zoom in-view of signal B in (a). (c) Signal of 1,1,1 trifluoroacetone (internal NMR standard) during GSH titration in (a).

In case if the fluorinated ligand (SG48D, shown as red rectangle) is occupying both the binding sites of the dimer (Figure 4.27a,I), GSH (shown as blue rectangle) appears to competing (weakly in NMR fast exchange regime) with the fluorinated reporter (SG48D) at one site (Figure 4.27a,II). This is reflected as a decrease in the signal intensity of the bound population of SG48D (B) in the spectrum. The binding of GSH also results in structural changes in *sd*Kef-QCTD, which is reflected as changes in the chemical shift of the residual SG48D-*sd*Kef-QCTD (B) signal (SG48D bound at the second site) in the spectrum.

Although, we have not tested SG48D (a GS-X) to determine whether it activates the K^+ efflux or not, however based on its structure, it is possible that when it binds to *sd*Kef-QCTD, the protein may be adapting the open conformation, resulting in the activation of K^+ efflux. And the binding of GSH perhaps is changing the conformation of the protein from open form to the closed form in order to inhibit the K^+ efflux. These conformational changes in the protein possibly decreases the affinity of the SG48D in the second binding site, eventually leading to its release from its binding pocket (Figure 4.27a,III). During titration of GSH, a point is reached whereby both the binding sites of the *sd*Kef-QCTD are saturated with GSH with no fluorinated ligand bound. This is reflected as a complete disappearance of the signal of bound population of SG48D (B) and a full recovery of the signal of free population of SG48D (F) in the spectrum.

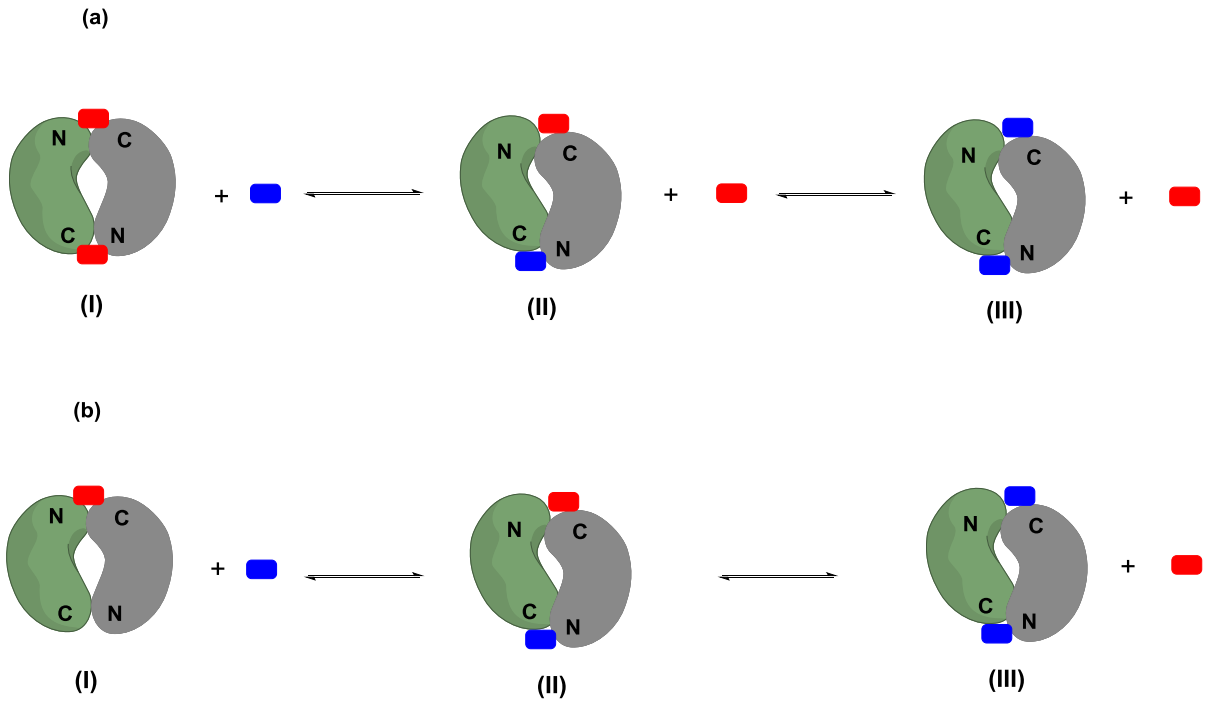


Figure 4.27 Cartoon representation of the proposed conformational changes that occur in *sd*Kef-QCTD due to glutathione binding.
(a) I: fluorinated ligand (SG48D, coloured red) is bound to both the binding sites of the *sd*Kef-QCTD dimer (green-grey object). II: GSH (coloured blue) competes (weakly) with SG48D at one binding site, resulting in its displacement. The binding of GSH also changes the conformation of *sd*Kef-QCTD possibly from the open form to the close form. III. Due to conformational changes in *sd*Kef-QCTD, the binding affinity of the SG48D at the second site is decreased, eventually leading to its release from the protein into solution. (b) I: SG48D (red) is occupying only one of the two binding sites of *sd*Kef-QCTD. II: GSH (blue) binds to the unoccupied site of *sd*Kef-QCTD dimer. The binding of GSH also changes the conformation of *sd*Kef-QCTD. III: The release of the SG48D from the second binding site as a result of conformational changes in the protein.

In case if the fluorinated ligand is occupying only one of the two binding sites of the dimer (Figure 4.27b,I), GSH appears to bind (weakly, NMR fast exchange regime) in the

unoccupied binding site (Figure 4.27b,II) and changes the conformation of *sd*Kef-QCTD possibly from open form to close form. Due to these conformational changes, the binding interaction of fluorinated ligand at the second site of *sd*Kef-QCTD dimer is weakened which then results in its release from its binding pocket into solution. The release of SG48D from the *sd*Kef-QCTD into solution is observed as an increase in the intensity of signal F, whereas, the conformational changes in the *sd*Kef-QCTD are reflected as changes in the chemical shift of the residual signal B. During GSH titration a point is reached whereby both the binding sites of the *sd*Kef-QCTD dimer are fully occupied by GSH (Figure 4.27b,III) (signal B disappears completely disappears).

Interestingly, by titrating ^{t}BuSG, a strong binder (K_D 0.400 µM,[24] and 0.54 ± 0.03 µM measured in this study by ^{19}F NMR) and activator of *sd*Kef-QCTD, we observed only the displacement of the SG48D (fluorinated reporter) without any changes in the chemical shift of residual signal of SG48D-*sd*Ked-QCTD (signal B) (Figure 4.28) unlike GSH. This supports the theory that with SG48D (fluorinated ligand) bound, *sd*Kef-QCTD is likely in the open conformation, that results in the activation of the K^+ efflux.

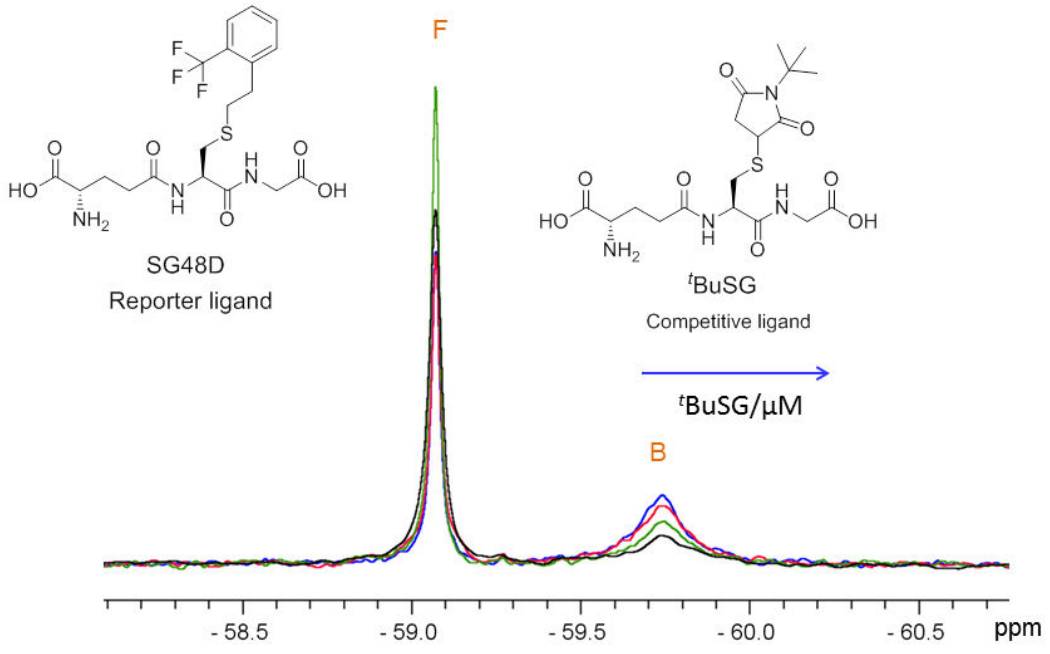


Figure 4.28 Displacement of SG48D by ^{t}BuSG as shown by ^{19}F NMR.
Titration of competitive ligand ^{t}BuSG to the NMR slow exchange binding SG48D to *sd*KefC-QCTD. The binding of ^{t}BuSG results in decrease of intensity of signal B and increase in intensity of signal F, with no change in the chemical shift (local environment) of residual B. The concentration of SG48D was 60 µM and *sd*KefC-QCTD 35 µM. The concentration of ^{t}BuSG used were 15 µM, 75 µM and 140 µM.

In this case if the reporter (red) is occupying both the binding sites (Figure 4.29a,I) of the *sd*Kef-QCTD dimer, it appears that ^{t}BuSG (blue) is simply competing with the reporter ligand at both the binding sites. As with both SG48D and ^{t}BuSG, *sd*Kef-QCTD is in one single conformation likely the open form, no changes in the chemical shift of residual signal of bound population of SG48D (B) are observed (Figure 4.29a,II). Similarly, if SG48D (red) is occupying only one of the two binding sites (Figure 4.29b,I) of the dimer, then in that case, the binding of ^{t}BuSG in the unoccupied site (Figure 4.29b,II) doesn't affect the bound population signal of SG48D (B), whereas its competition with the SG48D simply results in its displacement (Figure 4.29b,III) as demonstrated by decrease in intensity of signal B.

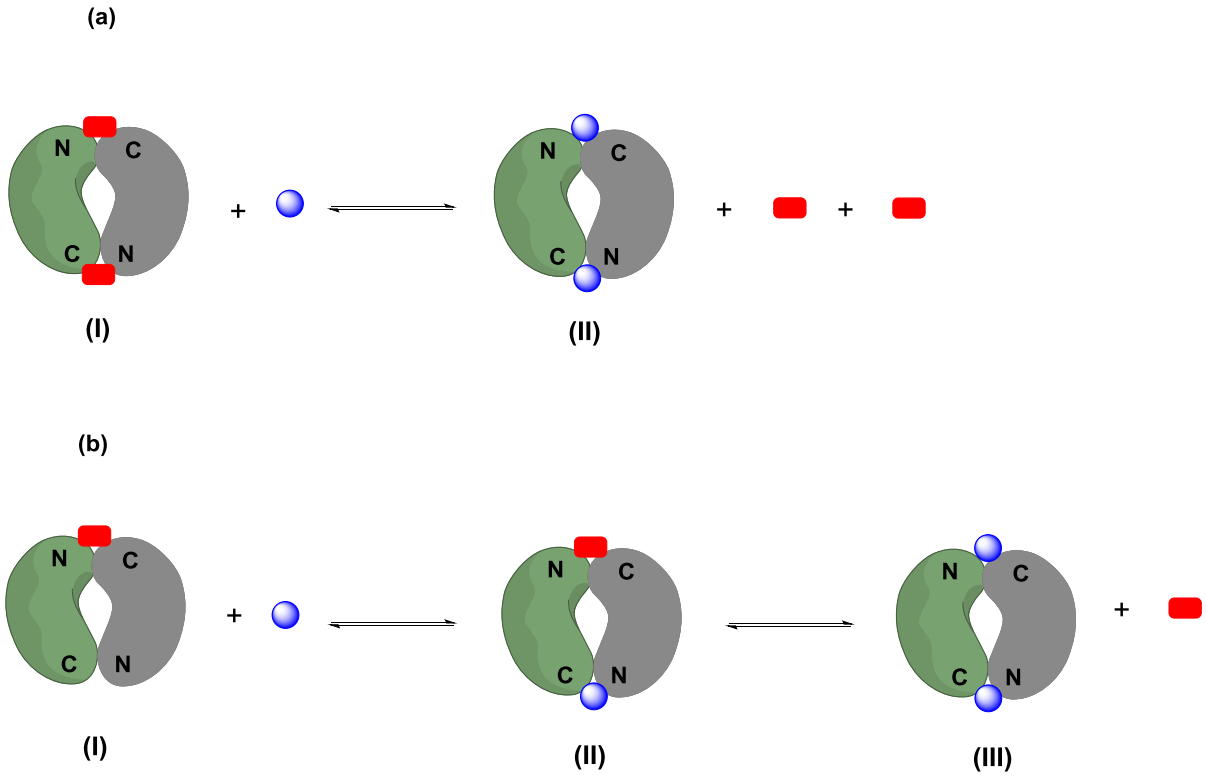


Figure 4.29 Cartoon representation of the binding of ^{t}BuSG with *sd*Kef-QCTD by ^{19}F NMR reporter ligand displacement assay.
(a) I: fluorinated reporter SG48D (red) occupies both the binding sites of the dimer *sd*Kef-QCTD. II: competitive ligand ^{t}BuSG (blue) competes with SG48D in both the binding sites. (b) I: SG48D (red) occupies only one of the two binding sites of the *sd*Kef-QCTD dimer. II: binding of ^{t}BuSG in the unoccupied site does not affect SG48D. III: ^{t}BuSG competes with SG48D at the second binding site.

4.6 Investigation of the identity of the nucleotide ligand in *sd*Kef-QCTD

The central feature of the cytosolic domains RCK or KTN, involved in bacterial Kef systems is the presence of a characteristic and highly conserved Rossmann fold in their structure.

Throughout biological systems, the Rossmann folds are associated with the binding of nucleotide ligand, especially NADH in dehydrogenases, where they were first reported.[47-50] Amongst the total 17 crystal structures reported for KTN domains, 5 contain NADH bound, while the remaining contain nucleotide molecules including AMP, ADP or ATP. In the previously reported *apo*-crystal structure *of E. coli* KefC-CTD in complex with ancillary protein Kef-F (PDB ID: 3EYW), the electron density in the nucleotide binding site was assigned as a NAD^+ because of the presence of this nucleotide ligand in the crystallization mixture.[10] The possibility of this electron density as an AMP molecule was ruled out on the basis that its nicotinamide ring is highly mobile for making discrete electron density. However, upon reanalysing the structure by the same group, it was proposed that the binding of NAD^+ to the protein was not feasible, because the second phosphate group could not be built into the structure without causing substantial steric clashes and energetically unfavourable interaction in the protein. In the structure of other KTN domains with bound NAD(H), the positioning of the γ-glutamate carboxyl group of GSH and D472, which is at the similar position to a more accommodating glycine residue, prevents the binding of nucleotide larger than AMP[14,16] as observed in the crystal structure of *E. coli* KefC (PDB ID: 3L9W). Therefore, the electron density was assigned as AMP molecule. But the question was where this AMP is coming from; is it originating from the decomposition of NAD^+ or it is a contaminant in the NAD^+ stock solution, that competes with NAD^+ for binding in the nucleotide binding site. This ambiguity about the identity of this nucleotide ligand raises important questions about its role in the biological function of Kef proteins.

In exploring the identity of this nucleotide ligand and its potential role in the biology of KTN domains, we solved a crystal structure for the *apo-sd*Kef-QCTD and identified electron density in both the nucleotide binding pockets (of the dimer) (Figure 4.30). This electron density was best fitted and refined to be an AMP molecule. We then carried out additional studies using techniques other than X-ray crystallography to support our results that the ligand bound in the nucleotide binding pocket is indeed an AMP molecule, and to support our theory that AMP has a structural role in the KTN domains i.e stabilising the dimeric form of Kef protein.

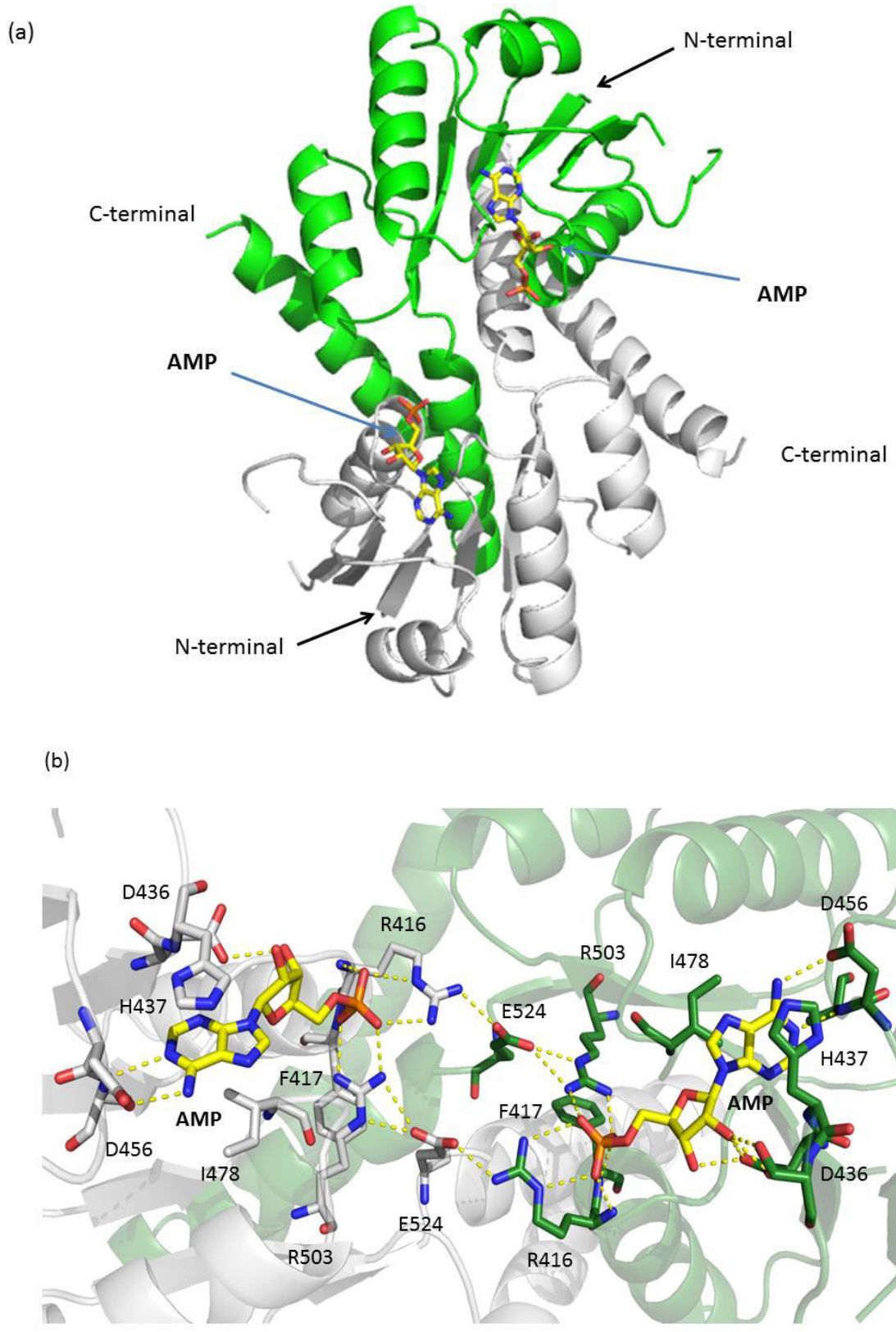


Figure 4.30 *apo*-crystal structure of *sd*Kef-QCTD.
(a) *sd*Kef-QCTD dimer; the two monomers are shown as green and grey; each with an AMP molecule (yellow sticks) in the nucleotide binding pocket. (b) Zoom in view of the AMP binding site in the two subunits (green and grey). The bridging E524 connects the two subunits with each other. The AMP phosphate oxygens make interaction with R503 and R416, both of which directly interacts with the bridging glutamates E524, thus increases the stability of the dimeric *sd*Kef-QCTD.

Several key interactions can explain the binding of AMP in the nucleotide binding pocket (Figure 4.30). The adenosine part of the AMP is located in a pocket formed by residues I478 and H437. H437 (3.9-4.0 Å) appears to form π-π stacking interactions with the purine rings. The amino group of the purine (aniline like) forms hydrogen bonding interaction with D456. The two –OH groups of ribose of AMP are positioned close to D436, forming hydrogen bonding interactions with this residue. Most of the AMP interactions with the protein are from the AMP phosphate group, which builds salt bridges to R416 and R503, and hydrogen bonds to the backbone NH of R416 and F417. A comparison of the X-ray crystal structure of the *E. coli* KefC-CTD (PDB ID: 3L9W) with that of *sd*Kef-QCTD (Figure 4.31) reveals that the key residues involved in the binding of AMP are conserved between these two proteins. In the *E. coli* KefC-CTD, an additional interaction between H434 and the AMP ribose oxygen is visible, however, in the *sd*KefQCTD crystal structure; the density for this residue is not well resolved.

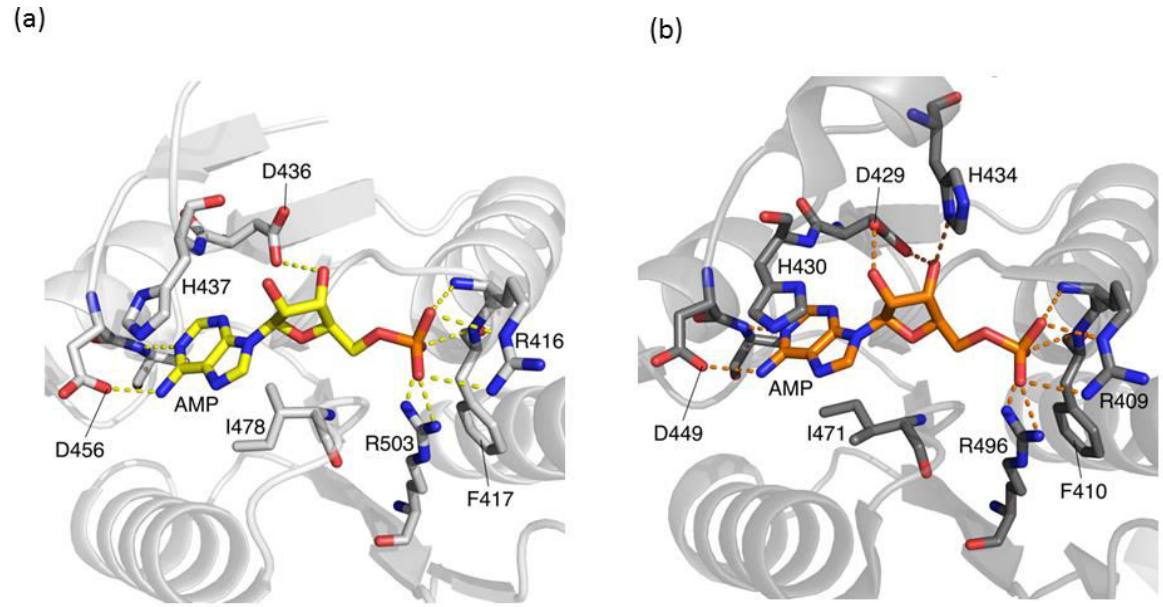


Figure 4.31 Comparison of the AMP binding pocket in the crystal structure of (a) *sd*Kef-QCTD and (b) *E. coli* KefC-CTD (PDB ID: 3L9W)

4.6.1 Confirmation of nucleotide ligand as AMP by ^{1}H NMR spectroscopy

We conducted CPMG edited ^{1}H NMR experiments both on the native and thermally denatured (by heating for 3 h at 80 °C) *sd*Kef-QCTD samples. The native *sd*Kef-QCTD spectrum showed no resonances corresponding to either NADH, AMP and ADP, suggesting

that any nucleotide present was either bound to the protein or there may be no nucleotide natively bound to the protein at all (Figure 4.32a). However, in the denatured *sd*Kef-QCTD spectrum sharp signals were observed reporting the presence of a free ligand in solution (Figure 4.32b). Comparison of the spectrum of the denatured *sd*Kef-QCTD to that of the spectrum of reference AMP (Figure 4.32c) suggests that the free ligand that was released in solution from denaturation of *sd*Kef-QCTD was AMP (Table 4.4). These results suggest that AMP was natively bound to the protein.

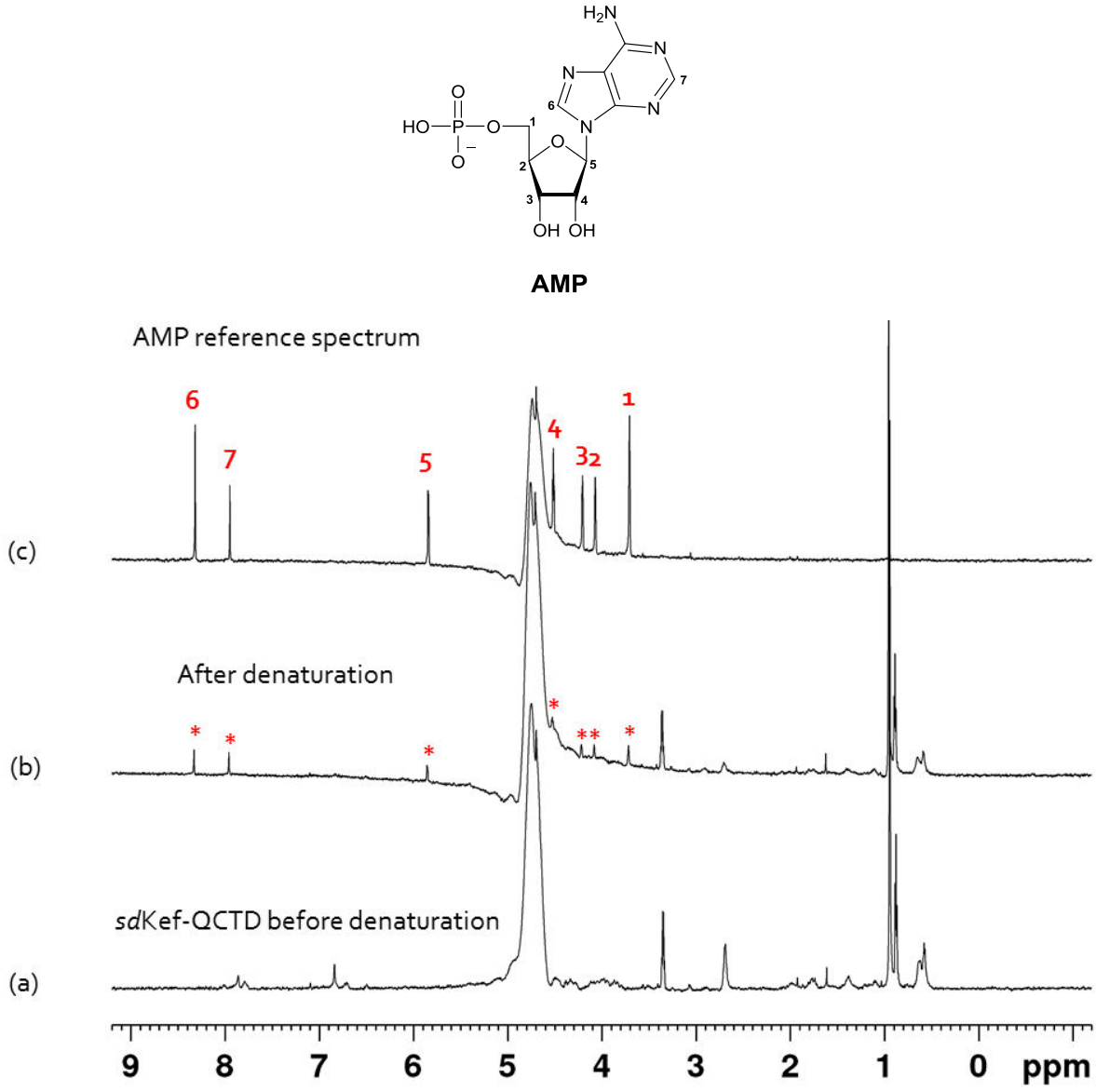
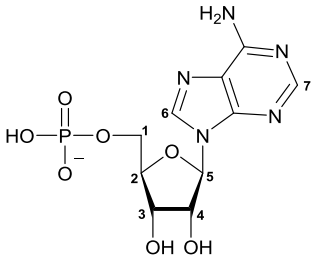


Figure 4.32 ^{1}H NMR conformation of the nucleotide ligand as AMP in the *apo*-crystal structure of sdKef-QCTD.
Overlay of CPMG edited 1H NMR spectra of the native and denatured sdKef-QCTD protein with reference AMP spectrum. (a) In the spectrum of native sdKef-QCTD (331 µM), the broad resonances of the protein are removed by the CPMG filter. Though, still some sharp peaks are seen, however these are most likely from small molecules impurities containing in the protein solution. (b) The red stars denote the signals of the free AMP that releases into solution as a result of thermal denaturation of sdKef-QCTD, suggesting that AMP was natively bound to sdKef-QCTD. (c) Reference spectrum of AMP. The total echo time used in these experiments were 96 ms. The assay solutions were prepared in 50 mM sodium phosphate buffer, pH 7.4, in D2O. The spectra were recorded at 278 K.



AMP

Table 4.4 ^{1}H chemical shift comparison of the peaks observed in the denatured protein spectrum (Figure 4.32b) compared to the AMP reference spectrum (Figure 4.32c).

Peak No.	Detected compound ^{1}H chemical shifts	Standard AMP ^{1}H chemical shifts
1	3.72 ppm	3.71 ppm
2	4.09 ppm	4.07 ppm
3	4.22 ppm	4.20 ppm
4	4.53 ppm	4.51 ppm
5	5.85 ppm, d	5.84 ppm, d
6	8.33 ppm, s	8.32 ppm, s
7	7.96 ppm, s	7.95 ppm, s

To further confirm that AMP is the ligand that release into solution as a result of *sd*Kef-QCTD denaturation, external AMP was added to the sample of thermally denatured *sd*Kef-QCTD (Figure 4.33b). In the spectrum, it was observed that the addition of external AMP increases the intensity of the AMP (signals) released in solution as a result of thermal denaturation of *sd*Kef-QCTD (Figure 4.33c).

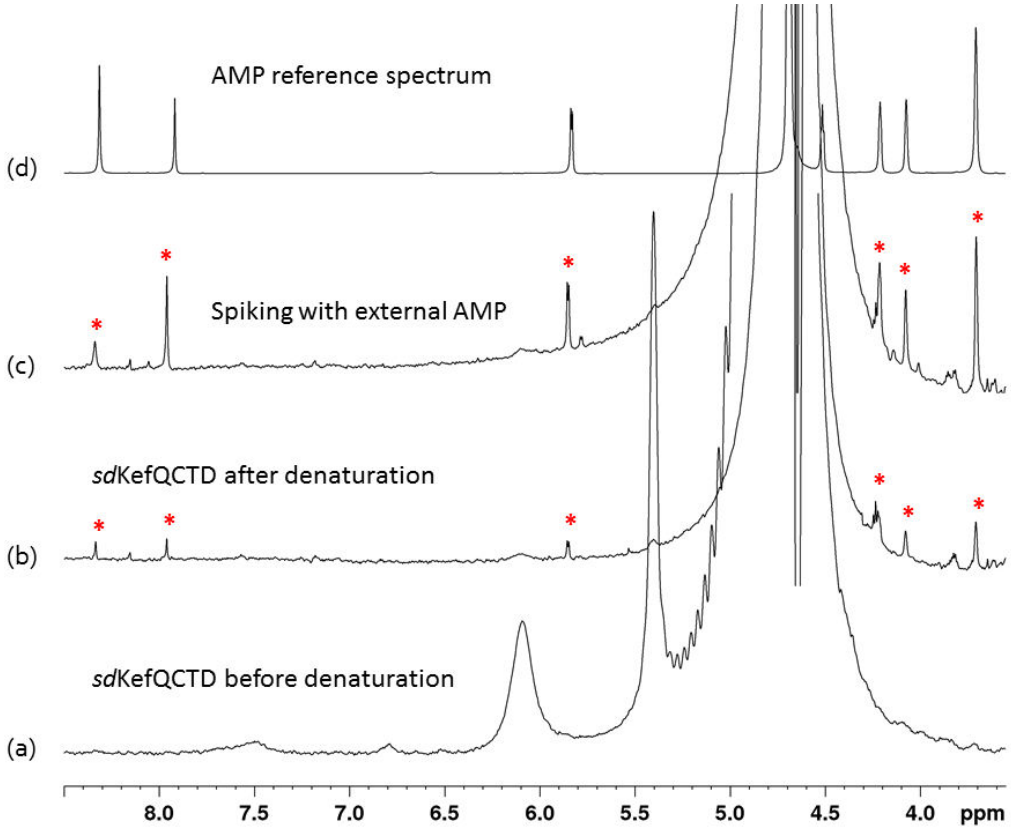


Figure 4.33 NMR spiking technique to further confirm that the ligand released as a result of thermal denaturation of *sd*Kef-QCTD was AMP.
Addition of external AMP to the solution of thermally denatured *sd*Kef-QCTD. Overlay of CPMG edited ^{1}H NMR spectra of the (a) Native *sd*Kef-QCTD (b) Thermally denatured *sd*Kef-QCTD (5 µM). Signals with red asterisk represent AMP that was natively bound to *sd*Kef-QCTD, which upon thermal denaturation of the protein released into solution. (c) After the addition of external AMP (15 µM) to sample of b. (d) Reference AMP spectrum. In spectrum (c), the intensity of the AMP signals is increased upon the addition of external AMP, further confirming that the ligand release after thermal denaturation of *sd*Kef-QCTD is the natively bound AMP. The total echo time used in these experiments were 96 ms. The assay solutions were prepared in 50 mM sodium phosphate buffer, pH 7.4, in D_2O. The spectra were recorded at 278 K.

To confirm that AMP was not produced by the hydrolysis of NADH or ADP upon heat treatment, we exposed NADH and ADP to the same conditions used to denature the *sd*Kef-QCTD (heating for 3 h at 80 °C). The results of these experiments are highlighted in Figure 4.34 & Figure 4.35, respectively. Although in both cases, new peaks were observed after heat treatment, however, AMP did not appear to be formed.

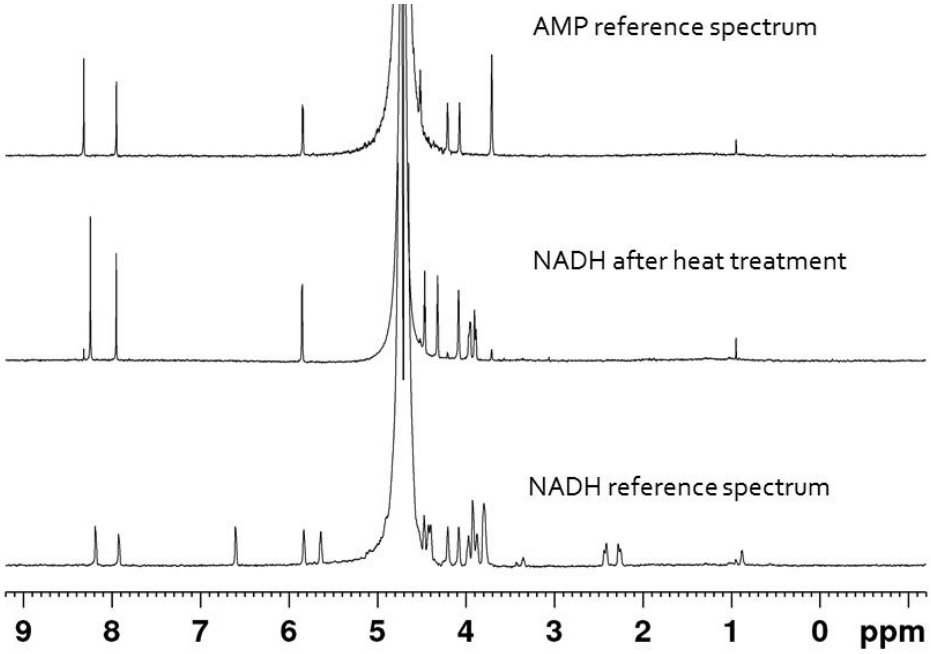


Figure 4.34 Experiment to prove that AMP was not formed from NADH thermal degradation.
Overlay of AMP reference ^{1}H NMR spectrum with NADH before and after being heated at 80 °C for 3 hours.

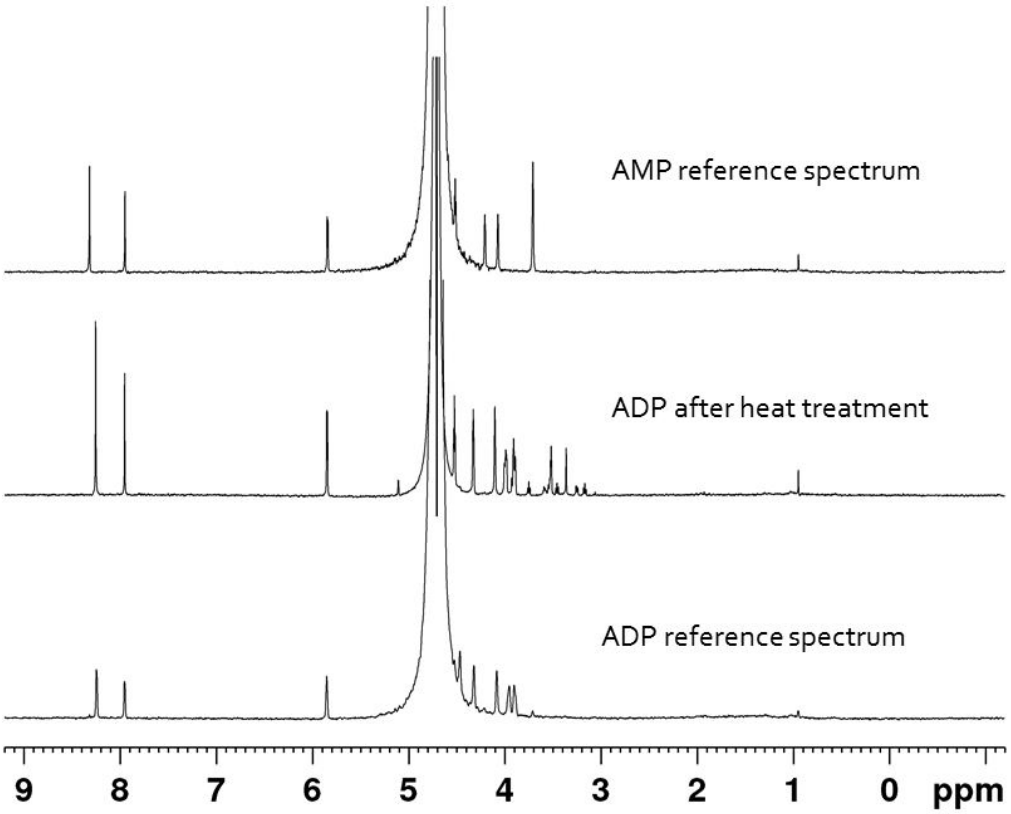


Figure 4.35 Experiment to prove that AMP was not formed from ADP thermal degradation.
Overlay of AMP reference ^{1}H NMR spectrum with ADP before and after being heated at 80 °C for 3 hours.

AMP was also subjected to the same heat treatment as *sd*Kef-QCTD and was found to be stable under these conditions (Figure 4.36). This was further confirmed by overlaying the AMP spectra with that of an (-)Adenosine reference spectrum i.e AMP was not degrading to (-)adenosine under denaturation conditions (Figure 4.37).

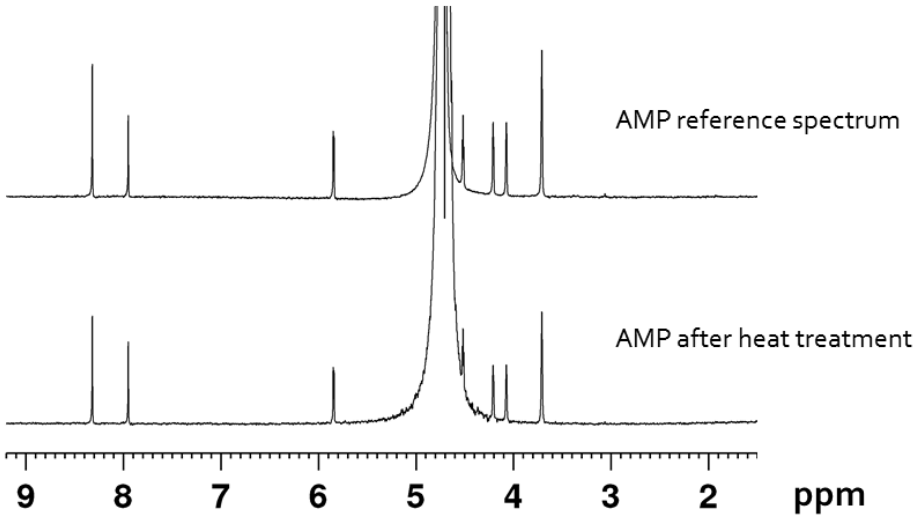


Figure 4.36 Experiment to prove that AMP remains stable under thermal denaturation conditions used. Overlay of AMP reference spectrum with AMP after being heated at 80 °C for 3 hours.

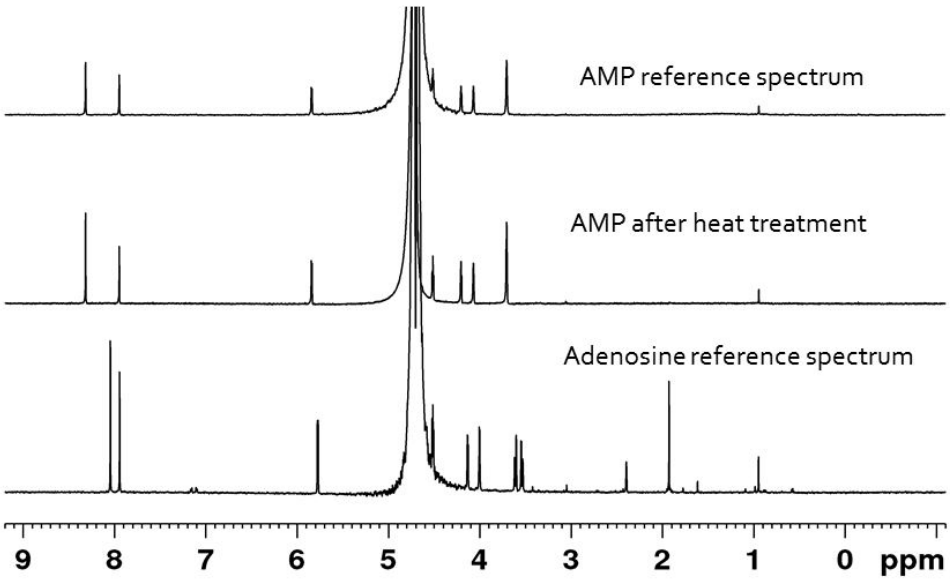


Figure 4.37 Experiment to prove that AMP was not degrading to adenosine under thermal denaturation conditions used. Overlay of AMP reference spectrum with AMP after being heated at 80 °C for 3 hours and an adenosine reference spectrum.

Thus our ^{1}H NMR based studies of *sd*Kef-QCTD clearly demonstrated that the electron density observed in the nucleotide binding pocket of *sd*Kef-QCTD crystal structure is unambiguously an AMP molecule that is naturally bound to the *sd*Kef-QCTD.

4.7 Role of AMP binding in *sd*Kef-QCTD

Several roles have so far been proposed about the binding of nucleotides in the Kef domains, for example it may be stabilizing the dimeric form of Kef proteins (domains),[6] or it may be involved in regulating the K^+ flow, or its binding may be inducing the required conformational changes in the KTN octamers, necessary for gating. In the two reported crystal structures of *E.coli* KefC,[6] the structure in complex with inhibitory ligand glutathione (PDB ID: 3L9W) shows an AMP molecule in the nucleotide binding site, whereas, the structure in complex with the activating ligand ESG (PDB ID: 3L9X), shows a sulphate ion likely from the crystallisation mixture instead of AMP. In the former structure, it is believed that the binding of AMP may perhaps be stabilizing the closed conformation of KefC-CTD, whereas, in the KefC-CTD-ESG structure, ESG may have displaced AMP somehow, in order for the cannels to open. There are however, no other experimental evidences to support this theory.

To explore the theory that AMP binding stabilizes the dimeric form of *sd*Kef-QCTD, we investigated the stability of *sd*Kef-QCTD by using differential scanning fluorimetry (DSF), mutagenesis studies and HPLC analysis.

4.7.1 Differential Scanning Fluorimetric studies

Differential Scanning Fluorimetry (DSF) is a technique commonly used to study the stability of a protein in its native form. It does this by measuring the heat changes associated with the molecule's thermal denaturation when heated at a constant rate. Macromolecules such as proteins can form folded structures that undergo thermally-induced conformational changes. These structural rearrangements result in the absorption of heat caused by the redistribution of non-covalent bonds. Differential scanning fluorimeters measure this heat uptake.

In this assay (Figure 4.38) the protein is heated in the presence of a fluorescent dye (SYPRO Orange), which binds non-specifically to hydrophobic surfaces. Water strongly quenches its fluorescence. When the protein unfolds upon heat treatment, the dye non-specifically binds to the exposed hydrophobic surfaces, resulting in an increase in fluorescence by excluding water. The stability curve and its midpoint value (melting temperature, T_m also known as the

temperature of hydrophobic exposure, T_h) are obtained by gradually increasing the temperature to unfold the protein and measuring the fluorescence at each point. Curves are measured for protein only and protein + ligand, and ΔT_m is calculated.

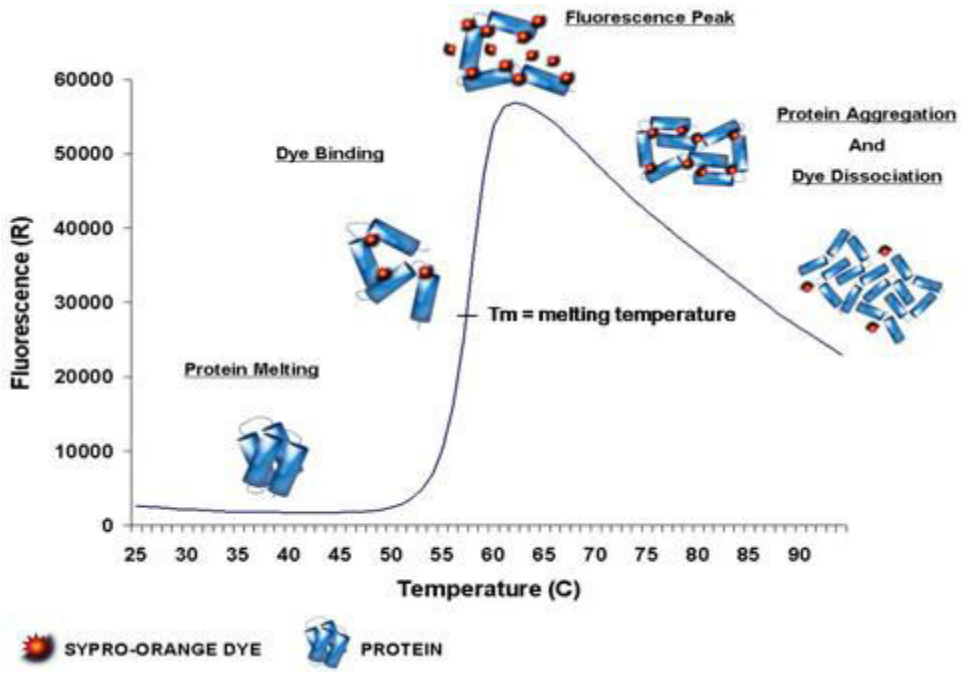


Figure 4.38 Cartoon representation of differential scanning fluorimetry (DSF) technique.

A measurement of ΔT_m is also often considered as a qualitative assessment of the binding affinity of a ligand to a protein. A high ΔT_m indicate high affinity ligand binding and *vice versa*. A ΔT_m = 15 ^oC was obtained for AMP, 7 ^oC for ADP, while ATP and adenosine had no effect on T_m. Similarly, NADH showed ΔT_m = 3 ^oC, whereas NADP and NAD^+ had little effect on T_m (Figure 4.39a). These results demonstrate that AMP gives more stabilizing effect to *sd*Kef-QCTD compared with the other nucleotides. It was further shown that binding of activating ligand ESG was additive with AMP. In the presence of ESG, a ΔT_m = 18 ^oC was observed for AMP (Figure 4.39b) suggesting that there are distinct nucleotide and peptide binding sites in the protein. The fact that adenosine does not bind (stabilizes) *sd*Kef-QCTD suggests that the AMP binding to *sd*Kef-QCTD is predominantly due to its phosphate group in agreement with the crystallographic results.

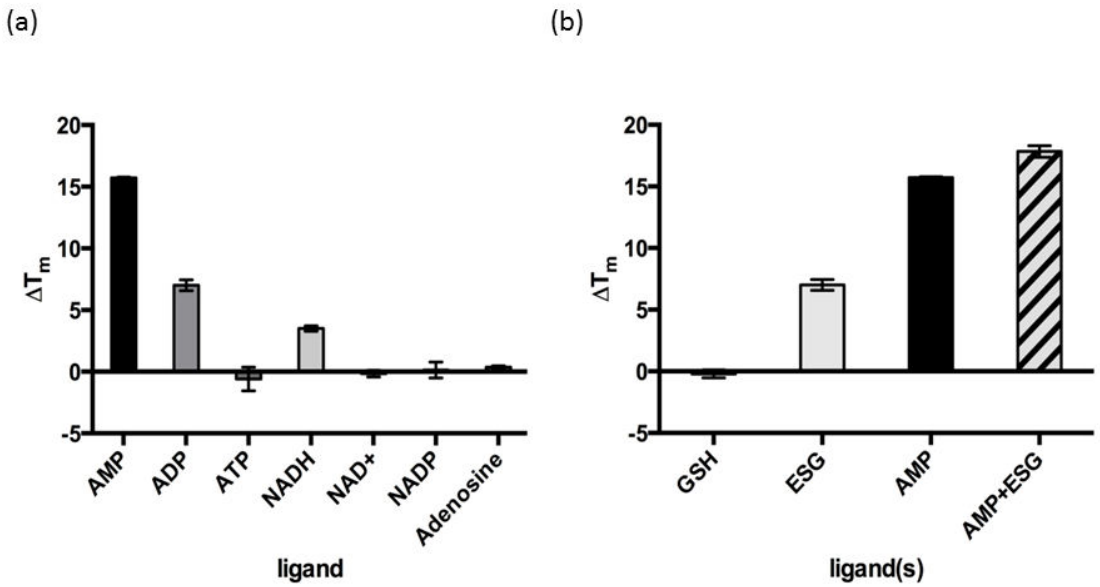


Figure 4.39 DSF experiments to determine the stabilising effect of nucleotide binding on *sd*Kef-QCTD.
(a) It is shown that AMP is most effective at stabilizing *sd*Kef-QCTD with ΔT_m = 15 °C. (b) DSF experiments to determine the effect of both AMP and ESG on the stabilization of *sd*Kef-QCTD. Negligible stabilization is provided by GSH due to its weak binding, whereas ESG shows ΔT_m = 7 °C. The effect of AMP and ESG binding simultaneously is additive. (Image provided by Dr Samuel C. Grayer)

4.7.2 Mutagenesis studies

To provide additional support to the DSF results, that AMP gives extraordinary stabilizing effect to *sd*Kef-QCTD, mutations studies of the key residues involved in the binding of AMP (Figure 4.40) was performed.

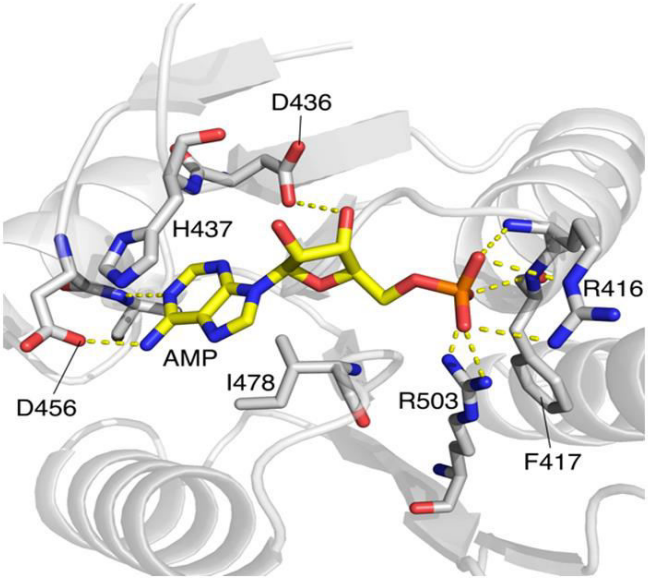


Figure 4.40 View of the AMP binding site in the *apo*-crystal structure of *sd*KefC-QCTD.

The first residue mutated was H437, which forms the top of the AMP binding pocket, located approximately 3.5 Å close to the purine ring of AMP. It is possible that the π-π stacking between two aromatic rings could be involved in the binding of AMP to *sd*Kef-QCTD (Figure 4.40), however, we found that H437A and H437N mutations have only very little effect on the expression of both mutated *sd*Kef-QCTD protein, suggesting that both these mutants *sd*Kef-QCTD proteins are likely possessing the bound AMP molecule necessary for protein stabilization (dimerization). Both H437A and H437N *sd*Kef-QCTD were also proved active in the K^+ efflux assay (performed by Silvia Ekkerman). H437N *sd*Kef-QCTD showed activity similar to wild type *sd*Kef-QCTD, whereas H437A *sd*Kef-QCTD displayed 50% of wild type activity. These results demonstrate that the AMP purine ring does not appear to play a critical role in its binding with *sd*Kef-QCTD protein.

The influence of D436 on the binding of AMP was then assessed. D436 forms hydrogen bonding interactions with the two hydroxyl groups on the ribose ring of AMP (Figure 4.40). It was found that D436E mutant *sd*Kef-QCTD, which was expected to form similar interactions with AMP as D436 *sd*Kef-QCTD, also resulted in a wild type expression levels, suggesting that mutant D436E *sd*Kef-QCTD perhaps also have bound AMP molecule required for protein stabilizing. In contrast, D436A and D436N resulted in significantly reduced level of protein expression, possibly due to the lack of bound AMP molecule. In the K^+ efflux assay, the activity of D436E was also not substantially different from the wild type, whereas for both mutant D436A and D436N *sd*Kef-QCTD, a significant reduction in their activity of K^+ efflux activation was observed consistent with the reduced level expression of these two mutants *sd*Kef-QCTD. These results suggest that D436 plays a crucial role in the binding of AMP, hence in the stability (expression) of *sd*Kef-QCTD.

After H437 and D436, the effect of R416 on AMP binding was then investigated. According to the X-ray crystal structure of *sd*Kef-QCTD (Figure 4.40), R416 appears to form interactions with the AMP phosphate group. The interactions between an arginine and a phosphate group are typically very strong and contribute significantly in ligand binding. R416A and R416M mutant *sd*Kef-QCTD were both expressed in very low levels compare to the wild type *sd*Kef-QCTD, likely due to the absence of bound AMP molecule. In the K^+ efflux assay, R416A retained some activity, R416M resulted in significant reduced activity, while R416E was completely inactive. With respect to R416A, AMP may still be accommodated, whereas steric occlusion to the binding site would be expected from R416M.

R416E, would likely to be more disruptive to the binding of AMP due to the repulsion between the γ-carboxylate of glutamate and the phosphate group of AMP.

Furthermore, the CPMG edited ^{1}H NMR analysis demonstrated that thermal denaturation of R416E *sd*Kef-QCTD does not result in release of the AMP molecule in solution (Figure 4.41), suggesting that there was no AMP natively bound to this mutant *sd*Kef-QCTD.

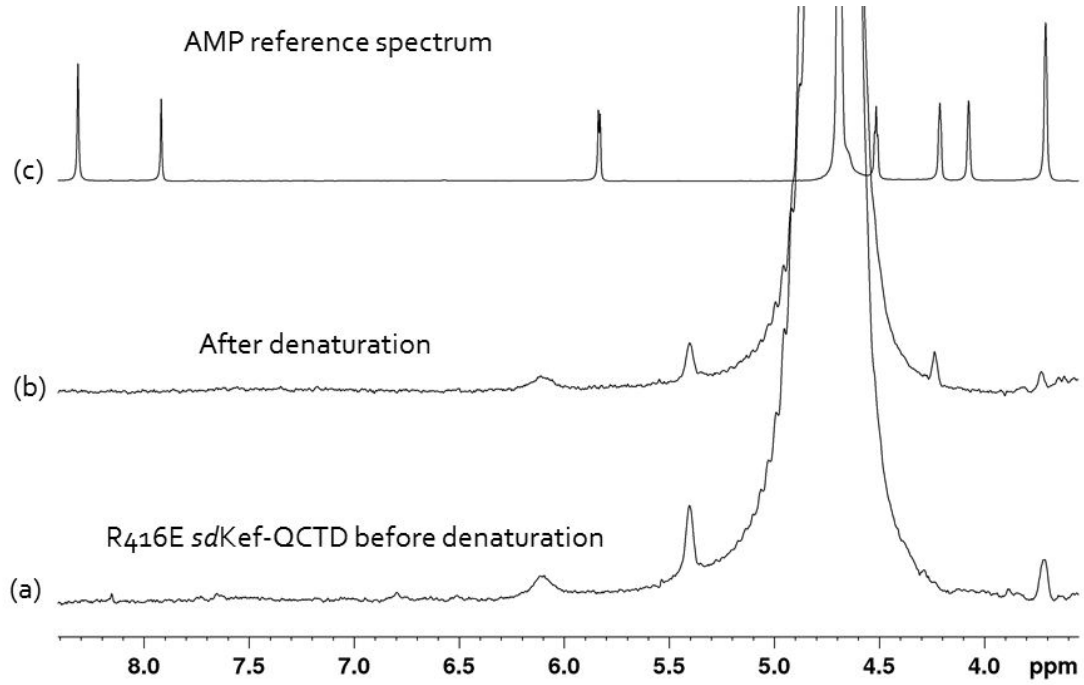


Figure 4.41 ^{1}H NMR experiment to prove that mutant R416E *sd*Kef-QCTD does not contain naturally bound AMP molecule.
(a) Experiment to prove that R416 (Figure 4.40) is the most crucial residue for the binding of AMP to *sd*Kef-QCTD. (a) CPMG ^{1}H NMR of R416E mutant *sd*Kef-QCTD (5 μM) before thermal denaturation (b) After thermal denaturation, no AMP signals are seen in this spectrum (c) AMP reference spectrum. The total echo time used in these experiments were 96 ms. The assay solutions were prepared in 150 mM sodium phosphate buffer, pH 7.4, in D_2O. The spectra were recorded at 278 K.

4.7.3 HPLC analysis

To provide additional experimental support to the finding that the variation in *sd*Kef-QCTD expression (stability) and function (Kef activation) is correlated to the degree of AMP in these proteins, we performed HPLC analysis (by Dr Anthony K. N. Chan) on the solutions of the thermally denatured formed of the mutant D436E, D436A (Figure 4.42) and R416E (Figure 4.43) *sd*Kef-QCTD to assess the amount of AMP that was co-purified with these proteins. It was observed that mutant D436E *sd*Kef-QCTD contained the same levels of AMP

as the wild type, whereas, D436A contained only 20% of AMP to the wild type. In case of R416E, the level of AMP detected was significantly less. These results are in good agreements with the mutations studies as described in section 4.7.2 and with the NMR analysis (R416E) that the presence of bound AMP helps in the *sd*Kef-QCTD stabilization (expression).

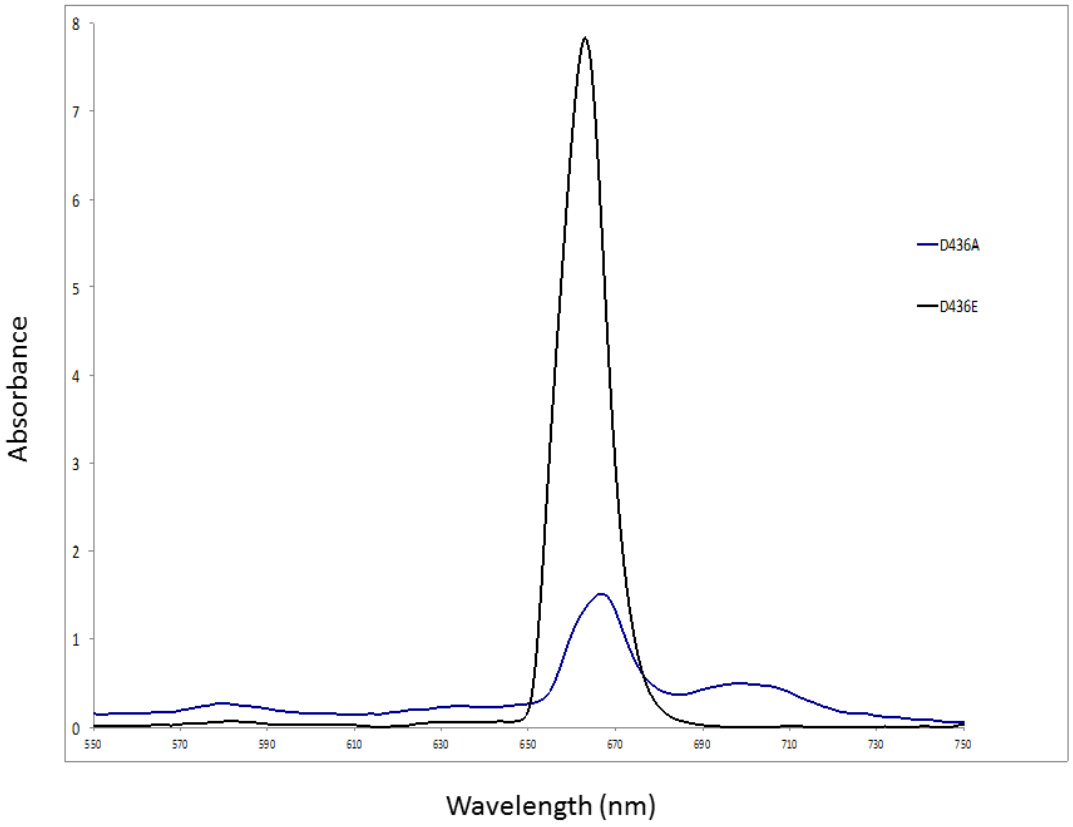


Figure 4.42 HPLC analysis for the presence of naturally bound AMP in the mutant D436A and D436E *sd*Kef-QCTD. AMP percentage of retention was calculated by comparing the absorption at 280 nm (A280) of the denatured proteins and AMP standards. D436E mutant *sd*Kef-QCTD was found to contain high level of natively bound AMP (black peak), whereas, D436A mutant *sd*Kef-QCTD was found to contain significantly low level of native bound AMP molecule (blue peak) (only 20% to that of D436E).

.

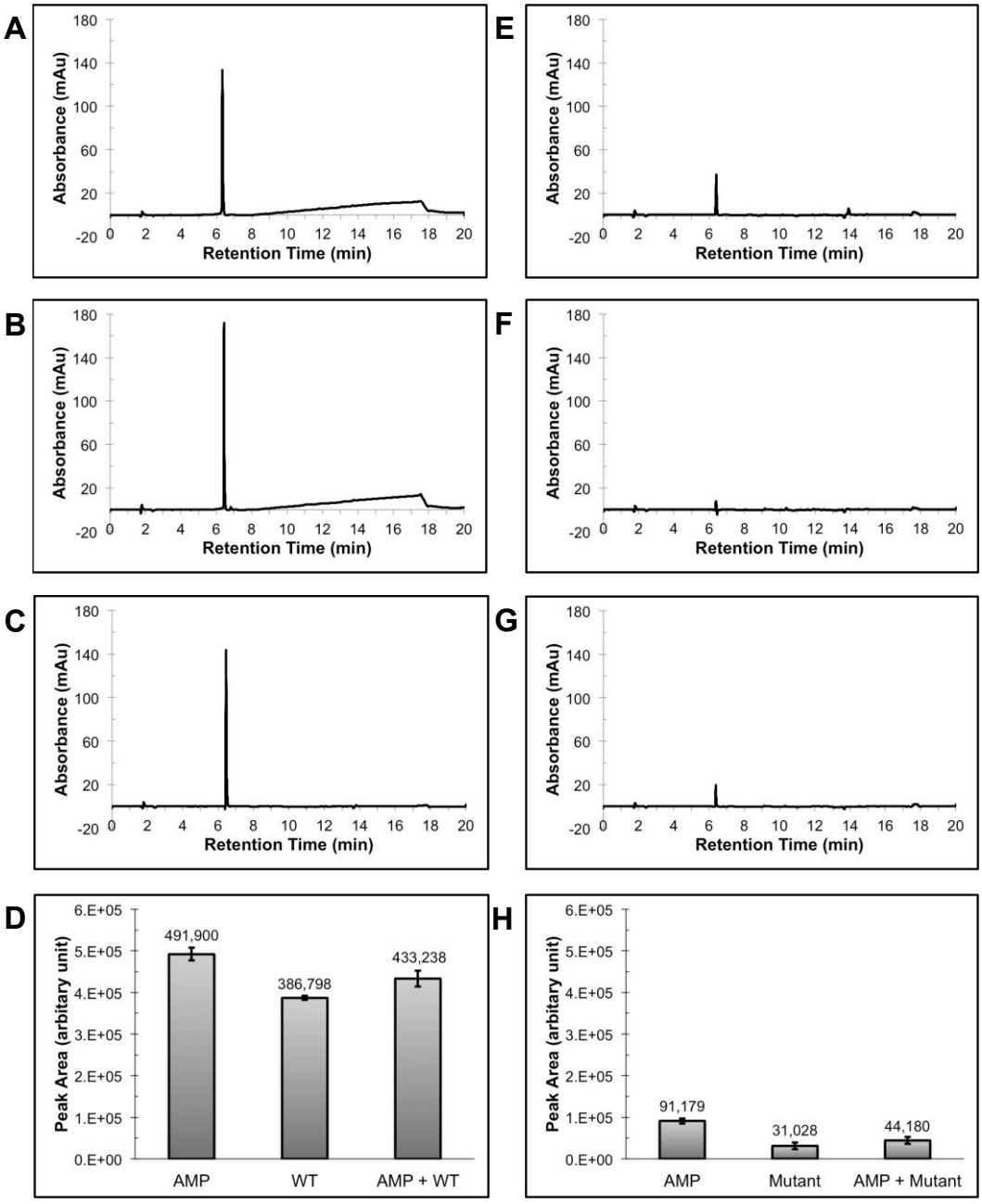


Figure 4.43 HPLC analysis of AMP released from thermally-denatured wild-type *sd*Kef-CTD and *sd*Kef-CTD (R416E). (A) HPLC profile of 50 µM pure adenosine monophosphate (AMP). (B) HPLC profile of 50 µM of denatured wild-type *sd*Kef-CTD. (C) Spiking experiment containing equal concentration (25 µM each) of pure AMP and denatured wild-type *sd*Kef-CTD. (D) Quantification of HPLC peak areas from A to C. Data shown are mean ± SD values from three different experiments (n = 3). (E) HPLC profile of 8.6 µM of pure AMP. (F) HPLC profile of 8.6 µM of denatured mutant *sd*Kef-CTD (R416E). (G) Spiking experiment containing equal concentration (4.3 µM each) of pure AMP and denatured mutant *sd*Kef-CTD (R416E). (H) Quantification of HPLC peak areas from E to G. Data shown are mean ± SD values from three different experiments (n = 3).

4.8 Summary

The work described in this chapter studied the binding interaction of Kef activating ligand glutathione-S-*N*-tertiary butylsuccinimide (^{t}BuSG) with *sd*Kef-QCTD by CPMG edited ^{1}H

NMR direct ligand-observation binding technique. Using this technique, we measured the binding dissociation constant value of ^{t}BuSG as 0.54 µM, which is good agreement to its reported literature value of 0.4 µM. With the aim of developing a potential novel non-peptidic antibiotic of Kef, which is a antibiotic target, we investigated the influence of each functional group of ^{t}BuSG on its binding interaction with *sd*Kef-QCTD. We have shown that the glutamate part of ^{t}BuSG except its ε-carbonyl group (more strictly carbonyl oxygen) is not essential for its binding interaction with *sd*Kef-QCTD, whereas, the rest of its functionalities i.e cysteine and glycine moieties are however, crucial for its binding interaction with *sd*Kef-QCTD. In exploring the size of the substituent group that is attached to the maleimide nitrogen of ^{t}BuSG, and which is crucial for activation of the K^+ efflux by displacing the F441 (*E.coli* KefC), we showed that a ligand with a bulky hydrophobic group of fluorescein also binds s*d*Kef-QCTD with an affinity (1.4 µM) close to ^{t}BuSG (K_D 0.4 µM). Based on this structure-activity relationship study, we incorporated a non-peptidic tetrazole functionality in the ^{t}BuSG and interestingly the ligand was found as a good binder of *sd*Kef-QCTD with an affinity of 7 µM. Towards our ultimate goal, selected compounds tested in this study have also shown interesting antibacterial activities.

In this study, we also applied ligand-observed ^{19}F NMR spectroscopy to study small molecule binding interaction with *sd*Kef-QCTD. We have shown that a slow exchanging ligand binding system on the NMR time scale is a potentially useful technique that can be used to probe the competitive binding of another ligand of interest through the NMR reporter displacement assay. More importantly, this technique allows calculating a ligand binding dissociation constant from a single shot experiment instead of a titration. The validity of this assay to *sd*Kef-QCTD system was exemplified with selected natural and synthetic Kef ligands and the values of their binding dissociation constants obtained with this technique were in good correlation to their reported literature values measured by other biophysical techniques.

Finally, we also demonstrated that the electron density that was found in the nucleotide binding pocket in crystal structure of *apo-sd*Kef-QCTD is unambiguously an AMP molecule that is natively bound to the protein; it is neither produced from the breakdown of NAD^+ nor it comes from the crystallisation mixture as a contaminant in the NAD^+ stock solution. We propose that his AMP molecule must be binding with *sd*Kef-QCTD with significantly high affinity in order to be retained by the protein during purification process. In addition, by using thermal shift studies, we showed that AMP stabilises the *sd*Kef-QCTD (dimeric form)

to a much greater extent than other nucleotides particularly NADH. Similarly, we also showed that the mutation of the key residues involved in the AMP binding caused loss of this ligand from the protein (*sd*Kef-QCTD) with simultaneous loss in the protein stability (expression).

4.9 References

1. Epstein, W. Osmoregulation by potassium transport in *Escherichia coli*. *FEMS Microbiol. Lett.* **1986**, 39, 73-78.
2. Booth, I. R. Regulation of cytoplasmic pH in bacteria. *Microbiol. Rev.* **1985**, 49, 359-78.
3. Nakamura, T.; Yuda, R.; Unemoto, T.; Bakker, E. P. KtrAB, a new type of bacterial K^+-uptake system from Vibrio alginolyticus. *J. Bacteriol.* **1998**, 180, 3491-4.
4. Derst, C.; Karschin, A. Evolutionary link between prokaryotic and eukaryotic K^+ channels. *J. Exp. Biol.* **1998**, 201, 2791-9.
5. Bossemeyer, D.; Borchard, A.; Dosch, D. C.; Helmer, G. C.; Epstein, W.; Booth, I. R.; Bakker, E. P. K^+-transport protein TrkA of *Escherichia coli* is a peripheral membrane protein that requires other trk gene products for attachment to the cytoplasmic membrane. *J. Biol. Chem.* **1989**, 264, 16403-10.
6. Roosild, T. P.; Castronovo, S.; Healy, J.; Miller, S.; Pliotas, C.; Rasmussen, T.; Bartlett, W.; Conway, S. J.; Booth, I. R. Mechanism of ligand-gated potassium efflux in bacterial pathogens. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, 107, 19784-9.
7. Ferguson, G. P.; Nikolaev, Y.; McLaggan, D.; Maclean, M.; Booth, I. R. Survival during exposure to the electrophilic reagent *N*-ethylmaleimide in *Escherichia coli*: role of KefB and KefC potassium channels. *J. Bacteriol.* **1997**, 179, 1007-12.
8. Ferguson, G. P.; McLaggan, D.; Booth, I. R. Potassium channel activation by glutathione-S-conjugates in *Escherichia coli*: protection against methylglyoxal is mediated by cytoplasmic acidification. *Mol. Microbiol.* **1995**, 17, 1025-33.
9. Vuilleumier, S. Bacterial glutathione S-transferases: what are they good for? *J. Bacteriol.* **1997**, 179, 1431-41.
10. Roosild, T. P.; Castronovo, S.; Miller, S.; Li, C.; Rasmussen, T.; Bartlett, W.; Gunasekera, B.; Choe, S.; Booth, I. R. KTN (RCK) domains regulate K^+ channels and transporters by controlling the dimer-hinge conformation. *Structure* **2009**, 17, 893-903.
11. McLaggan, D.; Rufino, H.; Jaspars, M.; Booth, I. R. Glutathione-dependent conversion of *N*-ethylmaleimide to the maleamic acid by *Escherichia coli*: an intracellular detoxification process. *Appl. Environ. Microbiol.* **2000**, 66, 1393-9.
12. Ferguson, G. P.; Totemeyer, S.; MacLean, M. J.; Booth, I. R. Methylglyoxal production in bacteria: suicide or survival? *Arch. Microbiol.* **1998**, 170, 209-18.

13. Ye, S.; Li, Y.; Chen, L.; Jiang, Y. Crystal structures of a ligand-free MthK gating ring: insights into the ligand gating mechanism of K^+ channels. *Cell* **2006**, 126, 1161-73.

14. Albright, R. A.; Ibar, J. L.; Kim, C. U.; Gruner, S. M.; Morais-Cabral, J. H. The RCK domain of the KtrAB K^+ transporter: multiple conformations of an octameric ring. *Cell* **2006**, 126, 1147-59.

15. Dong, J.; Shi, N.; Berke, I.; Chen, L.; Jiang, Y. Structures of the MthK RCK domain and the effect of Ca^{2+} on gating ring stability. *J. Biol. Chem.* **2005**, 280, 41716-24.

16. Roosild, T. P.; Miller, S.; Booth, I. R.; Choe, S. A mechanism of regulating transmembrane potassium flux through a ligand-mediated conformational switch. *Cell* **2002**, 109, 781-91.

17. Jiang, Y.; Lee, A.; Chen, J.; Cadene, M.; Chait, B. T.; MacKinnon, R. Crystal structure and mechanism of a calcium-gated potassium channel. *Nature* **2002**, 417, 515-22.

18. Jiang, Y.; Pico, A.; Cadene, M.; Chait, B. T.; MacKinnon, R. Structure of the RCK domain from the *E. coli* K^+ channel and demonstration of its presence in the human BK channel. *Neuron* **2001**, 29, 593-601.

19. Miller, S.; Ness, L. S.; Wood, C. M.; Fox, B. C.; Booth, I. R. Identification of an ancillary protein, YabF, required for activity of the KefC glutathione-gated potassium efflux system in *Escherichia coli*. *J. Bacteriol.* **2000**, 182, 6536-40.

20. Ness, L. S.; Booth, I. R. Different foci for the regulation of the activity of the KefB and KefC glutathione-gated K^+ efflux systems. *J. Biol. Chem.* **1999**, 274, 9524-30.

21. Miller, S.; Douglas, R. M.; Carter, P.; Booth, I. R. Mutations in the glutathione-gated KefC K^+ efflux system of *Escherichia coli* that cause constitutive activation. *J. Biol. Chem.* **1997**, 272, 24942-7.

22. Ferguson, G. P.; Munro, A. W.; Douglas, R. M.; McLaggan, D.; Booth, I. R. Activation of potassium channels during metabolite detoxification in *Escherichia coli*. *Mol. Microbiol.* **1993**, 9, 1297-303.

23. Elmore, M. J.; Lamb, A. J.; Ritchie, G. Y.; Douglas, R. M.; Munro, A.; Gajewska, A.; Booth, I. R. Activation of potassium efflux from *Escherichia coli* by glutathione metabolites. *Mol. Microbiol.* **1990**, 4, 405-12.

24. Healy, J.; Ekkerman, S.; Pliotas, C.; Richard, M.; Bartlett, W.; Grayer, S. C.; Morris, G. M.; Miller, S.; Booth, I. R.; Conway, S. J.; Rasmussen, T. Understanding the

structural requirements for activators of the Kef bacterial potassium efflux system. *Biochemistry* **2014**, 53, 1982-92.

25. Manzenrieder, F.; Frank, A. O.; Kessler, H. ^{31}P NMR spectroscopy as a versatile tool for compound library screening. *Angew. Chem., Int. Ed.* **2008**, 47, 2608-11.
26. Dalvit, C.; Fagerness, P. E.; Hadden, D. T.; Sarver, R. W.; Stockman, B. J. Fluorine-NMR experiments for high-throughput screening: theoretical aspects, practical considerations, and range of applicability. *J. Am. Chem. Soc.* **2003**, 125, 7696-703.
27. Leung, I. K. H.; Demetriades, M.; Hardy, A. P.; Lejeune, C.; Smart, T. J.; Szollossi, A.; Kawamura, A.; Schofield, C. J.; Claridge, T. D. W. Reporter ligand NMR screening method for 2-oxoglutarate oxygenase inhibitors. *J. Med. Chem.* **2013**, 56, 547-55.
28. Swann, S. L.; Song, D.; Sun, C.; Hajduk, P. J.; Petros, A. M. Labeled ligand displacement: Extending NMR-Based screening of protein targets. *ACS Med. Chem. Lett.* **2010**, 1, 295-299.
29. Cobb, S. L.; Murphy, C. D. ^{19}F NMR applications in chemical biology. *J. Fluor. Chem.* **2009**, 130, 132-143.
30. Mizukami, S.; Takikawa, R.; Sugihara, F.; Hori, Y.; Tochio, H.; Walchli, M.; Shirakawa, M.; Kikuchi, K. Paramagnetic relaxation-based ^{19}F MRI probe to detect protease activity. *J. Am. Chem. Soc.* **2008**, 130, 794-5.
31. Dalvit, C. Theoretical analysis of the competition ligand-based NMR experiments and selected applications to fragment screening and binding constant measurements. *Concepts in Magn. Reson. Part A* **2008**, 32A, 341-372.
32. Jackson, J. C.; Hammill, J. T.; Mehl, R. A. Site-specific incorporation of a (19)F-amino acid into proteins as an NMR probe for characterizing protein structure and reactivity. *J. Am. Chem. Soc.* **2007**, 129, 1160-6.
33. Kreutz, C.; Kahlig, H.; Konrat, R.; Micura, R. Ribose 2'-F labeling: a simple tool for the characterization of RNA secondary structure equilibria by ^{19}F NMR spectroscopy. *J. Am. Chem. Soc.* **2005**, 127, 11558-9.
34. Tengel, T.; Fex, T.; Emtenas, H.; Almqvist, F.; Sethson, I.; Kihlberg, J. Use of ^{19}F NMR spectroscopy to screen chemical libraries for ligands that bind to proteins. *Org. Biomol. Chem.* **2004**, 2, 725-31.
35. Dalvit, C.; Flocco, M.; Stockman, B. J.; Veronesi, M. Competition binding experiments for rapidly ranking lead molecules for their binding affinity to human serum albumin. *Comb. Chem. High Throughput Screen.* **2002**, 5, 645-50.

36. Matei, E.; Andre, S.; Glinschert, A.; Infantino, A. S.; Oscarson, S.; Gabius, H. J.; Gronenborn, A. M. Fluorinated carbohydrates as lectin ligands: dissecting glycan-cyanovirin interactions by using ^{19}F NMR spectroscopy. *Chemistry (Easton)* **2013**, 19, 5364-74.

37. Goudreau, N.; Coulombe, R.; Faucher, A. M.; Grand-Maitre, C.; Lacoste, J. E.; Lemke, C. T.; Malenfant, E.; Bousquet, Y.; Fader, L.; Simoneau, B.; Mercier, J. F.; Titolo, S.; Mason, S. W. Monitoring binding of HIV-1 capsid assembly inhibitors using ^{19}F ligand-and ^{15}N protein-based NMR and X-ray crystallography: early hit validation of a benzodiazepine series. *ChemMedChem* **2013**, 8, 405-14.

38. Jenkins, B. G.; Lauffer, R. B. Detection of site-specific binding and co-binding of ligands to human serum albumin using ^{19}F NMR. *Mol. Pharmacol.* **1990**, 37, 111-8.

39. Dugad, L. B.; Gerig, J. T. NMR studies of carbonic anhydrase-4-fluorobenzenesulfonamide complexes. *Biochemistry* **1988**, 27, 4310-4316.

40. Morris, G. A. Reference Deconvolution. In *eMagRes*, John Wiley & Sons, Ltd: **2007**.

41. Metz, K. R.; Lam, M. M.; Webb, A. G. Reference deconvolution: A simple and effective method for resolution enhancement in nuclear magnetic resonance spectroscopy. *Concepts Magn Reson* **2000**, 12, 21-42.

42. Arntson, K. E.; Pomerantz, W. C. Protein-observed ^{19}F NMR: A bioorthogonal approach for small molecule discovery. *J. Med. Chem.* **2015**, DOI: 10.1021/acs.jmedchem.5b01447.

43. Marsh, E. N. G.; Suzuki, Y. Using ^{19}F NMR to probe biological interactions of proteins and peptides. *ACS Chem. Biol.* **2014**, 9, 1242-1250.

44. Yu, J. X.; Hallac, R. R.; Chiguru, S.; Mason, R. P. New frontiers and developing applications in ^{19}F NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, 70, 25-49.

45. Chen, H.; Viel, S.; Ziarelli, F.; Peng, L. ^{19}F NMR: a valuable tool for studying biological events. *Chem. Soc. Rev.* **2013**, 42, 7971-82.

46. Kitevski-LeBlanc, J. L.; Prosser, R. S. Current applications of ^{19}F NMR to studies of protein structure and dynamics. *Prog. Nucl. Magn. Reson. Spectrosc.* **2012**, 62, 1-33.

47. Hanukoglu, I. Proteopedia: Rossmann fold: A beta-alpha-beta fold at dinucleotide binding sites. *Biochem. Mol. Biol. Educ.* **2015**, 43, 206-209.

48. Lesk, A. M. NAD-binding domains of dehydrogenases. *Curr. Opin. Struct. Biol.* **1995**, 5, 775-783.

49. Hanukoglu, I.; Gutfinger, T. cDNA sequence of adrenodoxin reductase. *Eur. J. Biochem.* **1989**, 180, 479-484.

50. Rao, S. T.; Rossmann, M. G. Comparison of super-secondary structures in proteins. *J. Mol. Biol.* **1973**, 76, 241-256.

5 Thesis Summary and Future Work

In summary, the work described in this thesis is focused on three projects;

1) γ-Butyrobetaine hydroxylase (GBBH or commonly referred as BBOX; from BBOX gene in human), is a 2-oxoglutarate (2OG) and Fe^{II} dependent oxygenase, currently a drug target for the treatment of myocardial infarction. It was demonstrated that ^{1}H NMR-based direct ligand observation is a simple and efficient NMR technique for studying small molecules binding interaction with BBOX. In this assay the paramagnetic relaxation enhancement (PRE) effect of Mn^{II} (an unreactive substitute for catalytic Fe^{II}) was used to enhance the NMR sensitivity of ligand binding to BBOX with significantly low amount of protein. Based on this approach, a ^{1}H NMR-based competition assay was developed using both natural substrate γ-butyrobetaine (GBB) and co-substrate 2OG simultaneously as reporter ligands. This technique readily assesses from a single NMR spectrum whether a competitive inhibitor of interest competes with GBB or 2OG or both. The applicability of this technique was exemplified with a structure-activity relationship study using isoquinoline-based ligands and a variety of other ligands containing the quinoline, pyrimidine and pyridine-based ligands. This assay is potentially useful in the development of selective inhibitors of BBOX over other 2OG oxygenases of biological interests.

In this assay in order to overcome the potential problem of signals overlap (^{1}H NMR), a ^{19}F NMR-based assay is proposed for future work, by employing a fluorinated GBB analogue such as GBBF or GBBNF and a fluorinated 2OG analogue as dual-reporter in the assay.

It was also shown that the positive charge on the nitrogen atom in the trimethylammonium group of GBB plays a critical role in psBBOX substrate recognition. GBB analogues which lack this positive charge are not accepted as substrates by psBBOX. It is still unknown whether all the aromatic cage residues (Tyr177, Tyr-194, Trp-181, Tyr-205 and Tyr-366) of BBOX (where GBB trimethylammonium group binds) are critical for this cation-π interaction or may be its one or two of these residues that are critically involved in cation-π interaction as observed in the case of acetylcholine (neurotransmitter) i.e cation-π interaction between the positively charged nitrogen of acetylcholine and a single tryptophan residue. Hence for future work, a mutation study of these aromatic cage residues and its influence upon GBB binding is proposed. The GBB carboxylate -OH group forms interactions with Asp204 and Asn292 (human BBOX); studies by Rydzik *et al.* suggest that GBB analogue lacking the carboxylate

-OH group (GBB-CH_3) is not accepted as substrate by both human and psBBOX; therefore for a future work we propose a mutation study to establish whether both these residues are critical for GBB binding or only one of them is crucial.

We also observed that a GBB mimic substrate of psBBOX including the commercially available clinical drug mildronate cannot fully displace GBB from its binding pocket. We propose that upon GBB/2OG/metal co-factor binding, substantial conformational changes occur in BBOX that possibly locks the GBB in one monomer and which cannot be outcompeted by another substrate mimic ligand, giving rise to the possibility that at a time only one monomer may involve in catalysis. For future work, we therefore propose protein-observed ^{19}F NMR studies (fluorine labelling of aromatic cage residues) to investigate that why GBB cannot be fully displaced from its binding pocket by another competitive substrate.

2) The activity of many 2OG oxygenases is enhanced by the presence of ascorbate in the assay mixture; the exact mode of action of which still remains unknown. We have shown that in the assay solution, without a 2OG oxygenase, ascorbate slowly undergoes oxidative degradation and cause concomitant 2OG decarboxylation into succinate. We speculate that the enhanced uncoupled 2OG turnover by a 2OG oxygenase in the presence of ascorbate may have potential artifacts of ascorbate-induced 2OG turnover instead formed by catalysis. We carried out a systematic study to investigate the influence of different experimental conditions including the ascorbate concentration, pH, buffer, effect of metals etc. on this ascorbate induced 2OG decarboxylation process. Our observations that other compounds such as dithiothreitol (DTT) are as good reducing agents as ascorbate in this process, suggest that these different compounds (ascorbate, DTT etc.) possibly interact and stimulate a 2OG oxygenase in a compound specific manner and not through a common mechanism i.e by reducing Fe^{III} to catalytic Fe^{II} in solution to activate the enzyme. For future work, we therefore propose the investigation of the binding interaction of ascorbate and some of its analogues such as D-isoascorbate, isopropylidine-ascorbate (IPPA), 6-deoxyascorbic acid, 6-aminodeoxyascorbic acid, 6-O-phenylascorbic acid, 6-S-phenylascorbic acid, 6-S-phenyl-2-aminoascorbic acid and imino-D-glucoascorbic acid to a variety of 2OG oxygenases and investigation of their simulation by these compounds. This study will help in understanding of how a particular 2OG oxygenase is stimulated by a specific reducing agent. For mapping of the ascorbate binding site, the technique of ^{1}H-^{15}N chemical shift perturbation is also proposed. Alternatively, the binding of ascorbate can also studied by sensitive techniques

such as ligand titration to fluorine-labelled protein. The binding mode of ascorbate to 2OG oxygenases can also be investigated *via* X-ray crystallographic studies.

3) Bacterial potassium ion efflux (Kef) is a ligand-gated K^+ channel system that protects bacterial against toxic electrophilic species. Bacterial Kef system is a novel target for the development of antibiotic agents that display novel mode of action. The strongest binder and activator of Kef reported so far is synthetic GS-X (glutathione-S-conjugate; X = electrophile) formed with the *N*-tertiary butylsuccinimide electrophile (K_D 0.4 μM). We identified the functional groups of this GS-X that form critical interactions with *sd*Kef-QCTD (*sd*: *Shewanella denitrificans*, QCTD: Q-linker C-terminal domain) using a structure-activity relationship (SAR) study. Based on this study a truncated (at glutamate) novel non-peptidic GS-X was synthesized in which the ligand glycine was replaced with a tetrazole functionality. It was then shown that this novel GS-X binds *sd*Kef-QCTD with an affinity in close range to the *N*-tertiary butylsuccinimide-based GS-X. Work is currently in progress in order to prove that this non-peptidic GS-X is a cell permeable antibiotic agent.

An efficient ligand-based ^{19}F NMR reporter displacement assay was developed that can be used for ligand screening as well as measuring their binding constant values from a single NMR spectrum. In this assay we observed the signals of both free and bound (to *sd*Kef-QCTD) populations of a fluorinated ligand; a slow exchange system on NMR time scale. The applicability of this system in competitive ligand screening was assessed using a range of reported GS-X and the results obtained were in reasonable correlations to their reported data obtained with other biophysical techniques. The slow exchange binding of this fluorinated ligand was then used to probe conformational changes that occur in Kef as a result of GSH (K^+ channels get closed) and GS-X (K^+ channels get opened) binding. With GSH and GS-X (*N*-tertiary butylsuccinimide) binding, different effects on the local chemical environments of the bound population of the fluorinated ligand were observed. These possibly reflect the closing of the K^+ channels upon GSH binding and the opening of the K^+ channels upon GS-X binding. In order to confirm these theories, for a future work, the critical F448 (F441 in *E.coli*) can be labelled with a fluorine and the titration of GSH and GS-X can be employed to probe these two different forms of Kef proteins.

We also proved that the electron density in the Rossmann fold of the *apo sd*Kef-QCTD is unambiguously an AMP molecule that is naturally bound to the protein and likely has a role in the stabilisation of the dimeric form of the Kef protein responsible for gating.

6 Experimental

6.1 Materials and chemicals

All the materials and chemicals used in the NMR experiments were from Sigma Aldrich or Alfa Aesar or Fluka, except those listed in Table 6.1, unless otherwise mentioned. These were used as provided by the supplier.

Table 6.1 suppliers of chemicals for NMR experiments

Chemicals	Suppliers
Tris-D_{11}	Cortectnet
D_2O	Apollo Scientific
DMSO-d_6	Merk Chemicals

6.2 Buffer for NMR experiments

Unless otherwise mentioned all the NMR experiments were run in buffer. In the psBBOX work (chapter 2), 50 mM Tris-D_{11}, pH 7.5 in D_2O was used. In the ascorbate work (chapter 3), 50 mM Tris-D_{11}, pH 7.5 in MilliQ H_2O was used. 10% D_2O was added to these samples for the spectrometer deuterium lock signal. Milli-Q H_2O was obtained from an Elix Reverse Osmosis system coupled with a Milli-Q Synthesis system and a 0.22 µm filter (Millipore). The Kef work (chapter 4) was performed with 50 mM sodium phosphate (NaH_2PO4/Na_2HPO4, supplemented with 150 mM NaCl) buffer, pH 7.4 in D_2O. For H_2O-based buffer, the pH was adjusted using HCl and NaOH, whereas, for the D_2O-based buffer, DCl and NaOD were used. All the buffers were filter sterilised through 0.22 µm filters (Millipore) and were stored in 50 mL falcon tubes (Sarstedt) at 4 ºC.

6.3 General sample preparation

All samples were prepared in a 1.5 mL Eppendorf microtubes and were transferred to the NMR tubes *via* either a micropipette (to conventional 5 mm diameter NMR tubes; Norell) or gel-loading tips (Fisher Scientific) (to 3 mm diameter Bruker MATCH NMR tubes). A sample volume of 160 µL and 500 µL was used for 3 mm and 5 mm NMR tubes, respectively. Unless otherwise mentioned, all stock solutions were prepared to the appropriate concentrations in the relevant NMR buffer or DMSO-D_6. Sodium, potassium, magnesium, calcium, zinc, copper, nickel, manganese and cobalt stock solutions were made

from their corresponding chloride salts (NaCl, KCl, $MgCl_2$, $CaCl_2$, $ZnCl_2$, $CuCl_2$, $NiCl_2$, $MnCl_2$, $CoCl_2$) in the relevant NMR buffer. The iron(II) stock solution (250 mM) was freshly prepared for each set of experiments using $(NH_4)_2Fe(SO_4)_2 \cdot 6H_2O$ in 20 mM HCl in H_2O, which was then diluted down to the appropriate concentration by the relevant NMR buffer just before adding to the reaction mixture. L-Ascorbate solutions were freshly prepared each day using sodium ascorbate salt in the buffer. All the assay mixtures studied for L-ascorbate degradation were incubated in a 1.5 mL transparent plastic eppendorf at room temperature on bench. The stock solutions of ascorbate and other reducing agents used in the ascorbate degradation assays were made in 25 mM concentration in H_2O-based Tris-D_{11} buffer. Similarly, the stock solutions of 2OG and other α-ketoacids used in these studies were made in 16 mM concentrations. All these mixtures were made with a total sample volume of 160 µL. 2OG stock solutions were prepared using 2OG monosodium salt in H_2O-based NMR buffer. GBB stock solutions were prepared using GBB chloride salt in the relevant buffer. psBBOX was produced by Dr Jurgen Brem according to the protocol reported by Rydzik *et al*[1]. *sd*Kef-QCTD was produced by Dr Anthony K. N. Chan according to the procedure reported by Healy *et al*.[2] The psBBOX used was incubated in 50 mM Tris buffer pH 7.5 in H_2O supplemented with 200 mM NaCl, whereas, *sd*Kef-QCTD used was incubated in with 50 mM sodium phosphate (NaH_2PO4/Na_2HPO4, supplemented with 150 mM NaCl) buffer, pH 7.4 in D_2O. The concentrations of purified proteins were determined by measuring the A280 using an ND-1000 spectrophotometer (NanoDrop Technologies). The Protein A280 and Other protein (E & MW) options were used. Protein buffer solution was used as blank. Protein molecular weights (MW) and extinction coefficients (E) were calculated using the ProtParam tool. Typically, 2 µL of the analyte was used per measurement. All measurements were performed at least three times. The catalase used was an aqueous suspension (10,000-40,000 units/mg protein) from bovine liver. In the anaerobic ascorbate degradation assay, all the stock solutions and buffer were first degassed and were then transferred into anaerobic chamber in sealed bottles. The assay was performed in the anaerobic chamber. After incubation in the anaerobic chamber the spectra were recorded within 5 minutes. Catalase was used as 3 µL from a 45mg/ml aqueous suspension (45×3/1000 = 135ug of protein; batch activity: 12852 units/mg, 12852×135/1000 = 1735units). The isoquinoline-based ligands were provided by Dr Anna M. Rydzik (synthesised according to the published protocols).[3,4] The GS-X compounds (glutathione-S-conjugates) were synthesised by Sam Grayer as per the published methodology.[2] Jos J. A. Kamps synthesised the GBB phospha, arsa and carba analogues.[5] Silvia Ekkerman performed the K^+ efflux assay.

6.4 Sample acquisition and processing

All NMR experiments were conducted at 298 K (unless otherwise mentioned) using either a Bruker Avance III 700 MHz spectrometer equipped with a 5 mm inverse TCI cryoprobe or a Bruker Avance II 500 MHz spectrometer equipped with a 5 mm inverse $^{1}H/^{19}F(^{13}C)$ TXI probe or a Bruker Avance III 600 MHz spectrometer equipped with a 5 mm BB-$^{19}F/^{1}H$ Prodigy N_2 cryoprobe. Unless otherwise stated, all the spectra were recorded at 298K. The spectra were processed by TopSpin 3.1 software. Spectra were processed typically with a line broadening of 0.3-1 Hz. The ^{1}H pulse tip-angle calibration was carried using the single notation method for the samples before running the actual experiments. The ^{19}F pulse was calibrated using 1,1,1-trifluoroacetone in the relevant NMR buffer. Unless otherwise stated, the area of the peaks was measured by integration. Each experiment was run with three separately prepared samples. Water suppression was achieved by excitation sculpting method (for samples in H_2O-based buffer) and presaturation.

6.5 Protein-titration experiments for ligand K_D measurements by CPMG edited ^{1}H NMR

All the CPMG edited ^{1}H NMR protein titration experiments for ligand K_D measurement were performed in 3 mm diameter Bruker MATCH NMR tubes on a Bruker Avance III 700 MHz spectrometer equipped with a 5 mm inverse TCI cryoprobe. In this assay separate samples were prepared for each protein concentration (containing a fixed ligand concentration). The CPMG experiments were recorded using the PROJECT – CPMG pulse sequence ($90°_x$ – [τ – $180°_y$ – τ – $90°_y$ – τ – $180°_y$ –τ]$_n$ – acq).[6] The total echo time (2τ) used were in the range of 32-88 ms. NMR parameters such as number of transients and acquisition time were judged on a per sample basis. The concentrations of the ligands used were typically 10-25 μM. The loss in the ligand signal intensity in the presence of protein (as a probe of binding interaction) was measured relative to the intensity (set to unit intensity) of the same signal in the ligand solution in the absence of protein. The K_D values were obtained from fitting the titration data in OriginPro 9.0 (Origin lab, Northampton, MA, USA) using the following quadratic equation,

$$y = A*(1/(2*C))*((B+x+C)-sqrt(((B+x+C)^2)-(4*x*C)))$$

Where y represents the loss in the ligand signal intensity (NMR response of ligand binding), A represents maximum NMR response of binding (when a ligand signal completely

attenuates in the presence of protein; saturation point), C represents fixed component (ligand) concentration and B represents the ligand K_D.

6.6 CPMG edited ^{1}H NMR dual-reporter displacement assay

The CPMG edited ^{1}H NMR dual-reporter ligand displacement assay as described in section 2.5 were performed in 3 mm diameter Bruker MATCH NMR tubes on a Bruker Avance III 700 MHz spectrometer equipped with a 5 mm inverse TCI cryoprobe. The experimental conditions used were, 25 µM GBB, 300 µM 2OG, 150 µM Mn^{II}, 80 mM KCl, 50 µM 1,1,1 trifluoroacetone (internal NMR standard), 15 µM psBBOX and 25 µM competitive ligand. The NMR parameters used were, acquisition time 1.46 s, relaxation delays 2 s, number of transients 32 and total echo time 32 ms. The competitive ligand was added to the assay mixture after the addition of protein. Each experiment was recorded with three separate samples.

6.7 Intrinsic fluorescence quenching binding assay for K_D measurement

The binding dissociation constant values (K_D) of the isoquinoline-based ligands with psBBOX described in section 2.5.1 were determined (performed by Robert K. Lesniak) from the quenching of psBBOX intrinsic tryptophan fluorescence measured at 294 K on a Pherastar FS plate reader (BMG labtech) with 96-well plates (Greiner, black, bottom: flat, clear). The experimental conditions used in these experiments were; excitation 280 nm, emission 350 nm, 5 µM psBBOX, 50 µM Mn^{II} and varying concentrations of the isoquinoline ligands to a final well volume of 50 µL in 50 mM Tris buffer pH 7.5, supplemented with 200 mM NaCl. Fluorescence readings were obtained in triplicate with errors represented as standard deviations. The binding dissociation constant (K_D) values were obtained by fitting the raw data (plotting inhibitor concentration versus $\Delta F_{obs}/F_{max}$, where ΔF_{obs} represents the decrease in the observed fluorescence signal and F_{max} the observed fluorescence signal without the isoquinoline ligand present) using Prism according the quadratic equation for K_D measurement.

6.8 ^{1}H NMR assay for BBOX inhibitor IC_{50} measurements

The BBOX turnover assay was carried out in a 1.5 mL eppendorf tube. The reaction was initiated with the addition of enzyme to the assay mixture and quenched at 60th second with the addition of 10 µL of 1M HCl. The concentrations of GBB turnover into L-carnitine were

measured by integration of their respective Me_3 signal from their trimethylammonium group in a standard ^1H NMR spectrum. The % enzyme activity was calculated as ($[CAR]_{P+I}$ / $[CAR]_P$) × 100 , where P = Protein (BBOX), I = Inhibitor. Inhibitors IC_{50} values were obtained from growth/sigmoidal dose response curves fitting (log of inhibitor concentration vs % enzyme activity) using OriginPro 9.0 (Origin lab, Northampton, MA, USA) according to the following equation.

$$y = A1 + \frac{A2 - A1}{1 + 10^{(LOGx0 - x)p}}$$

Where; y = response (% activity), $\log x0$ = the centre of the curve, p = hill slope, $A2$ = maximum activity, $A1$ = minimum activity

The experimental conditions used were; 200 µM GBB, 600 µM 2OG, 100 µM Fe^{II} which was prepared as a 250 mM stock solution in 20 mM HCl in H_2O (which was further diluted to a 2.5 mM solution in MilliQ purified water), 500 µM L-ascorbate, 80 mM KCl and 0.140 µM psBBOX in 50 mM Tris-D_{11}, pH 7.5 in D_2O. Each experiment was performed with three separately prepared samples.

6.9 Fluoride ion release assay for BBOX inhibitors IC_{50} measurements

Fluoride ion release assay was applied to measure the IC_{50} values of the isoquinoline-based ligands (performed by Robert K. Lesniak) described in section 2.6. The fluorescence signal was detected using a Pherastar FS plate reader (BMG labtech) fitted with a FITC FP 485/30 (485 nm, bandwidth 30 nm) and FITC FP 520/40 emission (520 nm, bandwidth 40 nm) filters, using 384-well plates (black, clear, flat bottom, Grenier BioOne). The TBS-protected fluorescein probe and 3-(*S*)-fluoro-GBB (GBBF) were synthesised according to the reporter method.[7] The assay conditions used were; 50 µM GBBF, 500 µM 2OG, 250 µM sodium ascorbate, varying concentrations of isoquinoline-ligands, 50 µM Fe^{II} (from $Fe(NH_4)_2(SO_4)_2$ salt. The $Fe(NH_4)_2(SO_4)_2$ solution was prepared as a 100 mM stock solution in 20 mM HCl (in H_2O), which was further diluted to a 250 µM solution in MilliQ purified water before being added at the start of the measurements. The psBBOX catalysed reaction was initiated by adding the protein (1 µM) at the end to the assay mixture in each well containing all inhibitors and cofactors. Assays were performed to a final well volume of 10 uL in 50 mM Tris buffer supplemented with 200 mM NaCl, pH 7.5 at 294 K. The reaction was quenched by the addition of 40 µL of TBS-protected fluorescein probe (final concentration 5 µM) after

10 mins. The plates were then sealed and incubated at room temperature for 1 hour. This was followed by the addition of 10 µL of HEPES buffer, pH 7.0 and the fluorescence signal read up for 5 minutes. The fluorescence data were recorded as quadruplicates and errors are shown as standard deviations. The fluorescence signals were normalized by subtracting the fluorescence values recorded from control wells, which contained all the assay components except psBBOX or without an inhibitor. The IC_{50} values of the inhibitors were obtained by fitting the raw data using the Prism four parameter logistic dose-response model (sigmoidal, variable slope).

6.10 Steady state fluorescence measurements (Fluorescence Competition Assay)

All measurements were conducted on a Perkin Elmer LS-50B Spectrometer (performed by Dr Samuel C. Grayer). The samples were prepared in 150 µL volume containing *sd*Kef-QCTD (6 µM) and DNGSH (5 µM) in 50 mM phosphate buffer, pH 7.4 supplemented with 150 mM NaCl. To each sample, the competing ligand of interest (1 mM) was added. Samples were excited with light of 340 nm and the emission spectra measured from 375-670 nm (the step was 1 nm, the dwell time 0.3 s, 3 repeats, the excitation slit was 1 nm, the emission slit 2 nm and the temperature was held at 20 °C). A reduction in the fluorescence intensity of the reporter ligand DNGSH after the addition of the competing ligand was interpreted as an indication of binding interaction of the competitive ligand with *sd*Kef-QCTD. The binding of the competitive ligands are expressed as ratios of the fluorescence intensities of the DNGSH before and after the addition of the tested competitive ligands at 525.5 nm ($DNGSH_{bound}$:$DNGSH_{ligand}$ or $(F_B/F_L - 1)$.

6.11 Differential scanning fluorimetric (DSF) assay

The DSF experiments (performed by Dr Samuel C. Grayer) were performed on a Stratagene Mx3005P qPCR (Expt filter set, ex. 492 nm, em. 568 nm). The initial temperature value was set to 25 °C (held for 5 minutes), increasing in increments of 1 °C for 55 cycles (held for 1 minute 30 seconds per cycle). Ligands stock solutions were prepared in 100 mM concentration in 50 mM sodium phosphate buffer, supplemented with 150 mM NaCl, pH 7.4. The 100 mM stock solutions were then diluted to 10 mM in the same buffer. A protein master mixture was prepared containing *sd*Kef-QCTD (13.3 µM) and Sypro Orange dye (2.2×, Invitrogen) in 50 mM sodium phosphate buffer supplemented with 150 mM NaCl, pH 7.4.

Ninety-six well plates (Axygen) were prepared using the protein master mixture (22.5 µL per well; 12 µM final concentration of protein and 2× final concentration of dye) and the appropriate ligand (2.5 µL per well; 1 mM final concentration). The plate was centrifuged at 1000 revolution per minute (rpm) for 3 minutes before being run. Controls experiments were performed with dye alone, ligand + dye, and the *sd*Kef-QCTD alone. The T_M (melting temperature) was identified by fitting the data to the Boltzmann equation (Prism 5).[8] The change in unfolding temperature (ΔT_M) of *sd*Kef-QCTD was calculated as the shift in T_M relative to the T_M of the *sd*Kef-QCTD + 2.5 µL buffer, in the absence of any ligand. A student T-test was performed to ensure that the changes were statistically significant.

6.12 Binding of Mn^{II} to psBBOX by bulk water T_1 measurements

The assay was performed according to the technique reported by Leung *et al.*[9] The bulk water T_1 was measured by saturation recovery method ($90°_x - G_1 - 90°_y - G_2 - 90°_x - G_3 - \tau$ –acq). NMR parameters used were: number of transients 1, relaxation delay 22 s. The receiver gain (RG) value was set minimum of 1 in order to prevent radiation damping.[10,11] Samples were made in 50 mM Tris-D_{11}, pH 7.5, 20% H_2O: 80% D_2O. The assay mixture contained 50 µM Mn^{II}, 1.2 mM 2OG, 100 mM KCl and varying concentration of psBBOX (psBBOX titration). The experiments were recorded on a Bruker Avance III 600 MHz spectrometer equipped with a 5 mm BB-$^{19}F/^1H$ Prodigy N_2 cryoprobe in 3 mm diameter Bruker MATCH NMR tubes with a total sample volume of 160 µL. Bulk water T_1 values were obtained using the Bruker T_1/T_2 Relaxation option and peak area was used for curve fitting. The curves were fitted according to *uxnmrt1* function as below:

$$I_{(\tau)} = I_{(0)}(1 - Ae^{-\tau/T1})$$

Where $I_{(\tau)}$ represents the signal intensity at time τ, and $I_{(0)}$ is the fully relaxed signal intensity, A is the correction factor to compensate for imperfection saturation and τ is the time delay. Each measurement was made with three freshly prepared separate samples. The Mn^{II} K_D value was obtained by fitting the raw data (psBBOX concentration against bulk water T_1) using OriginPro 9.0 (Origin lab, Northampton, MA, USA) according to the quadratic K_D equation.

6.13 ^{1}H NMR experiments in ascorbate degradation studies

All the ascorbate degradation analysis were carried out by recording standard ^{1}H NMR experiments at 298K on a Bruker Avance III 700 MHz spectrometer equipped with a 5 mm inverse TCI cryoprobe in 3 mm diameter Bruker MATCH NMR tubes using a total sample volume of 160 µL. At the end of incubation period 10% of D_2O was added to each sample for spectrometer deuterium locking. All the spectra were recorded within five minutes after the incubation times.

6.14 ^{19}F NMR experiments

All the ^{19}F NMR experiments described in section 4.4 and section 4.5 were conducted on a Bruker Avance III 600 MHz spectrometer equipped with a 5 mm BB-^{19}F/^{1}H Prodigy N_2 cryoprobe in 5 mm diameter NMR tubes with a total sample volume of 500 µL. In all these experiments 1,1,1-trifluoroacetone (-86.53 ppm) was used as an internal NMR standard. The spectra were processed with a line broadening of 20 Hz and the areas of the peaks were measured by Lorentz Gauss deconvolution method[12,13] automatically using the Bruker Topspin[3.2]. All the samples were prepared in 50 mM sodium phosphate (NaH_2PO4/Na_2HPO4, supplemented with 150 mM NaCl) buffer, pH 7.4 in D_2O. Typical NMR parameters used were; spectral width 60 ppm, acquisition time 1.92 s, relaxation delays 2 s and number of transients 1312.

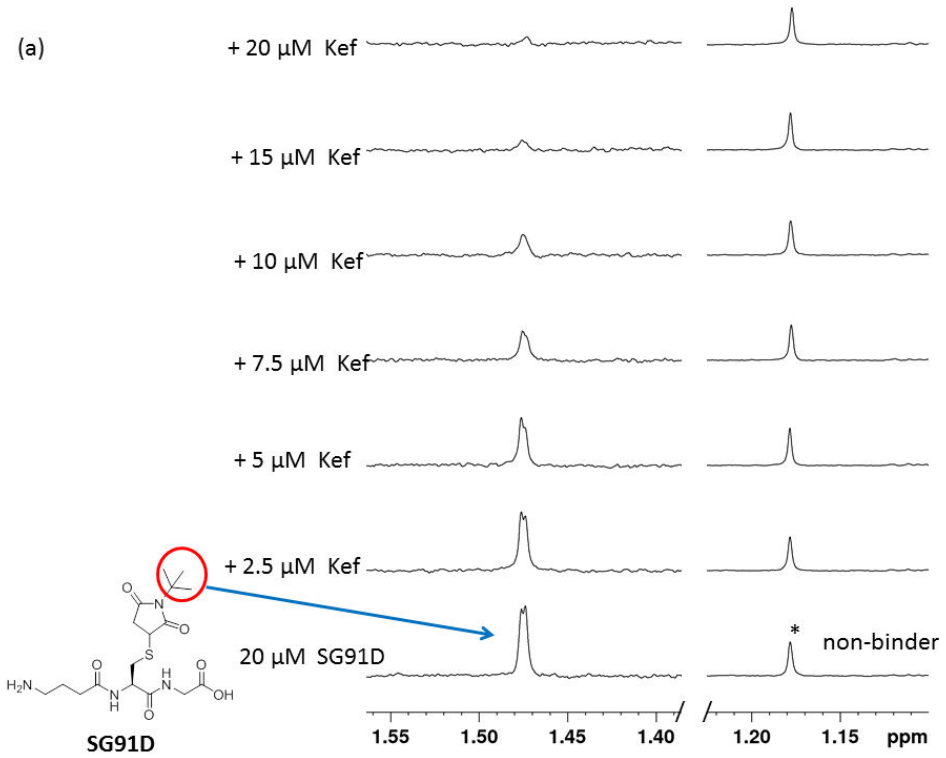
6.15 References

1. Rydzik, A. M.; Leung, I. K.; Kochan, G. T.; Loik, N. D.; Henry, L.; McDonough, M. A.; Claridge, T. D.; Schofield, C. J. Comparison of the substrate selectivity and biochemical properties of human and bacterial gamma-butyrobetaine hydroxylase. *Org. Biomol. Chem.* **2014**, 12, 6354-8.
2. Healy, J.; Ekkerman, S.; Pliotas, C.; Richard, M.; Bartlett, W.; Grayer, S. C.; Morris, G. M.; Miller, S.; Booth, I. R.; Conway, S. J.; Rasmussen, T. Understanding the structural requirements for activators of the Kef bacterial potassium efflux system. *Biochemistry* **2014**, 53, 1982-92.
3. Rydzik, A. M.; Chowdhury, R.; Kochan, G. T.; Williams, S. T.; McDonough, M. A.; Kawamura, A.; Schofield, C. J. Modulating carnitine levels by targeting its biosynthesis - selective inhibition of [gamma]-butyrobetaine hydroxylase. *Chemical Science* **2014**, 5, 1765-1771.
4. van Berkel, S. S.; Brem, J.; Rydzik, A. M.; Salimraj, R.; Cain, R.; Verma, A.; Owens, R. J.; Fishwick, C. W.; Spencer, J.; Schofield, C. J. Assay platform for clinically relevant metallo-beta-lactamases. *J. Med. Chem.* **2013**, 56, 6945-53.
5. Khan, A.; Lesniak, R. K.; Brem, J.; Rydzik, A. M.; Choi, H.; Leung, I. K. H.; McDonough, M. A.; Schofield, C. J.; Claridge, T. D. W. Development and application of ligand-based NMR screening assays for [gamma]-butyrobetaine hydroxylase. *MedChemComm* **2016**, 1270-1276.
6. Aguilar, J. A.; Nilsson, M.; Bodenhausen, G.; Morris, G. A. Spin echo NMR spectra without J modulation. *Chem. Commun. (Camb.)* **2012**, 48, 811-3.
7. Rydzik, A. M.; Leung, I. K. H.; Kochan, G. T.; Thalhammer, A.; Oppermann, U.; Claridge, T. D. W.; Schofield, C. J. Development and application of a fluoride-detection-based fluorescence assay for gamma-butyrobetaine hydroxylase. *Chembiochem* **2012**, 13, 1559-63.
8. Niesen, F. H.; Berglund, H.; Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.* **2007**, 2, 2212-21.
9. Leung, I. K. H.; Flashman, E.; Yeoh, K. K.; Schofield, C. J.; Claridge, T. D. W. Using NMR solvent water relaxation to investigate metalloenzyme-ligand binding interactions. *J. Med. Chem.* **2010**, 53, 867-75.

10. Mao, X.-a.; Guo, J.-x.; Ye, C.-h. Radiation damping effects on spin—lattice relaxation time measurements. *Chem. Phys. Lett.* **1994**, 222, 417-421.
11. Mao, X.-a.; Guo, J.-x.; Ye, C.-h. Competition between radiation damping and transverse relaxation effects on NMR signal intensities. *Chem. Phys. Lett.* **1994**, 218, 249-253.
12. Morris, G. A. Reference Deconvolution. In *eMagRes*, John Wiley & Sons, Ltd: **2007**.
13. Metz, K. R.; Lam, M. M.; Webb, A. G. Reference deconvolution: A simple and effective method for resolution enhancement in nuclear magnetic resonance spectroscopy. *Concepts Magn Reson* **2000**, 12, 21-42.

7 Appendix

7.1 Binding constant measurements by ^{1}H NMR (CPMG edited) direct ligand observation



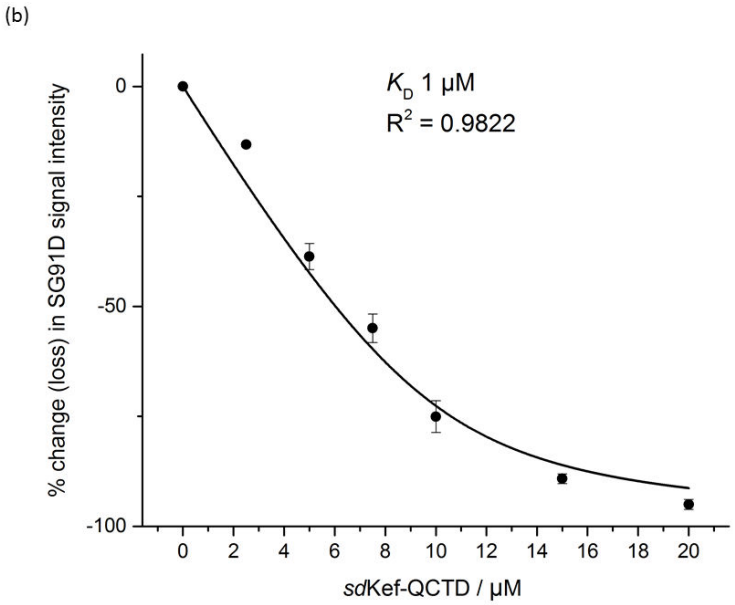
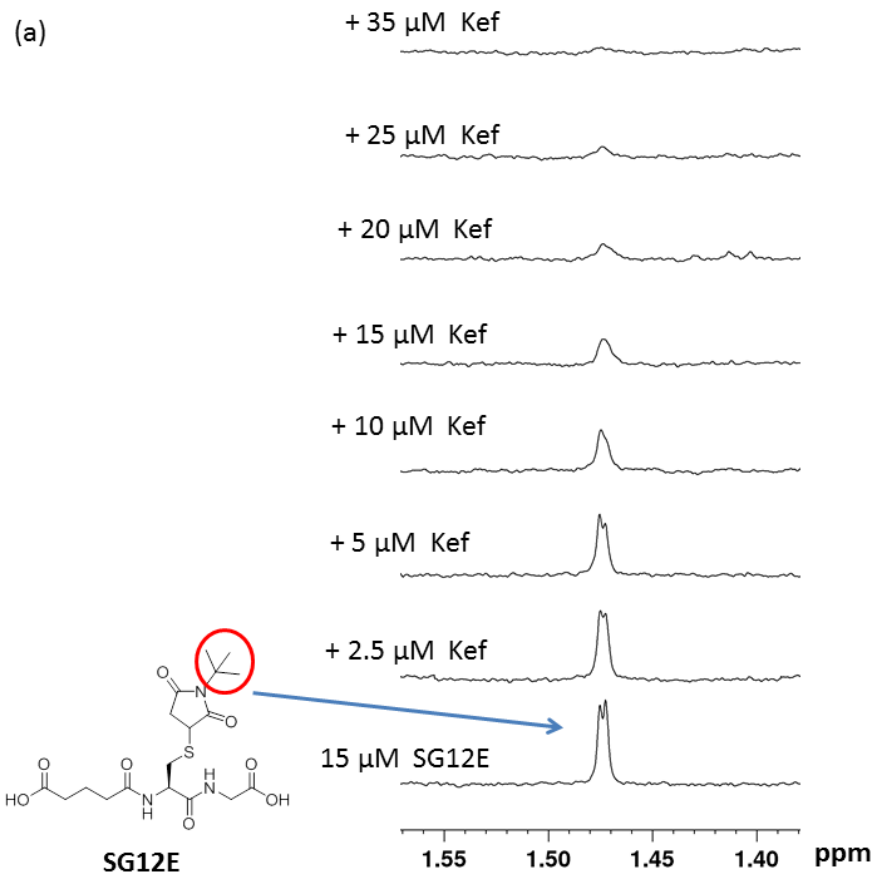


Figure 7.1 Measurement of the binding constant (K_D) of SG91D with *sd*Kef-QCTD by CPMG edited ^{1}H NMR direct ligand observation.
The ligand solution used was a racemic mixture. The asterisks represent the signal of ^{t}**Bu**OH added to the assay mixture as an internal NMR control. The K_D obtained was 1 ± 0.1 µM. The error bars represent standard deviation from three separate measurements.



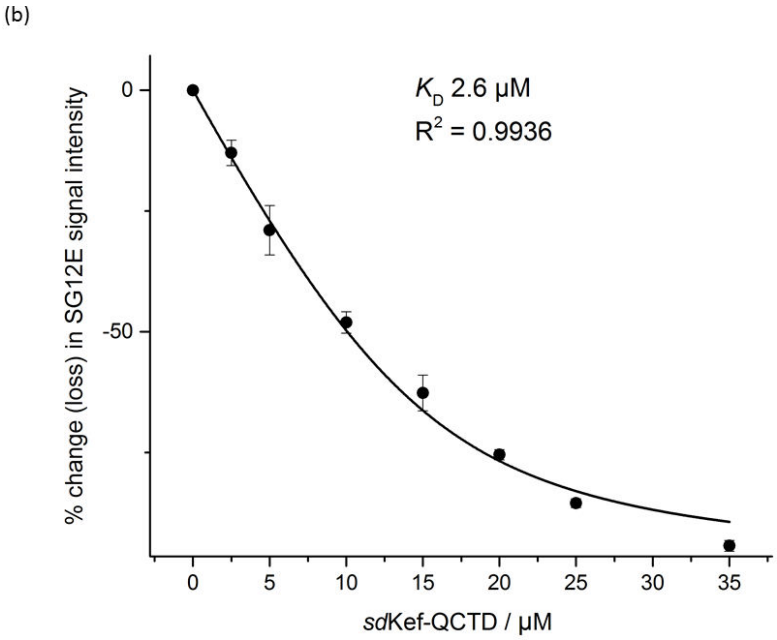
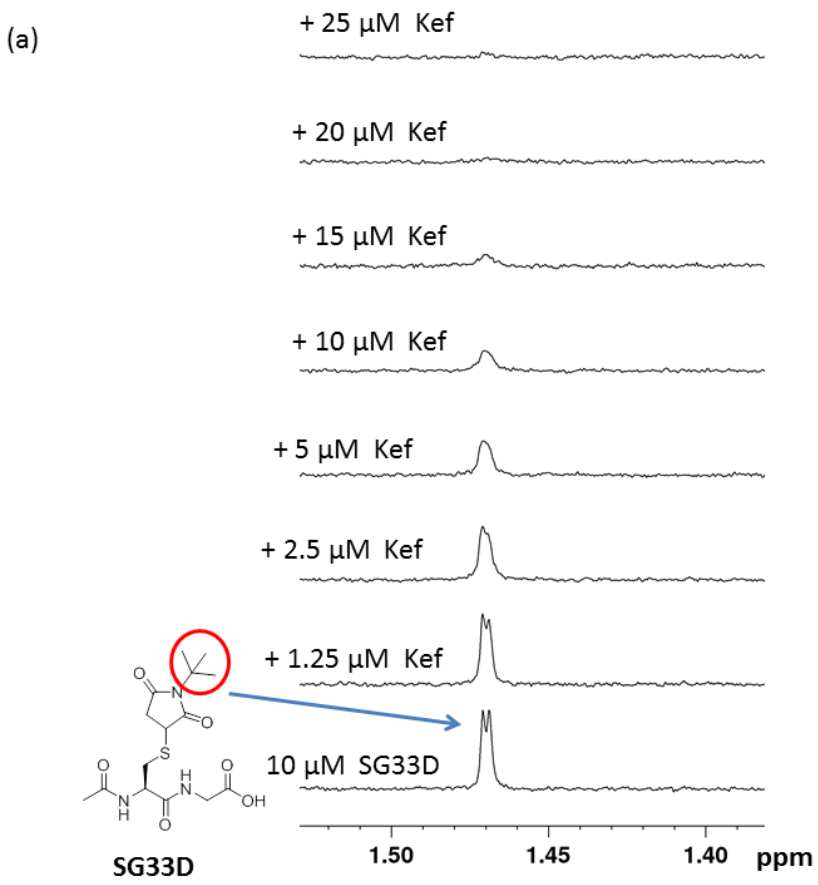


Figure 7.2 Measurement of the binding constant (KD) of SG12E with sdKef-QCTD by CPMG edited ^{1}H NMR direct ligand observation.

The K_D obtained was 2.6 ± 0.3 µM. The error bars represent standard deviation from three separate measurements. The ligand solution used was a racemic mixture.



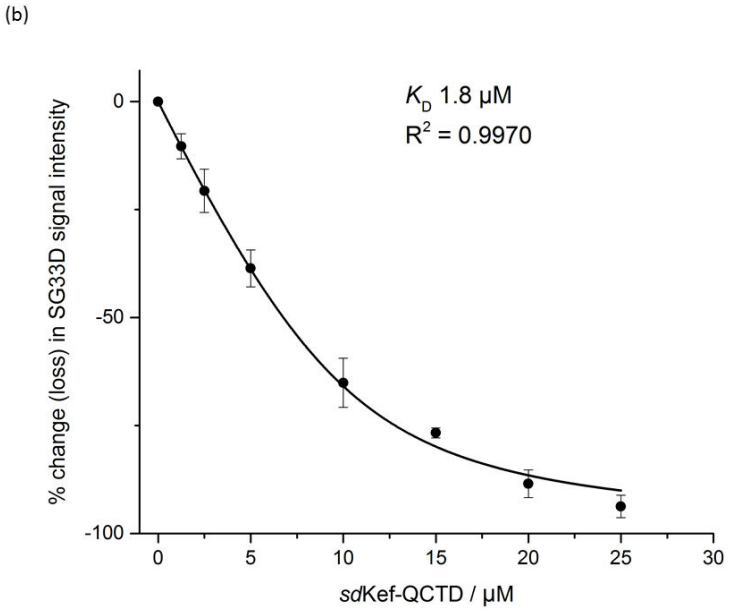
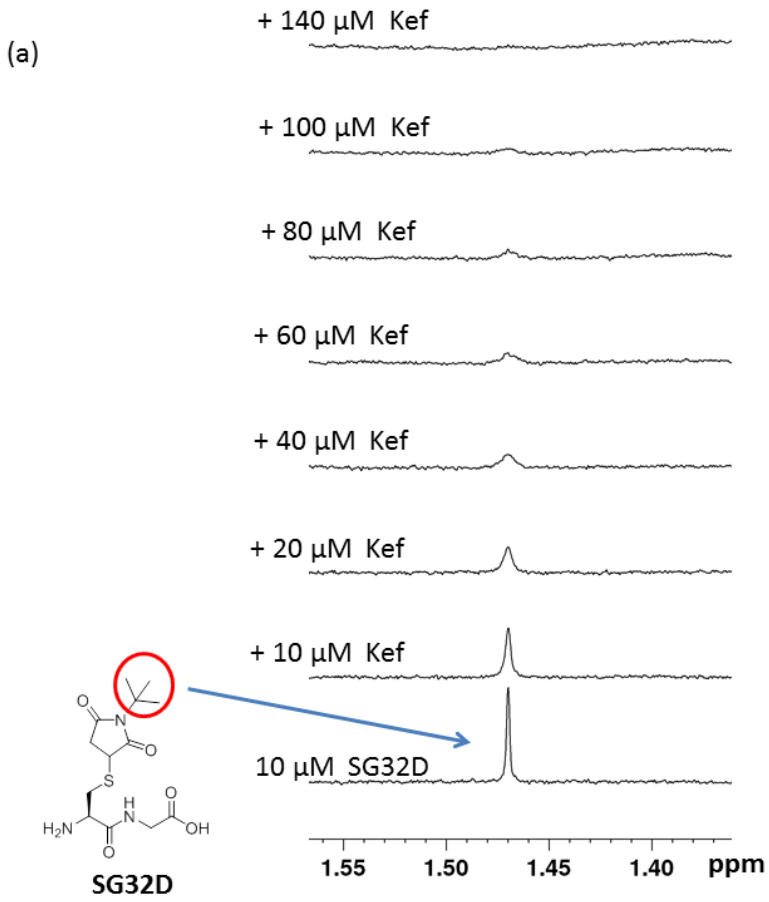


Figure 7.3 Measurement of the binding constant (K_D) of $SG_{33}D$ with *sd*Kef-QCTD by CPMG edited ^{1}H NMR direct ligand observation.

The K_D obtained was 1.8 ± 0.2 µM. The error bars represent standard deviation from three separate measurements. The ligand solution was used a racemic mixture.



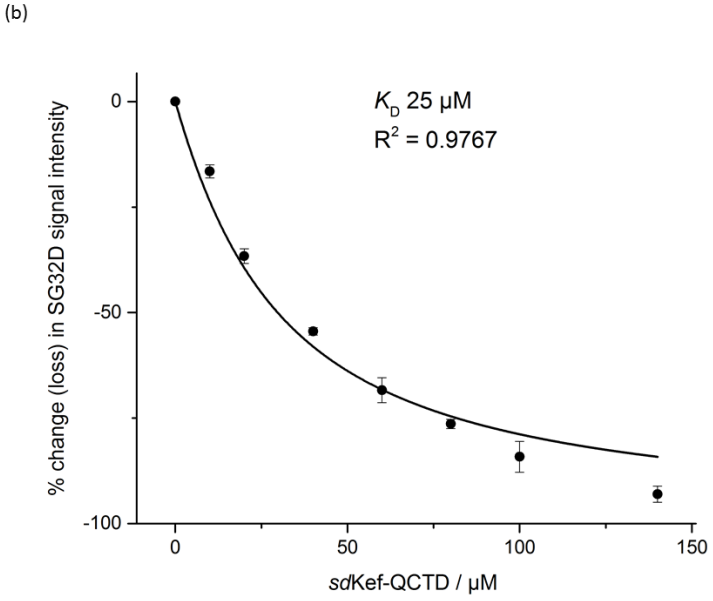
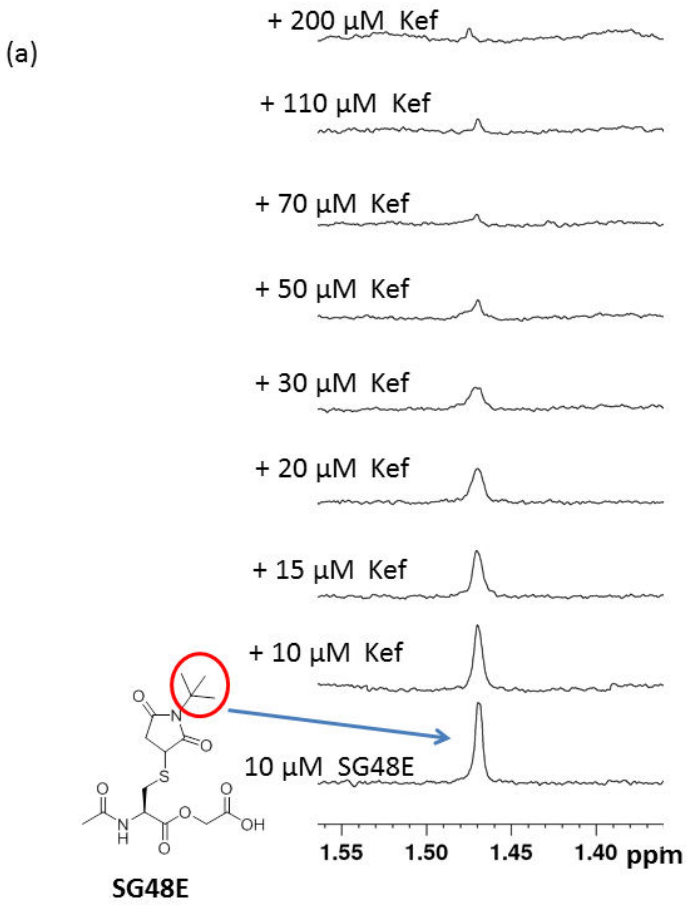


Figure 7.4 Measurement of the binding constant (K_D) of SG32D with *sd*Kef-QCTD by CPMG edited ^{1}H NMR direct ligand observation.

The K_D obtained was 25 ± 0.8 µM. The error bars represent standard deviation from three separate measurements. The ligand solution was used as a racemic mixture.



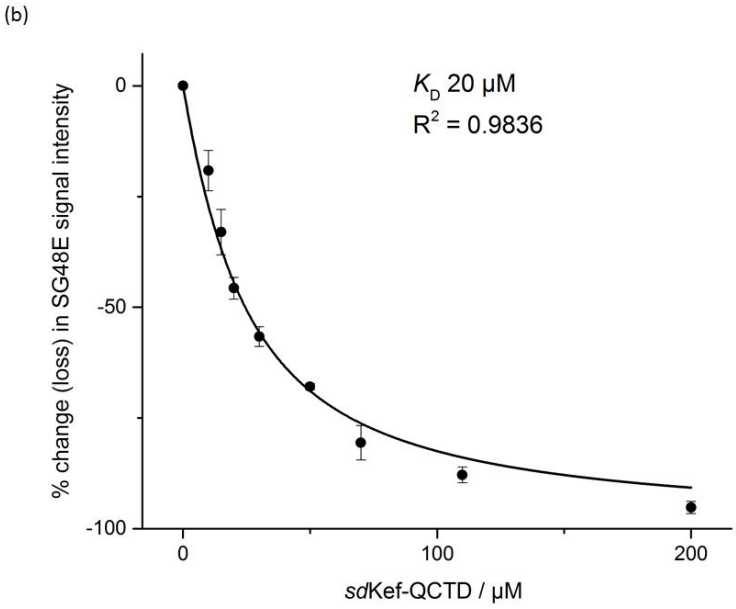


Figure 7.5 Measurement of the binding constant (K_D) of SG48E with *sd*Kef-QCTD by CPMG edited ^{1}H NMR direct ligand observation.

The K_D obtained was 20 ± 0.8 µM. The error bars represent standard deviation from three separate measurements. The ligand solution was used as a racemic mixture.

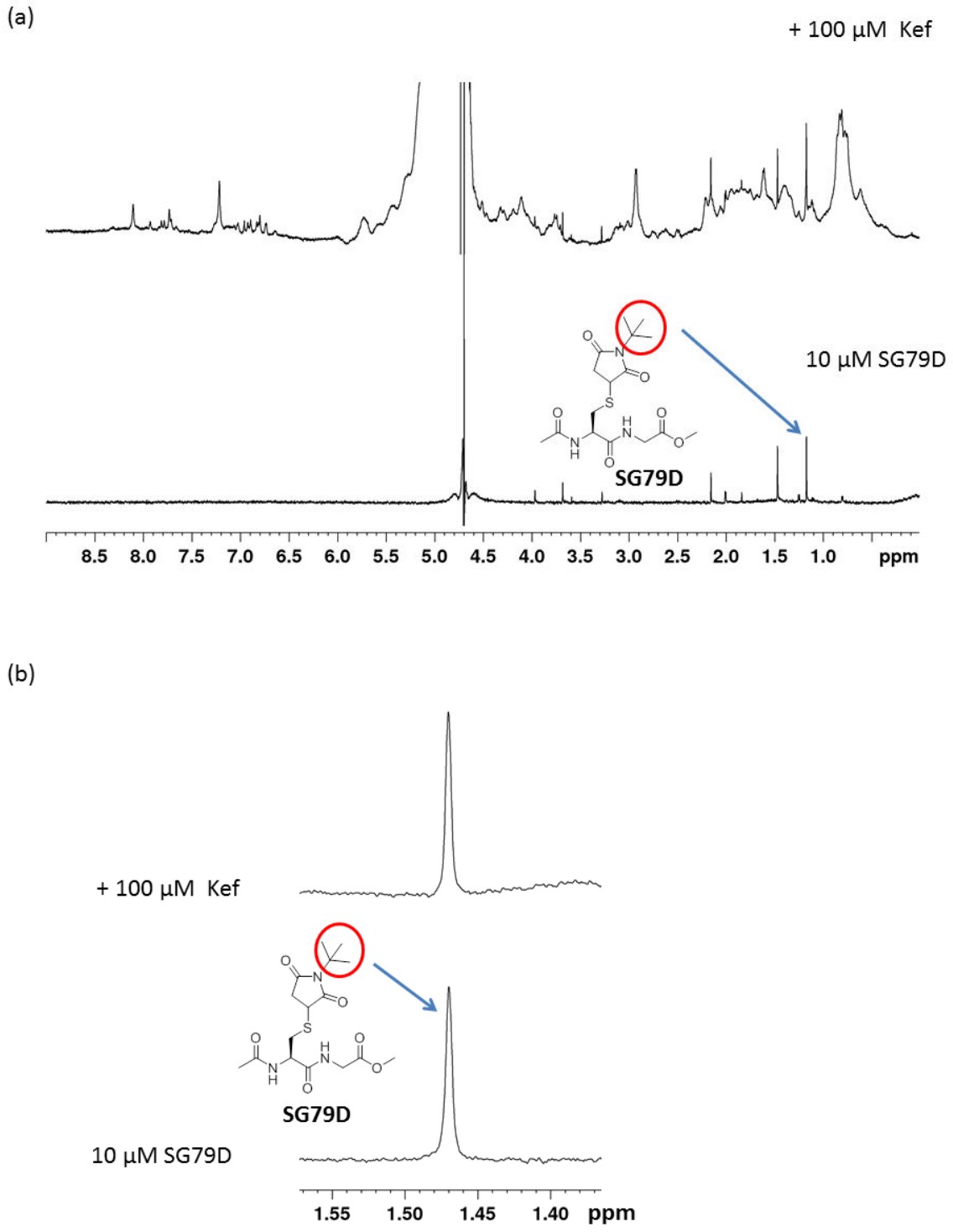
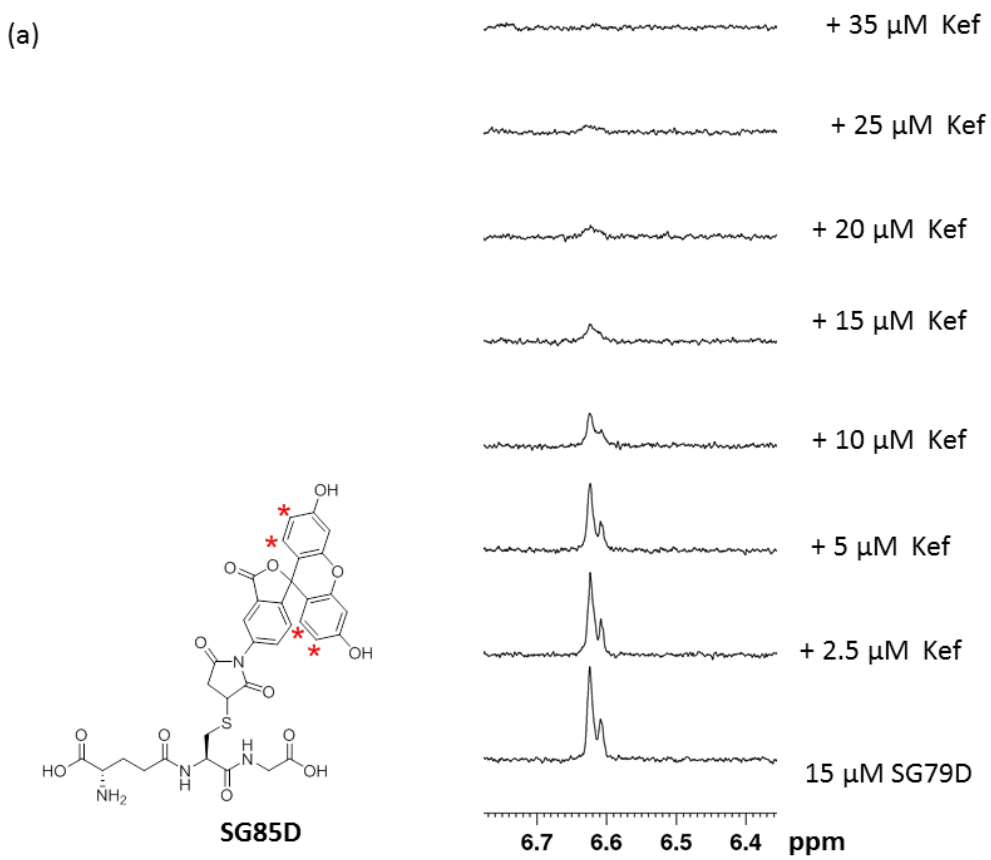


Figure 7.6 Studying the binding interaction of SG79D with *sd*Kef-QCTD by ^{1}H NMR direct ligand observation.
(a) Overlay of ^{1}H NMR spectra of a solution of SG79D without (bottom) and with *sd*Kef-QCTD. Signals of small molecules impurities are also seen in the spectra. After the addition of *sd*Kef-QCTD, the signal (^{t}Bu) of SG79D does not broaden or attenuate suggesting that it does not binding with the *sd*Kef-QCTD. (b) Corresponding CPMG edited ^{1}H NMR spectra. The ligand solution used was a racemic mixture.



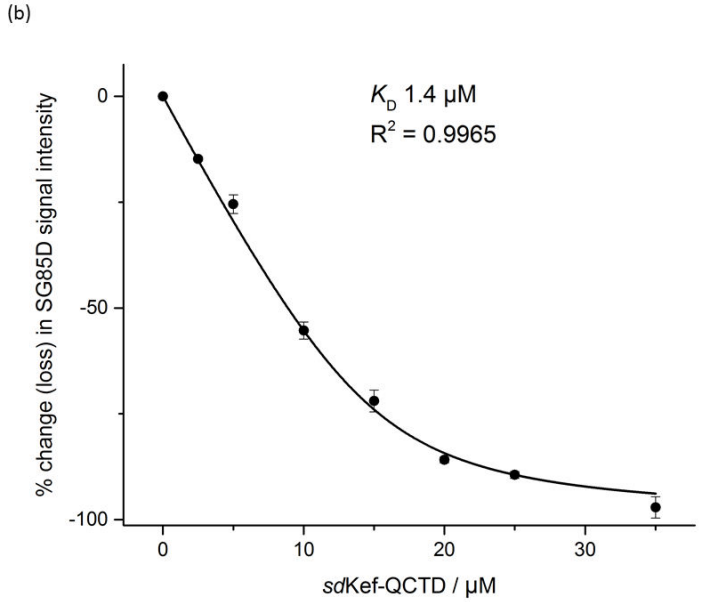
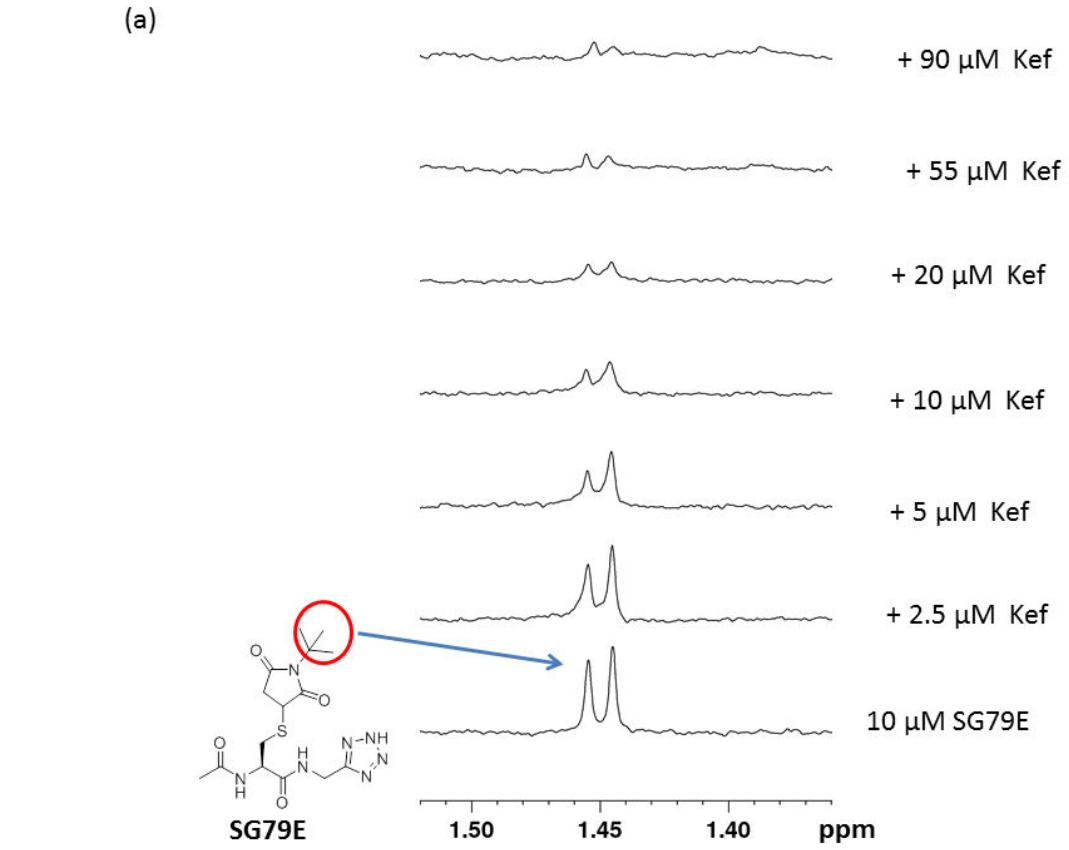


Figure 7.7 Measurement of the binding constant (K_D) of SG85D with *sd*Kef-QCTD by CPMG edited ^{1}H NMR direct ligand observation.

The ligand 4H (shown as asterisks) signal is followed for binding interaction The K_D obtained was 1.4 µM. The error bars represent standard deviation from three separate measurements. The ligand solution used was a racemic mixture with the average response measured.



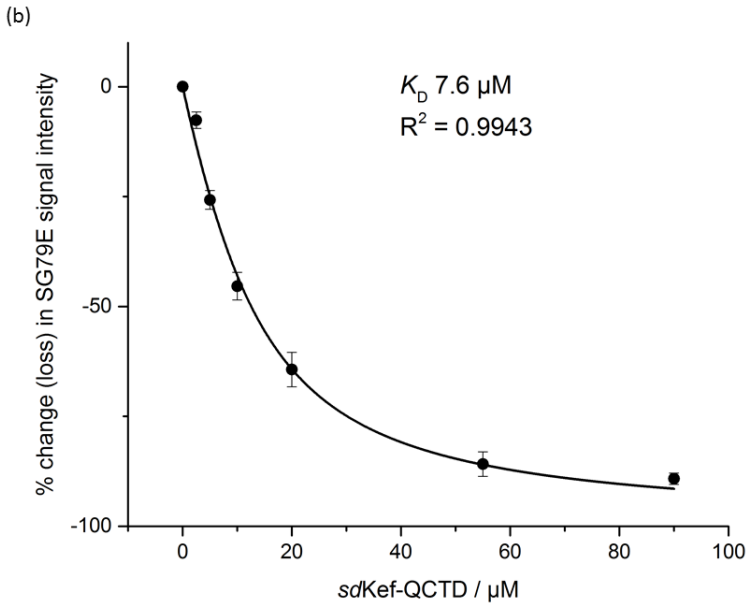
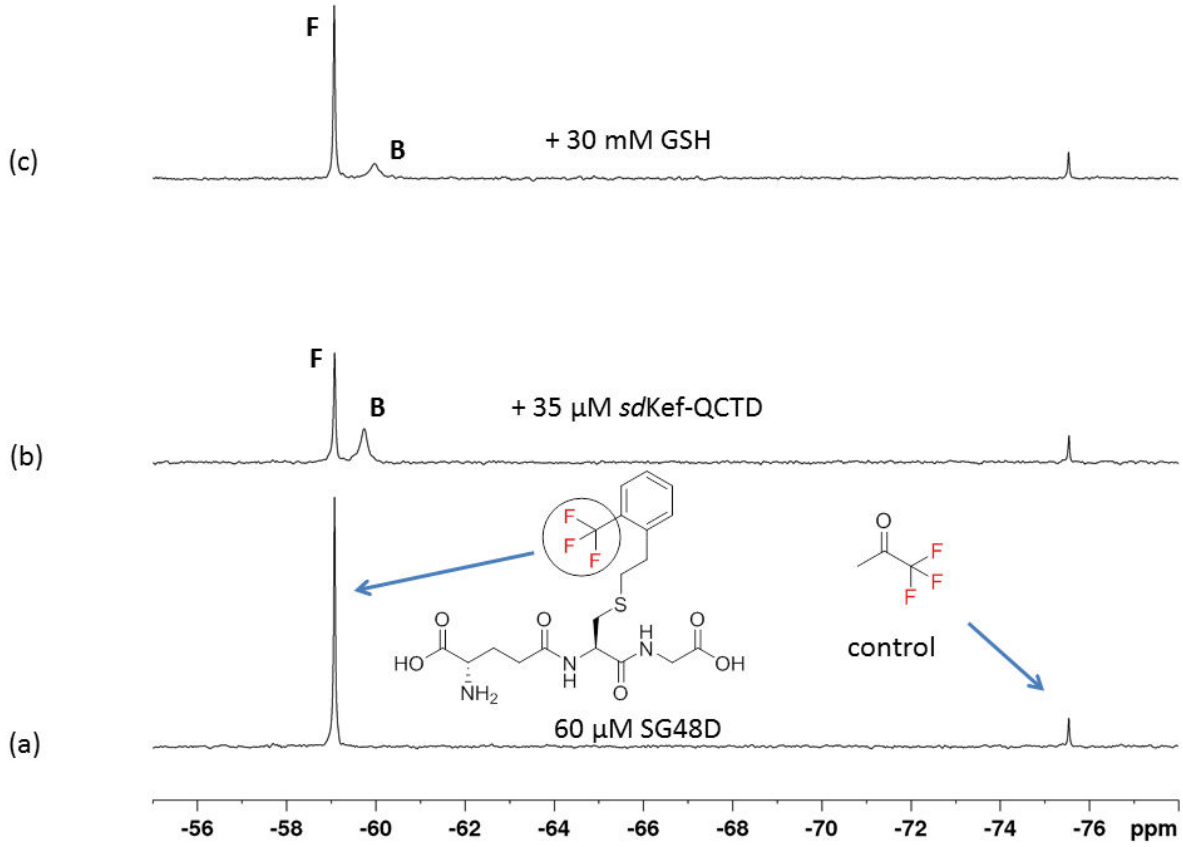


Figure 7.8 Measurement of the binding constant (K_D) of SG79E with *sd*Kef-QCTD by CPMG edited ^{1}H NMR direct ligand observation.

The K_D obtained was 7.6 ± 0.5 µM. The error bars represent standard deviation from three separate measurements. The ligand solution used was a racemic mixture with the average response measured.

7.2 Kef ligand screening by ^{19}F NMR reporter ligand displacement assay

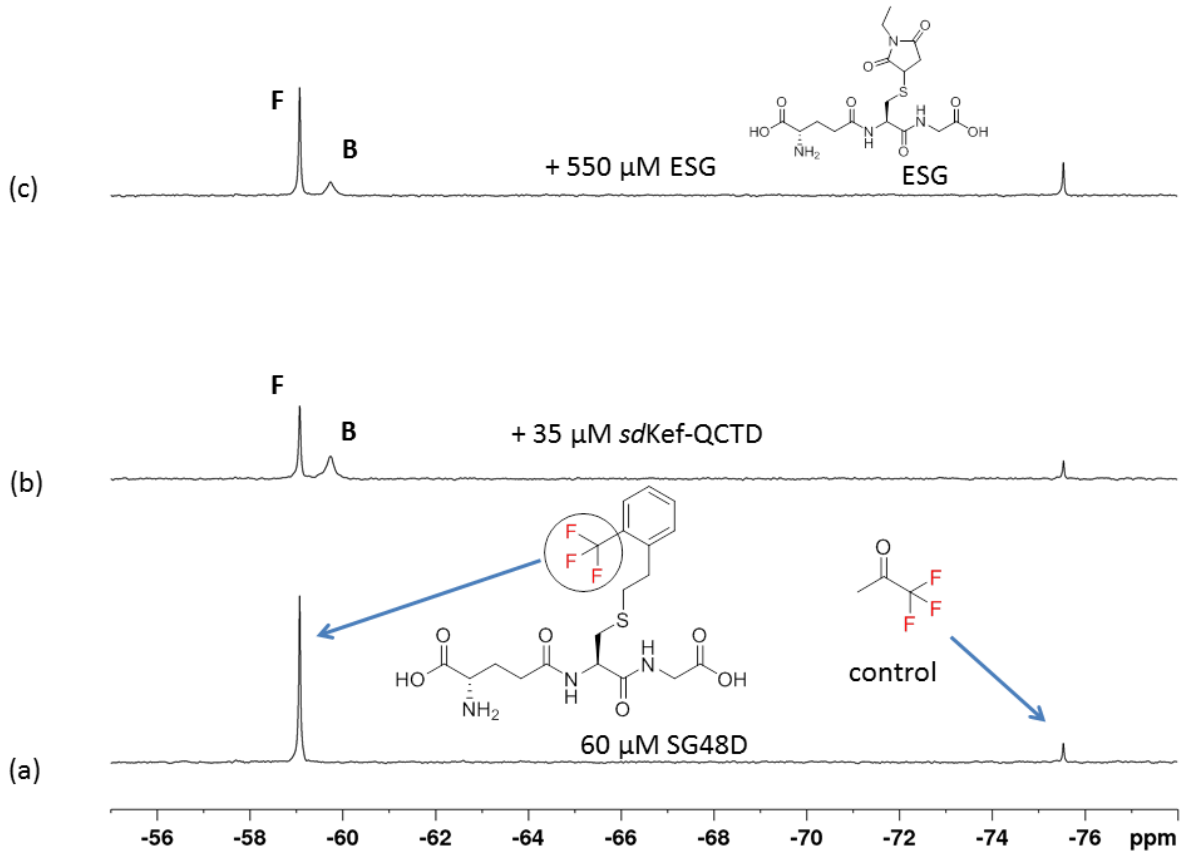


$[P_T]$ = Protein concentration used in the assay = 35 µM

$[L_T]$ = Ligand concentration used in the assay = 60 µM

Sample No.	[L] / µM (from peak area of **F**)	[PL] / µM (from peak area of **B**)	[P] = $[P_T]$-[PL] / µM (protein not complexed with SG48D)	$K_{D\ app} = \frac{[P][L]}{[PL]}$ / (µM)	$K_{D\ comp} = \frac{[C]K_{D\ rep}}{K_{D\ app} - K_{D\ rep}}$ K_D rep = 1.36 µM [C] = 30000 µM
1	42.10	17.90	17.10	40.23	1009
2	44.71	15.29	19.71	57.63	704
3	42.68	17.32	17.68	43.57	931
					Ave. K_D = 881 ± 159 / µM

Figure 7.9 Measurement of the binding constant (K_D) of glutathione (GSH) by ^{19}F NMR reporter displacement assay.

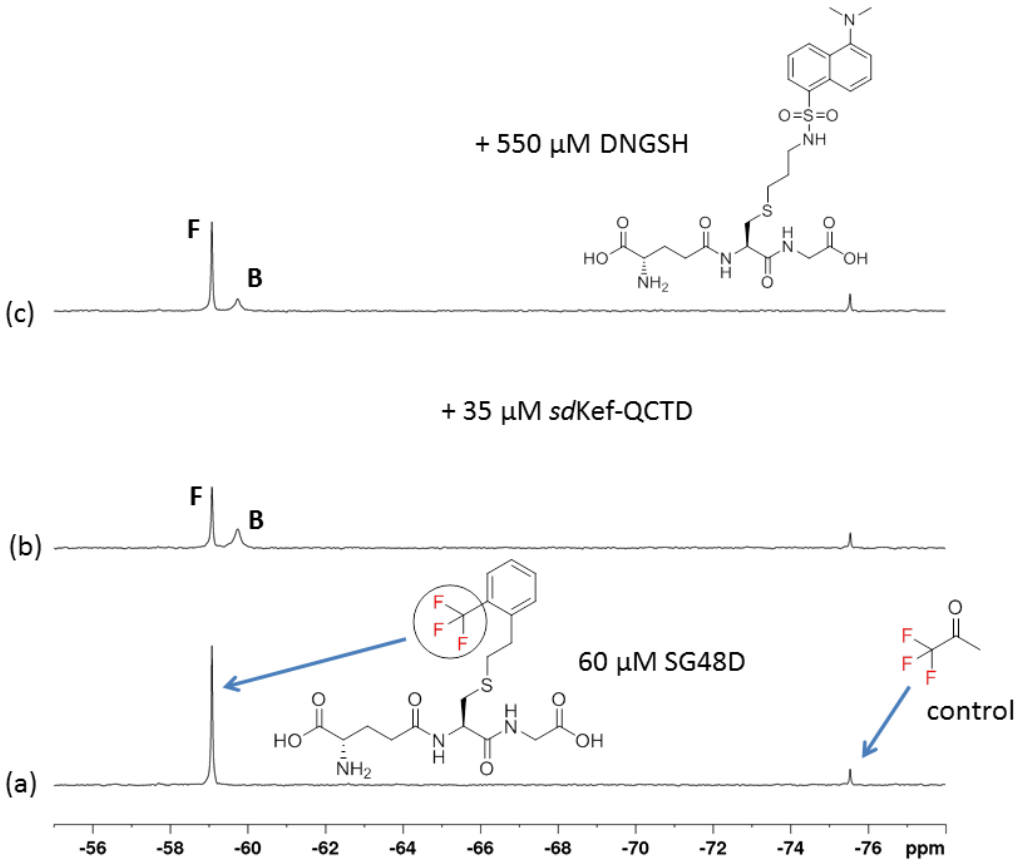


$[P_T]$ = Protein concentration used in the assay = 35 µM

$[L_T]$ = Ligand concentration used in the assay = 60 µM

Sample No.	[L] / µM (from peak area of **F**)	[PL] / µM (from peak area of **B**)	[P] = $[P_T]$-[PL] / µM (protein not complexed with SG48D)	$K_{D\ app} = \frac{[P][L]}{[PL]}$ / (µM)	$K_{D\ comp} = \frac{[C]K_{D\ rep}}{K_{D\ app} - K_{D\ rep}}$ K_D rep = 1.36 µM [C] = 550 µM
1	40.32	19.68	15.32	31.38	22.40
2	40.56	19.44	15.56	32.46	21.61
3	41.15	18.85	16.15	35.27	19.78
					Ave. K_D = 21 ± 1 / µM

Figure 7.10 Measurement of the binding constant (K_D) of ESG by ^{19}F NMR reporter displacement assay.

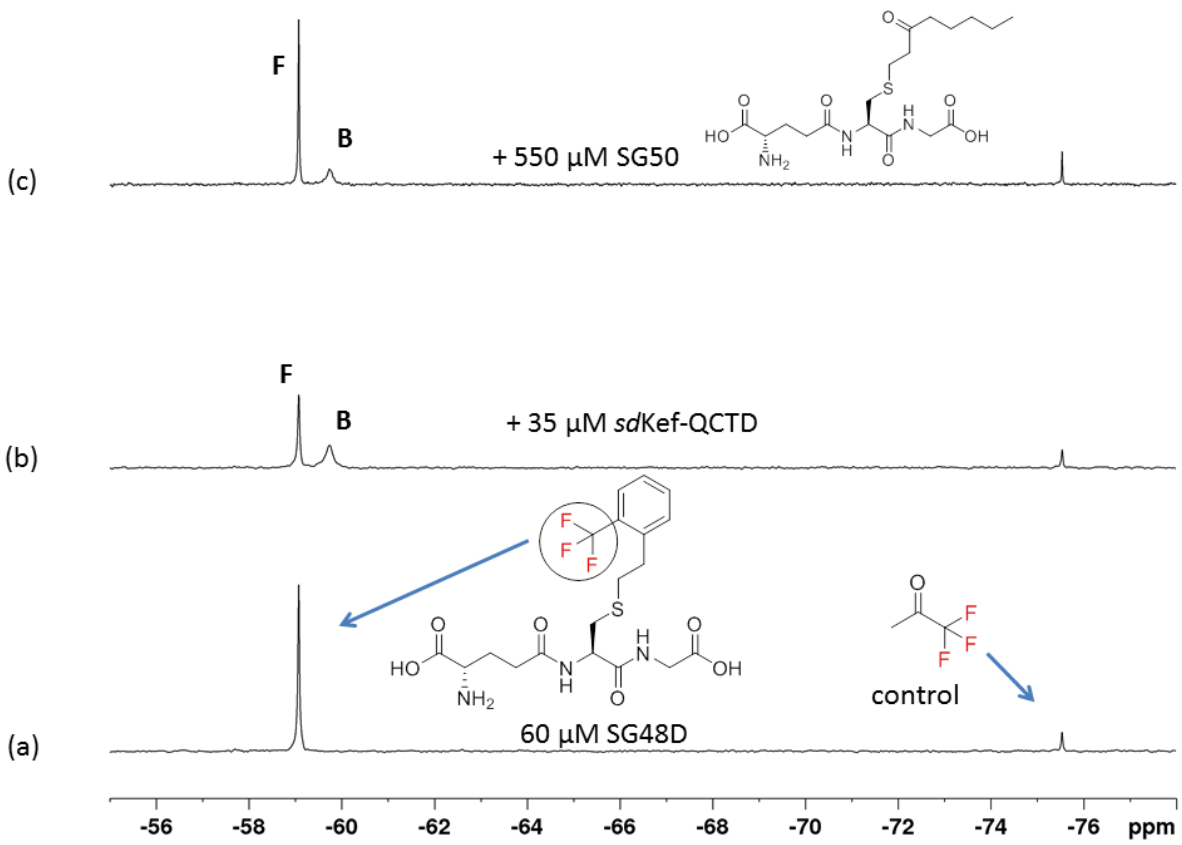


$[P_T]$ = Protein concentration used in the assay = 35 µM

$[L_T]$ = Ligand concentration used in the assay = 60 µM

Sample No.	[L] / µM (from peak area of **F**)	[PL] / µM (from peak area of **B**)	[P] = $[P_T]$-[PL] / µM (protein not complexed with SG48D)	$K_{D\ app} = \frac{[P][L]}{[PL]}$ / (µM)	$K_{D\ comp} = \frac{[C]K_{D\ rep}}{K_{D\ app} - K_{D\ rep}}$ K_D rep = 1.36 µM [C] = 550 µM
1	40.99	19.01	15.99	34.49	20.26
2	41.27	18.73	16.27	35.84	19.44
3	42.06	17.94	17.06	40	17.28
					Ave. K_D = 19 ± 1.5 / µM

Figure 7.11 Measurement of the binding constant (K_D) of DNGSH by ^{19}F NMR reporter displacement assay.

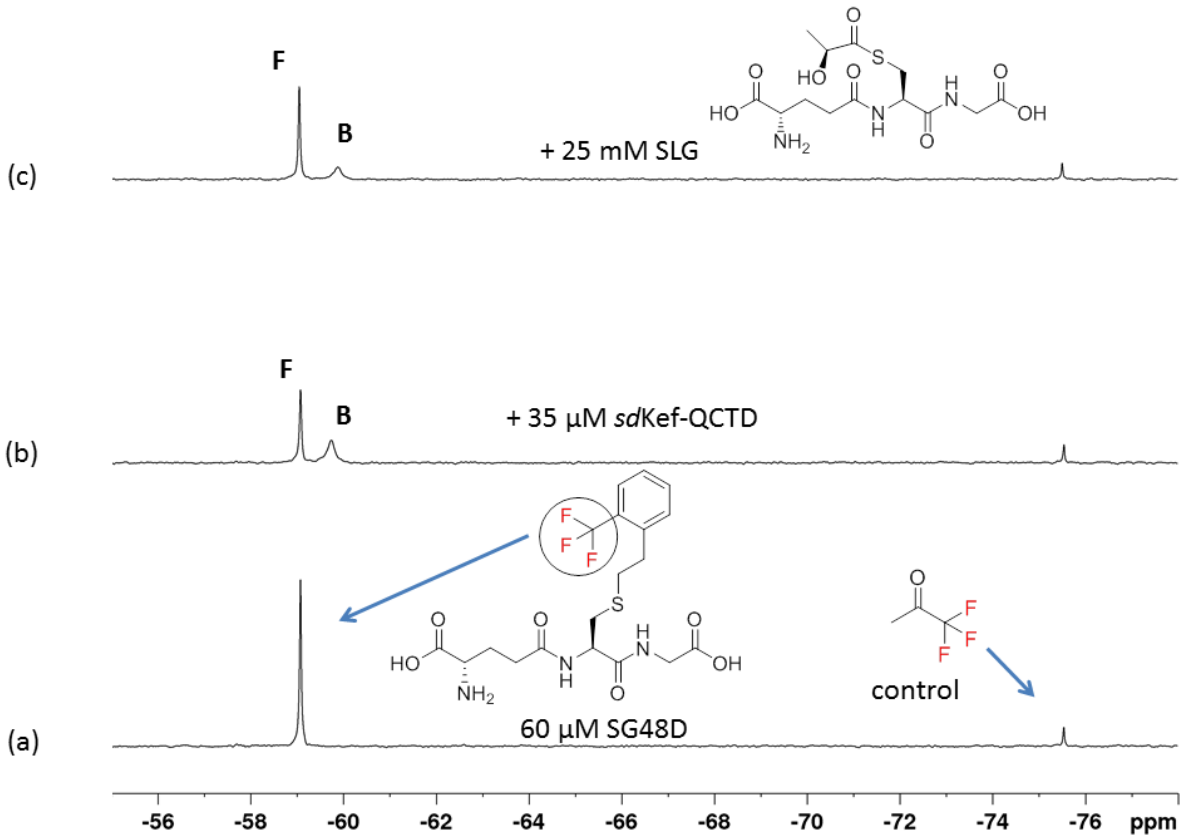


$[P_T]$ = Protein concentration used in the assay = 35 µM

$[L_T]$ = Ligand concentration used in the assay = 60 µM

Sample No.	[L] / µM (from peak area of **F**)	[PL] / µM (from peak area of **B**)	[P] = $[P_T]$-[PL] / µM (protein not complexed with SG48D)	$K_{D\ app} = \frac{[P][L]}{[PL]}$ / (µM)	$K_{D\ comp} = \frac{[C]K_{D\ rep}}{K_{D\ app} - K_{D\ rep}}$ K_D rep = 1.36 µM [C] = 550 µM
1	40.56	19.44	15.56	32.48	21.59
2	39.71	20.29	14.71	28.79	24.53
3	39.92	20.08	14.92	29.65	23.79
					Ave. K_D = 23 ± 1.5 / µM

Figure 7.12 Measurement of the binding constant (K_D) of SG50 by ^{19}F NMR reporter displacement assay.



$[P_T]$ = Protein concentration used in the assay = 35 µM

$[L_T]$ = Ligand concentration used in the assay = 60 µM

Sample No.	[L] / µM (from peak area of **F**)	[PL] / µM (from peak area of **B**)	[P] = $[P_T]$-[PL] / µM (protein not complexed with SG48D)	$K_{D\ app} = \frac{[P][L]}{[PL]}$ / (µM)	$K_{D\ comp} = \frac{[C]K_{D\ rep}}{K_{D\ app} - K_{D\ rep}}$ K_D rep = 1.36 µM [C] = 25000 µM
1	39.11	20.89	14.11	26.40	1281.90
2	40.19	19.81	15.19	30.82	1097.94
3	39.87	20.13	14.87	29.45	1148.96
					Ave. K_D = 1176 ± 95 / µM

Figure 7.13 Measurement of the binding constant (K_D) of SLG by ^{19}F NMR reporter displacement assay.

7.3 Publications

Parts of the work described in this Thesis have been published as follows:

1. A. Khan, R. K. Leśniak, J. Brem, A. M. Rydzik, H. Choi, I. K. H. Leung, M. A. McDonough, C. J. Schofield and T. D. W. Claridge; Development and applications of ligand-based NMR screening assays for γ-butyrobetaine hydroxylase, *Med. Chem. Commun.,* **2016**, DOI: 10.1039/c6md00004e. (**Hot Article**)

2. Kamps, J. J.; Khan, A.; Choi, H.; Lesniak, R. K.; Brem, J.; Rydzik, A. M.; McDonough, M. A.; Schofield, C. J.; Claridge, T. D.; Mecinovic, J. Cation-pi Interactions Contribute to Substrate Recognition in gamma-Butyrobetaine Hydroxylase Catalysis. *Eur. Chem. J.;* **2016**, 22, 1270-1276. DOI: 10.1002/chem.201503761. (**Hot Article**)

MedChemComm

RESEARCH ARTICLE

View Article Online
View Journal

Cite this: DOI: 10.1039/c6md00004e

Development and application of ligand-based NMR screening assays for γ-butyrobetaine hydroxylase†‡

A. Khan,§ R. K. Leśniak,§ J. Brem, A. M. Rydzik, H. Choi, I. K. H. Leung,¶ M. A. McDonough, C. J. Schofield* and T. D. W. Claridge*

Received 6th January 2016,
Accepted 15th February 2016

DOI: 10.1039/c6md00004e

www.rsc.org/medchemcomm

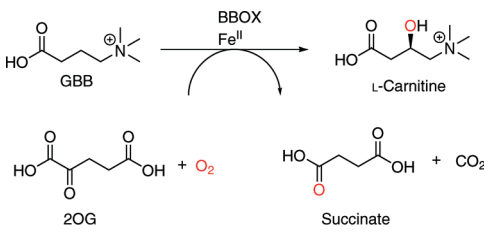
γ-Butyrobetaine hydroxylase (BBOX) is a 2-oxoglutarate (2OG) dependent oxygenase that catalyses the stereoselective C-3 hydroxylation of γ-butyrobetaine (GBB) to give L-carnitine. L-Carnitine is involved in fatty acid metabolism in all animals and in some prokaryotes, and BBOX is a current drug target for the treatment of myocardial infarction. We describe the development and application of ^{1}H NMR GBB/2OG reporter based assays employing paramagnetic relaxation enhancement to monitor inhibitor binding to the BBOX active site. In a single experiment, the method assesses inhibitors for competitive binding with 2OG or GBB, or both. The method was exemplified with a set of isoquinoline-based inhibitors; the results reveal structure–activity relationships that were not predicted from crystallographic studies, with some inhibitors competing 2OG only and some competing both 2OG and GBB. The method will also be applicable to work on the inhibition of other 2OG oxygenases.

Introduction

Ferrous ion (Fe^{II}) and 2-oxoglutarate (2OG) dependent oxygenases play multiple biologically important roles in humans and other animals, including in lipid metabolism.[1,2] 2OG oxygenases play a central role in the metabolism of the chlorophyll metabolite phytanic acid[1] and in the biosynthesis of L-carnitine,[1,2] which is required for fatty acid transport into mitochondria.[3] Two steps in L-carnitine biosynthesis are catalysed by 2OG oxygenases, *i.e.* the C-3 hydroxylation of trimethyllysine (TML) and of γ-butyrobetaine (GBB), which are catalysed by trimethyllysine hydroxylase (TMLH) and γ-butyrobetaine hydroxylase (BBOX), respectively (Scheme 1).[3] BBOX (and, maybe, TMLH) are inhibited by mildronate (THP or Met88),[4] which is given to patients after myocardial infarction to suppress fatty acid metabolism.[5] As shown by NMR and MS studies, mildronate is a competitive BBOX substrate, undergoing oxidation and subsequent degradation including *via* a Stevens type rearrangement;[6] it may also inhibit uptake of dietary L-carnitine. There is, therefore, interest in the development of other types of BBOX inhibitors, for example compounds acting as reversibly binding inhibitors that are competitive with 2OG and/or GBB, but which are not substrates.

We are interested in developing solution-based NMR methods to complement the extensive crystallographic studies on 2OG oxygenases, in order to help enable the development of clinically useful inhibitors.[2] Work using NMR spectroscopy on the 2OG dependent hypoxia inducible factor (HIF) prolyl hydroxylase PHD2, has shown that crystallographically observed inhibitor binding modes do not always reflect those occurring in solutions[7] and that inhibitors which might be expected to compete with substrates do not always do so.[7–9]

Direct ligand observation, monitoring the attenuation or broadening of ligand NMR resonances in the presence of a target protein, can be an efficient method for analysing the



Scheme 1 BBOX catalyses the stereoselective C-3 hydroxylation of GBB to give L-carnitine. The co-substrates are 2OG and oxygen and the co-products are succinate and CO_2.

Department of Chemistry, Chemistry Research Laboratory, University of Oxford, 12 Mansfield Road, Oxford, OX1 3TA, UK.
E-mail: christopher.schofield@chem.ox.ac.uk, tim.claridge@chem.ox.ac.uk;
Fax: +44 (0)1865 285002; Tel: +44 (0)1865 275658

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c6md00004e

‡ The authors declare no competing interests.

§ These authors contributed equally to this work.

¶ Current address: School of Chemical Sciences, The University of Auckland, Private Bag 92019, Auckland 1142, New Zealand.

binding of ligands to a protein and may employ conventional ^{1}H NMR[10–12] and/or transverse relaxation-edited ^{1}H NMR.[13,14] Site specific ligand binding may also be assessed through ligand competition, *i.e.* by monitoring the sharpening and recovery of the NMR resonances of displaced ligands, often referred to as spy or reporter molecules.[15]

We describe the development of an NMR assay for recombinant bacterial BBOX from *Pseudomonas sp.* AK1 (psBBOX)[16] that simultaneously uses both 2OG and GBB as reporter molecules. We validate the method using a set of isoquinoline-based derivatives, which are established inhibitors of 2OG oxygenases, including human BBOX (hBBOX).[2,17] The results reveal that even within the same series of 2OG oxygenase inhibitors, unexpected structure–activity relationships can emerge. The results should help in efforts to develop new BBOX inhibitors and to promote the use of NMR spectroscopy in 2OG oxygenase medicinal chemistry. In developing the assay we employed psBBOX rather than hBBOX, since it can be readily prepared on a large scale and is thought to be structurally closely related to hBBOX.[16]

Several crystal structures have been reported for hBBOX,[18–20] although none for psBBOX. hBBOX dimerises in an unusual manner using its N-terminal zinc binding domain to interact with the oxygenase domain of the other monomer in the dimer.[19,20] The general active site architecture of BBOX is very similar to those of other 2OG oxygenases with the active site Fe^{II} (to which the oxalyl group of 2OG binds in a bidentate manner) being ligated by three protein residues.[2] GBB binds *via* interactions involving both its trimethylammonium[21] and carboxylate groups (Fig. 1). The trimethylammonium group is located in an aromatic cage formed by the side chains of Tyr-177, Tyr-194, Trp-181, Tyr-205 and Tyr-366, while the carboxylate group is positioned to interact with the side chains of Asn-292, Asn-191 and the backbone amide of Tyr-205 (Fig. 1).

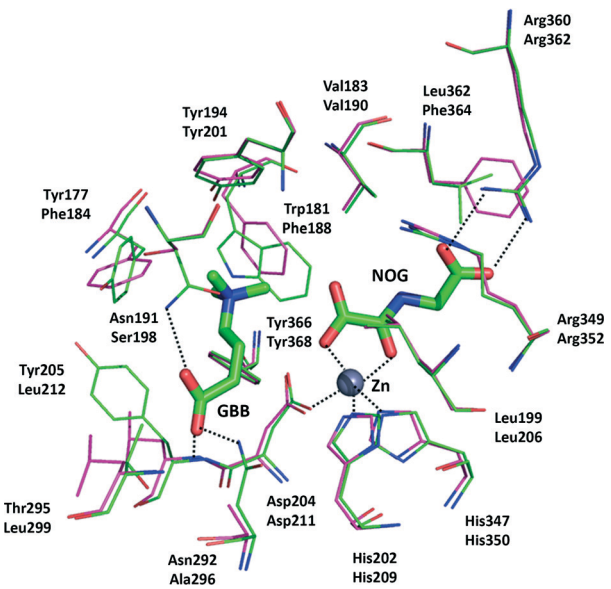


Fig. 1 View from a hBBOX crystal structure (green sticks, upper residues numbers, PDB ID: 3O2G)[20] and psBBOX model (pink sticks, lower residue numbers) showing binding of GBB and *N*-oxalylglycine (NOG), an unreactive 2OG analogue. Zn^{II} replaces native Fe^{II}.

Results and discussion

Initially, we optimised the assay conditions to enable detection of binding of both GBB and 2OG (as reporter molecules) to psBBOX in a single Carr-Purcell-Meiboom-Gill (CPMG)-edited ^{1}H NMR ligand-observe experiment monitoring line shape changes in the presence of psBBOX[13,14] (Fig. 2). The CPMG-edited ^{1}H NMR experiments were used in order to improve NMR sensitivity to ligand binding by enhancing signal attenuation of broadened resonances. To prevent psBBOX-catalysed GBB hydroxylation and uncoupled 2OG decarboxylation into succinate, we replaced the Fe^{II} with the 'non-catalytic' metals Zn^{II} or Mn^{II}. We observed that in the presence of psBBOX and Zn^{II}, the GBB and 2OG signals were broadened and attenuated, demonstrating their fast-exchange binding interaction with psBBOX (Fig. 2a). However, despite adding high psBBOX concentrations, only relatively moderate broadening was observed. We envisaged that greater sensitivity to ligand binding might be achieved by employing paramagnetic relaxation enhancement (PRE).[22,23] For this purpose either a paramagnetic metal co-factor such as Co^{II} (ref. 24) or Mn^{II} (ref. 25, 26) or a protein covalently tagged with a spin label[27] can be used. When a ligand binds to a protein containing a paramagnetic centre (within ~15 Å), its nuclear spin relaxation rates are enhanced and its NMR signals are significantly broadened. Accordingly, when paramagnetic Mn^{II}, rather than diamagnetic Zn^{II}, was added to the BBOX assay, significant increases in the line widths of both GBB and 2OG were observed in the presence of substantially lower amounts of psBBOX as compared to the Zn^{II} assay (Fig. 2b). Control experiments (without psBBOX) revealed that at the concentrations employed, free Mn^{II} in solution did not affect the GBB line shape; it did, however, broaden the resonances of 2OG since this directly ligates Mn^{II}. Despite this broadening, the resonance intensities (^{1}H NMR peak areas) remained similar to those prior to the addition of the metal (Fig. S1†). In a separate control experiment, when the GBB/2OG/Mn^{II}/psBBOX sample was heated (373 K, 5 min) to cause BBOX denaturation, recovery of the reporter signals was observed (Fig. S2†), implying that the enhanced broadening of the reporter signals upon the addition of psBBOX was due to binding to the protein. The binding affinities of GBB for both Zn^{II} and Mn^{II} forms of BBOX in the presence of 2OG were similar with K_D values of 5 ± 1 (Fig. S3†) and 4 ± 1 μM (Fig. S4†), respectively, as determined by titrating *apo*-psBBOX into a solution with constant GBB and metal concentrations. The Mn^{II} system was therefore chosen for further screening of inhibitors by the NMR dual-reporter displacement assay. In this, competition from a ligand for the GBB and/or 2OG binding sites could be established through the observed recovery of the reporter ligand

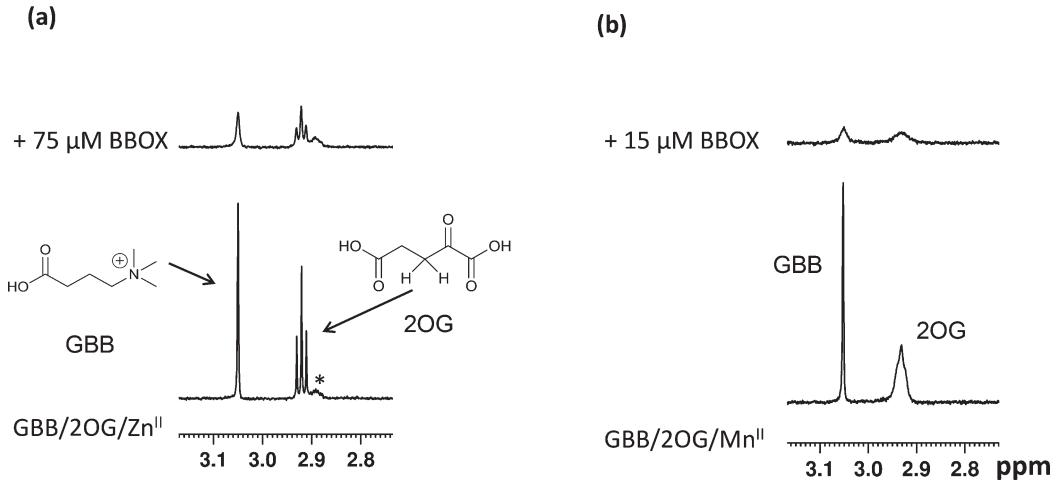


Fig. 2 Monitoring the binding of GBB and 2OG with psBBOX as determined by NMR direct ligand-observation. The GBB (Me_3 singlet at 3.05 ppm) and 2OG (CH_2 triplet at 2.91 ppm) signals from CPMG edited ^{1}H NMR experiments are shown. (a) GBB and 2OG binding to BBOX-Zn^{II} (b) GBB and 2OG binding to BBOX-Mn^{II}. The assay mixture contains 25 μM GBB, 300 μM 2OG, 150 μM M^{II}, 80 mM KCl, in 50 mM Tris-D_{11} buffer, pH 7.5, in D_2O. The asterisk indicates an impurity.

resonances upon addition of the inhibitor to the GBB/2OG/Mn^{II}/psBBOX sample.

A set of 15 isoquinoline-based compounds[17] were screened for their binding to psBBOX using simultaneously both GBB and 2OG as reporter molecules (Fig. 3). All these ligands shared the same isoquinoline core skeleton, but differed in their amino-acid derived side chains (Fig. 4). Exemplary results from the assay are presented in Fig. 3 for two of these isoquinoline-based ligands, one possessing the Gly (**1**) and the other the L-Trp (**11**) derived side chains. The spectra of GBB and 2OG with Mn^{II} were first recorded without and with psBBOX (Fig. 3a and b), then the isoquinoline-ligand was added to the latter sample. With **1**, only the signals of 2OG were observed to recover (Fig. 3c), implying that **1** competes efficiently with 2OG, but not GBB. In contrast, with **11**, we observed recovery of both 2OG and GBB signals (Fig. 3d), implying competition with both 2OG and GBB. The combined results of this assay for all the isoquinoline-ligands tested (at 25 μM) are summarised in Fig. 4.

The results imply that the isoquinoline-ligands containing the Gly (**1**) or Ala (both D, **2** and L, **3**) derived side chains compete with 2OG, but not GBB. The full displacement of 2OG by these ligands suggests that they are likely strong binders of psBBOX. The results also imply that binding of these compounds enhanced GBB binding relative to that in the presence of 2OG. On increasing the size of the steric bulk of the side chain, *i.e.* for those derived from Val (**4**, **5**) and Leu (**6**, **7**), the ligands appeared to bind less favourably as demonstrated by lower levels, or even a lack, of 2OG displacement compared to **1**, **2** and **3**. The apparent displacement of both 2OG and GBB by **5** and **7**, albeit at low levels, suggests that they attenuate GBB binding. Notably, binding was observed only for the L-forms (**5**, **7**), but not for the D-forms (**4**, **6**), of the Val and Leu derived compounds. This preference for binding by the L-enantiomer was preserved through most of the results where substantial reporter displacement was observed. These results encouraged us to analyse compounds with more hydrophobic or polar side chains (**7**–**15**). Ligands with the D- and L-Asp and Glu derived side chains (**12**–**15**) were mostly weak 2OG competitors relative to the inhibitors with hydrophobic side chains (**5**, **7**, **9** and **11**). **12** and **15** competed weakly with both 2OG and GBB, whereas, **13** competed with only 2OG, suggesting differences in binding modes, though care should be taken in interpreting results with weakly binding ligands. The results with the more bulky Phe (**8**, **9**) and Trp (**10**, **11**) derived side chain ligands were striking. The L-enantiomers **9** and **11** displaced both 2OG and GBB clearly contrasting with the ligands with smaller side chains (**1**–**3**).

The NMR binding results were then compared with those for the same compounds obtained using an intrinsic fluorescence-based binding assay.[17,28] The results of fluorescence based assay (Table 1) demonstrate that ligands with Gly **1**, D-Ala **2**, L-Ala **3**, L-Phe **9** and L-Trp **11** are the strongest binders, *i.e.* they agree with the ^{1}H NMR assay results. They also support the proposal that in general the L-enantiomers bind more tightly than the D-enantiomers. The combined NMR and fluorescence assay results reveal that the level of 2OG displacement correlates with the binding affinity, notably **1**, **2**, **3**, **9** and **11** all fully displace 2OG (within our limits of detection) and have approximately similar binding affinities (K_D = 22, 20, 15, 26 and 29 μM, respectively). Similarly, ligands **5**, **7** and **15**, all compete relatively moderately with 2OG compared to ligands **1**, **2**, **3**, **9** and **11**, and have similar binding affinities (K_D = 55, 55 and 64 μM respectively), which are weaker than **1**, **2**, **3**, **9** and **11**. Compounds **12** and **13**, which appear to displace 2OG to a similar extent, also have similar affinities (K_D = 101 and 91 μM respectively).

We then carried out catalytic turnover assays employing ^{1}H NMR[20] (Fig. 5) and fluoride ion release[29] assays to study

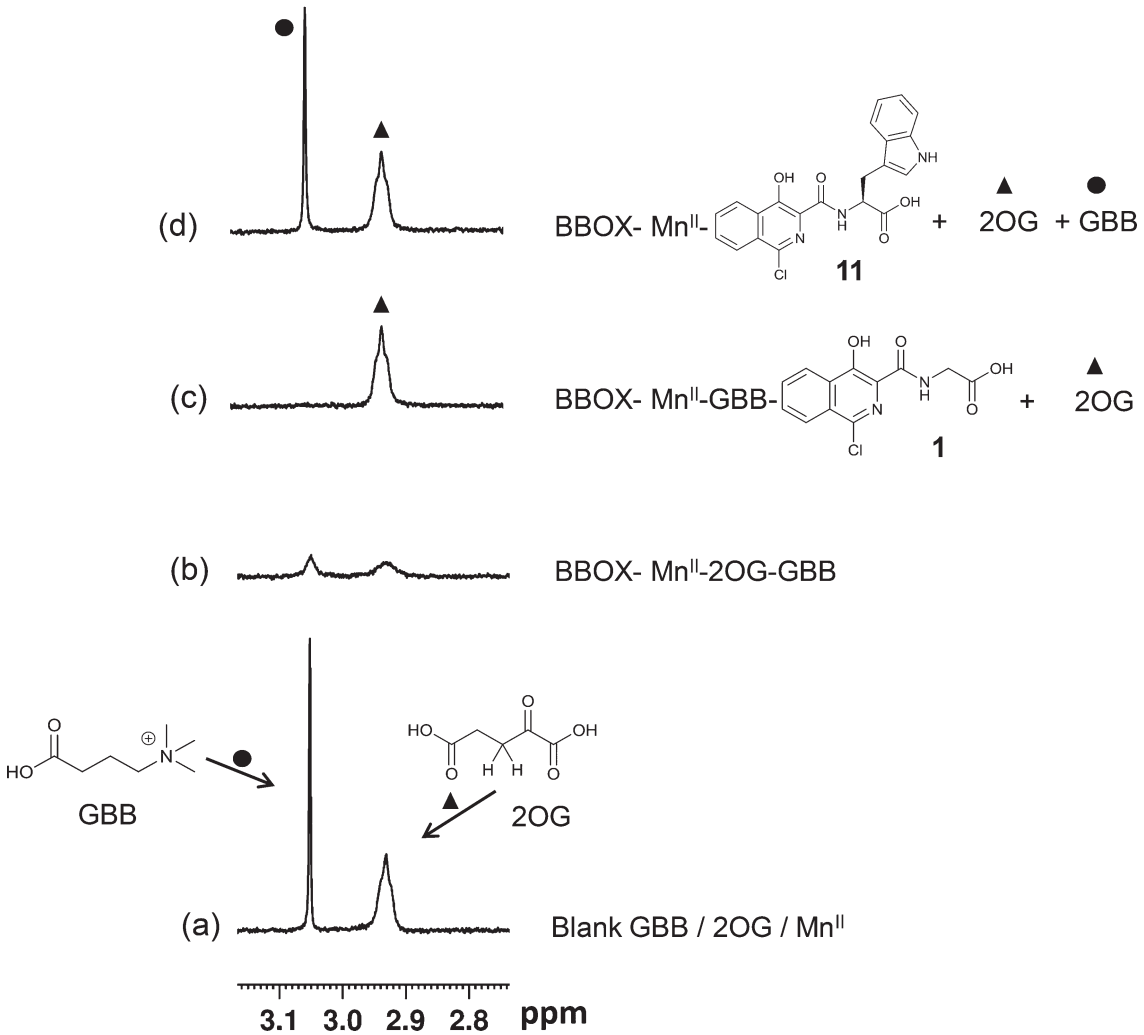


Fig. 3 Studying the binding of isoquinoline ligands **1** and **11** with psBBOX using the NMR dual-reporter displacement assay. 2OG and GBB are used as reporter ligands in CPMG edited ^{1}H NMR experiments. (a) Reporter ligands GBB and 2OG with Mn^{II} in solution in the absence of psBBOX. (b) After the addition of psBBOX. (c) Addition of **1** to sample of b. (d) Addition of **11** to sample of b. Symbols (●, ▲) indicate reporter signal recovery due to displacement. The final assay mixture contained 25 μM GBB, 300 μM 2OG, 150 μM Mn^{II}, 15 μM BBOX, 80 mM KCl and 25 μM isoquinoline-based ligand in 50 mM Tris-D_{11} buffer, pH 7.5, in D_2O.

BBOX inhibition by these isoquinoline-based compounds (Table 1). For the NMR assay, reactions (60 s) were quenched by the addition of 1 M HCl (Fig. S5†); the concentrations of GBB and L-carnitine were measured by integration of their Me_3 resonances (Fig. 5, S5†). The fluoride ion release assay[29] employs BBOX catalysed hydroxylation of (3*S*)-3-fluoro-4-(trimethylammonio)butanoate (GBBF) to give an unstable product which fragments to give a fluoride ion that can subsequently deprotect a *tert*-butyldimethylsilyl-protected fluorescein to provide measurable fluorescence. Differences in the absolute IC_{50} values of ^{1}H NMR and fluoride release assays, particularly for weak inhibitors, are likely due to the different experimental conditions used. The fluoride release assay was initially developed for use on hBBOX inhibition;[29] however, when applied to psBBOX we found the assay required higher concentrations of psBBOX to observe detectable fluorescence when compared to the hBBOX protocol (1 μM *vs.* 0.2 μM, respectively). This concentration difference is notable when considering the amount of enzyme used for the NMR assay (0.14 μM). In addition, the NMR assay uses the natural substrate GBB, whereas the fluoride assay necessarily uses GBBF whose affinity for psBBOX is notably different (K_M = 2.4 mM and 0.623 mM respectively).[16] Nonetheless the results from both assays (Table 1) suggest that isoquinoline-based ligands **1**, **2**, **3**, **9** and **11** are the strongest inhibitors in the series, consistent with their ability to compete with 2OG. Similarly, ligands **5**, **7** and **15** are medium potency inhibitors, whereas, ligands **4**, **6**, **8**, **10**, **12**, **13**, and **14** are weak inhibitors (Fig. 4, Table 1).

Although we have not yet been able to obtain a BBOX crystal structure with an isoquinoline-based inhibitor, manual docking and energy minimisation studies using a model of psBBOX based on a crystal structure of human BBOX (PDB ID: 3O2G)[20] suggest a possible structural explanation for the

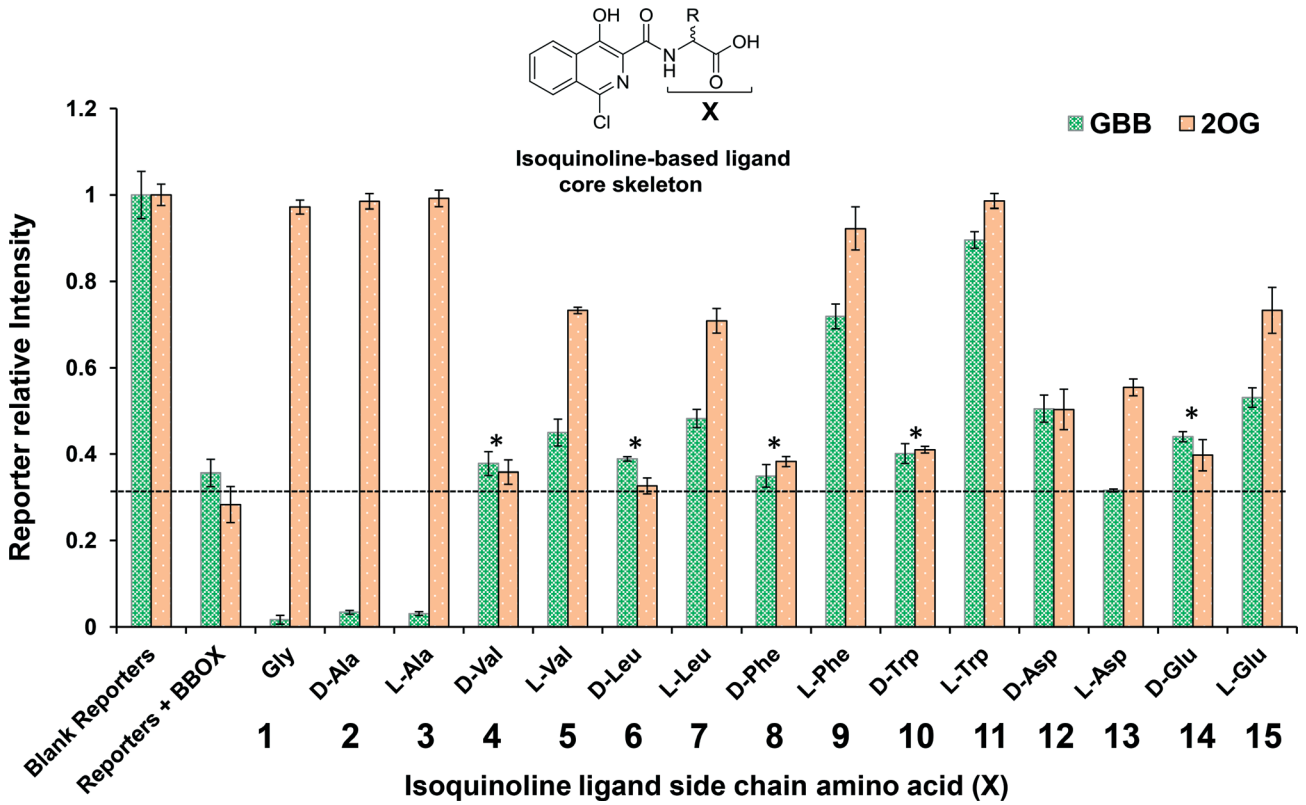


Fig. 4 The chart compares the signal intensity of GBB (green) and 2OG (orange) in the absence of psBBOX (1st left set of columns; intensity set to unit intensity) with their intensity after the addition of psBBOX (2nd left set of columns) and with their intensity after the addition of isoquinoline-based inhibitors **1**–**15** (side-chain amino acid X varying). The threshold line represents the relative intensity of the reporters in the presence of psBBOX. After the addition of isoquinoline-ligand, the level above this line represents reporter displacement; the level below represents enhancement of reporter binding. Asterisks highlight isoquinoline-ligands whose addition results in a negligible difference on GBB/2OG under these conditions. The final assay mixture contained 25 μM GBB, 300 μM 2OG, 150 μM Mn^{II}, 15 μM BBOX, 80 mM KCl and 25 μM isoquinoline-based ligand in 50 mM Tris-D_{11} buffer, pH 7.5, in D_2O. The error bars represent standard deviations from three separate measurements.

observed trends (Fig. 6). Studies by Rydzik *et al.*[17] on hBBOX using the fluoride ion release-based assay have shown that for hBBOX, as for psBBOX, the L-configured ligands **3** and **11** are better inhibitors than the corresponding D-forms (**2** and **10**) (IC_{50} 6 μM *vs.* 73 μM, and 11 μM *vs.* 33 μM, respectively). The docking results suggest that **1** binds psBBOX possibly by chelating the active site metal *via* its pyridinium nitrogen and amide carbonyl oxygen, as observed for other 2OG oxygenases, *i.e.* PHD2 (PDB ID: 2G1M)[30] and the fat mass and obesity protein FTO (PDB ID: 4 IE6),[31] although more than one chelation mode has been observed for related compounds with PHD2.[7] Removal of the isoquinoline inhibitor pyridinium nitrogen or the hydroxyl group leads to a significant loss in potency against hBBOX.[17] In our proposed binding mode for **1**, the phenyl ring of the isoquinoline projects towards the GBB binding pocket, causing movement of Phe184, but likely does not disrupt the pocket to the extent that GBB binding is ablated, *i.e.* it agrees with the NMR observations for **1**–**3** showing that 2OG but not GBB displacement occurs.

The observation that some ligands with larger hydrophobic side chains (**5**, **7**, **9**, **11**) compete with both 2OG and GBB, is interesting. Manual docking of **11** to the psBBOX model followed by energy minimisation suggests that the bicyclic ring system of these inhibitors is tilted with respect to that of the

Table 1 The binding dissociation constant (K_D) and inhibition constant (IC_{50}) values of the isoquinoline-based ligands tested in the study. Fluorescence K_D and ^1H NMR IC_{50} values are average of triplicates measurements. F^- release IC_{50} values show standard errors for quadruplicate measurements

Compound no.	Amino acid (X)	K_D (μM) by fluorescence assay	IC_{50} (μM) by ^1H NMR assay	IC_{50} (μM) by F^- release assay
1	Gly	22 ± 4	0.21 ± 0.01	1.2 ± 0.1
2	D-Ala	20 ± 4	0.20 ± 0.02	0.9 ± 0.03
3	L-Ala	15 ± 3	0.22 ± 0.01	0.7 ± 0.02
4	D-Val	144 ± 6	191 ± 6	>1000
5	L-Val	55 ± 3	23 ± 2	130 ± 14
6	D-Leu	103 ± 4	94 ± 8	>1000
7	L-Leu	55 ± 3	26 ± 2	125 ± 6
8	D-Phe	82 ± 3	38 ± 0.6	101 ± 6
9	L-Phe	26 ± 4	1 ± 0.03	13 ± 0.7
10	D-Trp	168 ± 37	18 ± 1	>1000
11	L-Trp	29 ± 5	0.9 ± 1	26 ± 2
12	D-Asp	101 ± 1	38 ± 1	160 ± 61
13	L-Asp	91 ± 2	76 ± 2	245 ± 62
14	D-Glu	89 ± 6	86 ± 5	>1000
15	L-Glu	64 ± 2	22 ± 2	147 ± 21

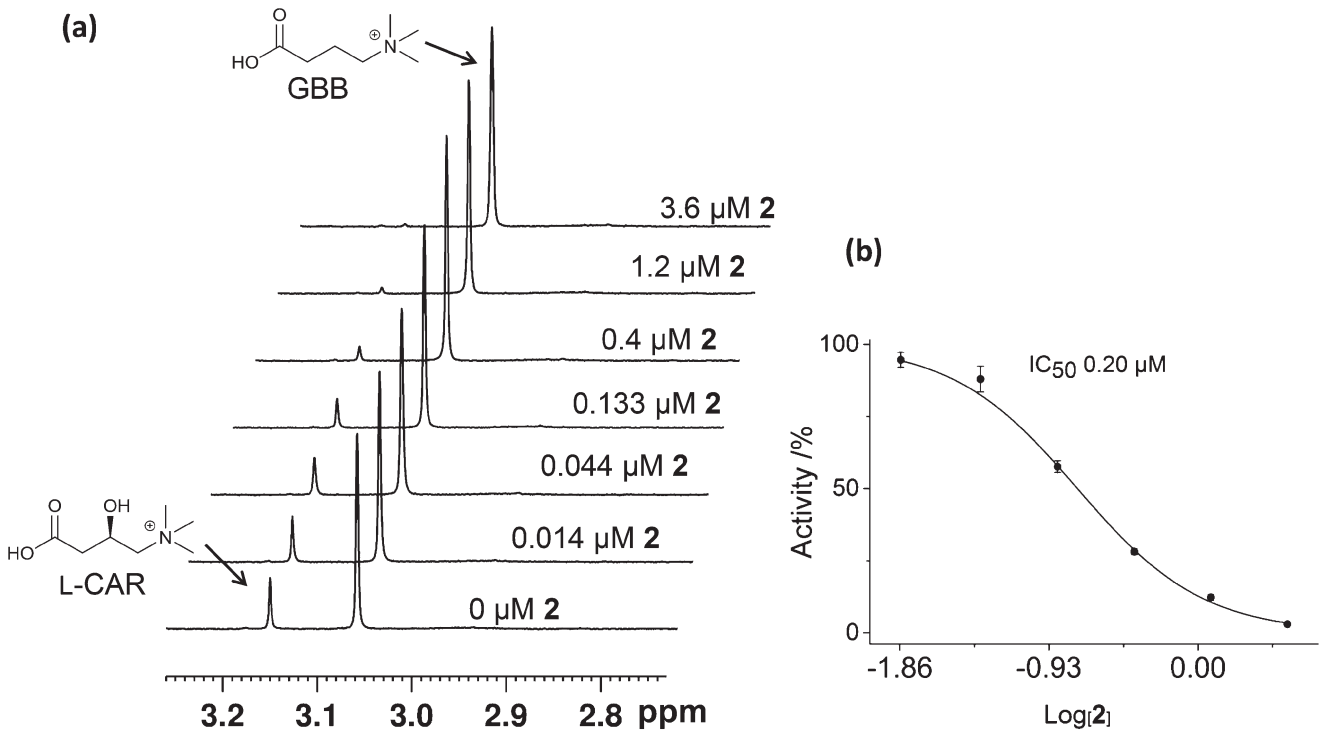


Fig. 5 Measurements of the IC_{50} value of the isoquinoline-ligand **2** by ^1H NMR. (a) Bottom to top; psBBOX catalysed GBB turnover into L-carnitine (L-CAR) in the presence of increasing amounts of **2**. (b) Dose response curve of (a). Error bars represent standard deviations from three separate measurements. The assay mixture contained 200 μM GBB, 600 μM 2OG, 100 μM Fe^{II}, 500 μM L-ascorbate, 80 mM KCl and 0.140 μM BBOX in 50 mM Tris-D_{11}, pH 7.5 in D_2O.

docked/minimised compound **1** while maintaining its metal coordination; this binding mode results in the chlorine and phenyl ring of the inhibitor protruding deeper into the GBB binding site. As a result the 'aromatic cage' residue Phe184 can move approximately 2.5 Å closer to the aromatic inhibitor thus changing the shape of the aromatic cage and also sterically blocking binding of the GBB trimethylamino group,[21] consistent with the NMR results, *i.e.* this inhibitor competes with 2OG and GBB. Although, alternative binding/chelation modes cannot be ruled out, the modelling studies imply that the larger side chains bind in a largely hydrophobic pocket adjacent to the 2OG pocket (defined by the side chains of Val-190, Ala-200, Leu-206, Ser-236, Leu-224, Leu-335, Trp-341, Phe-343, Arg-352 and Phe-364); they also imply that binding in this pocket prefers the L-, rather than the D-, stereochemistry as observed (Table 1, Fig. 6) due to a potential clash between the carboxylate of **10** (D-stereochemistry) and Leu-206. It is notable that a loop (residues 190–205) involved in GBB binding also forms part of the hydrophobic pocket and biophysical studies have revealed the occurrence of conformational changes in BBOX catalysis.[17] The binding of inhibitor **11** may thus elicit changes in the behaviour of this loop and so hinder GBB binding.

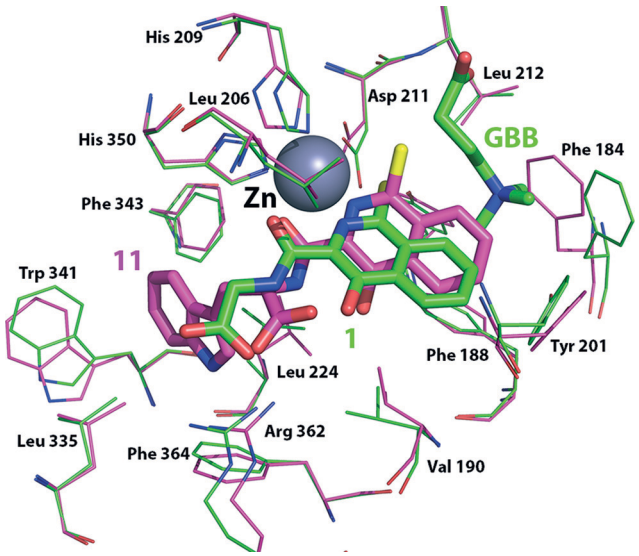


Fig. 6 Docking simulation showing possible conformational changes as a result of binding inhibitors **1** (green sticks, Cl in yellow) or **11** (pink sticks, Cl in yellow) to a model of the psBBOX active site based on a hBBOX X-ray crystal structure (PDB ID: 3O 2G).[20]

Conclusions

We have developed an efficient ligand-based ^1H NMR competitive binding assay for determining inhibitor binding to psBBOX that is operationally simple and requires amounts of materials comparable to other biophysical techniques. The method readily reveals whether an inhibitor disrupts the

binding of the co-substrate 2OG, the substrate GBB, or both. We envisage the assay will aid in the development of new BBOX inhibitors, including those selective for BBOX over other 2OG oxygenases. More generally, the results show how a readily applied NMR-based method can reveal mechanistic and structural insights not readily determined either by crystallography or by classical kinetic analyses.

Experimental

Materials

Chemicals were purchased from Sigma Aldrich, Alfa Aesar and Cortecnet unless otherwise stated. Details of the synthesis of the isoquinoline derivatives have been previously reported.[17,32]

Production and purification of psBBOX AK1

Recombinant psBBOX AK1 was produced and purified according to the protocol of Rydzik *et al.*[16]

NMR Experiments

All NMR experiments were performed at a temperature of 298 K using a Bruker Avance III 700 MHz spectrometer equipped with a TCI inverse cryoprobe or a Bruker Avance III 600 MHz spectrometer equipped with a 5 mm BB-F/^{1}H Prodigy N_2 cryoprobe. The PROJECT – CPMG pulse sequence $(90°_x - [\tau - 180°_y - \tau - 90°_y - \tau - 180°_y - \tau]_n - \text{acq})$[33] was used to attenuate broad resonances, using a total echo time of 32 ms. All spectra were processed with a Lorentzian line broadening of 0.3 Hz and were referenced to an internal standard (1,1,1 trifluoroacetone at 1.49 ppm). BBOX titration data for measuring the K_D of GBB were fitted using OriginPro 9.0 (Origin lab, Northampton, MA, USA). Bruker MATCH (3 mm diameter) and 5 mm NMR tubes with a total sample volume of 160 μL and 500 μL, respectively, were used. The solutions were buffered in 50 mM Tris-D_{11}·DCl, pH 7.5, in D_2O. For each sample the pulse tip-angle calibration was carried out using the single-pulse nutation method.[34] Each experiment was recorded with three separately prepared samples.

NMR IC_{50} measurements

The GBB turnover to L-carnitine was initiated by the addition of psBBOX to the assay mixture in a 1.5 mL plastic Eppendorf tube at ambient temperature (294 K), and was quenched after 60 s by addition of 10 μl 1 M HCl. The assay mixture contained 200 μM GBB, 600 μM 2OG, 100 μM Fe^{II} (from $Fe(NH_4)_2(SO_4)_2$ salt, which was prepared as a 250 mM stock solution in 20 mM HCl which was further diluted to a 2.5 mM solution in MilliQ purified water), 500 μM L-ascorbate, and 80 mM KCl in 50 mM Tris-D_{11}, pH 7.5 in D_2O, leading to a final psBBOX concentration of 0.14 μM. The psBBOX activity (%) at a given inhibitor concentration was measured as $([CAR]_{P+I}/[CAR]_P) \times 100$, where P = protein, I = inhibitor. Response curves were fitted using OriginPro 9.0 (Origin lab, Northampton, MA, USA).

Binding assay (K_D) – intrinsic fluorescence quenching[17,28]

K_D values were determined from the quenching of intrinsic tryptophan fluorescence measured at 294 K using a Pherastar FS plate reader (BMG labtech) with 96-well plates (Greiner, black, bottom: flat, clear). The following conditions were used: excitation 280 nm, emission 350 nm, 5 μM psBBOX, 50 μM Mn^{II} and varying concentrations of inhibitors to a final well volume of 50 μL in 50 mM Tris buffer pH 7.5, supplemented with 200 mM NaCl. Fluorescence readings were obtained in triplicate with errors represented as standard deviations. K_D values were obtained from the raw data by plotting inhibitor concentration against $\Delta F_{obs}/F_{max}$ (ΔF_{obs} is the decrease in observed fluorescence and F_{max} is the observed fluorescence signal without any inhibitor present).

Fluoride release assay (IC_{50})

IC_{50} measurements were obtained using a fluoride release assay, as previously applied to hBBOX.[29] The fluorescence signal was detected using a Pherastar FS plate reader (BMG labtech) fitted with a FITC FP 485/30 (485 nm, bandwidth 30 nm) and FITC FP 520/40 emission (520 nm, bandwidth 40 nm) filters, using 384-well plates (black, clear, flat bottom, Grenier BioOne). GBBF and the TBS-protected fluorescein probe were synthesised according to the published protocol.[29] The following conditions were used: 50 μM GBBF, 500 μM 2OG, 250 μM ascorbate, varying concentrations of inhibitors, 50 μM Fe^{II} (from $Fe(NH_4)_2(SO_4)_2$ salt, which was prepared as a 100 mM stock solution in 20 mM HCl which was further diluted to a 250 μM solution in MilliQ purified water before being added at the start of the measurements). Initiation of the assay was carried out by addition of psBBOX to each well containing all inhibitors and cofactors, to a concentration of 1 μM. Reactions were carried out in a final well volume of 10 μL in 50 mM Tris buffer pH 7.5 at 294 K, supplemented with 200 mM NaCl. After 10 minutes, the reaction was quenched by addition of 40 μL of TBS-protected fluorescein probe (final concentration 5 μM). The plates were then sealed and incubated for 1 hour at room temperature. Following this, 10 μL of HEPES buffer pH 7.0 was added and the fluorescence signal read up to 5 minutes after the addition of HEPES. Fluorescence data were recorded as quadruplicates and errors represented as standard deviations. Fluorescence signals were normalized by subtracting fluorescence values obtained from control wells which contained all reagents, but without enzyme or without inhibitor. IC_{50}s were obtained from the raw data using the Prism four parameter logistic dose–response model (sigmoidal, variable slope).

Acknowledgements

We thank the following for financial support: Abdul Wali Khan University, Mardan (AWKUM) Pakistan (A. K.), Biotechnology and Biological Sciences Research Council (BBSRC), the

British Heart Foundation, Wellcome Trust and Cancer Research UK.

References

1 *2-Oxoglutarate-Dependent Oxygenases*, ed. R. P. Hausinger and C. J. Schofield, Royal Society of Chemistry, Cambridge, 2015.
2 N. R. Rose, M. A. McDonough, O. N. King, A. Kawamura and C. J. Schofield, *Chem. Soc. Rev.*, 2011, **40**, 4364–4397.
3 F. M. Vaz and R. J. Wanders, *Biochem. J.*, 2002, **361**, 417–429.
4 B. Z. Simkhovich, Z. V. Shutenko, D. V. Meirena, K. B. Khagi, R. J. Mezapuke, T. N. Molodchina, I. J. Kalvins and E. Lukevics, *Biochem. Pharmacol.*, 1988, **37**, 195–202.
5 M. Dambrova, E. Liepinsh and I. Kalvinsh, *Trends Cardiovasc. Med.*, 2002, **12**, 275–279.
6 L. Henry, I. K. H. Leung, T. D. W. Claridge and C. J. Schofield, *Bioorg. Med. Chem. Lett.*, 2012, **22**, 4975–4978.
7 L. Poppe, C. M. Tegley, V. Li, J. Lewis, J. Zondlo, E. Yang, R. J. Kurzeja and R. Syed, *J. Am. Chem. Soc.*, 2009, **131**, 16654–16655.
8 M. C. Chan, O. Atasoylu, E. Hodson, A. Tumber, I. K. H. Leung, R. Chowdhury, V. Gomez-Perez, M. Demetriades, A. M. Rydzik, J. Holt-Martyn, Y. M. Tian, T. Bishop, T. D. W. Claridge, A. Kawamura, C. W. Pugh, P. J. Ratcliffe and C. J. Schofield, *PLoS One*, 2015, **10**, e0132004.
9 I. K. H. Leung, M. Demetriades, A. P. Hardy, C. Lejeune, T. J. Smart, A. Szollossi, A. Kawamura, C. J. Schofield and T. D. W. Claridge, *J. Med. Chem.*, 2013, **56**, 547–555.
10 M. D. Shortridge, D. S. Hage, G. S. Harbison and R. Powers, *J. Comb. Chem.*, 2008, **10**, 948–958.
11 M. Sarrazin, J. C. Sari, M. Bourdeaux-Pontier and C. Briand, *Mol. Pharmacol.*, 1979, **15**, 71–77.
12 J. J. Fischer and O. Jardetzky, *J. Am. Chem. Soc.*, 1965, **87**, 3237–3244.
13 C. Dalvit, A. D. Gossert, J. Coutant and M. Piotto, *Magn. Reson. Chem.*, 2011, **49**, 199–202.
14 P. J. Hajduk, E. T. Olejniczak and S. W. Fesik, *J. Am. Chem. Soc.*, 1997, **119**, 12257–12261.
15 W. Jahnke, P. Floersheim, C. Ostermeier, X. Zhang, R. Hemmig, K. Hurth and D. P. Uzunov, *Angew. Chem., Int. Ed.*, 2002, **41**, 3420–3423.
16 A. M. Rydzik, I. K. H. Leung, G. T. Kochan, N. D. Loik, L. Henry, M. A. McDonough, T. D. W. Claridge and C. J. Schofield, *Org. Biomol. Chem.*, 2014, **12**, 6354–6358.
17 A. M. Rydzik, R. Chowdhury, G. T. Kochan, S. T. Williams, M. A. McDonough, A. Kawamura and C. J. Schofield, *Chem. Sci.*, 2014, **5**, 1765–1771.
18 K. Tars, J. Leitans, A. Kazaks, D. Zelencova, E. Liepinsh, J. Kuka, M. Makrecka, D. Lola, V. Andrianovs, D. Gustina, S. Grinberga, E. Liepinsh, I. Kalvinsh, M. Dambrova, E. Loza and O. Pugovics, *J. Med. Chem.*, 2014, **57**, 2213–2236.
19 K. Tars, J. Rumnieks, A. Zeltins, A. Kazaks, S. Kotelovica, A. Leonciks, J. Sharipo, A. Viksna, J. Kuka and E. Liepinsh, *Biochem. Biophys. Res. Commun.*, 2010, **398**, 634–639.
20 I. K. H. Leung, T. J. Krojer, G. T. Kochan, L. Henry, F. von Delft, T. D. W. Claridge, U. Oppermann, M. A. McDonough and C. J. Schofield, *Chem. Biol.*, 2010, **17**, 1316–1324.
21 J. J. A. G. Kamps, A. Khan, H. Choi, R. K. Lesniak, J. Brem, A. M. Rydzik, M. A. McDonough, C. J. Schofield, T. D. W. Claridge and J. Mecinović, *Chem. – Eur. J.*, 2016, **22**, 1270–1276.
22 I. Bertini, C. Luchinat, G. Parigi and R. Pierattelli, *ChemBioChem*, 2005, **6**, 1536–1549.
23 I. Bertini, C. Luchinat and G. Parigi, *Solution NMR of paramagnetic molecules: applications to metallobiomolecules and models*, Elsevier, Amsterdam, 2001.
24 I. Bertini, M. Fragai, Y. M. Lee, C. Luchinat and B. Terni, *Angew. Chem., Int. Ed.*, 2004, **43**, 2254–2256.
25 I. K. H. Leung, E. Flashman, K. K. Yeoh, C. J. Schofield and T. D. W. Claridge, *J. Med. Chem.*, 2010, **53**, 867–875.
26 I. Bertini, M. Fragai, C. Luchinat and E. Talluri, *Angew. Chem., Int. Ed.*, 2008, **47**, 4533–4537.
27 W. Jahnke, S. Rudisser and M. Zurini, *J. Am. Chem. Soc.*, 2001, **123**, 3149–3150.
28 A. M. Rydzik, J. Brem, W. B. Struwe, G. T. Kochan, J. L. Benesch and C. J. Schofield, *Bioorg. Med. Chem. Lett.*, 2014, **24**, 4954–4957.
29 A. M. Rydzik, I. K. H. Leung, G. T. Kochan, A. Thalhammer, U. Oppermann, T. D. W. Claridge and C. J. Schofield, *ChemBioChem*, 2012, **13**, 1559–1563.
30 M. A. McDonough, V. Li, E. Flashman, R. Chowdhury, C. Mohr, B. M. Lienard, J. Zondlo, N. J. Oldham, I. J. Clifton, J. Lewis, L. A. McNeill, R. J. Kurzeja, K. S. Hewitson, E. Yang, S. Jordan, R. S. Syed and C. J. Schofield, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 9814–9819.
31 W. Aik, M. Demetriades, M. K. Hamdan, E. A. Bagg, K. K. Yeoh, C. Lejeune, Z. Zhang, M. A. McDonough and C. J. Schofield, *J. Med. Chem.*, 2013, **56**, 3680–3688.
32 S. S. van Berkel, J. Brem, A. M. Rydzik, R. Salimraj, R. Cain, A. Verma, R. J. Owens, C. W. Fishwick, J. Spencer and C. J. Schofield, *J. Med. Chem.*, 2013, **56**, 6945–6953.
33 J. A. Aguilar, M. Nilsson, G. Bodenhausen and G. A. Morris, *Chem. Commun.*, 2012, **48**, 811–813.
34 P. S. Wu and G. Otting, *J. Magn. Reson.*, 2005, **176**, 115–119.

ChemPubSoc Europe

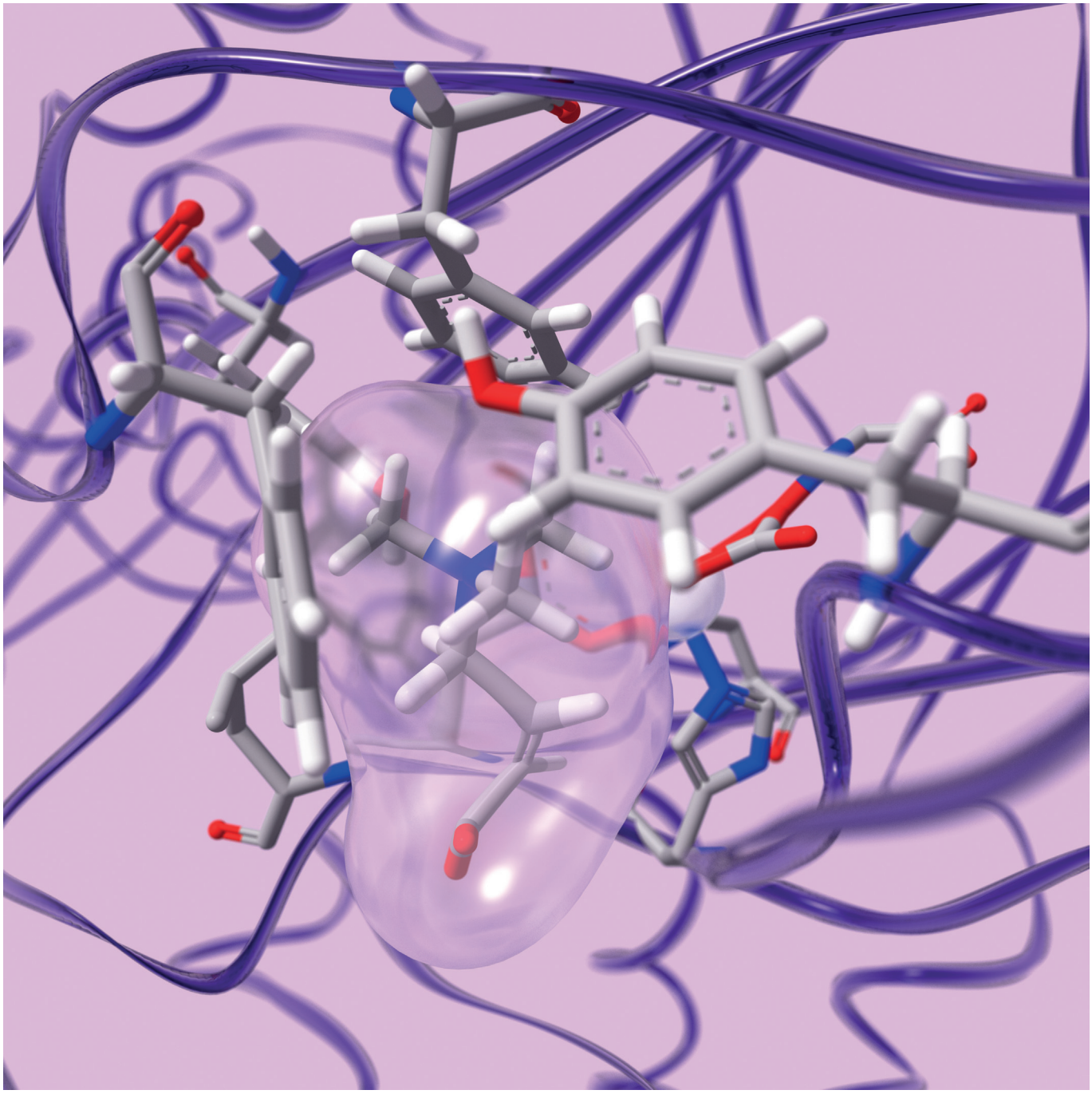
DOI: 10.1002/chem.201503761

CHEMISTRY
A European Journal
Full Paper

Biocatalysis | *Hot Paper*

Cation–π Interactions Contribute to Substrate Recognition in γ-Butyrobetaine Hydroxylase Catalysis

Jos J. A. G. Kamps,[a] Amjad Khan,[b] Hwanho Choi,[b] Robert K. Lesniak,[b] Jürgen Brem,[b] Anna M. Rydzik,[b] Michael A. McDonough,[b] Christopher J. Schofield,[b] Timothy D. W. Claridge,[b] and Jasmin Mecinović*[a]



Abstract: γ-Butyrobetaine hydroxylase (BBOX) is a non-heme Fe^{II}- and 2-oxoglutarate-dependent oxygenase that catalyzes the stereoselective hydroxylation of an unactivated C–H bond of γ-butyrobetaine (γBB) in the final step of carnitine biosynthesis. BBOX contains an aromatic cage for the recognition of the positively charged trimethylammonium group of the γBB substrate. Enzyme binding and kinetic analyses on substrate analogues with P and As substituting for N in the trimethylammonium group show that the analogues are good BBOX substrates, which follow the efficiency trend $N^+ > P^+ > As^+$. The results reveal that an uncharged carbon analogue of γBB is not a BBOX substrate, thus highlighting the importance of the energetically favorable cation–π interactions in productive substrate recognition.

Introduction

2-Oxoglutarate (2OG) and Fe^{II}-dependent oxygenases play important roles in human physiology, including in hypoxia sensing, DNA repair, chromatin modification and fatty acid metabolism.[1,2] γ-Butyrobetaine hydroxylase (BBOX), a 2OG oxygenase, catalyzes the stereoselective hydroxylation of γ-butyrobetaine (γBB, **1**) to form L-carnitine (L-CAR) in eukaryotes and some prokaryotes (Figure 1a).[3–6] L-Carnitine is required for the transport of fatty acids into the mitochondrial matrix, where they are converted into acetyl-CoA.[7] Structural analyses on human BBOX (hBBOX) reveal that the active site Fe^{II} is chelated by a 2His-1Asp triad and that the 2OG cosubstrate binds in a similar mode to other 2OG oxygenases (Figure 1b).[8,9] The BBOX active site contains an apparent "aromatic cage," which binds the γBB substrate's trimethylammonium group, and two asparagine residues that hydrogen bond with the γBB carboxylate (Figure 1b).[8,9] The chiral environment of the enzyme's active site enables hBBOX to catalyze the oxidative desymmetrization of achiral *N*,*N*-dialkyl piperidine-4-carboxylates.[10] hBBOX also catalyzes an unusual Stevens-type rearrangement of Mildronate (also known as Meldonium), a γBB competitive inhibitor that is clinically used in the treatment of myocardial infarction in order to inhibit fatty acid metabolism.[11,12] Recent studies revealed that subtle differences in the active sites of human and *Pseudomonas* sp. AK1 BBOX (hereafter psBBOX) can result in altered substrate-analogue selectivities.[13]

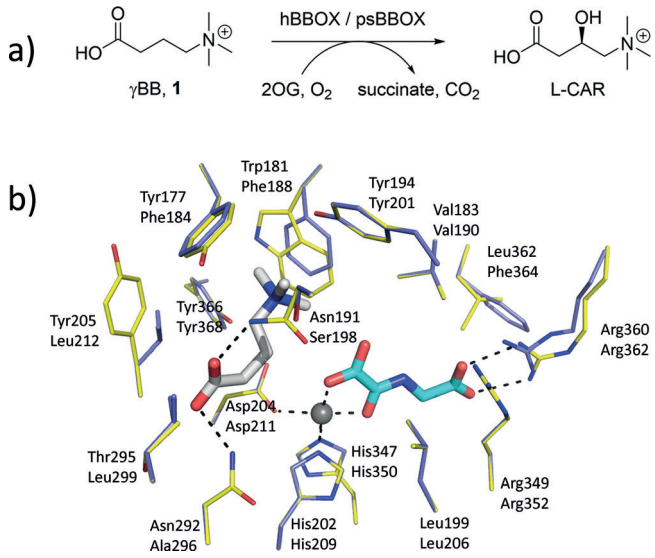


Figure 1. BBOX-catalyzed hydroxylation: a) hBBOX-/psBBOX-catalyzed hydroxylation of γBB **1**; b) view of the hBBOX active site (yellow sticks, upper residue numbers, PDB ID: 3O2G) and the psBBOX model (blue sticks, lower residue numbers) with γBB (white sticks), *N*-oxalylglycine (NOG, a 2OG analogue, cyan sticks) and Zn^{II} substituting for Fe^{II} (grey sphere).[13]

Functionally and structurally diverse proteins contain aromatic cages (or aromatic boxes) as recognition modules for substrate binding. Aromatic cages typically comprise the side-chains of 2–4 aromatic residues (Trp, Tyr, Phe), and are observed on both exposed and buried sites.[14–16] Work by Dougherty and coworkers has demonstrated that aromatic cages can recognize positively charged quaternary ammonium species *via* favorable cation–π interactions.[17–19]

Cation–π interactions are involved in associations of Cys-loop receptors and G-protein-coupled receptors with neurotransmitters, including acetylcholine, serotonin, dopamine, epinephrine, and histamine.[20] Studies on chromatin interactions involved in the regulation of gene expression reveal that "reader domain" proteins that recognize trimethyllysine-containing histone tails interact through cation–π interactions.[14,21] Thus, along with hydrophobic effects, hydrogen bonding, ion pairing, and van der Waals interactions, cation–π interactions are crucial noncovalent forces in protein–protein and protein–ligand associations.[22]

The BBOX aromatic cage contains the electron-rich aromatic residues Phe184, Phe188 and Tyr201; Tyr368 is also located in close proximity, with its side-chain OH group likely positioned to form an H-bond with Asp211 (Figure 1b).[13] The location of the γBB trimethylammonium group inside the aromatic cage

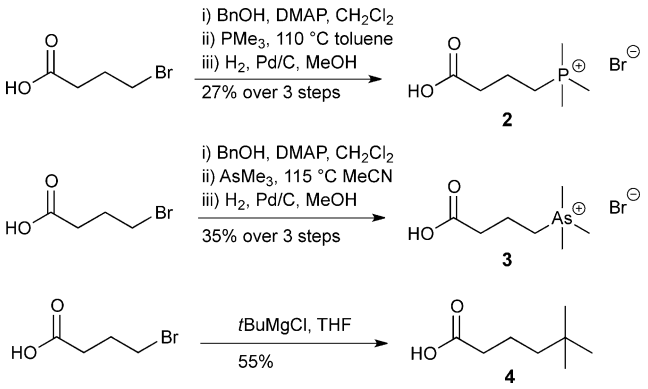
[a] *J. J. A. G. Kamps,[+] Dr. J. Mecinović*
Institute for Molecules and Materials, Radboud University Nijmegen
Heyendaalseweg 135, 6525 AJ Nijmegen (The Netherlands)
E-mail: j.mecinovic@science.ru.nl

[b] *A. Khan,[+] Dr. H. Choi, R. K. Lesniak, Dr. J. Brem, Dr. A. M. Rydzik, Dr. M. A. McDonough, Prof. C. J. Schofield, Prof. T. D. W. Claridge*
Chemistry Research Laboratory, Department of Chemistry
University of Oxford, 12 Mansfield Road, Oxford OX1 3TA (UK)

[+] *These authors contributed equally to this work.*

Supporting information and ORCID(s) from the author(s) for this article are available on the WWW under http://dx.doi.org/10.1002/chem.201503761.

suggests that the association between γBB and BBOX could be substantially mediated by energetically favorable cation–π interactions. In support of this proposition, previous work with recombinant psBBOX employing a radioactive-based assay with the C-analogue **4** of γBB (**1**), provided evidence that it is a poor substrate.[23] Herein, we report on the use of NMR and MS assays to investigate γBB analogues with P, As, and C substituting for N (**2**–**4**, respectively; Scheme 1) as psBBOX substrates; the results clearly support the proposal that cation–π interactions are crucial in the recognition of γBB by psBBOX.



Scheme 1. Syntheses of γBB analogues **2**–**4**. DMAP = 4-dimethylaminopyridine.

We envisaged that replacement of the trimethylammonium group of γBB by three closely analogous functionalities, that is, positively charged trimethylphosphonium and trimethylarsonium "Group V analogues" **2** and **3**, and the neutral *tert*-butyl analogue **4**, would provide insights into the role of cation–π interactions in BBOX catalysis (Scheme 1). Like γBB (**1**), the phosphorus (**2**) and arsenic (**3**) derivatives possess a "fixed" positive charge with respect to their trimethyl group, but are slightly larger than **1**, whereas the carbon analogue (**4**) has nearly the same size and shape, but lacks a positive charge. We hypothesized that direct comparison of binding and psBBOX-catalyzed hydroxylation of the positively charged γBB **1** and neutral **4** would inform on the interactions between the NMe_3^+ group and the aromatic cage of psBBOX. Thus, if **4** is a much poorer psBBOX ligand and substrate than **1**, the requirement of cation–π interactions in psBBOX catalysis would be implied. In contrast, if **4** is a better ligand and substrate for psBBOX, that would suggest that hydrophobic interactions dominate the psBBOX–γBB association.

Results and Discussion

The phosphorus (**2**) and arsenic (**3**) analogues of γBB (**1**) were synthesized in concise three-step sequences from 4-bromobutyric acid (Scheme 1 and Figure S1 in the Supporting Information). 5,5-Dimethylhexanoic acid (**4**) was synthesized from 4-bromobutanoic acid and a twofold excess of *tert*-butylmagnesium chloride (Scheme 1 and Figure S1).

We then used LC-MS analyses to test for psBBOX-catalyzed hydroxylation of the three γBB analogues. In the presence of psBBOX (1 μM), **1** was efficiently hydroxylated (complete conversion in 5 min; Figure 2a and Figure S2 in the Supporting Information). Phosphorus analogue **2** was also a good substrate, with about 70% conversion (Figure 2b); arsenic analogue **3** was less well hydroxylated (approximately 45%), but clear evi-

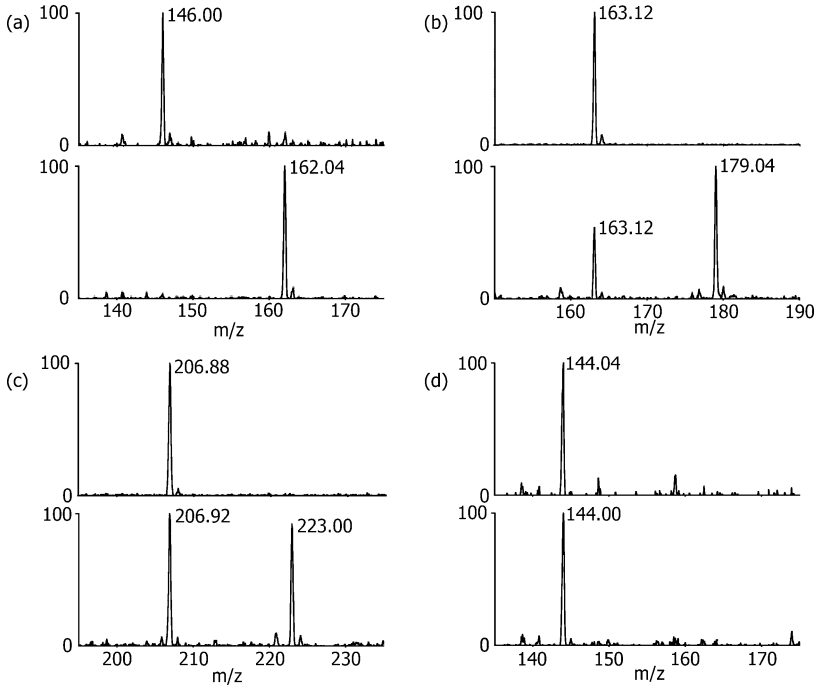


Figure 2. Mass spectrometry data for psBBOX-catalyzed hydroxylations: a) natural substrate γ-BB **1**; b) phosphorus analogue **2**; c) arsenic analogue **3**; d) neutral carbon analogue **4**. Top panel = starting substrate; bottom panel = psBBOX-catalyzed reaction [a–c) 1 μM psBBOX, 5 min; d) 10 μM psBBOX, 3 h].

dence for hydroxylation was obtained (Figure 2c). Importantly, the neutral analogue **4** was not hydroxylated by psBBOX within our limits of detection; use of an increased amount of psBBOX (10 μM) and prolonged incubation (3 h) did not result in the observation of hydroxylated product by LC-MS (Figure 2d). In the absence of psBBOX, no hydroxylation of **2** or **3** took place (see the Supporting Information, Figure S3). Time-course NMR studies were consistent with the order of efficiency as observed by LC-MS (**1** > **2** > **3**) and revealed that in each case psBBOX-catalyzed hydroxylation of **1**–**3** is tightly coupled to oxidation of 2OG to succinate (Figure 3a–c). Consistent with the LC-MS results, the NMR assays revealed that **4** was not hydroxylated, even in the presence of 10 μM psBBOX (Figure 3d). Controls with **2** and **3** showed a lack of hydroxylation without psBBOX (see the Supporting Information, Figures S4 and S5). Overall, these results reveal the importance of a positively charged XMe_3 substrate group for psBBOX catalysis.

COSY- and HSQC-based 2D NMR spectroscopy of the products implied that psBBOX-catalyzed hydroxylation of **2** and **3** occurs at C3 (see the Supporting Information, Figures S6–S11). These assignments were confirmed by stereoselective synthesis of the (3*R*)- and (3*S*)-hydroxylated phosphorus and arsenic derivatives **5**–**8** (Scheme 2 and Figure S12 in the Supporting Information). The intensity of ^{1}H NMR peaks that correspond to the hydroxylated products increased upon the addition of authentic (3*R*)-hydroxylated phosphorus- (**5**) and arsenic-containing (**6**) products into the respective reaction mixtures, confirming that psBBOX-catalyzed hydroxylation of the analogues occurs at C3 (see the Supporting Information, Figures S13 and S14). Together, these results imply that the psBBOX-catalyzed hydroxylation of **2** and **3** results in the C3-hydroxylated products **5** and **6**, likely with (3*R*) stereochemistry (Scheme 3).

LC-MS analyses testing (3*R*)- and (3*S*)-hydroxylated phosphorus derivatives **5** and **7** as substrates showed that the latter is a poor psBBOX substrate, giving a small amount of the 3-keto product in the presence of 20 μM psBBOX (as assigned on the basis of a −2 Da mass shift), whereas the (3*R*)-**5** enantiomer appears not to be converted within limits of detection for this

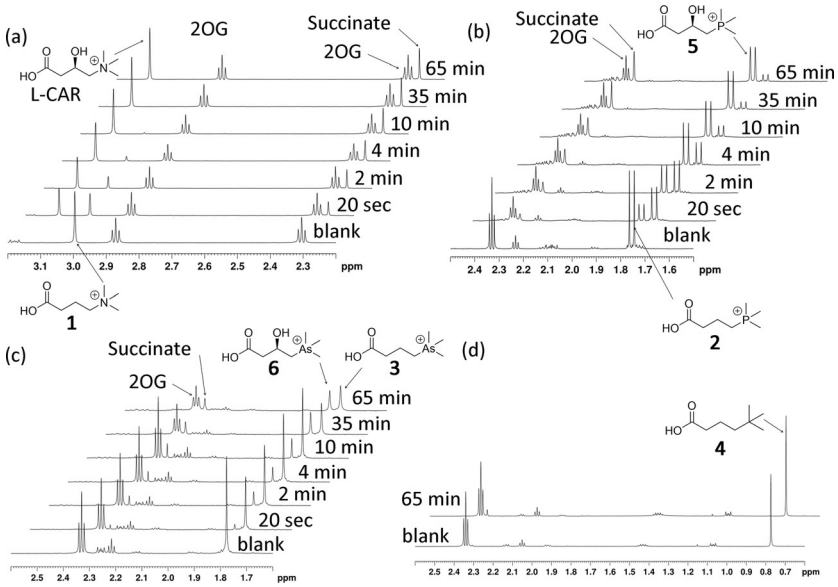
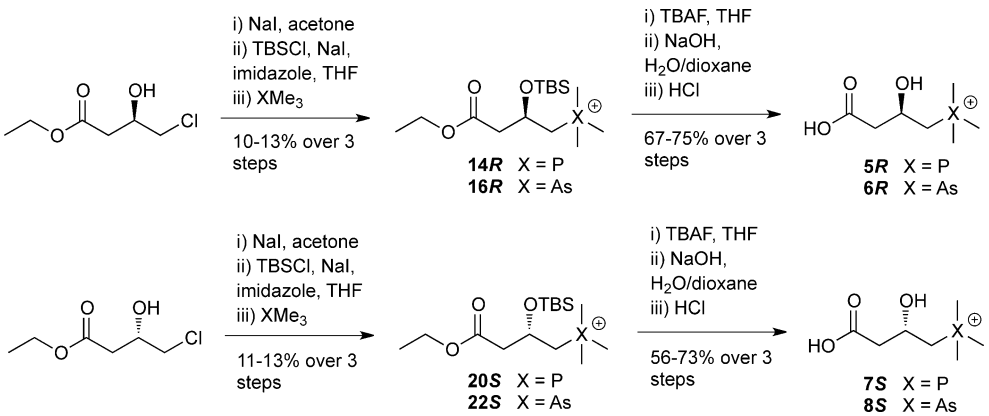
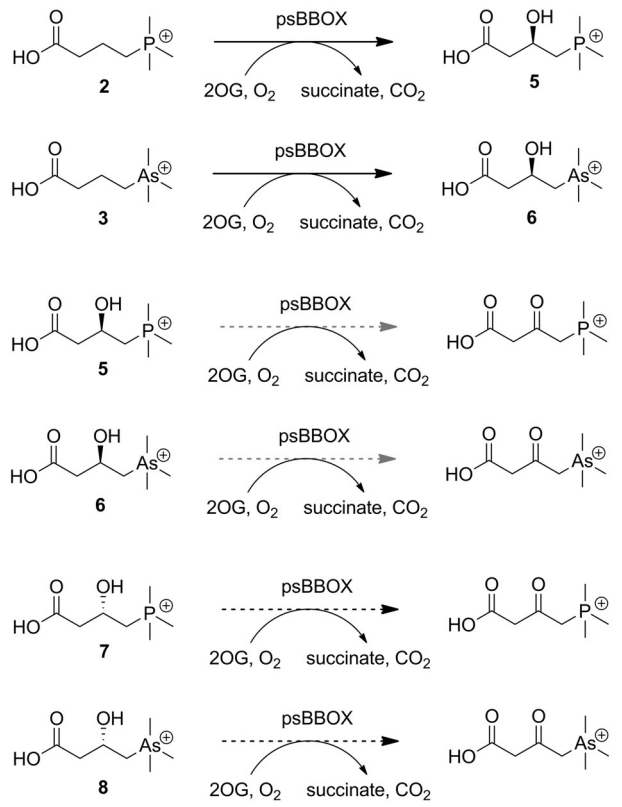


Figure 3. ^{1}H NMR monitoring of the hydroxylations by psBBOX in the presence of 2OG as a cosubstrate and Fe^{II} as a cofactor: a) **1**; b) **2**; c) **3**; d) **4** (blank = reaction mixture in the absence of psBBOX). The substrate/product ratio is measured for each spectrum as a function of reaction time.



Scheme 2. Stereoselective syntheses of (3*R*)- and (3*S*)-hydroxylated phosphorus- and arsenic-containing derivatives **5**–**8**.



Scheme 3. psBBOX-catalyzed stereoselective hydroxylation of positively charged phosphorus (**2**) and arsenic (**3**) analogues of γBB, and their (3*R*)- and (3*S*)-hydroxylated derivatives. Black dashed arrows and grey dashed arrows indicate poor and very poor conversions, respectively.

assay (see the Supporting Information, Figure S15). We did not detect ketone formation with the hydroxylated arsenic derivatives (3*R*)-**6** or (3*S*)-**8** by LC-MS (see the Supporting Information, Figure S16). However, NMR assays indicated that all four hydroxylated derivatives **5**–**8** are very poor psBBOX substrates with low levels of conversion being observed in the presence of 10 μM psBBOX, likely to the C3 ketone, which we could not fully characterize due to the low levels of conversion (Scheme 3 and Figures S17–S21 in the Supporting Information). The observation of a product with the same chemical shift for the PMe_3^+ group from incubation of **5** and **7** supports the proposed formation of the 3-ketone (as with hydroxylation of the enantiomeric arsenic derivatives **6** and **8**; Figure S17–S21). As for the carnitine enantiomers,[13] the (3*S*) enantiomers **7** and **8** were substantially better substrates than the (3*R*) enantiomers **5** and **6**, which afforded barely detectable products (36% vs. 7% for phosphorus derivatives **7** and **5**; 22% vs. 3% for arsenic derivatives **8** and **6**). These results are consistent with previous results of hBBOX- and psBBOX-catalyzed hydroxylation of D- and L-carnitine.[13]

Under (near) anaerobic conditions (Ar atmosphere), only traces (<5%) of hydroxylated products were observed for **2** and **3** (see the Supporting Information, Figure S22). Reactions of **2** and **3** in $H_2{}^{18}O$ (90% ^{18}O, Tris buffer, pH 7.5) afforded hydroxylated products with the same molecular mass as products from standard reaction in Tris buffer/$H_2{}^{16}O$ (see the Supporting Information, Figure S23). These results are consistent with the oxygen atom in the hydroxylated products from **2** and **3** deriving (at least predominantly) from atmospheric oxygen and not from water, as also found in oxygen-labeling studies on most other 2OG oxygenases.[24]

To quantify the relative efficiencies of the substrate analogues **2** and **3**, we carried out kinetic analyses using LC-MS. The results revealed that the natural NMe_3^+-containing substrate **1** is a superior substrate to the PMe_3^+- and $AsMe_3^+$-containing analogues **2** and **3**, by approximately 2- and 3-fold, respectively, as measured by k_{cat}/K_m values (Table 1 and Figure S24 in the Supporting Information). The differences in k_{cat} and K_m values are relatively small for **1**, **2**, and **3**, implying the importance of a positively charged XMe_3 group. There is a trend in decreasing k_{cat} from **1** to **3**, suggesting that the longer C–P and C–As bond lengths (1.9 Å and 2.0 Å, respectively) relative to the C–N bond length (1.5 Å) may cause substrates **2** and **3** to adopt a non-optimal binding mode with respect to the Fe^{IV}=O intermediate, as observed with substrate analogue studies on other 2OG oxygenases.[1] The K_m values for all three positively charged substrates were within error (1.00–1.37 mM).

Table 1. Kinetic parameters for conversion of γBB **1** and the phosphorus (**2**) and arsenic (**3**) analogues into the corresponding hydroxylated products by psBBOX.[a]

	1	**2**	**3**
V_{max} [μM s^{-1}]	2.81 ± 0.31	1.65 ± 0.16	1.04 ± 0.13
k_{cat} [s^{-1}]	7.02 ± 0.77	4.13 ± 0.40	2.61 ± 0.31
K_m [mM]	1.00 ± 0.21	1.20 ± 0.20	1.37 ± 0.28

[a] 400 nM psBBOX was used with varying concentrations of substrates from 50 μM to 1.5 mM.

Although the kinetic studies and time-course analyses imply that binding in the aromatic cage is important, the small differences in the k_{cat} and K_m values and the complexity of 2OG oxygenase catalysis motivated us to carry out NMR titration studies to obtain K_D values for the binding of **1**–**3** to the psBBOX·Zn^{II}·2OG complex (using Zn^{II} as an unreactive Fe^{II} substitute); the K_D values for analogues **1**, **2**, and **3** were 5 μM, 7 μM, and 17 μM, respectively (see the Supporting Information, Figures S25–S27). The NMR experiments also implied that **4** does not bind in the active site of psBBOX within detection limits (see the Supporting Information, Figure S28). Similar trends have been observed in inhibition studies on the serine protease factor Xa, which contains an aromatic cage in its "S4" substrate residue binding pocket. That is, a ligand possessing a quaternary ammonium moiety inhibits factor Xa to a similar extent as one with an analogous phosphonium group, whereas the neutral carba analogue is substantially (60-fold) less potent.[25,26]

To investigate the binding modes of **2** and **3** to psBBOX, we used an X-ray crystal structure of hBBOX in complex with Zn^{II}, NOG, and γBB [Protein Data Bank (PDB) ID: 3O2G] to make a homology model of psBBOX. All three positively charged substrates (**1**, **2**, **3**) are predicted to bind in a similar manner in

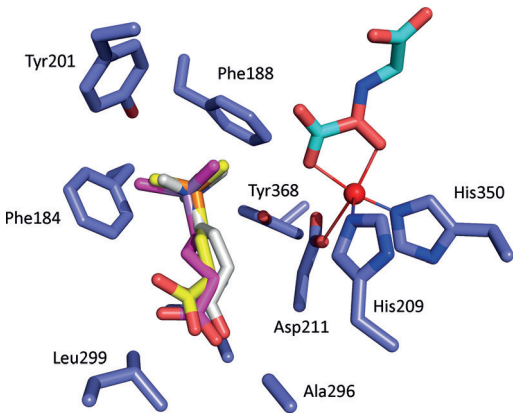


Figure 4. Modeled structure of psBBOX (blue) complexed with Zn^{II} (red), *N*-oxalylglycine (cyan), and substrates **1** (white), **2** (yellow), and **3** (magenta).

the psBBOX active site, consistent with the observation of C3 hydroxylation, and with the XMe_3^+ group positioned in the aromatic cage (Figure 4). Due to strong negative inductive effect of a large XMe_3^+ group, the C4 position (adjacent to the XMe_3^+ group) is activated, but also the most sterically hindered. The C2 position, although the most activated, is positioned away from the Fe^{IV}=O intermediate. Thus, of all three potential sites, psBBOX-catalyzed hydroxylation occurs at the C3 site of **1**–**3**.

We calculated the CHELPG atomic charges (see Experimental Section) of the X atom of the XMe_3 group and the attached three carbon and nine hydrogen atoms for the docked conformations for **1**–**3** and the minimized energy conformation for **4** (see the Supporting Information, Table S1). The average partial charges of the nine hydrogen atoms of XMe_3^+ show a very slight incremental trend of +0.1506 (for **3**), +0.1596 (for **2**), and +0.1652 (for **1**). The calculated partial charge of hydrogen atoms in the neutral *tert*-butyl group of **4** is +0.0884. These results are in agreement with trends in binding affinities as observed by NMR spectroscopy, that is, the more positively charged H atoms of the XMe_3^+ substrates result in stronger cation–π interactions with the psBBOX aromatic cage.

Conclusions

In conclusion, our substrate analogue studies employing both turnover and binding assays with purified recombinant enzyme clearly support the proposal that recognition and BBOX-catalyzed hydroxylation of γBB involve energetically favorable cation–π interactions between the positively charged trimethylammonium group of γBB and the aromatic cage of psBBOX. The observation that the neutral carbon analogue of γBB does not bind to psBBOX and does not undergo psBBOX-catalyzed hydroxylation in our assays further supports this view. Furthermore, the results reveal that the positively charged trimethylphosphonium and trimethylarsonium analogues of γBB are good substrate mimics and have a potential to act as small-molecule probes for functional studies of carnitine biosynthesis. Thus, for example, the enzymatic conversion of the phosphorus analogue could be probed by ^{31}P NMR spectroscopy, and the ^{32}P-labeled substrate might have a potential for utilization in radioactive tracing inside cells. Phosphorus and arsenic analogues of naturally occurring molecules that contain the quaternary ammonium groups might become useful probes for other genuinely important biomolecular processes that are driven by strong cation–π interactions. Given the central role of carnitine in eukaryotic fatty acid metabolism, our results also highlight biomedicinally important cation–π interactions.

Experimental Section

BBOX production and purification

Recombinant psBBOX was produced according to a previously described procedure.[13] In brief, cells were cultured in 2TY media supplemented with 50 μg mL^{-1} ampicillin until mid-log phase growth was achieved (OD600 0.7). Production of the recombinant proteins was then induced by addition of 0.2 mM IPTG and the cells were cultured for further 16 h at 15 °C. Cells were harvested by centrifugation (8 min, 8 g), then resuspended in lysis buffer (50 mM Tris pH 7.5/500 mM NaCl) supplemented with 0.2% Tween 20, DNAse, Lysosyme, and EDTA-free protease-inhibitors.

The cell lysates were loaded onto a 5 mL HisTrap HP column (GE Healthcare Life Sciences, Little Chalfont, UK), with 50 mM Tris pH 7.5/500 mM NaCl, containing 20 mM imidazole, then eluted with an imidazole gradient (up to 500 mM imidazole). Fractions containing the purified psBBOX protein were concentrated by centrifugal ultrafiltration (50 kDa cutoff membrane). The protein solution was then injected onto a Superdex S200 column (300 mL) and eluted with 20 mM Tris pH 7.5/200 mM NaCl supplemented with 10 mM EDTA. Fractions containing purified psBBOX were concentrated by centrifugal ultrafiltration (50 kDa cutoff filter) and buffer exchanged by using a PD-10 column to a Chelex 100-treated metal-free buffer (50 mM Tris pH 7.5/200 mM NaCl). The purity of the resulting fractions was ascertained to be >90% by SDS-PAGE analysis. Concentrations of the purified proteins were determined by using a ND-1000 NanoDrop spectrophotometer.

Enzyme kinetics experiments

Kinetics experiments were conducted at 296 K in Tris buffer (20 mM) and NaCl (200 mM) at pH 7.5. To a premixed solution of psBBOX (400 nM), $FeSO_4$ (50 μM), 2OG (1.5 mM), and ascorbate (5 mM) was added the substrate in a range of different concentrations. After 1 min, an aliquot (20 μL) of the reaction mixture was quenched with MeCN (80 μL). Subsequently the sample was analyzed by LC-MS. Each experiment was performed in duplicate.

NMR experiments

All NMR experiments were performed on a Bruker Avance III 700 MHz spectrometer equipped with a TCI inverse cryoprobe and on a Bruker Avance II 500 MHz spectrometer equipped with a 5 mm $^{13}C(^1H)$ dual cryoprobe at 298 K and data analyzed using Bruker Topspin 3.2. All spectra were processed with a Lorentzian line broadening of 0.3 Hz. The solutions were buffered in 50 mM Tris-D_{11}·HCl pH 7.5, in 90:10 H_2O/D_2O. Bruker MATCH 3 mm diameter and 5 mm NMR tubes, with total sample volumes of 160 μL and 500 μL, respectively, were used. For the psBBOX-catalyzed substrate-turnover experiments, the assay mixture was incubated in an Eppendorf tube and whenever necessary the reaction was quenched (stopped) with the addition of 1 M HCl (5 μL) and the spectrum was recorded for analysis. To measure the ligand binding

constant by psBBOX titration, separate samples were prepared. The PROJECT-CPMG pulse sequence (90°_x–[τ–180°_y–τ–90°_y–τ–180°_y–τ] n–acquisition), as described by Aguilar et al.,[27] was used to remove the broad resonances of the protein. The relaxation edited (CPMG) ^{1}H NMR experiments were recorded with a total filter time of 32 ms. Protein titration data were fitted using OriginPro 9.0 (Origin lab, Northampton, MA, USA) to calculate the ligand binding constant (K_D). Water suppression was achieved by presaturation.

Computational methods

For psBBOX protein-ligand docking simulation, the X-ray crystal structure of hBBOX in complex with NOG and γBB was employed (PDB entry: 3O2G) to make a homology structure of psBBOX using Modeller 9v4.[28] Careful attention was paid to the assignment of protonation states for Asp, Glu, His, and Lys residues. We calculated partial charge distribution by using quantum mechanical calculation at PBE1PBE/LanL2DZ level of theory.[29] Their partial charge is assigned with charges from electrostatic potentials using a grid-based method (CHELPG).[30] After reassigning CHELPG partial charges to X–γBB using the above quantum mechanical calculation, docking simulations of the X–γBB substrate with psBBOX were carried using the empirical AutoDock[31] scoring function improved by implementation of a new solvation model. The modified scoring function has the following form [Eq. (1)]:

$$\Delta G^{aq}_{bind} = W_{vdW}\sum_{i=1}\sum_{j>1}\left(\frac{A_{ij}}{r_{ij}^{12}}-\frac{B_{ij}}{r_{ij}^{6}}\right)+W_{hbond}\sum_{i=1}\sum_{j>1}E(t)\left(\frac{C_{ij}}{r_{ij}^{12}}-\frac{D_{ij}}{r_{ij}^{10}}\right)$$
$$+W_{elec}\sum_{i=1}\sum_{j>1}\frac{q_1q_2}{\varepsilon(r_{ij})r_{ij}}+W_{tor}N_{tor}+\Delta G_{sol}$$
$$=\sum_{i}^{atoms}\left\{S_i\left(O_i^{max}-\sum_{j}^{i\neq j}V_j e^{\frac{r_{ij}^2}{2\sigma^2}}\right)+P_i\sum_{j}^{i\neq j}V_j e^{\frac{r_{ij}^2}{2\sigma^2}}\right\} \quad (1)$$

where W_{vdW}, W_{hbond}, W_{elec}, W_{tor}, and W_{sol} are the weighting factors for the van der Waals, hydrogen-bond and electrostatic interactions, the torsional term, and the desolvation energy of the inhibitors, respectively. r_{ij} represents the interatomic distance and A_{ij}, B_{ij}, C_{ij}, and D_{ij} are related to the depths of the potential energy well and the equilibrium separations between the two atoms. The hydrogen bond term has an additional weighting factor, $E(t)$, representing the angle-dependent directionality. A cubic equation approach was applied to obtain the dielectric constant required to compute the interatomic electrostatic interactions between psBBOX and X–γBB. In the entropic term, N_{tor} is the number of sp^3 bonds in the ligand. In the desolvation term, S_i, P_i, and V_i are the solvation parameter, self-solvation parameter, and fragmental volume of atom i, respectively, whereas Occ_i^{max} is the maximum atomic occupancy.[32,33]

Acknowledgements

We thank the Netherlands Organisation for Scientific Research, the Wellcome Trust, Biotechnology and Biological Sciences Research Council, Cancer Research UK, the European Union for financial support. A.K. thanks Abdul Wali Khan University, Mardan, Pakistan for financial support.

Keywords: cation–pi interactions · C–H oxidation · enzyme catalysis · molecular recognition · oxygenases

[1] N. R. Rose, M. A. McDonough, O. N. F. King, A. Kawamura, C. J. Schofield, *Chem. Soc. Rev.* **2011**, *40*, 4364–4397.
[2] C. Loenarz, C. J. Schofield, *Nat. Chem. Biol.* **2008**, *4*, 152–156.
[3] S. Englard, J. S. Blanchard, C. F. Midelfort, *Biochemistry* **1985**, *24*, 1110–1116.
[4] K. Strijbis, F. M. Vaz, B. Distel, *IUBMB Life* **2010**, *62*, 357–362.
[5] C. J. Rebouche, H. Seim, *Annu. Rev. Nutr.* **1998**, *18*, 39–61.
[6] H. P. Kleber, *FEMS Microbiol. Lett.* **1997**, *147*, 1–9.
[7] J. Kerner, C. Hoppel, *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* **2000**, *1486*, 1–17.
[8] I. K. H. Leung, T. J. Krojer, G. T. Kochan, L. Henry, F. Von Delft, T. D. W. Claridge, U. Oppermann, M. A. McDonough, C. J. Schofield, *Chem. Biol.* **2010**, *17*, 1316–1324.
[9] K. Tars, J. Rumnieksa, A. Zeltins, A. Kazaks, S. Kotelovica, A. Leonciks, J. Sharipo, A. Viksna, J. Kukac, E. Liepinsh, M. Dambrova, *Biochem. Biophys. Res. Commun.* **2010**, *398*, 634–639.
[10] A. M. Rydzik, I. K. H. Leung, G. T. Kochan, M. A. McDonough, T. D. W. Claridge, C. J. Schofield, *Angew. Chem. Int. Ed.* **2014**, *53*, 10925–10927.
[11] M. Dambrova, E. Liepinsh, I. Kalvinsh, *Trends Cardiovasc. Med.* **2002**, *12*, 275–279.
[12] L. Henry, I. K. H. Leung, T. D. W. Claridge, C. J. Schofield, *Bioorg. Med. Chem. Lett.* **2012**, *22*, 4975–4978.
[13] A. M. Rydzik, I. K. H. Leung, G. T. Kochan, N. D. Loik, L. Henry, M. A. McDonough, T. D. W. Claridge, C. J. Schofield, *Org. Biomol. Chem.* **2014**, *12*, 6354–6358.
[14] S. D. Taverna, H. Li, A. J. Ruthenburg, C. D. Allis, D. J. Patel, *Nat. Struct. Mol. Biol.* **2007**, *14*, 1025–1040.
[15] L. M. Salonen, M. Ellermann, F. Diederich, *Angew. Chem. Int. Ed.* **2011**, *50*, 4808–4842; *Angew. Chem.* **2011**, *123*, 4908–4944.
[16] G. N. Nagy, L. Marton, A. Contet, O. Ozohanics, L. M. Ardelean, Á. Révész, K. Vákey, F. D. Irimie, H. Vial, R. Cerdan, B. G. Vértessy, *Angew. Chem. Int. Ed.* **2014**, *53*, 13471–13476.
[17] D. A. Dougherty, D. A. Stauffer, *Science* **1990**, *250*, 1558–1560.
[18] J. P. Gallivan, D. A. Dougherty, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 9459–9464.
[19] J. C. Ma, D. A. Dougherty, *Chem. Rev.* **1997**, *97*, 1303–1324.
[20] H. A. Lester, M. I. Dibas, D. S. Dahan, J. F. Leite, D. A. Dougherty, *Trends Neurosci.* **2004**, *27*, 329–336.
[21] J. J. A. G. Kamps, J. Huang, J. Poater, C. Xu, B. J. G. E. Pieters, A. Dong, J. Min, W. Sherman, T. Beuming, F. M. Bickelhaupt, H. Li, J. Mecinović, *Nat. Commun.* **2015**, *6*, 8911.
[22] D. A. Dougherty, *Acc. Chem. Res.* **2013**, *46*, 885–893.
[23] D. L. Ziering, R. A. Pascal, *J. Am. Chem. Soc.* **1990**, *112*, 834–838.
[24] R. W. D. Welford, J. M. Kirkpatrick, L. A. McNeill, M. Puri, N. J. Oldham, C. J. Schofield, *FEBS Lett.* **2005**, *579*, 5170–5174.
[25] L. M. Salonen, M. C. Holland, P. S. J. Kaib, W. Haap, J. Benz, J. L. Mary, O. Kuster, W. B. Schweizer, D. W. Banner, F. Diederich, *Chem. Eur. J.* **2012**, *18*, 213–222.
[26] K. Schärer, M. Morgenthaler, R. Paulini, U. Obst-Sander, D. W. Banner, D. Schlatter, J. Benz, M. Stihle, F. Diederich, *Angew. Chem. Int. Ed.* **2005**, *44*, 4400–4404; *Angew. Chem.* **2005**, *117*, 4474–4479.
[27] J. A. Aguilar, M. Nilsson, G. Bodenhausen, G. A. Morris, *Chem. Commun.* **2012**, *48*, 811–813.
[28] A. Šali, T. L. Blundell, *J. Mol. Biol.* **1993**, *234*, 779–815.
[29] J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
[30] F. Martin, H. Zipse, *J. Comput. Chem.* **2005**, *26*, 97–105.
[31] G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, A. J. Olson, M. E. T. Al, *J. Comput. Chem.* **1998**, *19*, 1639–1662.
[32] H. Kang, H. Choi, H. Park, *J. Chem. Inf. Model.* **2007**, *47*, 509–514.
[33] H. Choi, H. Kang, H. Park, *J. Cheminf.* **2013**, *5*, 8–21.

Received: September 18, 2015
Published online on December 14, 2015

