THE UNIVERSITY *of* EDINBURGH

Structural Dynamics and Ligand Binding in Kynurenine-3-monooxygenase

Martin Wilkinson

Supervised by Dr Chris Mowat

Doctor of Philosophy
University of Edinburgh
2013

Declaration

The work presented in this thesis is the original work of the author, except where specific reference is made to other sources. It has not been submitted in part, or in whole, for any other degree. The results are currently being prepared for publication.

Martin Wilkinson

Acknowledgements

I would like to acknowledge the excellent support, encouragement and supervision I received from my supervisors Dr Chris Mowat and Dr Simon Daff and I am very grateful I had the opportunity to work with them and their respective group members. Thanks to all the past and present members of the groups, especially including Dr George Pantouris, Dr Sarah Thackray, Dr Laura Campbell and Dr Ben Gazur, have always been quick and enthusiastic in offering advice and assistance.

I would also like to thank the members of the structural biology groups headed by Prof. Malcolm Walkinshaw and Dr Julia Richardson for being very helpful when I visited and for integrating me into their synchrotron data collection trips. In addition I am very grateful to Dr Iain McNae for help and guidance during the process of solving my first protein crystal structure.

I am very grateful for the friends and colleagues I have met during my time in Edinburgh, experiments tend to work better in a positive environment! In particular I am grateful to a very special person I have met who has been a constant source of inspiration and has reminded me that there's more to life than lab work.

Finally I am grateful to family and friends who have encouraged and supported me throughout every stage of my education to date.

Abstract

Kynurenine 3-monooxygenase is a FAD-dependent aromatic hydroxylase (FAH) which is a widely suggested therapeutic target for controlling the balance of bioactive metabolite levels produced by the mammalian kynurenine pathway (KP). Prior to starting this work no structural information was known for the enzyme, with studies of the human form complicated by the presence of a C-terminal transmembrane helix. The bacterial *Pseudomonas fluorescens* enzyme (*Pf*KMO) lacks the transmembrane region and has been previously characterised by Crozier-Reabe and Moran [1, 2]. Therefore *Pf*KMO, which shares 32 % sequence identity with the human enzyme, was selected as a target for structure solution. Initial substrate bound *Pf*KMO crystals showed poor X-ray diffraction. Subsequent growth optimisation and the generation of a C252S/C461S *Pf*KMO mutant (dm2) yielded crystals suitable for structure solution. Selenomethionine-labelled substrate bound dm2 crystals were used to solve the first structure to a resolution of 3.40 Å. With just one protein molecule per asymmetric unit, a high solvent content was responsible for the poor diffraction properties of this crystal form. The overall fold resembled that of other FAH enzymes with a Rossmann-fold based FAD-binding domain above a buried substrate binding pocket. Interestingly *Pf*KMO possesses an additional, novel C-terminal domain that caps the back of the substrate-binding pocket on the opposite side to the flavin. Residues proposed to be involved in substrate binding were identified and shown to be highly conserved among mammalian KMO sequences.

Subsequently single crystals of substrate-free dm2 *Pf*KMO were obtained and showed significantly stronger diffraction due to new lattice packing in an orthorhombic space group bearing four molecules per asymmetric unit. The structure was solved to a resolution of 2.26 Å and revealed a clear conformational change of the novel C-terminal domain. This movement opens a potential route of substrate/product exchange between bulk solvent and the active site. The investigation of a set of C-terminal mutants further explored the relevance and mechanics of the conformational change. In addition the presence of chloride ions in the substrate-free crystal growth solution caused a small number of localised subtle alterations to the structure, with a potential chloride binding

site identified adjacent to the flavin cofactor. This may have relevance for the observed inhibition of *Pf*KMO activity by monovalent anions – a feature widely common to FAH enzymes [3].

The first discovered KMO inhibitors were analogues of the substrate L-Kyn, however one such compound (*m*-NBA) was recently shown to instigate uncoupled NADPH oxidation leading to the release of cytotoxic hydrogen peroxide [1]. A set of substrate analogues were tested and characterised for inhibition of *Pf*KMO. The picture was shown to be complex as some substrate analogues completely inhibited the enzyme whilst the binding of some still stimulated low-levels of NADPH oxidation. Crystallographic studies with *m*-NBA and 3,4-dichlorobenzoylalanine (3,4-CBA) bound revealed indistinguishable structures from that of substrate-bound *Pf*KMO. These studies suggest that the analogue 3,4CBA is a potent *Pf*KMO inhibitor whose therapeutic potential may be re-visited. The previous most potent KMO inhibitor whose structure was not analogous to the substrate was Ro 61-8048 [4], which unfortunately did not pass pre-clinical safety tests. A novel series of 1,2,4-oxadiazole amides based on the structure of Ro 61-8048 was created by Gavin Milne (PhD, University of St Andrews) and tested on *Pf*KMO. Rounds of refinement led to the discovery and refinement of low nanomolar competitive inhibitors of the bacterial enzyme. *Pf*KMO was co-crystallised with each of the four most potent compounds forming a third different lattice arrangement, which yielded structures to resolutions of 2.15-2.40 Å. The structures displayed conformational changes resembling the substrate-free fold possibly caused by displacement of a crucial substrate-binding residue, R84.

Overall the wealth of structural data obtained may be transferable to predictions about the structural features of human KMO and to the rational design of therapeutic inhibitors. The potent novel inhibitors tested may additionally present a new exciting development for the therapeutic inhibition of human KMO.

Abbreviations and units

1. Amino acids

The single and three letter amino acid codes listed below are used. Mutations are described by the symbol of the initial amino acid, the residue number and then the new residue symbol, e.g. R386T.

Amino acid	Three letter code	Symbol
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartate	Asp	D
Cysteine	Cys	C
Glutamate	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

2. Nucleotides

When referring to nucleotide sequences, the single letter abbreviations A, C, G and T are used for adenine, cytosine, guanine and thymine respectively.

3. Abbreviations

3-HK	3-hydroxykynurenine
3-HAA	3-hydroxyanthranilic acid
3,4-CBA	3,4 dichlorobenzoylalanine
°	Degrees (angles)
°C	Degree Celsius
λ	Wavelength
Å	Angstrom
A_{xxx}	Absorbance at *xxx* nm wavelength
ΔA_{xxx}	Change in absorbance in a set time period
AP	Acute pancreatitis
ATP	Adenosine-5'-triphosphate
AVH	Aklavinone-11-hydroxylase
B-factor	Temperature factor – measure of mobility of atoms within a structure
βME	β-mercaptoethanol
CNS	Central nervous system
Da (kDa)	Daltons (kilodaltons)
DNA	Deoxyribonucleic acid
dm2	C252S/C461S double mutant
DHPH	2,6-dihydroxypyridine 3-hydroxylase
DMSO	Dimethyl sulphoxide
DTT	Dithiothreitol
E.coli	*Escherichia coli*
EG	Ethylene glycol
FAD	Flavin adenine dinucleotide
FAH	Class A flavin-dependent aromatic hydroxylase
g (ng/μg/mg/kg)	Grams (nano/micro/milli/kilograms)
GSH	Glutathione
h	Hours
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulphonic acid
HpxO	Urate oxidase
HS	Data for the highest resolution shell
I/σI	Measure of the ratio of signal:noise of recorded reflections
IC_{50}	The concentration of inhibitor corresponding to 50 % enzyme activity

IPTG	Isopropyl β-D-1-thiogalactopyranoside
KAT	Kynurenine aminotransferase
k_{cat}	Measure of the number of turnovers catalysed per second per enzyme
K_m	Michaelis-Menten kinetic parameter for substrate affinity
KMO	Kynurenine-3-monooxygenase
KP	Kynurenine pathway
KYNA	Kynurenate
LB	Luria broth
L-Kyn	L-kynurenine
M (nM/μM/mM)	Molar (nano/micro/millimolar)
m (nm/mm)	Meter (nano/millimetre)
mg	Milligrams
MHPCO	2-Methyl-3-hydroxypyridine-5-carboxylic acid oxygenase
mins	Minutes
***m*-HBH**	3-Hydroxybenzoate hydroxylase
***m*-NBA**	*Meta*-nitrobenzoylalanine
MOPEG	Methoxy poly-ethyleneglycol
MPD	3-Methyl-1,5-pentadiol
MWCO	Molecular weight cut-off
NAD/NAD^+	Nicotinamide adenine dinucleotide
NADPH/$NADP^+$	Nicotinamide adenine dinucleotide phosphate
NMDAr	N-methyl-D-aspartic acid receptor
OD_{xxx}	Optical density at *xxx* nm
pdb	Protein databank
PEG*x*	Poly-ethyleneglycol, average molecular weight of *x*
Pf	*Psuedomonas fluorescens*
***p*-HBH**	*Para*-hydroxybenzoate hydroxylase
PhH	Phenol hydroxylase
pIC_{50}	Equivalent to value of $Log(1/IC_{50})$
PMSF	Phenylmethanesulphonylfluoride
QUIN	Quinolinate
R_{merge}	Measure of agreement of multiple measurements of the same reflection
R	Measure of agreement between observed and calculated crystal data
R_{free}	As above but for a sub-set of reflections excluded from refinement

rmsd	Root means square deviation
s	Seconds
SDS	Sodium dodecyl sulphate
SeMet	L-selenomethionine
TetX	Tetracycline monooxygenase
TLS	Translation libration screw-motion
UV	Ultraviolet
Vis	Visible
V_{max}	Michaelis-Menten kinetic parameter for the maximal reaction rate

Contents

Chapter 1 – Introduction

1.1 The kynurenine pathway

1.1.1 Overview of the kynurenine pathway

Tryptophan is an amino acid that the human body must obtain externally. In mature mammals, the cellular tryptophan pool for protein synthesis is mostly maintained by protein degradation. This leads to >90 % of dietary tryptophan being catabolised via the kynurenine pathway (KP) (figure 1.1), into NAD^+, with the rest utilised for the synthesis of serotonin or tryptamine. NAD^+ is imperative to numerous cellular processes including ATP production and poly(ADP-ribose)polymerase activity. The KP is the major route for NAD^+ synthesis. Along the pathway numerous stable intermediates are generated, several of which have physiological and/or pathophysiological implications. This combined with the presence of pathway branch points has meant that control of the activities of enzymes acting along the pathway has proved crucial for maintaining a healthy balance of the metabolites. As a result the KP has been implicated in the pathophysiology of a diverse range of diseases and disorders including neurodegenerative disorders.

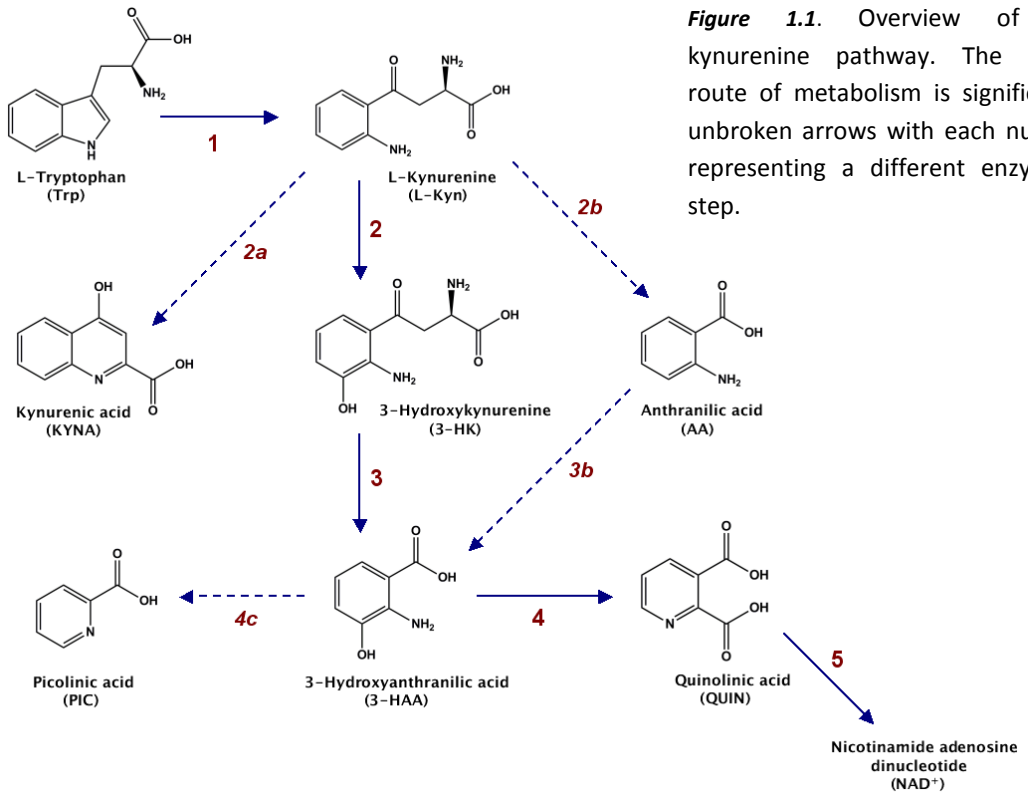


Figure 1.1. Overview of the kynurenine pathway. The major route of metabolism is signified by unbroken arrows with each number representing a different enzymatic step.

1.1.2 Regulation of the ratio of Trp:L-Kyn

The first step of the KP is the enzymatic cleavage of the indole-ring of tryptophan (Trp) into the transient intermediate formylkynurenine (FKyn). FKyn is rapidly converted into the first stable intermediate of the pathway, L-kynurenine (L-Kyn). The oxidation of Trp is catalysed by the enzymes tryptophan 2,3-dioxygenase (TDO) in the liver and indoleamine 2,3-dioxygenase 1 (IDO-1) and 2 (IDO-2) extra-hepatically. TDO and IDO may both be heme-proteins that catalyse the same reaction however their sequences and structures are unrelated with differential properties regarding: substrate selectivity, specific inhibition and activation/up-regulation. Certain cytokines and pro-inflammatory molecules released during an immune response have been found to stimulate the up-regulation of IDO-1 expression and activity. The most potent stimulant was found to be interferon gamma (IFN-γ) [5, 6] with stimulation also evident with IFN-β, tumour necrosis factor α (TNF-α), lipopolysaccharides, amyloid β peptide 1-42, human immunodeficiency virus (HIV) proteins and interleukin-Iβ [7-11]. IDO activation depletes Trp levels and has been shown to be critically involved in the mechanism of antimicrobial and antitumoral defence [12, 13].

In the past decade it has become clear that inflammation is a primary pathophysiological mechanism in many chronic illnesses [14, 15]. Raised levels of pro-inflammatory factors and cytokines have been associated with and shown to contribute to the development of dementia [16, 17], Alzheimer's disease (AD) [18, 19] amyotrophic lateral sclerosis [20] and major depression [21] among others. The immune-stimulated up-regulation of IDO-1 reduces serum Trp concentrations by 25-50 %, which raises concentrations of L-Kyn whilst proportionally decreasing the amount of Trp available for the generation of serotonin [22, 23]. The subsequent decrease in Trp levels has been shown to efficiently inhibit T-cell proliferation [24, 25] leading to an implication in the development of immunodeficiency and immunotolerance. Lowered Trp and raised L-Kyn concentrations have been discovered in patients with end-stage renal disease [26] and ovarian cancer [27], and have been directly correlated with decreased levels of $CD4^+$ T cells in HIV patients [28-30]. In addition to this, up-regulated IDO-1 was found to be specifically associated with neurofibrillary tangles and senile plaques [31] – both hallmarks of the pathophysiology of AD – and with severity of chronic kidney disease [32] among others.

Inflammatory immunoactivation can also influence the activities of more KP enzymes downstream of IDO, as shown in several human cell lines [33-35]. This altered KP regulation also affects the balance of levels of other KP intermediates that might be influential in pathophysiology, as we will see in the following sections. A complicating factor in the study of the KP in models of disease though is the fact that expression and activation of the KP enzymes varies between: species, cell types and even throughout the development of certain cell lines [35-39]. As a result care must be taken in assessing the therapeutic relevance of findings relating to the KP in animal models of disease. Such models are still useful to further our broader understanding of the physiological and pathophysiological involvements of KP metabolites.

1.1.3 Neurotoxicity versus neuroprotection; QUIN and KYNA

Interest in KP metabolites was prompted when kynurenate (KYNA) and quinolinate (QUIN) were found to act as antagonist and agonist respectively on the excitatory glutamate receptor NMDAr in the CNS [40, 41]. NMDAr over-stimulation causes an influx of calcium ions that causes disruption of mitochondrial functions and leads to the generation of cytotoxic reactive oxygen and nitrogen species (ROS and RNS) [42-44]. This process is known as excitotoxicity and ultimately leads to cell death. The question arose – could one KP metabolite QUIN contribute to excitotoxicity whilst another KYNA offered a counteracting neuroprotective role? With the two metabolites on separate branches of the KP, misregulation of the enzymes governing each route and therefore the imbalance of these two molecules could be a factor in the pathophysiology of numerous diseases. If true this would create a therapeutic opportunity to restore this balance by restricting the QUIN-generating route of the KP as discussed extensively elsewhere [45-49].

1.1.3.1 KYNA and neuroprotection

The biggest query regarding the hypothesis above was surrounding the hoped neuroprotective ability of KYNA. At high concentrations, KYNA blocks both the glutamate-binding site of the NMDAr, however at low concentrations it can block the

allosteric glycine site of the NMDAr as well as the α7-nicotinic acetylcholine receptor (α7-nAChr) [47]. The IC_{50} against the α7-nAChr was determined to be only 7 μM. The strength of KYNA antagonism at the NMDAr has been published as low to high micromolar values – for example the IC_{50} for KYNA competing with glycine was determined to be 238 μM [50] and around 8 μM [51] in separate studies with cultured neurons. In the rat brain it was determined that 10-40 μM KYNA was required to substantially inhibit hippocampal synaptic transmission [52]. Under physiological conditions, KYNA is present in the human brain at nanomolar concentrations (up to 1 μM) and in rat brains it was measured at <100 nM [53]. As a result speculation arose as to whether KYNA levels in the brain ever reach levels sufficient to significantly inhibit the NMDAr and counter the stimulation by QUIN.

Numerous studies have noted the neuroprotective effect of small increases in levels of KYNA with a reduction in locomotor activity, convulsion strength and frequency and excitotoxic neuronal damage [54-59]. Using *in vivo* and *ex vivo* models, post-ischemic brain damage was significantly reduced as a result of two to three-fold increases in KYNA concentrations [60, 61]. An increase in KYNA at low nM concentrations directly and drastically reduced extracellular brain Glu content in the rat caudate [62]. It has therefore been speculated that in fact the local synaptic KYNA concentrations may be significantly higher than the overall measurements made from whole tissue homogenizations [52]. Therefore in specific areas levels of KYNA may reach sufficient levels to show significant inhibition of excitatory amino acid receptors and block synaptic transmission. Otherwise it may be possible that KYNA additionally acts via an alternate NMDAr-independent mechanism to decrease concentrations of Glu. It has been reported that KYNA selectively activates a G-coupled protein receptor (GPR35) [63], which may contribute towards the observed reduction of Glu release [64].

Decreased levels of KYNA have been associated with the pathophysiology of several diseases including major depression [65], Alzheimer's disease [66] and renal cancer [67]. After twenty-two years of following patients who suffered from schizophrenia, those who attempted suicide had lower concentrations of KYNA in the CSF (0.8 nM compared to 1.7 nM) [68]. Outside of the CNS, KYNA concentrations in the kidney were found to be over three-times higher than healthy controls in kidney tumour tissue [67]. Moreover, *in*

vitro experiments showed that high micromolar concentrations of KYNA could significantly inhibit proliferation, DNA synthesis and migration of renal cancer Caki-2 cells. The mechanism of this inhibition was shown to involve KYNA-dependent over-expression of the cyclin-dependent kinase inhibitor p21 Waf1/Cip1 [69]. So outside of the CNS artificially-raised KYNA levels may be beneficial in the therapeutic treatment of cancers.

On the other hand, raised levels of KYNA have also been associated with several diseases including: the pathophysiology of schizophrenia [42, 70, 71], the high invasiveness and rapid progression of specifically adenocarcinoma among lung cancer patients [72], complex partial seizures in epilepsy patients [73] and human cataractous lenses [74]. In another example, a study of patients with long-term cognitive impairment associated with tick-borne encephalitis discovered dramatically raised levels of KYNA [25]. Sufferers had mean and maximum CSF KYNA concentrations of 5.3 nM and around 70 nM respectively compared to healthy age-matched controls who averaged 0.99 nM with a maximum concentration of 4 nM. Therefore there appears to be a need to maintain a balance between too little and too much KYNA depending on the specific subcellular location - but particularly in the brain - as recently reviewed elsewhere [75].

1.1.3.2 QUIN-mediated cytotoxicity

Studies of the potential dangers associated with QUIN have been more conclusive. QUIN has grown into a notorious cytotoxic villain since it was discovered to directly cause necrosis upon intracerebral injection [76] and when added to *in vitro* neuron cultures [77]. Raised levels of QUIN have since been shown to directly contribute to cytotoxicity in cases of AIDS dementia complex caused by infection with HIV-1 [78-80], Poliovirus brain infection [81], cerebral malaria [82], multiple sclerosis [83], cerebral ischemia [84] , epilepsy [85, 86], Huntington's disease [87, 88] and Alzheimer's disease [89, 90].

Physiological levels of QUIN have been determined to be nanomolar in the brain [91] and nano to micromolar in the CSF [92]. Sub-physiological QUIN concentrations greater than 150 nM have been shown to be cytotoxic in primary cultures of human neurons [93].

These effects were significantly attenuated by treatment with the NMDAr blockers: MK801 and memantine, as well as with the nitric oxide synthase (NOS) inhibitor, L-NAME. Interestingly it has been suggested that the enzyme responsible for directly catabolizing QUIN – QPRTase – is easily saturated at low QUIN concentrations [94]. *In vitro,* QPRTase activity began to saturate at around 300 nM QUIN [95]. This could explain the apparent CNS accumulation of QUIN seen in the aforementioned diseases and disorders. The most dominant cells in the CNS – astrocytes – are unable to generate QUIN as they do not express the enzyme KMO in humans [34]. The majority of QUIN within the brain is in fact produced by infiltrating macrophages and immune-activated resident microglia. It is believed that the elevated L-Kyn surrounding astrocytes can be used by such neighbouring macrophages and microglia to produce extra QUIN, particularly during a local inflammatory response [34].

Raised levels of QUIN are known to be involved in several cytotoxic mechanisms including some independent of NMDAr stimulation, as extensively reviewed recently [94, 96]. Firstly, QUIN has been shown to form a complex with Fe^{2+}, which increased the generation of hydroxyl radicals that directly promoted lipid peroxidation and DNA damage [97]. Parallel to this, intrastriatal injection of QUIN in rats was shown to stimulate lipid peroxidation within two hours and this correlated with raised extracellular levels of hydroxyl radicals [98, 99]. Oxidative stress via free-radical generation has emerged as the major route of QUIN-mediated cytotoxicity [100]. Antioxidants including melatonin, 6-hydroxymelatonin, selenium, L-carnitine, spermine and statins, among others, suppress the destructive consequences of QUIN administration, whereas they are ineffective on the cytotoxicity induced by other NMDAr agonists and NMDA itself [99-106]. Methods of reducing QUIN concentrations via the inhibition of KP enzymes such as KMO may be an effective therapeutic supplement to reducing the pathophysiological symptoms of many neurological diseases and disorders.

1.1.4 Cytotoxicity mediated by 3-HK

The KP intermediate 3-HK, which is produced from L-Kyn by the enzyme KMO, has been shown to cause cytotoxicity by the generation of ROS. In neuronal hybrid cell line cultures, addition of 3-HK to concentrations greater than 100 μM was toxic to more than 85 % of the cells over a period of 24 hrs [107]. A role for oxidative stress – specifically via hydrogen peroxide build up and metal ions – was implicated as catalase, glutathione, peroxidase or iron chelator (desferrioxamine) treatment significantly attenuated the cytotoxicity caused by 3-HK [108]. In a later study, intrastriatal injection of just 50 nM 3-HK into rat brains induced local tissue damage with greater sensitivity than via injection of QUIN [109]. Moreover the potent NMDAr antagonist MK-801 attenuated only QUIN-induced toxicity whereas the antioxidant *N*-acetyl-L-cysteine reduced 3-HK induced damage. Interestingly co-administration of just 5 nM 3-HK and 15 nM QUIN – levels of each harmless by themselves – to rat brains caused substantial neuronal loss [110]. Under physiological conditions, 3-HK was shown to form reactive autoxidation products whilst also generating cytotoxic hydrogen peroxide and hydroxyl radicals [111]. These studies led to the suggestion that 3-HK-mediated cytotoxicity proceeded differently to the QUIN NMDAr activation mechanism via the generation of free-radical forming ROS and ultimately causes cell death with apoptotic features [112].

At physiological concentrations, 3-HK has been shown to have beneficial roles in UV absorption and as an antioxidant [113-117]. In humans the average serum 3-HK concentration under physiological conditions has been measured to be 383 nM [73]. In primary cultures of mixed cortical cells, the minimum concentration of 3-HK required to cause a significant degree of neurotoxicity was 100 μM and just 1 μM in relation to exposure times of 24 hrs and 72 hrs respectively [112]. Similarly 3-HK concentrations of 1-100 μM caused cell death to primary neuronal cultures from the rat striatum [118]. Significantly raised 3-HK levels have been correlated with: vitamin B6 deficiency [119, 120], the prevalence of increased oxidative stress and cardiovascular disease among patients with end-stage renal disease [26], chronic renal insufficiency [121], bacterial endotoxin exposure in rat CNS [122], patients with Alzheimer's disease [123], Huntington's disease [124-127], development of cataracts in human lens [128-131], HIV-

associated dementia [132], experimental malaria in mice [82] a decline in memory and learning in Drosophila [133] and an experimental rat model of multiple sclerosis [83].

Acute pancreatitis (AP) involves inflammation of the pancreas which can be mild and self-limiting, however in 20% of cases it leads to multiple organ failure with a 30-50% mortality rate. In 2008, Mole *et al* [134] showed the mesenteric lymph to be directly involved in the development of acute lung injury in AP. Moreover they showed a correlation linking raised mesenteric kynurenine and 3-HK levels to the severity of AP lung injury in rats. Raised kynurenine levels were also observed in the plasma of AP patients. The cytotoxicity of AP mesenteric lymph on lung HCT116 cells was reproduced *in vitro* with exogenous 3-HK. Whilst the cause of raised 3-HK levels was not established, it was suggested that decreasing mesenteric lymph 3-HK levels could be an effective way of preventing the development of multiple organ failure in AP patients.

1.1.5 Other bioactive KP metabolites

The next intermediate generated in the KP after 3-HK is 3-HAA. Under physiological conditions, 3-HAA has implications as an antioxidant in rat cerebral cortex [117] and has been suggested as a natural regulator of T-cell expansion and activation [135]. At different concentrations it significantly inhibited NF-κB activation (10-50 μM 3-HAA) and caused apoptosis (100-500 μM 3-HAA) specifically among CD4 T-cells. At low μg doses, 3-HAA treatment significantly attenuated the symptoms of experimental asthma in mice, linked to the suppression of T-helper cell mediated inflammation. Like 3-HK, 3-HAA can be oxidised to generate ROS in the presence of Cu(II) and molecular oxygen [136]. Reduced serum 3-HAA levels were associated with epilepsy in a mouse model and in patients [137]. This could possibly involve overactive 3-HAA oxygenase converting extra 3-HAA into the excitotoxin QUIN. A decreased ratio of 3-HAA:AA has been detected in a range of disorders as summarised in 2010 [138] and was speculated to be a protective response. Moreover AA is an effective competitive inhibitor of 3-HAA oxygenase [35] and so its build-up would largely prevent the generation of QUIN to reduce local cytotoxicity.

The product of 3-HAA oxygenase is actually a transient intermediate (2-amino-3-carboxymuconate-6-semialdehyde), which subsequently undergoes a non-enzymatic transformation into QUIN. This transient intermediate can also be competitively and enzymatically converted into another stable KP metabolite – PIC. Little was really known about the biological significance of this molecule a part from its metal chelating properties [139]. PIC has shown uncharacterised antagonistic properties towards the cytotoxicity of QUIN [140, 141]. High PIC concentrations of 1-4 mM have shown an enhancement of macrophage effector functions and anti-viral properties [142]. The likelihood of these effects being reproduced in humans is small as physiological PIC concentrations have been reported to be thousands of times smaller in the low to mid nanomolar range [143].

1.1.6 Summary of the KP

In all the KP presents a complex balance of biologically active intermediates with serious implications in the pathophysiology of numerous inflammatory diseases (see table 1.1). The pathway branches at the stable intermediate L-Kyn with three possible routes. Damaging irregularities in intermediate levels could be corrected by the inhibition of KMO and subsequently of its associated pathway branch. This would suppress levels of the cytotoxic 3-HK and QUIN whilst raising levels of the neuroprotective compound KYNA. Normal levels of NAD production can be maintained by the side-branch via the intermediate AA. Several reviews have suggested that KMO inhibition may provide a key therapeutic tool in the treatment of many of the aforementioned conditions associated with the KP [48, 49, 88, 144, 145].

KP metabolite	*In vivo* Properties	Raised levels associated with…	Decreased levels associated with…
KYNA	Reduces glutamate as an NMDAr antagonist and possibly by activating GPR35, increases expression of the cyclin-dependent kinase inhibitor p21 Waf1/Cip1	Neuroprotection, Schizophrenia, epilepsy, cognitive impairment, adenocarcinoma	Neurodegenerative diseases, depression, renal cancer
3-HK	UV absorption in lens, antioxidant, Cytotoxic via generation of ROS including hydrogen peroxide	Multiple organ failure in AP, ischemic damage, end-stage renal disease, neurodegenerative diseases, cataracts, malaria, multiple sclerosis	None known
3-HAA	Antioxidant, Inhibits NF-κB inducing apoptosis in CD4 T cells to restrain T cell expansion and activation		Epilepsy
PIC	Metal chelator, may have antiviral effect	None known	None known
QUIN	NMDAr agonist linked to excitotoxicity, cytotoxic via free radical generation, involved in lipid peroxidation	Ischemic damage, neurodegenerative diseases, malaria, multiple sclerosis, epilepsy	None known

***Table 1.1*.** Summary of the biological and pathophysiological properties of metabolites generated by the KP.

1.2 Flavin-dependent aromatic hydroxylases

1.2.1 Biological flavin cofactors

The biological flavins – flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD) – have key catalytic and electron transport roles in metabolic pathways. The key conserved flavin moiety is a tricyclic isoalloxazine ring, which can exist in several redox states. These are illustrated in figure 1.2. Interestingly flavins are able to accept and donate either one or two electrons and as such have been utilised in a variety of

biochemical processes. Oxidised flavins are photoreducible, with a yellow colour that turns to blue for the semiquinone radical and red for the anionic state. When totally reduced, flavins are colourless. They also store energy, which is released as high-energy electrons upon oxidation due to the stabilisation effect created by the formation of a more conjugated aromatic ring system. This process is utilised in the electron transport chain in the mitochondria. The different flavin redox states are also crucial to catalysing and controlling many metabolic reactions. For example when in the reactive reduced state flavins can react with oxygen as will be discussed in the following sections.

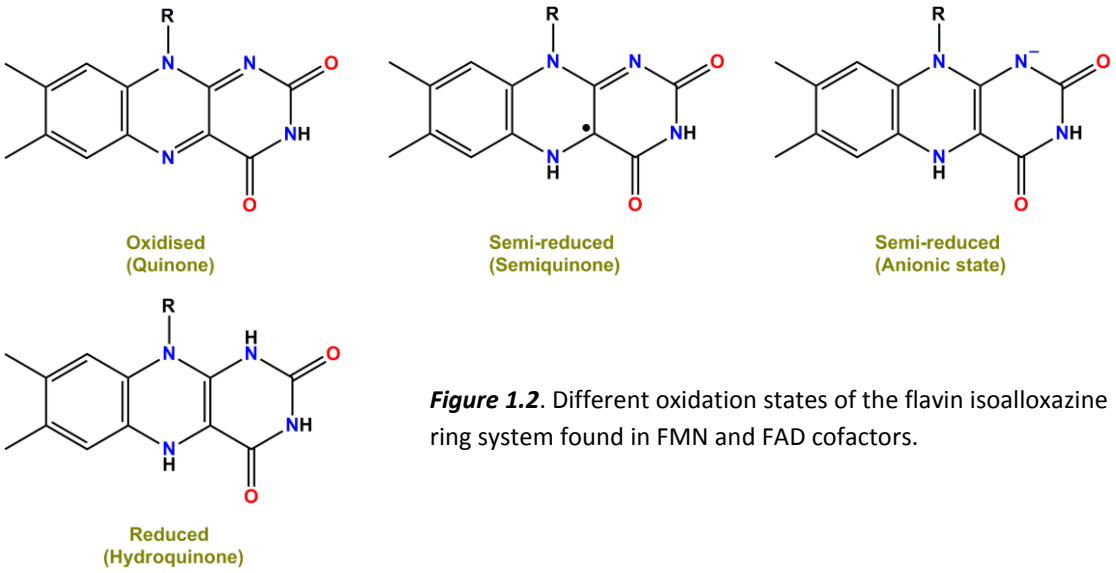


Figure 1.2. Different oxidation states of the flavin isoalloxazine ring system found in FMN and FAD cofactors.

1.2.2 Flavoprotein monooxygenases

In free solution, the reaction between reduced flavins and molecular oxygen generates hydrogen peroxide. The rate of reaction however is slow due to the need for oxygen spin inversion and the formation of unstable radical intermediates [146]. Molecular oxygen has a triplet spin in its ground state with two unpaired electrons whilst organic compounds such as flavins are diamagnetic singlets. Therefore in the reaction with oxygen, the highly reactive radical species superoxide and flavin semiquinones are formed. Flavoprotein monooxygenases/hydroxylases (FHs) have evolved to increase this rate of reaction by developing specialised local environments to stabilise particular oxy-flavin intermediates. As more FHs have been discovered and studied over the past

decade, they have been extensively reviewed and categorised [3, 147, 148]. These enzymes stabilise either a C4a-peroxyflavin or a C4a-hydroperoxyflavin intermediate to catalyse the respective reaction with their substrate(s). The very reactive nature of such intermediates is then utilised for varied catalytic tasks including hydroxylation, epoxidation, halogenation and more (see figure 1.3). Subsequently water is released to return the flavin to the oxidised state prior to the next round of catalysis. Without sufficient control over flavin reduction, the C4a-(hydro)peroxyflavin can decay to release the strong oxidiser hydrogen peroxide. To prevent this uncoupling of flavin reduction, FHs employ two different mechanisms whereby either flavin-reduction is ineffective until substrate is bound or after reduction the $NAD(P)^+$ remains tightly bound stabilizing the intermediate.

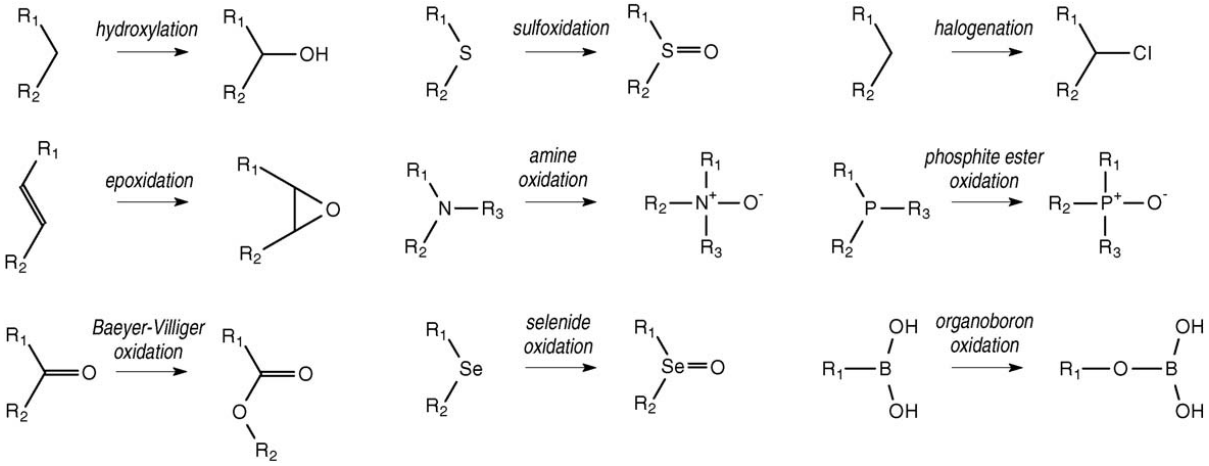


Figure 1.3. Summary of the different catalytic processes catalysed by flavoprotein hydroxylases [147]

All FHs non-covalently bind either an FAD or FMN flavin cofactor and most are classed as external as they reduce their flavin with electrons from free coenzymes in solution. Some rare examples of internal FHs exist where the flavins are reduced by the respective substrate, as is the case for lactate monooxygenase [149]. 'Life' is complicated for external FHs as the reactions they catalyse involve two chemical processes requiring very different environments. The first reaction involves flavin reduction and requires flavin exposure to solvent. Following from this is the flavin reaction with oxygen and subsequent catalysis. The reactive flavin-oxygen intermediate can only be maintained by shielding it from solvent. Two-component external FHs use two enzymes to overcome

this; with one involved in reducing the flavin cofactor and then transferring it to the other for catalysis [148]. Single-component external FHs can coordinate both requirements in one enzyme using different conformational states (See section **1.2.3**). In 2006, external FHs were categorised into six sub-classes A-F based on similarity in sequence or in structural characteristics [147].

1.2.3 Class A flavin-dependent aromatic hydroxylases

The enzyme KMO belongs to subclass A of the external FHs, known as flavin-dependent aromatic hydroxylases (FAHs). FAHs are encoded by a single gene with one dinucleotide binding domain that contains a tightly non-covalently bound FAD cofactor. Flavin reduction results in immediate release of $NAD(P)^+$. The intermediate formed is a C4a-hydroperoxyflavin which catalyses mainly hydroxylation reactions via electrophilic attack of an aromatic substrate which typically contains an activating hydroxyl or amino group. Flavin reduction is stimulated only upon substrate binding to prevent uncoupled degradation of the flavin-oxygen intermediate to release hydrogen peroxide. In *para*-hydroxybenzoate hydroxylase (*p*-HBH) the presence of substrate stimulated the rate of flavin reduction by NADPH by a factor of >100,000-fold [150]. Catalysis by FAHs displays a marked sensitivity to the presence of small monovalent anions such as Cl^-. In *p*-HBH, inhibition by halogen ions was studied in detail and shown to be competitive with NADPH and a mixture of competitive and non-competitive with respect to the substrate [151]. The strongest inhibition was seen with chloride ions with a K_i of 5.9 mM. The much-studied *p*-HBH is the prototypal FAH from which most of the known FAH structural and mechanistic information was elucidated.

1.2.3.1 Structural studies of para-hydroxybenzoate hydroxylase

The FAH enzyme *p*-HBH was first purified and studied from three bacterial *Pseudomonas* strains over forty years ago [152-154]. The enzyme was successfully crystallised with substrate in 1975 [155], with the resulting structure published in 1979 (depicted in figure 1.4) [156]. Interestingly this crystal growth was unrepeatable for a further decade until a

study on the influence of protein heterogeneity in 1989 [157]. It was discovered that strongly diffracting *p*-HBH crystals could only be generated from purely dimeric protein possessing a reduced SH group on residue C116. From this a structure emerged with substrate bound to a resolution of 1.9 Å [158]. Since then numerous mutagenic and structural studies of *p*-HBH have elucidated key features of the interactions and dynamics of FAHs.

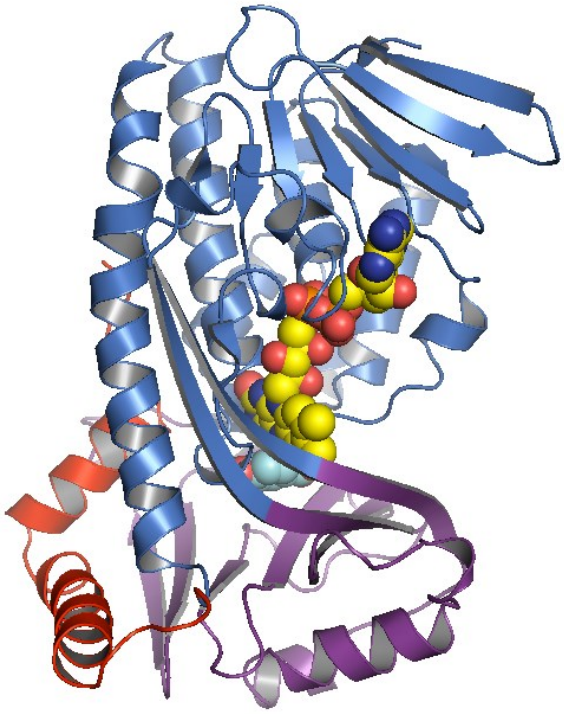


Figure 1.4. Cartoon representation of the structure of *p*-HBH (pdb code 1pbe) with FAD (yellow) and the substrate (cyan) shown in spheres. The protein is split into three domains coloured blue, purple and red.

The major developments that came from studies of *p*-HBH surrounded the discovery of a proposed 'wavin' flavin' conformational change to create the different solvent environments required during catalysis. An alternate flavin conformation was first seen in the structure of Y222F *p*-HBH in complex with substrate [159]. The electron density (figure 1.5) revealed a mixture of the previously seen 'in' and the new, more solvent exposed 'out' conformation for the flavin's isoalloxazine moiety. Moreover, in the structure of wt *p*-HBH complexed with a substrate analogue – 2,4-dihydroxybenzoate (DOB) - only the flavin's 'out' conformation was seen [160]. Here the out conformation of the flavin was stabilised by hydrogen bonds with the added hydroxyl on the substrate

analogue DOB and the side chain of Y222. Only the flavin appeared to move relative to the structure of wt substrate bound *p*-HBH (figure 1.6). The flavin movement was supported by spectroscopic measurements, which showed different spectral changes upon the addition of substrate or DOB, and later by Raman spectroscopy [161]. The flavin movement in solution was again investigated with a modified flavin cofactor, 6-azido-FAD, incorporated into *p*-HBH [162]. The azido group was photoactivated into the highly reactive nitrene group by visible light. When the flavin occupied the 'out' conformation this group could react with water whereas in the 'in' position it reacted and covalently cross-linked with the protein. In the absence of substrate the cross-linked yield was 80 % compared to 66 % with substrate and only 17 % when DOB was present. Further studies of the cross-linked protein showed that with the flavin fixed in the in position, both substrate exchange and reduction by NADPH occurred at a substantially decreased rate. In all these results correlated well to strongly confirm the crystallographically-observed flavin movement and imply its role in substrate entry and NADPH-mediated flavin reduction.

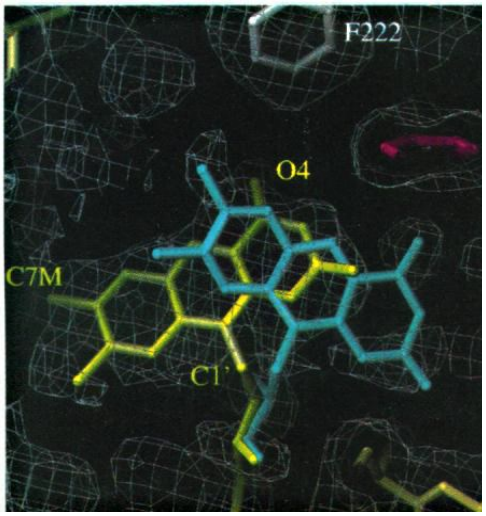


Figure 1.5. Electron densities for the two flavin conformations – *in* (blue) and *out* (yellow) in the Y222F *p*-HBH mutant (pdb code 1dob) as taken from Gatti *et al.*, 1994 [159]. Contouring was at 0.3 σ.

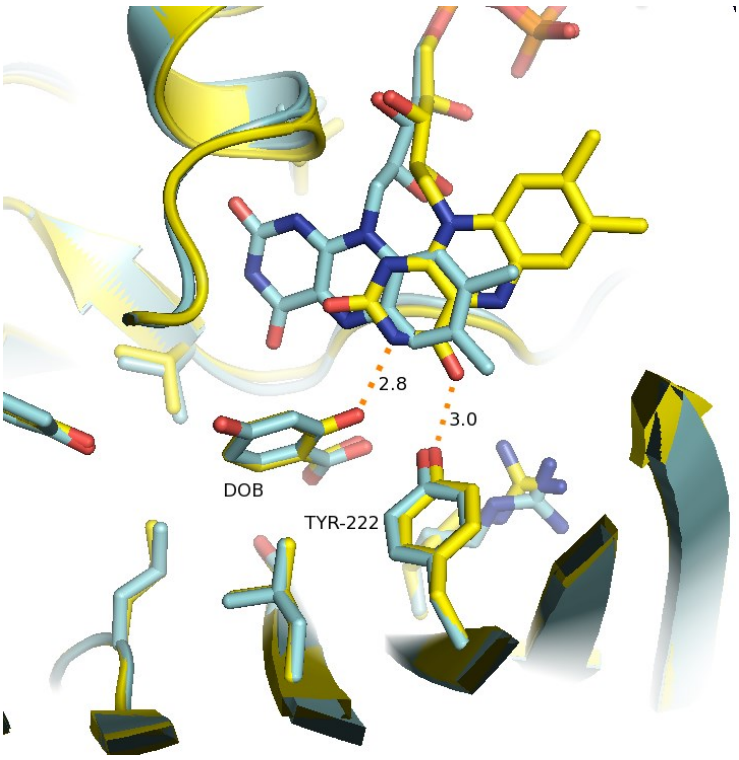


Figure 1.6. Two flavin conformations seen in the structures of substrate-bound (blue) and DOB-bound (yellow) wt *p*-HBH. The substrate analogue stabilises the flavin *out* conformation via the formation of two hydrogen bonds. Created using PyMol Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.

As FAH enzymes use external NADPH to reduce the flavin, there is no well-defined NADPH binding site. Instead NADPH transiently docks at the surface of the enzyme and is immediately released after flavin reduction. As a result, no native structures with NADPH bound have been obtained to date. In 2002, a NADPH-bound pHBH structure (figure 1.7) was obtained using a R220Q mutant form of the enzyme [163]. This substitution favoured the open conformation of the enzyme and dramatically diminished the substrate binding affinity. As a result the first substrate free pHBH structure was also obtained with this mutant. It was postulated that upon substrate binding the flavin swings out and the NADPH bends into a folded conformation to bring the isoalloxazine and nicotinamide moieties into close contact. A lot of the residues interacting with NADPH in the structure (including R44, F161, H162 & R269) had previously been shown to be important in NADPH binding from mutagenesis studies [164-167]. In contrast though, Y38 and R42 had also been implicated in the binding of the 2'-PO_4 of NADPH

[168, 169] whereas they were close but didn't directly interact in the presented structure. This left doubt as to whether the presented mutant enzyme's NADPH binding location and orientation accurately represented that of the native enzyme.

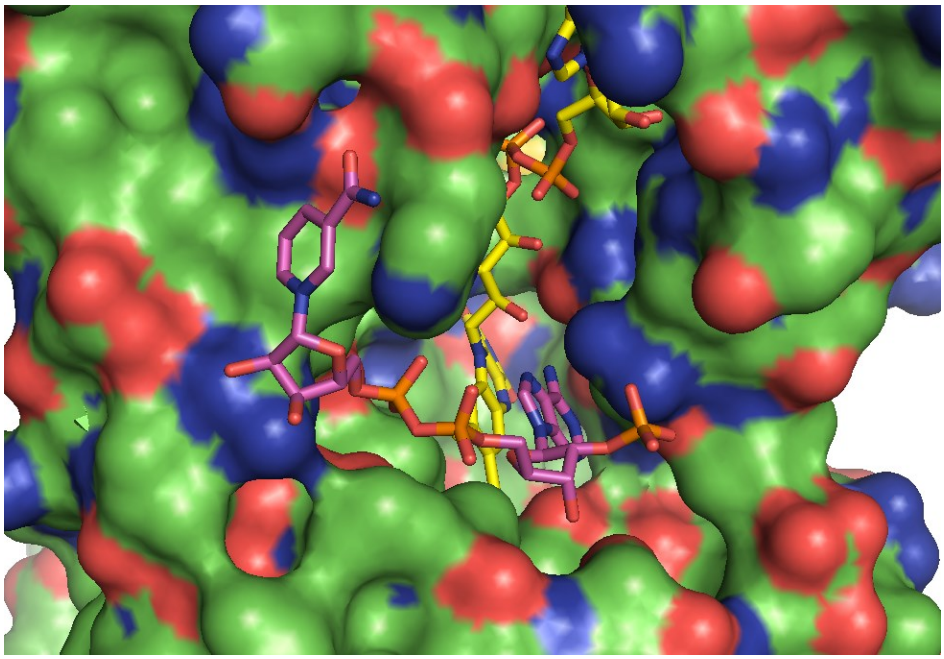
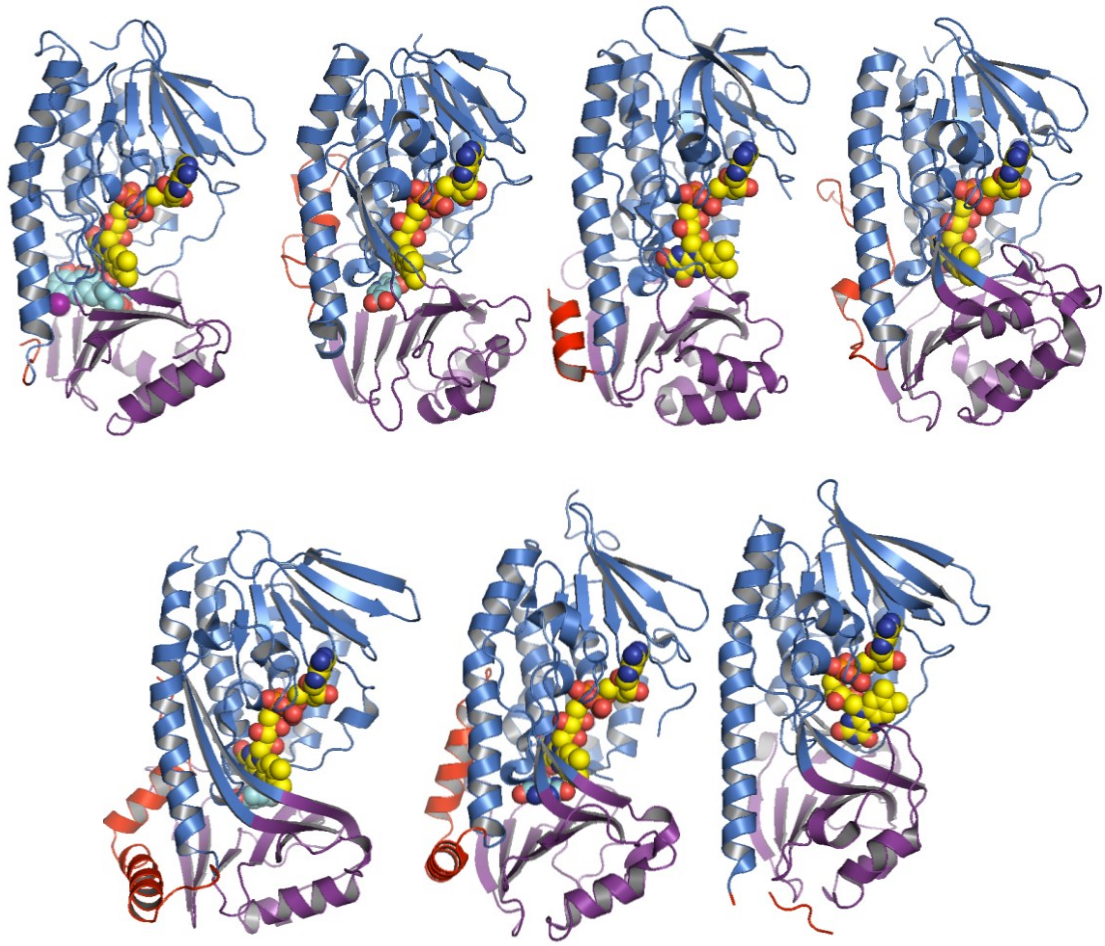


Figure 1.7. Surface view of the R220Q mutant *p*-HBH structure (pdb code 1k0j) with NADPH (magenta sticks) bound in a groove adjacent to the FAD-binding site. Figure created using PyMol.

1.2.3.2 Structural information from other FAHs

In addition to *p*-HBH, numerous FAH crystal structures have recently emerged including those of: phenol hydroxylase (PhH) [170], 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase (MHPCO) [171], RebC [172], , 2,6-dihydroxypyridine 3-hydroxylase (DHPH) [173], aklavinone-11-hydroxylase (AVH) [174], *m*-hydroxybenzoate hydroxylase (*m*-HBH) [175], tetracycline monooxygenase (TetX) [176], two aromatic polyketide hydroxylases (PgaE and CabE - [177]), urate oxidase (HpxO) [178] and two FAD-dependant monooxygenases with unknown substrates (PQSL - [179] and phzS - unpublished). The structures share low sequence identities averaging at 10-20 % but with conservative rmsd values of 2.0-3.5 $Å^2$.

As can be seen in figure 1.8, the structures in fact share very similar overall folds with highly conserved structural elements for the two domains decribed for *p*-HBH. Some FAH enzymes possess additional thioredoxin-like C-terminal domains with no known functional roles. The conserved first two domains comprise up to around 370-400 residues with the C-terminal extensions ranging from a further 20-250 residues. In all cases the flavin cofactor is tightly bound in a superimposable cleft near the surface of the enzymes with the respective substrate bound in a buried pocket beneath the isoalloxazine moeity. The previously mentioned different flavin conformations are represented in different FAH structures but are not dependent simply upon the presence/absence of substrate. In some cases, binding of the substrate immediately appears to cause flavin movement to the out position. Where this does not happen it is possible that the actual docking of NADPH directly causes the flavin movement. In such cases this still must only be allowed to happen in the presence of substrate to prevent uncoupled NADPH oxidation and hydrogen peroxide release.



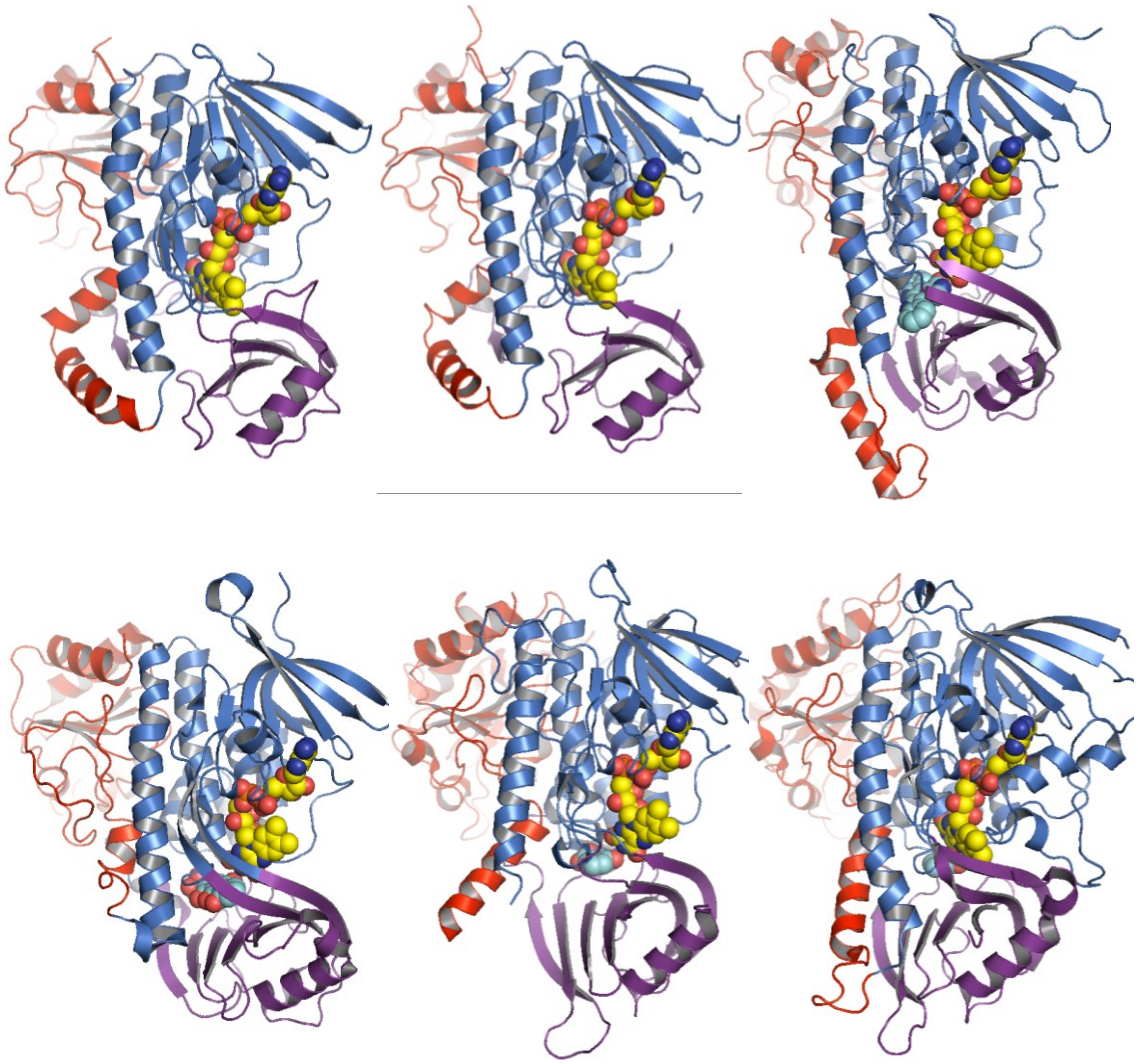


Figure 1.8. Cartoon representation of published FAH crystal structures with FAD (yellow) and the respective substrate (blue) shown in spheres. The structures are arbitrarily split into three domains in blue, purple and red respectively to show the similarity in the fold of the first two domains but the variation in the layout of the third, C-terminal domain. The respective pictures represent: row 1 = TetX (pdb code: 2y6q), MHPCO (3gmc), PQSL (2x3n), DHPH (2vou), row 2 = *p*-HBH (1pbe), HpxO (3rp7), PhzS (3c96), row 3 = CabE (2qa2), PgaE (2qa1), RebC (2r0p) and row 4 = AVH (3ihg), *m*-HBH (2dkh), PhH (1pn0). Created using PyMol.

From these structures and from studies of *p*-HBH, three conserved motifs were discovered for the identification of FAHs [166]. In all FAH enzymes the motifs are present in relatively similar locations in terms of primary sequence and three-dimensional structure. The first motif is conserved among the majority of single-component flavoenzymes, however the second and third motifs are typically only found among the FAH subclass. As can be seen in table 1.2, all three are conserved in the sequence of human KMO.

Motif	Location*	Function	Human KMO sequence
V*hhh***G***s***G***hh***G***hhhs*	Near N-terminus	Binding ADP moiety of FAD - βαβ-fold [180]	**VAVIGGGLVG***S***L***Q***A**
*chhhss***DG***xc***S***xh***R**	~ position 150-200	In contact with riboflavin moiety of FAD	**DLIVGCDGA***Y***STVR**
*hh***L***h***GDAAH***xxx***P***xx***G***x***G***x***N**	~ position 300	Important to FAD and NADPH binding	**CVLLGDAAHAIVPFFGQGMN**

Table 1.2. Description of the three conserved sequence motifs found among FAH enzymes and the relative sequence composition in human KMO. *h* = any hydrophobic residue, *s* = any small residue (G/A/C/S), *c* = any charged residue and *x* = any residue. Residues underlined are strictly conserved. The human KMO sequence residues that are not bold and in italic font do not fit with the generalised motif. *The motif location is relative to the amino acid sequence.

1.3 Kynurenine 3-monooxygenase (KMO)

1.3.1 Overview of KMO

KMO catalyses the hydroxylation of L-kynurenine into 3-hydroxykynurenine whilst consuming external NADPH and molecular oxygen. It has been classified as a member of the FAH family of enzymes due to the presence of the three aforementioned highly conserved motifs and from mechanistic studies [1]. KMO is a key player in the mammalian kynurenine pathway and acts in certain degradation pathways present in some species of bacteria and fungi. All KMO proteins are between 437-513 residues long with the eukaryotic sequences all appearing to contain a C-terminal trans-membrane helix. Whilst no KMO structures had been solved prior to this investigation, the enzyme had been purified and studied from several sources.

1.3.2 Studies of mammalian KMOs

Kynurenine 3-monooxgenase (KMO) was first isolated in a partial purification from rat liver mitochondria in 1957 [181]. The enzyme was shown to convert L-kynurenine (L-Kyn) into 3-hydroxykynurenine (3-HK) in the presence of NADPH and O_2. An experiment with ^{18}O confirmed that the oxygen atom incorporated into the product came from atmospheric oxygen and not from water. A further decade passed until further characterization confirmed that the enzyme was exclusively localised in the outer mitochondrial membrane (OMM) in preparations from rat liver [182] and that it required flavin adenine dinucleotide (FAD) [183]. In 1969, KMO from rat liver mitochondria solubilised and partially purified with ammonium sulphate fractionation was shown to be independent of OMM cytochromes and also able to utilise NADH for catalysis [184]. By 1975, a complete yet complex purification procedure was developed to isolate pure solubilized rat liver KMO [185]. It was concluded that addition of 1 % Triton X-100, 1 mM DTT and 5 mM FAD stabilised the enzyme during purification, that the protein was pure with a rough molecular weight of 200 kDa, and that the coenzyme FAD was easily dissociable. Further advances in chromatography simplified the procedure as it was found that the enzyme successfully bound to NADP-agarose but not to L-kynurenine or AMP coupled columns [186]. Characterization of the enzyme was still proving difficult though due to its membrane association, apparent oligomerization in solution and co-purification with haem-containing cytochromes. As a result assessments of the monomeric molecular weight and flavin content of KMOs had been wildy variable.

Almost two decades passed until real progress was seen in the characterization of mammalian KMO enzymes. At this stage PCR amplification and DNA sequencing were becoming phenomenal tools in developing the understanding of enzymes and their relationships. Firstly in 1997, DNA encoding human liver KMO was isolated from hybridization screening using a probe containing exon 2 of *Drosophila melanogaster* KMO [187]. After sequencing the 2.0 kb gene was predicted to encode for a 486 residue protein with a molecular mass of 55.7 kDa. Transfection of the DNA in human embryonic kidney fibroblast cells resulted in successful expression of active enzyme, verified by tests on crude cell homogenates. A K_m for L-Kyn of 13.0 $\pm$ 3.3 μM was determined and,

as before, the majority of specific KMO activity was seen in the mitochondrial fraction. A similar approach was later used with hybridised human KMO involving transfection and expression with COS-1 cells (expression cell line derived from monkey kidney tissue) and the addition of a cleavable glutathione S-transferase (GST) tag [188]. The GST tag provided greater expression levels of soluble human KMO. In 1998, pig liver KMO was fully purified to obtain the first absorption spectra without contamination by cytochromes (displayed in figure 1.9) [189, 190]. It was finally established in these three studies that the enzyme was a simple flavoprotein with an experimental monomeric molecular weight of around 49-59 kDa and one FAD molecule bound per monomer. The native enzyme was still present as an oligomer with an apparent molecular weight of around 300 kDa based on gel filtration. An interesting finding from the investigation with GST tagged human KMO was that normally NADPH oxidation by KMO in the absence of substrate did not occur except when the product 3-HK was present. This suggested that the product stimulated uncoupled NADPH oxidation with hydrogen peroxide formation, presenting a possible cause for the apparent *in vivo* neurotoxicity associated with elevated 3-HK levels (see section **1.1.4**).

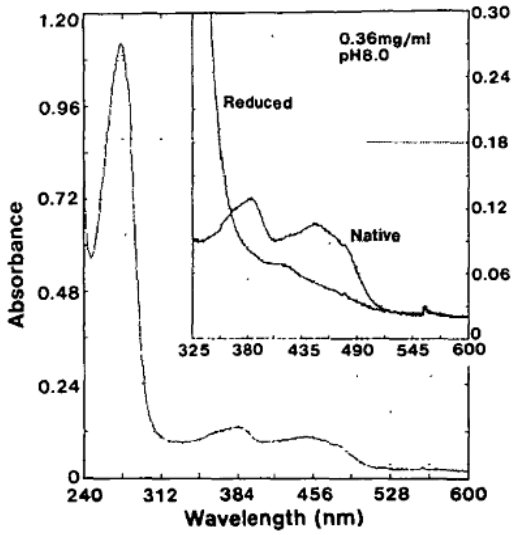


Figure 1.9. Absorbance spectra of fully purified pig liver KMO as taken from Uemura *et al.*, 1998 [189] with native and reduced KMO-FAD peaks shown inset.

In 2010, a series of truncation/mutation studies showed that the C-terminal region of pig liver KMO was involved in targeting the enzyme to the outer mitochondrial membrane and was required for enzyme activity [191]. A series of C-terminal truncates were created with 10, 20, 30 and 50 residues removed respectively, as shown relative to the

amino acid sequence in figure 1.10. Interestingly the intracellular localization of the truncates – determined by fluorescence microscopy – varied depending on the length of the truncation. The wt and ΔC10 truncate were both localised solely in the mitochondria whereas all the longer truncates were localised in the cytosol. This provided strong evidence that the highlighted hydrophobic patch between residues -25 and -47 (relative to the C-terminus) contained a transmembrane helix. The fact that the ΔC20 truncate was also localised in the cytosol suggested that either:

i) the transmembrane region extended past position -20, which was unlikely due to presence of arginines

ii) a mitochondrial membrane targeting signal was present in between positions -20 and -10. It has emerged that tail-anchored OMM proteins generally contain a signalling sequence including a positive patch of residues after the hydrophobic trans-membrane region [192]. This has been shown to be essential to the OMM targeting of Tom20 [193], a splice-isoform of vesicle-associated membrane protein-1 [194] and cytochromes b_5 [195] among others. Therefore the two consecutive arginines at position -18 in pig liver KMO were speculated to be responsible.

To determine whether the two arginines were in fact essential to the OMM targeting of pig liver KMO they were both substituted by serine. Subsequently the protein was localised mainly in the cytosol but with some still visible in the mitochondria. This indicated that the two positive residues were involved, although not solely responsible, in the mitochondrial targeting of the enzyme. Both arginines are conserved in the human enzyme although rat KMO contains a serine in the second position.

Figure 1.10. The relative sequence location of the introduced truncations to pig KMO as taken from Hirai *et al.*, 2010 [191]. Basic residues are labelled by a plus sign and the potential transmembrane region is highlighted in grey.

The truncations also affected the enzyme activity with 100 % expression-corrected activity up to truncate ΔC20 and then 160 % for ΔC30 and just 10 % for ΔC50 (For

summary see figure 1.11). The key information from this was that truncation to completely remove the transmembrane region effectively inactivated the enzyme. Soluble fully active enzyme could, however, be obtained from the removal of just the sequence up to the RR targeting signal as with the ΔC20 truncate. So in summary it was discovered that the C-terminal residues of pig KMO were linked to the enzyme activity and contained an OMM-localised transmembrane helix with a similar arrangement and signalling sequence to tail-anchored OMM proteins. The discoveries were thought to be transferable to human KMO based on sequence conservation of these key features.

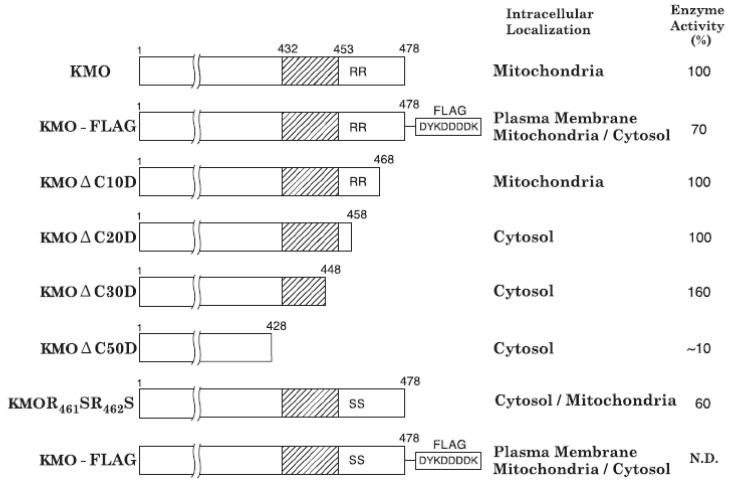


Figure 1.11. Schematic representation, subcellular localisation and enzymatic activity of the pig KMO constructs, taken from Hirai *et al*., 2010 [191]. The predicted transmembrane region is highlighted by hatched boxes and no activity was determined for the FLAG-attached SS form, KMO-FLAG (N.D.).

1.3.3 Studies of KMOs from other eukaryotic sources

Whilst studies of mammalian KMOs were under way, KMO activity was also detected in various insect tissue preparations [196], in *Drosophila melanogaster* [197], *Anopheles stephensi* [198], *Aedes aegypti* [199], the Mediterranean flour moth *Ephestia kuhniella* [200] and the red bread mould *Neurospora crassa* [201]. As with mammalian KMOs these insect and fungal KMOs showed an association with the outer mitochondrial membrane based on specific activity location and electron microscopy. The studies from *Anopheles stephensi* and *Aedes aegypti* involved purified and characterised recombinant

KMO where similar features were again seen with: pH optima of 7.0-7.5, narrow substrate specificity, monomeric molecular weights of around 55 kDa, binding of FAD and preference of NADPH over NADH as a coenzyme. An interesting side-story developing in the study of insect KMOs was the discovery that a mutation in the KMO gene was linked to development of a white-eye phenotype in *Aedes aegypti* [202, 203]. Going back to the work on recombinant KMO from the same species it was demonstrated that this mutant gene contained a 3' deletion translating to the loss of residues 373-427 near the C-terminus of the protein [199]. This mutant form of the enzyme was inactive but it was speculated that this might be because the deletion disrupted a predicted long α-helix. Whatever the reason this was further evidence that the C-terminus of KMOs had some relevance to the enzyme's catalytic activity.

1.3.4 Studies of bacterial KMOs

A small number of bacteria possess limited pathways for tryptophan catabolism that include KMO. Critically, sequence analysis shows that these bacterial forms of KMO do not appear to contain any predicted transmembrane regions. In 2007 a bacterial KMO from *Psuedomonas fluorescens* strain 17400 (*Pf*KMO) was cloned and expressed in *E.coli* with a high yield of soluble protein [2]. *Pf*KMO was purified with great success using a process of: streptomycin sulphate precipitation, ammonium sulphate fractionation and anion exchange chromatography. With the absence of a transmembrane region, the protein was soluble without the need for any detergents and so it could be characterised to a further extent than any previously studied KMO. Despite this, there were still a few minor difficulties that had to be overcome. Firstly no activity was detected when normal expression protocols involving addition of IPTG and growth at 37 ^{o}C were used but this was resolved by slower cell growth at 22 ^{o}C for 16 hours without induction with IPTG. It was thought that either the rate of folding or instability of the enzyme was restricting the production of active protein, particularly during overexpression with the inducing agent. The second, possibly related, problem involved the rapid loss of the purified enzyme's activity over time in the absence of reducing agent. Over a period of 30 hrs at 4 ^{o}C, up to 90 % of the enzyme activity was lost with the majority recovered by addition of

dithiothreitol (DTT), a strong reducing agent. *Pf*KMO contains six cysteine residues and it was shown that this loss of activity correlated with the oxidation of two of these residues based on a DTNB-based assay for free thiol groups (Figure 1.12). SDS-PAGE confirmed the reversible dimerization of *Pf*KMO over time depending on the presence of reducing agent. Therefore during purification and storage of the enzyme a reducing agent was required to prevent aggregation via disulphide bridge formation.

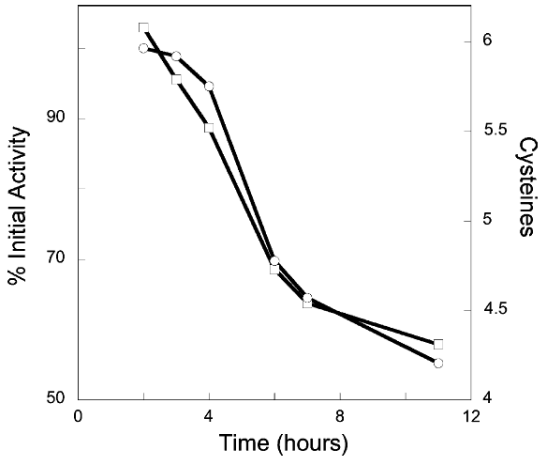


Figure 1.12. Graph taken from Crozier & Moran 2007 [2] showing loss of *Pf*KMO activity (○) over time relative to the number of free thiols/monomer (□).

Parameter	Value at 5 °C	Value at 25 °C
Turnover number (s^{-1})	1.27 ± 0.002	4.9 ± 0.1
$K_{L\text{-}Kyn}$ (μM)	11.8 ± 3.3	53 ± 7
$V/K_{L\text{-}Kyn}$ ($\mu M^{-1}s^{-1}$)	0.10 ± 0.02	0.08 ± 0.01
K_{NADPH} (μM)	8.5 ± 0.7	91 ± 6
V/K_{NADPH} ($\mu M^{-1}s^{-1}$)	0.15 ± 0.01	0.053 ± 0.004
K_{O2} (μM)	34.2 ± 4.0	71 ± 5
V/K_{O2} ($\mu M^{-1}s^{-1}$)	0.029 ± 0.003	0.008 ± 0.003

Table 1.3. Michaelis-Menten kinetic constants determined for *Pf*KMO at pH 7.5, adapted from Crozier & Moran 2007 [2].

The Michaelis-Menten kinetic parameters for *Pf*KMO were calculated and displayed in table 1.3 and it was shown that the enzyme had maximal activity at pH 8.0-8.5 and maximal stability at pH 6.0–7.0. The KMO-FAD-L-Kyn complex had a dissociation constant of 17.4 ± 1.4 μM and of 133 ± 6 μM at 5 and 25 °C respectively. The first real detailed mechanistic picture of KMO was then revealed in a second study on the bacterial enzyme [1] that firmly confirmed the association of KMO with the FAH enzyme family. As seen with other FAHs, NADPH oxidation in the absence of substrate was largely prohibited with subsequent flavin reduction occurring over long time intervals (around 5000 s). The process was stimulated 10^3-fold in the presence of L-Kyn. The reaction

catalysed by *Pf*KMO was shown to consist of several observable steps as summarised in the proposed mechanism in figure 1.13. For each transition the rate constant was measured using stopped-flow spectrophotometry and it was found that product release was the rate-limiting step. Building on previous observations on the sensitivity of KMOs to the presence of anions [188, 199, 200], it was shown that *Pf*KMO activity was sensitive to low concentrations of anions. Absorbance spectra confirmed that binding of anions directly influenced the flavin, with spectral changes similar to those seen in the presence of the product 3-HK. It was suggested that the occupancy of the flavin's multiple positions was sensitive to local electrostatic changes as caused by the binding of anions in the proximity of the active site.

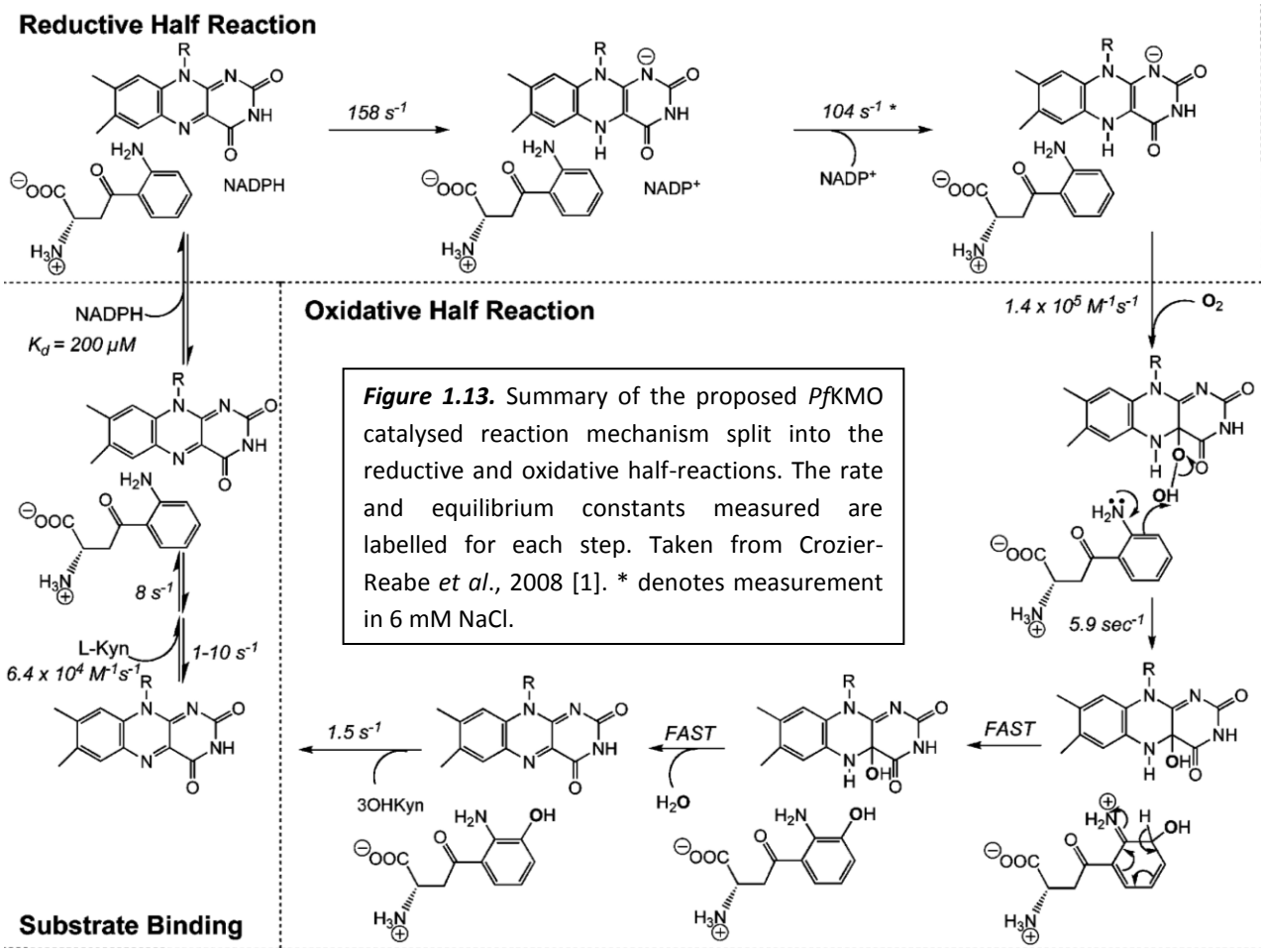


Figure 1.13. Summary of the proposed *Pf*KMO catalysed reaction mechanism split into the reductive and oxidative half-reactions. The rate and equilibrium constants measured are labelled for each step. Taken from Crozier-Reabe *et al*., 2008 [1]. * denotes measurement in 6 mM NaCl.

1.4 Inhibition of KMO

1.4.1 Discovery of KMO inhibitors

Partially purified yeast KMO was inhibited by α–keto acid analogues of branched chain amino acids, pyruvate and α-ketoglutarate [204] with K_i values in the low millimolar range. The keto acid analogues were found to show non-competitive inhibition with respect to both L-Kyn and NADPH whereas pyruvate and α-ketoglutarate were mixed-type inhibitors. Unfortunately such levels of inhibition were not reproduced when the compounds were tested on partially purified rat liver KMO [205]. The first potent KMO inhibitors synthesised: *o*-, *m*- & *p*-nitrobenzoylalanine (NBA), were analogues of the substrate L-kynurenine [54, 206]. The compound *m*-NBA in particular seemed to be a very potent competitive inhibitor, with an IC_{50} of 0.8 μM against KMO from resuspended, homogenised rat organs. Going further, administration of *m*-NBA to rats showed anticonvulsant activity against electroshock-induced seizures and increased the brain, blood and liver concentrations of both KYN and KYNA [54, 206]. Another group explored more manipulations of substrate analogues to develop a stronger inhibitor in the shape of 3,4-dichlorobenzoylalanine (3,4-CBA or FCE 28833). This compound showed similar, yet more potent benefits in rats compared to *m*-NBA and had an IC_{50} of 0.2 μM [207]. Further variations of substrate analogues with similar levels of inhibition were characterised by Giordani *et al.,* 1998 [208] and Drysdale *et al*., 2000 [209]. Again these compounds showed promising therapeutic potential by preventing IFN-γ induced synthesis of QUIN in primary cultures of human monocyte-derived macrophages [209]. Recently however, the therapeutic potential of substrate analogues was dismissed when it was shown that they actually cause uncoupled flavin reduction by NADPH which leads to the generation of cytotoxic hydrogen peroxide [1]. Thus, such inhibitors have been re-classified as non-substrate effector molecules.

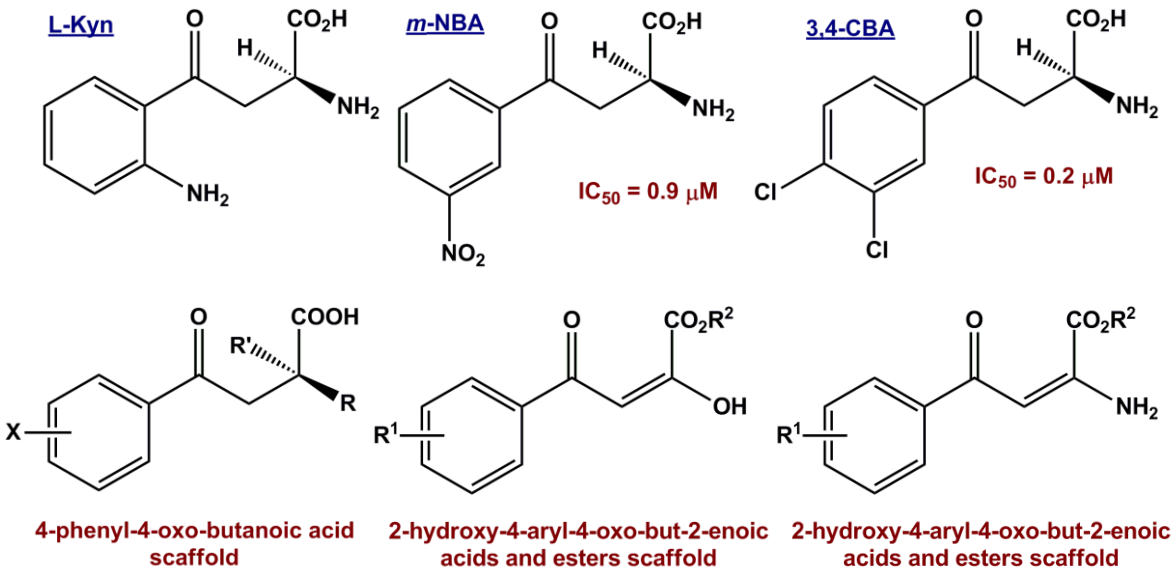


Figure 1.14. Summary of the molecular structures of the range of substrate analogue KMO inhibitors synthesised and tested compared to the substrate L-Kyn. The 4-phenyl-4-oxo-butanoic acid series of compounds were synthesised by Giordani *et al.* 1998 [208], whilst both the 4-aryl-4-oxo-but-2-enoic series were synthesised by Drysdale *et al*. 2000 [209].

Shortly after the discovery of substrate analogue inhibitors, the search for more specific and more potent KMO inhibitors moved on to novel and more complex organic compounds. In 1997, screening of a chemical library based on bioisosters of the L-Kyn carboxyl group rapidly identified a sulphonamide as a good starting scaffold [4]. Subsequently a series of *N*-(4-phenylthiazol-2-yl)benzenesulfonamides (figure 1.15) were synthesised and shown to be promising high-affinity competitive inhibitors of KMO. Of those tested *in vitro* on KMO from rat kidney mitochondria, the strongest compounds had low IC_{50} values of 20-40 nM, which were up to 10 times lower than the substrate analogues. A selection of the compounds was tested *ex vivo* on different homogenised organs after oral administration in both rats and gerbils. In all cases strong inhibition was seen and with significantly greater potency than 3,4-CBA when directly compared. The ED_{50} values determined for the two strongest inhibitors in the different organ tissues ranged between 0.5-5.5 μmol/kg compared to 22-222 μmol/kg for 3,4-CBA. Finally, administration of one of the compounds (Ro 61-8048) to rats was also shown to increase the extracellular hippocampal fluid concentration of KYNA by 7.5 times at a dose of 100 μmol/kg. Subsequently Ro 61-8048, which contained a 3-NO_2 R' group and a 3,4-OMe R''

group, was used in several studies to further investigate the therapeutic potential of inhibiting KMO.

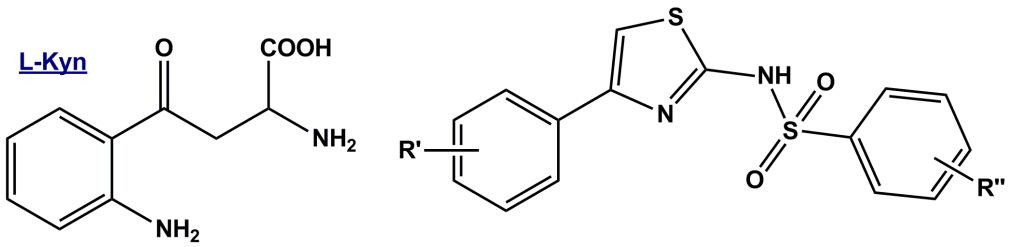


Figure 1.15. Comparison of the structure of the series of novel KMO inhibitors synthesised by Rover *et al*. 1997 [3] with that of the substrate L-Kyn.

Since this discovery an alternate unrelated group of specific KMO inhibitors has emerged. In 2012 a natural product lanthellamide A (figure 1.16) – isolated from the Australian marine sponge *lanthella quadrangulata* – showed relatively potent inhibition of KMO [210]. The IC_{50} from *in vitro* assays with rat KMO was determined to be 1.5 μM for the inhibition of 3-HK production. The inhibition was specific for KMO as only weak inhibition of KYNase was detected at concentrations of up to 1 mM lanthellamide A. Moreover co-administration of L-Kyn and the inhibitor in the hippocampus of anesthetised rats caused up to a peak three-fold increase in extracellular KYNA levels over L-Kyn only controls. Due to its recent discovery, little is known about the wider therapeutic potential of lanthellamide A and any possible related compounds.

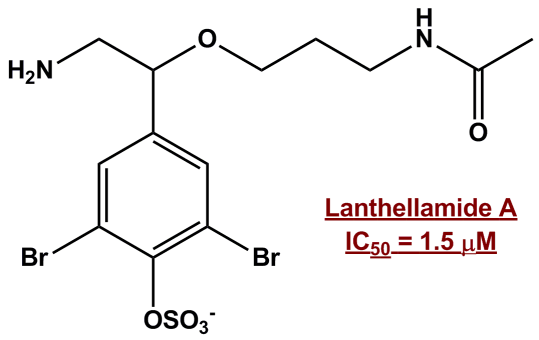


Figure 1.16. Structure of the marine sponge natural product that was found to inhibit KMO by Feng *et al*. 2012 [210].

1.4.2 Therapeutic effect of KMO inhibition in models of disease

One of the questions surrounding the therapeutic potential of KMO inhibitors was whether QUIN synthesis could continue regardless via KMO-independent pathways? In mice, large doses of *m*-NBA & Ro 61-8048 increased blood and brain KYNA whereas they only decreased QUIN levels in immune-stimulated animals [211]. It was suggested that during an immune response raised QUIN levels were achieved as a result of upregulation of the KMO branch, whereas anthranilic acid hydroxylation was normally sufficient to generate QUIN under normal conditions. These data were encouraging: minimising the effect of KMO inhibitors in healthy cells whilst significantly reducing pathophysiological levels of QUIN. Research involving induced rat and gerbil brains showed that the two KMO inhibitors protected against induced brain lesions [61]. The percentage of lesioned neurones reduced from around 90 % to 10% in the presence of the inhibitors and again a significant increase in the concentration of KYNA was noted. Ro 61-8048 has also been shown to offer partial protection against induced convulsions both extending the time before first convulsion and reducing the total number of convulsions in rats [212]. Aside from the main focus of controlling the ratio of QUIN:KYNA, there has also been cases where inhibition of KMO has been targeted for direct reduction of its reaction product 3-HK. It was thought that an excessive accumulation of 3-HK in the eye lens was linked to cataract formation [131]. In mitochondrial extracts from the rabbit iris and lens, Ro 61-8048 inhibited KMO activity and subsequently 3-HK production with IC_{50} values of 0.1 & 1.0 μM respectively (compared to 1.3 & 7.0 μM for *m*-NBA) [131].

There have been several examples of KMO inhibition resulting in observed neuroprotective qualities and raised KYNA levels [49, 60, 61, 144, 213-216]. There are mixed reports however as to whether the level of increase in [KYNA] caused by KMO inhibition is sufficient to reduce activity of EAA receptors and thus to suppress excitotoxicity. One group found evidence against the notion by studying anaesthetised rats where NMDAr depolarisation was measured after direct administration of QUIN, KYNA and the KMO inhibitor Ro 61-8048 [214, 217] . Whilst KMO inhibition with 100 mg/kg Ro 61-8048 increased the basal [KYNA] 11 times to around 30 nM, there was no evidence of an associated reduction in NMDAr response. Another study found that addition of 30-100 μM *m*-NBA or 1-10 μM Ro 61-8048 to hippocampal slice cultures

reduced post-ischemic neuronal death after 30 mins of exposure to oxygen and glucose deprivation [60]. KYNA levels increased but to only 50 nM, considered to be below the concentration needed to significantly affect EAA receptors. Due to the observed protective effect being counteracted by addition of 3-HK or QUIN, the authors proposed that KMO inhibition offered neuroprotection in this case by reducing the local synthesis of these two metabolites. A couple of years later it was found that local application of 30-100 nM KYNA to rats was enough to significantly decrease the concentration of GLU in the caudate and cortical dialysates but had no effect in the hippocampus [218]. Further, administration of 4-40mg/kg Ro 61-8048 increased the concentration of KYNA in dialysate collected from cortex, caudate and hippocampus but only reduced the concentration of GLU in the basal ganglia. A complex picture was emerging whereby raised KYNA levels to the extent seen with inhibition of KMO seem to be able to suppress NMDAr activity in some selective regions of the CNS. It was still not clear whether the definitive neuroprotective properties associated with KMO inhibition were directly linked to increased KYNA, reduced QUIN and/ or reduced 3-HK levels.

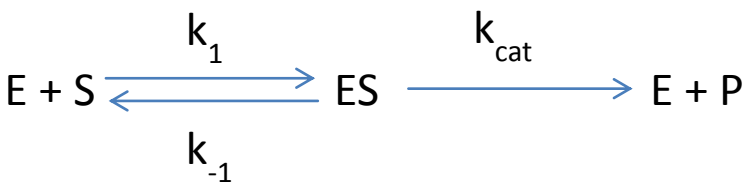
1.5 Characterising enzyme-catalysed reactions

1.5.1 Michaelis-Menten kinetics

The Michaelis-Menten equation (equation 1.1) describes the relationship between substrate concentration and the rate of an enzymatic reaction. Michaelis-Menten kinetics applies to all enzyme-catalysed reactions represented by scheme 1.1 and does not apply for reactions where there is cooperative or allosteric binding of the substrate or product.

$$V = \frac{V_{max}[S]}{K_m + [S]}$$

Equation 1.1. Michaelis-Menten relationship. V = initial reaction rate, [S] = concentration of substrate.



Scheme 1.1. Generalised reaction scheme defined by Michaelis-Menten kinetics. E = Enzyme, S = Substrate, P = product, k = rate constants.

The V_{max} is the maximum rate of reaction obtained with infinite substrate and is variable with respect to the enzyme concentration. The Michaelis-Menten constant k_{cat} (turnover number) represents the V_{max} per enzyme active site and reveals a particular enzyme's maximal rate of turnover per unit time. Another important constant in the characterisation of an enzyme is the K_m value for a respective substrate. The K_m is the substrate concentration required to achieve 50% of k_{cat}, which in reality reveals how tightly the substrate is bound in the active site. The efficiency of an enzyme is estimated by calculating the second order rate constant k_{cat}/K_m. The Michaelis-Menten constants for an enzyme can be determined experimentally from a plot of the initial reaction rate (v) against substrate concentration, as shown in figure 1.17.

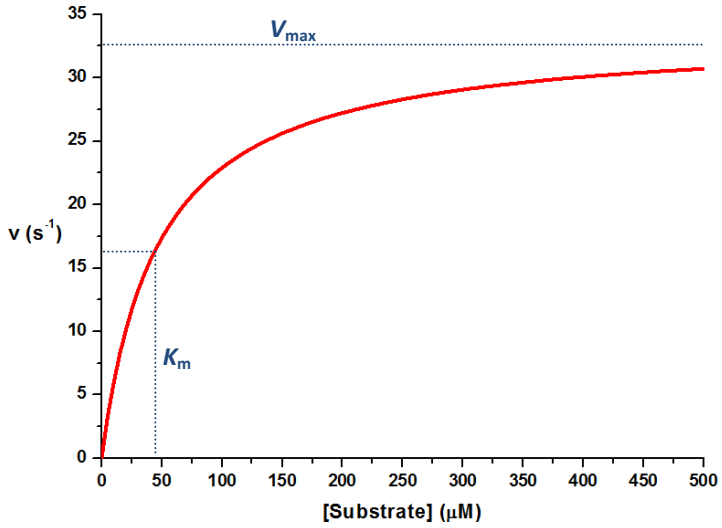


Figure 1.17. Example of Michaelis-Menten plot for the determination of V_{max} (therefore k_{cat}) and K_m.

1.5.2 Enzyme inhibition

The activity of an enzyme may be inhibited by a molecule in one of the following ways:

i) Competitive inhibition – whereby the inhibitor reversibly competes with the substrate for binding in the active site. With increasing [inhibitor] V_{max} is unchanged but K_m is increased as only substrate binding is affected.

ii) Uncompetitive inhibition – whereby the inhibitor binds to the enzyme-substrate complex to prevent catalysis. K_m and V_{max} are both lowered with increasing inhibitor.

iii) Mixed inhibition – whereby the inhibitor binds both to the free and to the substrate-bound enzyme with different binding affinities. Inhibitor binding affects binding of the substrate. The inhibitor binds in a site other than the substrate binding site either within the active site or at an alternate location (allosteric site). Both K_m and V_{max} are increased in the presence of inhibitor.

iv) Non-competitive inhibition – a form of mixed inhibition where the inhibitor binds only to the enzyme-substrate complex. With increasing inhibitor concentration K_m is unchanged, but V_{max} is reduced.

v) Irreversible inhibition – whereby the inhibitor binding is irreversible, usually due to the formation of a covalent adduct with the enzyme. Binding can be competitive or uncompetitive.

Each type of inhibition can be distinguished experimentally due to their different effects on the apparent V_{max} and K_m. The Lineweaver-Burke plot of 1/v against 1/[S] can be used to distinguish between different types of inhibition as illustrated in figure 1.18. The data fit to a straight line defined by equation 1.2, with the y intercept providing $1/V_{max}$ and the x intercept giving $-1/K_m$.

$$\frac{1}{v} = \frac{K_m}{V_{max}} \frac{1}{[S]} + \frac{1}{V_{max}}$$

Equation 1.2. Equation for the Lineweaver-Burke plot.

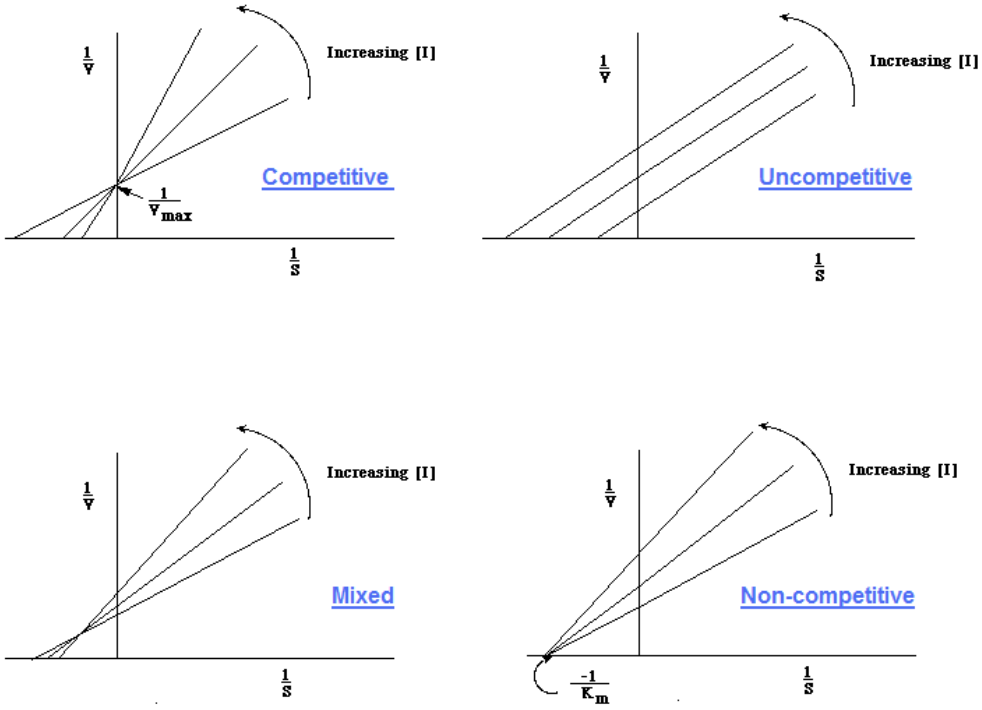


Figure 1.18. Lineweaver-Burke plots in the presence of increasing concentrations of inhibitor for different types of inhibition.

The Lineweaver-Burke plot is effectively a double-reciprocal plot of kinetic data, therefore it significantly amplifies the influence of errors within the data. More accurate plots to determine the K_i and identify the type of inhibition are plots of 1/v and of [substrate]/v *versus* [inhibitor] [219]. A summary of the characteristic plots observed with different types of inhibition is shown in figure 1.19.

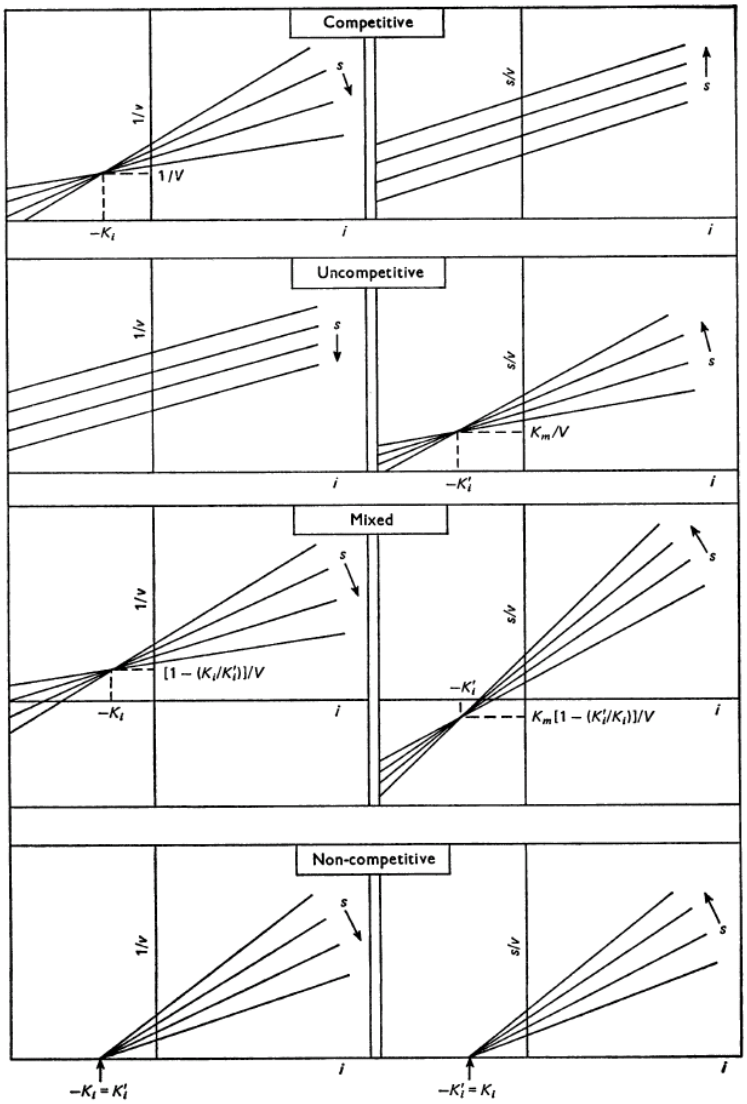


Figure 1.19. On the left plots of 1/v against [inhibitor] and on the right plots of [substrate]/v against [inhibitor] for different types of inhibition. Taken from Cornish-Bowden, 1974 [219].

1.6 X-ray crystallography of proteins

1.6.1 X-ray diffraction

There are two major techniques for determining the three dimensional structure of a protein – X-ray crystallography and NMR-spectroscopy. NMR-spectroscopy has the advantage of allowing the characterisation of a protein in solution, however the technique is typically limited to polypeptides of <30 kDa. As a result X-ray crystallography has remained a powerful method for observing the 3-D structure of proteins regardless

of their size. The technique utilises the phenomenon of diffraction. Diffraction occurs when the path of a wave is regularly distorted upon encountering an object of similar size to the respective incident wavelength. The size of covalent bonds in proteins is 1.2-1.5 Å and matches to wavelengths in the range of X-ray radiation (0.1-100 Å).

Bragg diffraction was discovered when patterns of diffracted spots (reflections) were observed upon passing a beam of X-rays through a crystalline solid. Bragg diffraction was quantified using an understanding of trigonometry, which lead to the derivation of Bragg's Law (equation 1.3). This showed that the positions of reflections were related to the interplanar spacing between each repeating unit cell in the crystal. Reflections are only recorded from constructive (in phase) diffraction from identical waves generated from parallel sets of symmetrical planes in the crystal lattice (see figure 1.20). The equation showed that this constructive diffraction at each interplanar distance (*d*) only occurs from waves that impinge at angles (θ) that give integer values for *n*. Therefore as a crystal is rotated different sets of reflections are recorded so that information about the whole 3-D structure within the repeating crystal lattice is collected. An important consequence of this process is that the formation of ordered, homogenous protein crystals is essential to achieving strong diffraction for complete structure solution. The mechanism of data collection and analysis will be addressed in the following sections.

$$n\lambda = 2d\sin\theta$$

***Equation 1.3*. Bragg's Law**
n = an integer, λ = X-ray wavelength, *d* = interplanar spacing, θ = incident angle of wave

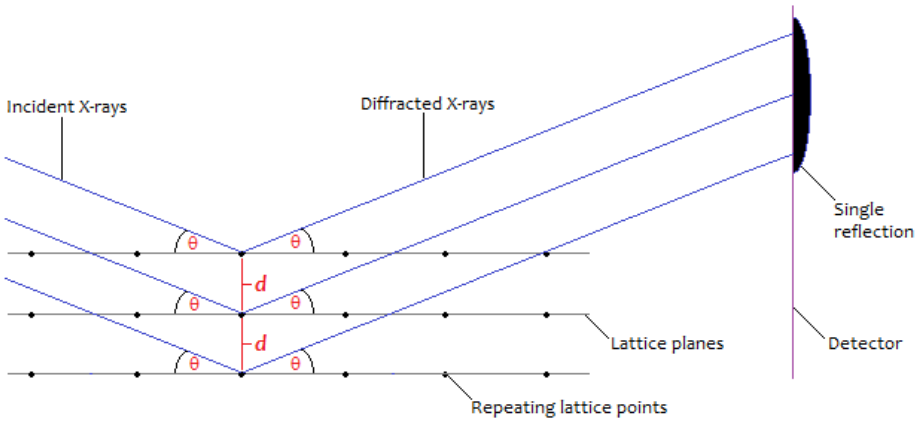


***Figure* 1.20.** Diffraction of X-rays by parallel sets of repeating planes in a crystal lattice results in Bragg diffraction and collection of a reflection.

It is important to remember that the protein structure determined is still just a model; a static representation of the average position of each atom from $>10^{13}$ protein molecules within a crystal. The structural model generated may be missing some of the amino acid residues of a polypeptide sequence, typically the terminal residues which are usually the most disordered. It is also important to ensure that the modelled structure of the crystalline protein is a true representation of the aqueous/functional structure of the protein. There are two main pieces of evidence for this. In the majority of cases, the structure of a protein in solution from NMR spectroscopy is very similar to the structure obtained from X-ray crystallography. Also many crystallised enzymes are still catalytically active, albeit at a reduced rate of turnover due to reduced accessibility of the active site or limited conformational flexibility.

1.6.2 Growth of protein crystals

Protein crystallisation is a form of ordered aggregation at super-saturating concentrations of the protein and may be facilitated by precipitants such as salts, solvents and/or PEGs. Initially the solution contains a precipitant concentration that is lower than that required to precipitate the protein. The controlled evaporation of water from this solution slowly raises the concentration of both the protein and precipitant increasing the likelihood of precipitation. The diffusion of water is controlled by vapour diffusion techniques. In hanging drop vapour diffusion (h.d.v.d), an aliquot of protein solution is mixed with an aliquot of precipitant solution on a cover slip. The cover slip is then placed and sealed over a well containing just the precipitant solution. In sitting drop vapour diffusion (s.d.v.d) the drop is set up in a raised well above the precipitant solution and the chamber is sealed by sticky tape. In both cases the rate of diffusion and the initial concentration of precipitant and protein greatly influence the potential to form good quality crystals. These methods provide a simple way of screening various crystallisation conditions whilst also allowing some control of the rate of diffusion from the protein drop.

There are two aspects to crystal formation - nucleation and growth. Crystal nucleation is the event of protein molecules associating together in an ordered arrangement. Crystal

growth then proceeds as new protein molecules are added to the site of nucleation. Crucially, as protein molecules start coming together, the protein concentration starts falling and so does the occurrence of further nucleation. Figure 1.21 describes the region where nucleation occurs and shows the metastable zone, where crystal growth, but not nucleation, occurs. So, to summarise, the conditions begin to the left of the solubility curve, then the concentrations rise until the nucleation zone is reached and this is followed by a fall in protein concentration to enter the metastable zone. Protein crystals are stable for long periods in the metastable zone.

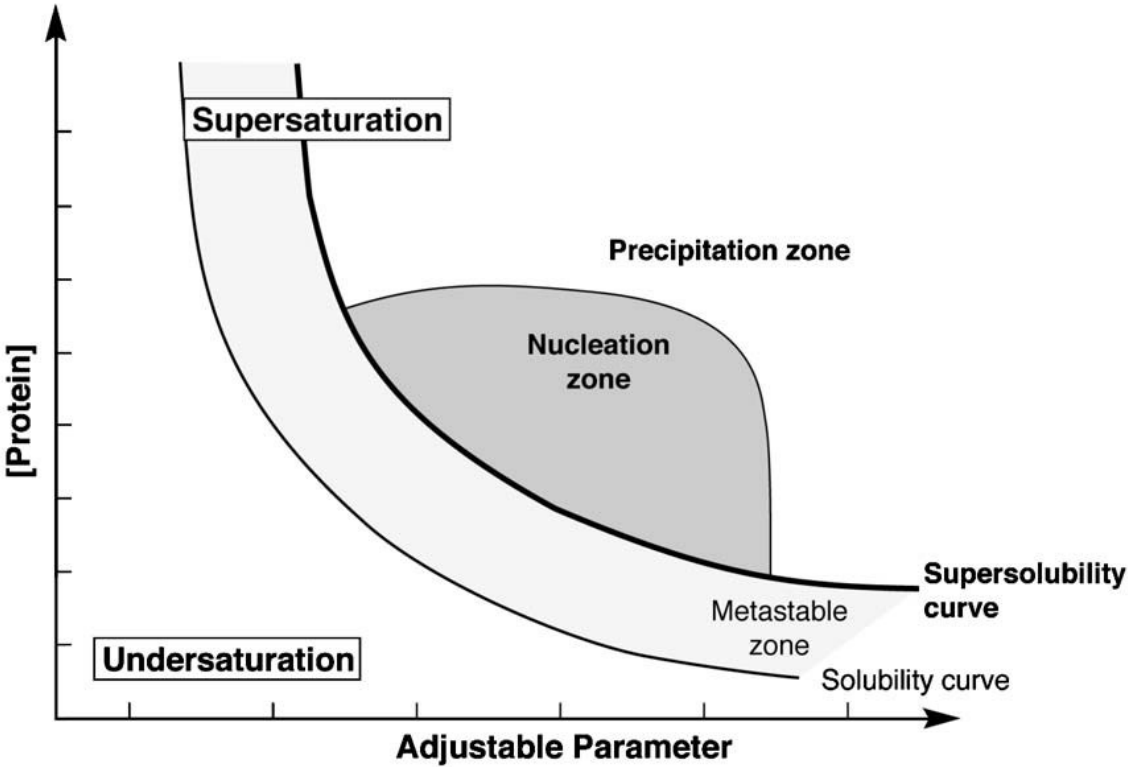


Figure 1.21. Graphical representation of the effect of varying [Protein] or an adjustable parameter (such as pH or [precipitant]) on protein solubility. Taken from N.E. Chayen, 2005 [220].

The nucleation zone depicted figure 1.21 is deceivingly large. In reality, as supersaturation is reached, uncontrolled aggregation/precipitation of the protein molecules occurs more often than not. It is important that this transition towards supersaturation occurs slowly to give the protein the highest chance to associate in a controlled and ordered way. For different proteins, different pH, temperature,

precipitants and concentrations (of protein and the precipitants) may be required to allow the formation of crystals. The purity and homogeneity of a protein solution is directly linked to the likelihood of forming good quality crystals. Not all proteins will easily form diffracting crystals and the ability to do so is largely dependent on the sequence and structural composition of the protein. For example, membrane proteins can be notoriously difficult to crystallise because of their general tendency to aggregate quickly in solution. This is mostly due to large hydrophobicity of trans-membrane domains, which have evolved to associate with the hydrophobic fatty acid chains found on the interior of cell lipid membranes.

Protein crystals are typically 50 % water with the surface of the protein coated by water molecules. Therefore the actual forces holding a crystal together are mostly hydrogen bonds between the protein amino acid residues and solvent molecules. This makes protein crystals much more fragile than crystals of small molecules. The large proportion of solvent creates an additional problem when freezing protein crystals. Protein crystals are typically flash-frozen in liquid nitrogen prior to being exposed to an X-ray source to protect the crystals from damage by the X-rays themselves. During freezing, the solvent molecules can form small ice crystals which disrupt the ordered lattice of the crystal. A clue to this can be seen by the appearance of solid rings on the X-ray diffraction images, however this can be simply down to ice on the surface of the crystal. Therefore protein crystals are typically soaked in mother liquor supplemented by a cryoprotectant such as EG, glycerol or MPD. The cryoprotectant displaces some of the water molecules to buffer against ice formation.

1.6.3 Techniques for improving crystal quality

Protein crystallisation is frequently the bottleneck in the structure solution process. Some proteins are unsuitable for crystallization either through not forming crystals or by forming crystals with only poor diffraction properties. Several strategies have emerged to counteract this to allow the solution of previously unattainable protein structures. These strategies include post-crystallisation treatments such as annealing, chemical cross-linking, dehydration and seeding as well as pre-crystallization engineering of the

protein itself such as the removal of flexible regions or mutagenesis of surface residues. Often the required strategy depends on the problem itself, however in most cases this is hard to elucidate and so all approaches should be considered.

i) Seeding

The most common post-crystal treatments initially explored involve types of seeding. Seeding involves the introduction of a nucleating agent into the crystallization mix to provide a seed to facilitate crystal growth. The seed itself can be heterogenous or homogenous. The most basic approach, which is applicable to situations where crystals have not been obtained, is to stroke across the crystallization drops with a hair or whisker during tray set up (streak seeding). It is thought that streak seeding leaves various nucleants in the drop from which crystals can grow where they otherwise would not have formed. Any hair/whisker can be tried – it was presented that some researchers routinely use their own human hair (unpublished, CCP4 northern protein structural workshop, Carlisle, Aug 2011). At the same conference, Dr Allan D'Arcy (Novartis) described that any solid material can be tried for seeding, from amorphous precipitates to microcrystals. Further, when crystallization conditions cannot be found, such crystalline material is added to the crystallization drops in sparse-matrix screening. Here ideally the seed comes from a previous drop with the protein of interest but a heterogeneous crystalline material from a different protein source can also be tried.

In the event that crystals can be obtained but of poor quality, homogenous seeding offers a powerful and reasonably successful tool. This involves the introduction of either whole crystals (macroseeding) or broken up crystal fragments (microseeding) into the crystallization drop to effectively grow second generation crystals. Without the need to nucleate, the original conditions can be refined to favour slow crystal growth from just the added homogenous seeds. Typically the precipitant concentration is lowered slightly to avoid new nucleation events. In this process the right seed concentration is essential as it determines the amount of crystallization sites in the drop. Typically serial dilutions of the seed crystals/crystal fragments are made and tested and can be stored as aliquots at -20 to -80 ^{o}C for future use. To create a microseed solution, a drop containing crystals of any size is selected and 10 μl of stabilizing solution is added. The crystal/s are

physically broken up either by pipetting up and down or by using a specialised mixing tube. The seed solution is then microcentrifuged at full speed to remove large crystal fragments and only leave the small stable crystal cores which act as effective microseeds. Subsequently serial ten-fold dilutions are made using the stabilising solution. The stabilising solution used can be the original mother liquor itself, usually with the precipitant concentration slightly raised. These techniques are also regularly used to obtain crystals of the SeMet-labelled form of a protein using a heterogenous seed from crystals of the respective native protein.

ii) Post-crystal annealing and dehydration

On some occasions poor crystal diffraction occurs as a result of a high lattice solvent content. In some cases two techniques have been found to counteract this and improve the crystal's diffraction power: annealing and dehydration. Annealing involves the exposure of the crystal to brief bursts of extreme temperatures. One way of doing this is by 'annealing on the loop' where the cryostream is removed from a mounted crystal with one or multiple exposures lasting 1-5s. Also a more subtle approach is where frozen crystals are soaked back into the cryoprotectant solution and then re-frozen. Either way annealing is thought to improve the crystal lattice packing by forming a more energetically favourable and ordered state. Successful cases of annealing are highlighted in the review of post-crystallisation treatments by Heras & Martin [221], which includes the improvement of the resolution achieved from crystals of DmpFG [222] and arsenate reductase [223] from around 6.0 Å to 2.1 and 2.3 Å respectively.

Crystal dehydration is a technique again aimed at improving crystal lattice packing but as a direct result of reducing the crystal solvent content. Crystals are exposed to a higher precipitant concentration for a length of time before being flash-frozen and tested for diffraction. The exposure can consist of soaking the crystal in a solution of mother liquor supplemented with extra precipitant for between 5s to several days. A more gentle dehydration involves simply replacing the reservoir solution with one with a higher precipitant concentration and the drop is then left to re-equilibrate over several hours to days. In both cases a series can be set up with increasing precipitant concentration to discover the optimum point before the crystals are damaged. Success stories of

dehydration are again reviewed by Heras & Martin [221] and include DsbG where the resolution improved from around 10.0 Å to 2.0 Å [224].

iii) Post-crystal cross-linking

Some crystals are particularly fragile and can suffer irreversible damage during handling including upon ligand soaking, cryoprotection and/or cooling. A technique that has been shown to improve crystal strength is the chemical cross-linking of lysine side-chains in protein crystals using glutaraldehyde [225, 226]. The technique is not widely used due to a tendency to reduce the diffraction quality of crystals. This can be due to the formation of varying chain lengths in the cross-links [227]. It was recently reported that an improvement in the success of chemical cross-linking was achieved by gently introducing the gluteraldehyde into the crystal drop using vapour diffusion [228].

iv) Protein engineering to improve crystal contacts

Unlike with small molecule crystallography, the crystallisation of protein molecules is inherently challenging due to their long chain lengths and high flexibility. It has been estimated that fewer than one in four single-domain prokaryotic proteins will yield crystals suitable for X-ray analysis using commercial screens [229, 230]. Protein engineering has provided a powerful tool for the successful crystallisation of numerous proteins which were otherwise resistant to the formation of a regular crystal lattice. The aim is usually to modify/remove surface features that inhibit the formation of stable, homogenous crystal contacts. More specifically high entropy patches are targeted to minimise the energy barrier associated with the formation of a regular crystal lattice out of protein molecules in solution. The full range of protein engineering strategies available has been extensively reviewed elsewhere [231]. These include:

- surface entropy reduction (SER) – this involves the mutation of residues with high entropy side chains (Lys, Glu & Gln) into low entropy residues (Ala, Ser, Thr, Tyr) [232].

- Cysteine mutation – cysteine residues are able to form covalent disulphide adducts, which can cause aggregation and hinder crystallisation. The mutation of cysteines for serines has widely been utilised in the successfull growth of

diffracting protein crystals such as with mitogen-activated protein (MAP) kinase p38α [233], foot-and-mouth disease virus 3C protease [234] and an RNA-binding protein, She2p [235], among others.

- Truncation/loop removal – Often, the N- and C- termini of proteins are flexible and disordered in many protein structures. As such their removal by truncation of the protein can increase the likelihood of crystallisation. Additionally flexible loop regions or even whole domains may also hinder the formation of stable crystal contacts and can be removed by cloning or by limited proteolysis [236]. In limited proteolysis the protein is incubated with a small amount of trypsin (around 50:1 of protein:trypsin) for a short time period (1-24 hours) either prior to or during the growth of crystals.

1.6.4 X-ray data collection

When collecting X-ray data for a crystal, to obtain all the information about the 3-D composition of the repeating crystal lattice unit (the unit cell) reflections must be recorded over different points of the surface of the crystal. Therefore the crystal is incrementally rotated by 0.2-2.0 ^{o} oscillations with full sets of reflections collected at each stage of rotation. Due to the symmetrical nature of the different Bravais lattices that can be formed, often as little as a 45 ^{o} wedge may be sufficient to obtain a complete dataset. Collecting more data than the minimum required for the space group is beneficial to improve the signal:noise (intensity) of reflections which are measured more than once. Care must be taken to avoid introducing radiation damage and so data quality/resolution must be balanced with the collection of a complete homogenous data set. Each image of reflections is indexed and integrated by computer programs (such as MOSFLM) to record the position and intensity of every spot. At the start of this process the pattern of spot positions observed is determined by the crystal symmetry/space group with the distance between spots inversely proportional to the unit cell dimensions. One value to keep an eye on during indexing is the mosaicity that is a measure of the minimum angle of rotation over which an individual reflection is

observed. It essentially reveals the regularity of the molecular packing within a crystal and as such the lower the data the better the quality of data that will be obtained.

There are three commonly used X-ray sources in protein structure determination: i) X-ray tubes/cathode-ray tubes, ii) rotating anode tubes and iii) particle accelerators (synchrotron sources). With X-ray tubes, electrons from a hot cathode filament are accelerated by electrically charged plates and collide with a water-cooled anode made of a specific metal. The rearrangement of electrons within the metal emits characteristic X-ray wavelengths with Kα and Kβ peaks specific to the metal. X-ray absorbing material is used to guide and filter the majority of the Kα radiation towards the crystals. Typically copper and sometimes molybdenum, with average Kα radiation wavelengths of 1.54 and 0.71 Å respectively, are used to probe atomic structures. The output of X-ray tubes is limited by the amount of heat that can be dissipated from the anode by water. To overcome this, rotating anode tubes were created whereby the metal is rapidly spun to spread the incoming electrons over a larger surface area. Now most in house X-ray sources involve adaptions of rotating anode tubes with a maximum resolution limit achievable limited by the wavelength of the Kα radiation.

In recent years the field of high resolution protein structure solution has been greatly advanced by the development of synchrotron radiation sources. These involve giant rings lined with powerful magnets with electrons circulating close to the speed of light. When the path of the electrons is bent by the magnets, high energy synchrotron radiation is emitted as X-rays. The X-rays are focused into powerful monochromatic beams with tuneable wavelengths and beam sizes. Synchrotron facilities offer increased levels of radiation energy (smaller X-ray wavelengths) than conventional copper anodes for higher resolution data collection. The process is also much faster with whole datasets collected in minutes rather than over hours with in house X-ray tube sources.

1.6.5 Structure solution

Each reflection formed can be described as the result of a complicated wave, which is defined by the summation of many different simple waves. The function that describes this complicated wave for a reflection with indices *h,k,l* is the structure factor F_{hkl}. The structure factors for all measured reflections are converted into a function of the electron density (ρ) present at different real lattice coordinates (x, y, z) by the Fourier Transform described by equation 1.4. To build an electron density map from a complete set of measured reflections, three sets of information are needed about the contributing waves: the amplitude, the frequency and the phase. The amplitude is proportional to the square root of the reflection intensity and the frequency is that of the X-ray source, however the phase is dependent upon the atom locations and cannot be directly measured.

$$\rho_{(x,y,z)} = 1/V \sum_{h} \sum_{k} \sum_{l} F_{h,k,l} \, e^{[-2\pi i(hx + ky + lz)]}$$

Equation 1.4. Fourier transform for calculating the electron density from the summation of the structure factors of each reflection. V = unit cell volume in Å, ρ = electron density, *h,k,l* represents the reflection indices and x,y,z represents the real unit cell coordinates.

There are ways to get around this though and calculate values for the phase so that crystal structures can be solved. Firstly atoms with large numbers of electrons (heavy atoms) can be regularly introduced into the crystal lattice either by soaking or by protein-labelling. At a source x-ray wavelength near the respective atom's characteristic absorption edge, these heavy atoms absorb the incoming radiation and re-emit it as a new wave with altered amplitude and phase. This is known as anomalous scattering. As a result of this scattering the symmetry-related reflections known as Friedal pairs have completely different phases whereas for just the un-labelled protein they would possess exactly opposite values (i.e +1 and -1). The change in intensity measured due to the effect of heavy atoms is known and can be looked up for each element at the source particular wavelength. By drawing the reflections as vectors with length relative to intensity and direction relative to the phase, the difference between measured intensities with and without anomalous scattering can be used to calculate the phases. Once the electron density can be calculated it is typically viewed as a map on graphics

programs such as Coot [237]. To reduce the bias of the model in the electron density output a $2F_o$-F_c map is usually displayed to aid model building/editing. A map of F_o-F_c is may also be shown to highlight significant discrepancies between the model and the observed data. Here positive areas (typically coloured green) represent regions where something is missing from the model and negative areas (typically coloured red) represent regions where the model contains additional atoms whose positions are not justified by the observed diffraction data.

There are two ways of collecting the different datasets at the different required beam wavelengths, the first is using two different crystals (one heavy and one native) in multiple isomorphous replacement (MIR). One problem with MIR is that the heavy crystal must be isomorphous with the native one but heavy atom binding may cause structural changes within the crystal. Therefore a more reproducible approach is known as multiwavelength anomalous dispersion (MAD) where multiple datasets are collected with different incident X-ray wavelengths from a single, heavy atom bound crystal. The technique utilises the measured different absorption/fluorescence patterns of each heavy metals at particular X-ray wavelengths. For this purpose the heavy atom is regularly introduced into the crystal by selenomethionine labelling of the protein during expression and translation in the host cells. At high resolution the phases can be solved from just one dataset in single-wavelength anomalous dispersion (SAD).

In addition the phases can be estimated using molecular replacement (MR) with the program Molrep. This requires a suitable template model from a structurally similar protein structure. The atom coordinates of the model are used to calculate a set of phases initially which are combined with the measured intensities from the diffraction data. For this to work, the model structure must be superposed over the exact location and orientation of the target protein in the unit cell of the collected dataset. To make this task easier the orientation and position searches are conducted separated in Molrep. A Patterson map which is calculated without phase information as the Fourier transform of just the reflection intensities rather than the reflection structure factors, is used to identify the correct orientation of the model. The map displays a pattern of peaks which represent the interatomic distance vectors of all the atoms in the unit cell. Following this, the correctly orientated model is translated through the measured unit

cell until an output is found where the model-calculated reflection amplitudes align with the observed measured amplitudes. This is measured by the R factor (R) (equation 1.5).

1.6.6 Structure refinement and validation

Typically the resolution from measured X-ray data is insufficient to allow free placement of the individual atoms of the protein molecule. Therefore model building/editing is strictly restrained by known plausible values for particular bond angles and lengths. The fit of the progressive structural model to the experimental electron density is refined and assessed by refinement programs such as Refmac 5 [238]. As part of this assessment a value for the R factor (R) is determined and this typically finalises around ten times the value of the high resolution limit (i.e. for a 2.2 Å resolution dataset, an R value of around 22 % or 0.22 could be expected). The problem with the R factor is that the values used to determine it are biased during refinement as the program tries to fit the data to the model. To counter this, an R_{free} value is also determined using a small (~ 5 %) set of relections that are excluded from the refinement process. The R_{free} will always be slightly higher than the R value, however the two should never diverge too far from each other in a good model. Translation libration screw-motion refinement (TLS) [239] can be a way of providing a better fit between the model and the observed diffraction data, lowering the values of R and R_{free}. It works by breaking the protein up into connected rigid body segments that can each move slightly in 3-D space within the structure.

$$R = \frac{\Sigma \left| F_o \right| - \left| F_c \right|}{\Sigma \left| F_o \right|}$$

Equation 1.5. Definition of the R-factor (R). F_o represents the amplitudes observed with the diffraction data whereas F_c represents the calculated amplitudes based on the model structure.

In addition to atomic coordinates, the final model also contains an associated temperature factor (the B-factor) for each atom. The B-factor is a measure of the mean area of displacement of an atom from its identified coordinates in $Å^2$. This is determined by the definition of the electron density surrounding the atom and indicates its flexibility within the crystal structure. Therefore flexible protein regions such as surface loops can

be typically identified by areas with relatively high B-factors. For lower resolution datasets the average B-factor of atoms is naturally going to be significantly higher than for the same structure solved at a higher resolution.

1.7 Aims of this work

This investigation was stimulated on the back of the initiation of collaboration between Mr Damian Mole (University of Edinburgh centre for inflammation research, Queen's medical research institute), Dr Scott Webster (BHF centre for cardiovascular science, Queen's medical research institute), Dr Nigel Botting (School of chemistry, St Andrews University, R.I.P) and Dr Chris Mowat. The main research goal was the development of effective KMO inhibitors for potential therapeutic use. To achieve this several separate research areas were combined: i) the organic synthesis of new compounds, ii) *in vitro* and *ex vivo* testing of compounds on the activity of human KMO, iii) *ex vivo* testing of compounds in models of pathophysiology and iv) structural characterization of the active site and inhibitor binding. The research undertaken in this investigation under the supervision of Dr Chris Mowat, primarily concerns area iv).

The first and most critical aim was to solve the 3-D structure of a KMO to obtain a detailed picture of the enzyme's active site and potential ligand interactions. Mammalian KMOs have been shown to contain a C-terminal transmembrane region that has made them difficult targets for *in vitro* isolation and characterization. To achieve our goals, the bacterial *Pf*KMO was selected as the most promising candidate for study due to its previously known solubility and successful purification. Following the successful study of this enzyme the determined bacterial structure could be used as a model for the prediction/solution of the structure of the human enzyme. Also the screening and characterization of new compounds as potential KMO inhibitors could be quickly undertaken with *Pf*KMO in *in vitro* assays. Additional aims dependent on the success of the project would include solving the structure with the novel compounds and published inhibitors bound to facilitate the rational design of KMO inhibitors. In studying the bacterial enzyme, it will be essential at all times to assess the suitability and transferability of *Pf*KMO as a model for the study of human KMO.

2.1 Cloning and mutagenesis

2.1.1 Cloning primers

Primers were manually designed, checked for formation of secondary structures with the online server OligoAnalyzer (Integrated DNA technologies) and purchased desalted from Sigma-Aldrich or Invitrogen. A list of all the primer sequences used during cloning in this project is shown in table 2.1.

	Primer	Sequence 5'-3' (Restriction sites/mutations)
1	T7 Promoter	TAATACGACTCACTATAGGG
2	T7 Terminator	GCTAGTTATTGCTCAGCGG
3	*Pf*KMO_NTerm_*Nde*_F	GAAGGAGATATACATATGACTGCTACAGACAACGC
4	*Pf*KMO_CT_Trunc1_451_R	AAGCTTTCAGGTCACTTCGTGCTCCAC
5	*Pf*KMO_CT_Trunc2_434_R	AAGCTTTCAGTTGGCCACTGCGAATTTCAG
6	*Pf*KMO_CT_Trunc3_409_R	AAGCTTTCAGAAGGTGACCATGCTGTAACGC
7	*Pf*KMO_C168S_F	GATGGCAGCAACTCGGCGGTGC
8	*Pf*KMO_C252S_F	CAGCCCGTCTTTTGCGCAACTGG
9	*Pf*KMO_C453S_R_*Xho*	GGATCTGCTCGAGTCAGCACAGGTGGCTCAAGGGTGGCAGGCTGCGGGTC
10	*Pf*KMO_C461S_R_*Xho*	GGATCTGCTCGAGTCAGCTCAGGTGGCTC
11	*Pf*KMO_dm2_R_*Xho*	GGATCTGCTCGAGTCAGCTCAGGTGGCTCAAGGGTGGCAGGCTGCGGGTC
12	*Pf*KMO_E372T_F	GGAAAACTACGTGACGATGAGCAGTAAAGTGG
13	*Pf*KMO_Y382F_F	CCACCTTCTTGCTGGAGCGCGAAC
14	*Pf*KMO_R386K_F	CACCTACTTGCTGGAGAAAGAACTGG
15	*Pf*KMO_R386T_F	CACCTACTTGCTGGAGACCGAACTGG
16	*Pf*KMO_LOOP1_R	GATCACCTTGGTTTGGCAGGAACAACGTCACC
17	*Pf*KMO_LOOP1_F	CAAACCAAGGTGATCCGTCTTTTGCACAAACTGGTG

Table 2.1. List of primer sequences used during the investigation. Mutated bases are highlighted in magenta with restriction sites (*Nde*I, *Hind*III & *Xho*I) in cyan.

2.1.2 Polymerase chain reaction (PCR)

The recombinant WT *Pf*KMO gene incorporated into the vector pET17b was constructed and kindly donated by Dr Graham R. Moran (University of Wisconsin-Milwaukee, USA). This plasmid contained ampicillin resistance and utilised the T7 RNA polymerase transcription system. Mutations within 30 bp of the gene termini were introduced on the respective flanking primer in a 50 μL PCR. Prior to setting up PCRs, 0.2 mL thin-wall PCR tubes, 1.5 mL eppendorfs and all pipette tips were sterilised in an autoclave at 120 ^{o}C for 20 mins. All PCR reaction components (see table 2.2) excluding the polymerase were assembled in a sterile 0.2 mL thin-wall PCR tube, vortexed to mix and placed in a thermo cycler express PCR machine (Thermo Fisher). Cycling conditions consisted of a step at 96 ^{o}C for 2 mins followed by 25 cycles of 96 ^{o}C for 1 min, 50-60 ^{o}C for 1 min and 72 ^{o}C for 1.5 mins before a final step at 72 ^{o}C for 5 mins. For every PCR run, the cycle was paused upon entering the first annealing step and the polymerase was added directly into each tube before continuing. The polymerase was not initially added with the rest of the reaction mixture due to its 3'-5' exonuclease activity which was capable of digesting the single stranded primers. Templates, primers and annealing temperatures for each PCR are listed in table 2.3.

Reaction Component	**[Final]**
Vent DNA polymerase (3U/μL)	1.5 U/PCR
10x 'Thermopol Reaction buffer'	1x
dNTPs (2.5 mM)	250 μM
Template DNA – plasmid extraction	≈ 1 ng/μL
Forward primer (10 μM)	0.4 μM
Reverse primer (10 μM)	0.4 μM
EB (10 mM Tris-HCl, pH 8.5)	up to 50 μL

Table 2.2. Final solution components for each 50 μl PCR reaction.

Mutant	**Mutations Introduced**	**Template plasmid**	**Primers**	**T_m (^{o}C)**
C453S	C453S	WT – p17b	3 + 9	
C461S	C461S	WT – p17b	3 + 10	
dm1 (C453S/C461S)	C453S/C461S	WT – p17b	3 + 11	
dm2 (C252S/C461S)	C461S	C252S – p17b	3 + 10	
Trunc_409	CT-truncation after pos. 409	WT – p17b	3 + 6	
Trunc_434	CT-truncation after pos. 434	WT – p17b	3 + 5	
Trunc_451	CT-truncation after pos. 451	WT – p17b	3 + 4	

Table 2.3. Template, primer and annealing temperatures used to construct each mutant *Pf*KMO form.

2.1.3 Megaprimer PCR

Mutations within the *Pf*KMO gene were introduced using the megaprimer technique which involved two separate PCR amplifications. The first PCR was set up as directed in section **2.1.2** amplifying from a primer containing the required mutation to the reverse flanking primer. The PCR product was run on a 1 % (w/v) agarose gel and purified via gel extraction (See section **2.1.5**). The amplifed gene segment bearing the desired mutation served as a 'megaprimer' to complement the forward flanking primer in a second round of PCR (PCR 2). Contents of a typical reaction mixture for the second PCR are shown in table 2.4. Addition of a small amount of Mg^{2+}, commonly used to lower DNA annealing temperatures, was found to slightly enhance the second PCR if initial yields were low. Cycling conditions were designed to intially extend only from the megaprimer for a few cycles before lowering the annealing temperature to allow the flanking primer to bind. An intial step at 96 ^{o}C for 2 mins was followed by 4 cycles of 96 ^{o}C for 1 min, 75 ^{o}C for 45 s and 68 ^{o}C for 45 s before 26 cycles of 96 ^{o}C for 1 min, 60 ^{o}C for 1.5 min and 72 ^{o}C for 1 min. The PCR concluded with a step of 5 mins at 72 ^{o}C. Templates, primers and annealing temperatures for each individual PCR are specified in table 2.5.

PCR 2 Reaction Component	**[Final]**	**Volume (μL)**
Vent DNA polymerase (3U/μL)	1.5 U/PCR	0.5
10x 'Thermopol Reaction buffer'	1x	5
dNTPs (2.5 mM)	250 μM	5
Template DNA – plasmid extraction	≈ 1 ng/μL	1
Forward primer (10 μM)	0.4 μM	2
Purified 'megaprimer'	-	15
$MgSO_4$ (25 mM)	(0/0.5 mM)	(0/1)
EB (10 mM Tris-HCl, pH 8.5)	-	up to 50 μL

Table 2.4. Final solution components for each 50 μl PCR2 reaction. The purified megaprimers were at different concentrations based on the respective yields.

Mutant	Mutations Introduced	Template plasmid	PCR 1 Primers	PCR 2 Primer
C168S	C168S	WT-p17b	7 + 2	3
C252S	C252S	WT-p17b	8 + 2	3
R386T/C461S	R386T	C461S-p17b	15 + 10	3
R386T dm2 (R386T/C252S/C461S)	C252S	R386T/C461S-p17b	8 + 10	3
E372T dm2 (E372T/C252S/C461S)	E372T	dm2-p17b	12 + 2	3
Y382F dm2 (Y382FC252S/C461S)	Y382F	dm2-p17b	13 + 2	3
R386K dm2 (R386K/C252S/C461S)	R386K	dm2-p17b	14 + 2	3

Table 2.5. Template, primer and annealing temperatures used to construct each mutant *Pf*KMO form.

2.1.4 Overlap extension PCR

The potential loop region HHQSPAAQPAS (aa 240-250) was mutated to PNQGD using overlap extension PCR as summarised in figure 2.1. Two initial PCRs A + B were set up as outlined in section **2.1.2**. PCR A amplified from the N-terminal flanking primer (primer 3) to a reverse primer with a terminal overhang containing the new sequence (primer 16). PCR B then amplified from a forward primer (primer 17) bearing the complementary overhang to that generated in A to a reverse flanking primer (primer 11). The products were run on a 1 % (w/v) agarose gel, excised and purified (see section **2.1.5**) and combined in a final overlap extension PCR as detailed in table 2.6. The cycling conditions were as described for PCR 2 in section **2.1.3**.

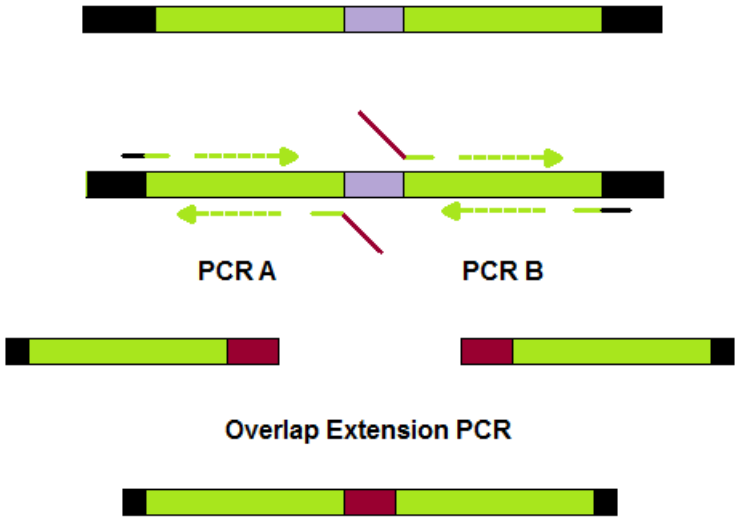


Figure 2.1. Schematic overview of the overlap extension PCR process. PCR A and PCR B generate two separate halves of the gene (green) with complimentary overlapping ends containing the new altered sequence (red). These then combine in the final step to recreate the whole gene. Non-coding plasmid regions are shown in black with the original target sequence in grey.

Overlap extension PCR Component	[Final]	Volume (μL)
Vent DNA polymerase (3U/μL)	1.5 U/tube	0.5
10x 'Thermopol Reaction buffer'	1x	5
dNTPs (2.5 mM)	250 μM	5
Template - PCR A product	-	5
Template - PCR B product	-	5
F flanking primer 3 (10 μM)	0.2 μM	1
R flanking primer 11 (10 μM)	0.2 μM	1
EB (10 mM Tris-HCl, pH 8.5)	-	up to 50 μL

Table 2.6. Final solution components for final 50 μl recombination reaction.

Mutant	Template plasmid for PCR A + B	Primers PCR A	Primers PCR B
'LOOP1 dm2' (HHQSPAAQPAS->PNQGD/C252S/C461S)	dm2-p17b	3 + 16	17 + 11

Table 2.7. Template, primer and annealing temperatures used to construct the mutant.

2.1.5 Agarose gel electophoresis

Agarose gel electrophoresis is a technique used to separate DNA based on its size. The separation properties of a gel were varied by changing the concentration of agarose. Typically 0.7 % (w/v) agarose gels were used for plasmid DNA (5-8 kb) whereas 1.0 % (w/v) agarose gels were used for PCR products (up to 1.5 kb). Molecular biology grade agarose (Fisher) was dissolved in 50 mL TAE buffer (40 mM Tris-acetate pH 8.0, 1 mM EDTA) with 2 mins heating in a microwave. After a couple of minutes cooling at room temperature, ethidium bromide was added to a concentration of 0.5 μg/mL. The edges of the gel tank were initially sealed with 1 mL of the molten agarose before the rest was poured in and combs were added to form wells. The gel was left for 20 minutes to set and then covered with TAE buffer prior to the gentle removal of combs. Samples were mixed with the appropriate 6x loading buffer (see table 2.8) and 5-20 μL was loaded into the wells of the gel alongside 5 μL of 1kb full scale DNA ladder (Fisher). A steady current of 50-80 μA with a voltage of around 80-100 V was applied to the gel until the DNA had migrated sufficiently far. DNA was visualised with UV irradiation in a sealed box. The target DNA bands of the correct expected size were excised and purified using a

PureLink™ gel extraction kit (Invitrogen) with a final elution volume of 40 μL of the supplied elution buffer.

6x Loading Buffer (made with distilled water)	DNA range (kb)
0.25 % w/v Bromophenol blue, 40 % w/v sucrose	> 0.5
0.25 % w/v Xylene cyanol, 40 % w/v sucrose	< 2.0

Table 2.8. DNA loading buffers for gel electrophoresis with rough usable size range for the target DNA band.

2.1.6 Restriction digestion and ligation

All 50 μL of each completed PCR was split across two adjacent wells on a 1 % (w/v) agarose gel and the products excised and purified (See section **2.1.5**). The respective NEB restriction buffer and 40 U of each NEB restriction endonuclease were added, briefly vortexed and incubated in a water bath at 37 ^{o}C for 4 hours. The plasmid vectors pET17b (Amp resistance, no tag) and pET28b (Kan resistance, N-terminal hexa-His tag) extracted from cell cultures (see section **2.1.8**) were digested similarly. The digested reactions were loaded onto a 0.7 % (w/v) agarose gel and the appropriate insert and vector DNA bands were gel extracted prior to ligation. Some digested PCR products were simply incubated at 70 ^{o}C for 10 mins to denature the restriction endonucleases prior to ligation – this saved time and resources at the cost of a slightly reduced cloning efficiency. Ligation reactions were set up as 10 μL mixtures in sterile 1.5 mL eppendorfs, briefly vortexed and incubated at 4 ^{o}C overnight for 16-20 hours.

2.1.7 Transformation of chemically competent hosts

Ligation reactions were directly used to transform chemically competent BL21(DE3) cells (Stratagene) using the heat shock method. The competent cell stock was thawed on ice and to a 50 μL aliquot of cells 4 μL of ligation reaction or 0.5 μL of plasmid solution was added. After 30 mins on ice, the mixture was incubated at 42 ^{o}C for 20 s before being transferred back into the ice bath for 2 minutes. Addition of 0.75 mL of SOC media (see

section **2.2.1**) preceeded incubated at 37 ^{o}C for 1 hr. An aliquot of 50-100 μL was spread onto LB agar (see **2.2.1**) plates containing the respective antibiotic for screening and incubated at 37 ^{o}C for at least 16 hours. Individual colonies were selected and restreaked onto fresh agar plates. For long-term storage, a loop full of restreaked cells were gently suspended in a 1 mL cryo-vial (Nunc) containing 900 μL of LB (see **2.2.1**) and 100 μL of DMSO and stored at -80 ^{o}C for up to a year.

Occasionally when low yields of insert/vector were experienced, initial transformation of ultra-competent NEB5α cells with ligation reactions was required. Subsequently successful clones were transferred into BL21(DE3) cells for expression and protein production. For production of SeMet labelled protein, clones were moved into the methionine auxotroph B834(DE3) pLys S. When transferring a successful clone to a different *E.coli* strain, the transformation protocol was the same except that 0.5 mL of the purified plasmid (see section **2.1.8**) was used in place of the ligation reaction.

2.1.8 Extraction of plasmids and Sanger sequencing

Restreaked transformants were used to inoculate 10 mL of LB plus the respective antibiotic and grown for 16 hours at 37 ^{o}C in an orbital incubator at around 200 rpm. After harvesting cells at 3-4000 g for 10 mins plasmids were extracted and purified using a QIAprep spin miniprep kit (Qiagen). Pure plasmids were eluted with 40 μL of elution buffer into sterile 1.5 mL eppendorfs. The Genepool facility carried out BigDye reactions and Sanger sequencing for two samples from each clone. To prepare samples, 5 μL of pure plasmid (≈ 500 ng) was mixed with 1 μL of 3.2 μM T7 promoter primer for one sample and with 1 μL of 3.2 μM T7 terminator primer for the second. Typically 700-950 bp of sequence could be read from each flanking primer and combined to provide the whole gene sequence of a clone.

2.2 Cell Growth and Protein Production

2.2.1 Media

All media/agar used was sterilised by autoclaving at 120 ^{o}C for 20 mins and used within two days. Antibiotics were added to media/agar at 1x concentration after autoclaving but once the solution had cooled down to below 50 ^{o}C. When at a similar temperature, roughly 25 mL of sterile agar was poured into each Petri dish and allowed to set before being stored at 4 ^{o}C. Prior to use, agar plates were dried for an hour in a 37 ^{o}C incubator. The media recipes used during the study are summarised in table 2.9.

LB media	Tryptone 10 g/L, Yeast extract 5 g/L, NaCl 8 g/L (pH 7.0)
LB agar	LB media + Agar 15 g/L
SOC media	Tryptone 20 g/L, Yeast extract 5 g/L, NaCl 0.58 g/L, KCl 0.19 g/L, $MgCl_2$ 10 mM, Glucose 20 mM (pH 7.2)
10x M9 Minimal salts (Autoclaved)	NH_4Cl 20 g/L, KH_2PO_4 60 g/L, Na_2HPO_4 (anhydrous) 136 g/L
1000x Metals mix (Filter Sterilised)	0.05 M $FeCl_3.6H_2O$, 0.02 M $MnCl_2.4H_2O$, 0.02 M $ZnSO_4.7H_2O$, 4 mM $CoCl_2.6H_2O$, 4 mM $NiCl_2.6H_2O$
1000x Vitamins mix (Filter Sterilised)	1 mg/mL Riboflavin, Niacinamide, Pyridoxine monohydrochloride & Thiamine.
100x Amino acid mix 1 (Filter Sterilised)	4 mg/mL amino acids: Ala, Arg, Asn, Asp, Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Pro, Ser, Thr, Val
100x Amino acid mix 2 (Filter Sterilised)	4 mg/mL aromatic amino acids: Phe, Trp, Tyr (pH 8.0 with NaOH)
Minimal media **glucose autoclaved separately as 20 % (w/v)*	1x M9 minimal salts, 1x metals mix, 1x vitamins mix, 1x amino acid mix 1, 1x amino acid mix 2, 0.4 % (w/v) D-glucose*, 2 mM $MgSO_4$, 40 mg/L L-SeMet (pH 7.2)
1000x Amp	60 mg/mL Ampicillin
1000x Chloramphenicol	25 mg/mL Chloramphenicol *dissolved in ethanol*
1000x Kan	50 mg/mL Kanamycin

Table 2.9. Media and antibiotic recipes for cell growth.

2.2.2 Growth of BL21(DE3) cultures for protein production

Starter cultures of 10 mL LB plus Amp (pET17b clones) or Kan (pET28b clones) were inoculated with the respective frozen DMSO cell stock and grown for 24 hours at 37 ^{o}C, with shaking at 200 rpm. Each starter culture was used to inoculate separate 1.0 L baffled shake flasks containing 1.0 L of LB plus the respective antibiotic. The cultures were subsequently incubated for 15-17 hours at 22 ^{o}C with shaking at 180 rpm. At the end of incubation the A_{600nm} was recorded of a mixture of 50 μL of cells with 950 μL of H_2O. Multiplication by the dilution factor of the mixture gave the OD of the culture. Cultures were harvested by centrifugation for 45 mins at 30000 g at 4 ^{o}C when they reached an OD of 1.0-2.0. The cell pellets were stored at -20 ^{o}C until required.

2.2.3 Growth of B834(DE3) pLys S cultures

B834(DE3) pLys S cells were grown in minimal media supplied with all amino acids except methionine and to allow production of SeMet labelled *Pf*KMO. Starter cultures were grown in LB media plus Amp/Kan and chloramphenicol (pLys S) as described in section **2.1.3**. Subsequently each starter culture was used to innoculate 1.0 L of minimal media (see section **2.2.1**) containing the required antibiotics. The cultures were initially grown at 30 ^{o}C with shaking at 180 rpm until an OD of 0.6-1.0. At this point IPTG was added to 0.1 mM, the temperature lowered to 22 ^{o}C and the cultures were grown for a further 4 hours. The cells were then harvested as described in section **2.2.2**.

2.3 Protein Isolation and Purification

2.3.1 Purification buffers

The composition of the buffers used during this investigation is described in table 2.10. During purification of protein for crystallography, HPLC grade water was used to make the S-75 buffer. All buffers were made fresh for each purification and stored at 4 ^{o}C. DTT was added to buffers immediately before use to minimise the oxidation of cysteines.

Resuspension Buffer	20 mM HEPES pH 7.5 10 mM NaCl 1.0 mM DTT
100x PMSF	0.2 M PMSF dissolved in ethanol
Q-Seph Elution Buffer	20 mM HEPES pH 7.5 110 mM NaCl 1.0 mM DTT
Ammonium Sulphate Buffer	20 mM HEPES pH 7.5 75 % saturated ammonium sulphate 1.0 mM DTT
G-25 Buffer (Assays)	20 mM HEPES pH 6.8 10 mM NaCl 1.0 mM DTT
S-75 Buffer (Crystallography)	20 mM HEPES pH 6.8 150 mM Na acetate 1.0 mM DTT

Table 2.10. Purification buffer compositions.

2.3.2 Cell lysis

Frozen cell pellets were thawed at room temperature and resuspended to a concentration of 1 g per 10 mL of resuspension buffer plus 1x PMSF. Aliquots of resuspended cells of 30-50 mL were lysed by sonication on ice with 6 rounds of 20 s pulses. The lysate was clarified by 4 ^{o}C centrifugation at 30000 g for 50 mins and the insoluble pellet discarded. The supernatant containing soluble proteins, nucleic acids and other cytosolic compounds was filtered using a 0.22 μm syringe filter. The filtered crude protein solution was used immediately for purification.

2.3.3 Purification of *Pf*KMO for assays

The filtered crude protein solution was loaded onto a 60 mL FPLC Q-sepharose 26/10 column pre-equilibrated with resuspension buffer. The flowthrough of unbound proteins was discarded. The column was then washed with Q-sepharose elution buffer which contained 110 mM NaCl. After 100 mL of washing, *Pf*KMO and free FAD would consistently elute as a yellow band with absorbance peaks at 280, 375 & 450 nm. Fractions of 5 mL elute were collected in test tubes using an automatic fraction collector. Fractions with absorbance at 450 nm were assayed for *Pf*KMO activity as described in section **2.5.1**. Those fractions containing activity greater than a ΔA_{340}/min of 0.500 were pooled.

The pooled Q-sepharose eluate was slowly brought to 50 % ammonium sulphate saturation by addition of ammonium sulphate buffer with constant, but very gentle, stirring. After 10 mins of equilibration, the solution/suspension was centrifuged at 20000 g for 15 mins to pellet pure, precipitated *Pf*KMO. The supernatant was discarded. Pelleted *Pf*KMO, visibly yellow, was incubated for 10 mins in G-25 buffer and then gently resuspended by inversion. A total volume of around 60 – 80 mL of G-25 buffer was used, with roughly 15 mL per pellet. The pooled solution of resuspended enzyme was concentrated to <5 mL using Millipore 30 kDa MWCO spin columns centrifuged at 1,000 g. Concentration was performed as quickly as possible to minimise precipitation of *Pf*KMO due to residual ammonium sulphate from the previous step.

The concentrated solution was filtered with a 0.22 μm syringe filter and loaded onto a pre-equilibrated 20 mL G-25 column to remove residual ammonium sulphate. The column was pre-equilibrated with G-25 buffer. The yellow, *Pf*KMO band was followed and collected in a sterile plastic bijou tube. The concentration of *Pf*KMO was estimated (see section **2.4.2**) and the protein solution was diluted to 10 μM using G25 buffer, with the addition of glycerol to a final concentration of 5 % (v/v). The *Pf*KMO solution was aliquotted out into 1 mL cryovials and flash-frozen in liquid nitrogen. The protein was then stored at -80 ^{o}C until required.

2.3.4 Purification of *Pf*KMO for X-ray crystallography

For protein purified for X-ray crystallography, an additional size-exclusion chromatography step was added to the purification protocol described in section **2.3.3.** This step not only improved protein purity but more importantly it improved the homogeneity of the *Pf*KMO solution, separating monomeric protein from higher molecular weight aggregates**.** The protein solution eluted from the G-25 column was concentrated to a volume of approximately 1 mL and injected onto a 280 mL HiLoad 26/10 S-75 FPLC column pre-equilibrated with S-75 buffer. The absorbance at 280, 363 & 450 nm of the eluted solution was monitored throughout. The yellow protein peak that eluted at the volume expected for monomeric 50 kDa *Pf*KMO was collected as 1 mL fractions with an automated fraction collector. The expected volume was determined from the calibration curve shown in figure 2.2.

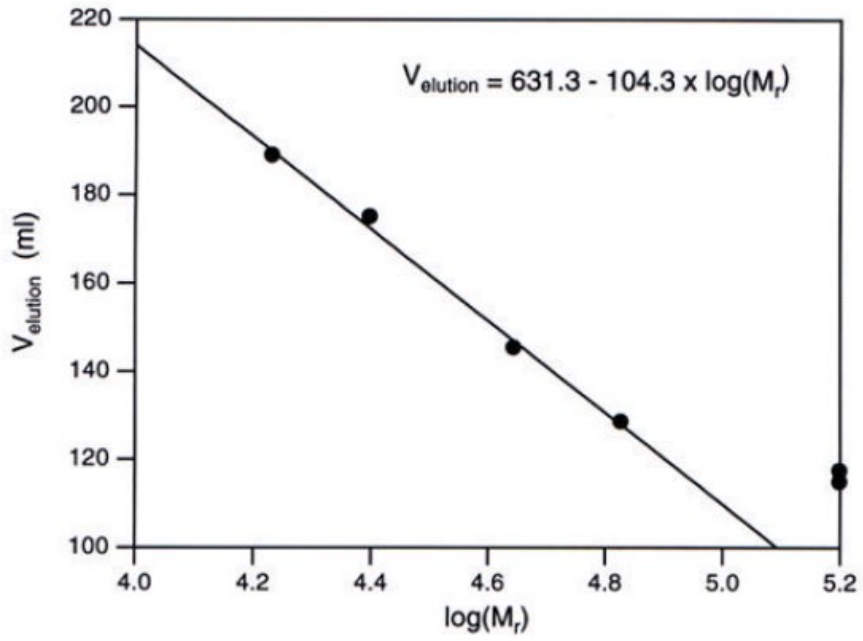


Figure 2.2. Calibration curve for S-75 chromatography showing the volume of eluate corresponding to the log(molecular weight) for several standard proteins.

Fractions with the smallest ratio of A_{280nm}:$A_{450nm,}$ representing the ratio of total protein:*Pf*KMO, were pooled. The protein concentration was estimated (see section **2.4.2**) and the solution was then concentrated with 30 kDa MWCO spin columns to the concentration required for crystallography, typically 5-15 mg/ml. It was observed that substrate-bound crystal forms required greater protein concentrations than substrate-free forms to grow crystals. During concentration, the protein solution was diluted with S-75 buffer dilution buffer to dilute the Na acetate from 150 mM to around 20-30 mM. For the growth of L-Kyn soaked crystals, L-Kyn was dissolved in S-75 dilution buffer (typically as a 10 mM stock) and added to the protein solution to a final concentration of 1 mM. Aliquots of 25 or 50 μL were flash-frozen and stored at -80 ^{o}C. Addition of glycerol prevented crystal formation at the protein concentrations tested.

2.4 Analysis of enzyme concentration and purity

2.4.1 SDS poly-acrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE is a method for separating proteins based on their molecular weight. The anionic detergent SDS denatures the proteins in the sample and binds to the polypeptide chain at regular intervals to apply a negative charge. The protein solution is boiled in order to help completely denature the proteins in the sample. During electrophoresis a current is passed through the gel so that negatively charged SDS migrates towards the anode. The longer the polypeptide chain, the slower the migration through the gel so that chains are effectively separated based on their size. Comparison with standard known molecular weight proteins allows the estimation of molecular weight based on distance migrated in a set time period.

20 μL of protein solution was mixed with 20 μL of 1× NuPAGE sample buffer in an eppendorf tube and left in boiling water for 5 minutes. Up to 20 μL was loaded into the wells of a pre-poured 12 % NuPAGE bis-tris gel (Invitrogen) alongside 5 μL of Seeblue Plus 2 (Invitrogen) molecular marker. The gel was pre-placed and then run in 1× NuPAGE MOPS running buffer (Invitrogen). A voltage of 150 V (120 mA, 60 W) was applied for approximately one hour until the dye front reached the bottom of the gel. The gel was

then removed and soaked in Coomassie blue stain (40 % methanol, 10 % acetic acid, 1 % Coomassie brilliant blue) with agitation for 20-25 mins. The stain was removed and the gel was left in hot water or in destain (50 % water, 40 % methanol, 10 % acetic acid) until blue protein bands were clearly visible.

2.4.2 Spectrophotometry

When Crozier & Moran purified *Pf*KMO in 2007 [2] they calculated the extinction coefficient (ε) for KMO-bound FAD at 450 nm to be 12300 $M^{-1}cm^{-1}$ at pH 7.0. Throughout the investigation, this value was used to estimate the concentration of the *Pf*KMO in solution by manipulating the equation **A = εcL** (**A** = absorbance, **ε** = extinction coefficient, **c** = concentration & **L** = path-length, i.e. width of cuvette). The 250-800 nm absorption spectra of protein purification samples using a quartz cuvette was recorded to identify fractions containing *Pf*KMO. Such spectra also provided an instant estimation of the purity of *Pf*KMO by measuring the ratio of A_{280}:A_{450}, with values < 7.0 indicating a suitable purity for crystallographic study.

2.5 PfKMO kinetic analysis

2.5.1 Measuring the rate of turnover

Figure 2.3 shows the UV-vis absorbance spectrum of each of the substrates L-kynurenine and NADPH, and the products 3-HK and $NADP^+$ of *Pf*KMO turnover. Notably the absorbance peak of 3-HK is only slightly blue-shifted by 17 nm from that of the starting compound, L-kynurenine. On the other hand, generation of $NADP^+$ results in complete loss of the large absorbance peak seen with NADPH. The UV-vis spectra in figure 2.4 show the change in absorbance over time during typical *Pf*KMO catalysis with excess substrates. There are two visible changes occurring, the large fall at 340 nm as NADPH is oxidised and a relatively small increase between 380-450 nm due to conversion of L-kynurenine to 3-HK.

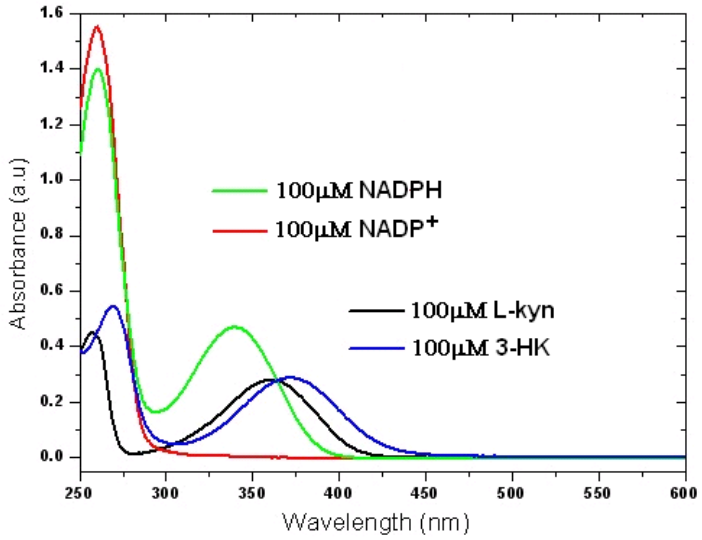


Figure 2.3. UV-vis absorption spectra for the substrates and products involved in the reaction catalysed by KMO.

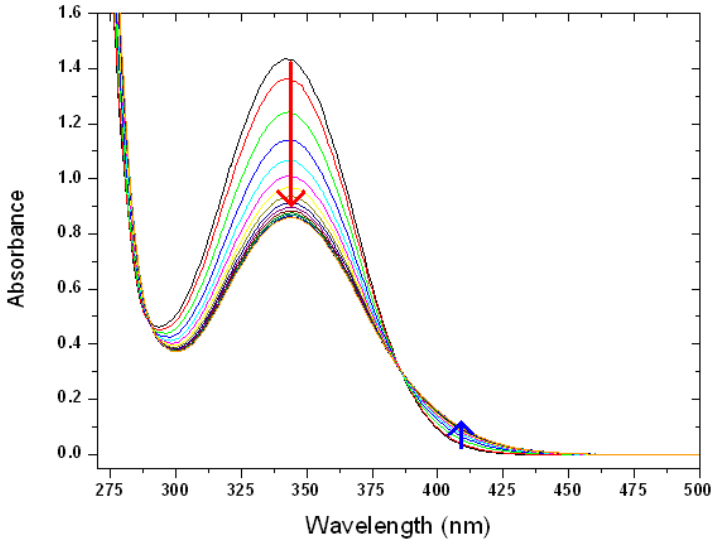


Figure 2.4. Absorbance spectra taken at regular time intervals during a reaction of *Pf*KMO with excess L-kynurenine and NADPH.
The red arrow indicates the fall in absorbance as NADPH is oxidised to $NADP^+$
The blue arrow shows the increase in absorbance as 3-HK is generated from L-kynurenine

To quantify *Pf*KMO-catalysed turnover of L-kynurenine, NADPH oxidation was measured by monitoring the absorbance change at a single wavelength of 340nm for a 1mL reaction mixture in a quartz cuvette. The typical 1x assay buffer consisted of 20 mM HEPES pH 8.0 unless stated otherwise. The enzyme was added last to initiate the reaction (with gentle inversion to mix) and the absorbance was recorded for one minute. The reaction rate was measured from the initial, linear absorbance change (the first 5-30 s depending on enzyme/substrate concentrations). In this timescale the pseudo-steady-state hypothesis is valid and the reaction obeys Michaelis-Menten kinetics. Morevover, the second substrate NADPH was present in excess and did not limit the overall reaction rate. Where use of excess substrates are detailed, 250 µM NADPH and 200 µM L-

kynurenine were the concentrations used. Generally, turnover was measured with 0.1 μM pure *Pf*KMO or 50 μL of crude protein samples (during purification). The rate of absorbance change per minute was converted into turnover (s^{-1}) using equation 2.1. For negative controls, the rate was measured in the absence of L-kynurenine and in the absence of enzyme before commencing a set of assays. All reactions were conducted at 25 ^{o}C but the enzyme and NADPH were stored on ice to prevent denaturation/degradation.

$$\textbf{Turnover } (s^{-1}) = \Delta A_{340}/\text{min} \,/\, (60 \times \varepsilon_{NADPH,\,340nm} \times \text{enzyme concentration})$$

Equation 2.1. Converting change in absorbance into a value for NADPH turnover.

$\varepsilon_{NADPH,\,340nm}$ = 6,200 $M^{-1}cm^{-1}$ and enzyme concentration is in M

2.4.2 Determination of kinetic constants k_{cat} and K_m

The rate of turnover of *Pf*KMO over varying concentrations of L-kynurenine was measured to determine the catalytic constants k_{cat} and K_m. The standard assay buffer was 20 mM HEPES pH 7.5, 2 mM DTT, 10 mM NaCl. Excess NADPH (250 μM) was present and the concentration of L-kynurenine was varied from 0-250 μM (> 15 different values). The final concentration of *Pf*KMO present in each assay was 0.1 μM. The k_{cat} and K_m were determined from a Michaelis-Menten curve fitted to a plot of turnover against concentration of L-kynurenine, using Origin 8 graphing software (OriginLab Corporation). Each set of assays was run three times to ensure the data were reliable, with final values for the kinetic constants determined from the averages of the separate runs. A plot from a single run with dm2 *Pf*KMO is shown in figure 2.5. The error within measuring the initial rate data due to background noise/pipetting variations was estimated to be around 1-15 % depending on the extent of the absorbance change being measured.

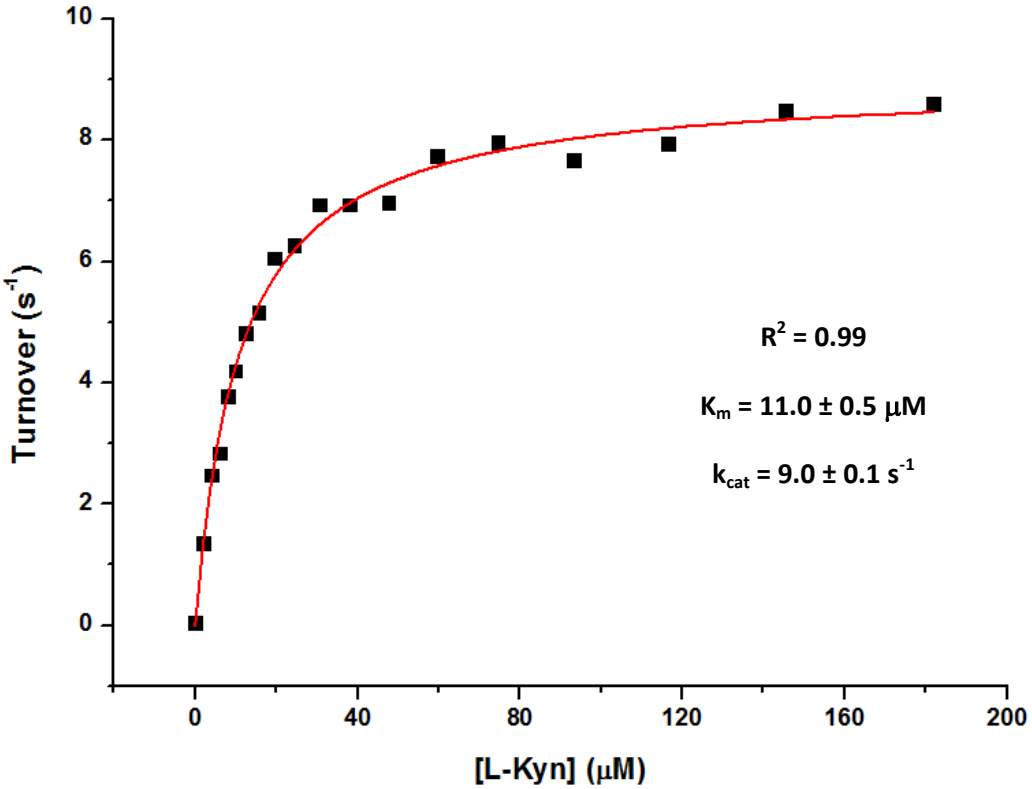


Figure 2.5. Michaelis-Menten plot for dm2 *Pf*KMO showing the relationship between the rate of turnover and substrate concentration. The R^2 value and determined kinetic constants for the plot are displayed.

2.5 Investigating Inhibition of PfKMO

2.5.1 Screening potential *Pf*KMO inhibitors

Initially 20 mM stocks of each compound were made up with water or with DMSO depending on solubility. Stocks were stored at -20 ^{o}C in between uses. Screening involved testing the rate of turnover of 50 μM L-Kyn by 0.1 μM *Pf*KMO in the presence of 0.5-100.0 μM of each inhibitor. This was assayed as described in section **2.4.1** and compared to turnover in the absence of inhibitor (replaced with water/ DMSO) which was measured in triplicate to get an accurate average. The relative percentage inhibition was calculated for each compound relative based on the measured turnover numbers.

2.5.2 96-well plate-based compound screening

As a greater number of compounds were arriving to be tested the need for a high-throughput 96-well plate screen arose. For this purpose a SpectraMax M5 multimode plate reader (Molecular Devices) was set to 25 ^{o}C with a 2 s plate mixing period before commencing data recording. For data collection, a multipipette was used to start the reaction in a whole column of eight wells simultaneously by addition of the substrate, L-Kyn. The reaction was run for three minutes with A_{340} values measured every 8s for each well. In addition before substrate was added the absorbance change was similarly recorded for two minutes to check the background rate of uncoupled NADPH oxidation and the A_{600} was measured to check for compound precipitation. Following the assay, the ΔA_{340} was plotted for each well and the reaction rate calculated from the initial linear period. The background rate was then subtracted before the percentage inhibition was determined relative to uninhibited controls. The first row of wells on the plate contained the negative controls, where the inhibitor was replaced by water or DMSO depending on the solvent used to suspend the compounds being tested. The final reaction components are detailed in table 2.11. DMSO was limited to 3 % v/v to balance compound solubility with enzyme activity.

Component	Volume	[Final]
NADPH (2.5 mM)	25.0 μl	250 μM
Inhibitor (333.3 μM)	7.5 μl	10 μM
Assay buffer (34.25 mM HEPES pH 8.0)	182.5 μl	25 mM
KMO (2.5 μM)	10.0 μl	0.1 μM
L-Kyn (0.5 mM)	25.0 μl	50 μM
	(250 μl)	

Table 2.11. Final reaction components each well of the 96-well plate for inhibitor screening.

2.5.3 Characterisation of *Pf*KMO inhibitors

The turnover of *Pf*KMO was measured over 6-9 concentrations of L-kynurenine (in the range 0-200 μM) and repeated in the presence of increasing concentrations of a potential inhibitor compound. Typically 5-7 inhibitor concentrations were selected in addition to a set of reactions with no inhibitor (substituted by water/DMSO). As a rough

guide a concentration of inhibitor that showed 20 % inhibition in screening would be around the K_i value for that compound. This provided a suitable middle concentration when choosing the range of inhibitor concentrations to be tested. Typically a broad range of substrate concentrations with a relatively narrow range of inhibitor concentrations provided the best data.

Both substrate and inhibitor concentrations were made up using dilution series from initial stocks. Inhibitors were diluted with whichever solvent was used to initially dissolve the respective compound. Sufficient buffer was made for the whole experiment and pre-incubated in 30 mL falcon tubes in a 25 ^{o}C water bath at least 30 mins before data collection commenced. NADPH and the enzyme were kept on ice during the entire experiment and concentrations were tested in a random order to minimise the influence of enzyme denaturation during proceedings. The final reaction components are summarised in table 2.12.

Component	**1 x**	***(Final conc.)***
NADPH (10 mM)	25 μL	*250 μM*
L-Kyn (stocks)	20 μL	*varied*
Inhibitor (stocks)	20 μL	*varied*
KMO (10 μM)	10 μL	*0.1 μM*
Assay buffer (section 2.5.1)	925 μL	

Table 2.12. Final reaction components for absorbance-based K_i assays in 1 ml cuvettes.

Data were plotted for each kynurenine concentration with 1/v (y-axis) against [inhibitor] (x-axis). Each series of kynurenine concentrations generated straight lines of best-fit that intersected at $-K_i$ on the x-axis and $1/k_{cat}$ on the y-axis. Due to experimental error within the data, the lines generally did not all intersect perfectly at just one location. Therefore, all the intersection points were considered and the average value used to estimate the K_i value. Standard deviation of all the intersections provided an estimate for the error of the K_i. The type of inhibition was determined from a plot of [S]/v (y-axis) against [inhibitor] (x-axis) as described in section **1.5.2**. Each set of assays were repeated three times, each with newly made buffer, NADPH, substrate stocks and inhibitor stocks and a fresh aliquot of enzyme to determine an accurate, average K_i value. An example of the two plot types is displayed in figures 2.6 and 2.7 from assay data obtained with the inhibitor GM769.

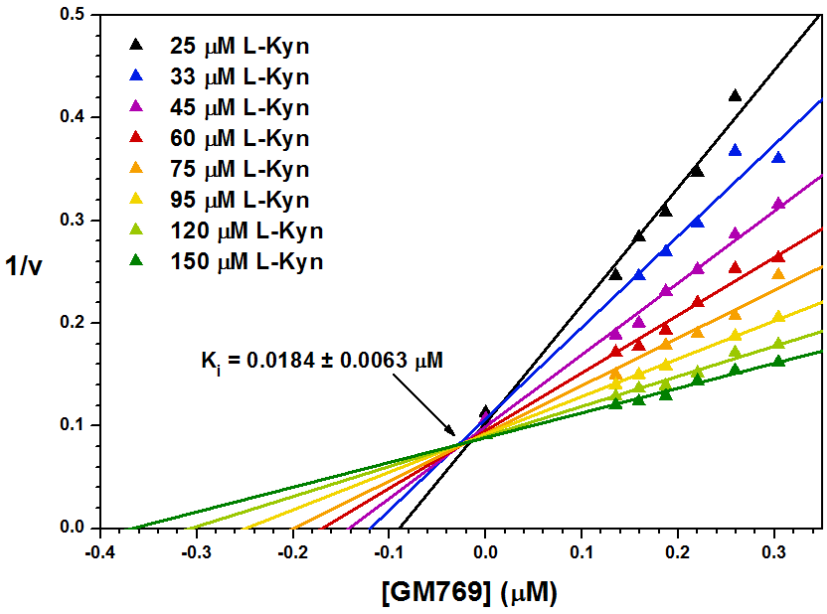


Figure 2.6. Dixon plot of 1/v against concentration of the inhibitor GM769 for different concentrations of substrate. The K_i was determined from the average of all measured intersection points between the lines of best fit.

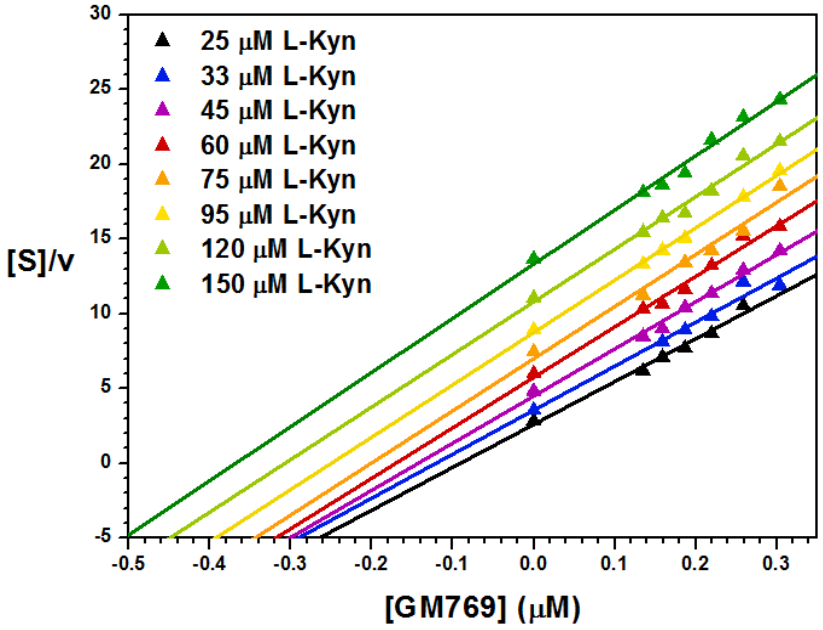


Figure 2.7. A plot of [S]/v against concentration of the inhibitor GM769 for different substrate concentrations. The relatively parallel nature of the best-fit lines suggests competitive inhibition.

2.6 Enzyme Crystallisation for Structure Determination

2.6.1 Crystallisation by vapour diffusion

Crystallisation trials with purified *Pf*KMO were set up using 24-well Linbro trays (Hampton research/Molecular Dimensions) and siliconised cover slips. Paraffin oil was used to grease the rims of the wells to generate effective seals with the cover slips. Either 500 μL or 1000 μL of different reservoir solutions was added to each well. On each cover slip, 1 μL of reservoir solution was pipetted into a 1 μL drop of protein solution. The cover slip was then inverted and placed over the respective well. Trays were sealed in aluminium foil to prevent photo-bleaching of the FAD cofactor and were incubated at 4^{o}C or 17^{o}C. The sealed drops were observed under a light microscope daily for 4 days and subsequently on a weekly basis.

Initial screening with *Pf*KMO involved setting up hanging drop vapour diffusion (h.d.v.d) trays with two drops per cover slip – one with *Pf*KMO and one with *Pf*KMO pre-soaked with 1mM L-kynurenine. Commercially-available Hampton Research screening kits (HR2-110, HR2-112, HR2-107, HR2-109, HR2-114, HR2-116 and HR2-126), with a range of different ready-made solutions were used to search for conditions that grow crystals. Trays were incubated at 4 ^{o}C or 17 ^{o}C. Sitting drop vapour diffusion (s.d.v.d) trays were also used during the growth of inhibitor-bound crystals. This involved the addition of 1 μl of both protein and reservoir solution in to a raised well above 500 μl of reservoir solution. Each well was sealed with clear tape.

2.7.2 Crystal growth using micro-seeding

Microseeding was combined with crystallisation screening trials and optimisation attempts to promote the growth of single crystals. To prepare the hetero-seed solution, a native substrate bound dm2 crystal was crushed, transferred to 100 μl of stabilising solution (0.1 M HEPES pH 7.0, 20 % v/v PEGMME 550, 10 % v/v 2-propanol) and micro-centrifuged to pellet large crystal fragments. The remaining supernatant was transferred to a new eppendorf and used to create serial dilutions via the transfer of 10 μl of seed solution into 90 μl of stabilising solution. During crystal growth the the different dilutions

were tried for the growth of single regular crystals. Dilutions were stored at -20 ^{o}C for re-use over a period of a few months. The procedure was also incorporated into commercial vapour-diffusion screening but failed to yield new conditions for the growth of *Pf*KMO crystals.

2.7.3 Flash-freezing crystals and data collection

Single protein crystals were removed from their drop by mounting in loops and flash-frozen in liquid nitrogen. To prevent the formation of damaging ice crystals, protein crystals were soaked in a cryoprotectant solution before freezing. The cryoprotectants initially explored were the mother liquor plus 10 or 20 % (saturation) ammonium sulphate, (w/v) ethyleneglycol, (w/v) xylitol, (w/v) MOPEG or (v/v) glycerol. Ultimately, complete cryoprotection was achieved with short 5-10s sequential soaks in mother liquor plus 10 % and then plus 20 % ethylene glycol (EG). Alternatively some crystals benefitted from addition of 1 μl of the cryoprotectant (mother liquor plus 20 % EG) to the crystal drop prior to soaking in the cryoprotectant. Frozen crystals were either immediately screened for diffraction using the in-house x-ray generator or stored in a Dewar vessel full of liquid nitrogen. For the collection of complete datasets, crystals were taken to several beamlines at the Diamond Light Source synchrotron in Oxford.

Initially crystals were screened to test their diffraction by collecting three single images 90^{o} apart using minimal beam intensity. Different edges of the crystal or at times whole crystal grid scans were used to identify the strongest diffracting region. Subsequently complete datasets were collected using a rotation range based on the strategy output from imosflm. Fine slicing with 0.2 or 0.3 ^{o} rotations per image provided the best data when matched with an equal exposure (in seconds per degree). A reduced beam intensity was used for the collection of a first dataset before higher intensities were attempted on neighbouring patches of the crystal. Large (>0.2 mm^{3}), substrate and non-substrate effector bound crystals required the maximal available beam intensities to maximise the resolution. Often the crystals displayed asymmetrical diffraction properties requiring testing of multiple angles to identify the most profitable region to collect. After collection, images were indexed and integrated with imosflm [240], tested for the best

Laue group with Pointless [241] and scaled with Scala [242] during which a random 5 % of the reflections were separated for the calculation of R_{free} during refinement. Inspection of the output log graphs and removal of images with high R_{merge} values (whilst maintaining complete datasets) helped to improve data quality. The process was coordinated with the CCP4 program suite [242].

2.7.4 Data collection with selenomethionine labelling

Selenomethionine (SeMet) labelled protein in a multiwavelength anomalous diffraction (MAD) experiment was used to solve the first *Pf*KMO structure. This involved the collection of four complete datasets using different regions of the same crystal. Different X-ray wavelengths were used as determined by a fluorescence scan of the mounted crystal controlled by the program CHOOCH [243]. An output scan for the SeMet crystal is displayed in figure 2.8. The four wavelengths chosen, known as: peak, inflection, high energy remote (HER) and low energy remote (LER), were selected by the X-ray wavelength corresponding to: the peak f'' energy, the peak f' energy, a value in the post-peak plateau and a value in the pre-peak plateau. Each dataset was separately indexed and integrated and separately scaled with anomalous pairs separated for merging statistics.

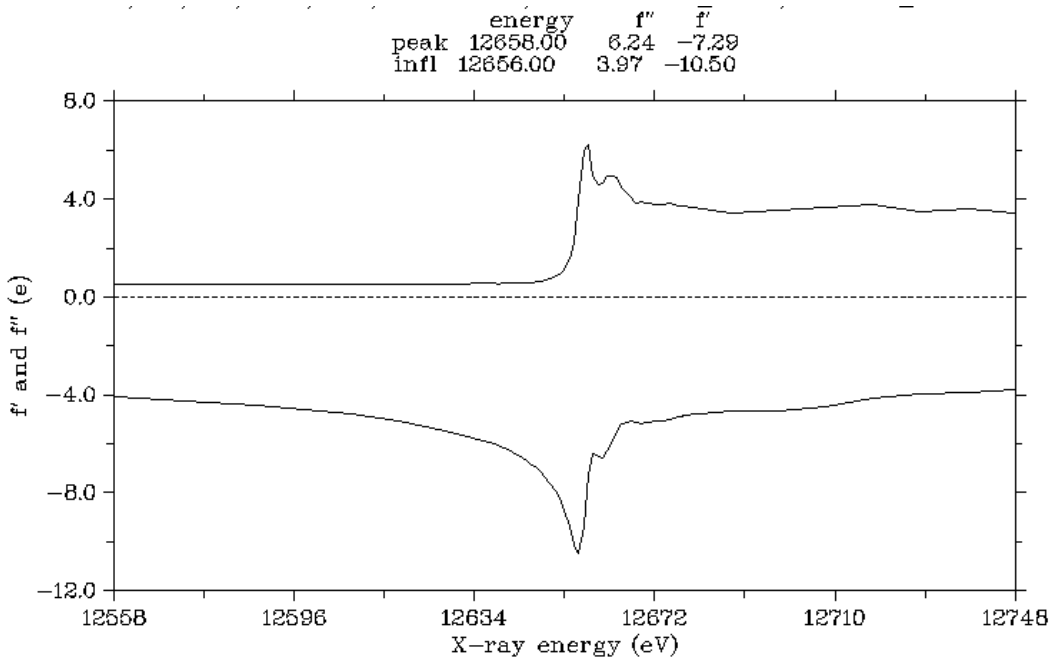


Figure 2.8. CHOOCH output displaying fluorescence scan with f'' (*above*) and f' (*below*) absorption at each X-ray energy. The peak and inflection values are labelled. The data were collected from the SeMet dm2 substrate-bound crystal used for final data collection at synchrotron beamline I04.

Chapter 3 – *Pf*KMO crystallisation and structure solution

3.1 Introduction

3.1.1 KMO from *Pseudomonas fluorescens*

Eukaryotic KMOs have been shown to contain a C-terminal trans-membrane helix that anchors them on to the outer mitochondrial membrane [182, 191] (see section **1.3**). It is widely known that any proteins with such membrane binding domains are intrinsically more difficult to study *in vitro* due to solubility issues. Whilst there are many detergents available these days for successfully purifying membrane proteins, there is still a large bottle-neck when it comes to trying to grow crystals suitable for X-ray diffraction and structure solution. With bacterial forms of KMO showing reasonable sequence similarity, but crucially lacking this C-terminal trans-membrane region, they offered a potentially easier route towards KMO structure solution. Therefore KMO from *Pseudomonas fluorescens* (*Pf*KMO) was selected for this investigation as this form had previously been cloned and was readily purified and extensively characterised by Crozier-Reabe & Moran [1, 2]. *Pf*KMO shares 32 % identity and 51 % sequence similarity with the entire human KMO protein as determined using an Emboss needle pairwise alignment (See appendix section **1.1** for full alignment). It is 461 residues long with a pI of 6.20 and functions as a monomeric 50.7 kDa class A flavin-dependent aromatic hydroxylase (FAH).

3.1.2 FAH crystallization

As highlighted in section **1.2**, there have been numerous FAH crystal structures that have been solved in recent years, and these are summarised in Table 3.1. A diagram depicting the 3-D structures of each of the enzymes was shown in section **1.2.3.2**. Each has a similar 3-D fold with conserved structural elements corresponding to the FAD and substrate binding domains. At the C-terminus of the enzymes the conservation breaks down, often with extensions consisting of whole new domains. Typically the shorter C-

termini simply seal the back of the substrate binding pocket from bulk solvent. The larger C-terminal extensions found in PhH, *m*-HBH, AVH, RebC, PgaE and CabE form a thioredoxin-like fold [170, 172, 174, 175, 177]. This third domain has no known function in the FAHs except occasional involvement in dimerization. From alignments constructed using the EMBOSS needle pairwise alignment tool, the most similar sequences to *Pf*KMO was those of PQSL and MHPCO with 23 and 22 % identity respectively.

								Pf*KMO alignment	
FAH	**pdb**	**Ligand**	**Flavin**	**Res. (Å)**	**Lattice**	**Res.**	**Dom.**	**% ID/cons**	**Score**
PQSL	2x3n	None	*Mixed*	1.75	$P\ 2_1\ 2_1\ 2_1$	399	2	23/32	227
MHPCO	3gmb	None	*In*	2.10	$C\ 2$	379	2	22/35	227
	3gmc	Substrate	*In*	2.10	$C\ 2$				
AVH	3ihg	Substrate	*Out*	2.49	$P\ 6_3$	535	3	19/28	185
TetX	2xdo	None	*In*	2.09	$P\ 1$	378	2	21/33	182
	2y6q	Substrate 1	*In*	2.37	$P\ 1$				
	2y6r	Substrate 2	*In*	3.10	$P\ 1$				
phzS	3c96	None	*Out*	1.90	$P\ 2_1\ 2_1\ 2$	402	2	20/29	160
p-HBH	1pbe	Substrate	*In*	1.90	$C\ 2\ 2\ 2_1$	394	2	20/32	159
PhH	1pn0	Substrate	*Mixed*	1.70	$P\ 2_1$	665	3	16/25	142
HpxO	3rp6	None	*In*	2.20	$P\ 2_1\ 2_1\ 2_1$	384	2	18/29	135
	3rp7	Substrate	*In*	2.04	$P\ 2_1\ 2_1\ 2_1$				
CabE	2qa2	None	*In*	2.70	$P\ 6_5\ 2\ 2$	489	3	17/27	117
PgaE	2qa1	None	*In*	1.80	$F\ 2\ 2\ 2$	491	3	21/33	113
DHPH	2vou	None	*In*	2.60	$P\ 3_2\ 2\ 1$	397	2	19/32	111
m-HBH	2dkh	Substrate	*Out*	1.80	$P\ 3_2\ 2\ 1$	639	3	16/27	92
RebC	2r0c	None	*Out*	1.80	$P\ 2_1$	529	3	18/26	89
	2r0g	Substrate	*In*	2.37	$P\ 2_1$				
	2r0p	Substrate Anal.	*Out*	2.10	$P\ 2_1$				

Table 3.1, A summary of current FAH crystal structures and their sequence similarity to *Pf*KMO. Res. = number of residues in chain, Dom. = number of domains in each single polypeptide chain structure. % ID/cons = % identical residues and % conserved residues, score = alignment strength assigned by program.
*EMBOSS needle pairwise alignment used for whole sequences relative to the length of the largest sequence. For references and full protein identification see section **1.2.3.2.**

Some of the structures present direct visualization of each of the two flavin conformations required during catalysis. Interestingly there is no conserved pattern between the flavin conformation and the presence/absence of substrate but the flavin is most commonly found in the *in* orientation. In four of the thirteen cases both substrate free and substrate bound forms of the respective enzyme have been obtained, each with no evident changes in the crystal lattice. The size and structure of the aromatic substrates varies in each case but they are always bound in a buried pocket between

domains 1 and 2. The substrate carbon that is hydroxylated by each enzyme is superimposable, positioned 4-5 Å from the flavin C_{4a} atom responsible for forming the catalytic hydroperoxy intermediate.

3.1.3 Structural information from other L-kynurenine catabolizing enzymes

As displayed in section **1.1**, the kynurenine pathway can diverge down three possible routes at the stable intermediate L-Kyn. Of course one route proceeds via KMO but this leaves two other enzymes that utilise L-Kyn as a substrate – kynureninase (KYNase) and kynurenine amino transferase (KAT). Whilst the enzymes are functionally and structurally unrelated to KMOs, information about how these enzymes bind L-Kyn may be useful in the understanding of L-Kyn binding in KMOs. Mammalian KATs I, II, III and IV are capable of catalyzing the transamination reaction that transforms L-Kyn into KYNA and belong to the class of fold-type I pyridoxal 5-phosphate (PLP)-dependent enzymes. Crystal structures of KAT enzymes with L-Kyn bound have recently emerged including those of human KAT II at 1.95 Å resolution [244], mouse KAT III at 2.8 Å resolution [245] and mouse KAT IV at 2.4 Å resolution [246]. The key protein-substrate binding interactions in each are displayed below in figure 3.1. Interestingly, some of the substrate interactions and even the substrate orientation vary between each form of KAT. In each active site though, there are conserved Asn and Arg side chains that form crucial interactions to stabilise the carboxyl group of L-Kyn.

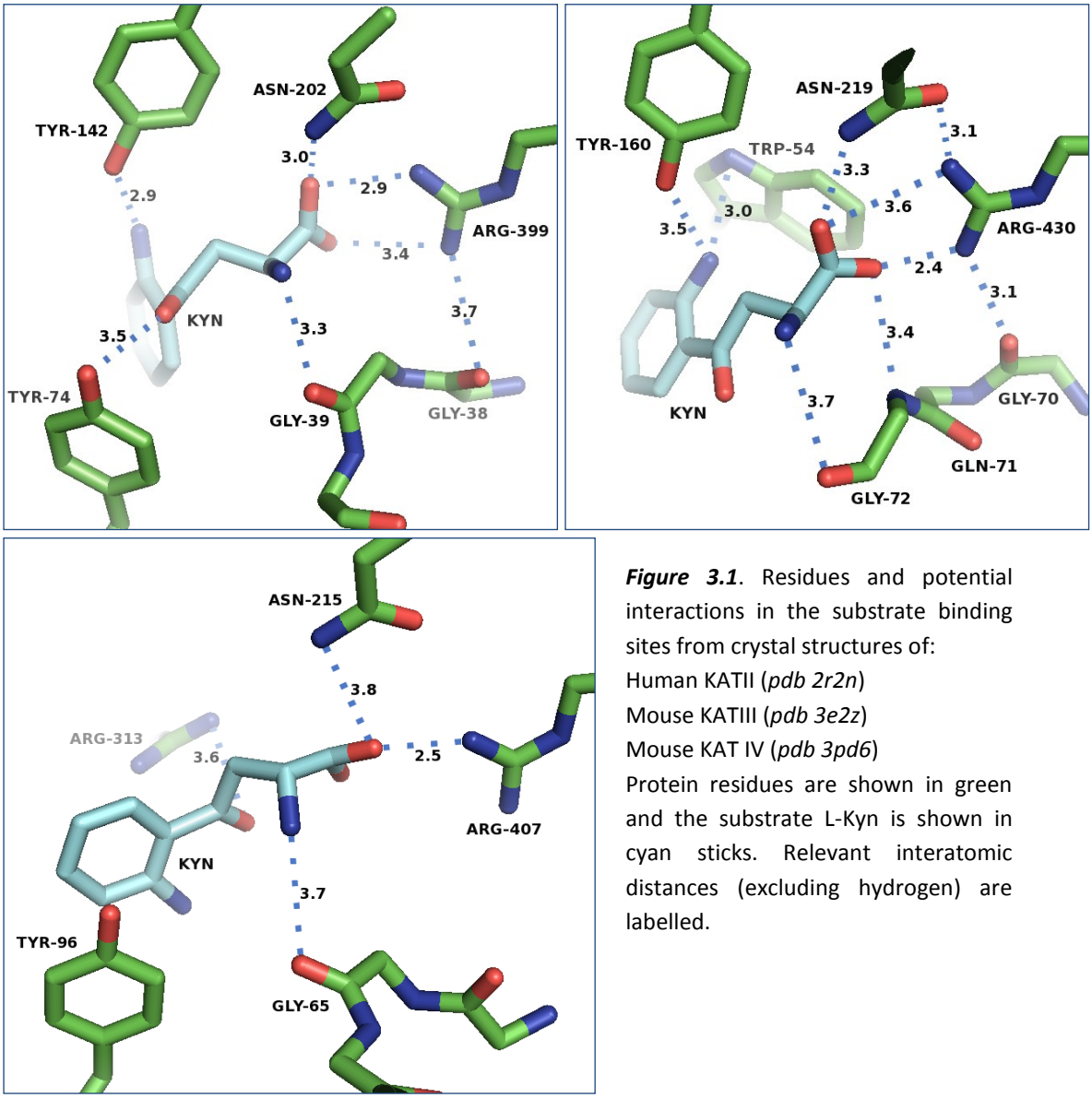


Figure 3.1. Residues and potential interactions in the substrate binding sites from crystal structures of:
Human KATII (*pdb 2r2n*)
Mouse KATIII (*pdb 3e2z*)
Mouse KAT IV (*pdb 3pd6*)
Protein residues are shown in green and the substrate L-Kyn is shown in cyan sticks. Relevant interatomic distances (excluding hydrogen) are labelled.

KYNase is also a PLP-dependent enzyme that catalyses the hydrolytic cleavage of 3-HK into 3-HAA and L-Ala. Bacterial and some fungal KYNases preferentially cleave L-Kyn over 3-HK to give AA and L-Ala. Crystal structures have been solved of bacterial forms of KYNase, but unfortunately not with L-Kyn bound. The only KYNase crystal structure with a ligand bound is of the human enzyme with 3-hydroxyhippuric acid (3XH) - an inhibitory analogue of 3-HK – bound, which is at a resolution of 1.70 Å [247]. The active site from this structure is displayed below in figure 3.2. As might be expected, there is again a conserved arginine side-chain that forms a salt bridge with the ligand's carboxyl tail. In human KYNase additional interactions with the carboxyl group are formed by the side-chain of a serine and a histidine residue. Whilst the substrate binding sites of KYNase and the KAT enzymes may not allow direct

modelling of the respective L-Kyn binding environment in KMO, they may be useful when assessing the selectivity of KMO inhibitors that are competitive with regards to L-Kyn binding.

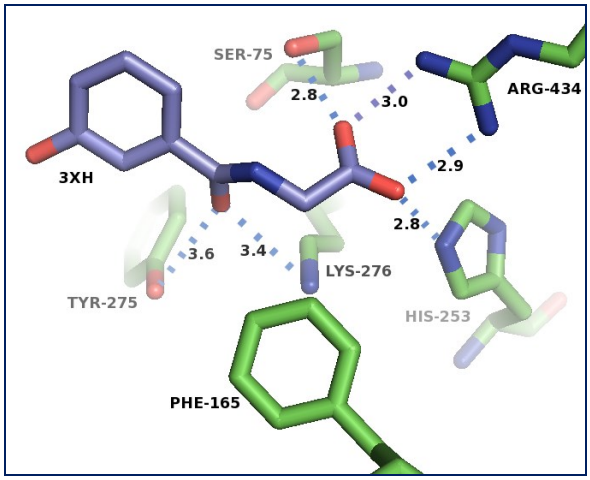


Figure 3.2. Residues and potential interactions in the substrate binding site from the crystal structure of Human KYNase (*pdb 3e9k*)
Protein residues are shown in green and the competitive inhibitor 3XH is shown in violet sticks. All relevant interatomic distances labelled.

3.1.4 Aims

The main and initial aim of this work was to solve a KMO crystal structure using the bacterial form from *Pseudomonas fluorescens*. The enzyme was to be produced in *E.coli* expression strains and purified without a tag using the protocols developed by Crozier & Moran [2]. Following successful purification, conditions were to be found and refined for the successful growth of strongly diffracting *Pf*KMO crystals. If possible, structures of both the substrate bound and substrate free forms of the enzyme were to be solved. More than just solving the 3-D structure, the aim was to significantly develop the structural understanding of the enzyme and assess the relevance of this information to human KMO. Moreover, knowledge of structural information regarding KMO enzymes and the key interactions involved in substrate binding would help the rational development of novel KMO inhibitors for therapeutic use.

*3.2 Pf*KMO – *Structural insights from bioinformatics*

3.2.1 Insights from *in silico* structure prediction tools

The webservers XtalPred [248], SERp [249] and foldIndex [250] were used to predict the possible likelihood of success of *Pf*KMO and human KMO crystallization and to predict potentially hindering features of their amino acid sequences. XtalPred uses analysis of patterns from protein sequences that either lead to successful or failed attempts at crystallization. SERp analyses protein sequences to identify patches of consecutive high entropy residues that are unfavourable and could hinder the formation of crystal contacts. Without structural knowledge there is no guarantee that patches identified will be positioned on the protein's surface to actually affect crystallization. Finally FoldIndex looks at the physicochemical properties of local groups of residues to identify potential areas of the sequence that are likely to be disordered. The output results from each were compared to values obtained from the current range of successfully crystallised FAH enzymes. The data are displayed in table 3.2.

XtalPred analysis of the *Pf*KMO sequence placed it in the lowest crystallization class, which represents sequences most suitable for crystallization. As can be seen from analysis of the FAH sequences, this assigned class is not definitive as the crystal structure of two enzymes (PhH & *m*-HBH) that were placed in class 4 out of 5 were solved to resolutions of 1.70 and 1.80 Å respectively. Despite these anomalies it was still pertinent to note that the sequence of human KMO was placed in the highest, most unsuitable class, which was likely to be due to the presence of a transmembrane region. The biophysical analysis of the *Pf*KMO sequence revealed no major abnormalities relative to the crystallised FAH enzymes. Human KMO did though possess relatively high pI, instability index and GRAVY hydrophobicity values. High entropy residues did not appear to be a major problem for potential crystallization of KMO, although it only takes a few strategically positioned to disrupt the formation of a favourable crystal contact. Interestingly both KMO sequences had a relatively high percentage of helices, which may reflect a prediction of the nature of the unconserved C-terminal residues of the enzyme. In all the bioinformatics analysis was not conceived to directly discover whether *Pf*KMO could be crystallised. Importantly though it revealed that there were no obvious

problems within the amino acid sequence of *Pf*KMO that needed addressing before starting the investigation.

	Results for crystallised FAH enzymes			
Analysis	**Lowest - highest values**	**Average**	***Pf*KMO**	**Human KMO**
Crystallization class[1]	1 (optimal) – 4 (diff.)	2 (suboptimal)	1 (optimal)	**5 (very diff.)**
Isoelectric point	4.9 – 6.8	5.7	6.2	**9.2**
N residues	378 – 665	413	461	486
Instability index	24.7 – 42.7	36.1	34.5	42.2
GRAVY index	-0.52 – -0.08	-0.23	-0.16	-0.08
Predicted Disordered	0 – 6 %	2 %	2 %	3 %
Predicted Coil	34 – 49 %	41 %	41 %	34 %
Predicted Helix	29 – 39 %	33 %	39 %	**44** %
Predicted Sheets	22 – 27 %	25 %	**20** %	22 %
Residues = K, E, Q	8.4 – 18.8 %	12.9 %	12.6 %	13.2 %
SERp outcome[2]	3 – 7 patches (2.4-7.64)	4 patches (4.0)	6 patches (2.4-3.2)	3 patches (5.0-5.7)
FoldIndex: unfoldability[3]	0.064 – 0.190	0.147	0.176	0.189
charge	0.006 – 0.045	0.022	0.017	0.027
phobic	0.442 – 0.491	0.474	0.483	0.491
Disordered	0 – 28 %	13 %	4 %	5 %

Table 3.2. Biophysical predictions and analysis from amino acid sequences using XtalPred, SERp and foldIndex servers. Results for sequences of FAH enzymes with solved crystal structures (including all those listed in Table 3.1) were collated with lowest, highest and average values displayed. For comparison the results based on the sequences of *Pf*KMO and human KMO are also displayed. The data is colour coded with low values (or more promising values) in green, medium in blue and high (or more unsuitable values) in red. Values outside the observed range for FAH enzymes are bold with yellow highlighting.

[1] Predicted propensity of sequence features towards successful crystallization, scale of 1 = optimal to 5 = very difficult.

[2] SERp server identifies patches of consecutive high entropy residues in the sequence that may hinder formation of stable crystal contacts. The value in parenthesis represents the lowest and highest scores assigned to selected patches with higher values indicating more harmful patches.

[3] The unfoldability value represents the average from all of the residues; a high value represents a greater propensity to being folded whereas a low values represents a greater likelihood that a residue is disordered.

3.2.2 Insights from primary sequence alignments

A colour-coded alignment of the sequence of *Pf*KMO with the sequences of known FAH enzymes revealed only strong conservational blocks around the three FAD-binding motifs GxGxxG, DG and GD as shown below in figure 3.3. As with the other FAH enzymes these motifs were present in relatively similar sequence locations indicating that *Pf*KMO may have a similar overall fold. Moreover the conserved proline just after the final motif, which has been shown to stack with the flavin ring when in the '*in*' position, is also present in the sequence of *Pf*KMO. With 461 residues, *Pf*KMO is longer than some of the two domain enzymes of length 380-400 aa, however it is not as long as the three-domain FAH enzymes of length 529-635. Moreover, the C-terminal 80 residues of *Pf*KMO do not share any sequence similarity with the longer FAH enzymes. As a result the enzyme may possess a third, C-terminal domain but one that is unlikely to resemble those evident in some of the current FAH structures.

Figure 3.3. Shows three sections taken from a multiple sequence alignment of *Pf*KMO with the sequences of the FAH enzymes whose structures have been solved. Blocks containing globally conserved FAH motifs are outlined by the three boxes. The alignment was generated with clustal omega and is colour-coded based on the biophysical properties of the amino

Even without structural information about KMO, much detailed information can be derived from alignments of known KMO sequences. Such alignments can trace the evolutionary divergence of the enzyme and can locate which areas of the sequence have been selectively protected from this process. A number of residues were strictly identical among all tested sequences of KMO. Looking at the KMO sequences alone, however, it is impossible to distinguish functionally important residues from conserved residues with only structural relevance (hydrophobic patches etc). Therefore the addition of structural information from FAH structures can help locate which of the KMO conserved residues are well positioned to partake in substrate binding. The full alignment is shown in appendix section **1.2** with the residues that directly interact with the flavin cofactor and respective substrate separately highlighted (in yellow and green respectively) on each FAH sequence. On the aligned *Pf*KMO sequence, residues that are conserved among all KMOs are highlighted blue. Barring some sporadic patches there were a few aligned segments of sequence that contained the majority of the respective substrate binding residues for the FAH enzymes. The major block is shown below in figure 3.4, which comprises a lot of the residues that line the β-sheet platform at the bottom of the substrate binding pocket. As can be seen there are several conserved KMO residues that align in relation to these and may therefore include several key substrate binding residues. In all a refined pool of forty-two conserved KMO residues potentially populating the substrate binding pocket was generated.

```
PfKMO           PNALHIWPHGDYMCIALPNLDRSFTVTLFLHHQSPAAQPASPCFAQLVDGHAARRFFQRQ 269
p-HBH           SHELIYANHPRGFALCSQRSATRSRYYVQVPLTEK----------VEDWSDERFWTELKA 245
PhH             --CAIHSAESGSIMIIPREN-NLVRFYVQLQARAE--KGGR----VDRTKFTPEVVIANA 317
PQSL            ----LYVDSQGGLAYFYPIGFDRARLVVSFPREEA--RELM----ADTRGESLRRRLQRF 251
MHPCO           IDMWNFWPRVQRILYSPCNENELYLGLMAPAADPR----------GSSVPIDLEVWVEMF 249
RebC            LHFFLMLSSSLRFFLHALDGRGLYRLIVGVDD-------------------ASKSTMDSF 255
phzS            DEHWSRLVAYPISARHAAEGKSLVNWVCMVPSAAVGQLDNEADWNRDGRLEDVLPFFADW 269
AVH             G-WYYLHHPEFKGTFGPTDRPDRHTLFVEYDP------DEG----ERPEDFTPQRCVELI 269
m-HBH           --VAIQSEQ-GNVLIIPREGGHLVRFYVEMDK-LD--ADER----VASRNITVEQLIATA 297
```

Figure 3.4. A portion of a sequence alignment between *Pf*KMO and FAH enzymes whose 3-D structures have been solved. For each FAH structure, residues interacting with FAD are highlighted yellow and those interacting with the respective substrate are in green. For *Pf*KMO residues conserved across all sequences of KMO are highlighted in cyan.
For full alignment see appendix section **1.2**.

The evolution and divergence of KMO sequences also provides clues as to which residues are more likely to be located on the surface with no functional significance. Separate multiple sequence alignments were created for mammalian, insect, fungal and bacterial KMO sequences with conserved residues from each compared. The overall results are summarised in appendix section **1.3**. This information can help further develop an understanding of the potential structural layout of the enzyme and can have further uses including: helping in the actual building and validation of a structural model and in identifying potential residues to substitute by site-directed mutagenesis for the improvement of crystal contacts. To enhance the reliability of these data only KMOs that were structurally similar to each other were considered. Therefore *Pf*KMO was aligned with all bacterial KMOs and not eukaryotic KMOs that were known to have additional transmembrane regions. For every residue in the alignment a percentage identity and percentage hydrophilicity were recorded. Logically any residue with low conservation but a high frequency towards hydrophilic substitutions was considered as a potential candidate and this accounted for 47/461 residues.

3.2.3 Prediction of *Pf*KMO secondary structure

The *Pf*KMO amino acid sequence was fed through numerous secondary structure prediction servers including: APSSP2, Baldig, GOR4, HNN, Psipred, Jnet and SSpro. The results were aligned and compared to assess the consistency of the predictions. In all the results showed relatively strong agreement with only small discrepancies regarding the start/stop position of predicted α-helices. The prediction of β-sheets was slightly more

troublesome in certain sequence areas but on the whole the predictions were similar. From these a consensus prediction was derived and compared to the known secondary structural elements of some published FAH structures (see appendix section **1.4**).

Interestingly, up to residues 370-390 there was a strong correlation between the locations of α-helixes and β-sheets in all the FAH sequences despite very limited sequence conservation. The blocks of secondary structure with the strongest conservation were those involved in the FAD-binding core of the proteins – residues 1-80 & 300-360. Otherwise a lot of the β-sheets were less rigidly conserved and often intertwined with several sequence insertions. The consensus prediction for the secondary structural layout of *Pf*KMO fitted very well with those from the aligned FAH structures, but again only up until around residue 370. For the C-terminal residues of *Pf*KMO, the consensus prediction did not match any of the other larger FAH enzymes and predicted a largely helical arrangement. In summary, secondary structure prediction programs seemed to potentially predict the location of such structural elements in the sequence of *Pf*KMO with reliable consistency. If the predictions are accurate then they suggest that the enzyme may possess a similar structural layout to the conserved first two domains seen in other solved FAH structures. The C-terminal 80-90 residues however may extend into a novel third domain relative to the current FAH structures.

3.2.4 Prediction of *Pf*KMO tertiary structure

The amino acid sequence of *Pf*KMO was fed into the homology-based protein tertiary structure prediction servers: I-TASSER [251], SWISS-MODEL [252], Phyre [253], and LOOPP [254]. A sample of the models can be seen below in figure 3.5. Every model that was constructed by the servers was based on the structure of one of the aforementioned published FAH structures. As a result of this each model was biased towards the respective template. This meant that different areas of different models appeared to be modelled better, i.e. were denser with secondary structure. Half of the models finish between residues 373-416 with only the I-TASSER models covering the whole 461 residue sequence. Where predicted, the C-terminal residues were often disordered strands or had little secondary structure and with few common consistencies. In the full

models the C-terminal 90 residues always started from the bottom of the long kinked α-helix (residues 346-376) but then extended out in different directions (figure 3.6).

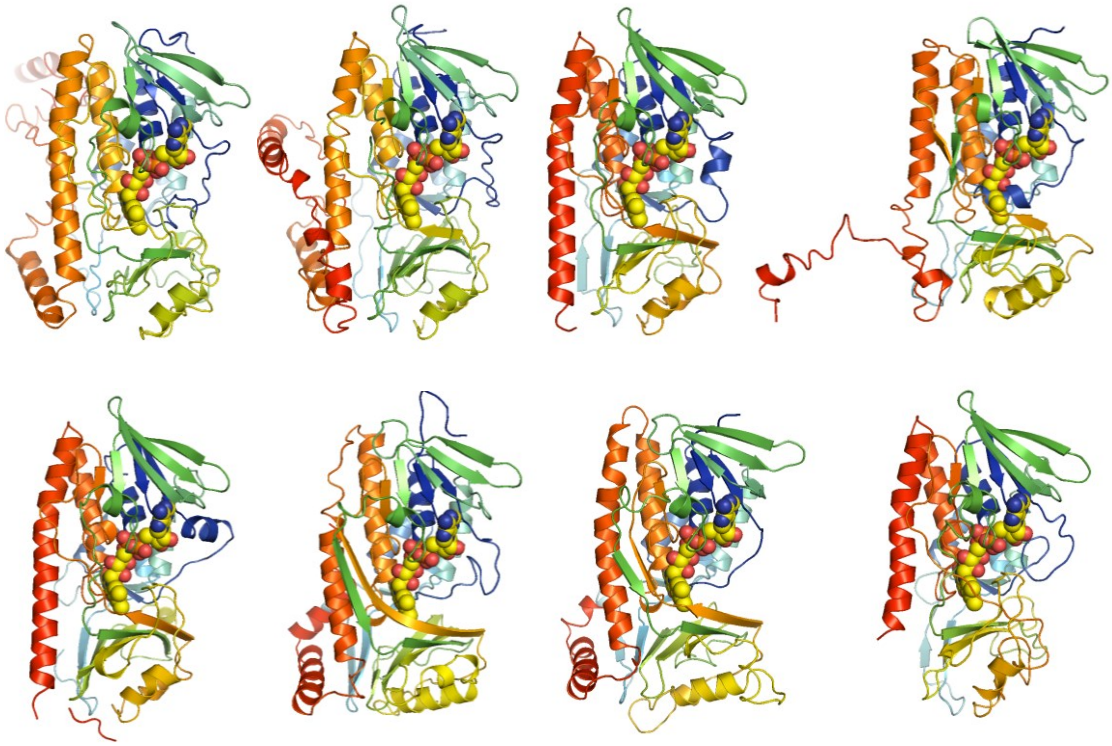


Figure 3.5. Cartoon representation for a sample of the *Pf*KMO tertiary structure models generated by the: I-TASSER, SWISS-MODEL, Phyre and LOOPP servers. The residues are coloured blue through to red going from the N- to the C-terminus with FAD shown in yellow spheres.

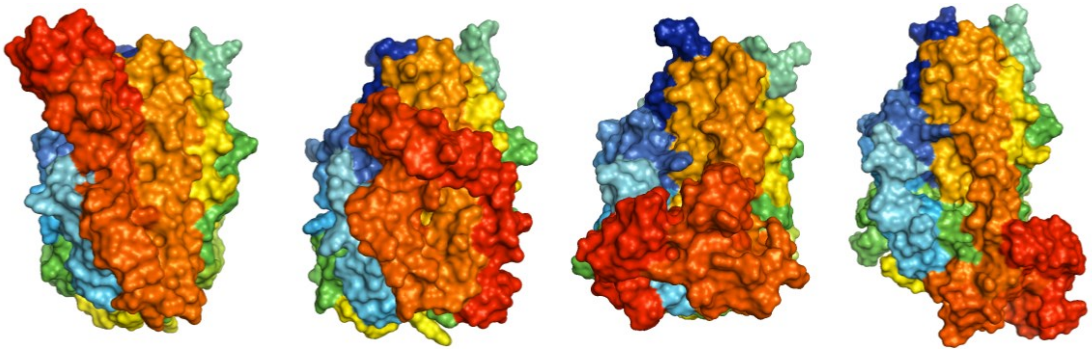


Figure 3.6. Surface representation showing four full length *Pf*KMO models generated by I-TASSER. Each is orientated similarly with the back of the protein – relative to the FAD binding site – facing out. The models are again colour-coded from blue at the N-terminus to red at the C-terminus and show the variability in the prediction of the arrangement of the C-terminal residues.

For the first 372 residues most of the secondary structure was similarly arranged however the exact 3-D positions of individual residues was variable. For example the side chains of only two residues were observed to be completely solvent exposed in every single model. As a result of this, the exact prediction/position of key residues was inconclusive from the models alone. The typical buried substrate binding chamber seemed to be accommodated by all of the models however its access route to bulk solvent could not be predicted. Some of the models did though offer potential starting templates for structure solution once suitable X-ray diffraction data had been obtained.

3.2.5 Summary

Bioinformatic analysis of simply the amino acid sequence of *Pf*KMO has revealed hidden structural information. Commonly used sequence analysis tools including XtalPred suggested that the enzyme's biophysical properties are suited to potentially favour successful crystallization. The combination of sequence alignments with known FAH structural information selected a refined pool of conserved KMO residues with potential functional relevance. They provided no evidence to suggest that the overall structural layout of *Pf*KMO will be significantly different from previously seen FAH structures. The C-terminal 60-80 residues of the enzyme appear to be arranged in a new layout though with an unpredictable structure relative to current FAH structures. From the alignments and tertiary structural models it is likely that these residues protrude from the rest of the structure and may form a new small domain. In all the information gained about the sequence of *Pf*KMO will help guide model building and refinement during structure solution - particularly where the X-ray diffraction resolution is ambiguous. In addition it can guide mutagenic studies either towards: i) functional analysis of the role of residues and ii) surface modification for altered crystallization properties.

3.3 Towards PfKMO structure solution

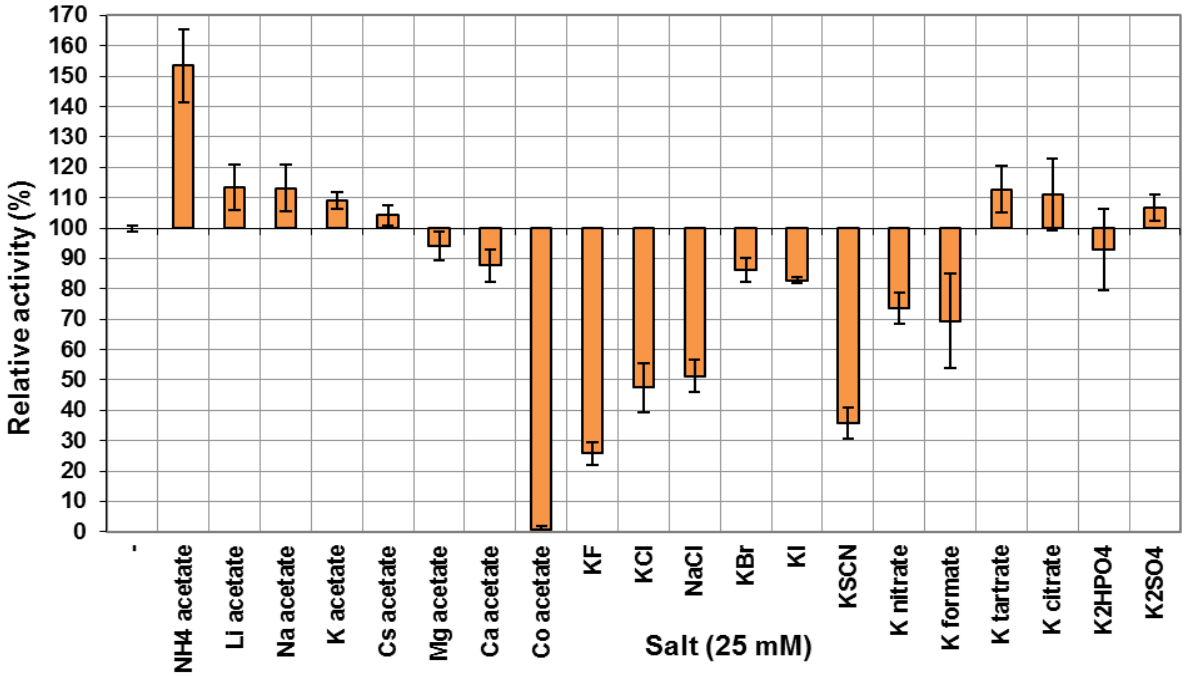
3.3.1 Influence of salts on *Pf*KMO

Numerous KMO studies have detailed an inhibitory effect of salt anions with a potential effect on the flavin orientation. As salts are common additives in protein crystallisation screens common anions and cations were assessed for their effects on *Pf*KMO activity. For this purpose, initial (pseudo-steady-state) *Pf*KMO reaction rates were determined based on the assay described in section **2.5.1** in the presence of excess L-KYN and NADPH. The results, as a percentage of *Pf*KMO activity measured in buffer absent of any added salts (relative activity = 100 x the rate observed in presence of respective salt/average rate observed in absence of salt), are shown in figure 3.7. The screen was repeated three separate times with the average values and their respective standard deviation (the error bars) displayed. The presence of just 25 mM salt in the reaction buffer showed drastic and varied consequences on the enzyme activity. The effect wasn't always negative, surprisingly ammonium acetate increased NADPH consumption by approximately 1.5-fold. Another apparent anomaly among the cations was the complete inhibition caused by cobalt acetate. Apart from these examples, the remaining cations had little effect except a 10-15 % increase in activity seen with the monovalent Li^+, Na^+ & K^+ ions. This could have possibly been due to improvement of *Pf*KMO solubility rather than a direct effect on activity.

Catalysis via flavin-dependent hydroxylases is known to be sensitive to the presence of small monovalent anions and *Pf*KMO does not break the trend. Of those tested fluoride showed the greatest inhibition at around 75 % followed by thiocyanate at 65 % and chloride with around 50 % inhibition. In fact the strength of inhibition appeared to decrease going down group 7 as bromide and iodide showed just 10-20 % inhibition. Nitrate and formate showed 25-30 % inhibition whilst the larger tartrate and citrate ions seemed to have a slightly beneficial effect as seen with acetate salts. The divalent anions had little effect on activity but the slight inhibition seen with phosphate is likely to be due to it competing with the phosphate moiety of NADPH.

In summary the screen particularly highlighted that salts containing Co^{2+}, nitrate, formate or small monovalent anions (such as fluoride and chloride) should be avoided in

the crystallisation of *Pf*KMO. Any ions that influence the dynamics of the enzyme may provide sources of heterogeneity within the crystal lattice. Interestingly bromide and iodide had only slightly inhibitory affects allowing the possibility of their use for anomalous dispersion and phase solution. Ammonium's substantial positive influence on NADPH consumption has provided an interesting route for further study. Ammonium



salts are very common in protein crystallisation but depending on the exact nature of the interaction with *Pf*KMO they may complicate crystallization of the enzyme.

The effect of ammonium and chloride ions was further investigated to determine their effect on the Michaelis-Menten kinetics of substrate binding in *Pf*KMO. The results are summarised below in table 3.3. The normal assay buffer as suggested by Crozier-Reabe & Moran [1, 2] contained 10 mM NaCl and so the assay was run with and without this. The omission of 10 mM NaCl from the assay buffer resulted in a 33 % increase in k_{cat} and around a 50 % decrease in the K_m for L-Kyn. It was clear that even just this small amount of chloride ions hampered the efficiency of the enzyme and so it was removed from all subsequent *Pf*KMO assay buffers. In addition the protein strorage buffer was altered to contain 10 mM Na acetate rather than NaCl to reduce possible interference with the substrate binding site. When 50 mM ammonium acetate was added to the assay buffer,

both the k_{cat} and the K_m significantly increased by 53 % and 183 % respectively. It is unclear how ammonium ions cause a weakened binding of substrate whilst increasing the rate of NADPH oxidation. This could be explained by ammonium ions stimulating the movement of FAD towards the out conformation, thus stimulating its reduction by external NADPH. More extensive analysis of the influence of ammonium ions would be required to provide detailed answers, however for the purpose of this investigation the key message learned was to avoid ammonium ions during the crystallisation of *Pf*KMO.

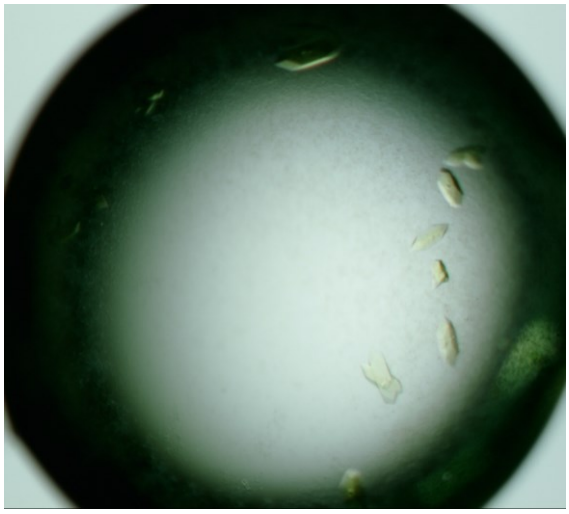
	Unsalted buffer	**+ 10 mM NaCl**	**+ 10 mM NaCl + 50 mM NH_4OAc**
K_m (μM)	14.6 ± 2.2	24.2 ± 1.8	68.4 ± 6.3
k_{cat} (s^{-1})	10.3 ± 0.6	7.5 ± 0.1	11.5 ± 0.4
k_{cat}/K_m ($s^{-1}\mu M^{-1}$)	0.7	0.3	0.2

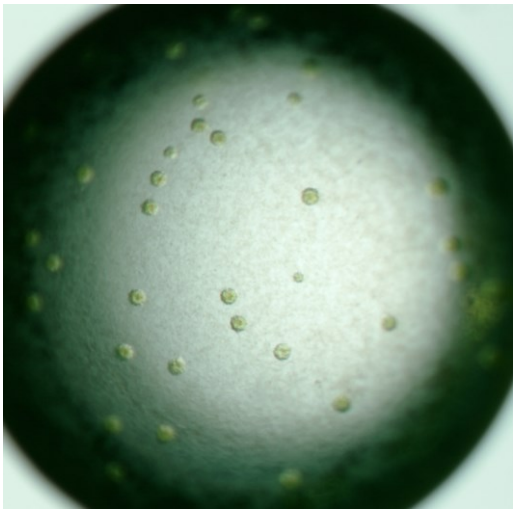
Table 3.3. Summary of the Michaelis-Menten kinetic constants for *Pf*KMO in the presence of NaCl and ammonium acetate.

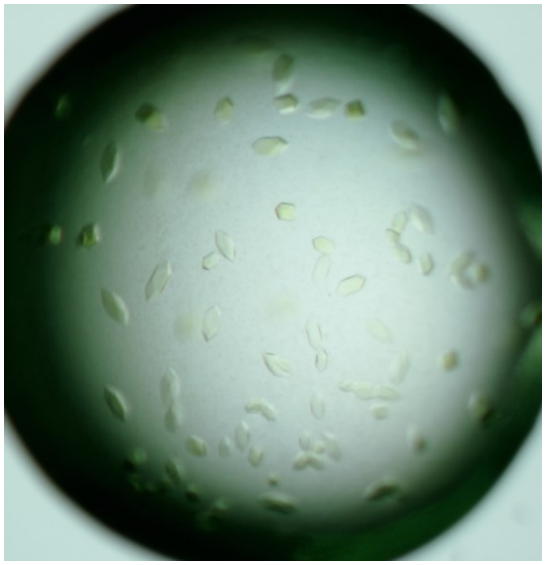
3.3.2 Crystallisation of wt *Pf*KMO

To search for conditions under which *Pf*KMO crystallised, hanging drop vapour diffusion screening with sparse matrix screens was set up using 5 mg/mL pure wild-type (wt) protein. After two days one set of conditions were observed to grow single crystals in the drop containing L-Kyn and clusters of microcrystals in the drop without substrate (figure 3.8). This difference was predicted to be due to decreased conformational flexibility with substrate bound and therefore all further trays were set up using only L-Kyn-bound *Pf*KMO. The well solution was condition 41 from the Hampton Research Crystal Screen 1 kit, which was comprised of 0.1 M HEPES pH 7.5, 20 % (w/v) PEG 4K and 10 % (v/v) 2-propanol. Crystals were reproduced at slightly different pH and concentration of PEG 4K when in-house reagents were used to make the well solution (figure 3.9). Crystals could be grown at both 4 ^{o}C and 17 ^{o}C but were significantly more abundant and robust when at 4 ^{o}C. Single crystals were flash-frozen in liquid nitrogen and tested for diffraction in-house and at the Diamond synchrotron source in Oxford. On the in-house generator only

minimal diffraction to 11 Å resolution was observed, whilst at the synchrotron reflections to around 8 Å were recorded (figure 3.10).







↑ ***Figure 3.8.*** *Pf*KMO 1:1 µl crystallisation drops from h.d.v.d screening with (*left*) and without substrate (*right*). With substrate single crystals formed but without clusters of microcrystals were visible.

← ***Figure 3.9.*** *Pf*KMO substrate-bound crystals reproduced using in-house lab reagents.

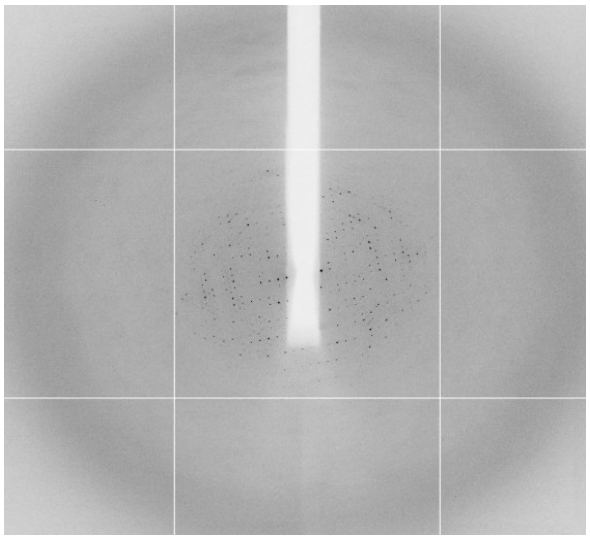


Figure 3.10. Diffraction from an initial substrate-bound *Pf*KMO crystal tested at the Diamond synchrotron beamline I04. Reflections extended to around 8 Å.

The first step towards improving the diffraction resolution limit was to improve protein homogeneity by introducing a gel filtration step at the end of the protein purification protocol. The major monomeric protein peak was collected for further concentration. As a result crystals were reproducibly grown over a much broader range of PEG concentrations and pH, but their diffraction was not substantially improved. The second focus was to try and slow crystal growth to try and improve lattice homogeneity and resulting diffracting power. When crystal trays were set up in the 4 ^{o}C room with all components thoroughly equilibrated, larger crystals were observed to grow. In combination the reservoir solutions were supplemented with numerous viscous additive solutions to try and further slow crystal growth. Addition of up to 10 % (v/v) glycerol and xylitol supported crystal growth whereas ethylene glycol and other PEG variants generated showers of microcrystals. Further additive screens using small compounds and salts were conducted using commercial kits including the 'silver bullets', however no significant improvements were observed. Different buffer pH values were also tested with crystals only growing between pH 6.5-8.5 and the best diffraction consistently observed around pH 7.0. When all of these new factors were combined in trays with different types of PEG precipitant, crystals with stronger diffraction were observed. Interestingly the crystallisation was sensitive to the different PEG variants and crystals could be grown with PEG 200-4000, PEGMME 550 and MOPEG 1500. At this stage larger crystals were being grown with dimensions of approximately 0.2 $\times$ 0.1 $\times$ 0.1 mm^3 and resulting diffraction extending to resolutions around 5-7 Å at the Diamond synchrotron.

While optimisation of crystal growth was on-going the process of crystal cryoprotection was also investigated. Ice rings were observed on the diffraction pattern of crystals grown in the lower PEG concentrations. For an initial test, crystals were moved to new drops of different soaking solutions and left to equilibrate for 24 hours to see if they 'survived'. The soaking solutions tested corresponded to the respective mother liquor plus up to 30 % of commonly used cryoprotectants such as EG, glycerol, MPD etc. Following on from this, different cryoprotectant solutions were designed and tested in soaks of 5 s up to 24 h prior to flash-freezing, mounting and testing for diffraction on the in-house generator. In the end, the results showed that soaking with mother liquor supplemented with 20 % EG provided the best diffraction results with a resolution limit

of around 6 Å in house and no ice rings. Some crystals still showed strong diffraction after 24 hrs but 5s soaks were sufficient. With further refinement the best results were obtained when the cryoprotectant was introduced gradually. This was either achieved by transferring the crystal to a drop containing 10 % and then 20 % EG cryoprotectant or by adding 1 μL of the 20 % EG cryoprotectant to the crystallisation drop before transferring the crystal directly into the 20% EG cryoprotectant. When effective cryoprotection was combined with the earlier improvements to the crystal growth process, crystal diffraction consistently extended to 4-5 Å. From the best crystal a complete dataset was collected to a final resolution of 4.2 Å (figure 3.11). The crystal and all previously diffracting substrate bound crystals belonged to the tetragonal body-centred Bravais lattice group (potentially the space group *I*4) with rough unit cell dimensions of a = b = 149.7, c = 271.6 Å, $\alpha = \beta = \gamma = 90.0^{o}$.

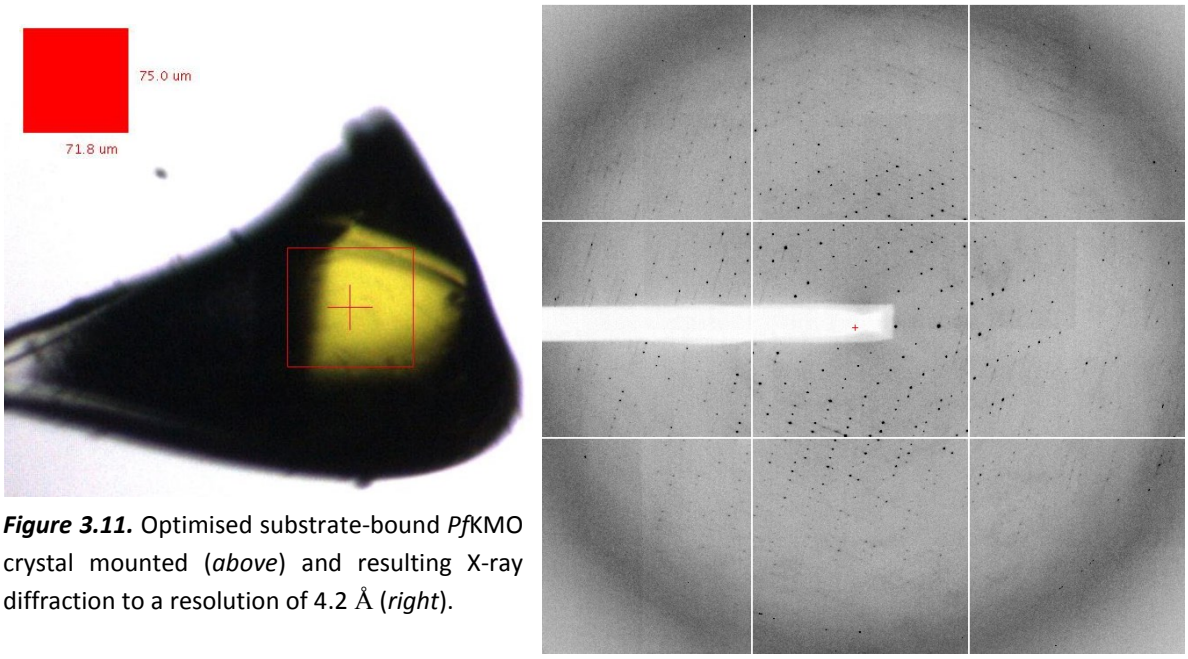


Figure 3.11. Optimised substrate-bound *Pf*KMO crystal mounted (*above*) and resulting X-ray diffraction to a resolution of 4.2 Å (*right*).

At this stage the resolution was not sufficient to be able to solve the protein's structure via molecular replacement and SeMet-substituted wt *Pf*KMO crystals only diffracted to around 5 Å resolution. Post-crystal treatments such as annealing, crystal dehydration and seeding were unsuccessful. Annealing on the loop via omission of the 100 K stream for any length of time destroyed crystals although subtle annealing by re-freezing the

crystal and transferring back to the cryoprotectant solution did not significantly affect diffraction. For dehydration the cover slips containing drops with crystals were moved over solutions with up to 20 % higher concentrations of PEG and left to equilibrate for a couple of days. Mostly the crystals survived but unfortunately their weak diffraction power also remained. Macroseeding with whole crystals/crystal chunks was also unsuccessful, but microseeding did manage to generate new crystals. The resulting 2^{nd} generation crystals were generally smaller than their relatives and again the X-ray diffraction power was not improved. To round-off a negative paragraph, attempts to further slow crystal growth by covering the drops with mineral oil or by adding up to 0.6 % (w/v) low-gelling agarose were unsuccessful. This led to the conclusion that some physical aspect of the enzyme itself may have been prohibiting the formation of an ordered crystal lattice. Therefore the next approach was to investigate the effects of protein engineering on the crystallisation of *Pf*KMO.

3.3.3 Investigation of *Pf*KMO truncation

3.3.3.1 Limited proteolysis

When trying to engineer a protein to favour effective crystallisation there have been numerous different successful approaches, as outlined in section **1.6.3**. The aim is to minimise the surface entropy of the protein to favour the formation of stable crystal contacts that allow an ordered lattice to form. With no prior structural information, trying to predict and mutate individual surface residues would be a time-consuming process and also only a handful of specific surface residues may be involved in forming crystal contacts. Therefore a broader approach was needed and so first limited proteolysis was explored to try to remove any potentially mobile or disordered loops/domains. Pure *Pf*KMO was mixed with a small concentration of bovine trypsin at a ratio of 50:1 and incubated at 25 ^{o}C. Prior to trypsin addition and then after specific time points an aliquot of the reaction was removed on to ice with PMSF (dissolved in ethanol) added to 2 mM final concentration. Each aliquot was immediately frozen and then all were thawed together and subjected to SDS-PAGE analysis in comparison with the undigested sample. The proteolysis experiment was run both without L-Kyn and in the

presence of 1 mM L-Kyn with the resulting gels displayed in figure 3.12. Firstly the presence of L-Kyn had no effect on the rate of digestion of *Pf*KMO. Interestingly two intermediates were captured at rough sizes of 46-48 KDa and 43-45 KDa compared to the 50.7 kDa native protein. The first was more transient, appearing after 30-60 mins and then mostly disappearing 30 mins later. The fact that it mostly disappeared when most of the full length protein still remains suggests that this intermediate may have been a digestion product of a contaminant originally hidden under the full length *Pf*KMO gel band. The second intermediate, however, was more stable as it started forming after around 60 mins and continued to accumulate 4 h later. After 24 hour exposure there was no full length protein visible but a small amount of the second intermediate remained. The next step was then to characterise the stable intermediate and determine whether it was worth attempting to crystallise it.

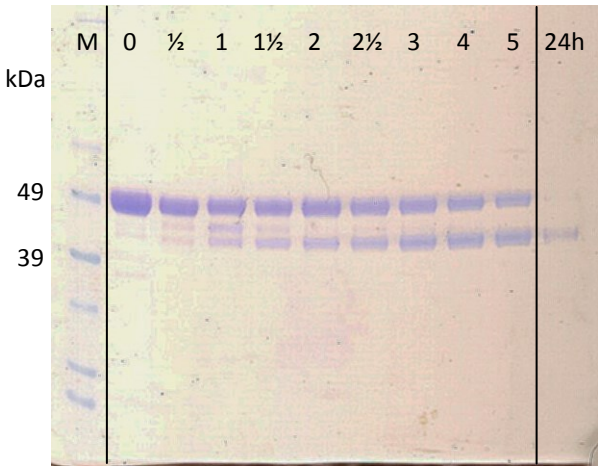


Figure 3.12. NuPAGE analysis of samples from limited proteolysis reaction stopped at various time points (hours). The uncut full length *Pf*KMO is shown in the zero column.

To try and isolate the stable intermediate for crystallography trials, a bigger volume of protein was digested and stopped as before after 4 hours of the reaction. This was then concentrated and subjected to size-exclusion chromatography on a 280 mL s-75 column to try and separate the fragment from the full size protein. Promisingly, the intermediate was stable throughout the chromatography and could be separated to some extent from the full size protein as shown in the protein gel in figure 3.13. This separation could have been further optimised by running the column more slowly etc. Unfortunately this was not necessary as it was clear that the isolated intermediate did not contain any FAD and

was inactive. The fraction was concentrated as much as possible to 3 mg/mL, soaked with FAD and used for brief crystallography trials but with no success.

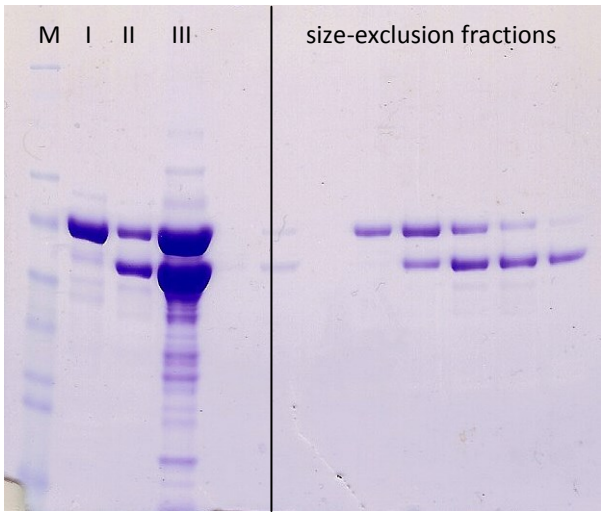


Figure 3.13. NuPAGE analysis of size-exclusion chromatography for digested *Pf*KMO.
M = Marker
I = pure *Pf*KMO
II = digested sample
III = concentrated digest
remaining lanes correspond to fractions eluted from the size exclusion column.

MALDI-TOF was carried out on excised and purified gel slices of both the undigested and stable intermediate *Pf*KMO bands to try and identify what section of protein had been removed. As part of the protocol both forms were completely digested with trypsin prior to fragment analysis. Each sample was run four times with an example peak spectrum from each form shown in figure 3.14 for the full protein and for the stable intermediate. The mass/charge of detected fragments was automatically correlated to predicted trypsin cleavage products from the *Pf*KMO protein sequence with the results for each protein sample summarised in figure 3.15.

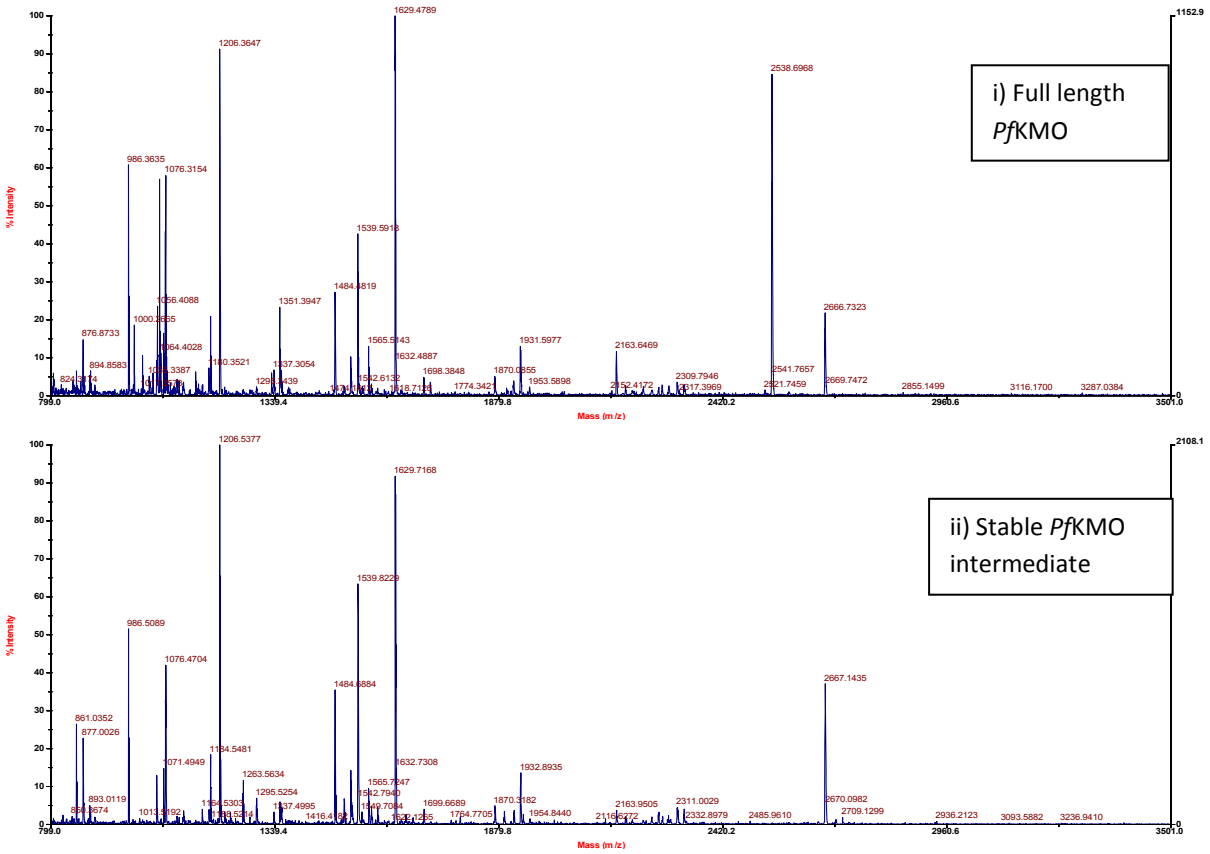


Figure 3.14. MALDI-TOF fragment analysis displaying mass/charge peaks from i) full length undigested *Pf*KMO and ii) the purified stable digest intermediate.

```
Full length    MTATDNARQVTIIGAGLAGTLVARLLARNGWQVNLFERRPDPRIETGARGRSINLALAER   60
Intermediate   MTATDNARQVTIIGAGLAGTLVARLLARNGWQVNLFERRPDPRIETGARGRSINLALAER   60

Full length    GAHALRLAGLEREVLAEAVMMRGRMVHVPGTPPNLQPYGRDDSEVIWSINRDRLNRILLD   120
Intermediate   GAHALRLAGLEREVLAEAVMMRGRMVHVPGTPPNLQPYGRDDSEVIWSINRDRLNRILLD   120

Full length    GAEAAGASIHFNLGLDSVDFARQRLTLSNVSGERLEKRFHLLIGADGCNSAVRQAMASVV   180
Intermediate   GAEAAGASIHFNLGLDSVDFARQRLTLSNVSGERLEKRFHLLIGADGCNSAVRQAMASVV   180

Full length    DLGEHLETQPHGYKELQITPEASAQFNLEPNALHIWPHGDYMCIALPNLDRSFTVTLFLH   240
Intermediate   DLGEHLETQPHGYKELQITPEASAQFNLEPNALHIWPHGDYMCIALPNLDRSFTVTLFLH   240

Full length    HQSPAAQPASPCFAQLVDGHAARRFFQRQFPDLSPMLDSLEQDFEHHPTGKLATLRLTTW   300
Intermediate   HQSPAAQPASPCFAQLVDGHAARRFFQRQFPDLSPMLDSLEQDFEHHPTGKLATLRLTTW   300

Full length    HVGGQAVLLGDAAHPMVPFHGQGMNCALEDAVALAEHLQSAADNASALAAFTAQRQPDAL   360
Intermediate   HVGGQAVLLGDAAHPMVPFHGQGMNCALEDAVALAEHLQSAADNASALAAFTAQRQPDAL   360

Full length    AIQAMALENYVEMSSKVASPTYLLERELGQIMAQRQPTRFIPRYSMVTFSRLPYAQAMAR   420
Intermediate   AIQAMALENYVEMSSKVASPTYLLERELGQIMAQRQPTRFIPRYSMVTFSRLPYAQAMAR   420

Full length    GQIQEQLLKFAVANHSDLTSINLDAVEHEVTRCLPPLSHLC   461
Intermediate   GQIQEQLLKFAVANHSDLTSINLDAVEHEVTRCLPPLSHLC   461
```

Figure 3.15. Identification of MALDI-TOF mass spec fragments of full length and the stable limited proteolysis intermediate digested by trypsin. Sequence in grey represents fragments too small or too large to be detected in the range of mass/charge values collected. Sequence in green represents the fragments successfully identified for each form of the protein.

The majority but not all of the possible fragments were detected for the full length sequence. It was clear that the residues removed by limited proteolysis related to the C-terminus as the N-terminal residues were detected in both protein forms. The peaks relating to the C-terminal segments located after K376 that were present in the full length protein, were absent in the stable limited proteolysis intermediate. Removal of the sequence after this residue would leave a protein core of 41 kDa, which roughly correlates to the change in position seen during gel electrophoresis. These data suggested that K376 is solvent exposed to be able to be cleaved during limited proteolysis, possibly located on a surface loop or turn. Importantly, removal of the C-terminal residues after this point was detrimental to the function of the enzyme due to the loss of flavin. In addition to this, C-terminal *Pf*KMO truncations after positions 409, 434 and 451 were engineered and found not to express any detectably active protein. From the tertiary structure models and prior sequence analysis it is unlikely that this sequence region is directly involved in flavin binding, more likely is that its removal destabilises the overall structure of the enzyme.

3.3.4 Cysteine mutagenesis

Cysteines are unique residues that can form covalent disulphide bridges with each other. Such bonds have proved essential in maintaining the internal structure of many proteins but when free cysteines are exposed on the surface of proteins they can cause unwanted disulphide-linked aggregation. *Pf*KMO contains six cysteine residues and has been shown to lose activity rapidly over time in correlation with the oxidation of two unidentified cysteines [2]. As mentioned in section **1.6.3**, the substitution of surface cysteines for serines to prevent disulphide-based aggregation has improved the crystallisation success for several proteins. Looking at the information from sequence alignments and homology modelling (Sections **3.2.2** & **3.2.4**) there was significant evidence that two of the cysteines in *Pf*KMO, C223 and C326 were likely to be buried. This left cysteines C168, C252 and the C-terminal C453 and C461 as potential surface targets. These four residues were separately substituted by serines and each construct was expressed in BL21(DE3) *E. coli* cells. The respective Michaelis-Menten properties and stabilities of each pure

mutant were tested in relation to wt enzyme as summarised in table 3.4. The single mutant proteins were purified and tested by the MChem project student Menglu Wang.

Protein	K_m (μM)	k_{cat} (s^{-1})	$t_{1/2}$ - RA (hours)	$t_{1/2}$ + RA (hours)
wt	14.5 ± 2.5	9.1 ± 1.6	7.9 ± 2.6	26.2 ± 2.0
***C168S**	21.6 ± 5.1	10.7 ± 0.2	25.0 ± 0.2	45.2 ± 16.3
***C252S**	18.7 ± 5.9	6.8 ± 0.6	53.5 ± 24.3	89.2 ± 7.9
***C453S**	19.9 ± 5.5	9.0 ± 0.8	29.4 ± 1.7	42.2 ± 8.7
***C461S**	15.5 ± 4.2	6.4 ± 0.5	31.5 ±2.3	47.3 ± 6.6

Table 3.4. Summary of Michaelis-Menten kinetics and enzyme half life ($t_{1/2}$) for wt and single cysteine mutant forms of *Pf*KMO. $t_{1/2}$ refers to the length of incubation at 25 ^{o}C required for half of the initial enzyme activity to be lost. RA indicates presence of reducing agent (1 mM DTT).
*All of the single mutant data was determined by the MChem student Menglu Wang.

The results showed that all four of the single mutations reduced the loss of activity over time, with the C252S mutation significantly providing the most protective effect. In the absence of reducing agent C252S had a half-life of around 54 hours – 7× greater than the 8 hours seen with the wt protein. The presence of reducing agent boosted the half-life of all the proteins by around a day. In addition to this, two double mutants were created – dm1 (C453S/C461S) and dm2 (C252S/C461S). The double mutants were also tested for stability. The measured $t_{1/2}$ values in the absence of reducing agent are compared to wt and the single mutants in figure 3.16. The half-life of dm1 was level to the values around 30 hours seen with the single mutants C453S and C461S. For dm2 however, the half-life was significantly improved further to 75.2 ± 5.2 h. With the high error associated with the data for the C252S mutant, this strongly confirmed the implication of this residue in destructive aggregation of wt *Pf*KMO. The Michaelis-Menten data for dm2 yielded a K_m of 8.8 ± 1.5 μM and a k_{cat} of 8.9 ± 1.1 s^{-1}. The kinetic data confirmed that all of the cysteine mutations had no significant effect on the enzyme's catalytic properties.

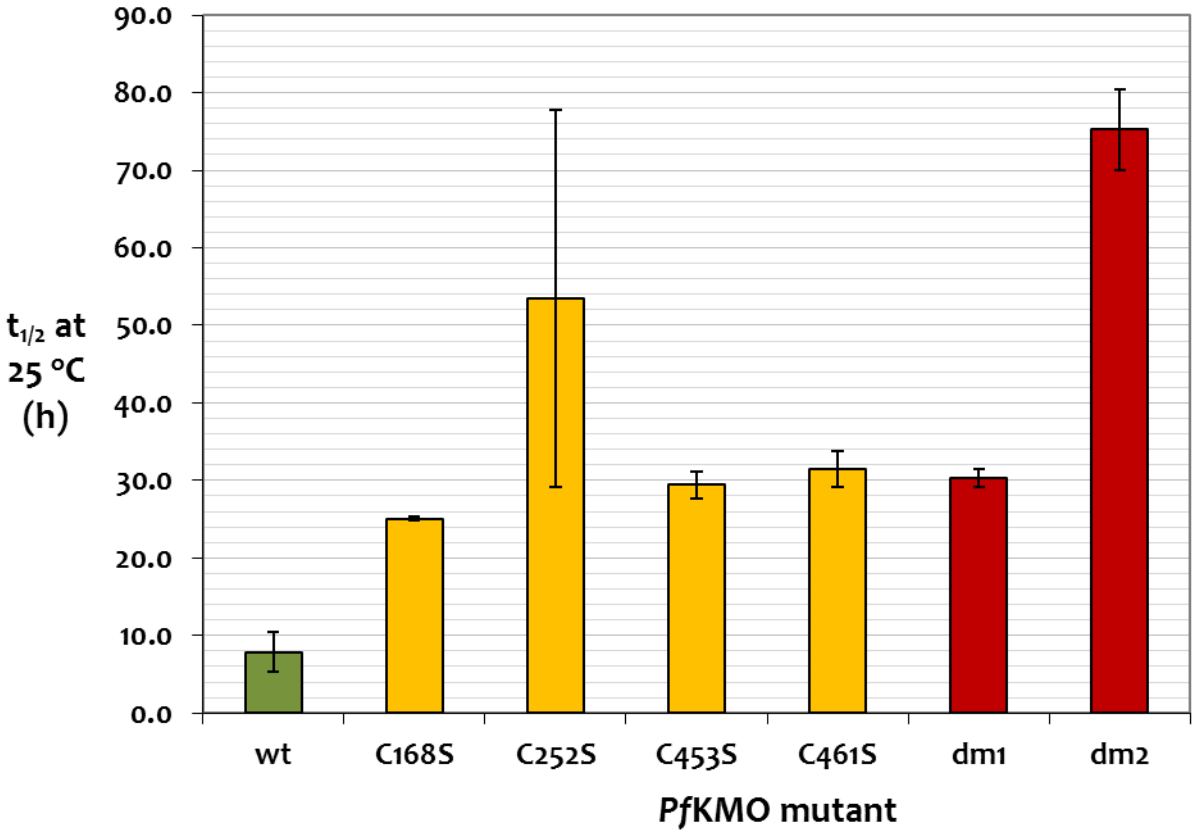
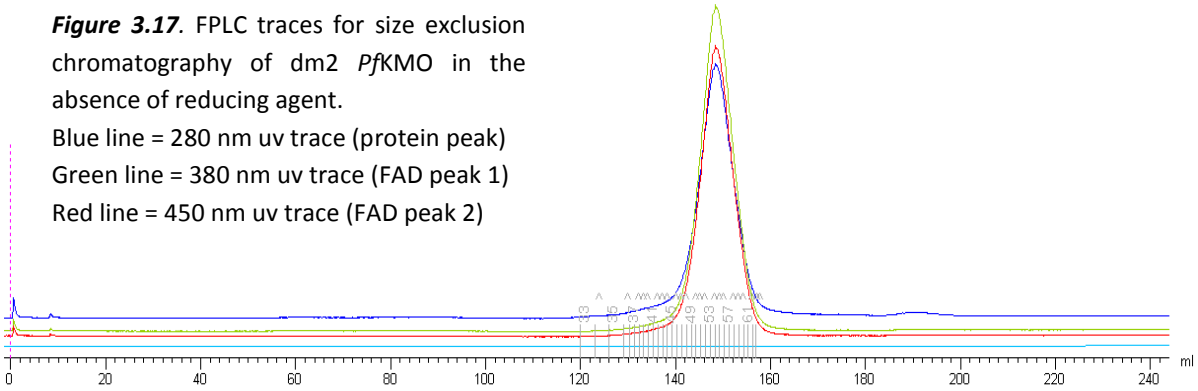


Figure 3.16. Graphical representation of the activity half-life ($t_{1/2}$) for each mutant when incubated at 25 ^{o}C in the absence of reducing agent. Made using Excel.

The dm2 protein was subsequently purified on a large-scale and used for crystallisation trials. It was noted that roughly 3-4× greater yield was seen compared to with wt. The protein ran as a single monomeric peak in size-exclusion chromatography even with the exclusion of reducing agent (see figure 3.17). Crystals could be grown from similar conditions to the wt protein but with higher concentrations of dm2 required. A search with the sparse-matrix crystallisation screens unfortunately failed to find any new leads for alternative conditions of crystal-growth. Crystallisation via h.d.v.d with 1 µl of 12.5 mg/ml dm2 (20 mM HEPES pH 6.8, 30 mM Na acetate, 1 mM DTT and 1 mM L-Kyn) mixed with 1 µl of reservoir solution reproducibly and more readily yielded larger crystals than with wt protein. The size of crystals reached up to 0.3 × 0.3 × 0.6 mm^3. After cryoprotection, crystals tested at the Diamond synchrotron readily diffracted below 4.0 Å resolution. A couple of the mounted crystals tested are shown in figure 3.18. The larger crystals could survive and in fact required the maximum available beam intensity to maximise the resolution of the reflections recorded. The best crystal

diffracted to give a 100.0 % complete dataset to a resolution of 3.35 Å with an R_{merge} of 0.085 (0.676 in the highest resolution shell - HS) and a I/σI of 15.8 (3.0 HS). This was achieved from 650 images with a rotation oscillation of 0.35 ° and exposure of 0.35 s/image. The crystal was still tetragonal ($I4_122$) with cell dimensions of a = b = 149.5, c = 271.9 Å and α = β = γ = 90.0 ° and grew from a mother liquor of 0.08 M HEPES pH 7.0, 6 % PEG 4K, 8 % glycerol, 12 mM NaCl, 10 % 2-propanol.

Figure 3.17. FPLC traces for size exclusion chromatography of dm2 *Pf*KMO in the absence of reducing agent.
Blue line = 280 nm uv trace (protein peak)
Green line = 380 nm uv trace (FAD peak 1)
Red line = 450 nm uv trace (FAD peak 2)



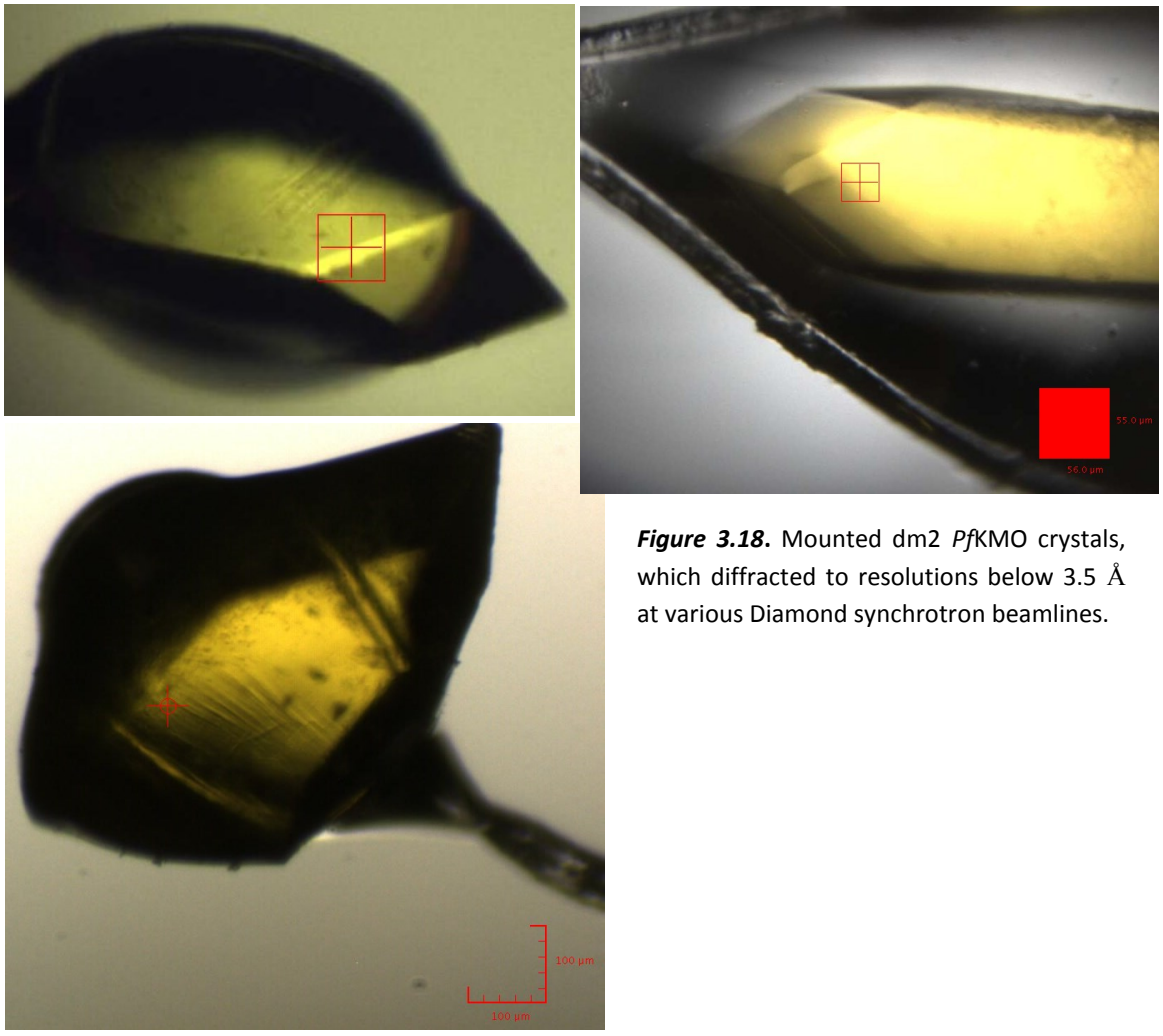


Figure 3.18. Mounted dm2 *Pf*KMO crystals, which diffracted to resolutions below 3.5 Å at various Diamond synchrotron beamlines.

3.3.5 Loop replacement for shorter homologous sequence

As mentioned in section **3.2.1**, there is a short six-residue segment of sequence in *Pf*KMO that is not present in other KMOs. Considering the small size and the conservation seen either side of this sequence it is probable that this corresponds to a surface loop. A section of a sequence alignment involving the closest evolutionary relatives to *Pf*KMO is displayed in figure 3.19. There is a shorter consensus sequence present in the homologous sequences of PNXGX (where X = any residue) that aligns to the residues HHQSPAAQPAS at positions 240-250 in *Pf*KMO. Interestingly there is a KMO from *Pseudoxanthomonas suwonensis* with a sequence similarity somewhere between *Pf* and the closely related *Xanthomonas* species. This intermediate species had a mixed sequence at the aforementioned position of PNQAAPGSGD. This presented a great opportunity to transfer a smaller sequence like those in the *Xanthomonas* species in to replace the larger sequence found in *Pf*KMO in the hope of minimizing the flexibility of this potential loop. Therefore the gene sequence corresponding to the eleven aa residues 240-250 in *Pf*KMO were replaced with sequence coding for the residues PNQGD by overlap extension PCR (see section **2.1.4**). The mutant construct created (LOOPr) also contained the C252S/C461S double mutation (see section **3.2.4**).

```
Sten.malt_KMO    237 LFLPNQ------GDPSFATVNTGAQAEALFAREFADTLPLIPDLRADWEQHPPGLLGTLT 293
Stre.spec_KMO    229 LFLPNQ------GDPSFATVNTGAQAEALFAREFADTLPLIPNLRADWEQHPPGLLGTLT 285
Xant.axon_KMO    237 LFLPNA------GEPSFATTRTGDEAFALFARDFPDALPLIPQLKEHWEEHPPGLLGTLT 293
Xant.oryz_KMO    237 LFLPNA------GEPSFATTRNGDEAFALFARDFPDALPLIPQLKQHWEEHPPGLLGTLT 293
Xant.gard_KMO    237 LFLPNT------GEPSFATTRTGDEALALFARDFPDALPLIPQLKEHWEEHPPGLLGTLT 293
Xant.vesi_KMO    237 LFLPNA------GEPSFATTRTGEEAHALFARDFPDALPLIPQLEEHWEEHPPGLLGTLT 293
Xant.Camp_KMO    238 LFLPNE------GMPSFATTRSGDEALALFARDFPDALPLIPQLKEHWEEHPPGLLGTLT 294
Xant.albi_KMO    236 LFLPNK------GEPSFASIRNGKEAVALFARDFADVLPLMPQLAEHWEQHPPGLLGTLC 292
Psxa.suwo_KMO    237 LFLPNQA-APGSGDPCFASVRSGAEVRALFARDFPDALPLIPDLESDWENHPPGLLGTLY 298
Pseu.fluo_KMO    237 LFLHHQSPAAQPASPCFAQLVDGHAARRFFQRQFPDLSPMLDSLEQDFEHHPTGKLATLR 299
                     *** :       . *.**    *  .  :* *:*.*  *:: .*  .:*.**.* *.**
```

Figure 3.19. Section from a sequence alignment of *Pf*KMO with its closest evolutionary homologs showing the potential loop insertion highlighted in red. The aligned bacterial species in full were: *Stenotrophomonas maltophilia*, *Streptococcus species*, *Xanthomonas axonopodis, Xanthomonas oryzae, Xanthomonas gardneri, Xanthomonas vesicatoria, Xanthomonas campestris, Xanthomonas albilineans, Pseudoxanthomonas suwonensis* and *Pseudomonas fluorescens*

LOOPr protein was produced and purified from the altered construct and comparable activity was seen in the crude cell extracts compared to expressed dm2 *Pf*KMO. The protein was purified as normal (figure 3.20) with a final step of size-exclusion

chromatography, where the mutant ran again as a single monomeric peak. The expected difference in size was too small to be observed in the gel filtration and on the NuPAGE gel. LOOPr had K_m and k_{cat} values of 12.5 ± 3.7 μM and 8.3 ± 0.4 s^{-1} respectively; similar to those observed previously with wt and dm2 proteins. Interestingly though, in limited proteolysis the protein digested much more quickly than wt enzyme as most of the full length protein band had disappeared after just 30 mins (figure 3.21). The stable intermediate still formed but again had a reduced lifespan and was fully degraded between 3-4 hrs. Following this, crystallisation trials of LOOPr (10 mg/mL plus 1 mM L-Kyn) using the commercial sparse-matrix screens failed to yield any crystals. Whilst the protein's catalytic activity was unaffected by the replacement, it appeared that local external changes may have perturbed the enzyme's resistance to trypsin digestion and ability to crystallise in the previously encountered form.

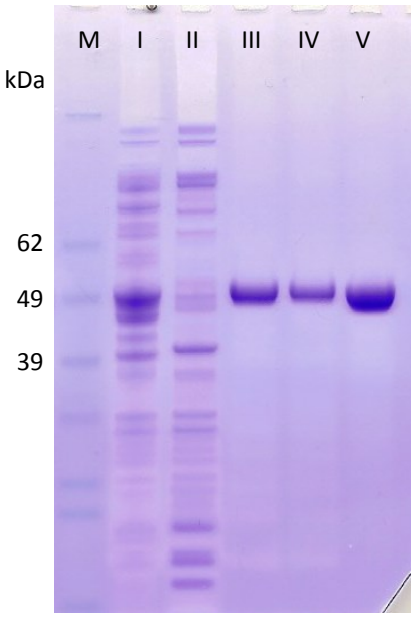


Figure 3.20. NuPAGE gel of LOOPr Purification.
I = sonication supernatant
II = Q-sepharose flow-through
III = Q-sepharose elute
IV = Ammonium sulphate pellet
V = size-exclusion pooled fractions

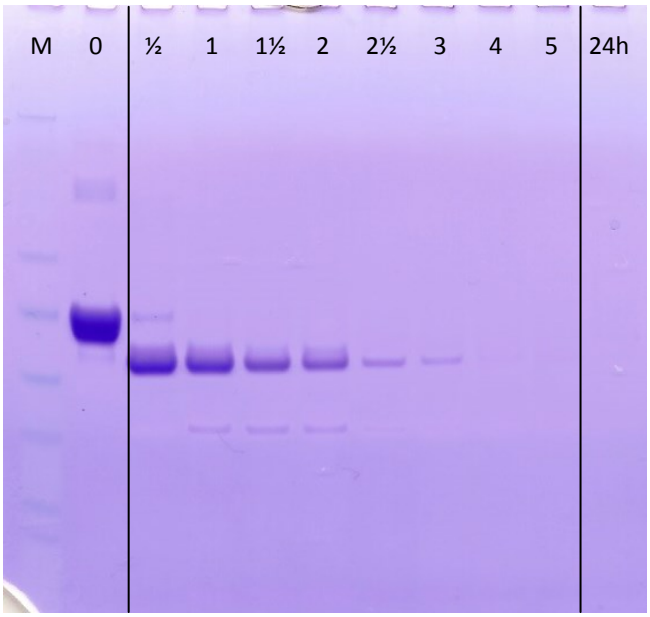


Figure 3.21. NuPAGE gel of LOOPr limited proteolysis samples removed after respective time intervals (hours).

3.3.6 Conclusions and summary

Before any experimental work was undertaken, a lot of significant information was obtained from the *Pf*KMO primary amino acid sequence alone. Evolutionary information, in combination with secondary and tertiary structure predictions, was analysed. The results of the secondary structure predictions in particular were largely consistent but struggled with certain potential sheets/loops. Predicted tertiary structural models were biased towards the respective FAH template structure used, however they may be relevant in this case due to the global fold relationship in all currently known FAH structures. In all, comparisons with other FAHs provided strong evidence that the first ~ 370 residues of *Pf*KMO were structurally related to the normal two domain core shape. The C-terminal 90 residues were unpredictable; these being additional compared to some FAH enzymes, and distinctly unrelated to the C-terminal domains found on the others. Looking at alignments of closely related KMOs there was clearly enough conservation of these residues to suggest that they have a stable structural arrangement. The conservation broke down when considering more evolutionary distant KMO sequences suggesting that the structure of the C-terminus has changed over time. This may be of particular relevance when trying to apply any *Pf*KMO structural information to other KMOs such as human KMO. Excluding the C-terminus, candidates for residues universally involved in substrate binding as well as those potentially on the enzyme's surface have been discovered. In all the only major obvious concern regarding the potential crystallisation success of the enzyme was the unknown nature of the C-terminal residues.

Since the early elation of discovering conditions that could grow *Pf*KMO crystals, a lot of work went in to trying to discern a productive direction to further the route towards structure determination. The diffraction limit of wt substrate bound crystals was significantly improved from around 8 Å to ~4 Å by optimising the methods of crystal growth and cryoprotection. At this point a wall was hit meaning that either new growth conditions and/or altered protein properties were needed. Screening for new conditions failed to find any fresh leads and so protein engineering was explored to try to alter/improve crystal packing. From the varied approaches tried the only success was encountered with cysteine mutagenesis. In particular the C252S/C461S double mutant

was shown to be significantly more stable than wt *Pf*KMO and more importantly the resulting crystals diffracted more strongly to give reflections below 3.50 Å. The resolution had now reached a stage where work towards the structure solution could begin, and this is described in the next chapter. Along the way though interesting insights have been discovered and dwelled upon below.

Throughout the crystallisation work so far a clear trend had emerged – larger substrate bound crystals give stronger diffraction. This in itself was interesting because if the problem with *Pf*KMO crystals was related to the malformation of a stable homogenous lattice than why would larger crystals help? Two possible theories could cover this: i) that the larger crystals had grown more slowly and so were able to form more homogeneous and stable lattices or ii) that the crystals were composed of significantly more solvent than protein meaning that larger crystals gave a greater amount of protein and thus diffraction power. Tests of dehydration and annealing which would have reduced the solvent content had no influence on the diffraction power but also neither did seeding which would have promoted slow controlled crystal growth. Either way the protein was clearly not forming stable crystal contacts in this particular lattice arrangement and so protein manipulation was needed.

The work on limited proteolysis and C-terminal truncation was itself limited by the loss of flavin cofactor and of activity of the enzyme. The removal of the C-terminus when forming the stable digestion intermediate may suggest that the protein had a similar arrangement of residues 1-376 to the two-domain core of other FAH enzymes. The C-terminal ninety residues that follow appear to form a removable extension that can leave a stable core. This core though was inactive and unoccupied by flavin so it must be inferred that the C-terminal residues were involved in maintaining functional, FAD-bound enzyme. With the location of the start of this region relative to the structures of other FAH enzymes, it appears unlikely that these residues could directly interact with the flavin. Therefore a reasonable conclusion is that the C-terminal residues may be largely exposed but still interact with and stabilise the main protein fold. Their removal could cause subtle structural rearrangements that perturb the flavin binding site without completely unfolding the protein which would then expose the remaining core to rapid trypsin digestion. Shortening of the potential loop at residues 240-250 had no effect on

the catalytic properties of the enzyme suggesting that this region may be on/near the protein surface as predicted in the homology models. The mutant was however more susceptible to trypsin digestion, implying some extent of local structural change to create greater exposure of basic residues to the trypsin. This loop may therefore as a result offer some shielding of K376 from exposure to the active site of trypsin. It was unclear though whether the proteolytic changes were a result of just a shorter loop at this position or whether the replacement disrupted local secondary structure and subsequently enhanced exposure of basic residues. Whilst some potentially useful clues were discovered here, no potential route towards better crystal packing was unearthed and true conclusions would only be reached in combination with structural data.

This section presented a lot of computational analysis and predicted inferences that may only find significant relevance during structure solution, particularly considering the poor resolutions currently attained. With the crystal X-ray diffraction resolution now below 3.5 Å the route towards structure solution could begin. Model building by molecular replacement would be challenging at such resolutions when taking into account the low sequence identities to known structures. Therefore L-selenomethionine (SeMet) labelled dm2 *Pf*KMO was generated and crystallised to directly calculate the phases and guide model building and refinement.

3.4 Crystal structure of selenomethionine-labelled substrate bound PfKMO

3.4.1 Crystal growth and data collection

SeMet-labelled dm2 *Pf*KMO was expressed in the methionine auxotroph B834 *E.coli* strain and the protein was purified as usual but in the presence of elevated levels of the reducing agent DTT. Size-exclusion chromatography was omitted from the purification to reduce the time of SeMet exposure to oxygen before crystallisation. The features of the final absorbance spectrum of the purified enzyme resembled wt *Pf*KMO (displayed in figure 3.22). Moreover, it was confirmed that the SeMet-labelling also did not affect the Michaelis-Menten kinetics of the enzyme. The K_m and k_{cat} values were determined to be

7.0 ± 0.7 μM and 8.5 ± 0.2 s^{-1} respectively compared to 8.8 ± 1.5 μM and 8.9 ± 1.1 s^{-1} for unlabelled dm2.

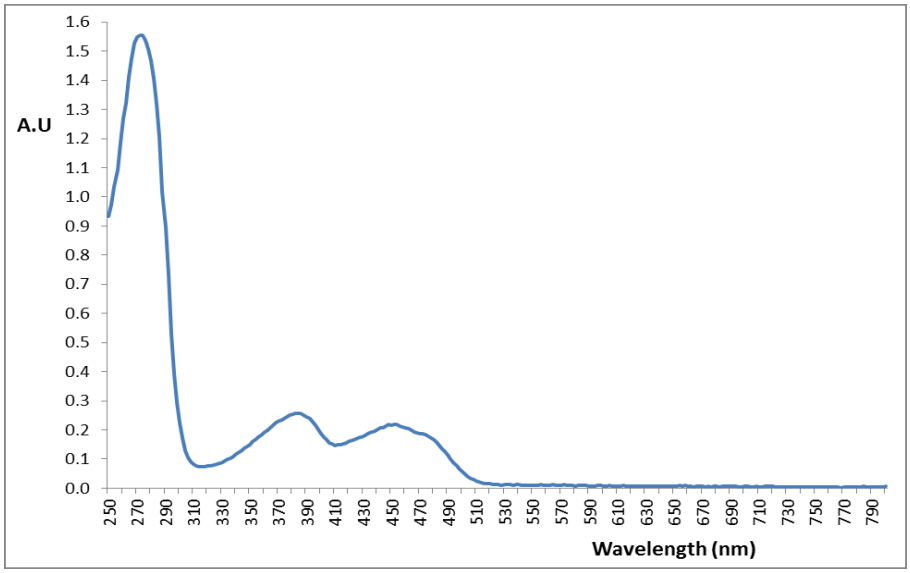


Figure 3.22. Absorbance spectrum for pure SeMet dm2 *Pf*KMO protein.

Crystals of the SeMet labelled protein were grown by h.d.v.d with hetero-microseeding. The drop contained 1.0 μl of protein solution (15 mg/ml in 20 mM HEPES pH 6.8, 20 mM K acetate, 1 mM L-Kyn, 2 mM DTT), 1.0 μl of the precipitant (0.1 M HEPES pH 7.0, 15 % w/v PEG 4000, 10 % v/v glycerol, 10 % v/v 2-propanol, 2 mM DTT & 1 mM L-Kyn) and 0.5 μl of hetero-seed solution. To prepare the seed solution, a native substrate bound dm2 crystal was crushed, transferred to 100 μl of stabilizing solution (0.1 M HEPES pH 7.0, 20 % v/v PEGMME 550, 10 % v/v 2-propanol) and micro-centrifuged. Chunky crystals resembling an octagonal prism but cut in half (see figure 3.23) with rough dimensions of 0.2 x 0.2 x 0.2 mm^3 grew in 3-5 days and were harvested for data collection after three weeks. Prior to flash-freezing in liquid nitrogen, 1 μl of 100 % (v/v) ethylene glycol was added directly to the drop and left to equilibrate for an hour.

The crystals were screened for diffraction quality at the Diamond synchrotron beamline I04 with several crystals diffracting to a resolution below 4.0 Å. Multi-wavelength anomalous dispersion (MAD) X-ray diffraction data were recorded with each dataset collected from a different area of a single crystal. The mounted crystal is shown in figure 3.23. Following a fluorescence scan each dataset was collected using X-ray wavelengths

of 0.97950 (Peak), 0.97965 (Inflection), 0.96860 (High Energy Remote) & 0.98730 Å (Low Energy Remote). Using an exposure of 0.5 s/frame of data, 220 frames were recorded with rotation increments of 0.5 º to yield complete data sets. The crystal was tetrahedral, space group $I4_122$, with averaged unit cell parameters of a = b = 150.0 Å, c = 273.4 Å, α = β = γ = 90.0 º. Indexing and integration of reflections was carried out in Mosflm before scaling in Scala with anomalous pairs separated. The four datasets were 98.8-99.9 % complete to a resolution of 3.40 Å with respective R_{merge} = 0.140, 0.137, 0.137 & 0.113 (highest resolution shell = 0.580-0.680). The full data collection statistics are tabulated in appendix section **3.1**. The phases were calculated using the Phenix program suite, which was able to locate all but the N-terminal SeMet residue.

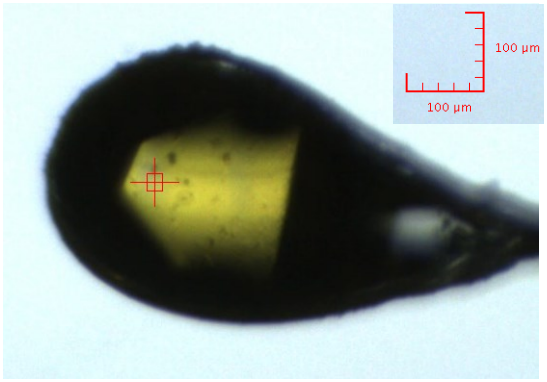


Figure 3.23. Mounted substrate-bound SeMet dm2 *Pf*KMO crystal used for four-wavelength MAD data collection. The scale is shown in the top-right corner.

3.4.2 Structure solution and refinement

A homology model of *Pf*KMO containing residues 7-380 produced by the Phyre server from a template structure of PhzS monooxygenase (pdb code 3C96) was selected to input as a starting scaffold to fit the phase-solved electron density map generated from Phenix. The cell displayed a high solvent content of 83.8 % with just one protein molecule in the asymmetric unit, which explained the need to grow large crystals to strengthen X-ray diffraction. The model was exhaustively rebuilt and extended by numerous rounds of manual editing in Coot [237] and restrained refinement with Refmac5 [255]. Guided by the positions of the thirteen SeMet residues and information from sequence and structural alignments, a suitable model was obtained containing residues 7-461 (figure 3.24). This revealed clear unfilled density for one molecule each of

non-covalently bound FAD and substrate L-Kyn to be included. The density for the substrate (figure 3.25) was slightly ambiguous but the final placement showed significantly the best geometric and structural fit to the data. Strong density was evident for the placement of the carboxyl tail but at this resolution density for the 2-aminobenzene ring was weak. The final model was refined with the addition of TLS to give R and R_{free} values of 0.248 and 0.281 respectively with a figure of merit of 0.854 and overall mean B factor of 75.7 $Å^2$ for 20381 significant reflections in the range of 77.86-3.40 Å. The root means square deviation (rmsd) of bond lengths and angles from the ideal are 0.005 Å and 0.95 ^{o} and all the residues are present in the most favoured or additionally allowed regions of a Ramachandran plot as determined using the Molprobity server [256]. Full refinement statistics are tabulated in appendix section **3.1**.

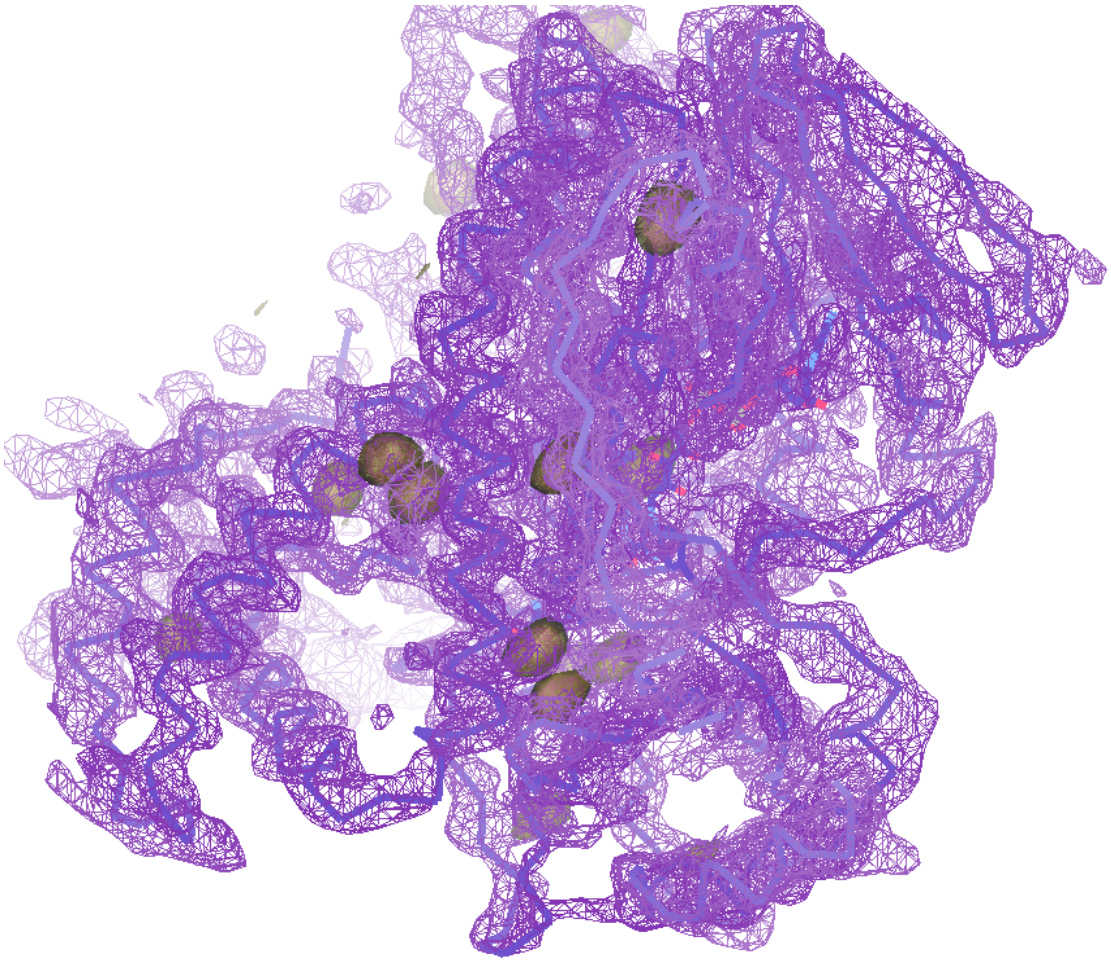


Figure 3.24. Representation of the fit of the final substrate-bound model to the refined $2F_o$-F_c map in purple (contoured at 1.8 σ) with anomalous difference map peaks – relating to the selenium positions – shown as gold blobs (contoured at 4.0 σ). Created in Coot.

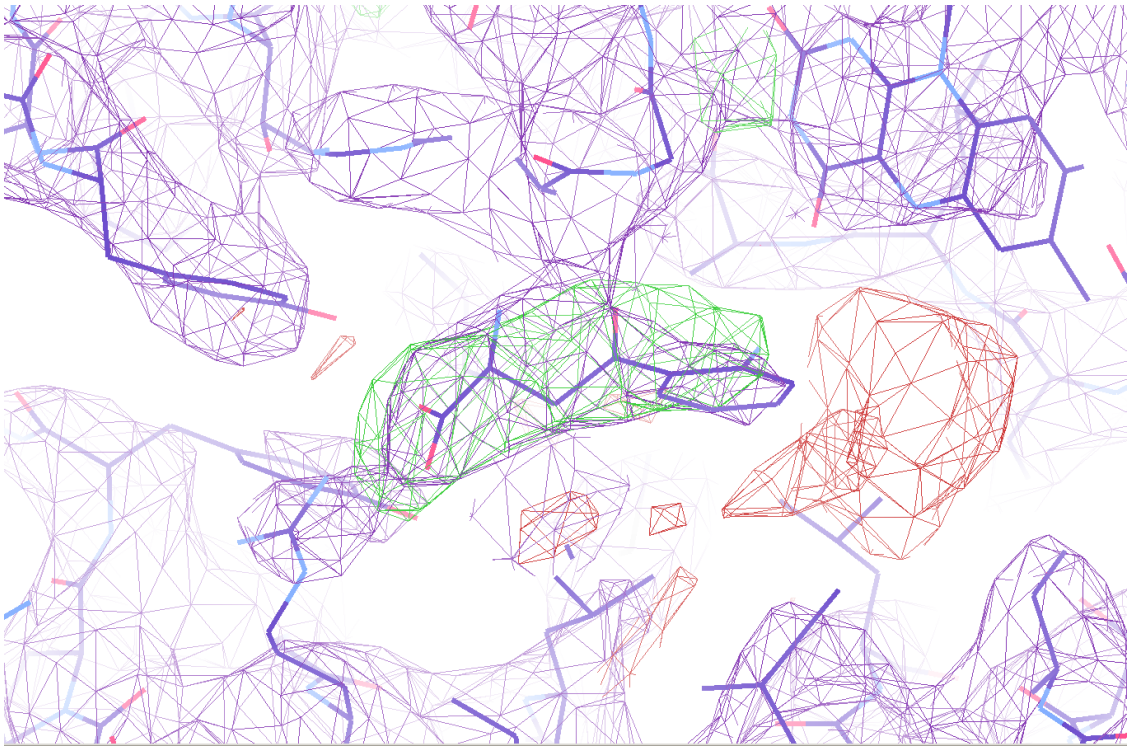


Figure 3.25. Diagram of the substrate-binding pocket with the refined $2F_o$-F_c map in purple (contoured at 1.5 σ) with the F_o-F_c map, calculated in the absence of L-Kyn, shown in green/red (contoured at 3.5 σ). Figure created in Coot.

3.4.3 Overall structure of *Pf*KMO

The protein can be split into three structural domains as shown in figure 3.26. Domain 1 is mainly involved in coordinating FAD binding with an N-terminal Rossmann-like fold and contains residues 7-78, 106-190 & 290-376. This fold consists of five parallel β-strands (β1, β2, β8, β12 & β19) flanked on top by three anti-parallel β-strands (β9, β10 & β11) and a single α-helix (α5), and below by a bundle of five α-helices (α1, α2, α3, α4 & α12) and two anti-parallel β-strands (β3 & β7). Domain 1 also contains two anti-parallel β-strands (β13 & β18) that link domains 1 and 2. Domain 2 at the bottom of the protein contains most of the substrate interactions. Here a platform of two anti-parallel β-strands (β5 & β6) parallel to a further four anti-parallel β-strands (β14, β15, β16 & β17) sits on top of four α-helices (α6, α8, α9 & α10).

The long, kinked α-helix α13 forms the spine of the protein through domains 1 & 2 and precedes the extension into domain 3. Domain 3 then extends out into four C-terminal α-helices (α14-17) which form a solvent-exposed bundle that caps the substrate pocket on the opposite side to the flavin. Domain 3 returns to contact Domain 2 of the protein at a patch involving helix α15 and the start of α16 which partially covers the back of the roof of the substrate binding pocket. At this patch, the highly conserved residues I362, V407, F409 & Y414 form hydrophobic contacts to anchor domain 3 to the main protein scaffold. The substrate L-Kyn is bound in a solvent-excluded pocket in the middle of all three domains.

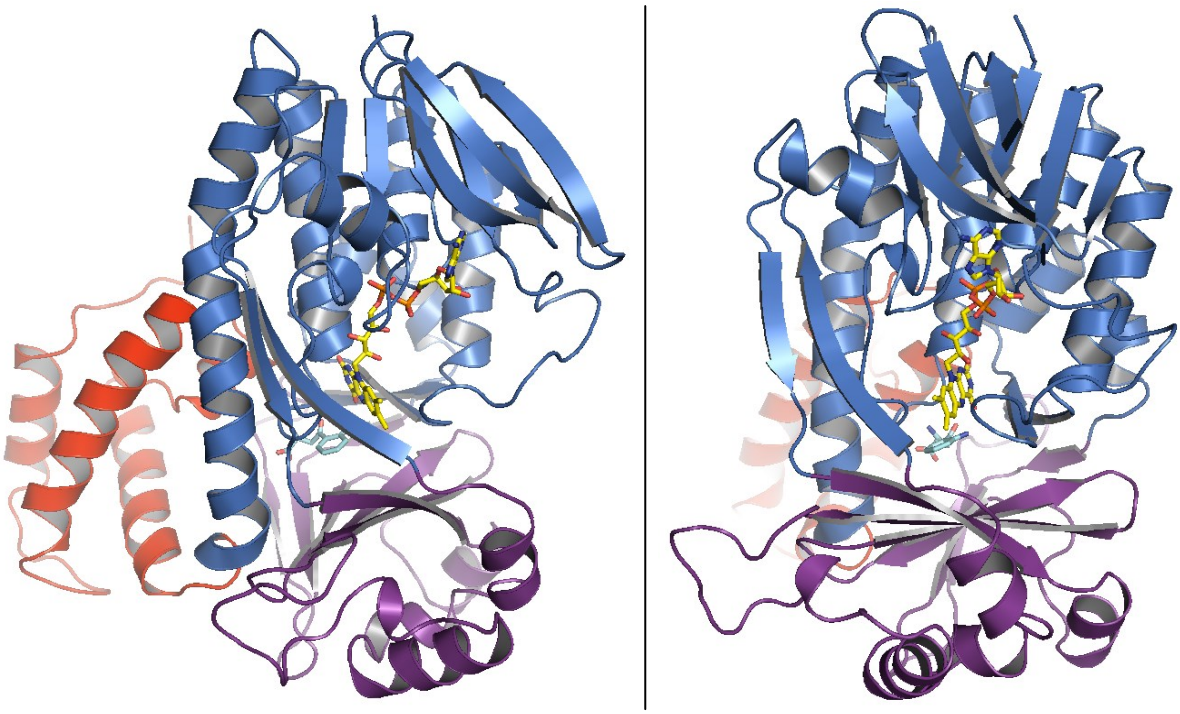


Figure 3.26. Two views of a cartoon representation of the final substrate-bound structure with domains 1-3 coloured blue, purple and red respectively. The substrate and FAD are shown as sticks, coloured cyan and yellow respectively. Figure created using PyMol.

The first 330-360 residues of the protein – effectively the entirety of domains 1 and 2 – align well with the structures of other flavin-dependent enzymes as shown in table 3.5. Of the most similar structures, all are confirmed or thought to belong to the FAH class. No conservation exists for the mobile domain 3 however, which appears to be completely novel. Where present, the third domain of other FAH enzymes is of varied size and structure with an external thioredoxin fold common to some cases. The FAH C-

terminal regions have been implicated in oligomerisation, simply capping the back of the substrate binding pocket or have unknown function. Figure 3.27 shows the *Pf*KMO structure in comparison with a few of the most similar FAH structures. The substrate is bound in a similar buried chamber beneath the isoalloxazine ring of FAD as in the displayed FAH structures. The flavin itself has previously been seen to occupy multiple conformations as described in section **3.1**. In substrate bound *Pf*KMO, the cofactor is present in the '*in*' conformation with a very similar local structural arrangement to the displayed structures of MHPCO, HpxO and *p*-HBH.

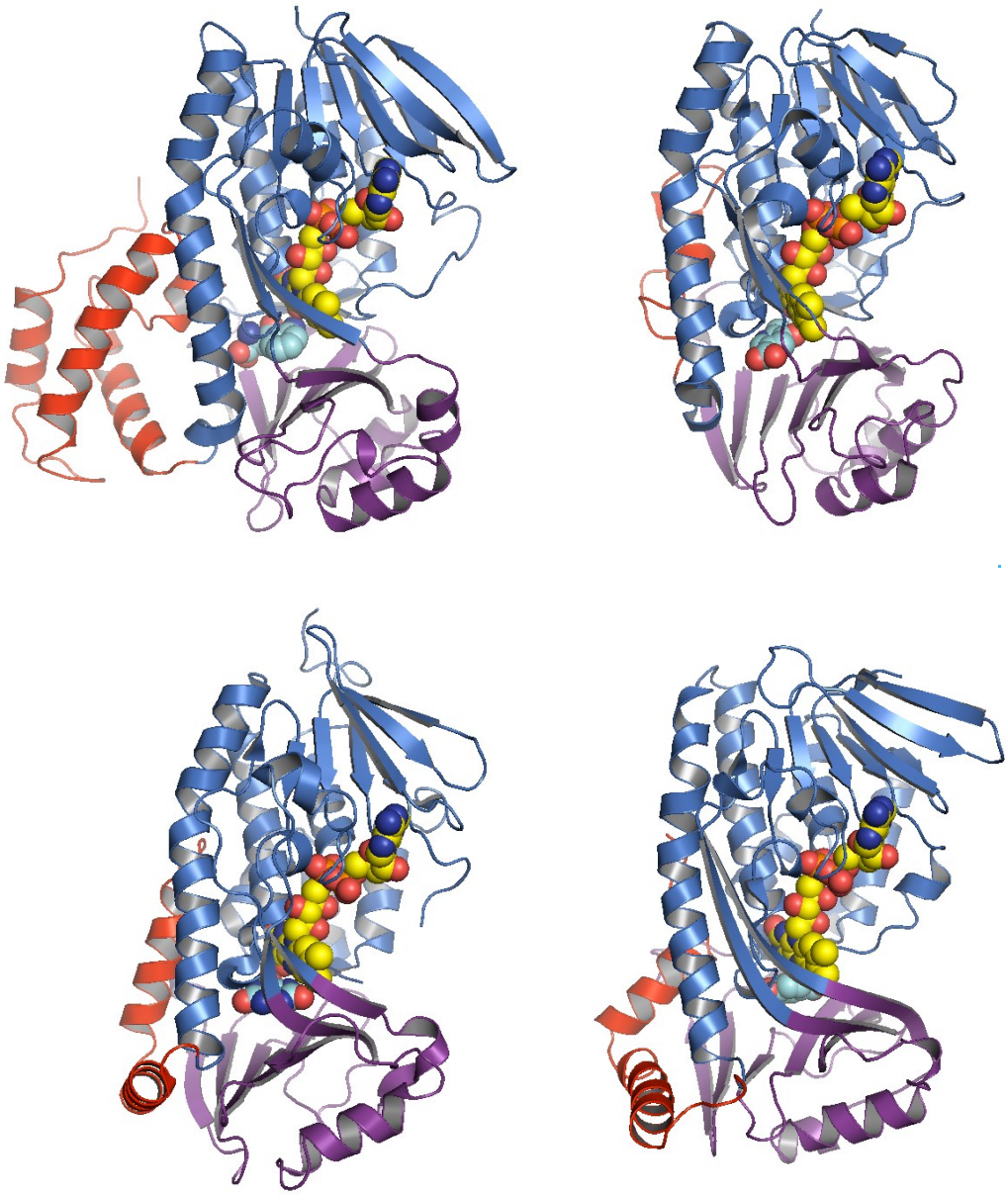


Figure 3.27. Cartoon representation of the final substrate-bound structure alongside the FAH structures of MHPCO (pdb code 3gmc), Reb C (3rp7) and p-HBH (1pbe). The relative sequence of domains 1-3 are coloured blue, purple and red respectively. The substrate and FAD are shown as spheres, coloured cyan and yellow respectively. Figure created using PyMol.

PDB	protein	Z score	rmsd (Å) for C_α aligned	residues aligned	sequence identity (%)
3ALJ	MHPCO	35.6	2.5	342	23
3RP8	HpxO	32.7	2.8	349	17
2X3N	PQSL	32.0	3.1	334	22
2Y6R	TetX	32.0	3.0	337	22
1CC4	*p*-HBH	31.4	3.5	359	19
2RGJ	PhzS	31.0	2.7	331	22
2VOU	DHPH	30.6	2.9	340	16

Table 3.5. Results of a search of the pdb for similar structures to substrate-bound dm2 *Pf*KMO using the Dali server [257].

A Dali server [257] alignment search using just the C-terminal eighty residues of *Pf*KMO revealed some structural resemblance to regions of a large number and variety of unrelated proteins. The C-terminus shows some relation to part of type I K homology (KH) domains, which are present in a variety of nucleic-acid binding proteins [258]. Figure 3.28 below shows a resulting alignment between the C-terminal domain of *Pf*KMO and the archaeal exosome RNA binding protein RRP4. The sequence identity for the alignment is only 10 % for 59 residues however there are conserved helical structural elements. The KH domain is however intertwined with additional β-sheets that are absent in the C-terminal domain of *Pf*KMO. It is possible that this is just a case of convergent evolution regarding *Pf*KMO whereby the different proteins have evolved towards similar arrangements for structural purposes. The C-terminal domain is essentially just a small four-helical barrel and such formations are likely to be commonly found among different protein structures.

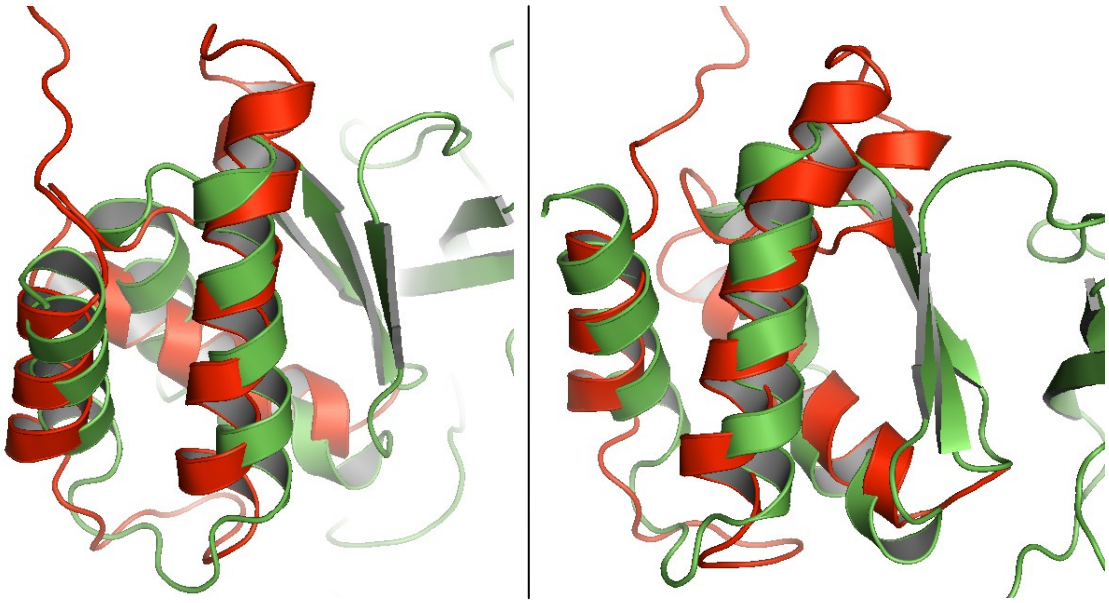
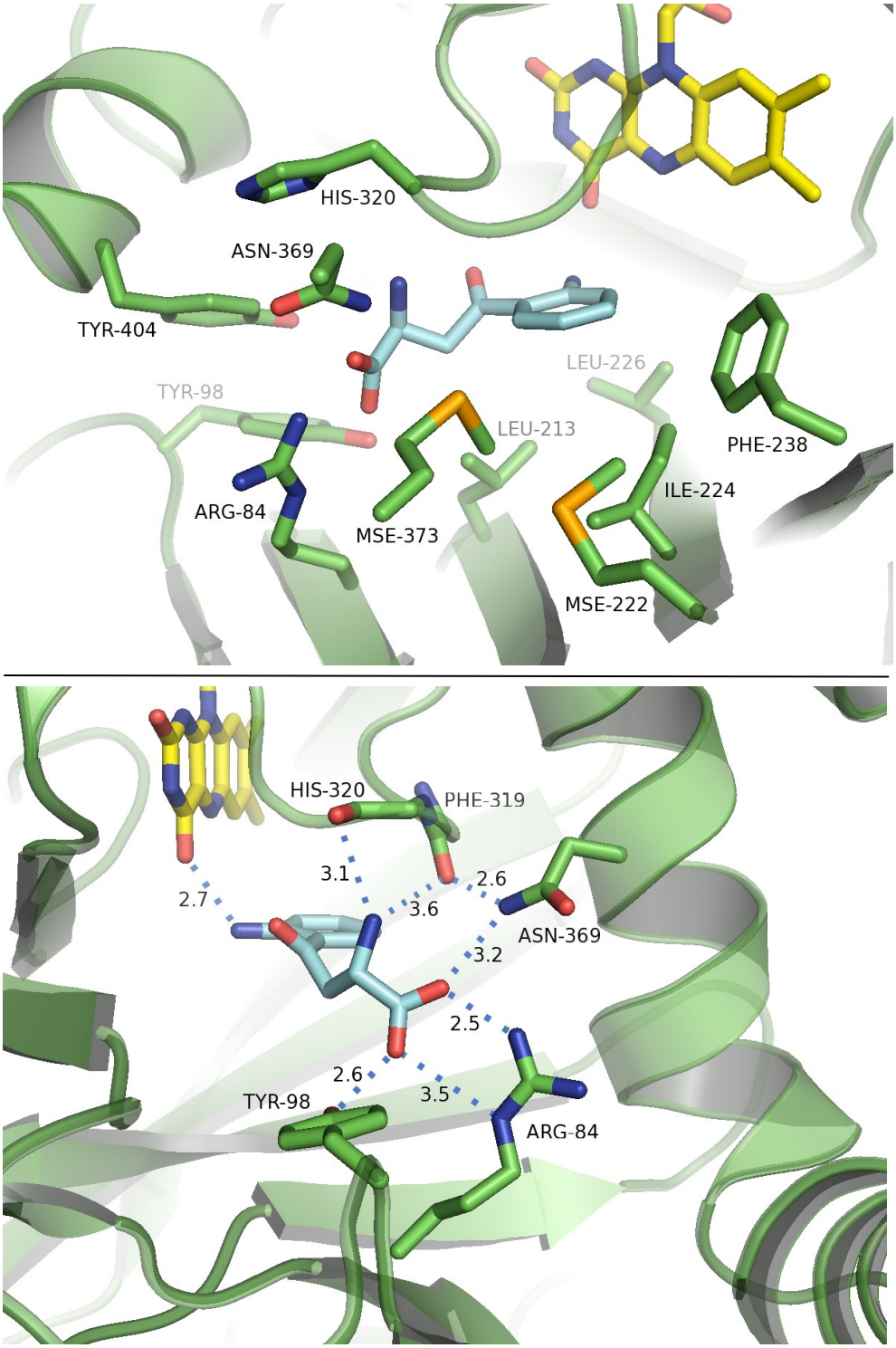


Figure 3.28. Stereo cartoon representation of the C-terminus of *Pf*KMO in red aligned with the archaeal exosome RNA binding protein RRP4 in green (pdb code 2ba0). Figure created using PyMol.

3.4.4 The substrate binding pocket

The substrate L-Kyn is tightly bound in between domains 1 and 2 with possible solvent access at the interface with domain 3. The binding site (see figure 3.29) forms a split hydrophobic and hydrophilic environment to meet the mixed demands of the substrate. The benzyl ring of L-Kyn is packed in a hydrophobic pocket below the isoalloxazine ring of FAD, lined with the side-chains of L213, I215, MSE222, I224, L226, F238, F319 & MSE373. On the other side its zwitterionic tail is surrounded by hydrophilic side chains including R84, Y98, H320, N369 & Y404. With the resolution of 3.4 Å one might typically expect around a 10 % positional error of 0.3-0.4 Å and so this was taken into account to filter plausible hydrogen bond distances. With this in mind, substrate specific binding may be coordinated by several hydrogen bonds (See figure 3.30). The majority position the carboxylate group of L-Kyn with potential hydrogen bonds to the side chains of R84 (2.5 & 3.5 Å), Y98 (2.6 Å) and N369 (3.2 Å) respectively. Additionally the tail ammonium group is well positioned to hydrogen bond with the backbone carbonyl of H320 (3.1 Å) and/or possibly with the backbone carbonyl of F319 (3.6 Å). Finally the 2-amino group of the phenyl ring appears to directly hydrogen bond with O4 of the isoalloxazine ring of FAD (2.7 Å). The distance between $C4_a$ of the FAD isoalloxazine ring – which forms the

catalytic hydroperoxy intermediate - and the hydroxylation site of L-Kyn is approximately 5.0 Å. This distance is feasible to allow the addition of molecular oxygen and subsequent hydroxylation of the substrate and is similar to that found in other FAH structures.



(Top) Figure 3.29. The mixed substrate-binding environment in the dm2 *Pf*KMO structure. L-Kyn (cyan), FAD (yellow) and key residues lining the pocket are shown in sticks.

(Bottom) Figure 3.30. Interatomic distances (excluding hydrogens) that could potentially represent hydrogen bonds involved in substrate binding. L-Kyn (cyan), FAD (yellow) and key interacting residues are shown in sticks. Figures made in PyMol.

3.5 Crystal structure of substrate-free PfKMO

3.5.1 Crystal growth and data collection

With wt *Pf*KMO, only microcrystals of substrate-free enzyme could be grown which were unsuitable for further study. After introduction of the double cysteine mutation, single crystals grew under similar conditions to those used with the enzyme co-crystallized with substrate. A salt additives screen (10 mM final salt concentration) established that larger crystals more reproducibly formed with tartrate or chloride ions. The morphology of the crystals was subtly different from those grown with substrate as can be seen with a substrate-free crystal displayed in figure 3.31. The drops contained 1.0 μl of protein solution (11.5 mg/ml in 20 mM HEPES pH 6.8, 20 mM K acetate) and 1.0 μl of the precipitant (0.1 M HEPES pH 7.0, 10 % w/v PEG 4000, 10 % v/v glycerol, 10 % v/v 2-propanol & 10 mM K tartrate or KCl). Crystals with a rounded wedge appearance and rough dimensions of 0.10 x 0.15 x 0.20 mm^3 grew in 3-7 days and were harvested for data collection after three weeks. Prior to flash-freezing in liquid nitrogen, 1 μl of cryoprotectant, mother liquor 20 % in EG, was added to the drop before the crystals were transferred directly into 2 μl of the cryoprotectant.

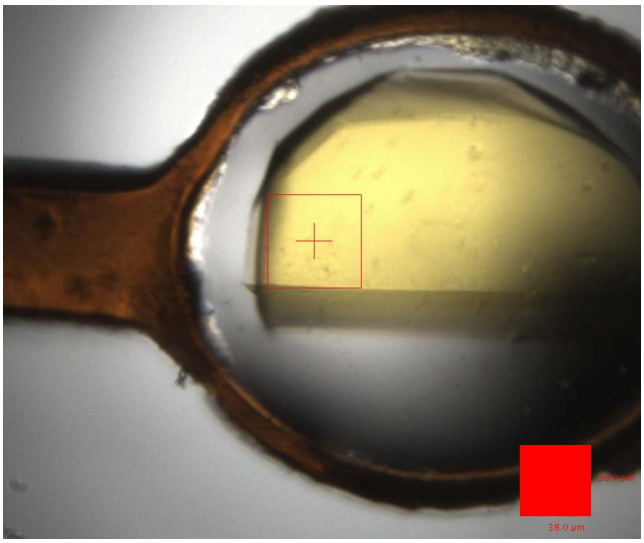


Figure 3.31. Mounted substrate-free crystal from the tartrate crystallisation drop which was used for data collection. The scale is labelled in the bottom right corner. The beam size (square box with a cross-hair in) was 50 x 50 μm.

Diffraction data were collected for each crystal at the Diamond synchrotron beamline I24 with an X-ray wavelength of 0.978 Å. All substrate-free crystals showed significantly stronger diffraction extending to below 2.5 Å as seen in figure 3.32. Using an exposure of 0.2 s and oscillation of 0.2 °, reflections were recorded over a wedge of 120 °. The crystals were packed in a new, orthorhombic space group - $P2_122_1$. The crystals had mean unit cell parameters of around a = 105.0, b = 134.0, c = 188.5 Å, α = β = γ = 90.0 °. Indexing and integration of reflections was carried out in Mosflm before scaling in Scala. The tartrate crystal dataset was 99.4 % complete to a resolution of 2.26 Å with R_{merge} = 0.055 overall (0.652 in the highest resolution shell) and average signal to noise of 13.5 (2.0). Similarly the chloride crystal dataset was 97.8 % complete to a resolution of 2.45 Å with R_{merge} = 0.063 overall (0.597) and average signal to noise of 12.4 (2.0). Full data collection statistics are tabulated in appendix section **3.2**. Based on the influence of salt ions on the reaction catalysed by *Pf*KMO described in section **3.3.1**, the data from the tartrate crystal were taken to initially represent the substrate-free structure.

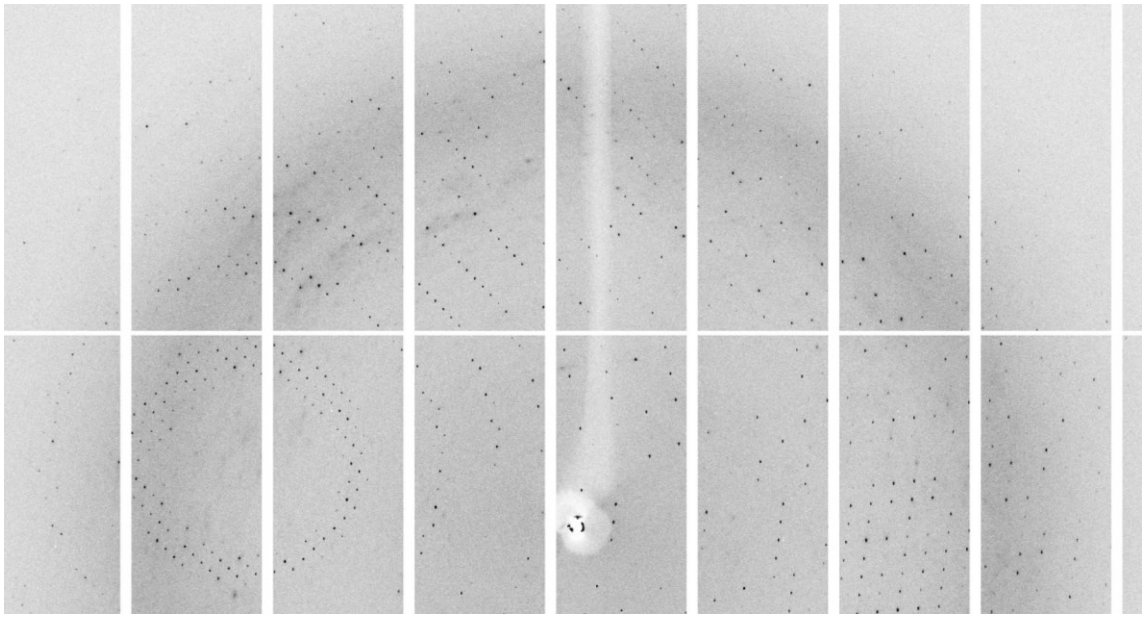


Figure 3.32. Diffraction image from a substrate-free crystal with reflections extending below 2.5 Å.

3.5.2 Refinement and overall structure

For the calculation of structure factors, a suitable model was generated using molecular replacement (Molrep) with the previously solved monomeric substrate bound structure (see section **3.3**). A definitive solution emerged with four protein molecules present in the asymmetric unit (see figure 3.33) yielding a solvent content of 62.9 %. The model was cycled through manual editing in Coot and restrained refinement in Refmac. For the first few cycles NCS restraints for each chain were applied in refinement. There was clear density for one FAD cofactor for each chain and only solvent in the substrate binding sites. For all four chains the residues 6-461 were included, except for two loop regions. For the loop involving residues 42-52 in front of the FAD binding site there was only partial density with additional density for a new, more solvent exposed loop position. Therefore these residues were removed from each chain in the structure. In addition, in chain D there was insufficient density to clearly place residues 375-378, which precede the C-terminal domain. Towards the end of refinement, coot was used to manually place 431 waters in the model.

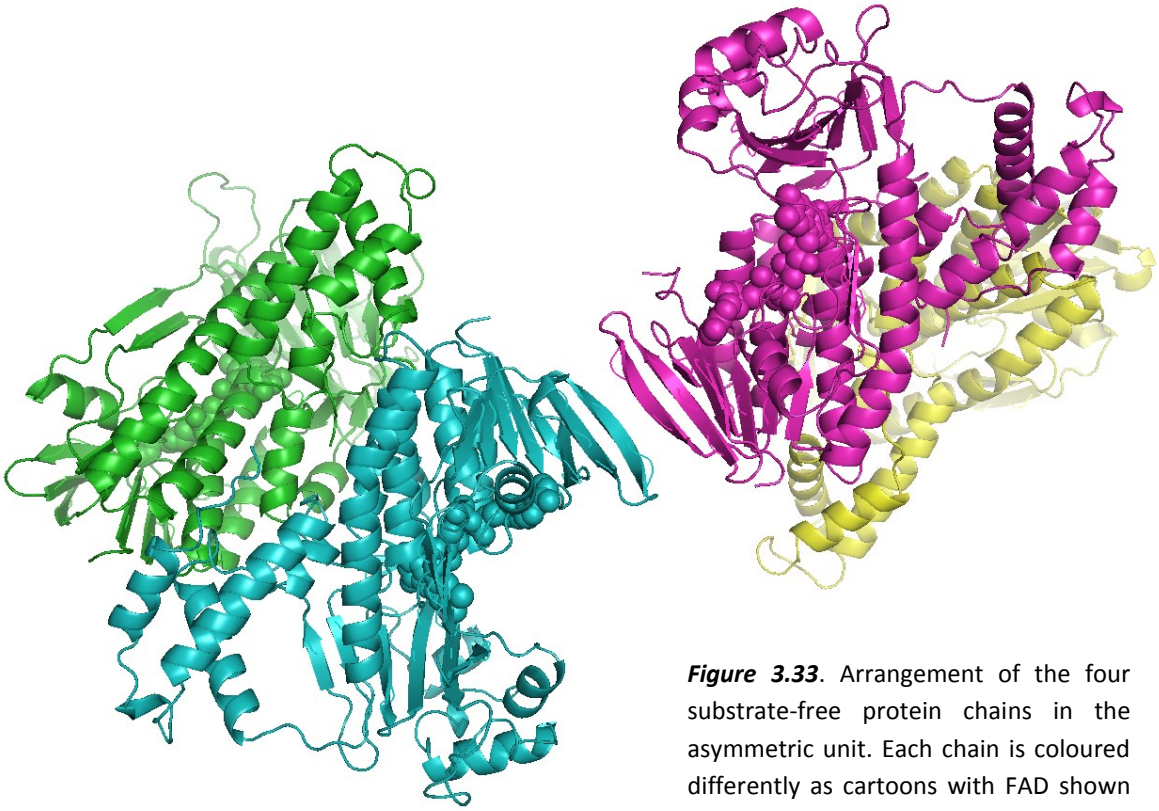


Figure 3.33. Arrangement of the four substrate-free protein chains in the asymmetric unit. Each chain is coloured differently as cartoons with FAD shown in spheres. Created using PyMol.

The final model was refined with the addition of TLS to give R and R_{free} values of 0.219 and 0.232 respectively with an overall mean B factor of 49.4 $Å^2$ for 117656 significant reflections in the range of 50.0-2.26 Å. The root means square deviation (rmsd) of bond lengths and angles from the ideal were 0.0050 Å and 0.94 ° with all the residues present in the most favoured or additionally allowed regions of a Ramachandran plot. The Molprobity server was used to validate the final model. Full refinement statistics are tabulated in appendix section **3.2**.

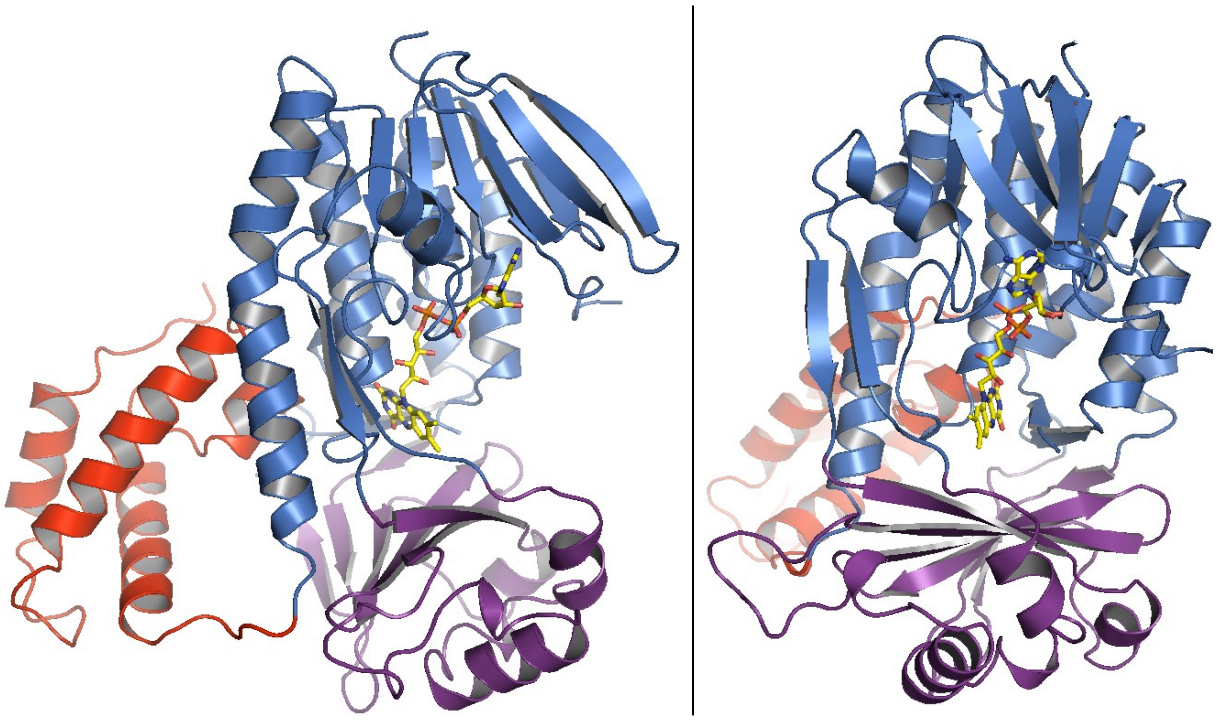


Figure 3.34. Two views of the cartoon layout of chain A of the substrate-free structure with the sequence representing each of the three domains coloured blue, purple and red. FAD is shown in yellow sticks. Figure created using PyMol.

The arrangement of the substrate-free structure showed a similar layout to substrate bound *Pf*KMO with the same residues forming three structural domains (See figure 3.34). Each polypeptide chain showed a rmsd of 1.4-1.5 $Å^2$ (for protein C_α atoms) when separately superimposed on to the substrate bound structure by the DaliLite pairwise server. Looking at an alignment of the two structures (see figure 3.35), it was clear that there was little to differentiate the first two domains bar the absence of substrate. The FAD cofactor was located in virtually the same position and the active site itself showed no clear differences. This was not the case for the whole protein though, as a clear

displacement of the novel C-terminal domain 3 is apparent. The helix α15 remains anchored to the back of domain 1 and acts as a pivot on which the helices α14, α16 & α17 swing away from the unfilled substrate binding pocket. In particular, the C_α backbone of helix α14 swings out by up to 6.1 Å at residue 381 in the substrate-free structure compared to with substrate bound. The two structures appear to display two conformations, a 'closed' conformation with substrate bound and an 'open' conformation in its absence. A different view of the movement in the structures (figure 3.36) shows that the open conformation creates a solvent-exposed tunnel leading directly to the substrate binding pocket.

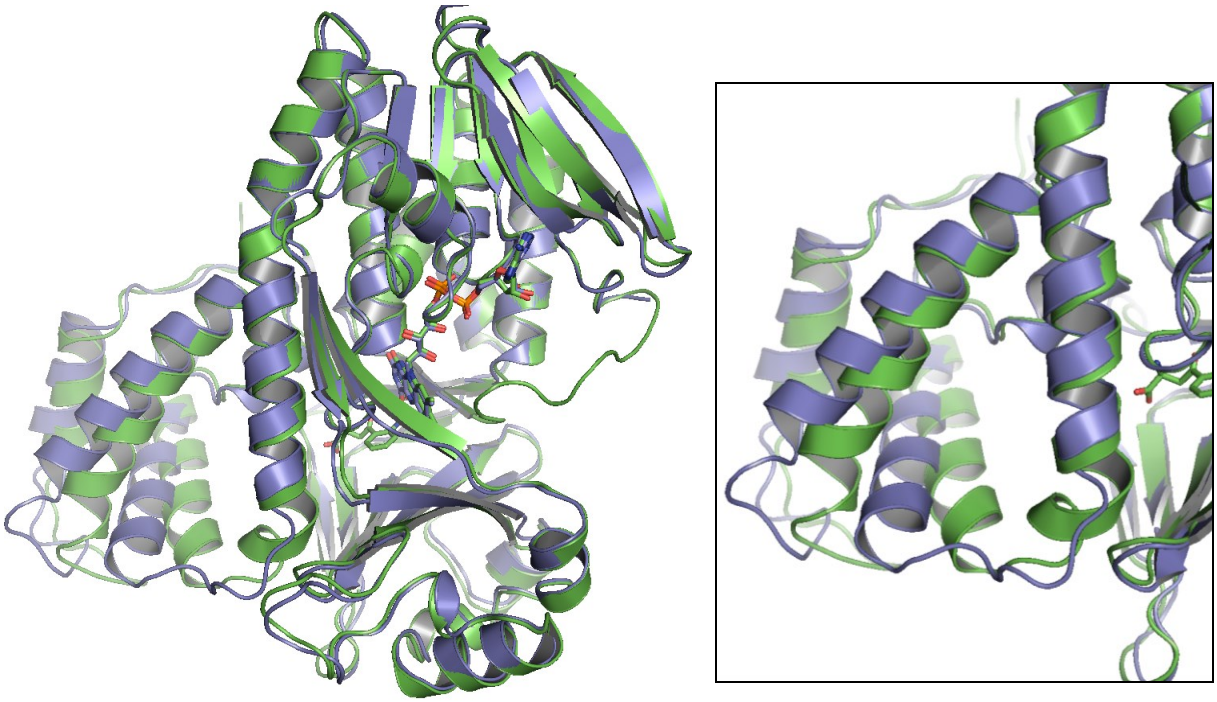


Figure 3.35. Alignment of chain A of the substrate-free (violet) and the substrate-bound dm2 *Pf*KMO structures. The overall alignment is shown on the left with a close-up of just the C-terminal domains shown on the right. The substrate and FAD are shown in sticks. Figures created with PyMol.

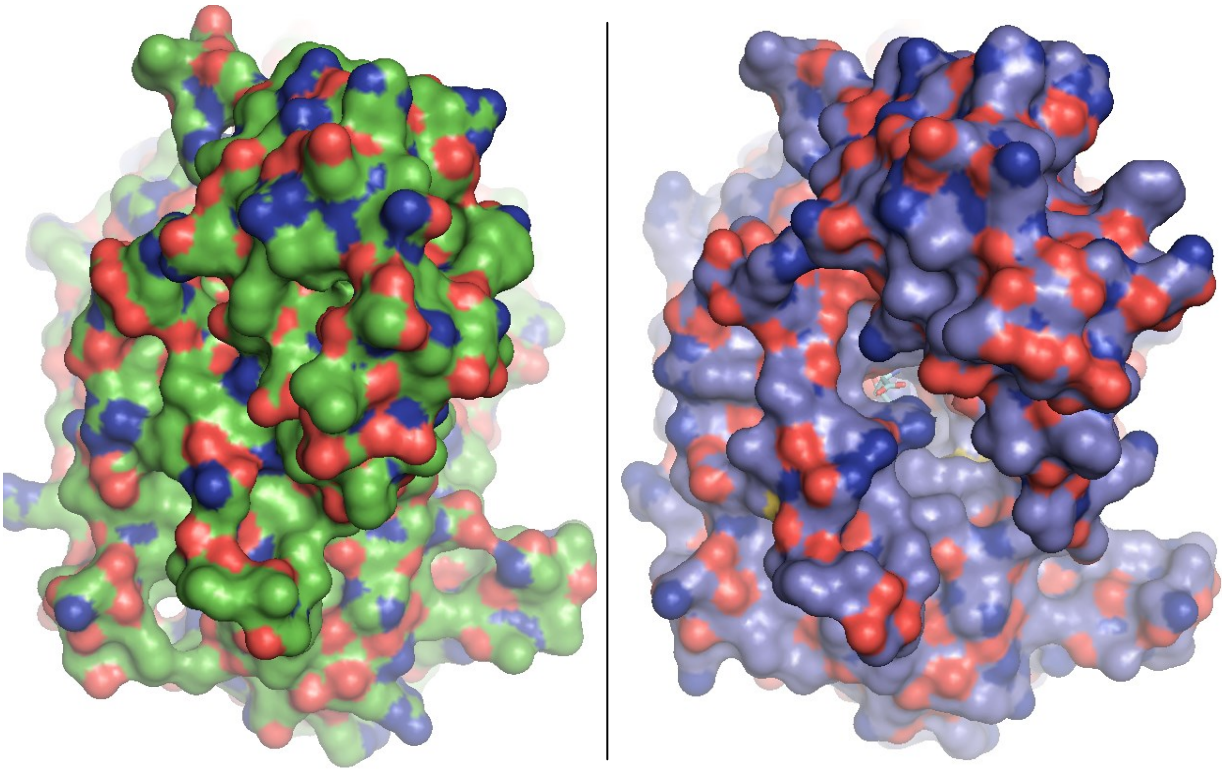


Figure 3.36. View of the translocation of the C-terminal domain which opens up a potential entry/exit tunnel to the substrate binding site. A surface view of the back of the substrate-bound structure (green) on the left adjacent to the corresponding surface view of the substrate-free structure (violet) on the right. The substrate (cyan sticks) is modelled into the substrate-free structure to illustrate the location of the binding pocket. Figures created with PyMol.

3.5.3 Structural differences in the presence of Cl^- ions

The presence of chloride ions in the crystal mother liquor caused subtle structural differences compared to the above tartrate-grown substrate-free structure. The final model possessed R/R_{free} values of 0.218/0.244 and a B factor of 46.2 $Å^2$ for 91114 significant reflections in the range 35.00-2.45 Å. It contained residues 6-461 and a molecule of FAD in the same four-chain arrangement as displayed above. On all four chains, a chloride ion was placed above the active site in a cleft adjacent to the flavin's isoalloxazine moiety as displayed in figure 3.37.

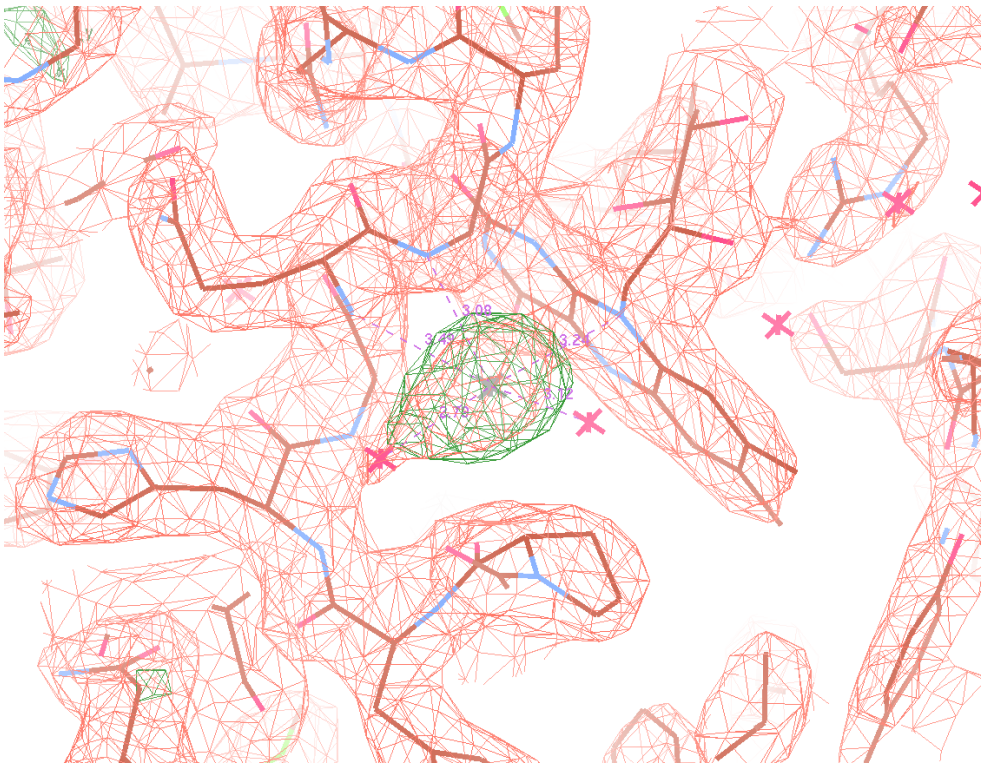


Figure 3.37. Binding site of a chloride ion (grey) in the chloride-bound structure above the substrate-binding pocket. The refined $2F_o$-F_c map is shown in orange (contoured at 1.5 σ) with the F_o-F_c map calculated in the absence of the ion shown in green (contoured at 3.5 σ). Figure created in Coot.

Overall the structural elements of each domain align well with the tartrate structure and with the C-terminal domain again in the open conformation. The average rmsd of the backbone C_α atoms of each chain to those in the tartrate structure ranges from 0.7-0.9 $Å^2$. Significant differences were evident in two distinct interlinking loop regions. Firstly the aforementioned external loop involving residues 45-53 showed strong density for just a single conformation – that of the substrate bound enzyme. This is shown in comparison to the tartrate structure in figure 3.38. Secondly the C-terminal end of the long helix α-13 and the adjoining loop into the start of domain 3 were pulled in to be more like the substrate bound structure. The two loop positions described for the tartrate and chloride structures are displayed in figure 3.39. This meant that the substrate-free 'chloride' structure provided a new mixed C-terminal conformation with elements of both the substrate bound and unbound 'tartrate' forms.

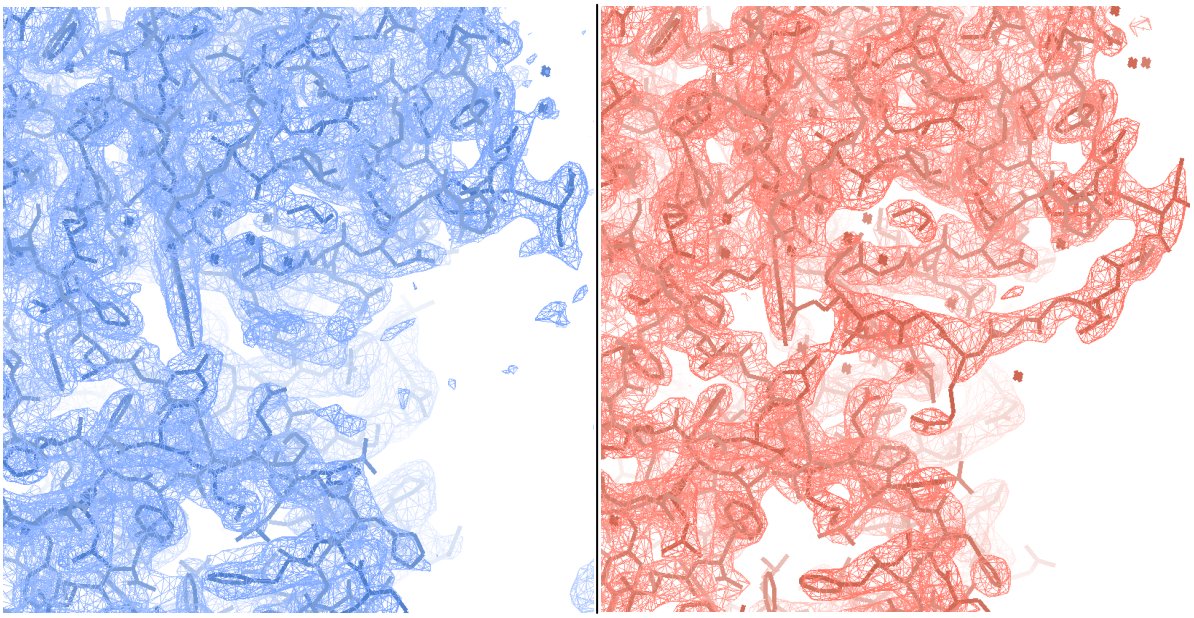


Figure 3.38. Loop involving residues 42-52 without density in the tartrate substrate-free structure (*left,* blue) but is clearly visible in the density for the chloride substrate-free structure (*right,* orange). The refined $2F_o$-F_c density is contoured at 1.5 σ with the figures created using Coot.

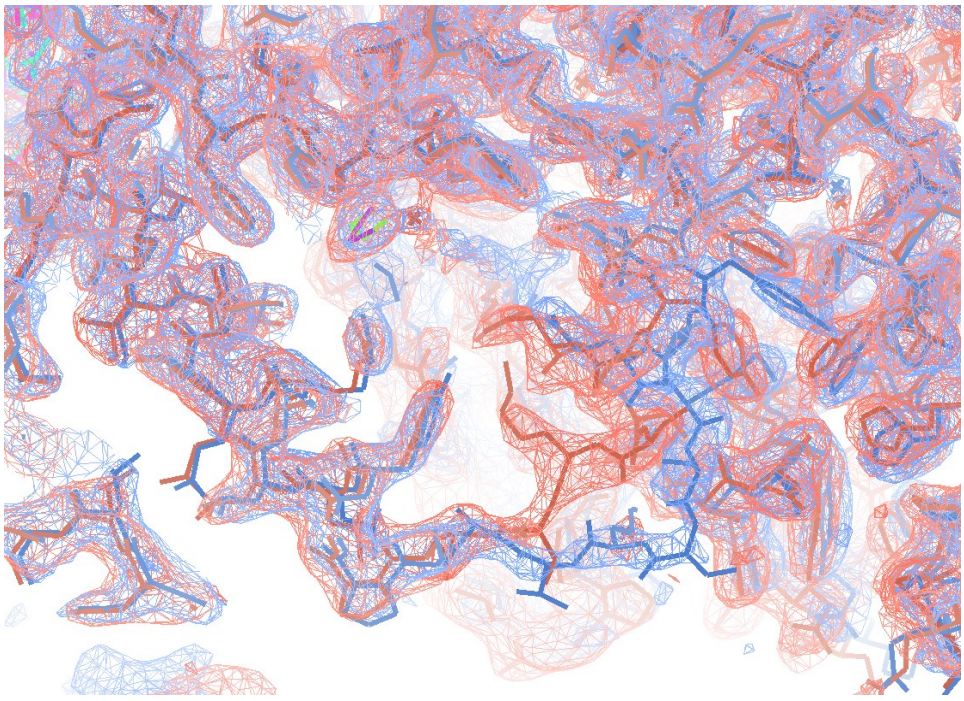


***Figure* 3.39**. $2F_o$-F_c density maps for different positions of the loop preceding the C-terminal domain in the chloride (orange) and tartrate (blue) substrate-free structures.
In the chloride structure the loop is pulled in towards the back of the substrate binding pocket as seen in the L-Kyn bound structure. Created using Coot with contouring at 1.5 σ.

At this stage the structural variations could either be dependent upon substrate binding, the influence of the respective salt ions or the packing of the crystal lattice. It is well known that small anions including chlorides strongly inhibit FAH enzymes such as pHBH (see section **1.2.3**), and have directly been shown to inhibit both *Pf*KMO and human KMO (see section **1.3**). Either way an interesting glimpse into the possible dynamics of the enzyme was developing. From this point onwards the tartrate substrate-free form is labelled as the unbound structure and the chloride form as the chloride-bound structure.

3.6 The cofactor binding site

As previously mentioned, the enzyme utilises a Rossmann-like fold to form the non-covalent FAD binding site. The cofactor density in the unbound and the chloride bound structures is very clear with average B-factors of 35.9 and 40.1 $Å^2$ respectively. In the lower resolution substrate bound structure the FAD B-factor is naturally higher at 104.2 $Å^2$. Comparing all the structures there are no significant differences in the orientation of the FAD or in the local protein environment. The high resolution unbound structure provides a detailed view of the interactions that govern flavin binding. The refined electron density calculated with and without FAD are highlighted below in figure 3.40.

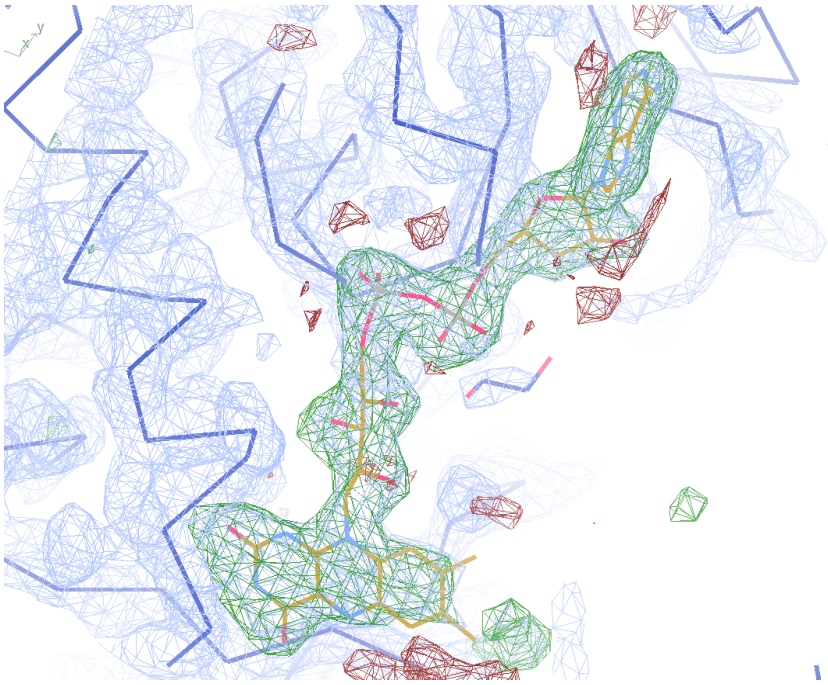


Figure 3.40. Fit of the FAD cofactor into the refined $2F_o$-F_c (blue) and F_o-F_c (green/red) density (calculated in the absence of FAD in the model) in the unbound (tartrate) structure with contouring at 2.0 and 4.0 σ respectively. Created using Coot.

The many potential hydrogen bonds associated with FAD binding in the unbound *Pf*KMO structure are summarized in figure 3.41. The adenine nucleotide is stabilised by interactions with the main-chain amide CO of L135 (3.3 Å), the amide NH of L135 (2.9 Å) and a water molecule (2.9 Å). Also the side chain of R38 stacks above the nucleotide partially shielding it from solvent. More extensive H-bonds are found with the sugar 2′ and 3′-OH groups and the perpendicularly stacked, highly conserved side-chains of E37 (2.5 & 2.7 Å) and R39 (2.8 & 3.1 Å). The glutamate interactions in particular are also highly conserved among other published FAH structures. Moving down, the di-phosphates form H-bonds with the amide backbone NH of residues A18 and D311 (both 2.9 Å) as well as with a few molecules of solvent. The side-chains of R111 (2.8 & 3.1 Å) and D311 (2.7 Å) H-bond with the next moiety – the ribityl chain. D311 is part of the strictly conserved GD motif among FAH enzymes. Finally, the flavin isoalloxazine ring is present in the 'in' conformation in all structures with substantial interactions with residues G321 and G323-N325. All the residues discussed are highly conserved in alignments of KMO sequences.

As mentioned in section **1.2.3.1**, FAH enzymes utilize a dynamic cofactor isoalloxazine ring to accommodate the different solvent requirements of flavin reduction and of hydroxylation of the substrate [148]. In all the *Pf*KMO structures in this section, the flavin cofactor is present in the *in* position in the buried active site. FAD captured in the *out* position in the structure of unbound *m*-HBH [175] was modelled into the structure of *Pf*KMO as displayed in figure 3.42. The loop involving residues 42-52, highlighted in green, clashes with the *out* conformation of the isoalloxazine ring. This loop is disordered in the unbound *Pf*KMO structure and so may be flexible enough to undergo the small displacement required to allow the predicted flavin movement. Counter intuitively to this, the B-factors for the residues in the loop are not that high in the substrate bound structure. It is possible that this loop is involved in forming the NADPH docking site but then may be pushed aside once NADPH docks.

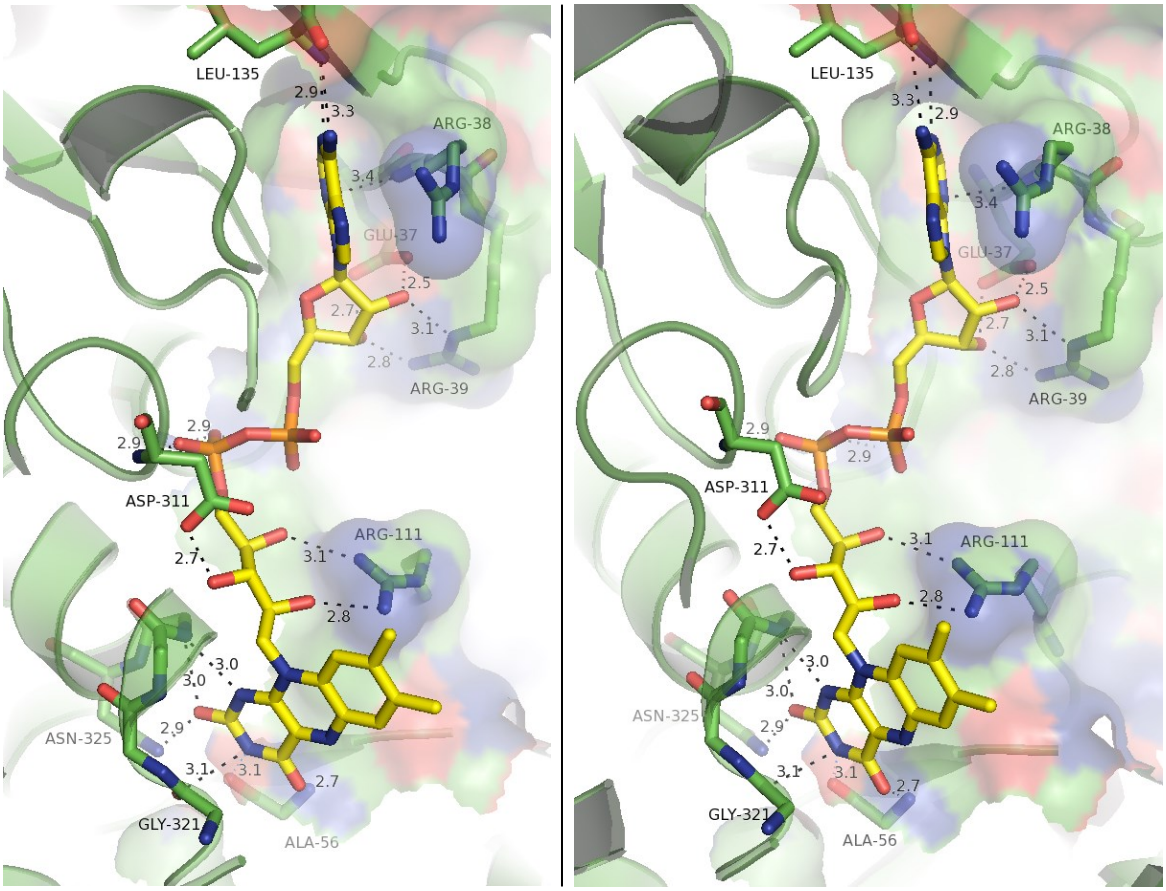


Figure 3.41. An overview of the interactions surrounding the non-covalent binding of the FAD cofactor in the unbound dm2 *Pf*KMO structure. Part of the FAD pocket is shown in surface view with key protein residues labelled and shown as sticks. Figure created with PyMol shown in stereo.

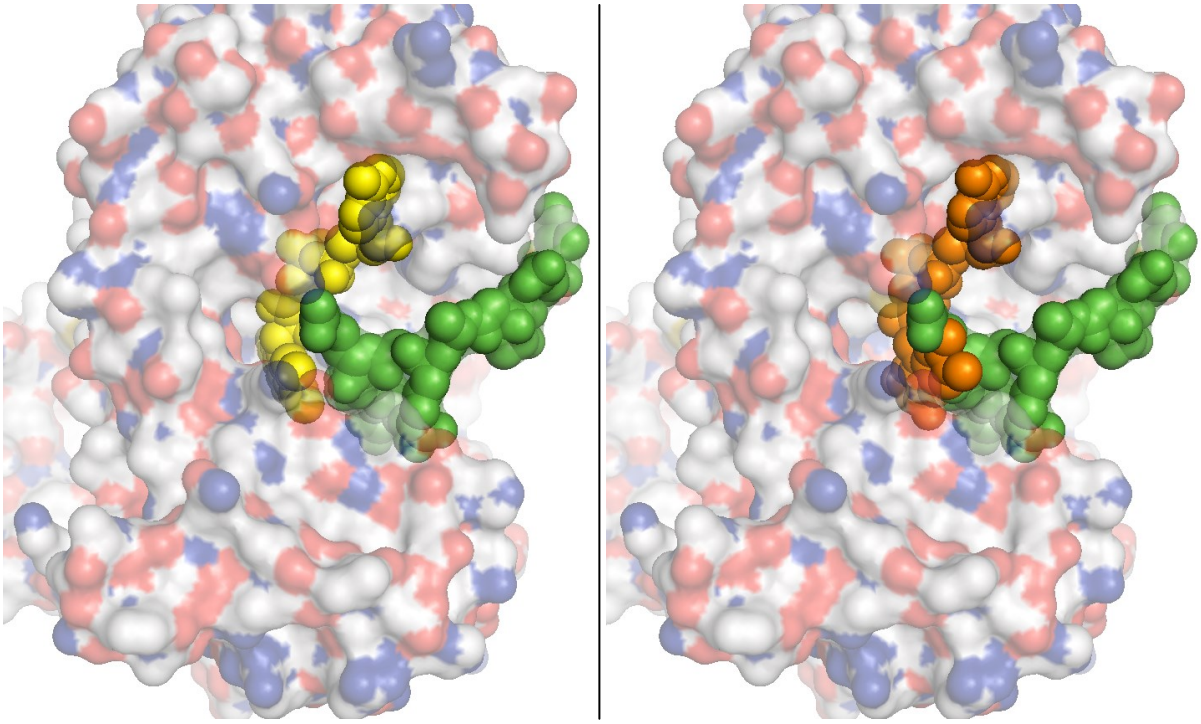


Figure 3.42. Potential clash observed between loop 42-52 (green) and FAD when it is modelled into the substrate-bound structure in the *out* position (orange) as taken from the *m*-HBH structure with pdb code 2dkh. On the left is the substrate-bound *Pf*KMO structure with FAD shown in yellow. Created using PyMol.

There is one FAH crystal structure that contains bound NADPH, which was obtained with an R220Q mutant version of *p*-HBH [163]. The mutation causes a slight displacement of the substrate binding domain, which reduces affinity for the substrate and destabilizes the flavin's *in* position. The orientation of the external NADPH cofactor in the structure results in an unrealistically large distance between the transferrable hydrogen atom and the flavin isoalloxazine ring. The authors proposed that for flavin reduction, NADPH adjusts into a folded conformation to bring the nicotinamide moiety towards the flavin isoalloxazine ring. The bound NADPH from this mutant *p*-HBH structure is aligned onto the structure of substrate bound *Pf*KMO below in figure 3.43. The modelled NADPH fits well onto the surface of the enzyme even with the 42-52 loop present. Also, residues conserved in all KMO sequences can be seen to cluster around the site of flavin movement and NADPH docking.

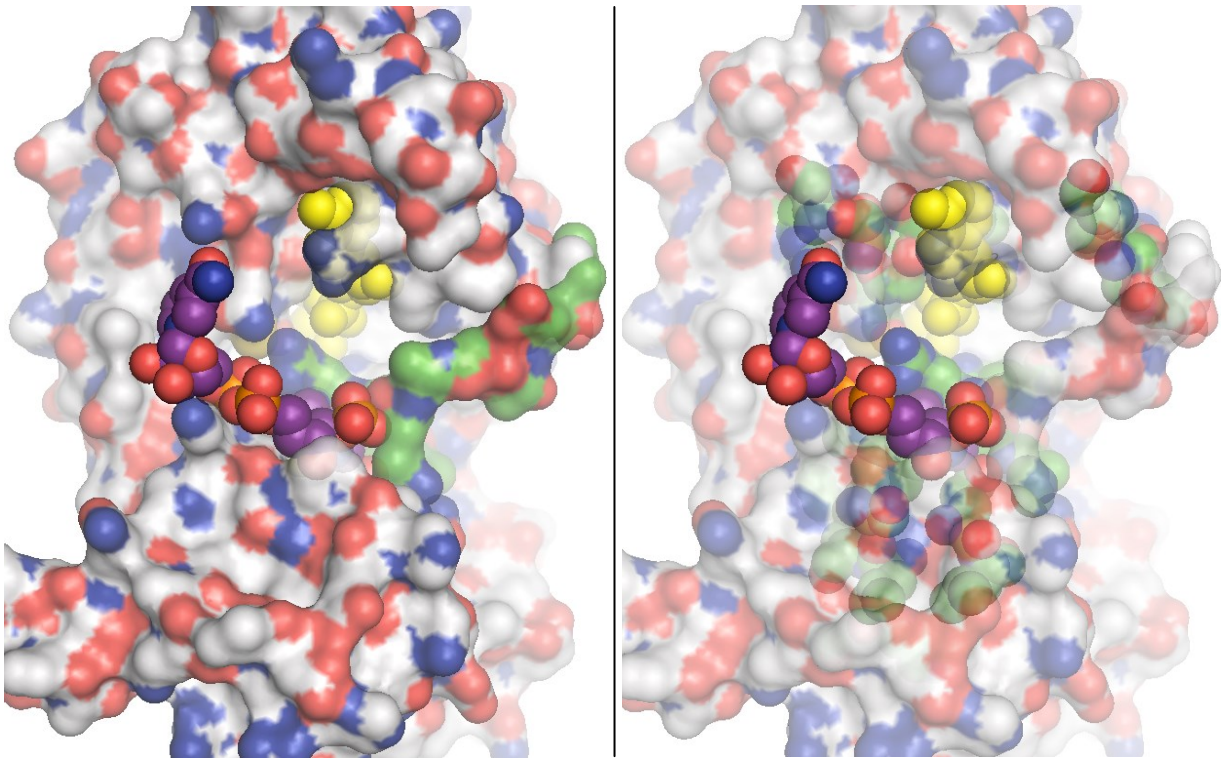


Figure 3.43. NADPH (purple spheres) docking into the substrate-bound *Pf*KMO structure (white). On the left loop 42-52 is highlighted green and on the right conserved residues among KMO sequences are highlighted green which appear to surround the modelled NADPH molecule. NADPH modelled based on the relative position in the structure of a *p*-HBH mutant structure (pdb code 1k0j). Created using PyMol.

3.7 Discussion and conclusions

In this investigation the first ever KMO structures have been solved of substrate-free and substrate-bound *Pf*KMO. The road to the substrate bound structure was long but progressive as the complex only crystallized in a lattice consisting of over 80 % solvent, resulting in very weak diffraction. Conditions were optimised to slowly grow larger crystals to give stronger diffracting power and with the additional mutagenesis of two cysteines to serines, resolutions of around 3.50 Å were reached. The substrate-bound structure was subsequently solved to a resolution of 3.40 Å using SeMet-labelled protein. Thirteen of the fourteen SeMet residues (missing the N-terminal residue) were found and whilst some were present in the vicinity of the active site, there was no obvious differences when the structure was modelled into data from unlabelled substrate-bound dm2 *Pf*KMO (data not shown). With the double mutation single crystals of unbound *Pf*KMO could be grown and without much optimisation showed significantly stronger diffraction despite growing from similar conditions and reaching much smaller sizes to those of the substrate bound enzyme. All unbound crystals were packed substantially differently in a new space group with a lower solvent content of around 60 %. Therefore these crystals were readily able to diffract to high resolutions, around 2.30 – 2.50 Å and add confidence and quality to the initial low resolution substrate bound model. The structures confirmed that both residues (252 and 461 respectively) are solvent exposed on the surface of the protein.

The two resulting structures revealed a three-domain structure with the first two domains arranged with a similar overall fold to all other published FAH structures. Also positioned in similar respective locations were the non-covalently bound FAD cofactor and substrate L-Kyn. The three-helical bundle of the C-terminal domain 3, however, is completely new to this class of enzyme. The domain extends out from the back of the substrate binding pocket and revealed an interesting translocation between the two structures, which appears to control the opening/closing of an access route to the substrate-binding pocket.

A substantial patch of free electon density was present in the active site to place L-Kyn into the substrate bound model. At 3.40 Å resolution the exact position of the functional

groups was not definitive however the side-chains of residues R84, Y98, N369 and possibly Y404 appear well positioned to interact with the bound substrate. It appears that the substrate's carboxyl group is involved in the majority of the interactions and could be critical to substrate recognition. As seen in other FAH structures, the substrate is directly below the isoalloxazine ring of the FAD in a large buried pocket. A similar distance of approximately 5.0 Å separates the catalytic C_{4a} of the flavin from the hydroxylated carbon of the substrate, fitting in with previous evidence that the reaction proceeds via a C_{4a}-hydroperoxyflavin intermediate. All of the residues that are located within 6 Å of the substrate are highly conserved among all KMO homologs including *h*KMO. The only residue in the active site that is not identical in a sequence alignment between *Pf*KMO and *h*KMO is H320, which is a phenylalanine in the human enzyme. This histidine is located directly above the carboxyl group of L-Kyn with its side-chain pointing out towards the possible route of substrate entry. There is over 5 Å separating this side-chain from L-Kyn suggesting that it does not directly interact with the substrate once bound. Looking at rat KMO, which has been the model for the synthesis of several published KMO inhibitors, there is one slight additional variation to both *Pf*KMO and *h*KMO. The residue relative to T408 of the bacterial enzyme is an alanine in the rat enzyme. The side-chain of T408 lines the substrate binding pocket at the opposite side to FAD, around 6 Å from the tail of L-Kyn.

Looking at other substrate bound FAH structures the orientation of the substrate itself seems to give away its route of entry in to the active site. Typically, as would be expected, the hydrophobic portion of the molecule is first to enter the buried site to evade interactions with solvent, leaving the hydrophilic end to point out towards solvent. L-Kyn can be clearly separated into a hydrophobic benzene ring and a hydrophilic zwitterionic tail as reflected in the split nature of the surrounding active site residues. The logical route of entry of the substrate into its buried binding site appears to be through a tunnel that leads to the back of the pocket opposite to the flavin. This tunnel is sandwiched between the structural elements β-6 of domain 2 and α-14 of domain 3. Interestingly the route to solvent is closed off in the substrate bound structure due to the closed conformation of domain 3. The apparent opening of this domain and subsequently of the tunnel in the unbound structure suggests a regulatory role of this

novel domain in substrate binding. At this stage of course it can't be established whether the change is directly dependent upon substrate binding or upon crystal lattice packing. The dynamics of domain 3 are investigated further in the next chapter.

In addition to unbound and substrate-bound forms, a chloride-bound structure was obtained. It had a similar crystal lattice and overall structural properties to unbound *Pf*KMO but presented a couple of loop elements that resembled the substrate bound structure. A chloride ion was found bound adjacent to N_{10} of the isoalloxazine moiety on the *si* side above the substrate binding pocket. The ion was positioned quite far from the two respective loop alterations though, raising questions as to its significance. Previously in *p*-HBH structures the binding of a bromide ion on the *re* side of the flavin was linked to the observed conformation of the flavin. In the case of *Pf*KMO the flavin conformation was not observed to be dependent on the presence/absence of the aforementioned chloride ion. In all the structures the FAD was positioned in the buried *in* position and additionally was shielded from solvent by loop 42-52 in the substrate bound structure. This loop would partially block flavin movement to the *out* position based on other FAH structures. Using NADPH bound in the structure of a *p*-HBH mutant [163], the coenzyme is able to dock in a similar location in *Pf*KMO. From analysis of the model and the KMO structures a possible hypothesis to explain the logistics of flavin reduction involves:

i) Flavin is positioned in the *in* position regardless of substrate binding.
ii) Substrate binding stimulates formation of the NADPH docking site which includes residues 42-52.
iii) Once NADPH docks it changes conformation and drives local movement to allow the FAD to swing to the *out* position.
iv) NADPH reduces FAD.
v) $NADP^+$ rapidly dissociates and the reduced flavin swings back to the active site to react with molecular oxygen.

In summary, the presented structures detail the environment of the active site of *Pf*KMO. With high sequence conservation of all the surrounding residues the structures should provide a suitable model to base investigations of the active site of human KMO. In particular this should be able to facilitate rational inhibitor design and screening for

the first time to enhance the discovery of therapeutically relevant compounds. Homology models for the human enzyme based on the presented *Pf*KMO structures can be found in the appendix section **2**. Sequence alignments and the structural models further suggest that the first two domains of the enzyme may be a relevant scaffold for those of human KMO. The predictions are of course biased by the template structure but it is important that there is no evidence in the primary sequences that KMO has undergone significant structural changes during the course of evolutionary divergence between the bacterial and mammalian enzymes. This sadly is not the case for the whole sequence of the enzyme though. Difficulties are encountered trying to predict the C-terminal domain of the human enzyme as the secondary structure prediction breaks down around residue 425 where mammalian KMOs have evolved a transmembrane helix.

The C-terminus, and particularly the apparent conformational changes of *Pf*KMO, will be explored further in the next chapter to try and see just how much information can be learned about the novel domain and its potential properties in human KMO.

Chapter 4 – Investigation of the dynamic, novel domain 3

4.1 Introduction

In chapter 3, the structures of substrate bound and substrate free *Pf*KMO were described, and revealed the presence of a novel dynamic C-terminal domain. Upon aligning the structures, a 7 Å swing of this domain from the back of the substrate-binding pocket was apparent between an in/closed and out/open conformation. This movement revealed the most obvious route of substrate entry/product release from the active site. Figure 4.1 shows the two structures with each residue colour-coded depending on their relative B-factor values. Looking at the unbound structure on the left, the section in red between the bottom of the vertical kinked α-helix and the first helix of the C-terminal domain (which includes the sequence with the greatest positional displacement between the two structures) has relatively very high B-factors. The corresponding, translated region of the substrate bound structure shows relatively lower B-factors indicating that the presence of the substrate stabilises the residues in the in/closed conformation. Moreover it should be kept in mind that the substrate bound structure was solved at a much lower resolution and so had much higher B-factors on average.

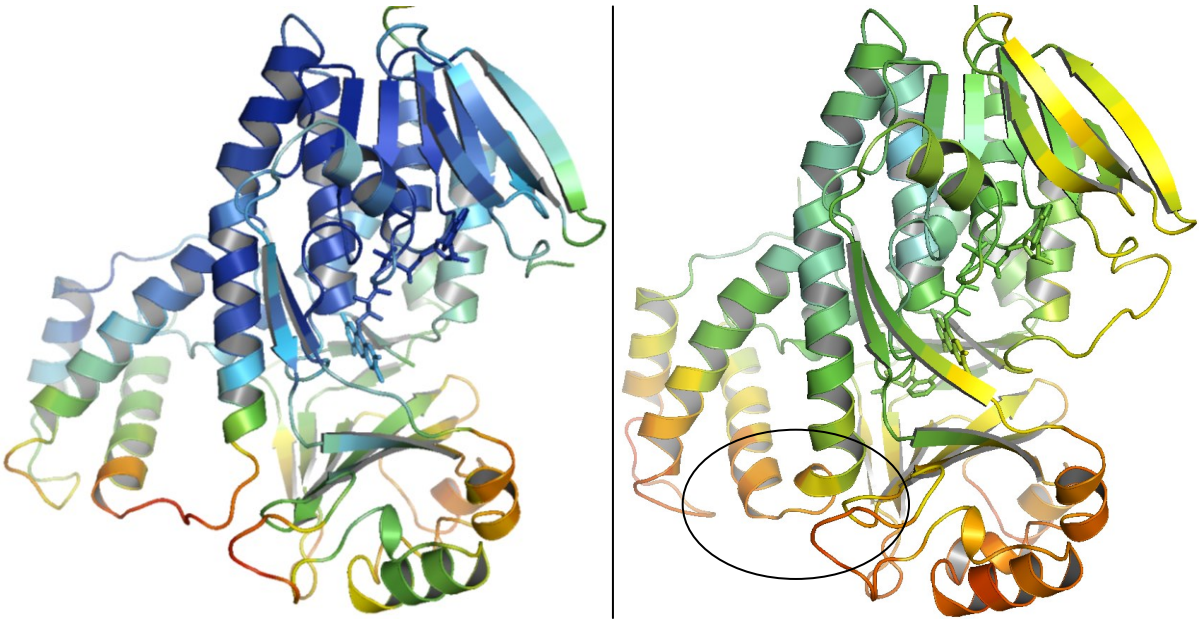


Figure 4.1. Cartoon representation of unbound (left) and substrate-bound (right) *Pf*KMO structures with B-factors coloured low-high as a spectrum of blue-red. The highly flexible C-terminal section in the unbound structure is highlighted by a black circle.

Looking at the structures from chapter **3** in more detail reveals that the side-chains of R84, E372, Y382 & R386 are well positioned to link the substrate to the closed helix α-14 - possibly via a hydrogen bond network. The network finishes with R386 linking domain 3 back to the main protein scaffold with a potential hydrogen bond to the backbone carbonyl of P97. In the open conformation observed in the unbound structure, Y382 and R386 are clearly moved away and the link breaks down (See figure 4.2). Therefore it is plausible that once the substrate has passed into the substrate binding pocket, it pulls domain 3 in behind via the aforementioned network. Whilst R84 and E372 are invariant among KMO sequences, Y382 and R386 are conserved but also present as F and K respectively - as is the case in human KMO. The interchanging of Y382 for a F, which would not be able to participate in the described hydrogen-bond network, is an interesting challenge for the hypothesis.

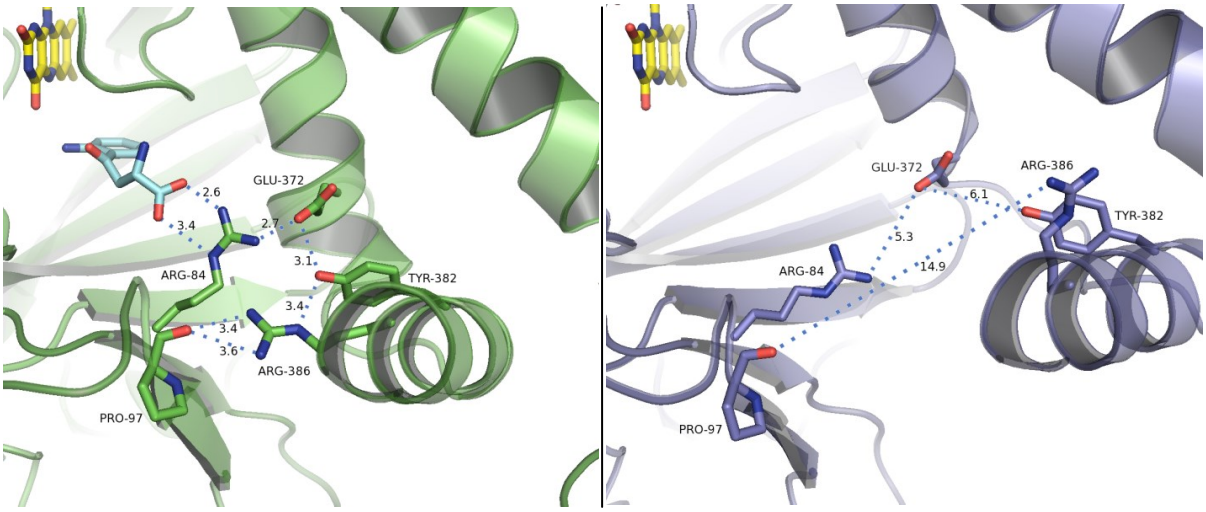


Figure 4.2. Cartoon representation of substrate binding pocket and start of C-terminal domain for substrate-bound (*left*) and substrate-free (*right*) *Pf*KMO structures.
Key residues that may be involved in the movement of domain 3 are shown in sticks in addition to FAD (yellow) and the substrate (cyan). Distances relating to the discussed potential hydrogen bond network are labelled in each structure.

Importantly the substrate-bound and substrate free enzyme forms displayed radically different lattices and crystal packing. As a result, the C-terminal differences described so far could have simply been an artefact caused by the different crystalline protein arrangements. C-terminal mutants (E372T, Y382F, R386K and R386T) were studied to

further characterize and validate the potential dynamics of domain 3 in *Pf*KMO, with the additional aim of determining whether a similar system could operate in human KMO.

4.2 Results

4.2.1 Crystallisation of Y382F dm2

The Y382F dm2 *Pf*KMO protein was expressed and purified as normal and used for h.d.v.d crystallization trials with L-Kyn bound. Crystals of substrate-free Y382F dm2 could not be obtained within the time constraints of the project. Large, hexagonal prism shaped substrate-bound crystals with rough dimensions of 0.2 x 0.2 x 0.5 mm^3 grew in 3 days and were immediately harvested for data collection. The final drop contained 1.0 μl of protein solution (15 mg/ml in 20 mM HEPES, 30 mM Na acetate, 1 mM DTT and 1 mM L-Kyn) with 1.0 μl of the precipitant solution (0.1 M HEPES pH 7.0, 10 % w/v PEG 4000, 10 mM KCl, 6 % v/v glycerol, 10 % v/v 2-propanol, 1 mM L-Kyn). Prior to flash-freezing in liquid nitrogen, 1 μL of cryoprotectant (mother liquor plus 20 % EG) was added to the drop before the crystal was transferred into a drop containing just the cryoprotectant for a few seconds.

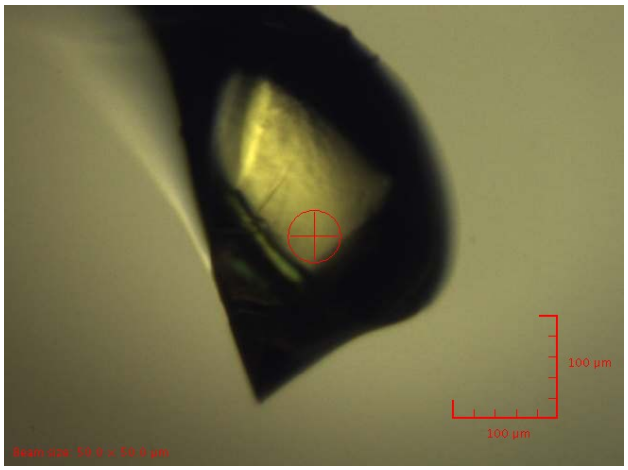


Figure 4.3. Close-up picture of the mounted Y382F crystal used for collection of the highest resolution dataset. The beam size and position is represented by the red circle and cross-hair.

Crystals were screened for diffraction quality at the Diamond synchrotron beamline I04-1, with several diffracting to a resolution around 3.4 Å. X-ray diffraction data were collected using an X-ray wavelength of 0.920 Å, a beam intensity of 90 % and an

exposure of 0.3 s with 0.3 ^{o} oscillations over a wedge of 180 ^{o}. The crystal (shown in figure 4.3) had similar lattice properties to native substrate bound crystals; space group $I4_122$ with mean unit cell parameters of a = b = 150.2, c = 273.6 Å, α = β = γ = 90.0 ^{o}. The dataset was 99.7 % complete to a resolution of 3.36 Å with R_{merge} = 0.095 overall (0.599 HS) and mean I/σI = 14.4 (3.6 HS). The full data collection statistics are tabulated in appendix section **3.3**.

For the calculation of structure factors, a suitable model was generated using molecular replacement (Molrep) with the previously solved monomeric substrate bound dm2 structure (see section **3.3**). Similarly, the Y382F mutant structure also contained one protein molecule in the asymmetric unit and a high solvent content of 83.8 %. A large displacement in the C-terminal domain was however evident and so the domain was rebuilt in a more open conformation as depicted in figure 4.4. The model was cycled through manual editing/building in Coot and TLS and restrained refinement in Refmac. There was clear density to place one molecule of FAD and one molecule of the substrate L-Kyn – both in the same position as seen before. The final model included residues 7-461 and is displayed in figure 4.5. It had R and R_{free} values of 0.247 and 0.277 respectively and an overall mean B factor of 103.1 $Å^2$ for 21461 significant reflections in the range of 65.93-3.36 Å. The rmsd of bond lengths and angles from the ideal were 0.005 Å and 0.900 ^{o} and all the residues were present in the most favoured or additionally allowed regions of a Ramachandran plot as determined using the Molprobity server. Full refinement statistics are tabulated in appendix section **3.3**.

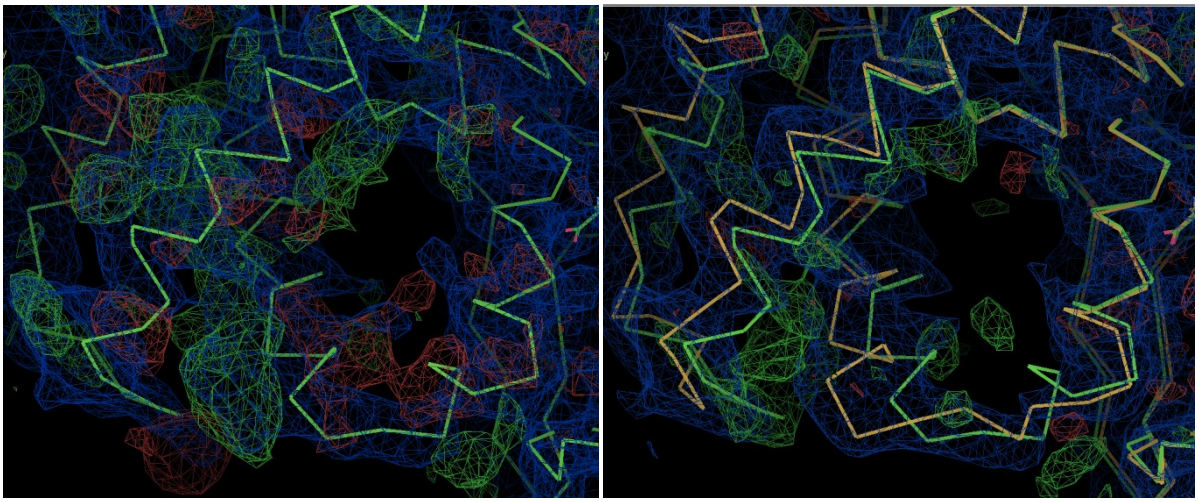


Figure 4.4. Y382F substrate bound $2F_o$-F_c density (blue) and F_o-F_c map (green/red) for C-terminal domain calculated: *(left)* with dm2 substrate bound model (green) and *(right)* with rebuilt more open loop position (yellow). Contouring, = 1.5, 3.0 σ

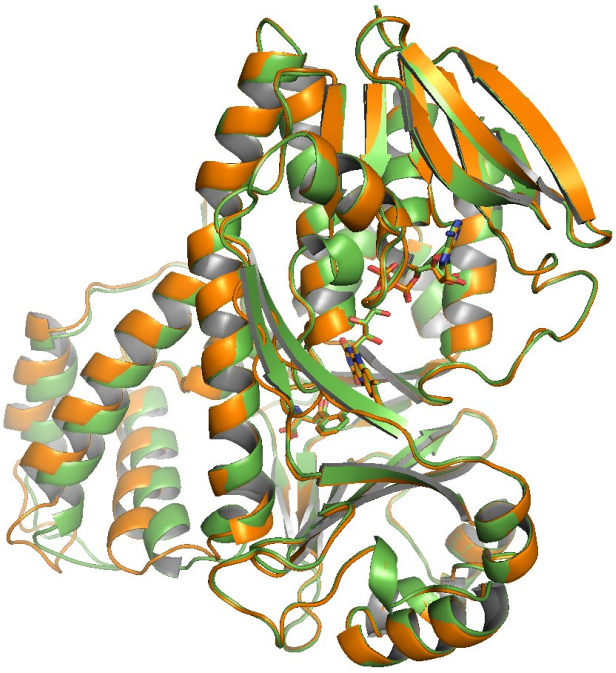


Figure 4.5. Cartoon representation of the superposed substrate-bound Y382F mutant (orange) and original dm2 (green) structures. FAD and the substrate L-Kyn are shown in sticks.

The first two domains of the Y382F substrate-bound structure, including the substrate and FAD-binding pockets, were identical to the original substrate-bound structure described in chapter **3.4**. The novel third domain however showed a relative displacement to a more open conformation reminiscent of the unbound *Pf*KMO structure from chapter **3.5**. Aligning the structure with the dm2 substrate-bound, chloride-bound and substrate-free structures revealed rmsd values of 0.8, 0.8 and 1.1 Å respectively (for protein C_α atoms). The overall substrate-bound structural alignment is displayed in figure 4.5, with a close-up section of the C-terminus including the original unbound structure shown in figure 4.6. The Y382F mutant C-terminal domain 3 displayed a mixed conformation similar to the dm2 chloride-bound structure described in section 3.5.3. The bottom of the long α-helix α-13 is pulled in towards the tail of the substrate but the C-terminus, from helix α-14 onwards, is swung out similar to the unbound structure. The extent of the swing (based on the position of the T381 C_α atom) is reduced at 3.8 Å compared to 6.1 Å in the fully open substrate-free conformation. The putative hydrogen-bond network described in figure 4.2 was speculated to link substrate binding to the closing of the C-terminal domain. Looking at the respective residue positions in the mutant substrate-bound structure in figure 4.7, this network is disrupted by the mutation of one its members leaving R386 out of hydrogen-bonding reach from E372

and P97. This may account for the observed opening of the C-terminal domain in the structure to help validate the hydrogen-bond network-hypothesis.

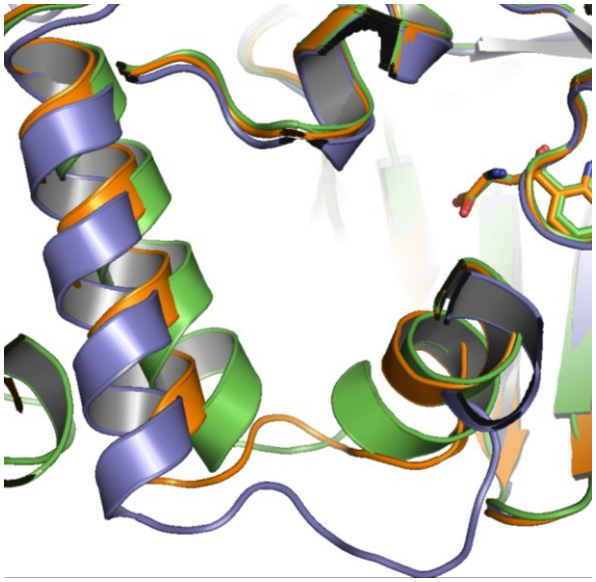


Figure 4.6. Representation of the mixed C-terminal conformation seen with the Y382F mutant. Includes a cartoon alignment of the Y382F substrate-bound structure (orange) with the dm2 substrate-bound (green) and substrate-free (violet) structures. The substrate L-Kyn is shown in sticks.

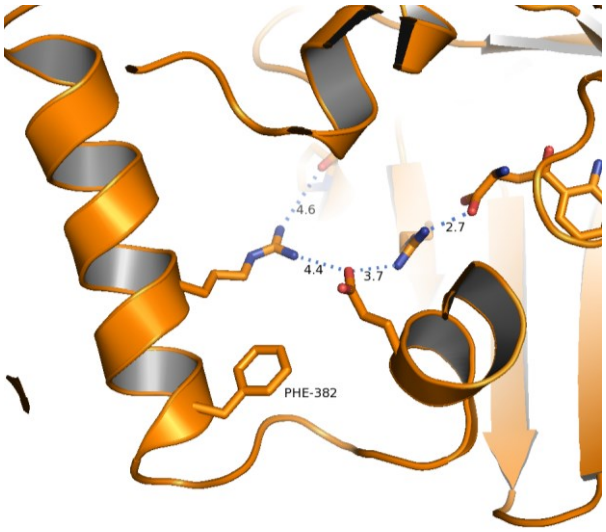


Figure 4.7. Altered positions of the residues previously implicated in a hydrogen-bonding network. The mutated residue Y382F is labelled along with the new distances between the residues.

To complement the structural data, the Michaelis-Menten kinetic constants were determined for the Y382F dm2 mutant enzyme. The K_m and k_{cat} values were 17.1 ± 1.0 μM and 12.7 ± 0.2 s^{-1} compared to 8.8 ± 1.5 μM and 8.9 ± 1.1 s^{-1} for dm2. Therefore the mutation did not have a massive effect on substrate binding and catalysis, despite the significant uncoupling of C-terminal domain movement from substrate binding.

4.2.2 Crystallisation of R386K dm2

The R386K dm2 *Pf*KMO protein was expressed and purified as normal and used for h.d.v.d crystallization trials with L-Kyn bound. The final drop contained 1.0 μl of protein solution (15 mg/ml in 20 mM HEPES, 30 mM Na acetate, 1 mM DTT and 1 mM L-Kyn) with 1.0 μl of the precipitant solution (0.08 M HEPES pH 7.0, 8 % w/v PEG 4000, 10 mM NaCl, 6 % v/v glycerol, 10 % v/v 2-propanol). Crystals shaped like hexagonal prisms with rough dimensions of 0.2 x 0.3 x 0.6 mm^3 grew in 2-3 days and were harvested for data collection after 10 days. Prior to flash-freezing in liquid nitrogen, 1.5 μl of cryoprotectant (mother liquor + 20 % EG) was added to the crystallisation drop for a couple of mins. The crystal was then transferred into a 2 μl drop of cryoprotectant for a few seconds.

Crystals were screened for diffraction quality at the Diamond synchrotron beamline I02 with several diffracting to a resolution around 3.4 Å. X-ray diffraction data were collected using an X-ray wavelength of 0.979 Å, a beam intensity of 80 % and an exposure of 0.3 s/frame of data, with 0.3 ^{o} oscillations over a rotation of 210 ^{o}. The crystal (displayed in figure 4.8) again had similar lattice properties to substrate bound wild type crystals; space group $I4_122$ with mean unit cell parameters of a = b = 149.7, c = 273.5 Å, α = β = γ = 90.0 ^{o}. The dataset was 100.0 % complete to a resolution of 3.30 Å with R_{merge} = 0.070 overall and 0.545 in the highest resolution shell. The full data collection statistics are tabulated in appendix section **3.3**.

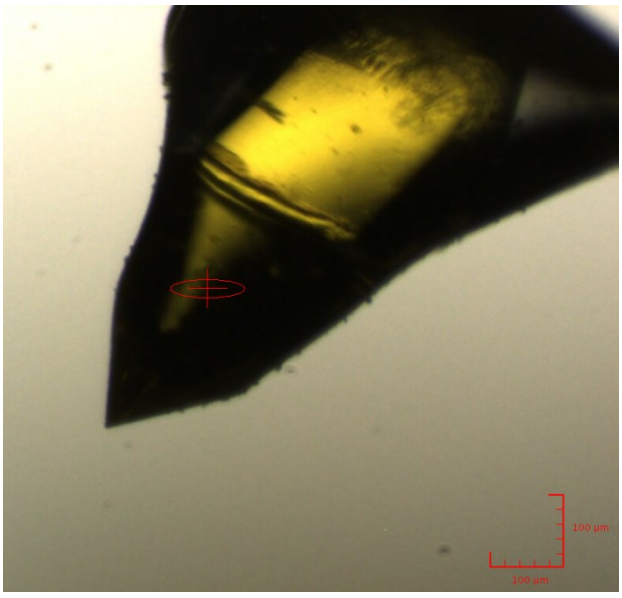


Figure 4.8. Close-up picture of the mounted R386K crystal used for collection of the highest resolution dataset. The beam size (100 x 25 μm) and position is represented by the red circle and cross-hair.

For the calculation of structure factors, a suitable model was generated using molecular replacement (Molrep) with the previously solved monomeric substrate bound structure (see section **3.3**). Few differences emerged with one protein molecule present in the asymmetric unit and a solvent content of 83.7 %. The residues at the start of domain 3 were again clearly displaced and rebuilt in a more open conformation (see figure 4.9). The model was cycled through manual editing/building in Coot and TLS and restrained refinement in Refmac. There was clear density to place one molecule of FAD and one molecule of the substrate L-Kyn – both in the same position as seen before. The final model included residues 7-461 with R and R_{free} values of 0.247 and 0.259 respectively and an overall mean B factor of 108.3 for 22568 significant reflections in the range of 53.96-3.30 Å. The rmsd of bond lengths and angles from the ideal were 0.005 Å and 0.953 ^{o} and all the residues were present in the most favoured or additionally allowed regions of a Ramachandran plot as determined using the Molprobity server. Full refinement statistics are tabulated in appendix section **3.3**.

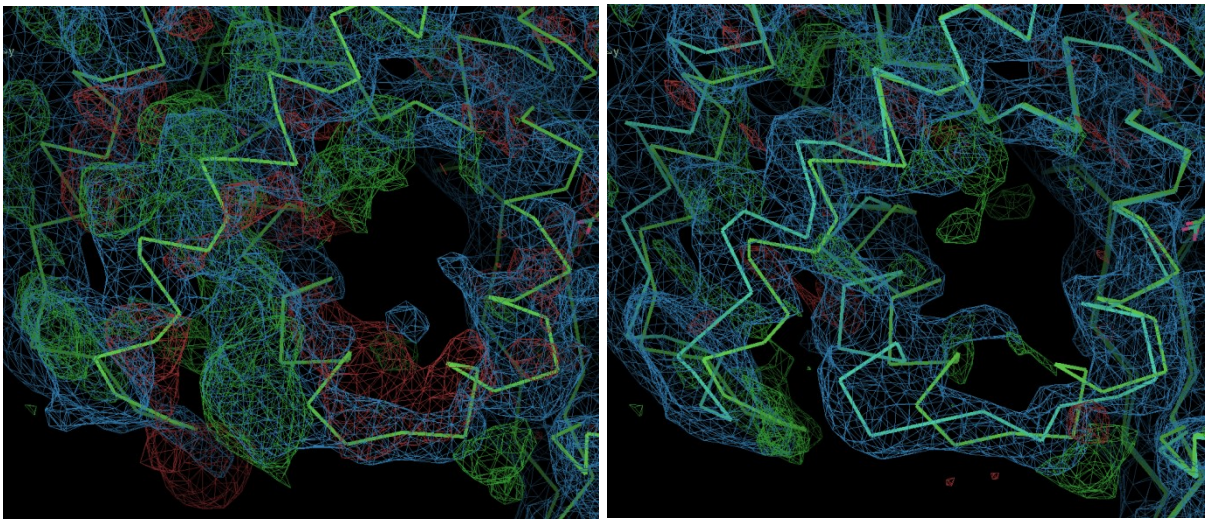


Figure 4.9. R386K substrate bound $2F_o$-F_c density (blue) and F_o-F_c map (green/red) for C-terminal domain calculated: *(left)* with dm2 substrate bound model (green) and *(right)* with rebuilt more open loop position (cyan). Contouring, = 1.5, 3.0 σ

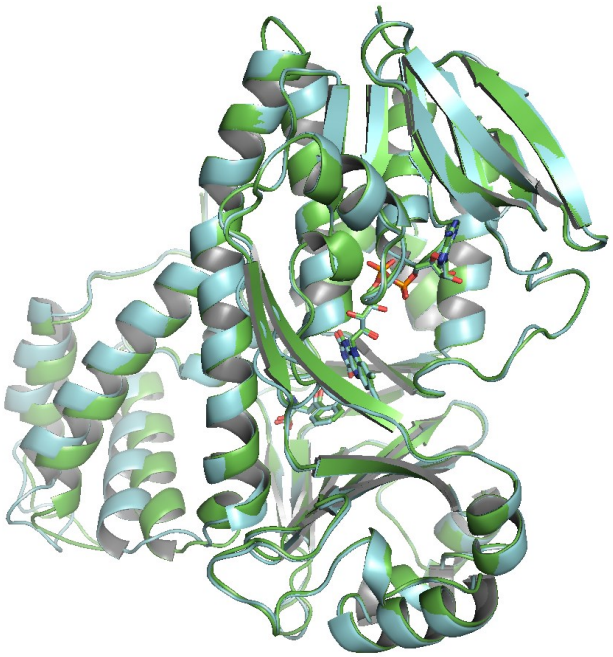


Figure 4.10. Cartoon representation of the aligned substrate-bound R386K mutant (cyan) and original dm2 (green) structures. FAD and the substrate L-Kyn are shown in sticks.

The R386K structure was very similar to that of Y382F with unchanged domains 1 and 2 and the same mixed C-terminal conformation. The rmsd of the substrate-bound structure superposed with that of Y382F, dm2, chloride-bound dm2 and substrate-free dm2 is: 0.2, 0.8, 0.7 and 1.0 Å respectively (for protein C_α atoms). The global superposition of the original substrate-bound dm2 structure is displayed in figure 4.10. The mixed C-terminal conformation is displayed more closely in figure 4.11 with the respective positions of the hydrogen-bond network residues indicated in figure 4.12. It is clear that the lysine residue introduced could not maintain the interactions involving R386 that closed the C-terminal domain upon substrate binding. Looking back to figure 4.2 showing the proposed native hydrogen-bond network, the R386K structure suggests that the interactions described from the low-resolution structure may be accurate. Lys is unable to coordinate the two separate hydrogen bonds predicted with the side-chain of Y382 and with the backbone carbonyl of P97. As a result the residue is left stranded in the structure at around 5.0 and 6.7 Å from Y382 and P97 respectively. The Michaelis-Menten kinetic constants were determined to be slightly elevated compared to the dm2 protein and were essentially identical to those for the Y382F mutant within the error limits (see table 4.1 for a complete summary of the data).

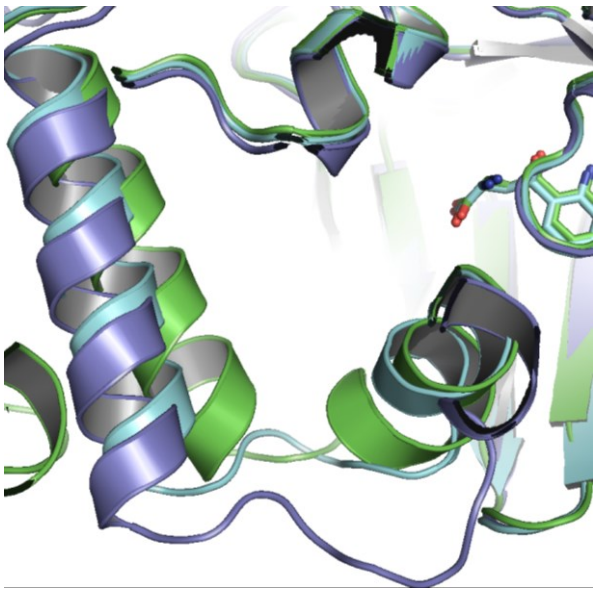


Figure 4.11. Representation of the mixed C-terminal conformation seen with the R386K mutant. Includes a cartoon alignment of the R386K substrate-bound structure (cyan) with the dm2 substrate-bound (green) and substrate-free (violet) structures. The substrate L-Kyn is shown in sticks.

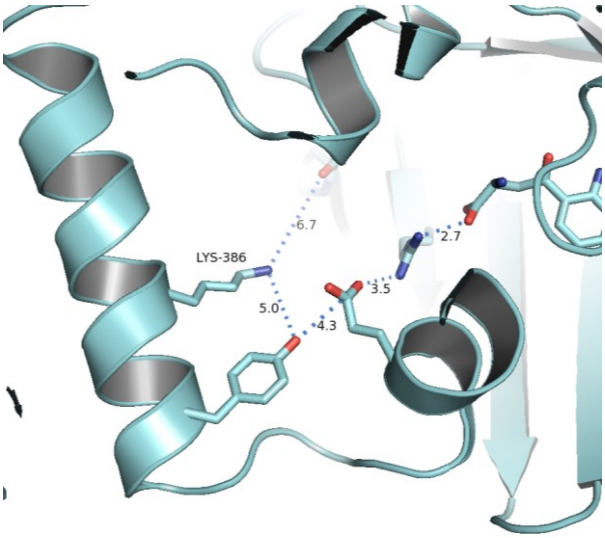


Figure 4.12. Altered positions of the residues previously implicated in a hydrogen-bonding network. The mutated residue R386K is labelled along with the new distances between the residues.

For a negative control, crystals of unbound R386K were generated from a 1:1 μl h.d.v.d drop of protein solution (12 mg/ml in 20 mM HEPES pH 6.8, 150 mM Na acetate, 1 mM DTT) and reservoir solution containing: 0.1 M HEPES pH 7.0, 12 % w/v PEG 4K, 10 % v/v glycerol, 10 mM Na tartrate, 10 % v/v 2-propanol. The crystal of rough dimensions 0.1 x 0.1 x 0.3 mm^3 grew in 3-5 days and was harvested after 6 days. Cryoprotection involved addition of 1.5 μl cryoprotectant (mother liquor plus 20 % EG) prior to transfer into the cryoprotectant for a few seconds. Diffraction data were collected with an exposure of 0.3 s per 0.3 ^{o} rotation over a total of 135 ^{o} at beamline I04 using a wavelength of 0.980 Å and beam intensity of 35 %. The averaged unit cell had dimensions of a = 105.2, b = 133.9 and c = 188.5 Å packed in the orthorhombic space group $P2_122_1$. Full data collection and refinement statistics are displayed in appendix section **3.4**. The structure was solved using molecular replacement with the dm2 unbound structure from section **3.5.2**. As previously seen there were four molecules in the asymmetric unit each with one molecule of FAD bound. Key refinement values include: Resolution = 45.00-2.63 Å,

R/R_{free} = 0.225/0.241, mean B-factor = 40.6 $Å^2$ and no Ramachandran plot outliers. The structure was virtually identical to the dm2 unbound structure with an overall rmsd of 0.1 Å, therefore confirming that the R386K mutation did not alter the general protein structure.

4.2.3 Crystallisation of R386T dm2

The R386T dm *Pf*KMO protein was expressed and purified as normal and used for h.d.v.d crystallization trials with L-Kyn bound. The final drop contained 1.0 µl of protein solution (13 mg/ml in 20 mM HEPES pH 6.8, 20 mM Na acetate, 1 mM DTT and 1.5 mM L-Kyn) with 1.0 µl of the precipitant solution (0.1 M HEPES pH 7.0, 6 % w/v PEG 4000, 10 mM NaCl, 6 % v/v glycerol, 10 % v/v 2-propanol, 1 mM L-Kyn). Crystals shaped like hexagonal prisms with rough dimensions of 0.1 x 0.2 x 0.3 mm^3 grew in 2-3 days and were harvested for data collection after 12 days. Prior to flash-freezing in liquid nitrogen, crystals were briefly soaked in successive drops of 10 % cryoprotectant (mother liquor plus 10 % EG) and then 20 % cryoprotectant (mother liquor plus 20 % EG).

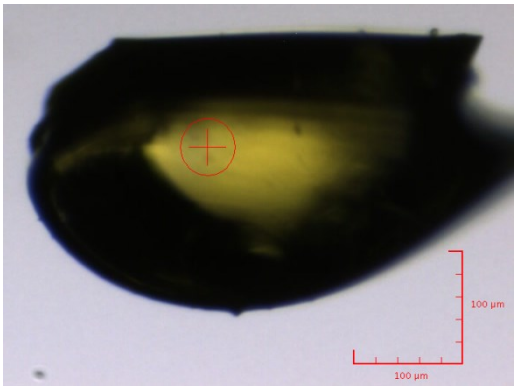


Figure 4.12. Close-up picture of the mounted R386T crystal used for collection of the highest resolution dataset. The beam size (50 x 50 µm) and position is represented by the red circle and cross-hair.

Crystals were screened for diffraction quality at the Diamond synchrotron beamline I03 with several diffracting to a resolution of around 3.4 Å. X-ray diffraction data were collected using: an X-ray wavelength of 0.976 Å, a beam intensity of 50 % and an exposure of 0.25 s/frame, with 0.25 ^{o} oscillations over a wedge of 87.5 ^{o}. The crystal again had similar lattice properties to substrate bound wild type crystals; space group $I4_122$ with mean unit cell parameters of a = b = 149.7, c = 272.1 Å, α = β = γ = 90.0 ^{o}. The

dataset was 99.9 % complete to a resolution of 3.40 Å with R_{merge} = 0.088 overall (0.644 HS) and mean I/σI = 9.9 (2.3 HS). The full data collection statistics are tabulated in appendix section **3.3**.

For the calculation of structure factors, a suitable model was generated using molecular replacement (Molrep) with the previously solved monomeric substrate bound structure (see section 3.3). Few differences emerged with one protein molecule present in the asymmetric unit and a solvent content of 83.6 %. The residues at the start of domain 3 were again clearly displaced and removed from the initial model. The model was cycled through manual editing/building in Coot and TLS and restrained refinement in Refmac. There was significant density to place one molecule of FAD and one molecule of L-Kyn into the structure. The final model included residues 7-461 with R and R_{free} values of 0.260 and 0.270 respectively, a figure of merit of 0.721 and an overall mean B-factor of 135.0 $Å^2$ for 20512 significant reflections in the range of 65.68-3.40 Å. The rmsd of bond lengths and angles from the ideal were 0.005 Å and 0.981 ^{o} and all the residues were present in the most favoured or additionally allowed regions of a Ramachandran plot as determined using the Molprobity server. Full refinement statistics are tabulated in appendix section **3.3**.

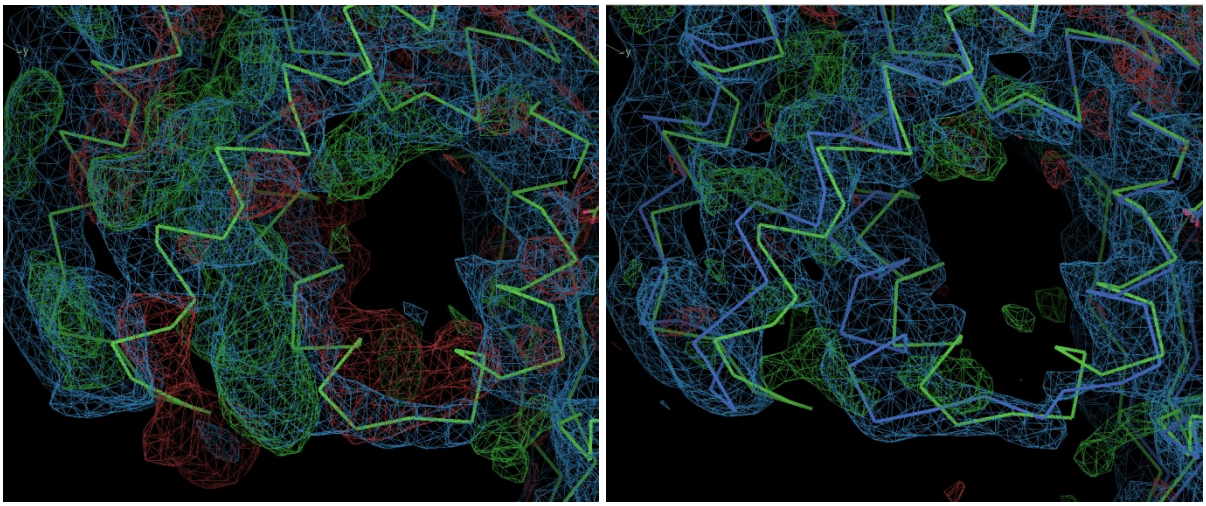


Figure 4.13. R386T substrate bound $2F_o$-F_c density (blue) and F_o-F_c map (green/red) for C-terminal domain calculated: *(left)* with dm2 substrate bound model (green) and *(right)* with rebuilt more open loop position (blue). Contouring, = 1.5, 3.0 σ

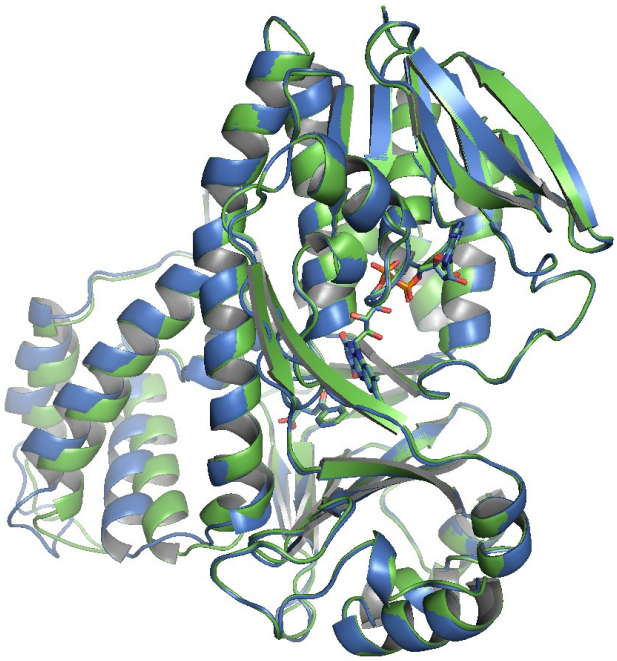


Figure 4.14. Cartoon representation of the aligned substrate-bound R386T mutant (blue) and original dm2 (green) structures. FAD and the substrate L-Kyn are shown in sticks.

An alignment of the R386T substrate-bound structure with the original dm2 substrate-bound structure is displayed in figure 4.14. The R386T mutant showed a similar mixed structural conformation to the previous Y382F and R386K substrate-bound structures as well as to the dm2 chloride-bound structure with rmsd values of 0.2, 0.2 and 0.8 Å respectively (for protein C_α atoms). The mixed C-terminal conformation is shown in comparison with the previously solved dm2 substrate-bound and substrate-free structures in figure 4.15. As expected, the R386T mutation disrupted the putative C-terminal hydrogen-bonding network, with the adjusted positions of the participating residue displayed in figure 4.16. The mutation caused no further changes to those seen with the R386K substitution. Despite this the Michaelis-Menten kinetic constants were found to reveal a slightly less efficient enzyme. The K_m and k_{cat} values for R386T were 63.9 ± 2.4 μM and 15.2 ± 1.3 s^{-1} compared to 17.0 ± 1.4 μM and 12.0 ± 0.9 s^{-1} for R386K.

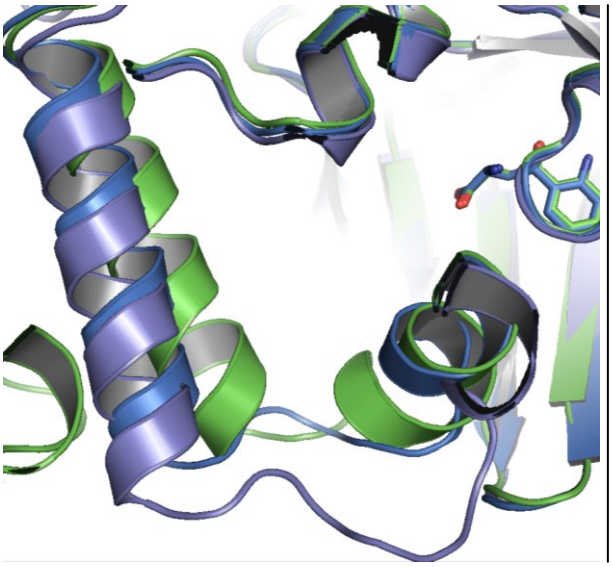


Figure 4.15. Representation of the mixed C-terminal conformation seen with the R386T mutant. Includes a cartoon alignment of the R386T substrate-bound structure (blue) with the dm2 substrate-bound (green) and substrate-free (violet) structures. The substrate L-Kyn is shown in sticks.

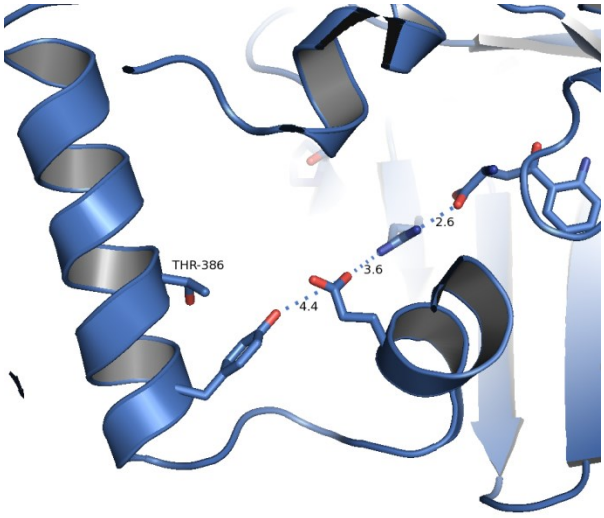


Figure 4.16. Altered positions of the residues previously implicated in a hydrogen-bonding network. The mutated residue R386T is labelled along with the new distances between the residues.

To confirm that the structural differences caused by the R386T mutation were specific to the binding of the substrate L-Kyn, a substrate-free structure was solved. Due to time constraints, only the chloride-bound crystal form could be obtained. Crystals with rough dimensions of 0.1 x 0.1 x 0.3 mm^3 grew in 3-5 days and were harvested after 6 days. They were generated from a 2.0:2.5 µl h.d.v.d drop of protein solution (14 mg/ml in 20 mM HEPES pH 6.8, 20 mM Na acetate, 1 mM DTT) and reservoir solution containing: 0.1 M HEPES pH 7.0, 9 % w/v PEG 4K, 10 % v/v glycerol, 10 mM NaCl, 10 % v/v 2-propanol. For cryoprotection 1.0 µl of cryoprotectant (mother liquor plus 20 % EG) was added to the crystal drop prior to transfer into cryoprotectant for a few seconds. A complete dataset was collected with an exposure of 0.3 s per 0.3 ^{o} rotation over a total of 150 ^{o} at beamline I03 using a wavelength of 0.976 Å and beam intensity of 35 %. The averaged unit cell had dimensions of 105.4, 134.0 and 189.5 Å packed in the orthorhombic space group $P2_122_1$. Full data collection and refinement statistics are displayed in appendix section **3.4**. The data was solved using molecular replacement with the dm2 chloride-bound structure from section **3.5.3**. As previously seen there were four molecules in the

asymmetric unit each with one molecule of FAD bound. Key refinement values include: Resolution = 44.71-2.52 Å, R/Rfree = 0.223/0.241, mean B-factor = 43.7 $Å^2$ and there were no Ramachandran plot outliers. The structure was identical to the corresponding dm2 structure confirming that the mutation did not affect the overall enzyme structure in the absence of substrate. The rmsd of the superimposed R386T and original dm2 chloride-bound structures is 0.1 Å for aligned C_α atoms.

4.2.4 Crystallisation of E372T dm2

The E372T mutation was also successfully introduced however unfortunately an additional unwanted mutation was gained of L367R. There was not enough time to repeat the cloning procedure so, as the site of the mutation seemed to be a relatively innocuous surface location in the previously solved structures, the mutant was characterised and crystallised. The E372T (L367R) dm2 *Pf*KMO protein was expressed and purified as normal, although precipitation during the final concentration step required the salt concentration to be increased. The Michaelis-Menten kinetic constants K_m and k_{cat} were determined to be very similar to those for dm2 at 10.3 ± 0.7 μM and 8.7 ± 0.1 s^{-1} respectively. Therefore it was likely that the additional L367R mutation did not cause any significant structural changes to the enzyme.

The pure protein was used for h.d.v.d crystallization trials with L-Kyn bound. The final drop contained 1.00 μl of protein solution (10 mg/ml in 20 mM HEPES, 50 mM Na acetate, 1 mM DTT and 1 mM L-Kyn) with 0.75 μl of the precipitant solution (0.1 M HEPES pH 7.0, 7 % w/v PEG 4000, 10 mM NaCl, 10 % v/v glycerol, 10 % v/v 2-propanol, 2 mM L-Kyn). Hexagonal prism shaped crystals with rough dimensions of 0.1 x 0.1 x 0.2 mm^3 grew in 3 days and were harvested for data collection after 1 week. Prior to flash-freezing in liquid nitrogen, 1 μL of cryoprotectant (mother liquor plus 20 % EG) was added to the drop before the crystal was transferred into the cryoprotectant solution for a few seconds.

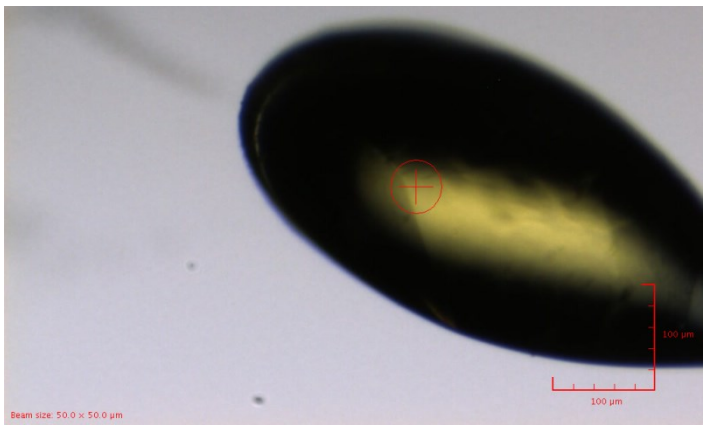
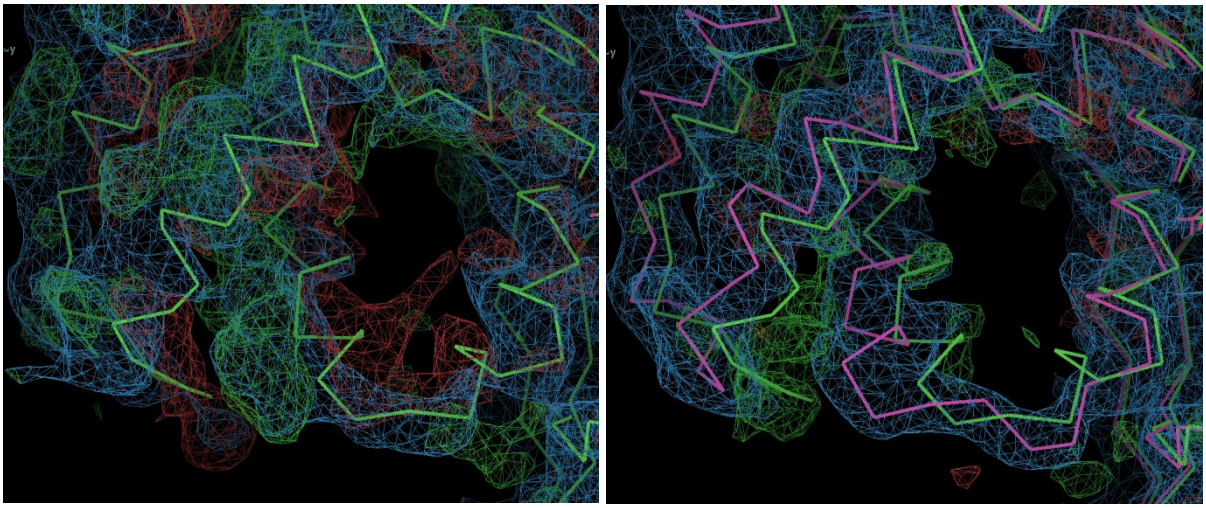


Figure 4.17. Close-up picture of the mounted crystal used for collection of the best resolution dataset. The beam size and position is represented by the red circle and cross-hair.

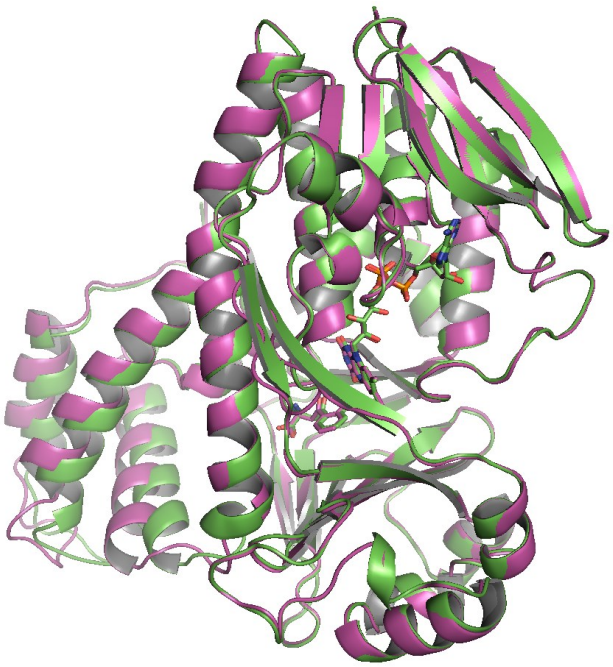
Crystals were screened for diffraction quality at the Diamond synchrotron beamline I03 and several diffracted to a resolution around 3.3 Å. With an X-ray wavelength of 0.976 Å, the best diffraction dataset was collected with a beam intensity of 50 % and an exposure of 0.3 s per 0.3 º oscillation over a 150 º wedge. The crystal (shown in figure 4.17) had similar lattice properties to native substrate bound crystals; space group $I4_122$ with mean unit cell parameters of a = b = 149.6, c = 270.7 Å, α = β = γ = 90.0 º. With the best resolution for substrate-bound data obtained to date of 3.24 Å, the dataset was 99.9 % complete with R_{merge} = 0.088 overall (0.691 in the highest resolution shell, HS) and mean I/σI = 14.2 (2.8 HS). The full data collection statistics are tabulated in appendix section **3.3**.

The structure was solved by molecular replacement using the previously solved monomeric substrate bound dm2 structure (see section **3.3**) as the template model. The E372T (L367R) mutant structure also contained one protein molecule in the asymmetric unit giving a high solvent content of 83.5 %. A large displacement in the C-terminal domain was however evident and so the domain was rebuilt in a more open conformation as depicted in figure 4.18. The model was cycled through manual editing/building in Coot and TLS and restrained refinement in Refmac. There was clear density to place one molecule of FAD and one molecule of the substrate L-Kyn – both in the same position as seen before. The final model displayed in figure 4.19 includes residues 7-461 with R and R_{free} values of 0.230 and 0.256 respectively for 23527

significant reflections in the range of 44.70-3.24 Å. In addition, the FOM was 0.786 and there was an overall mean B-factor of 96.7 $Å^2$. The rmsd of bond lengths and angles from the ideal were 0.005 Å and 0.966 ° and all the residues were present in the most favoured or additionally allowed regions of a Ramachandran plot as determined using the Molprobity server. Full refinement statistics are tabulated in appendix section **3.3**.



↑ ***Figure 4.18***. $2F_o$-F_c density (blue) for E372T (L367R) and F_o-F_c (green/red) calculated based on model the of substrate-bound dm2 structure (*left,* green ribbon) and of the final E372T (L367R) substrate-bound structure (*right,* magenta ribbon) with translocated C-terminal domain. Contouring, = 1.5, 3.0 σ



← ***Figure 4.19.*** Cartoon representation of final E372T (L367R) model (magenta) aligned with that of substrate bound dm2 *Pf*KMO (green). FAD and the substrate L-Kyn are shown in sticks.

Mutation of the highly conserved E372 to a T caused similar structural changes seen with the other C-terminal mutations in the presence of the substrate. The rmsd of the backbone C_α atoms compared to dm2, Y382F, R386K, R386T, chloride-bound dm2 and unbound dm2 (substrate-bound unless otherwise stated) are 0.8, 0.3, 0.3, 0.3, 0.7 and 1.1 Å respectively. As can be seen in figure 4.20, the C-terminus adopts the mixed conformation in the presence of L-Kyn with a relative swing of around 3.7 Å. As seen previously, the proposed hydrogen-bond network breaks down as shown in figure 4.21 providing further evidence for its role linking substrate-binding to the movement of the C-terminal domain. For this mutant however, it cannot be ruled that the additional L367R mutation by itself may have contributed to the uncoupling of C-terminal closing from substrate binding.

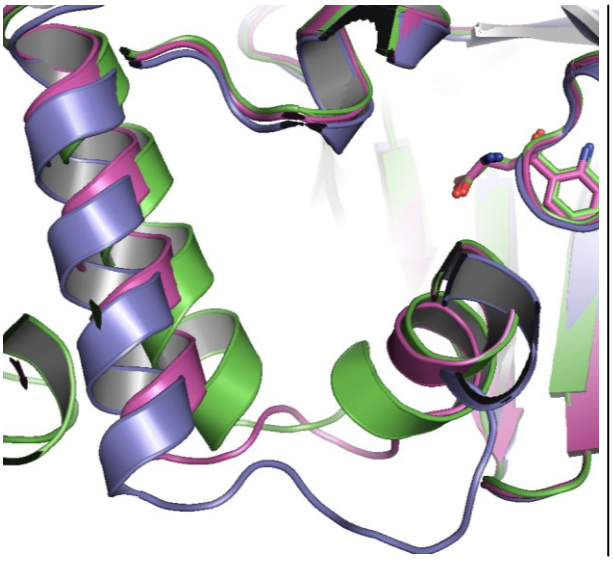


Figure 4.20. Cartoon structural representation of the mixed C-terminal conformation observed in the E372T (L367R) substrate-bound structure (magenta). The original dm2 substrate-bound and substrate-free structures are shown in green and violet with the substrate represented by sticks.

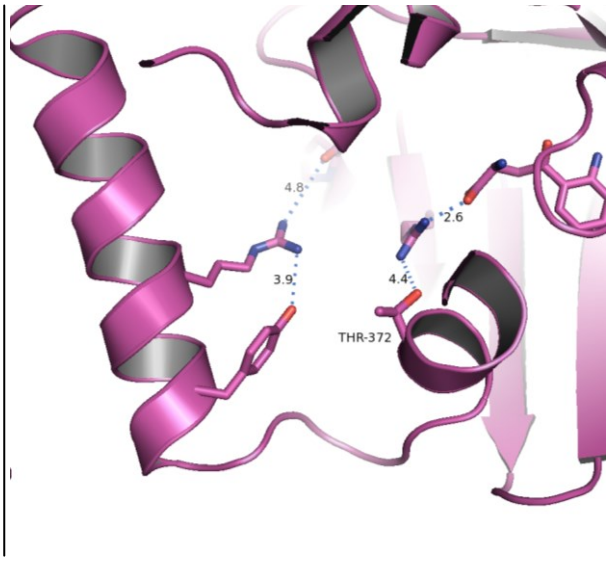


Figure 4.21. Breakdown of the putative C-terminal hydrogen-bond network in the E372T (L367R) mutant structure. The new distances between the participating residues as well as the mutated residue are labelled.

Interestingly, despite generally possessing smaller crystal dimensions than the previously seen strongest diffracting substrate-bound crystals, crystals of substrate-bound E372T (L367R) *Pf*KMO showed relatively stronger diffraction power. Looking at the structure in

more detail it is clear that the additional L367R mutation occurred in the heart of a crystal contact in the substrate-bound lattice arrangements. A comparison of this site in the mutant shown next to that in the original dm2 structure is shown in figure 4.22. The new arginyl side-chain fills a gap in the centre of the contact, which was presumably filled by solvent. Moreover this addition significantly displaced the side chain of R296 to a new conformation where it planar stacks with H185 and forms a hydrogen bond with Q242 from a symmetry-related protein molecule.

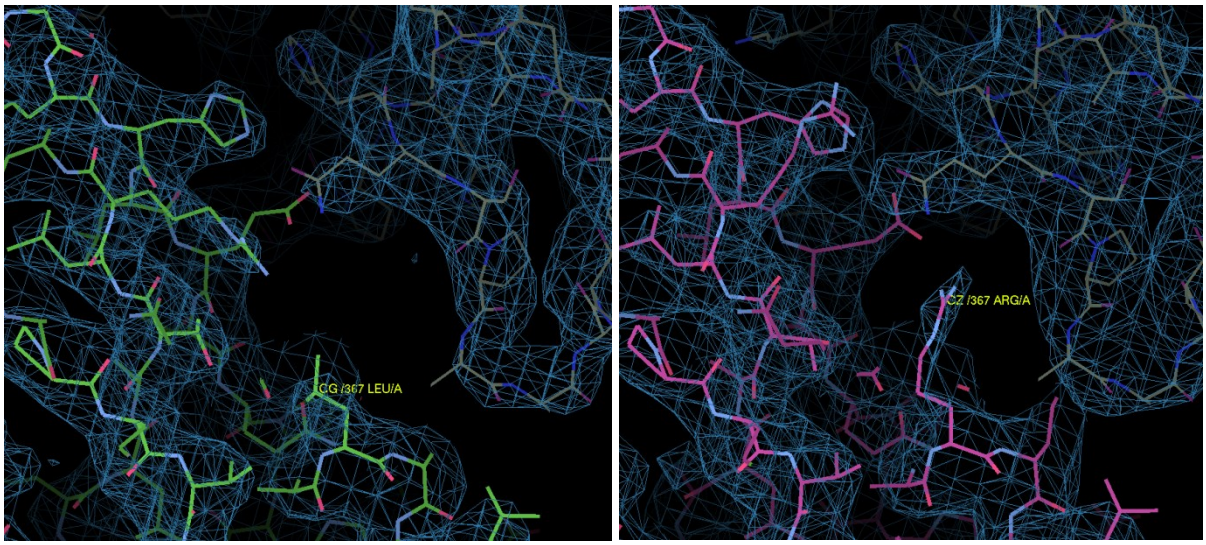


Figure 4.22. Improved crystal contact caused by the unintended L367R mutation. *Left,* crystal contact in substrate-bound dm2 *Pf*KMO structure and *right*, corresponding contact in substrate-bound E372T (L367R) mutant structure. Refined $2F_o$-F_c electron density shown in blue was contoured at 1.5 σ.

Due to the non-specific introduction of the L367R mutation, a substrate-free structure was also solved to act as a negative control. Due to time constraints, only the chloride-bound crystal form could be obtained. Crystals with rough dimensions of 0.1 x 0.1 x 0.2 mm^3 grew in 3-5 days and were harvested after 9 days. They were generated from a 2.0:1.5 μl h.d.v.d drop of protein solution (14 mg/ml in 20 mM HEPES pH 6.8, 20 mM Na acetate, 1 mM DTT) and reservoir solution containing: 0.1 M HEPES pH 7.0, 8 % w/v PEG 4K, 10 % v/v glycerol, 10 mM NaCl, 10 % v/v 2-propanol. For cryoprotection, the crystal was briefly soaked in successive drops of 10 % cryoprotectant (mother liquor plus 10 % EG) and 20 % cryoprotectant (mother liquor plus 20 % EG). A 99.9 % complete dataset

was collected from a 150 ^{o} segment at beamline I03. An exposure of 0.25 s per 0.25 ^{o} oscillation using a wavelength of 0.976 Å and beam intensity of 30 % were used. The averaged unit cell had dimensions of 104.7, 133.9 and 189.1 Å packed in the orthorhombic space group $P2_122_1$. Full data collection and refinement statistics are displayed in appendix section **3.4**. The data was solved using molecular replacement with the dm2 chloride-bound structure from section **3.5.3**. As previously seen there were four molecules in the asymmetric unit each with one molecule of FAD bound. Key refinement values include: Resolution = 50.51-2.32 Å, R/R_{free} = 0.221/0.242, mean B-factor = 45.4 $Å^2$ and there were no Ramachandran plot outliers. The structure was identical to the corresponding dm2 structure confirming that the mutation did not affect the overall enzyme structure other than the small localised alterations surrounding the L367R mutation. The average chain rmsd of the superimposed E372T (L367R) and original dm2 chloride-bound structures is 0.3 Å.

4.3 Discussion and conclusions

In chapter **3** the substrate-bound, chloride-bound and substrate-free dm2 *Pf*KMO structures were solved revealing a novel C-terminal domain present in three different conformations – in, mixed and out respectively. From the substrate-bound structure a hydrogen bond network appeared to link the bound substrate to the closed 'in' C-terminal conformation (see figure 4.2). This involved the carboxyl of the substrate, the side chains of R84, E372, Y382 and R386 as well as the backbone carbonyl of P97. Due to the poor resolution of this structure however, the exact positions of the participating residues is questionable. Therefore in this chapter mutants involving these residues were generated and tested to assess the accuracy/validity of the proposed hydrogen-bond network mechanism in the domain movement.

Each mutant was briefly characterised and crystallised with the results summarised in table 4.1. Each substrate-bound mutant structure revealed a mixed C-terminal conformation reminiscent of the chloride-bound dm2 structure from chapter **3.5.3**. Despite this the mutants all crystallised in the same high solvent tetragonal form as native substrate-bound crystals. This confirmed the relevance of the different C-terminal

structural conformations observed showing they were dependent solely on the protein rather than the crystal packing. The structures showed that mutation of each residue disrupted the interactions that closed the C-terminal domain after substrate binding with the native enzyme. These results correlate well to the putative hydrogen-bond network described in section **4.1** as the mutated residues would be unable to fulfil the precise interactions speculated in the native substrate-bound structure. This includes position 386 where the substituted lysine would be unable to coordinate both the two interactions involving the native arginine residue with Y382 and the backbone carbonyl of P97. Therefore whilst the exact mechanism that closes the C-terminal domain after substrate-binding may not be detailed without a high-resolution substrate-bound structure, the most likely explanation involves the formation of the aforementioned hydrogen-bond network.

Mutant	**K_m (μM)**	**k_{cat} (s^{-1})**	**k_{cat}/K_m (s^{-1}/μM)**	**C-term. conformation +L-Kyn/+Cl/unbound[(1)]**	**C-term. swing (Å) +L-Kyn/+Cl/unbound[(2)]**
dm2	8.8 ± 1.5	8.9 ± 1.1	1.0	IN/MIXED/OUT	0.0/5.1/6.1
E372T dm2*	10.3 ± 0.7	8.7 ± 0.1	0.8	MIXED/MIXED/n.d	3.7/4.9/n.d
Y382F dm2	17.1 ± 1.0	12.7 ± 0.2	0.7	MIXED/n.d/n.d	3.8/n.d/n.d
R386K dm2	17.0 ± 1.4	12.0 ± 0.9	0.7	MIXED/n.d/OUT	4.2/n.d/6.1
R386T dm2	63.9 ± 2.4	15.2 ± 1.3	0.2	MIXED/MIXED/n.d	4.3/5.2/n.d

Table 4.1. Summary of the kinetic and structural observations from each of the C-terminal mutants compared to the 'native' dm2 *Pf*KMO enzyme.
*E372T also contained an additional L367R mutation. [1] C-terminal conformations observed for the substrate-bound (+L-Kyn), chloride-bound (+Cl) and unbound structures based on those observed with dm2 in chapter **3**. [2] Swing distances represented the separation of the C_α atoms of T381 from aligned structures of the respective mutant and the original substrate-bound dm2 structure.

Barring E372T, the other mutations all caused the values of K_m and k_{cat} to increase by varying extents. The Y382F and R386K mutants both displayed a two-fold increase in the K_m compared to dm2, whereas the less conservative R386T mutant showed around a seven-fold increase. With the large distance separating the residues from the substrate binding site, these changes indicated that the native residues were in some way linked to the process of substrate binding. Thus, taking the structural evidence into account, the more open C-terminal conformations seen in all four mutant structures could account for the weakened K_m values. The ability of the C-terminus to shut after substrate binding

would prevent leakage back out to bulk solvent and help to keep the equilibrium towards binding into the enzyme. The further raised K_m seen with the R386T mutant comes despite a similar structural arrangement and C-terminal swing to the other mutants. This value can be accounted for by the following theories about the residue at position 386:

I) The long arginine or lysine side-chains provided a steric block preventing substrate egress after binding.

II) A basic residue may be involved in initially attracting the substrate into the entry tunnel preceding the binding pocket, specifically by interacting with the carboxyl group of L-Kyn.

As was previously mentioned the mutants also displayed increased k_{cat} values with Y382F and R386K slightly higher at around 12-13 s^{-1} and R386T 1.7-times higher at 15 s^{-1}. Interestingly Crozier-Reabe *et al.,* reported that product release was the rate-limiting step in the reaction catalysed by *Pf*KMO [1]. This suggested that the more open C-terminal positions seen with the mutants allowed faster release of the product to cause the increased reaction rate. The exception to the rule was the E372T mutant, which despite a more open C-terminus had a similar k_{cat} to dm2 of around 9 s^{-1}. Firstly it could be possible that the lower than expected reaction rate was caused by the additional L367R mutation. A more complex hypothesis could be that the negative E372 plays some part in ejecting the product – possibly by electrostatic repulsion of the substrate's carboxylic group – and so its removal may have slightly slowed the rate of product release to compensate the effects of the more open C-terminal domain. The same hypothesis but in reverse could apply to the stronger than expected K_m of the mutant for substrate binding, which again was similar to dm2 within the limits of error. Future work interrogating the separate catalytic steps by stopped-flow spectrophotometry would help to further elucidate the effects of the mutations.

Understanding of the dynamics of the C-terminus and substrate binding in *Pf*KMO seems to be developing well, however how does this translate to other KMOs including the human enzyme? As was previously mentioned R84 and E372 are strongly conserved across all KMO sequences, however Y382 and R386 are found as a mixture of Y/F and of R/K respectively. Interestingly human KMO possesses both the F and the K variations at

the two positions. This creates a problem as the data from the *Pf*KMO C-terminal mutants suggested that either of the Y382F or the R386K mutations separately suppressed the conformational change that closed the back of the substrate-binding pocket. These observations clearly cast doubt over the relevance of the C-terminal closing mechanism to the human enzyme.

Models of the structure of human KMO constructed based on the dm2 *Pf*KMO structures are shown in appendix section **2**. These show that from its primary sequence the human C-terminus may initially form the same structural domain to the novel C-terminus displayed by the bacterial enzyme. The similarity breaks down at around residue 330 located at the bottom of the penultimate helix (α-16) where the human sequence contains the predicted trans-membrane region. In addition key hydrophobic residues that maintain the C-terminal fold and link it to the main protein scaffold appear to be conserved in human KMO. Although the models are largely biased towards the template, they show that it is possible that human KMO may have a similar entry tunnel enclosed by the C-terminus before it breaks into the trans-membrane helix. Therefore it could still be possible that different interactions have evolved to drive the same type of substrate-related C-terminal dynamics seen in *Pf*KMO.

Chapter 5 – Structural investigation of inhibitor binding

5.1 Introduction

5.1.1 KMO Inhibition by substrate analogues

The first type of potent KMO inhibitors discovered were analogues of the substrate L-Kyn. Whilst they were indeed effective at preventing the formation of a hydroxylated product, it was recently shown that the binding of such compounds stimulates uncoupled NADPH oxidation [1]. This led to flavin reduction, reaction with oxygen and then decomposition to release cytotoxic hydrogen peroxide and return the flavin to its oxidized state. Consequently substrate analogues have been re-classified as non-substrate effectors and have no potential for therapeutic use. Despite this their study was still of significance as it revealed critical information about the substrate binding environment and the potential functional groups that would be useful for designing true inhibitors of KMO.

Two very potent substrate analogues emerged early on in the shape of 3,4-CBA [207] and *m*-NBA [206], which had IC_{50} values of 0.2 μM and 0.9 μM respectively against KMO from rat organ homogenates. From this Giordani *et al*. continued their exploration of the compounds by studying the effect on rat liver and brain KMO inhibition of varying the functional groups of 3,4-CBA [208]. Initially they found that the carboxylic acid group was crucial to inhibition as its removal increased the IC_{50} to millimolar levels. This fitted in with the observations from the substrate bound structure presented in chapter **3** where the carboxyl group of L-Kyn facilitates the majority of its interactions with the residues lining the substrate binding site. Removal of the tail amine group also reduced the inhibition strength but to a significantly lesser extent, with an IC_{50} of 3.9 μM and 0.9 μM on liver and brain KMO respectively. They then synthesized a 4-phenyl-4-oxo-butanoic acid backbone (figure 5.1) from which the R, R' and X functional groups were varied. The results regarding inhibition of liver and brain rat KMO are shown in table 5.1.

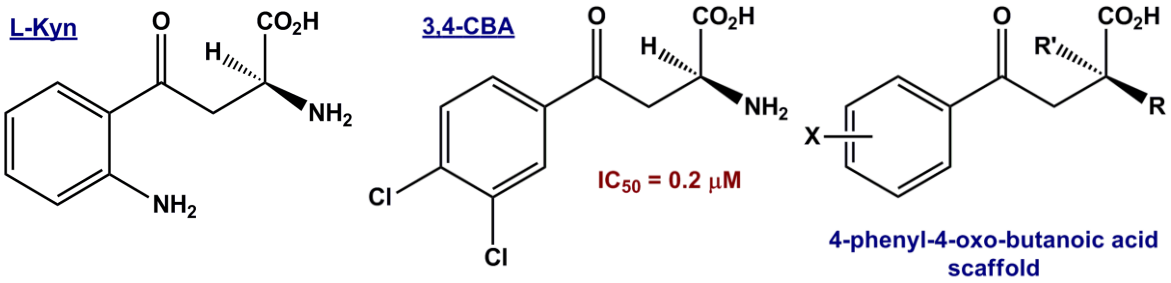


Figure 5.1. Structures of the substrate L-Kyn, the substrate analogue 3,4-CBA and the 4-phenyl-4-oxo-butanoic acid scaffold.

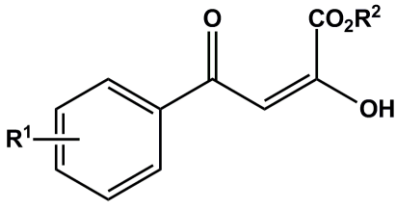
Table 5.1. Inhibition data from 4-phenyl-4-oxo-butanoic acid compounds based on scaffold in figure 5.1. Giordani *et al.,* [208]

Entry	R	R'	X	Rat liver KMO (IC_{50} ± SEM, μM)	Rat brain KMO (IC_{50} ± SEM, μM)
2	H	H	3,4-diCl	3.9 ± 0.2	0.9 ± 0.1
4	CH_3	H	3,4-diCl	6.9 ± 1.2	3.5 ± 0.2
5	OCH_3	H	3,4-diCl	6.9 ± 3.7	1.2 ± 0.1
6	$=CH_2$		3,4-diCl	36.0 ± 8.3	nt
7	Cl	H	3,4-diCl	14.8 ± 2.4	2.2 ± 0.2
8	OH	H	3,4-diCl	1.4 ± 0.3	0.30 ± 0.06
9	Ph	H	3,4-diCl	19.5 ± 1.8	nt
10	CH_2Ph	H	3,4-diCl	2.9 ± 1.2	0.18 ± 0.01
11	CH_3	CH_3	3,4-diCl	21 % @ 100 μM	nt
12	CH_3	NH_2	3,4-diCl	35 % @ 100 μM	nt
13	CH_3	OH	3,4-diCl	43 % @ 100 μM	nt
14	Cyclopropyl		3,4-diCl	20 % @ 100 μM	nt
20	OH	H	3-Cl	1.10 ± 0.3	0.48 ± 0.02
21	OH	H	3-F	9.1 ± 1.1	5.6 ± 0.9
22	OH	H	3-NO_2	11.5 ± 4.0	1.95 ± 0.15
23	OH	H	3,4-diF	3.0 ± 0.2	1.45 ± 0.08
24	OH	H	OCH_3	7 % @ 100 μM	nt

At the R position originally occupied by an amine group in the substrate and 3,4-CBA, quite varied changes could be introduced without causing a substantially large reduction in the strength of inhibition. However, significantly reduced potency was experienced when hydrophobic or large bulky groups were present. No drastic change was observed upon substitution with a hydroxyl group, on the other hand, suggesting that a hydrogen bond donor/acceptor may be required at this position. Indeed from the substrate bound structure solved in chapter 3, it appeared that the corresponding amine group of L-Kyn was involved as an H-bond donor to the backbone carbonyl of H320. The addition of functional groups at the R' position greatly diminished the strength of inhibition of the

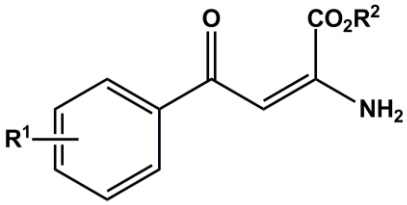
compounds in keeping with the fact that D-Kyn does not effectively bind to KMO. Finally, introduction of a substituent at the 3'-position of the aromatic ring showed that a hydrophobic OMe group abolished KMO inhibition leaving the order of inhibition strength as 3-Cl>3-F>3-NO_2. The importance of an added group at the 4' position was inconclusive as it improved the inhibition strength with the 3,4-difluoro but did not effect it with the 3,4-dichloro analogue.

A similar study was conducted in 2000 using 2-hydroxy- and 2-amino-4-aryl-4-oxo-butanoic acids and esters [209]. The results of inhibition assays using rat liver mitochondrial suspensions for the various compounds are detailed in tables 5.2 and 5.3. The results were quite inconsistent, particularly regarding differences between the 2-amino and their respective 2-hydroxy compounds (for example comparing compounds 3b to 5b and 2k to 4k). There was however strong evidence that substitution at the 2' position of the aromatic ring was largely ineffective compared to the inhibition potency achieved with substitutions at the 3' and 4' positions. In addition, the acid analogues showed significantly greater KMO inhibition than their ester analogues in the majority of cases. There was no clear evidence that the but-2-enoic chain produced stronger or weaker inhibitors than the previously used butanoic chain. Taken together, these studies highlighted some of the key features required for strong binding in the active site, the majority of which correlate well with the substrate binding site environment described in chapter **3**.



2-hydroxy-4-aryl-4-oxo-but-2-enoic acids and esters scaffold

compd	R^1	R^2	IC_{50} (μM) or % inhibn at 10 μM[a]
2a	H	Me	23%
3a	H	H	0%
2b	2-Cl	Me	31%
3b	2-Cl	H	33%
2c	2-NO_2	Me	1%
3c	2-NO_2	H	24%
2d	3-Cl	Me	1.9
3d	3-Cl	H	0.32
2e	3-Br	Me	2.1
3e	3-Br	H	1.7
2f	3-F	Me	1.6
3f	3-F	H	0.58
2g	4-Cl	Me	38%
3g	4-Cl	H	3.0
2h	4-Br	Me	1.9
2i	4-F	Me	22%
2j	4-Me	Me	12%
2k	2,3-di-Cl	Me	14%
3k	2,3-di-Cl	H	33%
2l	2,3-di-F	Me	8.1
2m	3,4-di-Cl	Me	1.9
3m	3,4-di-Cl	H	1.4
2n	3,4-di-F	Me	3.3
3n	3,4-di-F	H	0.12
2o	3-Cl,4-F	Me	2.6
3o	3-Cl,4-F	H	0.27
2p	3-Me,4-Cl	Me	4.9
3p	3-Me,4-Cl	H	1.7
2q	3-NO_2,4-Cl	Me	11.2
3q	3-NO_2,4-Cl	H	2.0
(±)-*m*-nicotinylbenzoylalanine			8.0
(*S*)-(+)-*m*-nicotinylbenzoylalanine			4.0



2-hydroxy-4-aryl-4-oxo-but-2-enoic acids and esters scaffold

compd	R^1	R^2	IC_{50} (μM) or % inhibn at 10 μM[a]
4a	H	Me	3%
5a	H	H	17%
4b	2-Cl	Me	31%
5b	2-Cl	H	19%
4d	3-Cl	Me	7.9
5d	3-Cl	H	2.4
4e	3-Br	Me	0%
4f	3-F	Me	0%
4g	4-Cl	Me	13%
5g	4-Cl	H	33%
4h	4-Br	Me	0%
4i	4-F	Me	1%
5i	4-F	H	41%
4j	4-Me	Me	4%
4k	2,3-di-Cl	Me	33%
5k	2,3-di-Cl	H	32%
4m	3,4-di-Cl	Me	37%
5m	3,4-di-Cl	H	1.0
4n	3,4-di-F	Me	3.6
5n	3,4-di-F	H	1.5
4o	3-Cl,4-F	Me	4%
5o	3-Cl,4-F	H	0.82
4p	3-Me,4-Cl	Me	0%
(±)-*m*-nicotinylbenzoylalanine			4.0
(*S*)-(+)-*m*-nicotinylbenzoylalanine			8.0

Tables 5.2 (left) & 5.3 (right). Rat KMO liver Inhibition results for compounds based on the structural scaffolds depicted above each respective table (taken from Drysdale et al 2000 [209]). [a] IC_{50} values taken as averages from multiple runs.

5.1.2 KMO inhibition by novel organic compounds

Shortly after the discovery of substrate analogue inhibitors, the search for more specific and more potent KMO inhibitors moved on to novel and more complex compounds. In 1997, screening of a chemical library based on bioisosters of the L-Kyn carboxyl group rapidly identified a sulphonamide as a good starting scaffold [4]. Subsequently a series of *N*-(4-phenylthiazol-2-yl)benzenesulphonamides were synthesized and shown to be promising high-affinity competitive inhibitors of KMO. The inhibition results are summarised in table 5.4. Of those tested *in vitro* on KMO from rat kidney mitochondria,

the strongest compounds had IC_{50} values of 20-40 nM, up to 10 times lower than the substrate analogues. Within error there was little difference to discriminate between lots of the compounds, however a few patterns were clear. On the R' ('head') ring the addition of functional groups at the 3-, the 4-, or both positions improved the strength of inhibition of the compound. Of the functional groups tested, hydrophilic groups were favoured over hydrophobic groups to provide the strongest KMO inhibition. As a result the order of potency ranged from H<OMe<Me<Cl/NO_2. Again the presence of functional groups at position 4-, or at 3- and 4- on the R'' ('tail') ring improved the inhibition strength, but here Cl groups were less effective than NH_2, Me or OMe. Compound 16, known as Ro 61-8048, was taken forward for potential therapeutic inhibition of human KMO. Whilst the inhibitor was an effective therapeutic tool in several *in vivo* and *ex vivo* disease models (see section **1.4.2**) it never went past pre-clinical trials as it failed the Ames mutagenic test.

Scaffold for Rover compounds

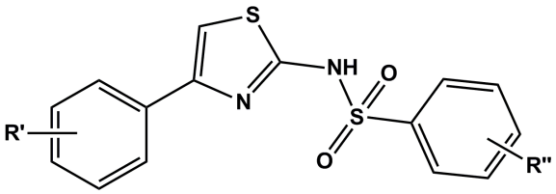


Table 5.4. Summary of inhibition results for series of benzenesulphonamides based around the displayed molecule scaffold. FCE 28833 is another name for 3,4-CBA. Taken from Rover et al., 1997 [].

compd	R'	R''	IC_{50} (nM)
6	H	4-Me	470 ± 190
7	4-OMe	4-Me	240 ± 41
8	4-Cl	4-Me	56 ± 8
9	4-Me	4-Me	76 ± 14
10	3-Cl, 4-Cl	4-Me	56 ± 8
11	3-NO_2	4-Me	48 ± 12
12	3-NO_2	H	200 ± 35
13	3-NO_2	3-Cl, 4-Cl	120 ± 36
14	3-NO_2	4-Cl	84 ± 29
15	3-NO_2	4-OMe	51 ± 16
16	3-NO_2	3-OMe, 4-OMe	37 ± 3
17	3-NO_2	4-NH_2	40 ± 3
18	3-NO_2	4-*i*-Pr	990 ± 20
19	2-F, 5-CF_3	3-OMe, 4-OMe	39 ± 17
20	2-F, 5-CF_3	4-NH_2	19 ± 2
m-NBA			774 ± 113
FCE 28833			237 ± 32

The aim of the work described in this chapter was to build on screening findings and use co-crystallisation to reveal information regarding the mechanism/interactions involved in the binding of non-substrate effectors and inhibitors. Part of this involves efforts to understand what causes the differential recognition of the two types of compound by KMO. With this information novel inhibitors could be rationally designed without the side-effect of uncoupled NADPH oxidation. As part of this the design and synthesis of a novel series of potential KMO inhibitors was carried out by Gavin Milne (PhD, St.

Andrews University). Using *in vitro* screening with pure recombinant *Pf*KMO, these compounds were rationally modified with the aim of increasing their potency. Novel potent inhibitors were co-crystallised with *Pf*KMO to obtain crucial structural information regarding the nature of inhibitor binding. Structural information obtained would rapidly speed up the development of targeted inhibitors of human KMO for potential therapeutic exploitation.

5.2 Binding of substrate analogues

5.2.1 Inhibition data

Substrate analogues may themselves be unsuitable as therapeutic tools, however a lot of useful information could still be gained from their study. The changes that these cause to the flavin spectrum upon binding, and their competition with L-Kyn binding indicate that it is likely that these compounds bind in the same specific location and orientation as the substrate itself. Therefore investigating the *Pf*KMO inhibition strength of different analogue compounds could aid the rational design of novel KMO inhibitors. The published substrate analogues benzoylalanine (BA), *m*-nitrobenzoylalanine (*m*-NBA) and 3,4-dichlorobenzoylalanine (3,4-CBA), as well as some additional compounds (labelled c39-c43), were synthesized by Gavin Milne (PhD, St Andrews University). The compounds, which are displayed in figure 5.2, were screened and partially characterized for *Pf*KMO inhibition with respect to L-Kyn binding.

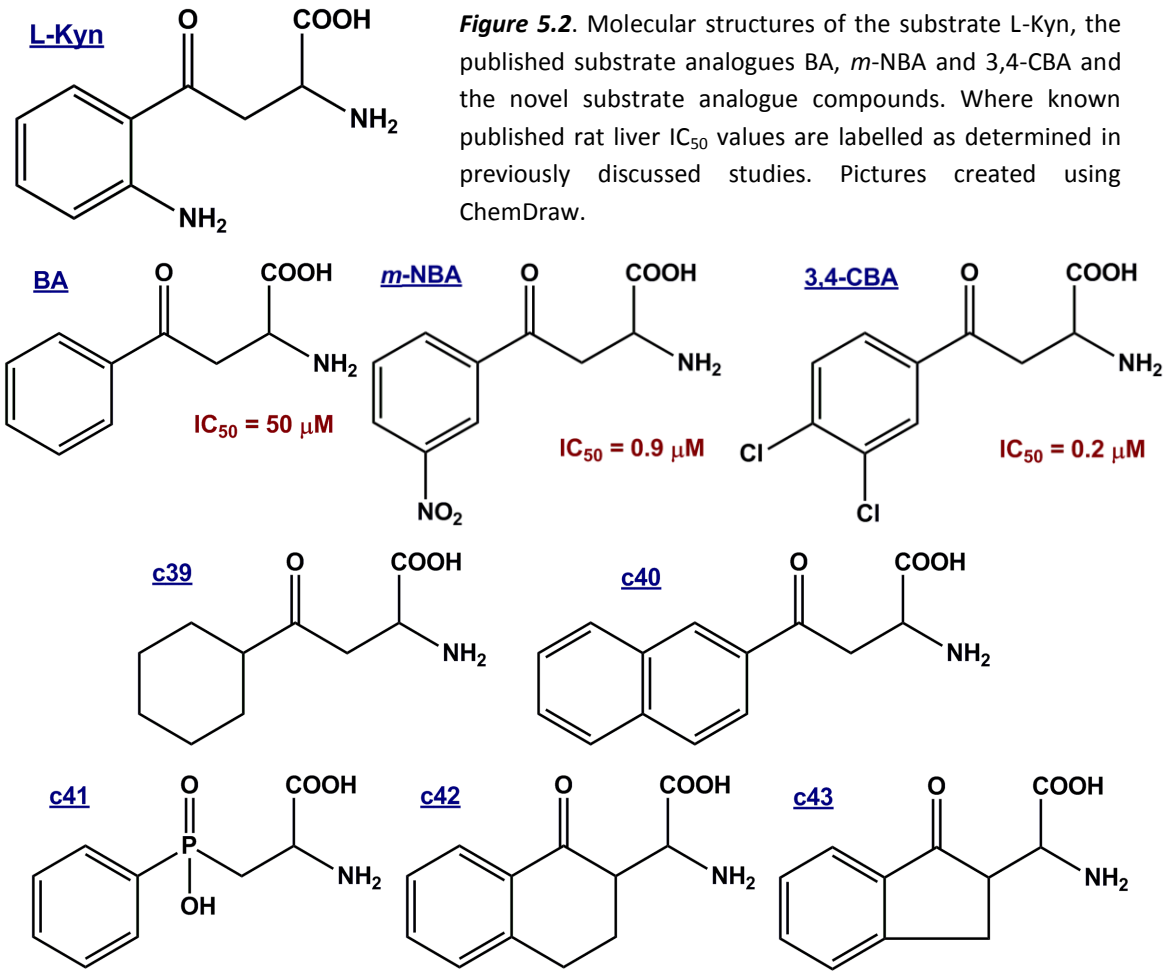


Figure 5.2. Molecular structures of the substrate L-Kyn, the published substrate analogues BA, *m*-NBA and 3,4-CBA and the novel substrate analogue compounds. Where known published rat liver IC_{50} values are labelled as determined in previously discussed studies. Pictures created using ChemDraw.

The published substrate analogues BA, *m*-NBA and 3,4-CBA were shown to be competitive inhibitors of L-Kyn binding of varying strength. The K_i values measured correlated well with the published IC_{50} values determined in assays with rat KMO. The unsubstituted BA had a moderate K_i of 7.0 ± 3.6 μM whereas addition of an *m*-nitro or 3,4-dichloro functional group dramatically increased the potency with K_i values of 42.0 ± 3.0 nM and 5.3 ± 1.0 nM respectively. For 3,4-CBA the inhibition constant was so low that it could not be accurately determined within the error of each individual K_i assay, however it was evidently in the very low nanomolar region. Example inhibition plots from one run with each compound *m*-NBA and 3,4-CBA are shown in figure 5.3. It was previously reported that the binding of these compounds stimulated uncoupled NADPH

oxidation leading to degradation of the reduced flavin to release the free radical generator hydrogen peroxide [1]. To confirm this each substrate analogue was tested for NADPH oxidation in the absence of substrate. It was observed that BA and *m*-NBA stimulated NADPH oxidation confirming them to be substrates or non-substrate effectors rather than inhibitors. However, in the presence of up to 100 μM of the compound 3,4-CBA there was no sign of background NADPH oxidation. With the incredibly low K_i value for this compound it seems likely that 3,4-CBA in fact acts as a competitive *Pf*KMO inhibitor. The K_i assays were all proportionally background corrected to remove the contribution of NADPH oxidation caused by BA and *m*-NBA binding from the measured rate readings

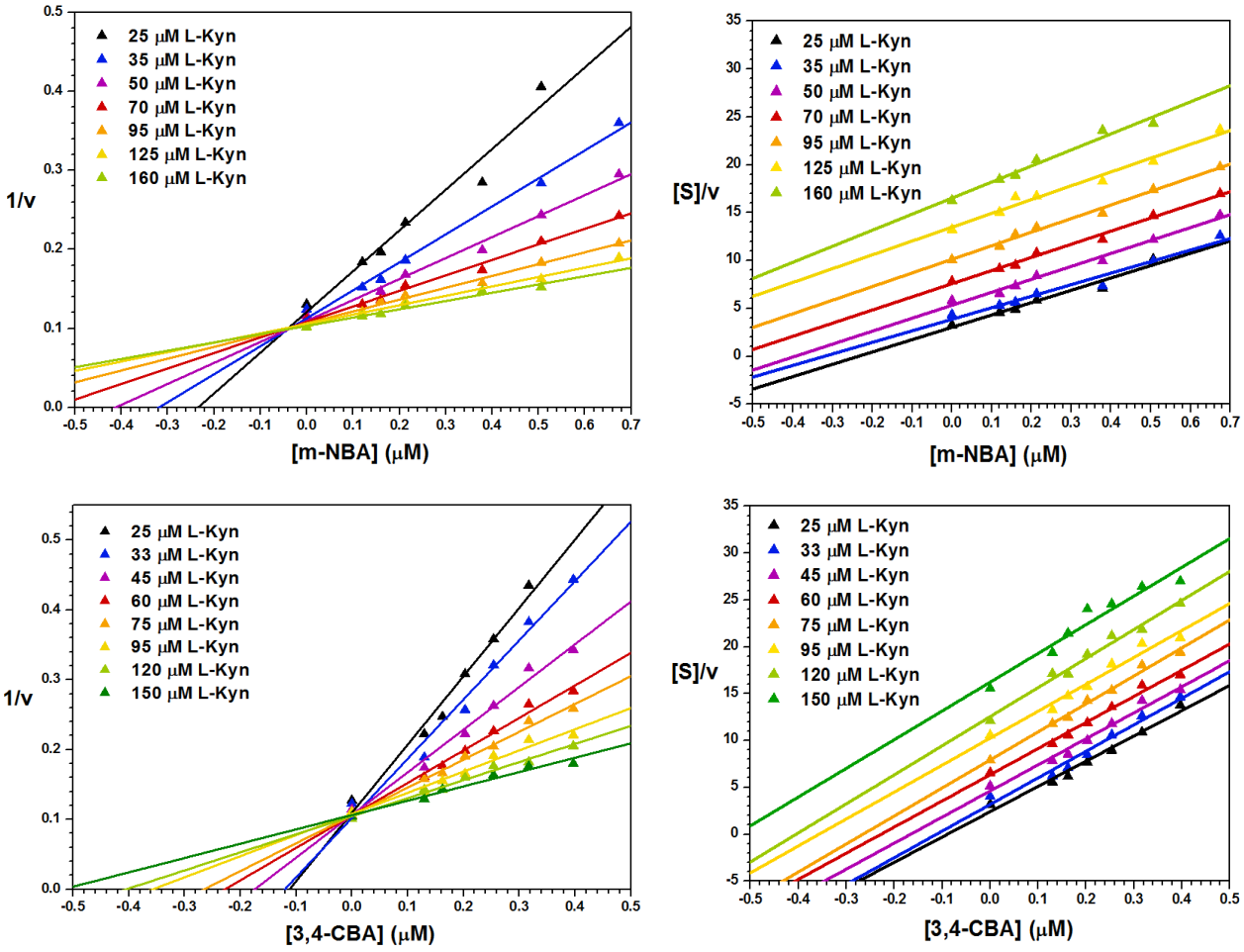


Figure 5.3. Inhibition plots for *m*-NBA and 3,4-CBA. The K_i is estimated from the average intersection point on the graph of 1/v against concentration of inhibitor (*left*) and competitive inhibition is shown by the parrellel best-fit lines on the graph of [S]/v (where [S] = concentration of L-Kyn) against concentration of inhibitor (*right*). Data plotted using Origin 8.0.

All of the novel substrate analogues tested had reduced potency of L-Kyn inhibition but to varying extents. The conversion of BA's phenyl ring into a fully saturated cyclohexyl ring in c39 reduced the inhibition potency by around five-fold. On the other hand conversion into a naphthalene ring in c40 only had no effect on the strength of inhibition. This was interesting as it suggested that there was a large hydrophobic pocket that could be filled at this end of the binding site. Additionally the pocket must have enough room to allow addition of functional groups at the 3' and 4' positions of the phenyl ring as seen with *m*-NBA and 3,4-CBA. These observations match with the nature of the substrate binding environment discovered in section 3. A large hydrophobic pocket surrounds the substrate's phenylamine ring but at with a relatively tight fit as the separation distance is only around 4.5 Å. Certainly if c40 was to bind in a similar location there would have to be a reasonable amount of local positional adjustments.

The other three compounds synthesized – c41, c42 and c43 – were very ineffective as *Pf*KMO inhibitors. This was to be expected if the compounds did bind in the substrate binding pocket due to the conformation of the substrate L-Kyn adopted in the bound structure. In particular c42 and c43 with their greatly reduced flexibility would be hindered in trying to position the amino and carboxyl groups to coordinate the key hydrogen bonds observed in the substrate binding pocket. With the weak binding of these compounds it was difficult to observe any stimulation of NADPH oxidation in the absence of substrate for the concentrations tested. The aforementioned compound c39 clearly showed background NADPH oxidation and so could be considered as either a substrate or a non-substrate effector similar to BA and *m*-NBA. The extent of uncoupling of NADPH oxidation with c40 needs to be tested further. All the substrate analogues tested showed competitive inhibition in plots of [S]/v against [analogue]. The full inhibition results for the substrate analogues are summarised in table 5.5.

	% Inhibition with x μM inhibitor					
Compound	**10**	**30**	**100**	***BG turnover**	**K_i (μM)**	**IC_{50} (μM)**
c39	6.3	10.0	17.1	10.2 %	35 ± 18	
c40	n.d	n.d	n.d	n.d	10.3 ± 2.3	
c41	0	0	4.1	0.0 %		
c42	0	5.7	8.9	0.0 %		
c43	1.0	6.3	9.0	2.9 %		
BA	25.6	46.2	71.2	13.5 %	7.0 ± 3.6	50
***m*-NBA**	92.1	96.1	96.5	13.0 %	0.042 ± 0.003	0.9
3,4-CBA	97.2	100.0	100.0	0.4 %	0.0053 ± 0.0010	0.2

Table 5.5. Summary table of the inhibition data measured for the substrate-analogue compounds. Screen values represent the % inhibition of the rate of NADPH oxidation with 50 μM L-Kyn for the stated analogue concentrations compared to uninhibited controls. n.d = no data. *BG turnover relates to the measured background NADPH oxidation in absence of substrate as a percentage of the normal rate of turnover with 50 μM L-Kyn, recorded in the presence of 100 μM of each compound. The BG was subtracted from screen values and in Ki measurements. K_i values are shown as average values from multiple runs.

5.2.2 The Structure of *Pf*KMO with *m*-NBA bound

Given the strong binding affinity of *m*-NBA and 3,4-CBA for *Pf*KMO, these compounds could be readily co-crystallized with the enzyme. As such crystals of *Pf*KMO with *m*-NBA bound were grown using hanging drop vapour diffusion (h.d.v.d) with a 1:1 μL drop composed of protein solution (14 mg/ml in 20 mM HEPES pH 6.8, 30 mM Na acetate, 1 mM DTT) and reservoir solution supplemented with the substrate analogue (0.08 M HEPES pH7.0, 6 % PEG 4K, 8 % glycerol, 8 mM NaCl, 10 % 2-propanol and 1 mM *m*-NBA). Large, hexagonal prism shaped crystals with rough dimensions of 0.2 × 0.3 × 0.5 mm^3 grew in 2-3 days and were harvested for data collection after 9 days. Prior to flash-freezing crystals in liquid nitrogen, 1.5 μl of cryoprotectant (mother liquor plus 20 % EG) was added to the crystallisation drop and then the crystal was briefly soaked in a drop of cryoprotectant.

The crystals were screened for diffraction quality at the Diamond synchrotron beamline I02 with several diffracting to a resolution around 3.5 Å. X-ray diffraction data were collected using an X-ray wavelength of 0.979 Å, exposure of 0.3 s/frame of data and beam intensity of 100 % with 0.3 ^{o} oscillations over a total of wedge of 67.2 ^{o}. The crystal

lattice was similar to substrate bound crystals; space group $I4_122$ with mean unit cell parameters of a = b = 149.7, c = 273.0 Å, α = β = γ = 90.0 °. The dataset was 96.1 % complete to a resolution of 3.55 Å with R_{merge} = 0.096 overall and 0.799 in the highest resolution shell. The crystal used for data collection is displayed in figure 5.4 with full data collection statistics tabulated in appendix section **3.5**.

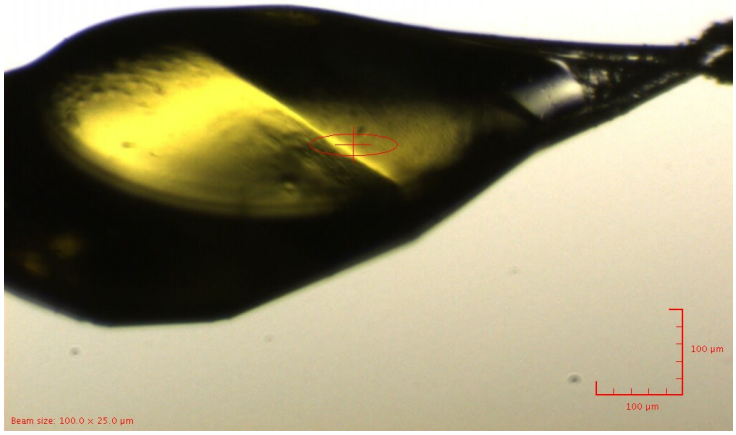


Figure 5.4. Close-up picture of the mounted m-NBA bound *Pf*KMO dm2 crystal used for data collection with scale and beam size labelled.

The structure was solved by molecular replacement using Molrep [reference], with the previously solved monomeric substrate bound structure used as the molecular replacement model (see section **3.3**). There was one protein molecule present in the asymmetric unit with a solvent content of 83.7 %. The model was cycled through manual editing in Coot and TLS and restrained refinement in Refmac. There was clear density to place one molecule of FAD and one molecule of the substrate analogue *m*-NBA. Refined electron density maps calculated in the presence and absence of *m*-NBA are shown in figure 5.5. The final model included residues 7-461 with R and R_{free} values of 0.247 and 0.254 respectively, a figure of merit of 0.776 and an overall mean B-factor of 133.4 $Å^2$ for 17259 significant reflections in the range of 44.76-3.55 Å. The root means square deviation (rmsd) of bond lengths and angles from the ideal were 0.005 Å and 0.919 ° and all the residues were present in the most favoured or additionally allowed regions of a

Ramachandran plot as determined using the Molprobity server. Full refinement statistics are tabulated in appendix section **3.5**.

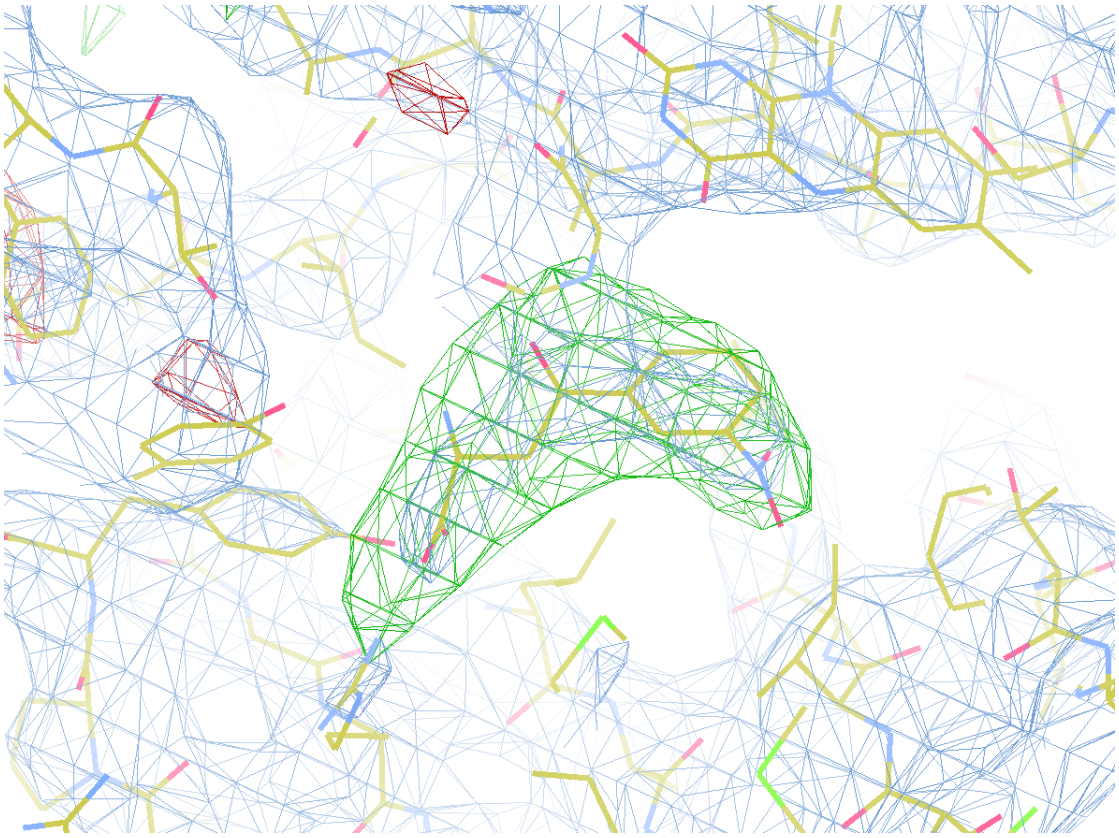


Figure 5.5. Electron density for the bound substrate-analogue *m*-NBA in the final model. The $2F_o$-F_c density map from the final round of refinement is shown in blue. The F_o-F_c map after refinement in the absence of *m*-NBA is shown in green/red. Created with Coot with contouring of 1.5 and 3.0 σ.

Overall the structure of *m*-NBA bound *Pf*KMO was visually indistinguishable from the previously described substrate bound structure. As can be seen in figure 5.6, the substrate analogue was bound in the same site as the substrate L-Kyn with the C-terminal domain also in the same closed conformation. This structural snapshot suggests that non-substrate effectors like *m*-NBA are initially recognised and bound as substrates, which presumably leads to the observed uncoupled NADPH oxidation and hydrogen peroxide release. Interestingly the *m*-nitro group is not orientated to directly block hydroxylation by a C_{4a}-hydroperoxyflavin. Instead it points away from the flavin and into the hydrophobic pocket lined by residues M222, I224, F238, F319 and M373. Therefore the failure of these compounds to act as substrates is likely to be due to the loss of the 2'

amine group compared to L-Kyn rather than due to steric hindrance from the 3' nitro group. At the resolution observed it is difficult to accurately compare the exact positions/orientations of the substrate and *m*-NBA. Therefore it is difficult to conclusively pinpoint the cause of the differential hydroxylation patterns. From the structures, however, it is plausible that either: i) the *m*-nitro goup slightly distorts the position of the substrate phenyl ring or ii) the missing 2'-amine group is directly involved in directing the flavin intermediate towards efficient hydroxylation over the uncoupled release of hydrogen peroxide.

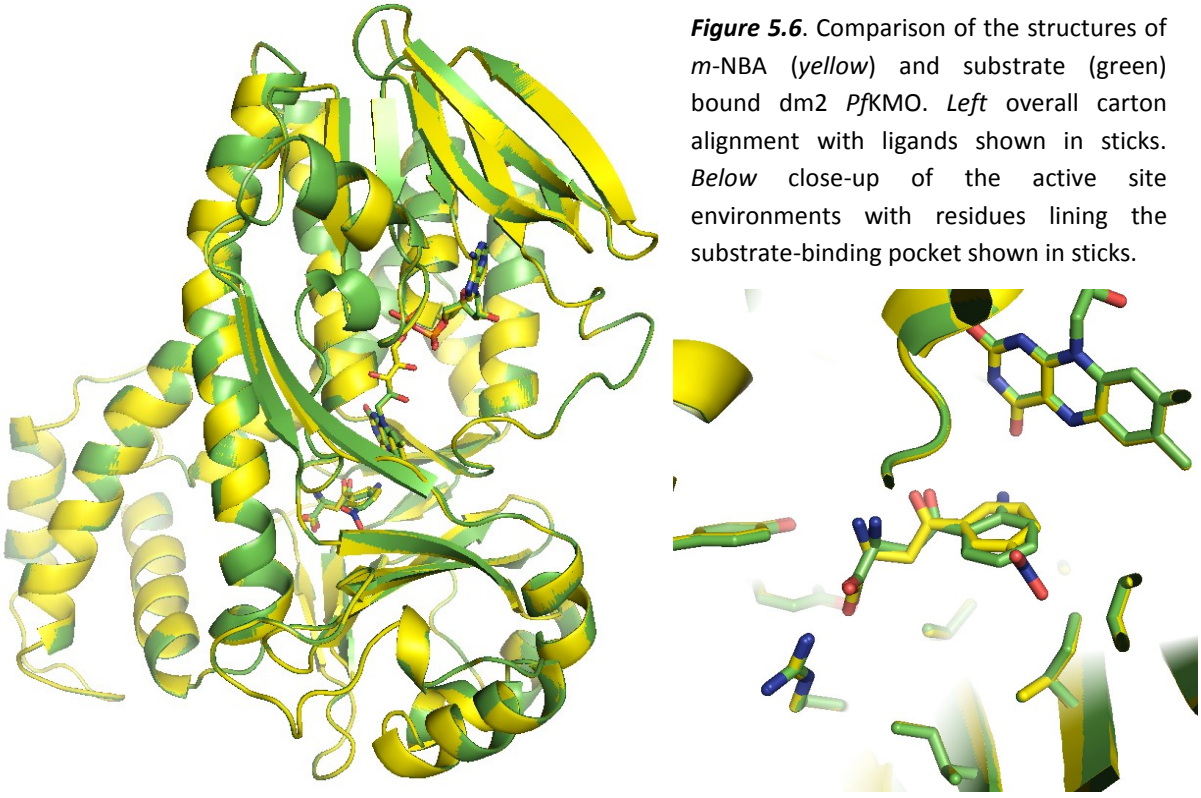


Figure 5.6. Comparison of the structures of *m*-NBA (*yellow*) and substrate (green) bound dm2 *Pf*KMO. *Left* overall carton alignment with ligands shown in sticks. *Below* close-up of the active site environments with residues lining the substrate-binding pocket shown in sticks.

5.2.3 The Structure of *Pf*KMO with 3,4-CBA bound

Crystals of *Pf*KMO with 3,4-CBA bound were grown using h.d.v.d with a 1:1 μL drop composed of protein solution (11.5 mg/ml in 20 mM HEPES pH 6.8, 20 mM K acetate) and reservoir solution supplemented with the substrate analogue (0.1 M HEPES pH 7.0, 9 % PEG 4K, 8 % glycerol, 10 mM NaCl, 10 % 2-propanol, 1 mM 3,4-CBA). Large crystals

shaped like hexagonal prisms (see figure 5.7) with rough dimensions of 0.2 × 0.2 × 0.6 mm^3 grew in 2-4 days and were harvested for data collection after 7 days. Crystals were flash-frozen after brief sequential soaks in mother liquor supplemented with 10 % and then 20 % EG. The crystals were screened for diffraction quality at the Diamond synchrotron beamline I24 with several of them diffracting to a resolution around 3.4 Å. X-ray diffraction data were collected using an X-ray wavelength of 0.978 Å, exposure of 0.2 s/frame of data and beam intensity of 60 % with 0.2 ^{o} oscillations over a total of wedge of 120 ^{o}. The crystal lattice was similar to substrate bound crystals; space group $I4_122$ with mean unit cell parameters of a = b = 149.4, c = 273.8 Å, α = β = γ = 90.0 ^{o}. The dataset was 96.4 % complete to a resolution of 3.37 Å with R_{merge} = 0.099 overall (0.651 HS) and a mean intensity of 11.2 (2.8 HS). The full data collection statistics are tabulated in appendix section **3.5**.

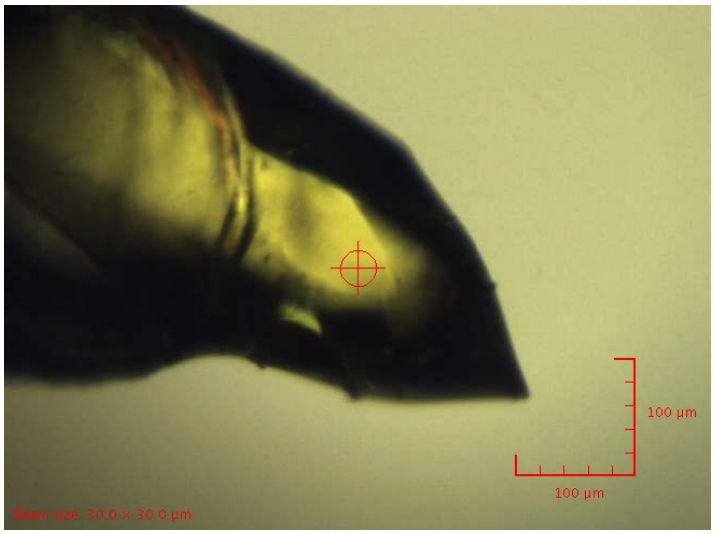


Figure 5.7. Close-up picture of one the mounted 3,4-CBA-bound *Pf*KMO crystals tested for diffraction with scale and beamsize are labelled

The structure was solved by molecular replacement using Molrep, with the previously solved monomeric substrate bound structure used as the molecular replacement model (see section **3.3**). As expected, there was found to be one protein molecule present in the asymmetric unit and a solvent content of 83.7 %. The model was cycled through manual editing in Coot and TLS and restrained refinement in Refmac. There was clear density to place one molecule of FAD and one molecule of the substrate analogue 3,4-CBA – both in the same position as seen before. Due to the low resolution of the data, it

was not conclusively clear which orientation of the compound's 3,4-chloro groups was correct. After refining both variations, one option was selected with a significantly better fit - as can be seen in the density maps in figures 5.8 and 5.9. At this resolution though, the multiple occupancy of both conformations cannot be ruled out. The final model included residues 7-461 with R and R_{free} values of 0.237 and 0.239 respectively, a figure of merit of 0.773 and an overall mean B-factor of 112.3 for 20256 significant reflections in the range of 34.56-3.37 Å. The rmsd of bond lengths and angles from the ideal were 0.004 Å and 0.882 ° and all the residues were present in the most favoured or additionally allowed regions of a Ramachandran plot as determined using the Molprobity server. Full refinement statistics are tabulated in appendix section **3.5**.

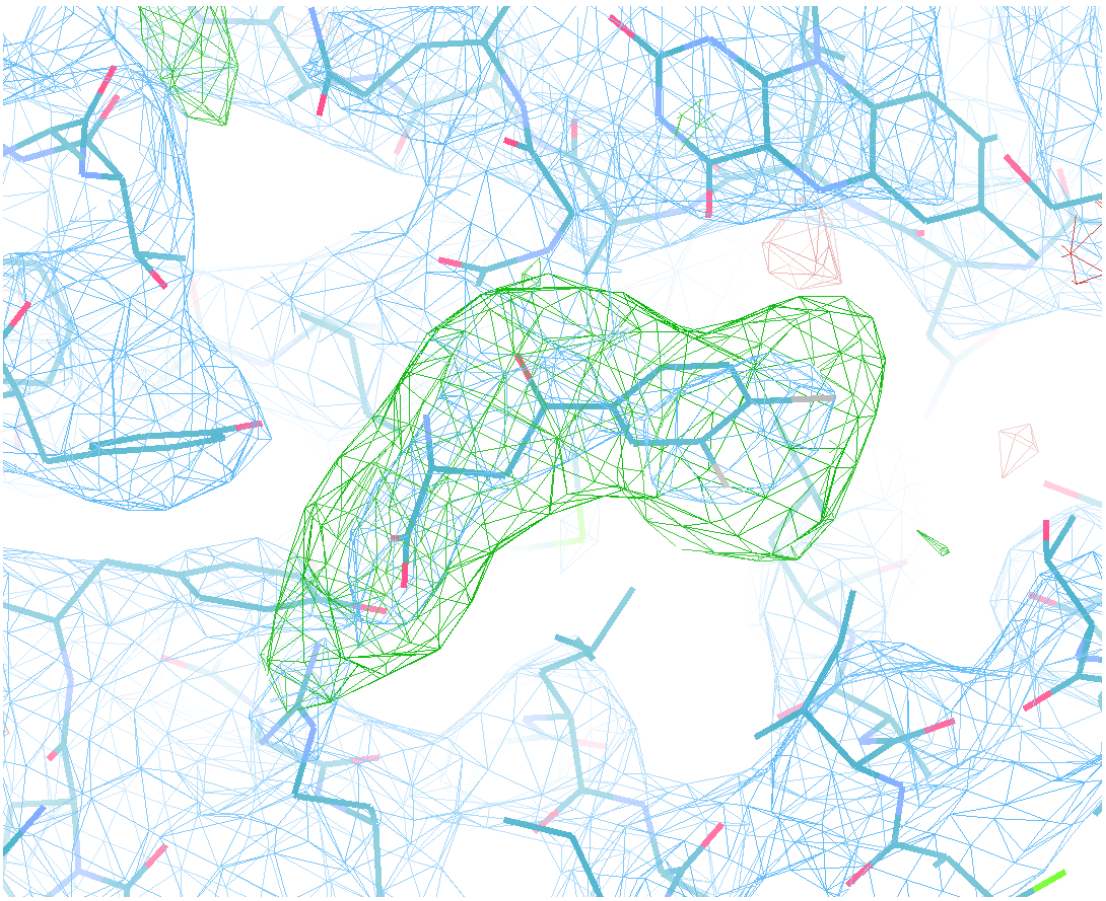


Figure 5.8. Refined $2F_o$-F_c (blue) and F_o-F_c (green/red) density for the final model of 3,4-CBA in the substrate-binding site. The F_o-F_c map was calculated after refinement without 3,4-CBA in the model. Created with Coot with contouring of 1.5 and 3.0 σ.

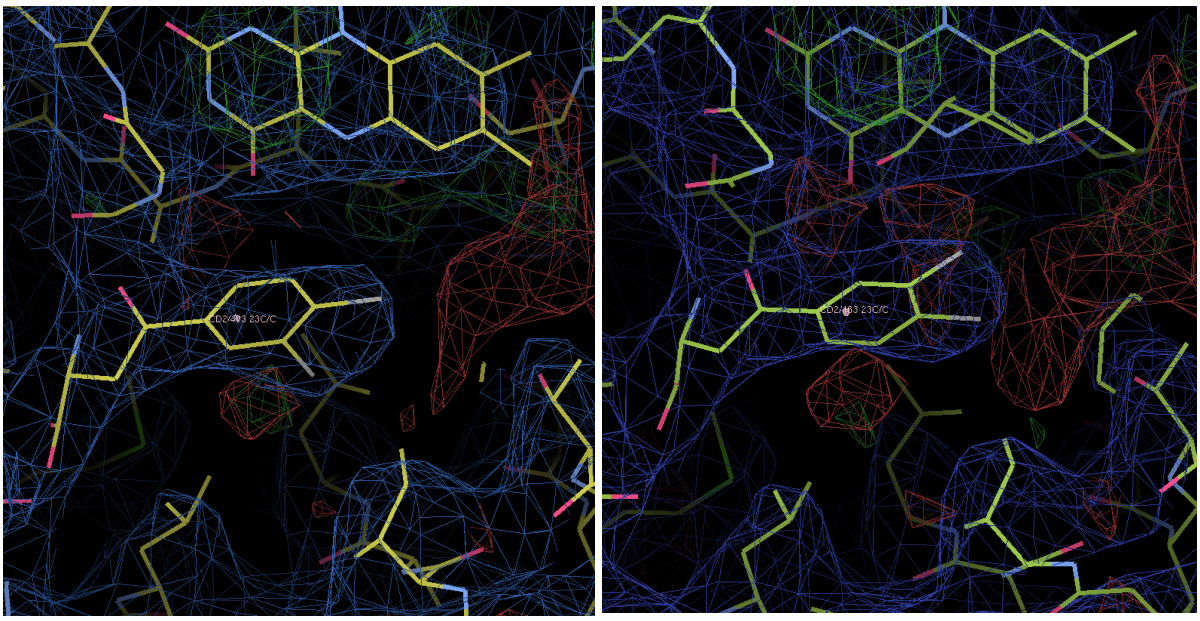


Figure 5.9. Refined $2F_o$-F_c density (blue) and F_o-F_c density (green/red) for the two orientations of 3,4-CBA in the substrate-binding site. Pictures created with Coot with contouring of 1.5 and 3.0 σ.

The overall structure of 3,4-CBA bound *Pf*KMO was again virtually indistinguishable from the previously described substrate bound structure. As can be seen in figure 5.10, the substrate analogue was bound in the same site as the substrate L-Kyn with the C-terminal domain also in the same closed conformation. As with *m*-NBA the additional functional groups of the analogue – in this case the 3,4-dichloro groups – were not orientated to directly block hydroxylation by a C_{4a}-hydroperoxyflavin. Instead they point away from the flavin and into the same hydrophobic pocket as mentioned before. Therefore the documented failure of these compounds to act as substrates is likely to be due to the loss of the 2' amine group compared to L-Kyn rather than due to direct steric hindrance from the added 3' functional groups. Interestingly the inhibition results suggested that 3,4-CBA was a competitive *Pf*KMO inhibitor rather than a substrate analogue like *m*-NBA. No obvious differences are evident between the structures and the binding nature of the two compounds. Instead their differing natures regarding uncoupled NADPH oxidation may be down to the addition of a functional group at the 4'-position of the phenyl ring in 3,4-CBA. Whether or not this position is critical for distinguishing substrate analogues that do and do not stimulate uncoupled NADPH oxidation, with 3,4-CBA bound the process of flavin reduction must be restricted. From

this it could be possible that either the movement of the flavin isoalloxazine moiety or the movement/formation of the NADPH docking site is prohibited.

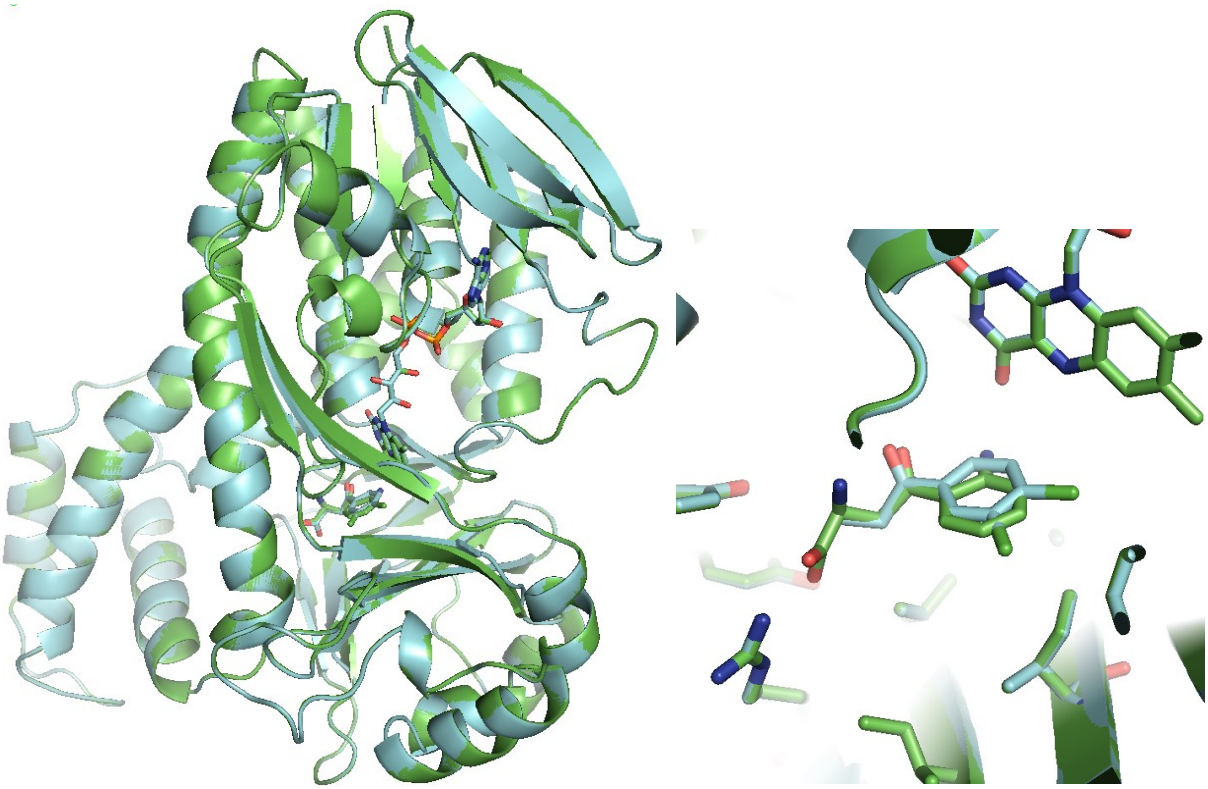


Figure 5.10. Comparison of the structures of 3,4-CBA (cyan) and substrate (green) bound dm2 *Pf*KMO. *Left* overall carton alignment with ligands shown in sticks. *Below* close-up of the active site environments with residues lining the substrate-binding pocket shown in sticks. Created using PyMol.

5.3 Binding of a novel inhibitor series

5.3.1 A novel series of *Pf*KMO inhibitors based on the 'Rover' compounds

The most successful KMO inhibitors discovered so far were the compounds synthesized by Rover *et al*. [4] , as discussed in section **5.1.2**. These compounds were therefore selected to provide both an initial design scaffold and a comparative level for a novel series of potential inhibitors. Gavin Milne (PhD, University of St Andrews) synthesized Rover compounds 16 (Ro 61-8048) and 19, which were subsequently tested for *in vitro*

*Pf*KMO inhibition. Both were shown to be reasonably strong competitive inhibitors with respect to L-Kyn with K_i values of **1.03 ± 0.27 μM** and **1.67 ± 0.23 μM** respectively. Inhibition plots for one of the runs with Rover 16 is shown in figure 5.11. With each compound, NADPH oxidation in the absence of substrate was not detected. The compounds were however very insoluble, and initial solid stocks were dissolved in 100 % DMSO and diluted to give up to a 2 % DMSO final assay concentration. The aim was to use a similar scaffold to that of the potent Rover compounds but with varied functional groups to improve the efficacy and safety of the compounds for potential therapeutic use. As Rover 16 was found to be slightly more potent than Rover 19, the *m*-nitro functional group was initially used in conjunction with a novel scaffold depicted in figure 5.12 to investigate the effect of varying the R_1 groups on the strength of KMO inhibition. The inhibition data is summarised in table 5.6.

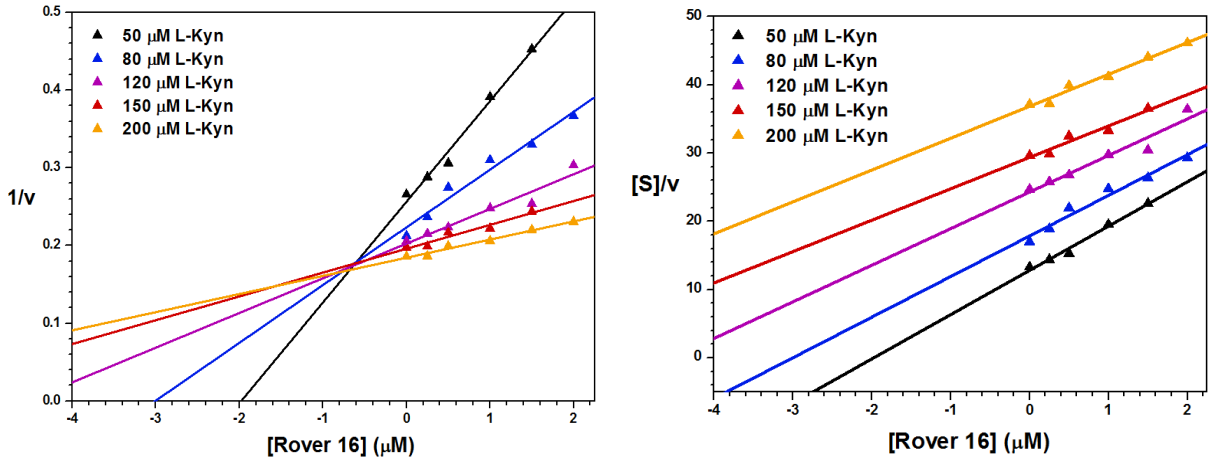


Figure 5.11. Inhibition plots for Rover 16 (Ro 61-8048). The K_i is estimated from the average intersection point on the graph of 1/v against concentration of inhibitor (*left*) and competitive inhibition is shown by the parrellel best-fit lines on the graph of [S]/v (where [S] = concentration of L-Kyn) against concentration of inhibitor (*right*). Data plotted using Origin 8.0.

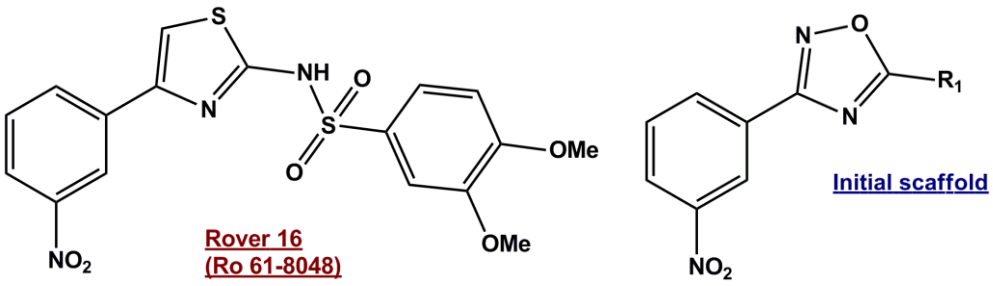


Figure 5.12. Molecular structures of Rover compound 16 as well as the initial scaffold for the novel inhibitor series synthesised. Created using ChemDraw.

		% Inhibition with x μM inhibitor				
Compound	**R_1 group**	**5**	**10**	**15**	**Screen**	**K_i (μM)**
GM283	COOH	0.0	0.0	0.0		
GM57	CH_2-CH_2NH_2	0.0	0.0	0.0		
GM42	CH_2-$CHNH_2COOH$	0.0	0.0	0.0		
GM41	CH_2-CH_2-$CHNH_2COOH$	0.0	0.0	0.0		
GM73	CH_2-CH_2-C_6H_5	1.6	8.4	16.2		
GM107	CH_2-CH_2-C_6H_3(3,4-diOMe)	2.5	6.1	12.8		
GM88	CH=CH-C_6H_3(3,4-diOMe)	ppt	ppt	ppt		
GM155	*CH-(O)-CH-C_6H_3(3,4-diOMe)	4.9	14.1	26.4		**2.88 ± 0.18**
GM213	CH=CNH_2-C_6H_3(3,4-diOMe)				20.3 ± 7.1	
GM266	CHOH-$CHNH_2$-C_6H_3(3,4-diOMe)				13.6 ± 8.2	
GM395	CHOH-$CHNH_2$-C_6H_3(3,4-diOMe)				5.1 ± 5.5	
GM511	*Cis* CHOH-CHOH-C_6H_3(3,4-diOMe)				0.0 ± 2.1	
GM512	*Trans* CHOH-CHOH-C_6H_3(3,4-diOMe)				4.5 ± 5.4	
GM229	CH_2-CO-C_6H_3(3,4-diOMe)				13.5 ± 4.2	
GM141	NH-CO-C_6H_3(3,4-diOMe)	**17.9**	**19.2**	**31.8**	**37.4 ± 9.9**	**1.30 ± 0.21**
GM195	NH-CO-CH_2-C_6H_3(3,4-diOMe)				**32.0 ± 11.2**	**1.82 ± 0.38**
GM289	CH_2-NH-CO-C_6H_3(3,4-diOMe)				10.1 ± 11.0	
GM444	CH_2-NH-SO_2-C_6H_5				1.5 ± 2.9	
GM473.1	2,2-diMe,4-C_6H_3(3,4-diOMe)-1,3-dioxolane				6.5 ± 11.3	
GM473.2	2,2-diMe,4-C_6H_3(3,4-diOMe)-1,3-dioxolane				0.0 ± 6.6	
Rover 16	NH-SO_2-C_6H_3(3,4-diOMe)	**12.1**	**26.3**	**32.7**	**43.6 ± 14.4**	**1.03 ± 0.27**

Table 5.6. Summary of *Pf*KMO inhibition data for initial set of compounds compared to Rover 16. The % inhibition values were measured relative to uninhibited enzyme in 1 ml cuvette assays. ppt indicates that the compound precipitated during the assay. For the screen values, the % inhibition represents an average value plus or minus the standard deviation from 3-6 readings in 96-well plate screens. K_i values were determined an average values plus or minus the standard deviation from multiple separate assays. More detailed experimental details are available in section **2.5**.

The first compounds tested did not yield any strong inhibitors of *Pf*KMO. All compounds with hydrophilic R groups, including GM283, GM42, GM57 and GM41, showed no inhibition at the concentrations tested. Some inhibition was seen where the R group contained a phenyl group, as in compounds GM73 and GM107. Compounds GM73, GM88 and GM107 had low solubility in the final assay mixture. GM73 and GM107 were kept soluble with the addition of 5-10 % DMSO but GM88 continued to precipitate. As a result of this finding it was presumed that a hydrophobic R group was required for inhibition, as is the case in the Rover compounds. This hydrophobic nature does though lead the compounds towards insolubility – another major feature of the Rover compounds. It was decided to fix the two phenyl ends of the molecule and investigate the positions in the adjoining alkyl chain with particular emphasis towards the addition of solubilizing hydrophilic functional groups.

When different organic linker chains were tested, different potencies of *Pf*KMO inhibition were detected. Levels of inhibition similar to the published KMO inhibitor Rover 16 were seen with a couple of the compounds that contained an amide linker group. GM141 was the most potent of the compounds tested with a K_i of **1.30 ± 0.21 μM**. GM195, with an extra CH_2 spacer in the R_1 linker showed a slightly weaker K_i of **1.82 ± 0.38 μM**. The position of the amide group was critical to the potency of the compound as GM289 – in which the amide was shifted along the chain by one place – showed significantly weaker *Pf*KMO inhibition. The carbonyl group by itself could not account for the increased potency over the previous round of compounds, as GM229 was not as effective during screening. The conservation of the backbone NH adjacent to the 5-membered heterocycle in the strongest inhibitors may implicate its involvement in a hydrogen bonding interaction with *Pf*KMO. The compounds with the amide linker were clearly more soluble than their predecessors, but still some DMSO was required. This round of refinement significantly furthered understanding in the development of this particular series of KMO inhibitors. At this point, no opportunity for significant deviation from the structure of the Rover compounds had been unearthed. In the final round of compound synthesis the attention turned towards varying the functional groups associated with the two phenyl rings (See figure 5.13).

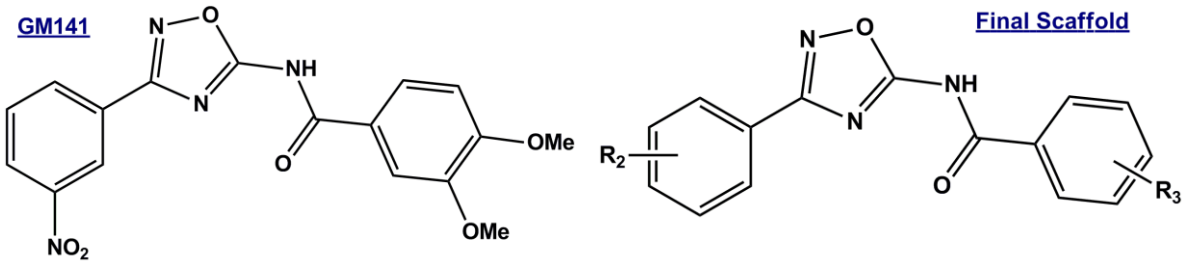


Figure 5.13. Molecular structures GM141 as well as the final compound scaffold for the synthesis of the last set of inhibitors. Created using ChemDraw.

The last round of compounds synthesized by Gavin Milne yielded an array of potent *Pf*KMO inhibitors as summarised in table 5.7. Substitution on the R_2 ring had the most significant impact on the potency of the compounds. From screening it appeared that inhibition strength increased with the following order of functional groups at the 3' position: H<Me<OMe<CF_3<F/NO_2<Cl/Br. The effect of swapping the NO_2 group of GM141 for a Cl or Br group was to slightly increase the potency of *Pf*KMO inhibition from 1.3 to around 0.8 μM. This increase was significantly amplified when a second Cl group was added at the 4' position in GM303 and GM347, which had nanomolar K_i values of 115.3 ± 10.7 nM and 65.0 ± 20.0 nM respectively. Monobromination at the 4' position in GM307 and GM469 yielded a similar level of inhibition to the same substituent in the 3' position. These findings lead to the suggestion that the presence of an electronegative/electron-withdrawing group at both positions is responsible for the potency of the strongest *Pf*KMO inhibitors.

			% Inhibition with x μM inhibitor					
Comp.	**R_2**	**R_3**	**0.5**	**5**	**20**	**50**	**Screen**	**K_i (μM)**
GM302	H	H	-	11.6	4.6	20.8	10.3 ± 9.5	-
GM339	3-Me	H	-	19.2	45.7	-	8.4 ± 9.6	-
GM348	3-OMe	H	-	23.8	49.1	-	7.0 ± 5.1	-
GM342	3-F	H	-	25.0	50.0	-	1.3 ± 8.0	-
GM304	3-CF_3	H	-	22.9	41.6	62.8	22.6 ± 12.1	-
GM307	4-Br	H	-	20.9	49.2	62.1	21.3 ± 10.9	-
GM340	3-Cl	H	-	29.9	57.2	-	22.0 ± 3.6	2.770 ± 0.430
GM308	3-Br	H	-	38.9	66.1	84.3	13.0 ± 9.2	1.581 ± 0.197
GM303	3,4-diCl	H	**37.0**	**87.0**	**95.6**	**96.6**	**95.0 ± 3.2**	**0.115 ± 0.011**
GM328	H	3,4-diOMe	-	-	-	-	5.4 ± 4.6	-
GM366	3-Me	3,4-diOMe	-	-	-	-	2.2 ± 7.7	-
GM467	3-OMe	3,4-diOMe	-	-	-	-	20.3 ± 5.1	-
GM367	3-F	3,4-diOMe	-	-	-	-	17.7 ± 9.1	-
GM457	3-CF_3	3,4-diOMe	-	30.0	-	-	39.7 ± 13.6	-
GM469	4-Br	3,4-diOMe	-	44.0	-	-	44.3 ± 13.6	-
GM466	3-Cl	3,4-diOMe	-	50.0	-	-	56.4 ± 9.6	0.782 ± 0.054
GM440	3-Br	3,4-diOMe	-	58.3	-	-	62.4 ± 10.9	-
GM347	3,4-diCl	3,4-diOMe	**50.9**	**95.9**	**96.9**	**-**	**94.4 ± 1.9**	**0.065 ± 0.020**
GM729	3,4-diCl	3,4-diOH	**37.9**	**90.3**	**-**	**-**	**93.4 ± 2.2**	-
GM760	3,4-diCl	3-$PO_2(OEt)_2$, 4-OMe	**78.4**	**99.7**	**-**	**-**	**91.6 ± 6.7**	**0.020 ± 0.004**
GM769	3,4-diCl	3-OMe, 4-$PO_2(OEt)_2$	**75.1**	**97.2**	**-**	**-**	**97.6 ± 0.1**	**0.017 ± 0.002**
GM141	3-NO_2	3,4-diOMe	-	-	-	-	37.4 ± 9.9	1.30 ± 0.21
Rover 16	3-NO_2	3,4-diOMe	-	29.9	64.7	87.3	43.6 ± 14.4	1.03 ± 0.27

Table 5.7. Summary of *Pf*KMO inhibition data for the final set of novel compounds compared to Rover 16. The % inhibition values were measured relative to uninhibited enzyme in 1 ml cuvette assays. For the screen values, the % inhibition represents an average value plus or minus the standard deviation from 3-6 readings in 96-well plate screens. K_i values were determined an average values plus or minus the standard deviation from multiple separate assays. For a selection of inhibition plots for novel compounds see appendix section **4.**

Substitution on the R_3 ring also influenced the strength of *Pf*KMO inhibition, although to a lesser extent than on the R_2 ring. As can be seen from the screening results, all the compounds with the 3,4-diOMe substitution showed greater inhibition than their respective unsubstituted counterparts. For example, the K_i of GM347 was around two-fold lower than GM303, and the K_i of GM466 was three to four-fold lower than that of GM340. All of the compounds with the 3,4-diOMe substitution were observed to be less soluble than those that were unsubstituted. Consequently some other R_3 functional

groups were introduced to try and improve the solubility of the most potent inhibitors GM303 and GM347. The 3,4-dihydroxy substituted compound (GM729) was more soluble and showed similar inhibition levels to GM303. It was the addition of a diethylphosphate group at either the 3' or 4' position that proved most exciting though. GM760 and GM769 both had very low K_i values of around 20 nM but also possessed improved solubility. The four nanomolar *Pf*KMO inhibitors synthesised as a result of these rounds of design, screening and refinement are displayed in figure 5.14. These compounds were ten to fifty-times as potent as Rover 16 (Ro 61-8048) in our assays with *Pf*KMO. A selection of inhibition plots can be found in appendix section **4**.

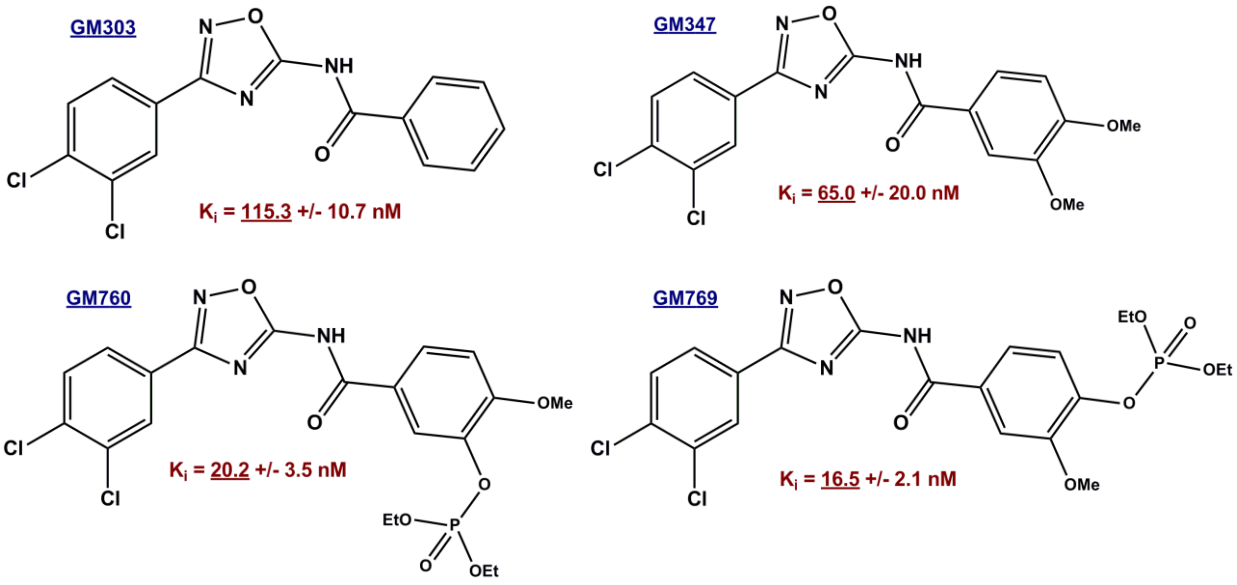


Figure 5.14. Molecular structures of the four most potent *Pf*KMO inhibitors tested. Created using ChemDraw.

Initial attempts to solve the structure of dm2 *Pf*KMO bound with the novel inhibitors synthesised by Gavin Milne were unsuccessful. The protein was mixed 1:1 μL with well solutions containing a number of different inhibitors at 1-2 mM. Crystals grew in the presence of several of the compounds (GM303, GM307, GM469) in the $I4_122$ space group with similar cell dimensions to those obtained for substrate bound *Pf*KMO crystals, but in every case the quality of the electron density was insufficient in order to confidently model the respective inhibitors. Some crystals also grew in the space group $P2_122_1$ in the presence of GM347 and GM443 but were similar to unbound enzyme in

every sense. Attempted sparse-matrix screening to search for alternate crystal growth conditions in the presence of GM347 were fruitless. A happy conclusion was reached though when the compounds were added directly to the thawed aliquot of enzyme for 5-20 mins (room temperature) or 15-30 mins (on ice) prior to the setting up of crystal trays. The inhibitors were added to 1 mM from a 20-50 mM stock that was made in 100 % DMSO. Concomitant with inhibitor addition 1.0 M HEPES pH 8.0 was added to 20 mM final concentration, as the compounds were found to be more soluble with a raised pH. Finally, a move to sitting drop vapour diffusion (s.d.v.d) reproducibly produced more and thicker crystals. All crystals of inhibitor-bound protein grew as needles of varying thickness as typified by the crystallisation drop shown in figure 5.15. Despite numerous efforts, crystals of KMO with either of the Rover compounds bound could not be obtained.

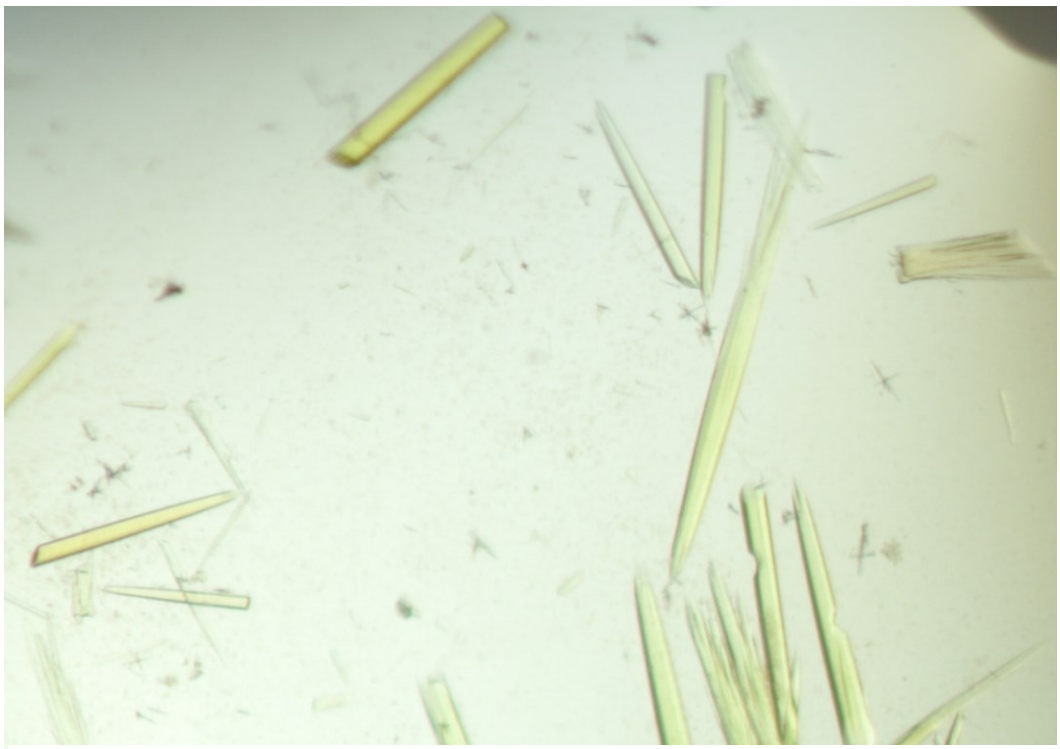


Figure 5.15. Picture of a s.d.v.d crystallisation drop containing optimised GM303-bound dm2 *Pf*KMO crystals.

5.3.2 The Structure of *Pf*KMO with GM303 bound

Crystals of GM303 bound dm2 *Pf*KMO were grown by the s.d.v.d. The drop contained 1.0 µl of GM303 soaked protein solution (13 mg/ml in 20 mM HEPES pH ~ 7.5, 30 mM Na acetate, 1 mM DTT and 1 mM GM303 – from an inhibitor stock of 20 mM in 100 % DMSO) and 1.0 µl of the precipitant solution (0.1 M HEPES pH 7.0, 12 % w/v PEG 4K, 5 % v/v glycerol, 10 mM Na tartrate 10 % v/v 2-propanol). The crystals were much thinner than the substrate-bound crystals, still with a hexagonal prism shape but with a different orientation. The hexagonal face represented the smallest surface whereas it was previously the largest face. The crystals with maximum dimensions of up to 0.05 × 0.05 × 0.4 mm^3 grew in 3-7 days and were harvested for data collection after 6 weeks. Prior to flash-freezing in liquid nitrogen, 3 µl of cryoprotectant containing mother liquor supplemented with 20 % (v/v) ethylene glycol and 1 mM GM303, was added directly to the crystallisation drop.

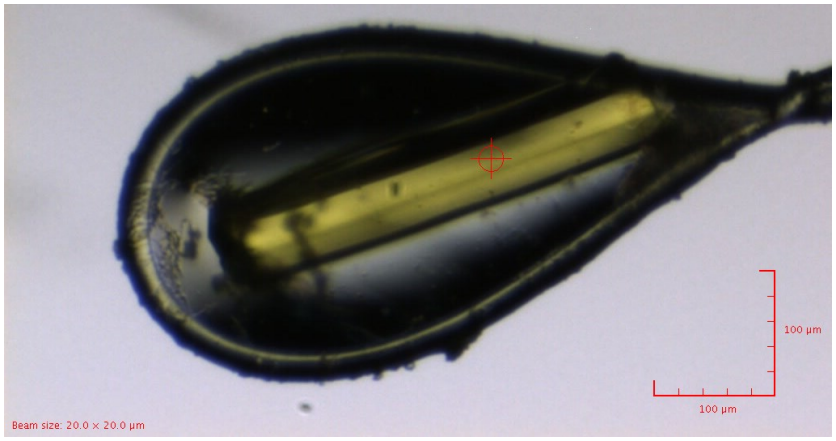


Figure 5.16. Picture of one of the mounted GM303-bound crystals tested with diffraction below 2.5 Å.

The scale and beam size are displayed.

The crystals (as shown in figure 5.16) were screened for diffraction quality at the Diamond synchrotron beamline I03 with several diffracting to a resolution below 2.5 Å with a new set of lattice properties. X-ray diffraction data were collected using an X-ray wavelength of 0.9763 Å, exposure of 0.5 s/frame of data and beam intensity of 80 % with 0.5 º oscillations over a total of 250 frames. The crystal was found to be hexagonal - space group $P6_122$ (after testing different point groups with Molrep) - with mean unit cell parameters of a = b = 117.2, c = 418.2 Å, α = β = 90.0, γ = 120.0 º. Due to the large unit cell dimensions on the crystal's c-axis, reflections were very close together which caused difficulties with automatic indexing in Mosflm. To overcome this, the mosaicity

and spot separation were entered manually. The dataset was 100.0 % complete to a resolution of 2.18 Å with R_{merge} = 0.136 overall and 0.547 in the highest resolution shell. The full data collection statistics are tabulated in appendix section **3.6**.

For the calculation of structure factors, a suitable model was generated using molecular replacement (Molrep) with the previously solved substrate bound structure (see section **3.3**). A definitive solution emerged with three protein molecules present in the asymmetric unit yielding a solvent content of 54.9 %. The model was cycled through manual editing in Coot and restrained refinement in Refmac. For the first few cycles NCS restraints for each chain were applied in refinement. There was clear density for one molecule of FAD and one molecule of the inhibitor GM303 for each chain. The electron density for the inhibitor is shown in figure 5.17 and with the relatively high resolution of the data it was possible to establish the correct orientation of the molecule with confidence. On all three chains the residues 6-461 were included, except that there was insufficient clear density to unambiguously place residues 378-381 which precedes the C-terminal domain in chain A. Towards the end of refinement, Coot was used to manually place 783 waters in the model. The final model was refined with the addition of TLS to give R and R_{free} values of 0.197 and 0.238 respectively, a figure of merit of 0.862 and an overall mean B factor of 29.5 for 85024 significant reflections in the range of 46.5-2.18 Å. The rmsd of bond lengths and angles from the ideal were 0.005 Å and 1.10 ^{o} and all the residues were present in the most favoured or additionally allowed regions of a Ramachandran plot as determined using the Molprobity server. Full refinement statistics are tabulated in appendix section **3.6**.

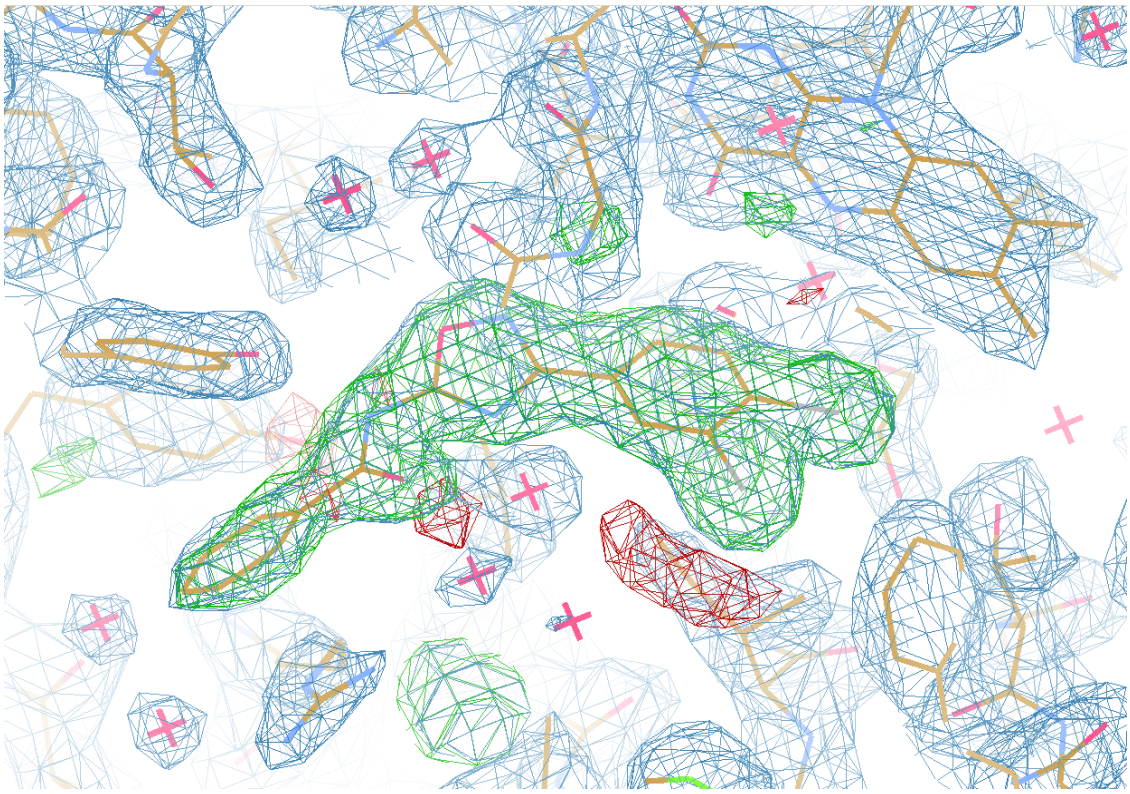


Figure 5.17. Fit of the bound molecule of GM303 in the active site of the structure. The refined 2 F_o-F_c density (contoured at 1.5 σ) for the final position of the molecule is shown in blue with the F_o-F_c map, calculated after refinement without the inhibitor (contoured at 3.5 σ), shown in green. Figure made using Coot.

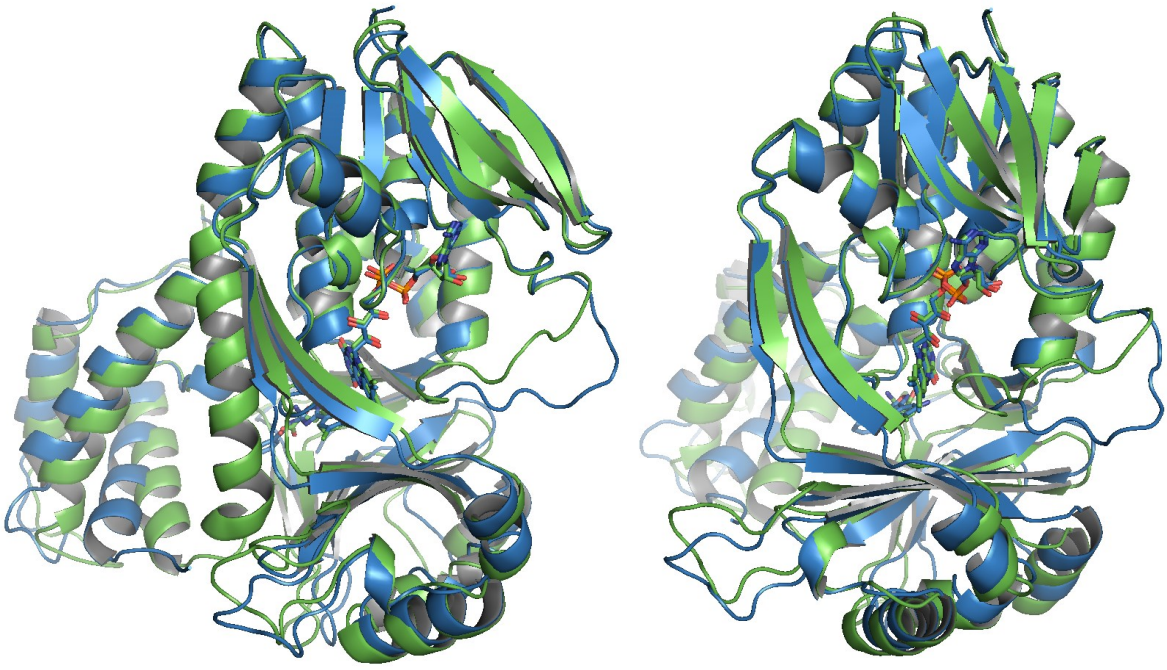


Figure 5.18. Two views of a cartoon overlay of the structures of substrate-bound (green) and GM303-bound (blue) dm2 *Pf*KMO with FAD and the ligands shown in sticks. Made using PyMol.

The GM303-bound *Pf*KMO structure is displayed aligned with the substrate-bound structure in figure 5.18. The average rmsd of C_α atoms of the chains from the GM303 bound structure was 1.8 Å compared to substrate (and substrate analogue) bound KMO and 0.8 Å compared to unbound *Pf*KMO. Therefore the structure in fact appeared to more strongly resemble the unbound structure from section **3.5**. Whilst the inhibitor was bound in the same pocket as the substrate/substrate analogues, the polypeptide chain had undergone several major alterations. Firstly the C-terminal domain was in the open, substrate-free conformation with a swing of 6.4 Å (comparing the C_α atoms of aligned T381 residues) relative to the original substrate bound conformation. The second major change was that the long external loop including residues 42-52 was clearly fixed in a more expansive conformation in all three chains. Density for such a conformation was also seen to exist in two of the four chains from the solved unbound structure (see section **3.5**). It is possible though that in the inhibitor bound structure the expansive conformation may have been an artefact of its new arrangement in terms of crystal lattice packing. As can be seen in figure 5.18, domain 2 at the bottom of the enzyme is slightly displaced relative to all the *Pf*KMO structures solved to date. It is possible that this change may have stabilized the loop in the second conformation.

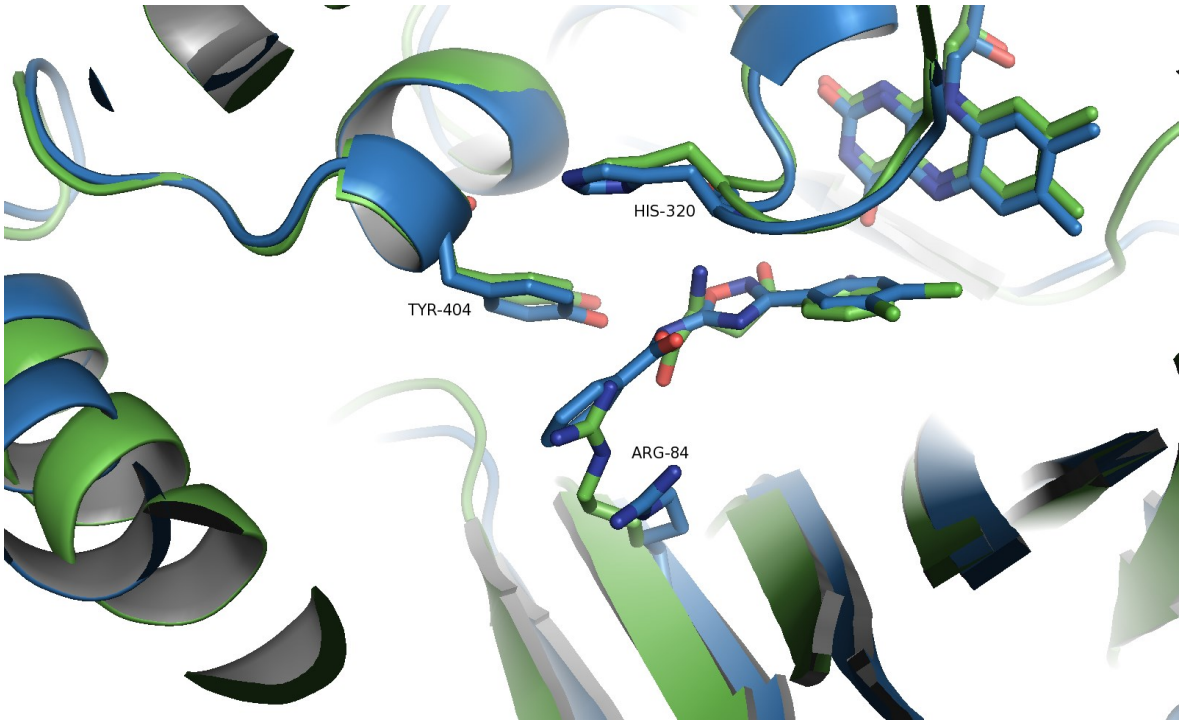


Figure 5.19. Comparison of the aligned active site environment from the substrate-bound (green) and GM303-bound (blue) structures. Ligands and key residues are shown in sticks. A glimpse of the more open C-terminus can be seen by the helix displacement on the left of the picture. Figure created using PyMol.

With a resolution of 2.18 Å the electron density was sufficient to accurately position the inhibitor and its functional groups. The 3,4-dichloro ring of GM303 was positioned and oriented just as was seen when 3,4-CBA was bound with the two chloro groups pointing away from the flavin. The carbon atom of GM303 corresponding to the hydroxylation site of L-Kyn is 4.1 Å away from the C4a of the flavin isoalloxazine ring. All of the potential hydrogen bonds that stabilize GM303 binding appear to involve the amide group that proved pivotal to compound inhibition during screening. The amide nitrogen shares a hydrogen bond (2.6 Å) with the hydroxyl of the side chain of Y404, which was previously not utilized when substrate was bound. The side chain of N369, which is involved in positioning the carbonyl group of L-Kyn, hydrogen bonds with the amide oxygen of GM303 *via* a water molecule (2.7 Å and 2.6 Å). GM303's unsubstituted phenyl ring then sits where the guanidinium group of R84 was located in the substrate bound structure. As can be seen in figure 5.19, R84 is pushed aside by around 3.5 Å leaving the phenyl ring of GM303 stacked between the side chains of R84 and Y404. This obstruction by the protruding end of GM303 may hinder the formation of the network of interactions that was responsible for closing domain 3 in the presence of substrate. With few other significant changes observed in the active site it may be this displacement of R84 that causes the other structural changes observed upon GM303 binding. Therefore it appeared from the structural features of the GM303-bound structure that the inhibitor was not recognized as a substrate as suggested by the lack of observable uncoupling of NADPH oxidation.

5.3.3 The Structure of *Pf*KMO with GM347 bound

Crystals of GM347 bound dm2 *Pf*KMO were grown by s.d.v.d. The drop contained 1.0 μl of GM347 soaked protein solution (13 mg/ml in 20 mM HEPES pH ~ 7.5, 30 mM Na acetate, 1 mM DTT and 1 mM GM347 – from an inhibitor stock of 20 mM in 100 % DMSO) and 1.0 μl of the precipitant solution (0.1 M HEPES pH 7.0, 12 % w/v PEG 4K, 8 % v/v glycerol, 20 mM Na Tartrate, 10 % v/v 2-propanol). Thin, hexagonal prism crystals (similar to with GM303-bound) with rough dimensions of up to $0.05 \times 0.05 \times 0.40$ mm^3 grew in 3-7 days and were harvested for data collection after 10 days. Prior to flash-

freezing in liquid nitrogen, 1.5 µl of cryoprotectant containing mother liquor supplemented with 33 % (v/v) ethylene glycol and 2 mM GM347, was added directly to the drop.

The crystals were screened for diffraction quality at the Diamond synchrotron beamline I04-1 with several diffracting to a resolution below 2.3 Å with similar lattice properties to the GM303 bound KMO crystals. An example crystal is shown in figure 5.20. X-ray diffraction data were collected using an X-ray wavelength of 0.920 Å, exposure of 0.3 s/frame of data with 0.3 º oscillations over a total of 131.1 º. The crystal was found to belong to the hexagonal space group $P6_122$ with mean unit cell parameters of a = b = 117.7, c = 420.1 Å, α = β = 90.0, γ = 120.0 º. The data were indexed and integrated in Mosflm and scaled in Scala as before. The dataset was 99.9 % complete to a resolution of 2.15 Å with R_{merge} = 0.137 overall (0.549 HS) and an intensity of 11.0 (2.7). The full data collection statistics are tabulated in appendix section **3.6**.

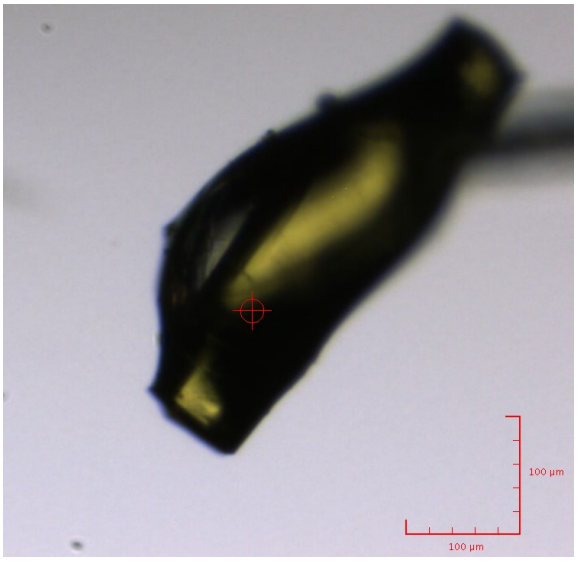


Figure 5.20. Picture of a mounted GM347-bound crystal, which diffracted to a resolution around 2.25 Å. Beam-size = 20 × 20 µm.

For the calculation of structure factors, a suitable model was generated using molecular replacement (Molrep) with the previously solved GM303 bound structure. A definitive solution emerged again with three protein molecules present in the asymmetric unit, yielding a solvent content of 55.4 %. The model was cycled through manual editing in Coot and restrained refinement in Refmac. For the first few cycles NCS restraints for each chain were applied in refinement. There was clear density for one molecule of FAD for each chain. There was also clear density to place the 3,4-dichloro R_2 ring, the 5-

membered heterocycle ring and the amide linker of GM347 in a similar location to that seen with GM303. Unfortunately the density for the 3,4-OMe R_3 ring was weak and ambiguous meaning that it could not be accurately modelled. Several datasets of similar resolution from different crystals harvested from different drops were processed, but still the same problem consistently remained. Therefore the R_3 ring was excluded from the final model and the refined electron density maps calculated in the presence and absence of this truncated inhibitor are shown in figure 5.21. On all three chains the residues 6-461 were included except that there was insufficiently clear density to unambiguously place residues 378-379, 375-378 and 377-381 in chains B, A and C respectively. Towards the end of refinement, Coot was used to manually place 816 waters in the model. The final model (shown in figure 5.22) was refined with the addition of TLS to give R and R_{free} values of 0.208 and 0.234 respectively and an overall mean B-factor of 31.4 $Å^2$ for 89610 significant reflections in the range of 45.00-2.15 Å. The rmsd of bond lengths and angles from the ideal were 0.005 Å and 1.12 ^{o} and all the residues were present in the most favoured or additionally allowed regions of a Ramachandran plot as determined using the Molprobity server. Full refinement statistics are tabulated in appendix section **3.6**.

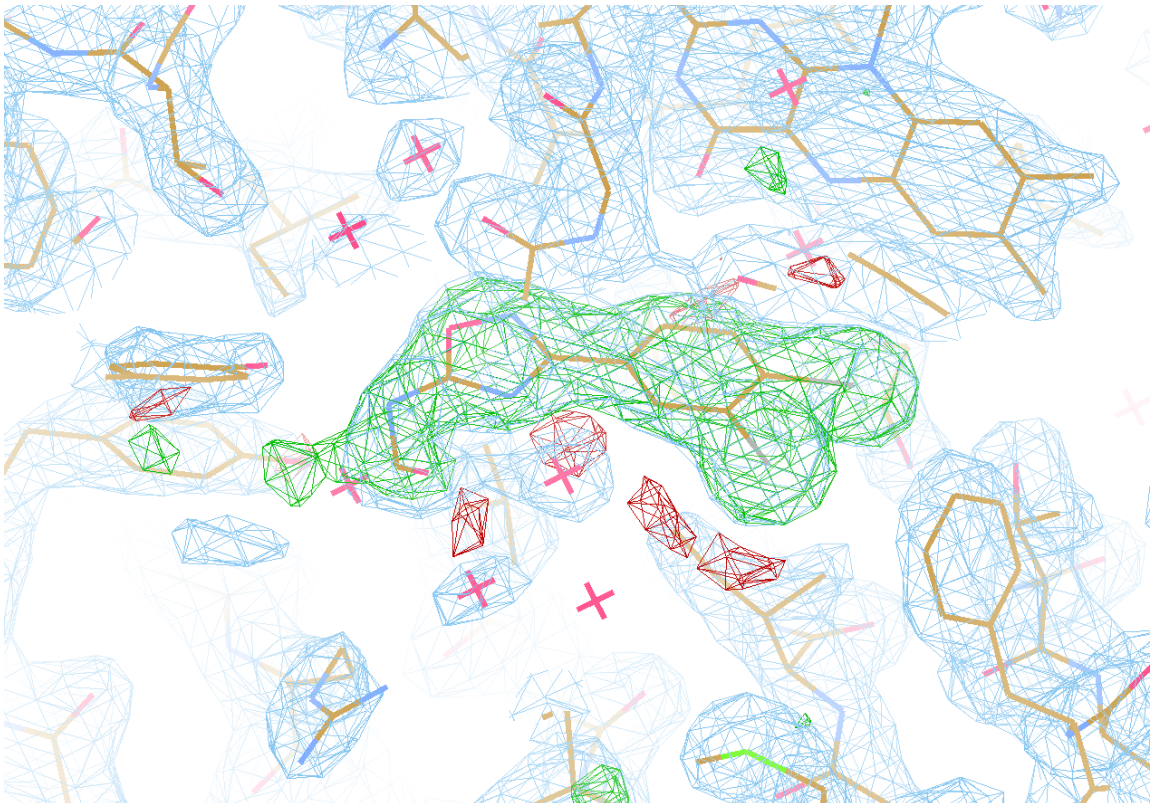


Figure 5.21. Fit of the truncated molecule of GM347 in the active site of the structure. The refined $2F_o$-F_c density (contoured at 1.5 σ) for the final position of the molecule is shown in blue with the F_o-F_c map calculated after refinement withouth the inhibitor present (contoured at 3.5 σ) shown in green. Figure made using Coot.

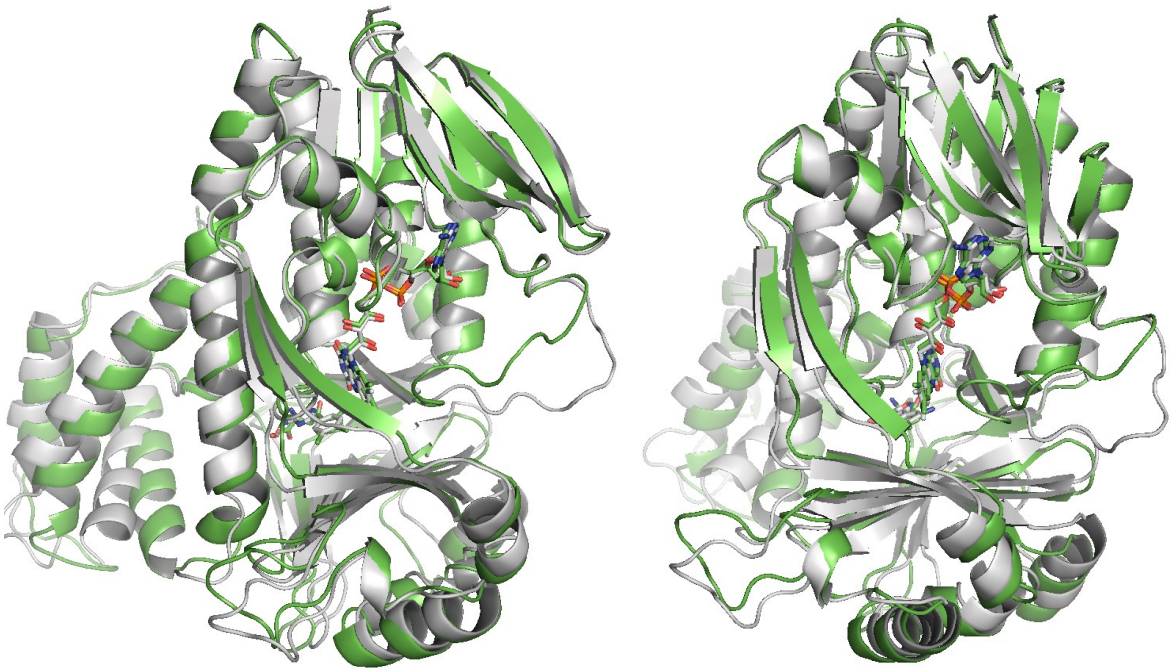


Figure 5.22. Side and front views of a cartoon overlay of the structures of substrate-bound (green) and GM347-bound (grey) dm2 *Pf*KMO with FAD and the ligands shown in sticks. Made using PyMol.

The overall structure of GM347 bound *Pf*KMO was very similar to the previous GM303 bound structure, which had structural features that more resembled unbound over substrate bound *Pf*KMO. The average rmsd of C_α atoms for all three chains of the GM347 bound structure was: 0.4 Å compared to GM303 bound, 1.8 Å compared to substrate (and substrate analogue) bound and 0.9 Å compared to unbound *Pf*KMO. As previously detailed with the GM303 bound structure, the C-terminal domain 3 and external 42-52 loop were both present in their respective open conformations. The parts of GM347 that could be unambiguously positioned were found to adopt the same orientation as observed for GM303, with the 3,4-dichloro groups clearly pointing out away from the flavin. It is unclear why the 3,4-OMe R_3 ring could not be located by clear density given the high resolution of the datasets tested. It is possible that either the compound started to decompose or that the ring occupied multiple conformations upon binding. The latter is more likely as the density for the neighbouring side-chain of Y404 was also relatively weak and no deterioration in the inhibition strength was noticed over time during inhibition assays and screening. With the basic R84 side-chain nearby it was hoped that binding of the two strongest inhibitors, which contained R_3 $PO_2(OEt)_2$ groups, would be more stable.

5.3.4 The Structure of *Pf*KMO with GM760 bound

Crystals of GM760 bound dm2 *Pf*KMO were grown by s.d.v.d. The drop contained 1.0 µl of GM760 soaked protein solution (13 mg/ml in 20 mM HEPES pH ~ 7.5, 30 mM Na acetate, 1 mM DTT and 1 mM GM760 – from an inhibitor stock of 50 mM in 100 % DMSO) and 1.0 µl of the precipitant solution (0.1 M HEPES pH 7.0, 13 % w/v PEG 4K, 6 % v/v glycerol, 10 mM Na tartrate, 10 % v/v 2-propanol). Thin, hexagonal prism crystals (similar to with GM303-bound) with rough dimensions of $0.04 \times 0.04 \times 0.80$ mm^3 grew in 3-7 days and were harvested for data collection after 10 days. Prior to flash-freezing in liquid nitrogen, 1 µl of 100 % (v/v) ethylene glycol was added directly to the drop.

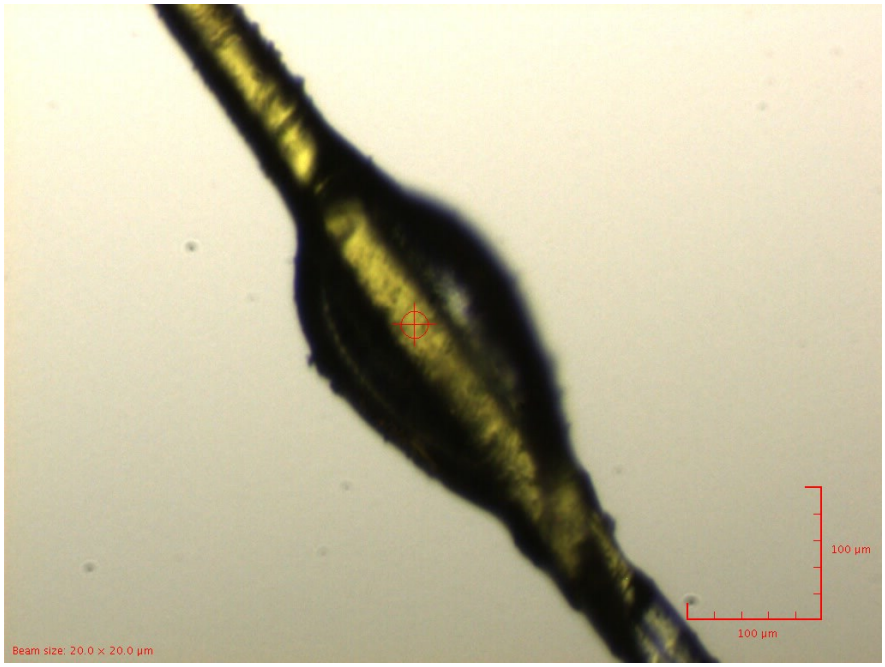


Figure 5.23. Picture of the mounted GM760-bound crystal used for data collection.

The scale and beam size are labelled.

The crystals were screened for diffraction quality at the Diamond synchrotron beamline I02 with several diffracting to a resolution below 2.5 Å with similar lattice properties to the GM303 bound crystals. X-ray diffraction data were collected using an X-ray wavelength of 0.979 Å, exposure of 0.3 s/frame of data and beam intensity of 50 % with 0.3 º oscillations over a total of 117 º. The crystal (displayed in figure 5.23) belonged to space group $P6_122$ with mean unit cell parameters of a = b = 117.3, c = 418.1 Å, α = β = 90.0, γ = 120.0 º. The data were indexed and integrated in Mosflm and scaled in Scala as

before. The dataset was 100.0 % complete to a resolution of 2.40 Å with R_{merge} = 0.156 overall (0.682 HS) and intensity 11.0 (2.6 HS). The full data collection statistics are tabulated in appendix section **3.6**.

The GM303 bound structure was used as a model to start refinement. The model was cycled through manual editing in Coot and restrained refinement in Refmac. For the first few cycles NCS restraints for each chain were applied in refinement. There was clear density for one molecule of FAD and one molecule of the inhibitor GM760 for each chain. The refined density with and withouth the inhibitor is shown in figure 5.24 and with the relatively high resolution of the data the position of the functional groups was established. The positions of the oxygen and OEt groups on the phosphate though could not be differentiated with 100 % confidence and may be interchangeable. On all three chains the residues 6-461 were included except that there was insufficient density to clearly place residues 376-382, which precede the C-terminal domain in chain A. Towards the end of refinement, Coot was used to manually place 940 waters in the model. The final model was refined with the addition of TLS to give R and R_{free} values of 0.203 and 0.248 respectively and an overall mean B-factor of 44.4 for 64181 significant reflections in the range of 50.0-2.40 Å. The rmsd of bond lengths and angles from the ideal were 0.005 Å and 1.08 ^{o} and all the residues were present in the most favoured or additionally allowed regions of a Ramachandran plot as determined using the Molprobity server. Full refinement statistics are tabulated in appendix section **3.6**.

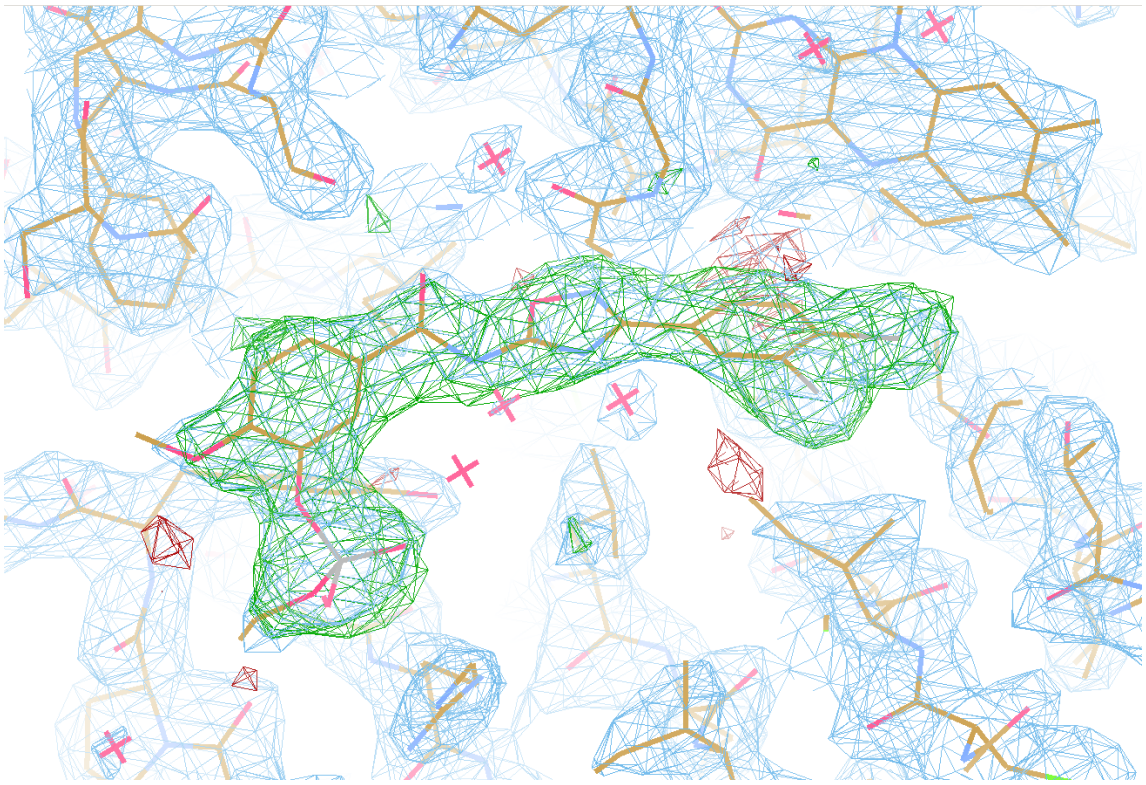


Figure 5.24. Fit of the bound molecule of GM760 in the active site of the structure. The refined $2F_o$-F_c (contoured at 1.5 σ) for the final position of the molecule is shown in blue with the F_o-F_c map (contoured at 3.5 σ) in the absence of inhibitor shown in green. Figure made using Coot.

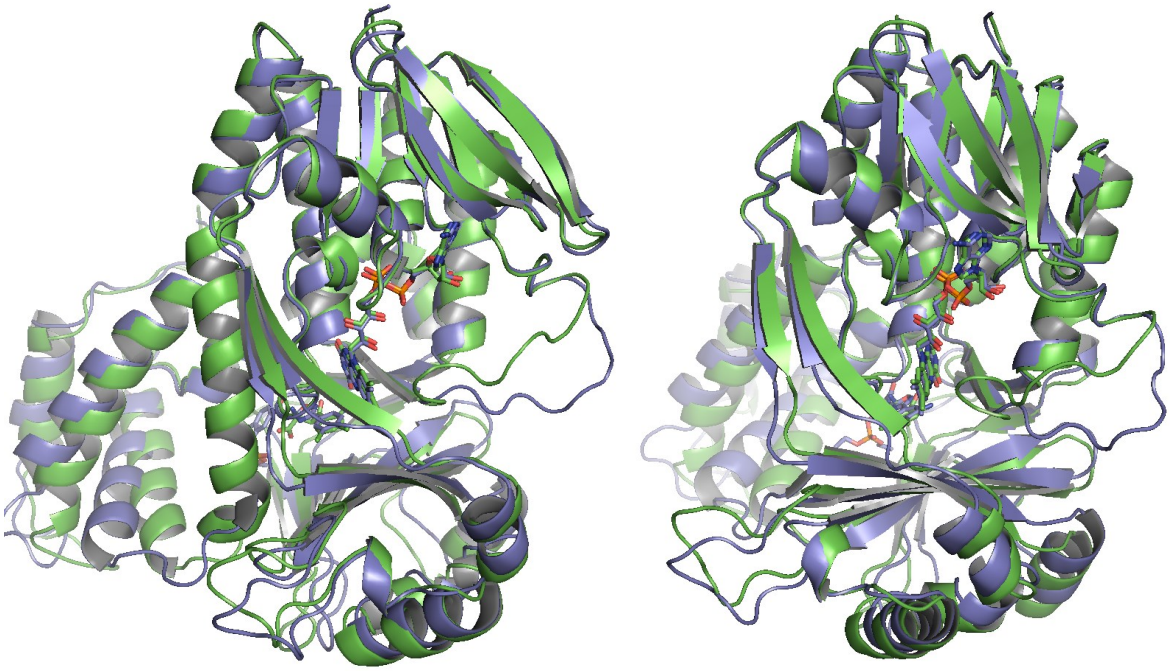


Figure 5.25. Side and front views of a cartoon overlay of the structures of substrate-bound (green) and GM760-bound (violet) dm2 *Pf*KMO with FAD and the ligands shown in sticks. Made using PyMol.

structural features that more resembled unbound over substrate bound *Pf*KMO. The rmsd of C_α atoms of a chain of the GM760 bound structure was: 0.5 Å compared to GM303 bound, 1.8 Å compared to substrate (and substrate analogue) bound and 0.9 Å compared to unbound *Pf*KMO. As previously detailed with the GM303 and GM347 bound structures, the C-terminal domain 3 and external 42-52 loop were both present in their respective open conformations. The inhibitor was again bound in the same pocket as before and is shown in more detail in figure 5.26.

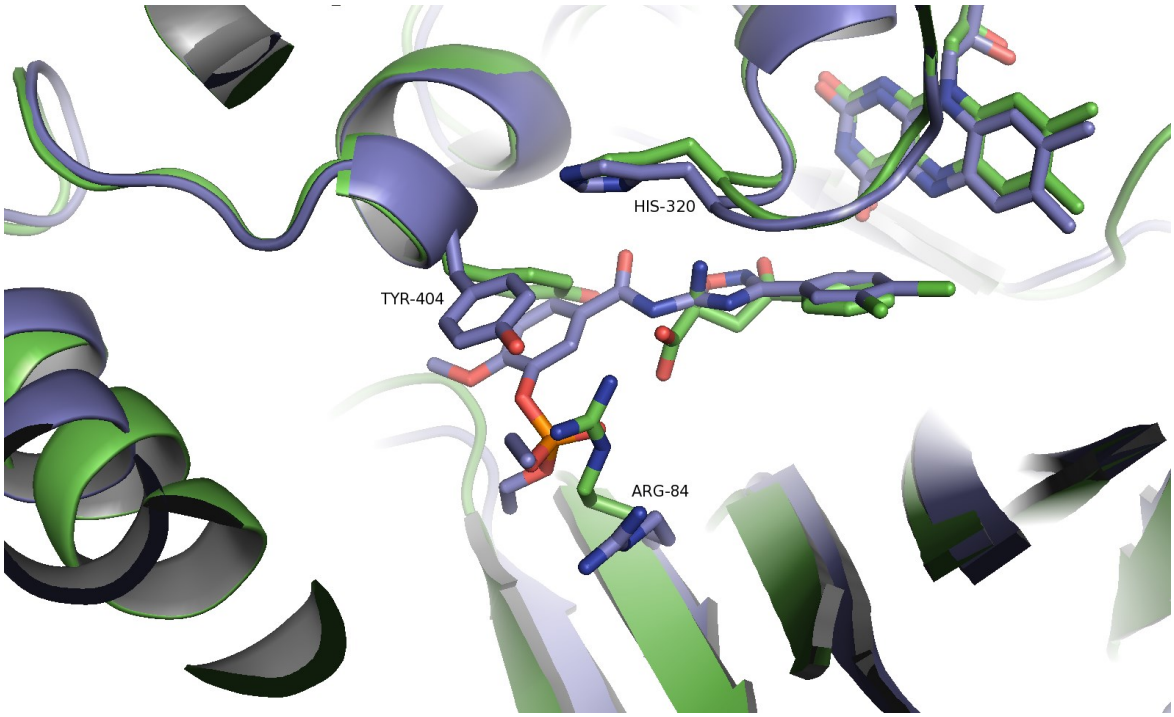


Figure 5.26. Comparison of the aligned active site environment from the substrate-bound (green) and GM760-bound (violet) structures. Ligands and key residues are shown in sticks. A glimpse of the more open C-terminus can be seen by the helix displacement on the left of the picture. Figure created using PyMol.

Interestingly, with the large 3'-$PO_2(OEt)_2$ group added to the R_3 ring of GM760, a couple of significant local changes were enforced compared to with GM303 bound. The 3'-$PO_2(OEt)_2$ group pointed out of the binding pocket to be adjacent to the imidazole group of R84 at a distance of Å. This pushed the R_3 ring deeper into the buried pocket, which presumably caused the observed out-swing of the side chain of Y404 by around 120 ^{o} relative to the substrate/substrate analogue bound, unbound and GM303 bound

structures. Therefore in the GM760 bound structure Y404 sits in the plane above the R_3 ring with a possible π-stacking interaction. To accommodate this translocation, the inhibitor's amide linker is rotated by around 150 ^{o} relative to GM303 and GM347 to point in to the hydrophobic pocket below the side chain of H320. The carbonyl oxygen atom now formed potential hydrogen bonds with the hydroxyl of T408 (2.6 Å) and a water molecule (3.0 Å). In addition two water molecules formed potential hydrogen bonds with the amide nitrogen atom (both at around 2.7 Å).

5.3.5 The Structure of *Pf*KMO with GM769 bound

Crystals of GM769 bound dm2 *Pf*KMO were grown by s.d.v.d. The drop contained 1.0 μl of GM769 soaked protein solution (13 mg/ml in 20 mM HEPES pH ~ 7.5, 30 mM Na acetate, 1 mM DTT and GM769 – from an inhibitor stock of 50 mM in 100 % DMSO) and 1.0 μl of the precipitant solution (0.1 M HEPES pH 7.0, 11 % w/v PEG 4K, 6 % v/v glycerol, 10 mM Na tartrate, 10 % v/v 2-propanol). Long, thin, hexagonal prism crystals (similar to with GM303-bound) with rough dimensions of 0.05 × 0.05 × 0.50 mm^3 grew in 3-7 days and were harvested for data collection after 3-4 weeks. Prior to flash-freezing in liquid nitrogen, 3 μl of cryoprotectant containing well solution supplemented with 20 % (v/v) ethylene glycol and 1 mM GM769, was added directly to the drop.

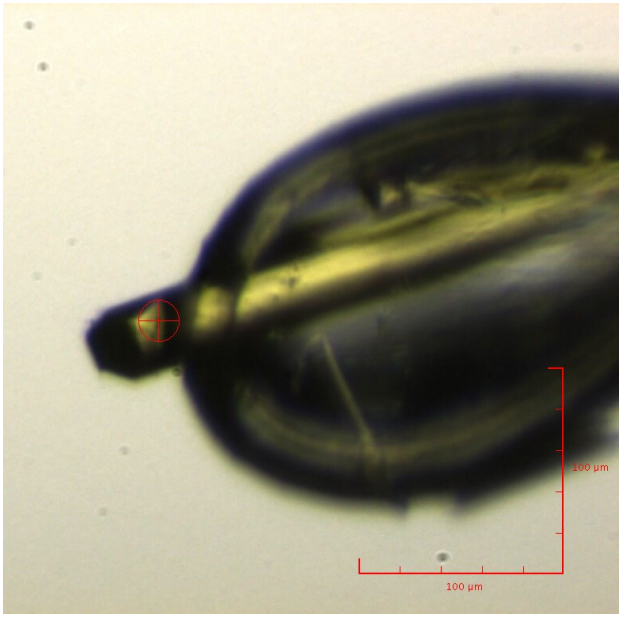


Figure 5.27. Picture of the mounted GM769-bound crystal used for data collection.

The beam size was 20 × 20 μm.

The crystals were screened for diffraction quality at the Diamond synchrotron beamline I02 with several diffracting to a resolution below 2.5 Å with similar lattice properties to the previous inhibitor bound crystals. X-ray diffraction data were collected from the crystal displayed in figure 5.27. An X-ray wavelength of 0.979 Å, exposure of 0.35 s/frame of data and beam intensity of 55 % with 0.3 º oscillations over a total of 153.6 º was used. The crystal was monoclinic, space group $P6_122$ with mean unit cell parameters of a = b = 117.5, c = 417.8 Å, α = β = 90.0, γ = 120.0 º. The data were indexed and integrated in Mosflm and scaled in Scala as before. The dataset was 99.9 % complete to a resolution of 2.30 Å with R_{merge} = 0.157 overall (0.675 HS) and I/σI 12.5 (3.2 HS). The full data collection statistics are tabulated in appendix section **3.6**.

The previously solved GM303 bound structure was again used to start refinement. The model was developed as described previously. There was clear density for one molecule of FAD and one molecule of the inhibitor GM769 for each chain. The refined density in the presence and absence of the inhibitor is shown in figure 5.28 and with the relatively high resolution of the data the position of the functional groups was established. The positions of the oxygen and OEt groups on the phosphate again could not be differentiated with 100 % confidence and may be interchangeable. On all three chains the residues 6-461 were included except that there was insufficient density to clearly place residues 377-379, 378-380 and 375-377 in chains B, A and C. Towards the end of refinement, Coot was used to manually place 820 waters in the model. The final model was refined with the addition of TLS to give R and R_{free} values of 0.202 and 0.230 respectively and an overall mean B-factor of 37.0 for 72900 significant reflections in the range of 38.80-2.30 Å. The rmsd of bond lengths and angles from the ideal were 0.005 Å and 1.05 º and all the residues were present in the most favoured or additionally allowed regions of a Ramachandran plot as determined using the Molprobity server. Full refinement statistics are tabulated in appendix section **3.6**.

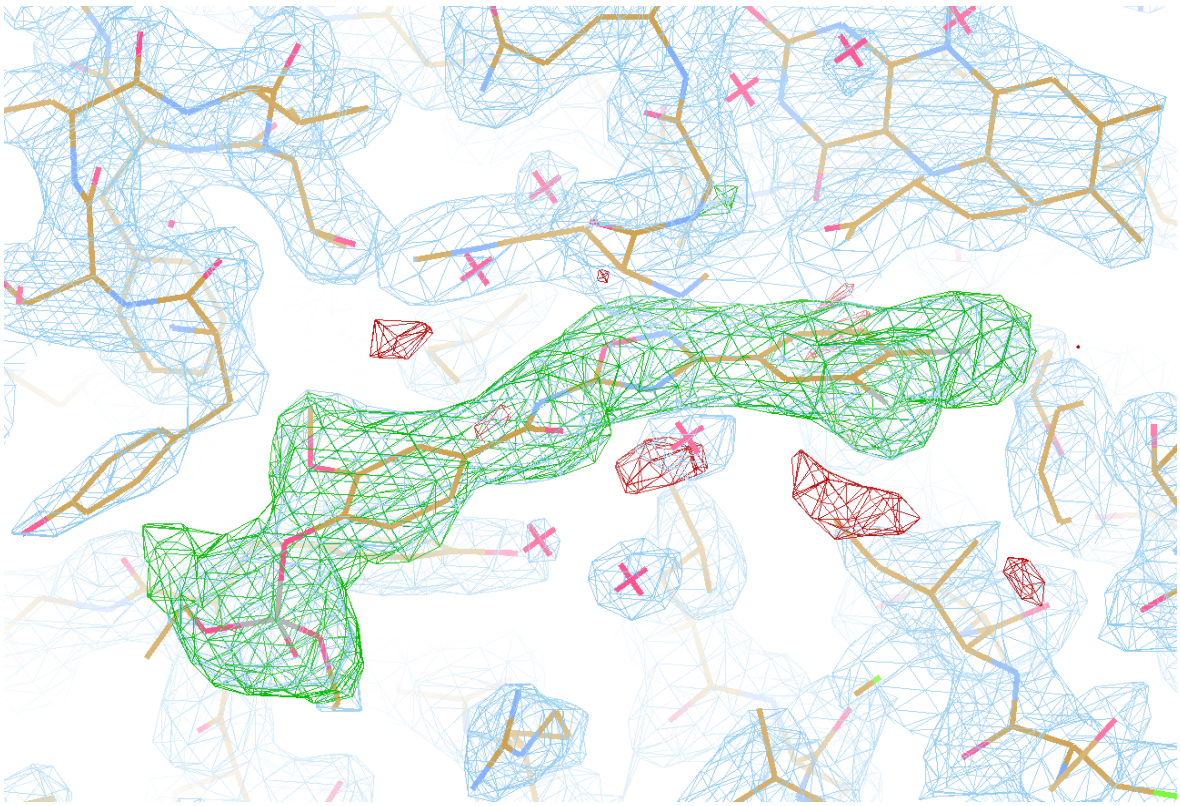


Figure 5.28. Fit of the bound molecule of GM769 in the active site of the structure. The refined $2F_o$-F_c density (contoured at 1.5 σ) for the final position of the molecule is shown in blue with the F_o-F_c map (contoured at 3.5 σ) calculated in the absence of GM769 shown in green. Figure made using Coot.

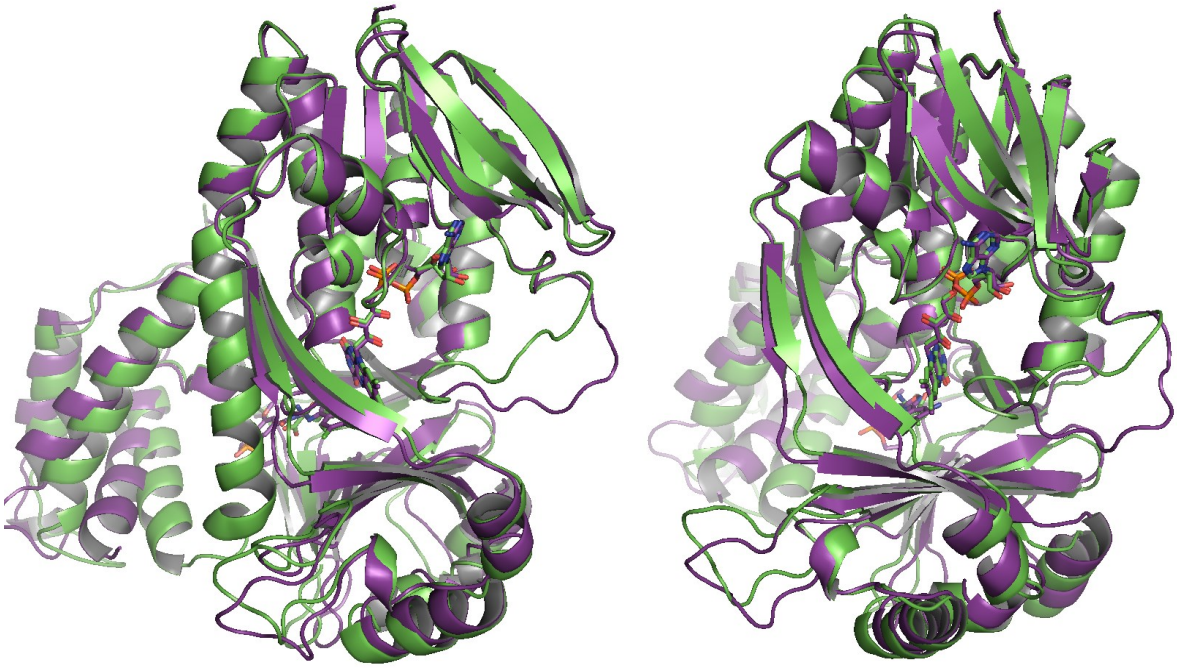


Figure 5.29. Side and front views of a cartoon overlay of the structures of substrate-bound (green) and GM769-bound (purple) dm2 *Pf*KMO with FAD and the ligands shown in sticks. Made using PyMol.

The overall structure of GM769 bound dm2 *Pf*KMO is aligned to the substrate-bound form in figure 5.29. The structure was very similar to the previous inhibitor bound structures, which had structural features that more resembled unbound over substrate bound *Pf*KMO. As discussed before with the other inhibitor bound structures, the C-terminal domain 3 and external 42-52 loop were both present in their respective open conformations. The rmsd of C_α atoms of a chain of the GM769 bound structure was: 0.4 Å compared to GM303 bound, 0.5 Å compared to GM760 bound, 1.8 Å compared to substrate (and substrate analogue) bound and 0.9 Å compared to unbound *Pf*KMO. The inhibitor was bound in the same pocket as before and is shown in more detail in figure 5.30.

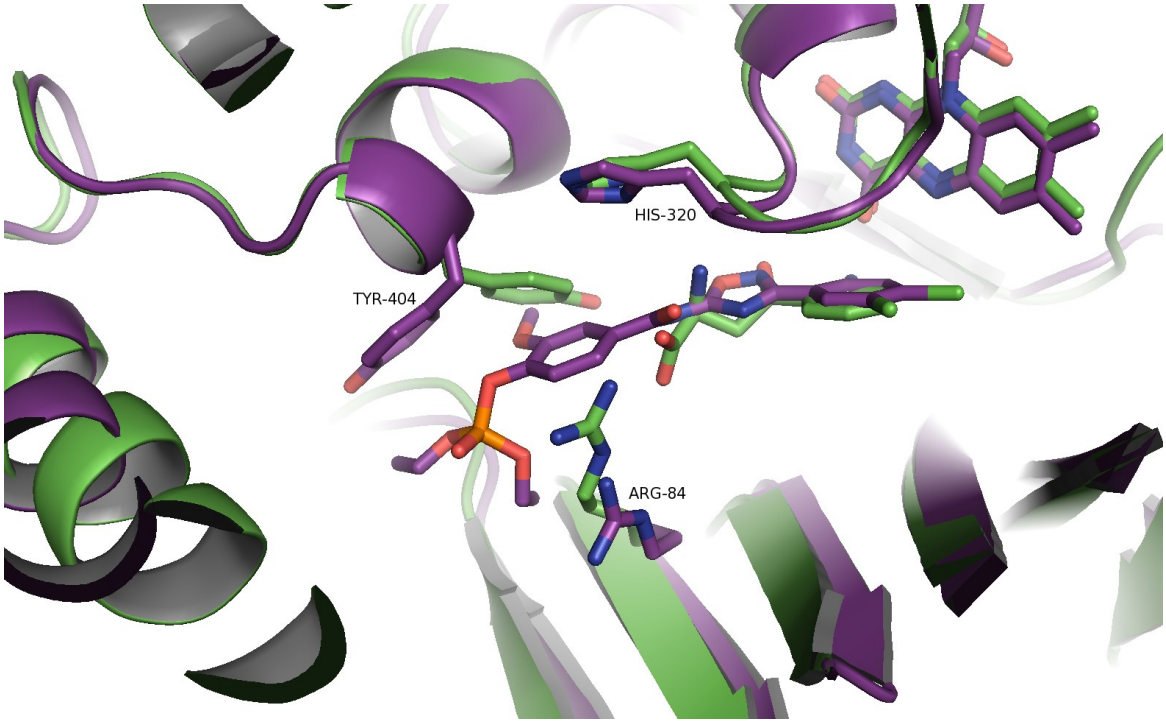


Figure 5.30. Comparison of the aligned active site environment from the substrate-bound (green) and GM769-bound (purple) structures. Ligands and key residues are shown in sticks. A glimpse of the more open C-terminus can be seen by the helix displacement on the left of the picture. Figure created using PyMol.

Interestingly, with the $PO_2(OEt)_2$ group on the 4' rather than the 3' position in GM760, the amide linker flips by 180 ˚ to resume the orientation occupied in bound GM303 (and

possibly GM347). Moreover a novel third conformation for the side-chain of Y404 is seen with a rotation of around -120 ˚ relative to the substrate/substrate analogue bound, unbound and GM303 bound structures. R84 is again displaced but now it was the 4'-$PO_2(OEt)_2$ group that was sandwiched between the side chains of R84 and Y404. The 3'-OMe group of GM769 points in to the vacant space left by the side-chain of Y404. This demonstrates the flexibility of the residues lining the active site to accommodate various alterations to the functional groups of the novel inhibitors.

5.4 Discussion and conclusions

In section **5.2** the binding characteristics of substrate analogues was investigated. Background NADPH oxidation in the absence of L-Kyn confirmed the analogues were non-substrate effectors rather than inhibitors as first discussed by Crozier-Reabe [1]. From *in vitro* assays with pure *Pf*KMO, it was found the addition of electronegative functional groups to the compound's phenyl ring drastically improved its binding affinity. Interestingly, little effect was seen upon conversion into a naphthalene ring system showing that there was a large hydrophobic pocket to be filled. This is consistent with the spacious hydrophobic pocket seen adjacent to the outer side of the benzene ring of the substrate in the previously described crystal structures. Crystal structures with the tight binding published substrate analogues *m*-NBA and 3,4-CBA bound revealed that the compounds bound in the same location and orientation as the substrate L-Kyn. The C-terminal domain 3 was positioned in the closed, substrate-bound position suggesting that the compounds were initially recognized as potential substrates upon binding. It is therefore possible that the apparent uncoupling of NADPH oxidation was due to either: i) the loss of an interaction from the amine group on L-Kyn with the catalytic flavin-intermediate and/or ii) a slight displacement in the position of the target benzene ring. The key information for taking on to the search for novel *Pf*KMO inhibitors was that the addition of functional groups such as *m*-nitro and 3,4-dichloro to the benzene ring greatly improved the strength of binding.

Following this information, a novel series of *Pf*KMO inhibitors was successfully discovered and refined. The compounds were based on the structure of the published

potent KMO inhibitor Ro 61-8048, synthesised by Rover *et al*. [4]. Ro 61-8048 initially showed therapeutic potential as discussed in section **1.4.2.** It is however, unsuitable for therapeutic use as it failed the Ames test, which was probably due to the presence of a nitro group. Given the low published IC_{50} values of around 20-50 nM for the Rover compounds, the strength of inhibition in our assays on *Pf*KMO was a little weaker than expected. The published values were determined using mitochondrial extracts from rat liver KMO and so there was a possibility that the differences were due to subtle variations between the bacterial and rat KMO active site environments. The strength of inhibition in *Pf*KMO was still reasonably strong though, and enough to suggest a high level of similarity between the binding site in the two enzymes. Therefore, together with the clear conservation seen with the active site residues in the previously solved bacterial KMO structures, it was concluded that *Pf*KMO presents a suitable model for developing potential inhibitors of human KMO. We therefore set out to generate a new set of compounds that could reproduce the inhibitory properties of Ro 61-8048 with alternative, safer functional groups.

Subsequently, Gavin Milne (PhD, University of St Andrews) designed an initial 1,2,4-oxadiazole-based scaffold to mimic the 4-phenylthiazol-2-yl of the Rover compounds. From this starting point, compound refinement and screening identified the key functional groups required for potent *Pf*KMO inhibition. The combination of a tail benzene ring with an amide linker successfully created a novel secondary scaffold with similar levels of *Pf*KMO inhibition as the Rover compounds. The inhibition constant K_i of the novel compound (GM141) was determined to be 1.30 ± 0.21 μM compared to 1.03 ± 0.27 μM and 1.67 ± 0.23 μM for Rover 16 (Ro 61-8048) and 19 respectively. After this point the addition/variation of functional groups associated with the 3' and 4' positions of both the head R_2 and tail R_3 benzene rings identified even more potent inhibitors with nanomolar inhibition constants. On the R_2 ring, the 3,4-dichloro moieties gave the strongest inhibition, similar to the pattern seen with the substrate analogues. At the tail R_3 ring the addition of a 3' or 4'-$PO_2(OEt)_2$ group further increased the inhibition strength. The process of screening and compound refinement finished with compounds GM760 and GM769, which had essentially identical K_i values of 20.2 ± 3.5 nM and 16.5 ± 2.1 nM respectively. Importantly, both compounds were largely soluble in water unlike

the majority of the compounds which did not contain the 3' or 4'-$PO_2(OEt)_2$ group. The plate-screening and Ki data for the majority of the compound series is summarised in the graph in figure 5.31.

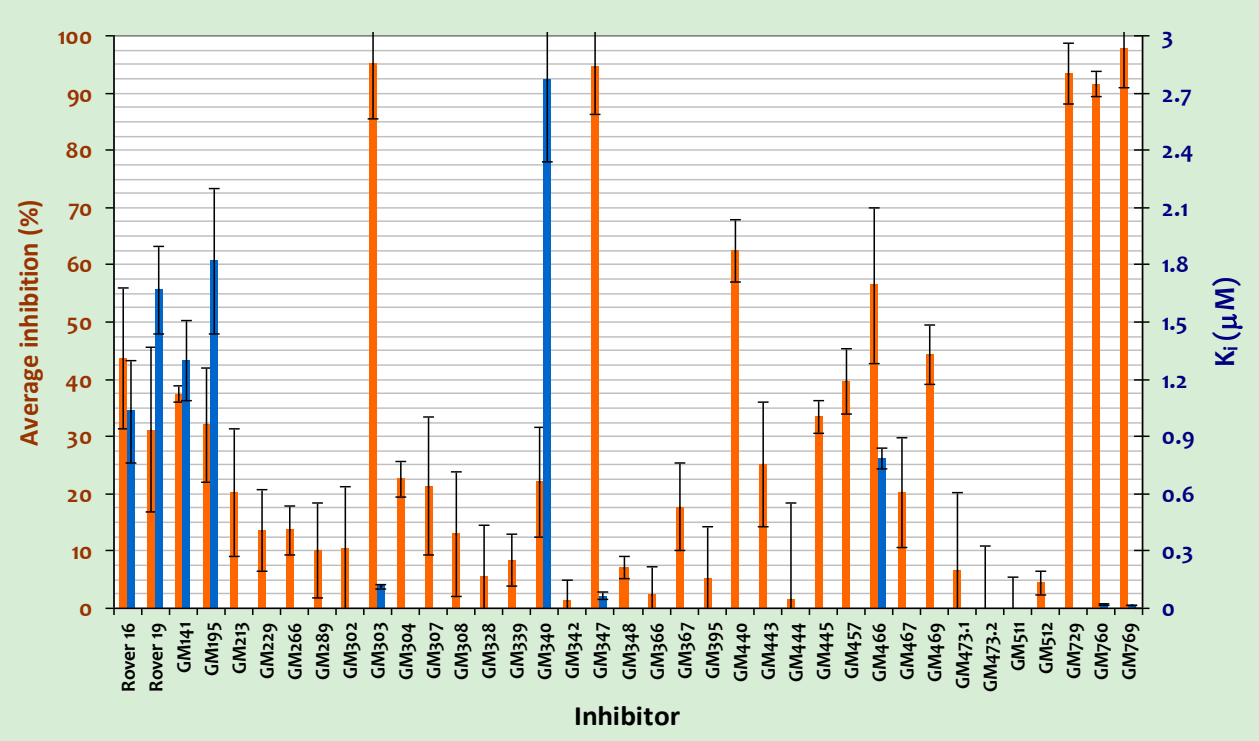


Figure 5.31. Graph summarising inhibition data for novel 1,2,4-oxadiazole-based series of *Pf*KMO inhibitors. Average plate-screen data with 10 μM of each inhibitor is displayed with orange bars (*left axis*) with K_i values determined plotted as blue bars (*right axis*). See sections **2.5** and **5.3.1** for full assay details. Created in Excel.

Following the successful crystallization and structure determination of dm2 *Pf*KMO (see Chapter **3**), co-crystallization with the more potent inhibitors was attempted. Eventually conditions were found from which crystals of inhibitor bound enzyme could be reproducibly grown leading to structures containing the four tightest binding novel inhibitors: GM303, GM347, GM760 and GM769. Unexpectedly a third new type of lattice packing was observed with three protein molecules in the asymmetric unit in the hexagonal space group $P6_122$ and with strong diffraction. The final resolution of the X-ray datasets was between 2.15 and 2.40 Å. From these structures the positions of the inhibitors (with the exception of GM347) could be accurately defined in the substrate-binding pocket. The tail of the compounds extended out of the pocket, disrupting the

interactions that previously linked the substrate with the C-terminal domain 3 via interaction with R84. As a result the inhibitor structures exhibited an open C-terminal conformation reminiscent of the unbound *Pf*KMO structure. This was a significant difference with the substrate analogues and may explain the effect of these compounds as true *Pf*KMO inhibitors rather than as non-substrate effectors.

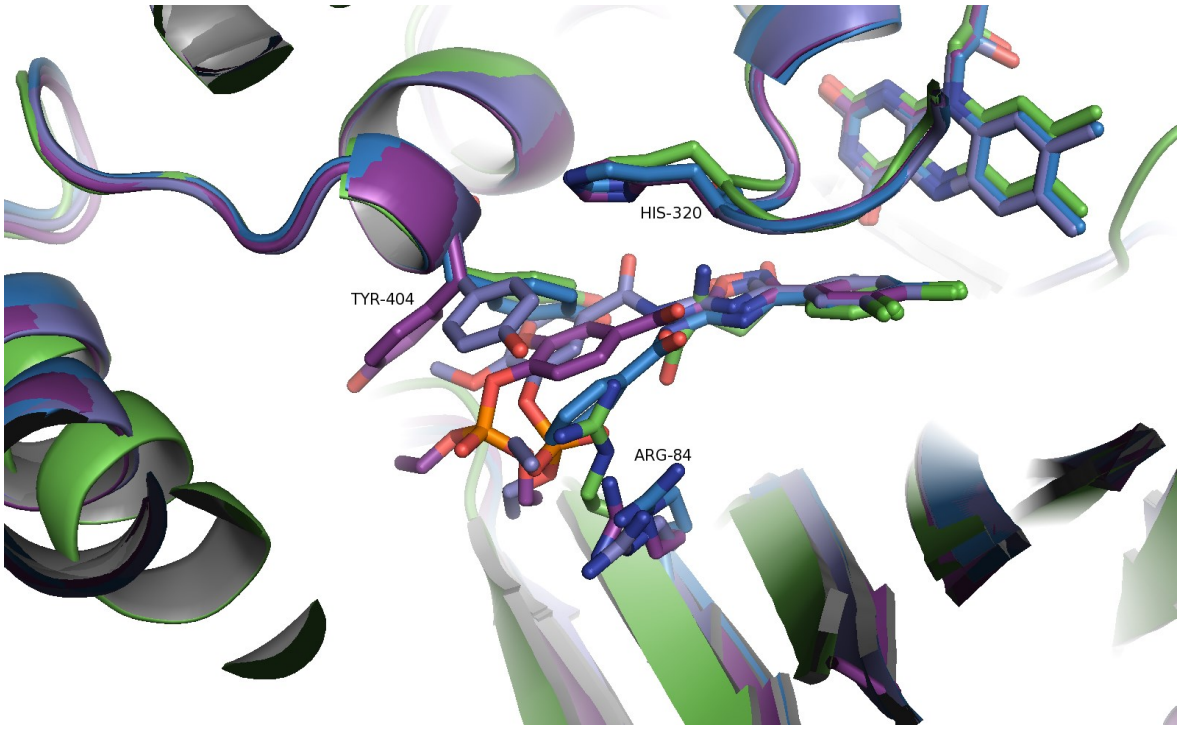


Figure 5.32. Structural overlay of the inhibitor binding sites for each novel compound with that of the substrate-bound structure (green). The inhibitor-bound structures are coloured blue, violet and purple for GM303, GM760 and GM769. Ligands and key protein residues are shown as sticks with the protein structure represented in cartoon. Figrue and alignmentcreated using PyMol.

A structural comparison of the inhibitor binding environments for GM303, GM760 and GM769 is shown in figure 5.32. In all three cases the side-chain of R84 is pushed aside accounting for the breakdown of interactions responsible for closing the C-terminal domain when substrate/substrate analogues bind (see Chapter **4**). A beautiful insight into the rotation of the Y404 side-chain to accommodate the three different inhibitor tails can be seen. The side-chain occupies three separate orientations to π–stack with each of the inhibitors. An additional variation among the compounds is the orientation

of the amide linker group. In the GM303 and GM769 bound structures the carbonyl points out towards N369 but in GM760 – with the $PO_2(OEt)_2$ group in the 3' position – it rotates to point into the hydrophobic pocket below H320. In all three cases the head 3,4-dichlorobenzene ring is virtually identical and adopts a position similar to that occupied by the substrate analogue 3,4-CBA. It therefore is likely that the previously mentioned difference in stimulation of flavin reduction between the inhibitors and substrate analogues is down to the tail of the molecule. This is possibly linked to the loss of the carboxyl group on 3,4-CBA and the substrate, and its replacement with a much larger tail arrangement.

In summary the aim of developing a novel series of potent *Pf*KMO inhibitors was achieved with great success. Moreover, a new set of compounds with further alternative tail R_3 groups have been synthesized by Gavin Milne ready for future testing and structure solution. The presented and ongoing new compounds may themselves present exciting therapeutic tools for the inhibition of human KMO. The only significant sequence difference in the predicted active site of the human enzyme is the substitution of H320 for a phenylalanine. As can be seen from the structures, this side chain does not influence inhibitor binding except with GM760. The amide carbonyl of GM760 is around 3.0-3.5 Å from a nitrogen atom in the imidazole ring of H320 but is unlikely to be involved in a hydrogen bond with it. This is because the carbonyl is already fully occupied by two <3.0 Å hydrogen bonds, and additionally the position of the imidazole nitrogen doesn't fit with either of these. Still, if the same inhibition properties do not transfer to clinical studies with the human enzyme, the presented inhibitor bound structures should still serve as a significantly useful model for the rational design and optimisation of inhibitors of human KMO.

Chapter 6 – Conclusions and Future Work

In this investigation, the first *Pf*KMO crystal structures were solved using an engineered cysteine double mutant C252S/C461S (dm2) with enhanced stability in solution. The mutations did not significantly affect the catalytic ability of the enzyme and the residues are located in surface areas of the structure distant from the substrate entry route. Therefore it is presumed that the structures presented and ideas discussed are transferable to the native enzyme. In all the structures of dm2 *Pf*KMO obtained include substrate bound, substrate-free (unbound), substrate-free with chloride ions (chloride), substrate analogue bound (*m*-NBA and 3,4-CBA), novel inhibitor bound (GM303, GM760 and GM769), Y382F substrate bound, R386K substrate bound and unbound, R386T substrate bound and substrate-free chloride, E372T/L367R substrate bound and chloride. In all cases the majority of the protein was modelled from residues 6-461 with a similar overall architecture which can be arbitrarily split into three structural domains. The first two domains (up to residue 376), comprising the FAD and substrate binding sites are structurally related to the fold seen with other FAH enzymes whose structures have been solved. The third, C-terminal domain is unrelated to anything else seen in the enzyme class to date and extends out from the first two domains from the back of the substrate binding pocket at the substrate's carboxyl tail. This domain occupies different conformations in the solved *Pf*KMO structures and this observation led to the investigation of its role/function. The second and main focus of the investigation as it developed was to quantify and characterise the nature of inhibitor binding in the substrate binding site.

6.1 Substrate binding and role of the C-terminal domain

The initial substrate-bound structure was solved to a resolution of 3.40 Å with SeMet-labelled protein. Again it was evident that the SeMet residues did not significantly alter the catalytic ability or the structure of the enzyme. The exact position of the substrate and the surrounding active site residues are slightly ambiguous given the resolution of the data. The substrate's orientation is strongly supported by the positioning of the carboxyl group within 3 Å of a guanidinium group from the highly conserved residue R84

and by the positioning of the phenyl ring of L-Kyn with its site of hydroxylation approximately 5 Å from the catalytic intermediate-forming C_{4a} atom of the flavin. This specific substrate-flavin distance correlates well with the reported values seen in structures of other FAH enzymes [158, 170-172, 174-176, 178, 259]. In addition the protein residues proposed to surround the bound substrate are all highly conserved among KMO homologs. Building on this structure the Mowat group will look to substitute these potential substrate interacting residues to fully characterise the nature of substrate binding.

The substrate analogue structures with 3,4-CBA and with *m*-NBA bound both further validated the substrate's orientation and position. Whilst the resolution of the structures was not improved, the added functional groups provided significant density to confirm the position of the ligand phenyl ring. Interestingly in 2008 Crozier-Reabe *et al.,* showed that BA and *m*-NBA acted as non-substrate effectors of *Pf*KMO as their binding stimulated uncoupled flavin reduction leading to hydrogen peroxide production [1]. The preliminary data in this study agreed with these findings but suggested that the analogue 3,4-CBA did not have this effect and was a true competitive inhibitor of *Pf*KMO. There were no structural differences evident between the binding of 3,4-CBA and *m*-NBA so further investigation is needed to confirm and explain the difference between the compounds. This could be achieved by stopped-flow work with 3,4-CBA and the use of fluorescence resonance energy transfer (FRET) along with possible substitution of the FAD cofactor for fluorescently-labelled analogues.

The solution of the structure of unbound *Pf*KMO to a much higher resolution of 2.26 Å provided a more defined and detailed structural model that additionally incorporated solvent molecules. The crystals exhibited different lattice packing in a different space group to substrate bound crystals. The resulting structure was virtually indistinguishable from the substrate bound form over the first two domains; however the novel C-terminal domain swung out from the rest of the enzyme by up to around 6 Å. This movement was shown to be dependent upon the binding of substrate rather than the nature of the crystal packing by the study of a series of C-terminal mutants – E372T, Y382F, R386T and R386K. In the substrate bound structure, with what was labelled as the closed C-terminal conformation, these residues appeared to be well positioned to

form a chain of non-covalent intermolecular interactions linking the C-terminus to the side-chain of a key substrate-interacting residue R84.

In the open conformation of the unbound *Pf*KMO structure the residues involved in this chain of interactions were clearly further apart from each other. A similar, partially open (mixed) C-terminal conformation was present in the substrate bound mutant structures. Whilst the resolution was unfortunately still poor in these structures at 3.24-3.40 Å, there was clear evidence for placing the C-terminus in a significantly more open conformation than with the initial substrate bound dm2 *Pf*KMO structure. The exact mechanism and catalytic relevance of the C-terminal movement needs further study. For example the mutants could be used to validate whether the rate of product release is faster when the C-terminal domain is more open using stopped flow spectrophotometry. In addition it would be interesting to see if the mutants experience raised levels of uncoupled flavin reduction generating hydrogen peroxide due to leakage of bound substrate from the NADPH-reduced enzyme-substrate complex. Finally the exact mechanism of substrate entry can be studied in greater detail to visualise the nature of the two-step substrate binding process observed by Crozier-Reabe *et al.,* [1]. Given the data presented here it is possible that the substrate could initially bind in the C-terminal domain – possibly involving residue R386 (as R386T possessed a four-fold raised K_m in comparison to the other mutants) – before being guided through into the buried substrate binding site. If this were true it could have implications for the rational design of a new series of inhibitors which target the C-terminal domain rather than the active site itself.

6.2 Inhibition and human KMO

A novel series of 1,2,4-oxadiazole amides designed and created by Gavin Milne (PhD, University of St Andrews) produced four nanomolar K_i competitive inhibitors of *Pf*KMO. Crystal structures with the compounds bound revealed that they filled the substrate binding pocket, prevented closure of the C-terminal domain and possibly disrupted the NADPH binding site. As previously mentioned the *Pf*KMO active site looks like a good model for that of the human enzyme when looking at the sequence conservation of the involved residues. The only difference in residues within 5 Å of the substrate is the

change of H320 from *Pf*KMO to phenylalanine in the human enzyme. To provide more direct evidence for this hypothesis, purified dm2 *Pf*KMO was recently sent to GlaxoSmithKline plc (GSK) to directly compare the inhibition pattern observed with a set of compounds to the pattern observed using overexpressed human KMO from isolated SF9 (a eukaryotic expression strain) cell membranes. The overall correlation between the pIC_{50} results with the two enzymes was good with an R^2 value of 0.83 (see figure 6.1). The results, which were based on the direct measurement of 3-HK production by RapidFire® mass spectrometry, suggest that the bacterial active site structural features and inhibition results presented in this work may be relevant in the search for inhibitors of the human enzyme.

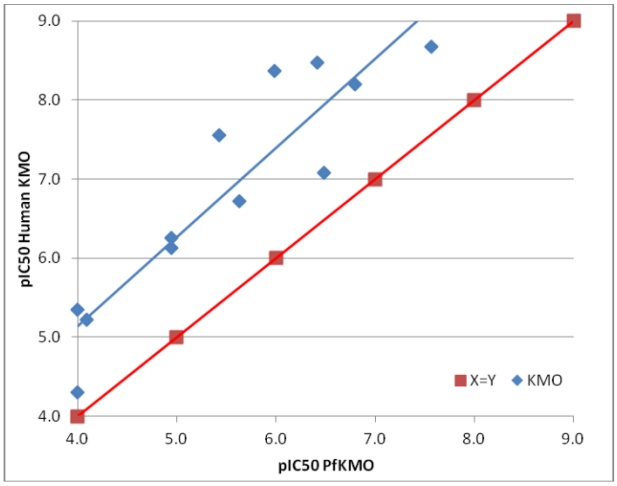


Figure 6.1. Correlation of pIC_{50} values between the *Pf*KMO and human KMO RapidFire® mass spectrometry assays. Data is expressed as mean values, the overall correlation between assays gives an R^2 value of 0.83.

With the mammalian enzyme still difficult to purify and study *in vitro*, the generation of H320F dm2 *Pf*KMO (currently ongoing in the group) may further help to establish *Pf*KMO as a useful surrogate target for the screening, characterisation and crystal structure solution of potential human KMO inhibitors. A more complicated approach may be to generate a fusion protein with the transmembrane-containing C-terminal domain of human KMO swapped out for those of the bacterial *Pf*KMO. Even with the structural information and homology models available this task may not be straight forward. Different fusions could be created at different points of the human sequence to ideally try and start the swap at a loop and not in the middle of a secondary structural element. Secondary mutations may be necessary to 'fix' patches of the fusion where *Pf*KMO residues that are different to the corresponding human residues conflict with the

properties of neighbouring residues. So it would take a lot of careful designing but may yield a soluble active protein that can be crystallised and co-crystallised with novel inhibitors in the human active site.

There are still more novel inhibitors to be tested and therefore more structures to solve. It would also be interesting to look into how these inhibitors prevent flavin reduction. The process of NADPH docking is still poorly understood so mutational studies regarding the mobile loop involving residues 43-52 may be a good start. It may also be worth trying to crystallise the substrate-free and the inhibitor bound enzyme with NADPH or $NADP^+$ or to soak these compounds into substrate-free, substrate bound and inhibitor bound crystals. The coenzyme's would likely have to be present at very high concentrations in order to fully occupy the transient docking site. In the case of the inhibitor bound crystals though it may be that an NADPH complex can be obtained more easily if it is the flavin dynamic movement presumed necessary for reaction with NADPH, rather than the NADPH docking site, that is perturbed. Again stopped-flow and fluorescence work with the inhibitors and *Pf*KMO may also be useful to understand the differences between their binding and that of the substrate and substrate analogues. For example with some sort of fluorescent labelling near the docking site or on the NADPH itself the relative enzyme affinity for NADPH could be compared when different ligands are bound.

In all there is a lot of work on the bacterial enzyme that has been started in this investigation and some selected areas may be more highly prioritised for further study. Due to the clinical interest the future of the enzyme lies with the mammalian forms and so it is conceivable that the discussed adaptions of *Pf*KMO to be more like human KMO may be the key and highest impact route for the continued study of the enzyme. For now though the set of structures presented here should have an immediate impact as they can be used to speed up human KMO structural modelling and subsequently the rational design of human KMO inhibitors for potential therapeutic use.

References

1. Crozier-Reabe, K.R., R.S. Phillips, and G.R. Moran, *Kynurenine 3-monooxygenase from Pseudomonas fluorescens: substrate-like inhibitors both stimulate flavin reduction and stabilize the flavin-peroxo intermediate yet result in the production of hydrogen peroxide.* Biochemistry, 2008. **47**(47): p. 12420-33.
2. Crozier, K.R. and G.R. Moran, *Heterologous expression and purification of kynurenine-3-monooxygenase from Pseudomonas fluorescens strain 17400.* Protein Expr Purif, 2007. **51**(2): p. 324-33.
3. Montersino, S., D. Tischler, G.T. Gassner, and W.J.H. van Berkel, *Catalytic and structural features of flavoprotein hydroxylases and epoxidases.* Adv. Synth. Catal., 2011. **353**: p. 2301-2319.
4. Rover, S., A.M. Cesura, P. Huguenin, R. Kettler, and A. Szente, *Synthesis and biochemical evaluation of N-(4-phenylthiazol-2-yl)benzenesulfonamides as high-affinity inhibitors of kynurenine 3-hydroxylase.* J Med Chem, 1997. **40**: p. 4378-4385.
5. Yasui, H., K. Takai, R. Yoshida, and O. Hayaishi, *Interferon enhances tryptophan metabolism by inducing pulmonary indoleamine 2,3-dioxygenase: its possible occurence in cancer patients.* Proc Natl Acad Sci U S A, 1986. **83**: p. 6622-6.
6. Werner-felmayer, G., E.R. Werner, D. Fuchs, A. Hausen, G. Reibnegger, and H. Wachter, *Characteristics of interferon induced tryptophan metabolism in human cells in vitro.* Biochim Biophys Acta, 1989. **1012**: p. 140-7.
7. Babcock, T.A. and J.M. Carlin, *Transcriptional activation of indoleamine dioxygenase by interleukin 1 and tumor necrosis factor alpha in interferon-treated epithelial cells.* Cytokine, 2000. **12**(6): p. 588-94.
8. Currier, A.R., M.H. Ziegler, M.M. Riley, T.A. Babcock, V.P. Telbis, and J.M. Carlin, *Tumor necrosis factor-alpha and lipopolysaccharide enhance interferon-induced antichlamydial indoleamine dioxygenase activity independently.* J Interferon Cytokine Res, 2000. **20**(4): p. 369-76.
9. Hamanaka, R.B. and N.S. Chandel, *Mitochondrial reactive oxygen species regulate cellular signaling and dictate biological outcomes.* Trends Biochem Sci, 2010. **35**(9): p. 505-13.
10. Zunszain, P.A., C. Anacker, A. Cattaneo, S. Choudhury, K. Musaelyan, A.M. Myint, S. Thuret, J. Price, and C.M. Pariante, *Interleukin-1beta: a new regulator of the kynurenine pathway affecting human hippocampal neurogenesis.* Neuropsychopharmacology, 2012. **37**(4): p. 939-49.
11. Heyes, M.P., C.Y. Chen, E.O. Major, and K. Saito, *Different kynurenine pathway enzymes limit quinolinic acid formation by various human cell types.* Biochem J, 1997. **326 (Pt 2)**: p. 351-6.
12. Demeter, I., K. Nagy, L. Gellert, L. Vecsei, F. Fulop, and J. Toldi, *A novel kynurenic acid analog (SZR104) inhibits pentylenetetrazole-induced epileptiform seizures. An electrophysiological study : special issue related to kynurenine.* J Neural Transm, 2012. **119**(2): p. 151-4.
13. Carlin, J.M., E.C. Borden, P.M. Sondel, and G.I. Byrne, *Interferon-induced indoleamine 2,3-dioxygenase activity in human mononuclear phagocytes.* J Leukoc Biol, 1989. **45**(1): p. 29-34.
14. Couzin-Frankel, J., *Inflammation bares a dark side.* Science, 2010. **330**(6011): p. 1621.
15. Block, M.L. and J.S. Hong, *Microglia and inflammation-mediated neurodegeneration: multiple triggers with a common mechanism.* Prog Neurobiol, 2005. **76**(2): p. 77-98.

16. Bruunsgaard, H., K. Anderson-Ranberg, B. Jeune, A.N. Pedersen, P. Skinhoj, and B.K. Pedersen, *A high plasma concentration of TNF-alpha is associated with dementia in centenarians.* J Gerontol A Biol Sci Med Sci, 1999. **54**: p. M357-364.
17. Bruunsgaard, H., *The clinical impact of systemic low-level inflammation in elderly populations. With special reference to cardiovascular disease, dementia and mortality.* Dan Med Bull, 2006. **53**: p. 285-309.
18. Vehmas, A.K., C.H. Kawas, W.F. Stewart, and J.C. Trocoso, *Immune reactive cells in senile plaques and cognitive decline in Alzheimer's disease.* Neurobiol Aging, 2003. **24**: p. 321-331.
19. Guerreiro, R.J., I. Santana, J.M. Bras, B. Santiago, A. Paiva, and C. Oliveira, *Peripheral inflammatory cytokines as biomarkers in Alzheimer's disease.* Neurodegener Dis, 2007. **4**: p. 406-412.
20. Chen, Y., R. Stankovic, K.M. Cullen, V. Meininger, B. Garner, S. Coggan, R. Grant, B.J. Brew, and G.J. Guillemin, *The kynurenine pathway and inflammation in amyotrophic lateral sclerosis.* Neurotox Res, 2010. **18**(2): p. 132-42.
21. Harroon, E., C.L. Raison, and A.H. Miller, *Psychoneuroimmunology meets neuropsychopharmacology: translational implications of the impact of inflammation on behaviour.* Neuropyschopharmacology Reviews, 2012. **37**(137-162).
22. Fuchs, D., A. Forsmann, L. Hagberg, and e. al., *Immune activation and decreased tryptophan in patients with HIV-1 infection.* J Interferon Res, 1990. **10**(6): p. 599-603.
23. Fuchs, D., A.A. Moller, G. Reibnegger, E. Stockle, E.R. Werner, and H. Wachter, *Decreased serum tryptophan in patients with HIV-1 infection correlates with increased serum neopterin and with nearologic/psychiatric symptoms.* J Acquir Immune Defic Syndr, 1990. **39**(9): p. 873-876.
24. Turski, M.P., P. Kaminski, W. Zgrajka, M. Turska, and W.A. Turski, *Potato- an important source of nutritional kynurenic acid.* Plant Foods Hum Nutr, 2012. **67**(1): p. 17-23.
25. Holtze, M., A. Mickiene, A. Atlas, L. Lindquist, and L. Schwieler, *Elevated cerebrospinal fluid kynurenic acid levels in patients with tick-borne encephalitis.* J Intern Med, 2012. **272**(4): p. 394-401.
26. Pawlak, K., T. Domaniewski, M. Mysliwiec, and D. Pawlak, *The kynurenines are associated with oxidative stress, inflammation and the prevalence of cardiovascular disease in patients with end-stage renal disease.* Atherosclerosis, 2009. **204**(1): p. 309-14.
27. Sperner-Unterweger, B., G. Neurauter, M. Klieber, K. Kurz, V. Meraner, A. Zeimet, and D. Fuchs, *Enhanced tryptophan degradation in patients with ovarian carcinoma correlates with several serum soluble immune activation markers.* Immunobiology, 2011. **216**(3): p. 296-301.
28. Werner, E.R., D. Fuchs, and A. Hausen, *Tryptophan degradation in patients infected by human immunodeficiency virus.* Biol Chem Hoppe Seyler, 1988. **369**: p. 337-340.
29. Fuchs, D., A.A. Moller, G. Reibnegger, E.R. Werner, G. Werner-felmayer, M.P. Dierich, and H. Wachter, *Increased endogenous interferon-gamma and neopterin correlate with increased degradation of tryptophan in human immunodeficiency virus type 1 infection.* Immunol Lett, 1991. **28**: p. 207-11.
30. Zangerle, R., B. Widner, G. Quirchmair, G. Neurauter, M. Sarcletti, and D. Fuchs, *Effective antiretroviral therapy reduces degradation of tryptophan in patients with HIV-1 infection.* Clin Immunol, 2002. **104**: p. 242-7.
31. Bonda, D.J., M. Mailankot, J.G. Stone, M.R. Garrett, M. Staniszewska, R.J. Castellani, S.L. Siedlak, X. Zhu, H.G. Lee, G. Perry, R.H. Nagaraj, and M.A. Smith, *Indoleamine*

2,3-dioxygenase and 3-hydroxykynurenine modifications are found in the neuropathology of Alzheimer's disease. Redox Rep, 2010. **15**(4): p. 161-8.
32. Schefold, J.C., J.P. Zeden, C. Fotopoulou, S. von Haehling, R. Pschowski, D. Hasper, H.D. Volk, C. Schuett, and P. Reinke, *Increased indoleamine 2,3-dioxygenase (IDO) activity and elevated serum levels of tryptophan catabolites in patients with chronic kidney disease: a possible link between chronic inflammation and uraemic symptoms.* Nephrol Dial Transplant, 2009. **24**(6): p. 1901-8.
33. Asp, L., A.S. Johansson, A. Mann, B. Owe-Larsson, E.M. Urbanska, T. Kocki, M. Kegel, G. Engberg, G.B.S. Lundkvist, and H. Karlsson, *Effects of pro-inflammatory cytokines on expression of kynurenine pathway enzymes in human dermal fibroblasts.* J Inflammation, 2011. **8**: p. 25-31.
34. Guillemin, G.J., S.J. Kerr, G.A. Smythe, D.G. Smith, V. Kapoor, P.J. Armati, J. Croitoru, and B.J. Brew, *Kynurenine pathway metabolism in human astrocytes: a paradox for neuronal protection.* J Neurochem, 2001. **78**(4): p. 842-53.
35. Guillemin, G.J., K.M. Cullen, C.K. Lim, G.A. Smythe, B. Garner, V. Kapoor, O. Takikawa, and B.J. Brew, *Characterization of the kynurenine pathway in human neurons.* J Neurosci, 2007. **27**(47): p. 12884-92.
36. Allegri, G., A. Bertazzo, M. Biasiolo, C.V. Costa, and E. Ragazzi, *Kynurenine pathway enzymes in different species of animals.* Adv Exp Med Biol, 2003. **527**: p. 455-63.
37. Allegri, G., C.V. Costa, A. Bertazzo, M. Biasiolo, and E. Ragazzi, *Enzyme activities of tryptophan metabolism along the kynurenine pathway in various species of animals.* Farmaco, 2003. **58**(9): p. 829-36.
38. Fujigaki, S., K. Saito, M. Takemura, H. Fujii, H. Wada, A. Noma, and M. Seishima, *Species differences in L-tryptophan-kynurenine pathway metabolism: quantification of anthranilic acid and its related enzymes.* Arch Biochem Biophys, 1998. **358**(2): p. 329-35.
39. Zarnowski, T., T. Choragiewicz, M. Tulidowicz-Bielak, S. Thaler, R. Rejdak, I. Zarnowski, W.A. Turski, and M. Gasior, *Ketogenic diet increases concentrations of kynurenic acid in discrete brain structures of young and adult rats.* J Neural Transm, 2012. **119**(6): p. 679-84.
40. Stone, T.W. and J.H. Connick, *Quinolinic acid and other kynurenines in the central nervous system.* Neuroscience, 1985. **15**(3): p. 597-617.
41. Stone, T.W. and M.N. Perkins, *Quinolinic acid: a potent endogenous excitant at amino acid receptors in CNS.* Eur J Pharmacol, 1981. **72**(4): p. 411-2.
42. Alexander, K.S., H.Q. Wu, R. Schwarcz, and J.P. Bruno, *Acute elevations of brain kynurenic acid impair cognitive flexibility: normalization by the alpha7 positive modulator galantamine.* Psychopharmacology (Berl), 2012. **220**(3): p. 627-37.
43. Fulop, F., I. Szatmari, J. Toldi, and L. Vecsei, *Modifications on the carboxylic function of kynurenic acid.* J Neural Transm, 2012. **119**(2): p. 109-14.
44. Pawlak, K., M. Mysliwiec, and D. Pawlak, *Hyperhomocysteinemia and the presence of cardiovascular disease are associated with kynurenic acid levels and carotid atherosclerosis in patients undergoing continuous ambulatory peritoneal dialysis.* Thromb Res, 2012. **129**(6): p. 704-9.
45. Stone, T.W., *Endogenous neurotoxins from tryptophan.* Toxicon, 2001. **39**(1): p. 61-73.
46. Pellicciari, R., L. Amori, G. Costantino, A. Giordani, A. Macchiarulo, L. Mattoli, P. Pevarello, C. Speciale, and M. Varasi, *Modulation of the kynurine pathway of tryptophan metabolism in search for neuroprotective agents. Focus on kynurenine-3-hydroxylase.* Adv Exp Med Biol, 2003. **527**: p. 621-8.

47. Chen, Y. and G.J. Guillemin, *Kynurenine pathway metabolites in humans: disease and healthy States.* Int J Tryptophan Res, 2009. **2**: p. 1-19.
48. Vamos, E., A. Pardutz, P. Klivenyi, J. Toldi, and L. Vecsei, *The role of kynurenines in disorders of the central nervous system: possibilities for neuroprotection.* J Neurol Sci, 2009. **283**(1-2): p. 21-7.
49. Stone, T.W., C.M. Forrest, and L.G. Darlington, *Kynurenine pathway inhibition as a therapeutic strategy for neuroprotection.* FEBS J. FIELD Full Journal Title:FEBS Journal, 2012. **279**(8): p. 1386-1397.
50. Hilmas, C., E.F. Pereira, M. Alkondon, A. Rassoulpour, R. Schwarcz, and E.X. Albuquerque, *The brain metabolite kynurenic acid inhibits alpha7 nicotinic receptor activity and increases non-alpha7 nicotinic receptor expression: physiopathological implications.* J Neurosci, 2001. **21**(19): p. 7463-73.
51. Parsons, C.G., W. Danysz, G. Quack, S. Hartmann, B. Lorenz, C. Wollenburg, L. Baran, E. Przegalinski, W. Kostowski, P. Krzascik, B. Chizh, and P.M. Headley, *Novel systemically active antagonists of the glycine site of the N-methyl-D-aspartate receptor: electrophysiological, biochemical and behavioral characterization.* J Pharmacol Exp Ther, 1997. **283**(3): p. 1264-75.
52. Scharfman, H.E., J.H. Goodman, and R. Schwarcz, *Electrophysiological effects of exogenous and endogenous kynurenic acid in the rat brain: studies in vivo and in vitro.* Amino Acids, 2000. **19**(1): p. 283-97.
53. Moroni, F., P. Russi, G. Lombardi, M. Beni, and V. Carla, *Presence of kynurenic acid in the mammalian brain.* J Neurochem, 1988. **51**(1): p. 177-80.
54. Carpenedo, R., A. Chiarugi, P. Russi, G. Lombardi, V. Carla, R. Pellicciari, L. Mattoli, and F. Moroni, *Inhibitors of kynurenine hydroxylase and kynureninase increase cerebral formation of kynurenate and have sedative and anticonvulsant activities.* Neuroscience, 1994. **61**(2): p. 237-244.
55. Erhardt, S., M. Hajos, A. Lindberg, and G. Engberg, *Nicotine-induced excitation of locus coeruleus neurons is blocked by elevated levels of endogenous kynurenic acid.* Synapse, 2000. **37**(2): p. 104-8.
56. Moroni, F., P. Russi, M.A. Gallo-Mezo, G. Moneti, and R. Pellicciari, *Modulation of quinolinic and kynurenic acid content in the rat brain: effects of endotoxins and nicotinylalanine.* J Neurochem, 1991. **57**(5): p. 1630-5.
57. Russi, P., M. Alesiani, G. Lombardi, P. Davolio, R. Pellicciari, and F. Moroni, *Nicotinylalanine increases the formation of kynurenic acid in the brain and antagonizes convulsions.* J Neurochem, 1992. **59**(6): p. 2076-80.
58. Vecsei, L. and M.F. Beal, *Comparative behavioral and pharmacological studies with centrally administered kynurenine and kynurenic acid in rats.* Eur J Pharmacol, 1991. **196**(3): p. 239-46.
59. Vecsei, L., J. Miller, U. MacGarvey, and M.F. Beal, *Kynurenine and probenecid inhibit pentylenetetrazol- and NMDLA-induced seizures and increase kynurenic acid concentrations in the brain.* Brain Res Bull, 1992. **28**(2): p. 233-8.
60. Carpenedo, R., E. Meli, F. Peruginelli, D.E. Pellegrini-Giampietro, and F. Moroni, *Kynurenine 3-mono-oxygenase inhibitors attenuate post-ischemic neuronal death in organotypic hippocampal slice cultures.* J Neurochem, 2002. **82**(6): p. 1465-71.
61. Cozzi, A., R. Carpenedo, and F. Moroni, *Kynurenine hydroxylase inhibitors reduce ischemic brain damage: studies with (m-nitrobenzoyl)-alanine (mNBA) and 3,4-dimethoxy-[-N-4-(nitrophenyl)thiazol-2yl]-benzenesulfonamide (Ro 61-8048) in models of focal or global brain ischemia.* J Cereb Blood Flow Metab, 1999. **19**(7): p. 771-7.

62. Carpenedo, R., A. Pittaluga, A. Cozzi, S. Attucci, A. Galli, M. Raiteri, and F. Moroni, *Presynaptic kynurenate-sensitive receptors inhibit glutamate release.* Eur J Neurosci, 2001. **13**(11): p. 2141-7.
63. Wang, J., N. Simonavicius, X. Wu, G. Swaminath, J. Reagan, H. Tian, and L. Ling, *Kynurenic acid as a ligand for orphan G protein-coupled receptor GPR35.* J Biol Chem, 2006. **281**(31): p. 22021-8.
64. Moroni, F., A. Cozzi, M. Sili, and G. Mannaioni, *Kynurenic acid: a metabolite with multiple actions and multiple targets in brain and periphery.* J Neural Transm, 2012. **119**(2): p. 133-9.
65. Myint, A.M., Y.K. Kim, R. Verkerk, S. Scharpe, H. Steinbusch, and B. Leonard, *Kynurenine pathway in major depression: evidence of impaired neuroprotection.* J Affect Disord, 2007. **98**(1-2): p. 143-51.
66. Hartai, Z., A. Juhasz, A. Rimanoczy, T. Janaky, T. Donko, L. Dux, B. Penke, G.K. Toth, Z. Janka, and J. Kalman, *Decreased serum and red blood cell kynurenic acid levels in Alzheimer's disease.* Neurochem Int, 2007. **50**(2): p. 308-13.
67. Walczak, K., M. Zurawska, J. Kis, R. Starownik, W. Zgrajka, K. Bar, W.A. Turski, and W. Rzeski, *Kynurenic acid in human renal cell carcinoma: its antiproliferative and antimigrative action on Caki-2 cells.* Amino Acids, 2012. **43**(4): p. 1663-70.
68. Carlborg, A., J. Jokinen, E.G. Jonsson, S. Erhardt, and P. Nordstrom, *CSF kynurenic acid and suicide risk in schizophrenia spectrum psychosis.* Psychiatry Res, 2013. **205**(1-2): p. 165-7.
69. Walczak, K., W.A. Turski, and W. Rzeski, *Kynurenic acid enhances expression of p21 Waf1/Cip1 in colon cancer HT-29 cells.* Pharmacol Rep, 2012. **64**(3): p. 745-50.
70. Erhardt, S., K. Blennow, C. Nordin, E. Skogh, L.H. Lindstrom, and G. Engberg, *Kynurenic acid levels are elevated in the cerebrospinal fluid of patients with schizophrenia.* Neurosci Lett, 2001. **313**(1-2): p. 96-8.
71. Erhardt, S., L. Schwieler, L. Nilsson, K. Linderholm, and G. Engberg, *The kynurenic acid hypothesis of schizophrenia.* Physiol Behav, 2007. **92**(1-2): p. 203-9.
72. Sagan, D., T. Kocki, J. Kocki, and J. Szumilo, *Serum kynurenic acid: possible association with invasiveness of non-small cell lung cancer.* Asian Pac J Cancer Prev, 2012. **13**(9): p. 4241-4.
73. Heyes, M.P., K. Saito, O. Devinsky, and N.S. Nadi, *Kynurenine pathway metabolites in cerebrospinal fluid and serum in complex partial seizures.* Epilepsia, 1994. **35**(2): p. 251-7.
74. Zarnowski, T., R. Rejdak, E. Zielinska-Rzecka, E. Zrenner, P. Grieb, Z. Zagorski, A. Junemann, and W.A. Turski, *Elevated concentrations of kynurenic acid, a tryptophan derivative, in dense nuclear cataracts.* Curr Eye Res, 2007. **32**(1): p. 27-32.
75. Szalardy, L., D. Zadori, J. Toldi, F. Fulop, P. Klivenyi, and L. Vecsei, *Manipulating kynurenic Acid levels in the brain - on the edge between neuroprotection and cognitive dysfunction.* Curr Top Med Chem, 2012. **12**(16): p. 1797-806.
76. Schwarcz, R., W.O. Whetsell, Jr., and R.M. Mangano, *Quinolinic acid: an endogenous metabolite that produces axon-sparing lesions in rat brain.* Science, 1983. **219**(4582): p. 316-8.
77. Kim, J.P. and D.W. Choi, *Quinolinate neurotoxicity in cortical cell culture.* Neuroscience, 1987. **23**(2): p. 423-32.
78. Heyes, M.P., B.J. Brew, A. Martin, R.W. Price, A.M. Salazar, J.J. Sidtis, J.A. Yergey, M.M. Mouradian, A.E. Sadler, J. Keilp, and et al., *Quinolinic acid in cerebrospinal fluid and serum in HIV-1 infection: relationship to clinical and neurological status.* Ann Neurol, 1991. **29**(2): p. 202-9.

79. Heyes, M.P., M. Gravell, M. April, D. Blackmore, W.T. London, J.A. Yergey, and S.P. Markey, *Quinolinic acid concentrations are increased in cerebrospinal fluid of Rhesus macaques (Macaca mulatta) naturally infected with simian retrovirus type-D.* Adv Exp Med Biol, 1991. **294**: p. 555-8.
80. Heyes, M.P., D. Rubinow, C. Lane, and S.P. Markey, *Cerebrospinal fluid quinolinic acid concentrations are increased in acquired immune deficiency syndrome.* Ann Neurol, 1989. **26**(2): p. 275-7.
81. Heyes, M.P., K. Saito, D. Jacobowitz, S.P. Markey, O. Takikawa, and J.H. Vickers, *Poliovirus induces indoleamine-2,3-dioxygenase and quinolinic acid synthesis in macaque brain.* Faseb J, 1992. **6**(11): p. 2977-89.
82. Sanni, L.A., S.R. Thomas, B.N. Tattam, D.E. Moore, G. Chaudhri, R. Stocker, and N.H. Hunt, *Dramatic changes in oxidative tryptophan metabolism along the kynurenine pathway in experimental cerebral and noncerebral malaria.* Am J Pathol, 1998. **152**(2): p. 611-9.
83. Chiarugi, A., A. Cozzi, C. Ballerini, L. Massacesi, and F. Moroni, *Kynurenine 3-mono-oxygenase activity and neurotoxic kynurenine metabolites increase in the spinal cord of rats with experimental allergic encephalomyelitis.* Neuroscience, 2001. **102**(3): p. 687-95.
84. Myint, A.M., M.J. Schwarz, R. Verkerk, H.H. Mueller, J. Zach, S. Scharpe, H.W. Steinbusch, B.E. Leonard, and Y.K. Kim, *Reversal of imbalance between kynurenic acid and 3-hydroxykynurenine by antipsychotics in medication-naive and medication-free schizophrenic patients.* Brain Behav Immun, 2011. **25**(8): p. 1576-81.
85. Vezzani, A., R. Serafini, M.A. Stasi, S. Caccia, I. Conti, R.V. Tridico, and R. Samanin, *Kinetics of MK-801 and its effect on quinolinic acid-induced seizures and neurotoxicity in rats.* J Pharmacol Exp Ther, 1989. **249**(1): p. 278-83.
86. Vezzani, A., M.A. Stasi, H.Q. Wu, M. Castiglioni, B. Weckermann, and R. Samanin, *Studies on the potential neurotoxic and convulsant effects of increased blood levels of quinolinic acid in rats with altered blood-brain barrier permeability.* Exp Neurol, 1989. **106**(1): p. 90-8.
87. Beal, M.F., R.J. Ferrante, K.J. Swartz, and N.W. Kowall, *Chronic quinolinic acid lesions in rats closely resemble Huntington's disease.* J Neurosci, 1991. **11**(6): p. 1649-59.
88. Giorgini, F., P. Guidetti, Q. Nguyen, S.C. Bennett, and P.J. Muchowski, *A genomic screen in yeast implicates kynurenine 3-monooxygenase as a therapeutic target for Huntington disease.* Nat Genet, 2005. **37**(5): p. 526-31.
89. Guillemin, G.J., B.J. Brew, C.E. Noonan, O. Takikawa, and K.M. Cullen, *Indoleamine 2,3 dioxygenase and quinolinic acid immunoreactivity in Alzheimer's disease hippocampus.* Neuropathol Appl Neurobiol, 2005. **31**(4): p. 395-404.
90. Guillemin, G.J., K.R. Williams, D.G. Smith, G.A. Smythe, J. Croitoru-Lamoury, and B.J. Brew, *Quinolinic acid in the pathogenesis of Alzheimer's disease.* Adv Exp Med Biol, 2003. **527**: p. 167-76.
91. Walczak, K., W. Dabrowski, E. Langner, W. Zgrajka, J. Pilat, T. Kocki, W. Rzeski, and W.A. Turski, *Kynurenic acid synthesis and kynurenine aminotransferases expression in colon derived normal and cancer cells.* Scand J Gastroenterol, 2011. **46**(7-8): p. 903-12.
92. Dolecka, J., T. Urbanik-Sypniewska, B. Skrzydlo-Radomanska, and J. Parada-Turska, *Effect of kynurenic acid on the viability of probiotics in vitro.* Pharmacol Rep, 2011. **63**(2): p. 548-51.
93. Braidy, N., R. Grant, S. Adams, B.J. Brew, and G.J. Guillemin, *Mechanism for quinolinic acid cytotoxicity in human astrocytes and neurons.* Neurotox Res, 2009. **16**(1): p. 77-86.

94. Guillemin, G.J., *Quinolinic acid, the inescapable neurotoxin.* FEBS Journal, 2012. **279**: p. 1356-1365.
95. Rahman, A., K. Ting, K.M. Cullen, N. Braidy, B.J. Brew, and G.J. Guillemin, *The excitotoxin quinolinic acid induces tau phosphorylation in human neurons.* PLoS One, 2009. **4**(7): p. e6344.
96. Perez-De La Cruz, V., P. Carrillo-Mora, and A. Santamaria, *Quinolinic acid, an endogenous molecule combining excitotoxicity, oxidative stress and other toxic mechanisms.* Int. J. Tryptophan Res. FIELD Full Journal Title:International Journal of Tryptophan Research, 2012. **5**: p. 1-8.
97. Platenik, J., P. Stopka, M. Vejrazka, and S. Stipek, *Quinolinic acid-iron(ii) complexes: slow autoxidation, but enhanced hydroxyl radical production in the Fenton reaction.* Free Radic Res, 2001. **34**(5): p. 445-59.
98. Santamaria, A., M.E. Jimenez-Capdeville, A. Camacho, E. Rodriguez-Martinez, A. Flores, and S. Galvan-Arzate, *In vivo hydroxyl radical formation after quinolinic acid infusion into rat corpus striatum.* Neuroreport, 2001. **12**(12): p. 2693-6.
99. Santamaria, A., R. Salvatierra-Sanchez, B. Vazquez-Roman, D. Santiago-Lopez, J. Villeda-Hernandez, S. Galvan-Arzate, M.E. Jimenez-Capdeville, and S.F. Ali, *Protective effects of the antioxidant selenium on quinolinic acid-induced neurotoxicity in rats: in vitro and in vivo studies.* J Neurochem, 2003. **86**(2): p. 479-88.
100. Behan, W.M., M. McDonald, L.G. Darlington, and T.W. Stone, *Oxidative stress as a mechanism for quinolinic acid-induced hippocampal damage: protection by melatonin and deprenyl.* Br J Pharmacol, 1999. **128**(8): p. 1754-60.
101. Kalonia, H., P. Kumar, and A. Kumar, *Comparative neuroprotective profile of statins in quinolinic acid induced neurotoxicity in rats.* Behav Brain Res, 2011. **216**(1): p. 220-8.
102. Maharaj, D.S., H. Maharaj, E.M. Antunes, D.M. Maree, T. Nyokong, B.D. Glass, and S. Daya, *6-Hydroxymelatonin protects against quinolinic-acid-induced oxidative neurotoxicity in the rat hippocampus.* J Pharm Pharmacol, 2005. **57**(7): p. 877-81.
103. Mallozzi, C., A. Martire, M.R. Domenici, A. Metere, P. Popoli, and A.M. Di Stasi, *L-NAME reverses quinolinic acid-induced toxicity in rat corticostriatal slices: Involvement of src family kinases.* J Neurosci Res, 2007. **85**(12): p. 2770-7.
104. Silva-Adaya, D., V. Perez-De La Cruz, M.N. Herrera-Mundo, K. Mendoza-Macedo, J. Villeda-Hernandez, Z. Binienda, S.F. Ali, and A. Santamaria, *Excitotoxic damage, disrupted energy metabolism, and oxidative stress in the rat brain: antioxidant and neuroprotective effects of L-carnitine.* J Neurochem, 2008. **105**(3): p. 677-89.
105. Vega-Naredo, I., B. Poeggeler, V. Sierra-Sanchez, B. Caballero, C. Tomas-Zapico, O. Alvarez-Garcia, D. Tolivia, M.J. Rodriguez-Colunga, and A. Coto-Montes, *Melatonin neutralizes neurotoxicity induced by quinolinic acid in brain tissue culture.* J Pineal Res, 2005. **39**(3): p. 266-75.
106. Velloso, N.A., G.D. Dalmolin, G. Fonini, V.D. Gindri Sinhorin, A. Ferreira da Silveira, M.A. Rubin, and C.F. Mello, *Spermine attenuates behavioral and biochemical alterations induced by quinolinic acid in the striatum of rats.* Brain Res, 2008. **1198**: p. 107-14.
107. Eastman, C.L. and T.R. Guilarte, *Cytotoxicity of 3-hydroxykynurenine in a neuronal hybrid cell line.* Brain Res, 1989. **495**(2): p. 225-31.
108. Eastman, C.L. and T.R. Guilarte, *The role of hydrogen peroxide in the in vitro cytotoxicity of 3-hydroxykynurenine.* Neurochem Res, 1990. **15**(11): p. 1101-7.
109. Nakagami, Y., H. Saito, and H. Katsuki, *3-Hydroxykynurenine toxicity on the rat striatum in vivo.* Jpn J Pharmacol, 1996. **71**(2): p. 183-6.
110. Guidetti, P. and R. Schwarcz, *3-Hydroxykynurenine potentiates quinolinate but not NMDA toxicity in the rat striatum.* Eur J Neurosci, 1999. **11**(11): p. 3857-63.

111. Vazquez, S., B. Garner, M.M. Sheil, and R.J. Truscott, *Characterisation of the major autoxidation products of 3-hydroxykynurenine under physiological conditions.* Free Radic Res, 2000. **32**(1): p. 11-23.
112. Chiarugi, A., E. Meli, and F. Moroni, *Similarities and differences in the neuronal death processes activated by 3OH-kynurenine and quinolinic acid.* J Neurochem, 2001. **77**(5): p. 1310-8.
113. Goshima, N., A. Wadano, and K. Miura, *3-Hydroxykynurenine as O2-. scavenger in the blowfly, Aldrichina grahami.* Biochem Biophys Res Commun, 1986. **139**(2): p. 666-72.
114. Luthra, M. and D. Balasubramanian, *3-Hydroxykynurenine and 3-hydroxyanthranilic acid may act as endogenous antioxidants in the eye lens.* Exp Eye Res, 1992. **55**(4): p. 641-3.
115. Truscott, R.J., J.A. Carver, A. Thorpe, and R.H. Douglas, *Identification of 3-hydroxykynurenine as the lens pigment in the gourami Trichogaster trichopterus.* Exp Eye Res, 1992. **54**(6): p. 1015-7.
116. Wood, A.M. and R.J. Truscott, *Ultraviolet filter compounds in human lenses: 3-hydroxykynurenine glucoside formation.* Vision Res, 1994. **34**(11): p. 1369-74.
117. Leipnitz, G., C. Schumacher, K.B. Dalcin, K. Scussiato, A. Solano, C. Funchal, C.S. Dutra-Filho, A.T. Wyse, C.M. Wannmacher, A. Latini, and M. Wajner, *In vitro evidence for an antioxidant role of 3-hydroxykynurenine and 3-hydroxyanthranilic acid in the brain.* Neurochem Int, 2007. **50**(1): p. 83-94.
118. Okuda, S., N. Nishiyama, H. Saito, and H. Katsuki, *Hydrogen peroxide-mediated neuronal cell death induced by an endogenous neurotoxin, 3-hydroxykynurenine.* Proc Natl Acad Sci U S A, 1996. **93**(22): p. 12553-8.
119. Takeuchi, F. and Y. Shibata, *Kynurenine metabolism in vitamin-B-6-deficient rat liver after tryptophan injection.* Biochem J, 1984. **220**(3): p. 693-9.
120. Guilarte, T.R. and H.N. Wagner, Jr., *Increased concentrations of 3-hydroxykynurenine in vitamin B6 deficient neonatal rat brain.* J Neurochem, 1987. **49**(6): p. 1918-26.
121. Topczewska-Bruns, J., D. Pawlak, E. Chabielska, A. Tankiewicz, and W. Buczko, *Increased levels of 3-hydroxykynurenine in different brain regions of rats with chronic renal insufficiency.* Brain Res Bull, 2002. **58**(4): p. 423-8.
122. Heyes, M.P., B.J. Quearry, and S.P. Markey, *Systemic endotoxin increases L-tryptophan, 5-hydroxyindoleacetic acid, 3-hydroxykynurenine and quinolinic acid content of mouse cerebral cortex.* Brain Res, 1989. **491**(1): p. 173-9.
123. Schwarz, M.J., G.J. Guillemin, S.J. Teipel, K. Buerger, and H. Hampel, *Increased 3-Hydroxykynurenine serum concentrations differentiate Alzheimer's disease patients from controls.* Eur Arch Psychiatry Clin Neurosci, 2012.
124. Guidetti, P., G.P. Bates, R.K. Graham, M.R. Hayden, B.R. Leavitt, M.E. MacDonald, E.J. Slow, V.C. Wheeler, B. Woodman, and R. Schwarcz, *Elevated brain 3-hydroxykynurenine and quinolinate levels in Huntington disease mice.* Neurobiol Dis, 2006. **23**(1): p. 190-7.
125. Guilarte, T.R. and C.L. Eastman, *Is 3-hydroxykynurenine an endogenous neurotoxin in Huntington's disease?* J Neurol Sci, 1993. **116**(2): p. 227-8.
126. Pearson, S.J. and G.P. Reynolds, *Increased brain concentrations of a neurotoxin, 3-hydroxykynurenine, in Huntington's disease.* Neurosci Lett, 1992. **144**(1-2): p. 199-201.
127. Reynolds, G.P. and S.J. Pearson, *Increased brain 3-hydroxykynurenine in Huntington's disease.* Lancet, 1989. **2**(8669): p. 979-80.
128. Tomoda, A., Y. Yoneyama, T. Yamaguchi, E. Shirao, and K. Kawasaki, *Mechanism of coloration of human lenses induced by near-ultraviolet-photo-oxidized 3-hydroxykynurenine.* Ophthalmic Res, 1990. **22**(3): p. 152-9.

129. Malina, H.Z. and X.D. Martin, *Deamination of 3-hydroxykynurenine in bovine lenses: a possible mechanism of cataract formation in general.* Graefes Arch Clin Exp Ophthalmol, 1995. **233**(1): p. 38-44.
130. Aquilina, J.A., J.A. Carver, and R.J. Truscott, *Oxidation products of 3-hydroxykynurenine bind to lens proteins: relevance for nuclear cataract.* Exp Eye Res, 1997. **64**(5): p. 727-35.
131. Chiarugi, A., E. Rapizzi, F. Moroni, and F. Moroni, *The kynurenine metabolic pathway in the eye: studies on 3-hydroxykynurenine, a putative cataractogenic compound.* FEBS Lett, 1999. **453**(1-2): p. 197-200.
132. Sardar, A.M., J.E. Bell, and G.P. Reynolds, *Increased concentrations of the neurotoxin 3-hydroxykynurenine in the frontal cortex of HIV-1-positive patients.* J Neurochem, 1995. **64**(2): p. 932-5.
133. Savvateeva, E., A. Popov, N. Kamyshev, J. Bragina, M. Heisenberg, D. Senitz, J. Kornhuber, and P. Riederer, *Age-dependent memory loss, synaptic pathology and altered brain plasticity in the Drosophila mutant cardinal accumulating 3-hydroxykynurenine.* J Neural Transm, 2000. **107**(5): p. 581-601.
134. Mole, D.J., N.V. McFerran, G. Collett, C. O'Neill, T. Diamond, O.J. Garden, L. Kylanpaa, H. Repo, and E.A. Deitch, *Tryptophan catabolites in mesenteric lymph may contribute to pancreatitis-associated organ failure.* Br J Surg, 2008. **95**(7): p. 855-67.
135. Hayashi, T., J.H. Mo, X. Gong, C. Rossetto, A. Jang, L. Beck, G.I. Elliott, I. Kufareva, R. Abagyan, D.H. Broide, J. Lee, and E. Raz, *3-Hydroxyanthranilic acid inhibits PDK1 activation and suppresses experimental asthma by inducing T cell apoptosis.* Proc Natl Acad Sci U S A, 2007. **104**(47): p. 18619-24.
136. Goldstein, L.E., M.C. Leopold, X. Huang, C.S. Atwood, A.J. Saunders, M. Hartshorn, J.T. Lim, K.Y. Faget, J.A. Muffat, R.C. Scarpa, L.T. Chylack, Jr., E.F. Bowden, R.E. Tanzi, and A.I. Bush, *3-Hydroxykynurenine and 3-hydroxyanthranilic acid generate hydrogen peroxide and promote alpha-crystallin cross-linking by metal ion reduction.* Biochemistry, 2000. **39**(24): p. 7266-75.
137. Arbatova, J., E. D'Amato, A. Vaarmann, A. Zharkovsky, and M. Reeben, *Reduced serotonin and 3-hydroxyanthranilic acid levels in serum of cystatin B-deficient mice, a model system for progressive myoclonus epilepsy.* Epilepsia, 2005. **46 Suppl 5**: p. 49-51.
138. Darlington, L.G., C.M. Forrest, G.M. Mackay, R.A. Smith, A.J. Smith, N. Stoy, and T.W. Stone, *On the Biological Importance of the 3-hydroxyanthranilic Acid: Anthranilic Acid Ratio.* Int J Tryptophan Res, 2010. **3**: p. 51-9.
139. Holmes, F. and W.R.C. Crimmin, *The stabilities of metal chelate compounds formed by some heterocyclic acids. I. Studies in aqueous solution.* J. Chem. Soc., 1955(Copyright (C) 2013 American Chemical Society (ACS). All Rights Reserved.): p. 1175-80.
140. Beninger, R.J., A.M. Colton, J.L. Ingles, K. Jhamandas, and R.J. Boegman, *Picolinic acid blocks the neurotoxic but not the neuroexcitant properties of quinolinic acid in the rat brain: evidence from turning behaviour and tyrosine hydroxylase immunohistochemistry.* Neuroscience, 1994. **61**(3): p. 603-12.
141. Cockhill, J., K. Jhamandas, R.J. Boegman, and R.J. Beninger, *Action of picolinic acid and structurally related pyridine carboxylic acids on quinolinic acid-induced cortical cholinergic damage.* Brain Res, 1992. **599**(1): p. 57-63.
142. Grant, R.S., S.E. Coggan, and G.A. Smythe, *The physiological action of picolinic Acid in the human brain.* Int J Tryptophan Res, 2009. **2**: p. 71-9.

143. Smythe, G.A., A. Poljak, S. Bustamante, O. Braga, A. Maxwell, R. Grant, and P. Sachdev, *ECNI GC-MS analysis of picolinic and quinolinic acids and their amides in human plasma, CSF, and brain tissue.* Adv Exp Med Biol, 2003. **527**: p. 705-12.
144. Ceresoli-Borroni, G., P. Guidetti, L. Amori, R. Pellicciari, and R. Schwarcz, *Perinatal kynurenine 3-hydroxylase inhibition in rodents: pathophysiological implications.* J Neurosci Res, 2007. **85**(4): p. 845-54.
145. Heyes, M.P., *The kynurenine pathway and neurologic disease. Therapeutic strategies.* Adv Exp Med Biol, 1996. **398**: p. 125-9.
146. Massey, V., *Activation of molecular oxygen by flavins and flavoproteins.* J Biol Chem, 1994. **269**(36): p. 22459-62.
147. van Berkel, W.J., N.M. Kamerbeek, and M.W. Fraaije, *Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts.* J Biotechnol, 2006. **124**(4): p. 670-89.
148. Ballou, D.P., B. Entsch, and L.J. Cole, *Dynamics involved in catalysis by single-component and two-component flavin-dependent aromatic hydroxylases.* Biochem Biophys Res Commun, 2005. **338**(1): p. 590-8.
149. Sutton, W.B., *Mechanism of action and crystallization of lactic oxidative decarboxylase from Mycobacterium phlei.* J Biol Chem, 1957. **226**(1): p. 395-405.
150. Husain, M. and V. Massey, *Kinetic studies on the reaction of p-hydroxybenzoate hydroxylase. Agreement of steady state and rapid reaction data.* J Biol Chem, 1979. **254**(14): p. 6657-66.
151. Steennis, P.J., M.M. Cordes, J.H. Hilkens, and F. Muller, *On the interaction of para-hydroxybenzoate hydroxylase from Pseudomonas fluorescens with halogen ions.* FEBS Lett, 1973. **36**(2): p. 177-80.
152. Hosokawa, K. and R.Y. Stanier, *Crystallization and properties of p-hydroxybenzoate hydroxylase from Pseudomonas putida.* J Biol Chem, 1966. **241**(10): p. 2453-60.
153. Yano, K., N. Higashi, and K. Arima, *P-hydroxybenzoate hydroxylase: conformational changes in crystals of holoenzyme versus holoenzyme-substrate complex.* Biochem Biophys Res Commun, 1969. **34**(1): p. 1-7.
154. Howell, L.G., T. Spector, and V. Massey, *Purification and properties of p-hydroxybenzoate hydroxylase from Pseudomonas fluorescens.* J Biol Chem, 1972. **247**(13): p. 4340-50.
155. Drenth, J., W.G. Hol, and R.K. Wierenga, *Crystallization and preliminary x-ray investigation of p-hydroxybenzoate hydroxylase from Pseudomonas fluorescens.* J Biol Chem, 1975. **250**(13): p. 5268-9.
156. Wierenga, R.K., R.J. de Jong, K.H. Kalk, W.G. Hol, and J. Drenth, *Crystal structure of p-hydroxybenzoate hydroxylase.* J Mol Biol, 1979. **131**(1): p. 55-73.
157. Van der Laan, J.M., M.B. Swarte, H. Groendijk, W.G. Hol, and J. Drenth, *The influence of purification and protein heterogeneity on the crystallization of p-hydroxybenzoate hydroxylase.* Eur J Biochem, 1989. **179**(3): p. 715-24.
158. Schreuder, H.A., P.A. Prick, R.K. Wierenga, G. Vriend, K.S. Wilson, W.G. Hol, and J. Drenth, *Crystal structure of the p-hydroxybenzoate hydroxylase-substrate complex refined at 1.9 A resolution. Analysis of the enzyme-substrate and enzyme-product complexes.* J Mol Biol, 1989. **208**(4): p. 679-96.
159. Gatti, D.L., B.A. Palfey, M.S. Lah, B. Entsch, V. Massey, D.P. Ballou, and M.L. Ludwig, *The mobile flavin of 4-OH benzoate hydroxylase.* Science (Washington, D. C.) FIELD Full Journal Title:Science (Washington, DC, United States), 1994. **266**(5182): p. 110-14.
160. Schreuder, H.A., A. Mattevi, G. Obmolova, K.H. Kalk, W.G. Hol, F.J. van der Bolt, and W.J. van Berkel, *Crystal structures of wild-type p-hydroxybenzoate hydroxylase complexed with 4-aminobenzoate,2,4-dihydroxybenzoate, and 2-hydroxy-4-*

aminobenzoate and of the Tyr222Ala mutant complexed with 2-hydroxy-4-aminobenzoate. Evidence for a proton channel and a new binding mode of the flavin ring. Biochemistry, 1994. **33**(33): p. 10161-70.

161. Zheng, Y., J. Dong, B.A. Palfey, and P.R. Carey, *Using Raman spectroscopy to monitor the solvent-exposed and "buried" forms of flavin in p-hydroxybenzoate hydroxylase.* Biochemistry, 1999. **38**(51): p. 16727-32.
162. Palfey, B.A., D.P. Ballou, and V. Massey, *Flavin conformational changes in the catalytic cycle of p-hydroxybenzoate hydroxylase substituted with 6-azido- and 6-aminoflavin adenine dinucleotide.* Biochemistry, 1997. **36**(50): p. 15713-23.
163. Wang, J., M. Ortiz-Maldonado, B. Entsch, V. Massey, D. Ballou, and D.L. Gatti, *Protein and ligand dynamics in 4-hydroxybenzoate hydroxylase.* Proc Natl Acad Sci U S A, 2002. **99**(2): p. 608-13.
164. Eppink, M.H., C. Bunthol, H.A. Schreuder, and W.J. van Berkel, *Phe161 and Arg166 variants of p-hydroxybenzoate hydroxylase. Implications for NADPH recognition and structural stability.* FEBS Lett, 1999. **443**(3): p. 251-5.
165. Eppink, M.H., H.A. Schreuder, and W.J. van Berkel, *Interdomain binding of NADPH in p-hydroxybenzoate hydroxylase as suggested by kinetic, crystallographic and modeling studies of histidine 162 and arginine 269 variants.* J Biol Chem, 1998. **273**(33): p. 21031-9.
166. Eppink, M.H., H.A. Schreuder, and W.J. Van Berkel, *Identification of a novel conserved sequence motif in flavoprotein hydroxylases with a putative dual function in FAD/NAD(P)H binding.* Protein Sci, 1997. **6**(11): p. 2454-8.
167. Eppink, M.H., H.A. Schreuder, and W.J. Van Berkel, *Structure and function of mutant Arg44Lys of 4-hydroxybenzoate hydroxylase implications for NADPH binding.* Eur J Biochem, 1995. **231**(1): p. 157-65.
168. Eppink, M.H., K.M. Overkamp, H.A. Schreuder, and W.J. Van Berkel, *Switch of coenzyme specificity of p-hydroxybenzoate hydroxylase.* J Mol Biol, 1999. **292**(1): p. 87-96.
169. Eppink, M.H., H.A. Schreuder, and W.J. van Berkel, *Lys42 and Ser42 variants of p-hydroxybenzoate hydroxylase from Pseudomonas fluorescens reveal that Arg42 is essential for NADPH binding.* Eur J Biochem, 1998. **253**(1): p. 194-201.
170. Enroth, C., *High-resolution structure of phenol hydroxylase and correction of sequence errors.* Acta Crystallogr D Biol Crystallogr, 2003. **59**(Pt 9): p. 1597-602.
171. McCulloch, K.M., T. Mukherjee, T.P. Begley, and S.E. Ealick, *Structure of the PLP degradative enzyme 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase from Mesorhizobium loti MAFF303099 and its mechanistic implications.* Biochemistry, 2009. **48**(19): p. 4139-49.
172. Ryan, K.S., A.R. Howard-Jones, M.J. Hamill, S.J. Elliott, C.T. Walsh, and C.L. Drennan, *Crystallographic trapping in the rebeccamycin biosynthetic enzyme RebC.* Proc. Natl. Acad. Sci. U. S. A. FIELD Full Journal Title:Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(39): p. 15311-15316.
173. Treiber, N. and G.E. Schulz, *Structure of 2,6-dihydroxypyridine 3-hydroxylase from a nicotine-degrading pathway.* J Mol Biol, 2008. **379**(1): p. 94-104.
174. Lindqvist, Y., H. Koskiniemi, A. Jansson, T. Sandalova, R. Schnell, Z. Liu, P. Mantsala, J. Niemi, and G. Schneider, *Structural basis for substrate recognition and specificity in aklavinone-11-hydroxylase from rhodomycin biosynthesis.* J Mol Biol, 2009. **393**(4): p. 966-77.
175. Hiromoto, T., S. Fujiwara, K. Hosokawa, and H. Yamaguchi, *Crystal structure of 3-hydroxybenzoate hydroxylase from Comamonas testosteroni has a large tunnel for substrate and oxygen access to the active site.* J Mol Biol, 2006. **364**(5): p. 878-96.

176. Volkers, G., G.J. Palm, M.S. Weiss, G.D. Wright, and W. Hinrichs, *Structural basis for a new tetracycline resistance mechanism relying on the TetX monooxygenase.* FEBS Lett, 2011. **585**(7): p. 1061-6.
177. Koskiniemi, H., M. Metsa-Ketela, D. Dobritzsch, P. Kallio, H. Korhonen, P. Mantsala, G. Schneider, and J. Niemi, *Crystal structures of two aromatic hydroxylases involved in the early tailoring steps of angucycline biosynthesis.* J Mol Biol, 2007. **372**(3): p. 633-48.
178. Hicks, K.A., S.E. O'Leary, T.P. Begley, and S.E. Ealick, *Structural and Mechanistic Studies of HpxO, a Novel Flavin Adenine Dinucleotide-Dependent Urate Oxidase from Klebsiella pneumoniae.* Biochemistry, 2013. **52**(3): p. 477-87.
179. Oke, M., L.G. Carter, K.A. Johnson, H. Liu, S.A. McMahon, X. Yan, M. Kerou, N.D. Weikart, N. Kadi, M.A. Sheikh, S. Schmelz, M. Dorward, M. Zawadzki, C. Cozens, H. Falconer, H. Powers, I.M. Overton, C.A. van Niekerk, X. Peng, P. Patel, R.A. Garrett, D. Prangishvili, C.H. Botting, P.J. Coote, D.T. Dryden, G.J. Barton, U. Schwarz-Linek, G.L. Challis, G.L. Taylor, M.F. White, and J.H. Naismith, *The Scottish Structural Proteomics Facility: targets, methods and outputs.* J Struct Funct Genomics, 2010. **11**(2): p. 167-80.
180. Wierenga, R.K., P. Terpstra, and W.G. Hol, *Prediction of the occurrence of the ADP-binding beta alpha beta-fold in proteins, using an amino acid sequence fingerprint.* J Mol Biol, 1986. **187**(1): p. 101-7.
181. Saito, Y., O. Hayaishi, and S. Rothberg, *Studies on oxygenases; enzymatic formation of 3-hydroxy-L-kynurenine from L-kynurenine.* J Biol Chem FIELD Full Journal Title:The Journal of biological chemistry, 1957. **229**(2): p. 921-34.
182. Okamoto, H., S. Yamamoto, M. Nozaki, and O. Hayaishi, *On the submitochondrial localization of l-kynurenine-3-hydroxylase.* Biochem Biophys Res Commun, 1967. **26**(3): p. 309-14.
183. Okamoto, H. and O. Hayaishi, *Flavin adenine dinucleotide requirement for kynurenine hydroxylase of rat liver mitochondria.* Biochem Biophys Res Commun, 1967. **29**(3): p. 394-9.
184. Okamoto, H. and O. Hayaishi, *Solubilization and partial purification of kynurenine hydroxylase from mitochondrial outer membrane and its electron donors.* Arch Biochem Biophys, 1969. **131**(2): p. 603-8.
185. Nisimoto, Y., F. Takeuchi, and Y. Shibata, *Isolation of L-dynurenine 3-hydroxylase from the mitochondrial outer membrane of rat liver.* J Biochem, 1975. **78**(3): p. 573-81.
186. Nishimoto, Y., F. Takeuchi, and Y. Shibata, *Purification of L-kynurenine 3-hydroxylase by affinity chromatography.* J Chromatogr, 1979. **169**: p. 357-64.
187. Alberati-Giani, D., A.M. Cesura, C. Broger, W.D. Warren, S. Rover, and P. Malherbe, *Cloning and functional expression of human kynurenine 3-monooxygenase.* FEBS Letters, 1997. **410**(2-3): p. 407-412.
188. Breton, J., N. Avanzi, S. Magagnin, N. Covini, G. Magistrelli, L. Cozzi, and A. Isacchi, *Functional characterization and mechanism of action of recombinant human kynurenine 3-hydroxylase.* Eur J Biochem, 2000. **267**(4): p. 1092-9.
189. Uemura, T. and K. Hirai, *L-kynurenine 3-monooxygenase from mitochondrial outer membrane of pig liver: purification, some properties, and monoclonal antibodies directed to the enzyme.* J Biochem, 1998. **123**(2): p. 253-62.
190. Uemura, T. and K. Hirai, *Purification of L-kynurenine 3-monooxygenase from mitochondrial outer membrane of pig liver.* Adv Exp Med Biol, 1999. **467**: p. 619-23.
191. Hirai, K., H. Kuroyanagi, Y. Tatebayashi, Y. Hayashi, K. Hirabayashi-Takahashi, K. Saito, S. Haga, T. Uemura, and S. Izumi, *Dual role of the carboxyl-terminal region of pig liver*

L-kynurenine 3-monooxygenase: mitochondrial-targeting signal and enzymatic activity. J Biochem, 2010. **148**(6): p. 639-50.
192. Wattenberg, B.W., D. Clark, and S. Brock, *An artificial mitochondrial tail signal/anchoring sequence confirms a requirement for a moderate hydrophobicity for targeting.* Bioscience Reports, 2007. **27**(6): p. 385-401.
193. Kanaji, S., J. Iwahashi, Y. Kida, M. Sakaguchi, and K. Mihara, *Characterization of the signal that directs Tom20 to the mitochondrial outer membrane.* JCB, 2000. **151**(2): p. 277-288.
194. Isenmann, S., Y. Khew-Goodall, J. Gamble, M. Vadas, and B.W. Wattenberg, *A splice-isoform of vesicle-associated membrane protein-1 (VAMP-1) contains a mitochondrial targeting signal.* Mol Biol Cell, 1998. **9**(7): p. 1649-1660.
195. Hwang, Y.T., S.M. Pelitire, M.P.A. Henderson, D.W. Andrews, J.M. Dyer, and R.T. Mullen, *Novel targeting signals mediate the sorting of different isoforms of the tail-anchored membrane protein cytochrome b5 to either endoplasmic reticulum or mitochondria.* Plant Cell, 2004. **16**(11): p. 3002-3019.
196. Linzen, B. and U. Hertel, *In vitro assay of insect hynurenine-3-hydroxylase.* Naturwissenschaften, 1967. **54**(1): p. 21.
197. Ghosh, D. and H.S. Forrest, *Enzymatic studies on the hydroxylation of kynurenine in Drosophila melanogaster.* Genetics, 1967. **55**(3): p. 423-31.
198. Hirai, M., M. Kiuchi, J. Wang, A. Ishii, and H. Matsuoka, *cDNA cloning, functional expression and characterization of kynurenine 3-hydroxylase of Anopheles stephensi (Diptera: Culicidae).* Insect Mol Biol, 2002. **11**(5): p. 497-504.
199. Han, Q., E. Calvo, O. Marinotti, J. Fang, M. Rizzi, A.A. James, and J. Li, *Analysis of the wild-type and mutant genes encoding the enzyme kynurenine monooxygenase of the yellow fever mosquito, Aedes aegypti.* Insect Mol Biol, 2003. **12**(5): p. 483-90.
200. Stratakis, E., *Submitochondrial localization of kynurenine 3-hydroxylase from ovaries of Ephestia kuhniella.* Insect Biochem, 1981. **11**(6): p. 735-741.
201. Schott, H.H., V. Ullrich, and H. Staudinger, *Enzymatic properties of L-kynurenine 3-hydroxylase (EC 1.14.1.2) in Neurospora crassa.* Hoppe Seylers Z Physiol Chem, 1970. **351**(1): p. 99-101.
202. Bhalla, S.C., *White eye, a new sex-linked mutant of Aedes Aegypti.* Mosquito News, 1968. **28**(380-385).
203. Cornel, A.J., M.Q. Benedict, C.S. Rafferty, A.J. Howells, and F.H. Collins, *Transient expression of the Drosophila melanogaster cinnabar gene rescues eye color in the white eye (WE) strain of Aedes aegypti.* Insect Biochem Mol Biol, 1997. **27**(12): p. 993-7.
204. Shin, M., K. Sano, and C. Umezawa, *Inhibition of L-kynurenine 3-hydroxylase from Saccharomyces carlsbergensis by alpha-keto acid derivatives of branched chain amino acids.* J Nutr Sci Vitaminol (Tokyo), 1982. **28**(3): p. 191-201.
205. Shin, M., O. Yamada, K. Sano, and C. Umezawa, *Effect of branched chain alpha-keto acids on kynurenine 3-hydroxylase from rat liver.* J Nutr Sci Vitaminol (Tokyo), 1982. **28**(3): p. 203-7.
206. Pellicciari, R., B. Natalini, G. Costantino, M.R. Mahmoud, L. Mattoli, B.M. Sadeghpour, F. Moroni, A. Chiarugi, and R. Carpenedo, *Modulation of the kynurenine pathway in search for new neuroprotective agents. Synthesis and preliminary evaluation of (m-nitrobenzoyl)alanine, a potent inhibitor of kynurenine-3-hydroxylase.* J Med Chem, 1994. **37**(5): p. 647-55.
207. Speciale, C., H.Q. Wu, M. Cini, M. Marconi, M. Varasi, and R. Schwarcz, *(R,S)-3,4-dichlorobenzoylalanine (FCE 28833A) causes a large and persistent increase in brain kynurenic acid levels in rats.* Eur J Pharmacol, 1996. **315**(3): p. 263-7.

208. Giordani, A., P. Pevarello, M. Cini, R. Bormetti, F. Greco, S. Toma, C. Speciale, and M. Varasi, *4-Phenyl-4-oxo-butanoic acid derivatives inhibitors of kynurenine 3-hydroxylase.* Bioorganic & medicinal chemistry letters, 1998. **8**: p. 2907-2912.
209. Drysdale, M.J., S.L. Hind, M. Jansen, and J.F. Reinhard, Jr., *Synthesis and SAR of 4-aryl-2-hydroxy-4-oxobut-2-enoic acids and esters and 2-amino-4-aryl-4-oxobut-2-enoic acids and esters: potent inhibitors of kynurenine-3-hydroxylase as potential neuroprotective agents.* J Med Chem, 2000. **43**(1): p. 123-7.
210. Feng, Y., B.F. Bowden, and V. Kapoor, *Ianthellamide A, a selective kynurenine-3-hydroxylase inhibitor from the Australian marine sponge Ianthella quadrangulata.* Bioorg Med Chem Lett, 2012. **22**(10): p. 3398-401.
211. Chiarugi, A. and F. Moroni, *Quinolinic acid formation in immune-activated mice: studies with (m-nitrobenzoyl)-alanine (mNBA) and 3,4-dimethoxy-[-N-4-(-3-nitrophenyl)thiazol-2yl]-benzenesul fonamide (Ro 61-8048), two potent and selective inhibitors of kynurenine hydroxylase.* Neuropharmacology, 1999. **38**(8): p. 1225-33.
212. Dale, W.E., Y. Dang, N. Amiridze, and O.R. Brown, *Evidence that kynurenine pathway metabolites mediate hyperbaric oxygen-induced convulsions.* Toxicol Lett, 2000. **117**(1-2): p. 37-43.
213. Chiarugi, A., R. Carpenedo, and F. Moroni, *Kynurenine disposition in blood and brain of mice: effects of selective inhibitors of kynurenine hydroxylase and of kynureninase.* J Neurochem, 1996. **67**(2): p. 692-8.
214. Urenjak, J. and T.P. Obrenovitch, *Kynurenine 3-hydroxylase inhibition in rats: effects on extracellular kynurenic acid concentration and N-methyl-D-aspartate-induced depolarisation in the striatum.* J Neurochem, 2000. **75**(6): p. 2427-33.
215. Samadi, P., L. Gregoire, A. Rassoulpour, P. Guidetti, E. Izzo, R. Schwarcz, and P.J. Bedard, *Effect of kynurenine 3-hydroxylase inhibition on the dyskinetic and antiparkinsonian responses to levodopa in Parkinsonian monkeys.* Mov Disord, 2005. **20**(7): p. 792-802.
216. Zwilling, D., S.Y. Huang, K.V. Sathyasaikumar, F.M. Notarangelo, P. Guidetti, H.Q. Wu, J. Lee, J. Truong, Y. Andrews-Zwilling, E.W. Hsieh, J.Y. Louie, T. Wu, K. Scearce-Levie, C. Patrick, A. Adame, F. Giorgini, S. Moussaoui, G. Laue, A. Rassoulpour, G. Flik, Y. Huang, J.M. Muchowski, E. Masliah, R. Schwarcz, and P.J. Muchowski, *Kynurenine 3-monooxygenase inhibition in blood ameliorates neurodegeneration.* Cell, 2012. **145**(6): p. 863-74.
217. Obrenovitch, T.P. and J. Urenjak, *In vivo assessment of kynurenate neuroprotective potency and quinolinate excitotoxicity.* Amino Acids, 2000. **19**(1): p. 299-309.
218. Moroni, F., A. Cozzi, R. Carpendo, G. Cipriani, O. Veneroni, and E. Izzo, *Kynurenine 3-mono-oxygenase inhibitors reduce glutamate concentration in the extracellular spaces of the basal ganglia but not in those of the cortex or hippocampus.* Neuropharmacology, 2005. **48**(6): p. 788-95.
219. Cornish-Bowden, A., *A simple graphical method for determining the inhibition constants of mixed, uncompetitive and non-competitive inhibitors.* Biochem J, 1974. **137**(1): p. 143-4.
220. Chayen, N.E., *Methods for separating nucleation and growth in protein crystallisation.* Prog Biophys Mol Biol, 2005. **88**(3): p. 329-37.
221. Heras, B. and J.L. Martin, *Post-crystallization treatments for improving diffraction quality of protein crystals.* Acta Crystallogr D Biol Crystallogr, 2005. **61**(Pt 9): p. 1173-80.
222. Manjasetty, B.A., N. Croteau, J. Powlowski, and A. Vrielink, *Crystallization and preliminary X-ray analysis of dmpFG-encoded 4-hydroxy-2-ketovalerate aldolase--*

aldehyde dehydrogenase (acylating) from Pseudomonas sp. strain CF600. Acta Crystallogr D Biol Crystallogr, 2001. **57**(Pt 4): p. 582-5.
223. Guan, Z., L. Hederstedt, J. Li, and X.D. Su, *Preparation and crystallization of a Bacillus subtilis arsenate reductase.* Acta Crystallogr D Biol Crystallogr, 2001. **57**(Pt 11): p. 1718-21.
224. Heras, B., M.A. Edeling, K.A. Byriel, A. Jones, S. Raina, and J.L. Martin, *Dehydration converts DsbG crystal diffraction from low to high resolution.* Structure, 2003. **11**(2): p. 139-45.
225. QUIOCHO, F.A. and F.M. RICHARDS, *INTERMOLECULAR CROSS LINKING OF A PROTEIN IN THE CRYSTALLINE STATE: CARBOXYPEPTIDASE-A.* Proc Natl Acad Sci U S A, 1964. **52**: p. 833-9.
226. Haas, D.J. and M.G. Rossmann, *Crystallographic studies on lactate dehydrogenase at-75 degrees C.* Acta Crystallogr B, 1970. **26**(7): p. 998-1004.
227. Peters, K. and F.M. Richards, *Chemical cross-linking: reagents and problems in studies of membrane structure.* Annu Rev Biochem, 1977. **46**: p. 523-51.
228. Lusty, C.J., *A gentle vapor-diffusion technique for cross-linking of protein crystals for cryocrystallography*. 1999: J. Appl. Cryst. p. 106-112.
229. Price, W.N., Y. Chen, S.K. Handelman, H. Neely, P. Manor, R. Karlin, R. Nair, J. Liu, M. Baran, J. Everett, S.N. Tong, F. Forouhar, S.S. Swaminathan, T. Acton, R. Xiao, J.R. Luft, A. Lauricella, G.T. DeTitta, B. Rost, G.T. Montelione, and J.F. Hunt, *Understanding the physical properties that control protein crystallization by analysis of large-scale experimental data.* Nat Biotechnol, 2009. **27**(1): p. 51-7.
230. Canaves, J.M., R. Page, I.A. Wilson, and R.C. Stevens, *Protein biophysical properties that correlate with crystallization success in Thermotoga maritima: maximum clustering strategy for structural genomics.* J Mol Biol, 2004. **344**(4): p. 977-91.
231. Derewenda, Z.S., *Application of protein engineering to enhance crystallizability and improve crystal properties.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 5): p. 604-15.
232. Derewenda, Z.S., *It's all in the crystals….* Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 4): p. 243-8.
233. Patel, S.B., P.M. Cameron, B. Frantz-Wattley, E. O'Neill, J.W. Becker, and G. Scapin, *Lattice stabilization and enhanced diffraction in human p38 alpha crystals by protein engineering.* Biochim Biophys Acta, 2004. **1696**(1): p. 67-73.
234. Birtley, J.R. and S. Curry, *Crystallization of foot-and-mouth disease virus 3C protease: surface mutagenesis and a novel crystal-optimization strategy.* Acta Crystallogr D Biol Crystallogr, 2005. **61**(Pt 5): p. 646-50.
235. Niessing, D., S. Hüttelmaier, D. Zenklusen, R.H. Singer, and S.K. Burley, *She2p is a novel RNA binding protein with a basic helical hairpin motif.* Cell, 2004. **119**(4): p. 491-502.
236. Fontana, A., G. Fassina, C. Vita, D. Dalzoppo, M. Zamai, and M. Zambonin, *Correlation between sites of limited proteolysis and segmental mobility in thermolysin.* Biochemistry, 1986. **25**(8): p. 1847-51.
237. Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics.* Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 12 Pt 1): p. 2126-32.
238. Zheng, S., B. Chen, X. Qiu, K. Lin, and X. Yu, *Three novel cytochrome P450 genes identified in the marine polychaete Perinereis nuntia and their transcriptional response to xenobiotics.* Aquat Toxicol, 2013. **134-135**: p. 11-22.
239. Winn, M.D., G.N. Murshudov, and M.Z. Papiz, *Macromolecular TLS refinement in REFMAC at moderate resolutions.* Methods Enzymol, 2003. **374**: p. 300-21.

240. Battye, T.G., L. Kontogiannis, O. Johnson, H.R. Powell, and A.G. Leslie, *iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM.* Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 4): p. 271-81.
241. Singh, R.P., R.N. Singh, M.K. Srivastava, A.K. Srivastava, S. Kumar, R.C. Dubey, and A.K. Sharma, *Structure prediction and analysis of MxaF from obligate, facultative and restricted facultative methylobacterium.* Bioinformation, 2012. **8**(21): p. 1042-6.
242. Winn, M.D., C.C. Ballard, K.D. Cowtan, E.J. Dodson, P. Emsley, P.R. Evans, R.M. Keegan, E.B. Krissinel, A.G. Leslie, A. McCoy, S.J. McNicholas, G.N. Murshudov, N.S. Pannu, E.A. Potterton, H.R. Powell, R.J. Read, A. Vagin, and K.S. Wilson, *Overview of the CCP4 suite and current developments.* Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 4): p. 235-42.
243. Evans, G. and R. Pettifer, *CHOOCH: a program for deriving anomalous-scattering factors from X-ray fluorescence spectra*. 2001: J. Appl. Cryst. p. 82-86.
244. Han, Q., T. Cai, D.A. Tagle, H. Robinson, and J. Li, *Substrate specificity and structure of human aminoadipate aminotransferase/kynurenine aminotransferase II.* Biosci Rep, 2008. **28**(4): p. 205-15.
245. Han, Q., H. Robinson, T. Cai, D.A. Tagle, and J. Li, *Biochemical and structural properties of mouse kynurenine aminotransferase III.* Mol Cell Biol, 2009. **29**(3): p. 784-93.
246. Han, Q., H. Robinson, T. Cai, D.A. Tagle, and J. Li, *Biochemical and structural characterization of mouse mitochondrial aspartate aminotransferase, a newly identified kynurenine aminotransferase-IV.* Biosci Rep, 2011. **31**(5): p. 323-32.
247. Lima, S., S. Kumar, V. Gawandi, C. Momany, and R.S. Phillips, *Crystal structure of the Homo sapiens kynureninase-3-hydroxyhippuric acid inhibitor complex: insights into the molecular basis of kynureninase substrate specificity.* J Med Chem, 2009. **52**(2): p. 389-96.
248. Slabinski, L., L. Jaroszewski, L. Rychlewski, I.A. Wilson, S.A. Lesley, and A. Godzik, *XtalPred: a web server for prediction of protein crystallizability.* Bioinformatics, 2007. **23**(24): p. 3403-5.
249. Goldschmidt, L., D.R. Cooper, Z.S. Derewenda, and D. Eisenberg, *Toward rational protein crystallization: A Web server for the design of crystallizable protein variants.* Protein Sci, 2007. **16**(8): p. 1569-76.
250. Prilusky, J., C.E. Felder, T. Zeev-Ben-Mordehai, E.H. Rydberg, O. Man, J.S. Beckmann, I. Silman, and J.L. Sussman, *FoldIndex: a simple tool to predict whether a given protein sequence is intrinsically unfolded.* Bioinformatics, 2005. **21**(16): p. 3435-8.
251. Zhang, Y., *I-TASSER server for protein 3D structure prediction.* BMC Bioinformatics, 2008. **9**: p. 40.
252. Arnold, K., L. Bordoli, J. Kopp, and T. Schwede, *The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling.* Bioinformatics, 2006. **22**(2): p. 195-201.
253. Kelley, L.A. and M.J. Sternberg, *Protein structure prediction on the Web: a case study using the Phyre server.* Nat Protoc, 2009. **4**(3): p. 363-71.
254. Teodorescu, O., T. Galor, J. Pillardy, and R. Elber, *Enriching the sequence substitution matrix by structural information.* Proteins, 2004. **54**(1): p. 41-8.
255. Murshudov, G.N., A.A. Vagin, and E.J. Dodson, *Refinement of macromolecular structures by the maximum-likelihood method.* Acta Crystallogr D Biol Crystallogr, 1997. **53**(Pt 3): p. 240-55.
256. Chen, V.B., W.B. Arendall, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, L.W. Murray, J.S. Richardson, and D.C. Richardson, *MolProbity: all-atom structure*

validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 1): p. 12-21.
257. Holm, L., S. Kääriäinen, P. Rosenström, and A. Schenkel, *Searching protein structure databases with DaliLite v.3.* Bioinformatics, 2008. **24**(23): p. 2780-1.
258. Valverde, R., L. Edwards, and L. Regan, *Structure and function of KH domains.* FEBS J, 2008. **275**(11): p. 2712-26.
259. Abe, I., K. Kashiwagi, and H. Noguchi, *Antioxidative galloyl esters as enzyme inhibitors of p-hydroxybenzoate hydroxylase.* FEBS Lett, 2000. **483**(2-3): p. 131-4.

Appendix 1 – Amino acid sequence analysis

1.1 Emboss needle pairwise alignment to human KMO

```
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 505
# Identity:     160/505 (31.7%)
# Similarity:   257/505 (50.9%)
# Gaps:          63/505 (12.5%)
# Score: 681.0


HumanKMO          1 ---MDSSVIQRKKVAVIGGGLVGSLQACFLAKRNFQIDVYEAREDTRVAT     47
                        .|::    ::|.:||.||.|:|.|..||:..:|::::|.|.|.|:.|
PfKMO             1 MTATDNA----RQVTIIGAGLAGTLVARLLARNGWQVNLFERRPDPRIET     46

HumanKMO         48 FTRGRSINLALSHRGRQALKAVGLEDQIVSQGIPMRARMIHSLSGKKSAI     97
                    ..|||||||||:.||..||:..|||.:::::.:.||.||:|......:..
PfKMO            47 GARGRSINLALAERGAHALRLAGLEREVLAEAVMMRGRMVHVPGTPPNLQ     96

HumanKMO         98 PYG-TKSQYILSVSRENLNKDLLTAAEKYPNVKMHFNHRLLKCNPEEGMI    146
                    ||| ..|:.|.|::|:.||:.||..||. ....:|||..|...:.....:
PfKMO            97 PYGRDDSEVIWSINRDRLNRILLDGAEA-AGASIHFNLGLDSVDFARQRL    145

HumanKMO        147 TV--LGSDKVPKDVTCDLIVGCDGAYSTVRSHLMKKPRFDYSQQYIPHGY    194
                    |:  :..:::.|  ...|::|.||..|.||..:..........:..||||
PfKMO           146 TLSNVSGERLEK--RFHLLIGADGCNSAVRQAMASVVDLGEHLETQPHGY    193

HumanKMO        195 MELTIPPK-NGDYAMEPNYLHIWPRNTFMMIALPNMNKSFTCTLFMPFEE    243
                    .||.|.|: :..:.:|||.|||||...:|.|||||::|||.|||:..:.
PfKMO           194 KELQITPEASAQFNLEPNALHIWPHGDYMCIALPNLDRSFTVTLFLHHQS    243

HumanKMO        244 ---------FEKLLTSNDVVDFFQKYFPDAIPLIGEKLLVQDFFLLPAQP    284
                            |.:|:..:....|||:.|||..|::..  |.|||...|...
PfKMO           244 PAAQPASPCFAQLVDGHAARRFFQRQFPDLSPMLDS--LEQDFEHHPTGK    291

HumanKMO        285 MISVKCSSFHFKSHCVLLGDAAHAIVPFFGQGMNAGFEDCLVFDELMDKF    334
                    :.:::.:::|.....||||||||.:|||.|||||...||.:...|.:...
PfKMO           292 LATLRLTTWHVGGQAVLLGDAAHPMVPFHGQGMNCALEDAVALAEHLQSA    341

HumanKMO        335 SNDLSLCLPVFSRLRIPDDHAISDLSMYNYIEMRAHVNSSWFIFQKNMER    384
                    :::.| .|..|:..|.||..||..:::.||:||.:.|.|..::.::.:.:
PfKMO           342 ADNAS-ALAAFTAQRQPDALAIQAMALENYVEMSSKVASPTYLLERELGQ    390

HumanKMO        385 FLHAIMPSTFIPLYTMVTFSRIRYHEAVQRWHWQKKVINKGLFFLGSLIA    434
                    .:....|:.|||.|:||||||:.|.:|:.|...|::::...:.....|.:
PfKMO           391 IMAQRQPTRFIPRYSMVTFSRLPYAQAMARGQIQEQLLKFAVANHSDLTS    440

HumanKMO        435 I---SSTYLLIHYMSPRSFLRLRRPWNWIAHFRNTTCFPAKAVDSLEQIS    481
                    |   :..:.:...:.|.|.|.
PfKMO           441 INLDAVEHEVTRCLPPLSHLC----------------------------    461

HumanKMO        482 NLISR    486

PfKMO           462 -----    461
```

1.2 Multiple sequence alignment with FAH structural data

Conserved residues across all KMO sequences are highlighted on the sequence of *Pf*KMO in cyan. From the FAH structures residues involved in FAD-binding are highlighted yellow and in substrate-binding are highlighted green. Created with Clustal W2.

```
PfKMO       -------------------------MTATDNARQVTIIGAGLAGTLVARLLARNG----- 30
p-HBH       ------------------------------MKTQVAIIGAGPSGLLLGQLLHKAG----- 25
PhH         -------------------------MTKYSESYCDVLIVGAGPAGLMAARVLSEYVRQKPD 36
PQSL        ----------------------------MTDNHIDVLINGCGIGGAMLAYLLGRQG----- 28
MHPCO       ----------------------MANVNKTPGKTRRAEVAGGGFAGLTAAIALKQNG----- 34
RebC        --------------------------MNAPIETDVLILGGGGPVGMALALDLAHRQ----- 29
phzS        ---------------------------MSEPIDILIAGAGIGGLSCALALHQAG----I 28
AVH         ---------------------------MNDHEVDVLVVGAGLGGLSTAMFLARQG----- 28
m-HBH       MQFHLNGFRPGNPLIAPASPLAPAHTEAVPSQVDVLIVGCGPAGLTLAAQLAAFP----D 56
                                            : * *  *   .  *

PfKMO       WQVNLFERRPDPRIET--GARG---RSINLALAERGAHALRLAG--LEREVLAEAVMMRG 83
p-HBH       IDNVILERQTPDYVLGRIRAGVLEQGMVDLLREAGVDRRMARDG------LVHEGVEIAF 79
PhH         LKVRIIDKRSTKVYNG--QADGLQCRTLESLKNLGLADKILSEANDMSTIALYNPD--EN 92
PQSL        HRVVVVEQARRERAIN--GADLLKPAGIRVVEAAGLLAEVTRRG----GRVRHELEVYHD 82
MHPCO       WDVRLHEKSSELRAFG--AGIYLWHNGLRVLEGLGALDDVLQGS---------HTPPTYE 83
RebC        VGHLVVEQTDGTITHP--RVGTIGPRSMELFRRWGVAKQIRTAG-WPGDHPLDAAWVTRV 86
phzS        GKVTLLESSSEIRPLG--VGINIQPAAVEALAELGLGPALAATA-----IPTHELRYIDQ 81
AVH         VRVLVVERRPGLSPYP--RAAGQNPRTMELLRIGGVADEVVRADDIRGTQGDFVIRLAES 86
m-HBH       IRTCIVEQKEGPMELG--QADGIACRTMEMFEAFEFADSILKEACWINDVTFWKPDPAQP 114
                : :                   :          :

PfKMO       RMVHVPGTPPNLQPYGRDDSEVIWSINR-----DRLNRILLDGAEAAGASIHFNLGLDSV 138
p-HBH       AGQRRRIDLKRLSGGKTVTVYGQTEVTR---------DLMEAREACGATTVYQAAEVRLH 130
PhH         GHIRRTDRIPDTLPGISRYHQVVLHQGR-IERR-ILDSIAEISDTRIKVERPLIPEKMEI 150
PQSL        GELLRYFNYSSVDARGYFILMPCESLRR---------LVLEKIDGEATVEMLFETRIEAV 133
MHPCO       TWMHNKSVSKETFNGLPWRIMTRSHLHD----------ALVNRARALGVDISVNSEAVAA 133
RebC        GGHEVYRIPLGTADTRATPEHTPEPDAIC-----PQHWLAPLLAEAVGERLRTRSRLDSF 141
phzS        SGATVWSEPRGVEAGNAYPQYSIHRGEL------QMILLAAVRERLGQQAVRTGLGVERI 135
AVH         VRGEILRTVSESFDDMVAATEPCTPAGWAMLSQDKLEPILLAQARKHGGAIRFGTRLLSF 146
m-HBH       GRIARHGRVQDTEDGLSEFPHVILNQAR-VHDH-YLERMRNS-PSRLEPHYARRVLDVKI 171


PfKMO       DFARQR--------LTLSNVS---------------------------------------- 151
p-HBH       DLQGER--------PYVTFER---------------------------------------- 143
PhH         DSSKAEDPEAYPVTMTLRYMSEDESTPLQFGHKTENGLFRSNLQTQEEEDANYRLPEGKE 210
PQSL        QRDERH---------AIDQVR---------------------------------------- 145
MHPCO       D---------------PVGRLT---------------------------------------- 140
RebC        EQRDDH------VRATITDLR---------------------------------------- 156
phzS        EERDGR--------VLIGARDG--------------------------------------- 149
AVH         RQHDDDAG----AGVTARLAG---------------------------------------- 163
m-HBH       DHGAAD----YPVTVTLERCD-----------------------------------AAH 191


PfKMO       -GERLEKRFHLLIGADGCNSAVRQAMAS-VVDLGEHLETQPHGYKELQITPEASAQFNLE 209
p-HBH       DGERLRLDCDYIAGCDGFHGISRQSIP--------AERLKVFERVYPFGWLGLLADTPPV 195
PhH         AGEIETVHCKYVIGCDGGHSWVRRTLG--FEMIG-EQTDYIWGVLD-AVPASNFPDIRSR 266
PQSL        LNDGRVLRPRVVVGADGIASYVRRRL----LDIDVERRPYPSPMLVGTFALAPCVAERNR 201
MHPCO       LQTGEVLEADLIVGADGVGSKVRDSIG-FKQDRWVSKDGLIRLIVPRMKKELGHGEWDNT 199
RebC        TGATRAVHARYLVACDGASSPTRKALG--IDAPPRHRTQVFRNILFRAPELRSLLGERAA 214
phzS        HGKPQALGADVLVGADGIHSAVRAHLHPDQRPLSHGGITMWRGVTEFDRFLDGKTMIVAN 209
AVH         PDGEYDLRAGYLVGADGNRSLVRESLG--IGRYGHGTLTHMVGVIF-DADLSGIMEPGTT 220
m-HBH       AGQIETVQARYVVGCDGARSNVRRAIG--RQLVG-DSANQAWGVMD-VLAVTDFPDVRYK 247
                      : ..**  .  *  :
```

```
PfKMO       PNALHIWPHGDYMCIALPNLDRSFTVTLFLHHQSPAAQPASPCFAQLVDGHAARRFFQRQ 269
p-HBH       SHELIYANHPRGFALCSQRSATRSRYYVQVPLTEK----------VEDWSDERFWTELKA 245
PhH         --CAIHSAESGSIMIIPREN-NLVRFYVQLQARAE--KGGR----VDRTKFTPEVVIANA 317
PQSL        ----LYVDSQGGLAYFYPIGFDRARLVVSFPREEA--RELM----ADTRGESLRRRLQRF 251
MHPCO       IDMWNFWPRVQRILYSPCNENELYLGLMAPAADPR----------GSSVPIDLEVWVEMF 249
RebC        LFFFLMLSSSLRFPLRALDGRGLYRLTVGVDD-------------------ASKSTMDSF 255
phzS        DEHWSRLVAYPISARHAAEGKSLVNWVCMVPSAAVGQLDNEADWNRDGRLEDVLPFFADW 269
AVH         G-WYYLHHPEFKGTFGPTDRPDRHTLFVEYDP------DEG----ERPEDFTPQRCVELI 269
m-HBH       --VAIQSEQ-GNVLIIPREGGHLVRFYVEMDK-LD--ADER----VASRNITVEQLIATA 297

PfKMO       FPDLSPMLDSLEQDFEHHPTGKLATLRLTTWHVGG----------QAVLLGDAAHPMVP 318
p-HBH       RLPAEVAEKLVTGPSLEKSIAPLRSFVVEPMQHGR----------LFLAGDAAHIVPP 293
PhH         KKIFHPYTFDVQQLDWFTA-YHIGQRVTEKFSKDE----------RVFIAGDACHTHSP 365
PQSL        VGDESAEAIAAVTGTSRFKGIPIGYLNLDRYWADN----------VAMLGDAIHNVHP 299
MHPCO       PFLEPCLIEAAKLKTARYDKYETTK--LDSWTRGK----------VALVGDAAHAMCP 295
RebC        ELVRRAVAFDTEIEVLSDSEWHLTHRVADSFSAGR----------VFLTGDAAHTLSP 303
phzS        DLGWFDIRDLLTRNQLILQYPMVDRDPLPHWGRGR----------ITLLGDAAHLMYP 317
AVH         GLALDAPEVKPELVDIQG--WEMAARIAERWREGR----------VFLAGDAAKVTPP 315
m-HBH       QRVLHPYKLDVKNVPWWSV-YEIGQRICAKYDDVADAVATPDSPLPRVFIAGDACHTHSP 356
                                                          : *** :   *

PfKMO       FHGQGMNCALEDAVALAEHLQSAADNASA----LAAFTAQRQPDALAIQAMALEN----- 369
p-HBH       TGAKGLNLAASDVSTLYRLLLKAYREGRG--ELLERYSAICLRRIWKAERFSW------- 344
PhH         KAGQGMNTSMMDTYNLGWKLGLVLTGRAK-RDILKTYEEERQPFAQALIDFDHQFSRLFS 424
PQSL        ITGQGMNLAIEDASALADALDLALRDACALEDALAGYQAERFPVNQAIVSYGH------- 352
MHPCO       ALAQGAGCAMVNAFSLSQDLEEGSSVEDA----LVAWETRIRPITDRCQALSG------- 344
RebC        SGGFGMNTGIGSAADLGWKLAATLRGWAG-PGLLATYEEERRPVAITSLEEAN------- 355
phzS        MGANGASQAILDGIELAAALARNADVAAA----LREYEEARRPTANKIILANR------- 366
AVH         TGGMSGNAAVADGFDLAWKLAAVLQGQAG-AGLLDTYEDERKVAAELVVAEAL------- 367
m-HBH       KAGQGMNFSMQDSFNLGWKLAAVLRKQCA-PELLHTYSSERQVVAQQLIDFDREWAKMFS 415
              . . . .  .   *   *            *  :

PfKMO       -----------------YVEMSSKVASPTYLLERELG-----------------QIMAQR 395
p-HBH       -----------------WMTSVLHRFPDTDAFSQRIQ----------------QTELEYY 371
PhH         GRPAKDVADEMGVSMDVFKEAFVKGNEFASGTAINYDENLVTDKKSSKQELAKNCVVGTR 484
PQSL        ----------------ALATSLEDRQRFAGVFDTALQ--------------------GSS 376
MHPCO       ----------------------------DYAANRSLS-------------------KGNM 357
RebC        -----------VNLRRTMDRELPPGLHDDGPRGERIR-----------------AAVAEK 387
phzS        ----------------------EREKEEWAAASRPK----------------------T 381
AVH         ---------------AIYAQRMAPHMAEVWDKSVGYP----------------ETLLGFR 396
m-HBH       --DPAKEGGQGGVDPKEFQKYFEQHGRFTAGVGTHYAPSLLTG-QASHQALASGFTVGMR 472

PfKMO       QPTRFIPRYSMVTFSRLPYAQAMARGQIQEQLLKFAVAN-HSDLTSINLDAVEHEVTRCL 454
p-HBH       LGSEAGLATIAENYVGLPYEEIE  -  394
PhH         FKSQPVVRHSEGLWMHFGDRLVT---DGRFRIIVFAGKATDATQMSRIKKFAAYLDSENS 541
PQSL        RTPEALGGERSYQPVRSPAPLG  -  398
MHPCO       FTPAALEAARYDPLRRVYSWPQ  -  379
RebC        LERSGARREFDAPGIHFGHTYRSSIVCGEPETEVATGGWRPSARPGARAPHAWLTPTTST 447
phzS        EKSAALEAITGSYRNQVERPR  -  402
AVH         YRSSAVLATDDDP-ARVENPLTP---SGRPG---FRGPHVLVSRHGERLSTVDLFGDGWT 449
m-HBH       FHSAPVVRVSDAKPLQLGHCGKA---DGRWRLYAFAGQNDLAQPESGLLALCRFLESDAA 529

PfKMO       PPLSHLC  -  461
PhH         -VISRYTPKGADRNSRIDVITIHSCHRDDIEMHDFPAPALHPKWQY---DFIYADCDSWH 597
RebC        LDLFGRG-----------FVLLSFGTTDGVEAVTR---------------AFADRHVPLE 481
AVH         LLAGELG---------ADWVAAAEAVSAELGVPVR---------------AYRVGAGLT 484
m-HBH       SPLRRFTPSGQDIDSIFDLRAIFPQAYTEVALETLPALLLPPKGQLGMIDYEKVFSPDLK 589
                                    :                          .

PhH         HPHPKSYQAWGVDETKGAVVVVRPDGYTSLVTDLEGTAEIDRYFSGILVEPKEKSGAQTE 657
RebC        TVTCHAPEIHALY--ERAHVLVRPDGHVAWRGD-HLPAELGGLVDKVRGAA  -  529
AVH         DPESAVSERYGIG--KAGASLVRPDGIVAWRTD-EAAADAAQTLEGVLRRVLDR  -  535
m-HBH       NAGQDIFELRGIDRQQGALVVVRPDQYVAQVLPLGDHAALSAYFESFMRA  -  639
                   :  ::   . : .:****  .:       * .   .  .
```

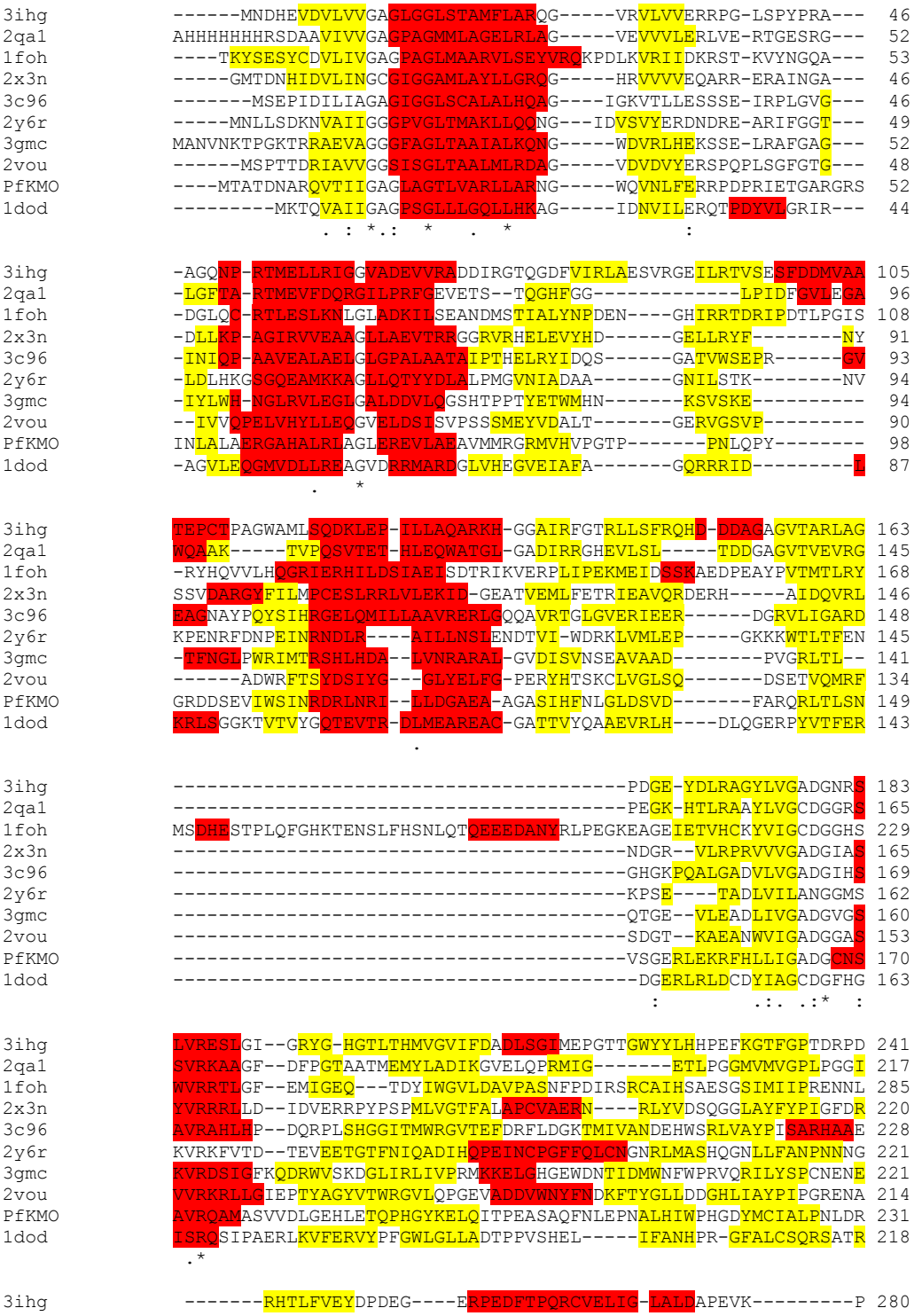
1.3 Analysis of KMO sequence alignments

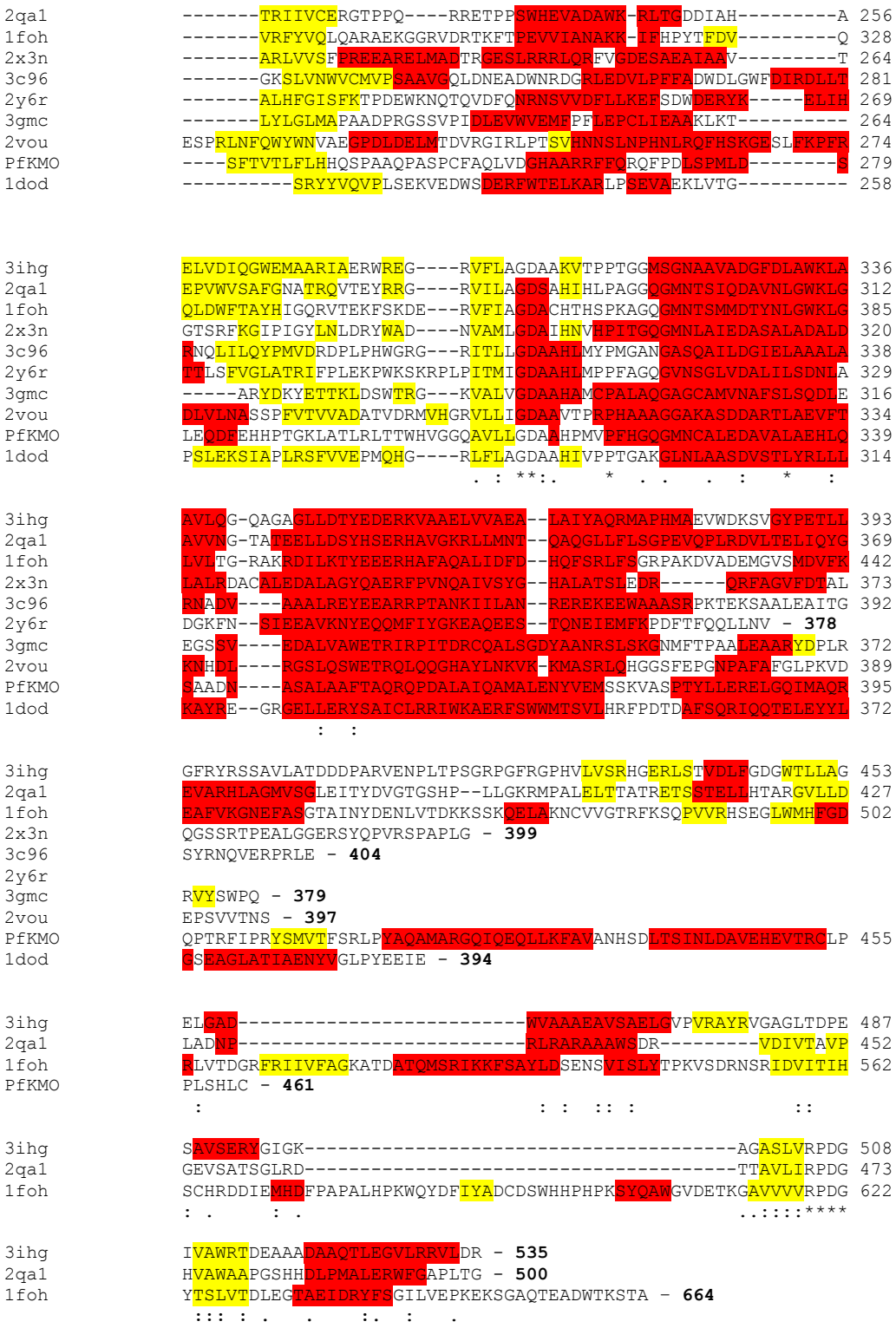
Residues that are 100 % identical (green) and 100 % conserved (yellow) from separate KMO sequence alignments are highlighted on the respective representative sequences.
H.sapiens KMO represents an alignment of mammalian KMO sequences (12 sequences).
C.elegans KMO represents an alignment of worm and insect KMOs (37 sequences).
S.cerevisiae KMO represents an alignment of fungi KMOs (35 sequences).
*Pf*KMO (All bact) represents an alignment of bacterial KMOs (43 sequences).
*Pf*KMO (close Hs) represents an alignment of just the most similar bacterial homologs to *Pf*KMO (10 sequences).

```
h.sapiens KMO     ------------MDSSVIQRKKVAVIGGGLVGSLQACFLAKRNFQIDVYEAREDTRVATF 48
C.elegans KMO     ------------MPS-------VAIAGAGLVGALNACFFAQKGWDVSVYEFRKDIRTMKH 41
S.cerevisiae KMO  ------------MSES------VAIIGAGLVGCLAALAYSKEGYNVTLFDYRQDPRLDTT 42
PfKMO (All Bact)  -------------MTATDNARQVTIIGAGLAGTLVARLLARNGWQVNLFERRPDPRIETG 47
PfKMO (Close Hs)  -------------MTATDNARQVTIIGAGLAGTLVARLLARNGWQVNLFERRPDPRIETG 47

h.sapiens KMO     TRG--RSINLALSHRGRQALKAVG--LEDQIVSQGIPMRARMIHSLSGKK-SAIPYGTKS 103
C.elegans KMO     VQG--RSINLALSQRGKSALEAVG--LKEYIVNQGVPLYARLIHNKDGKTYSRQPYGKPG 97
S.cerevisiae KMO  KNKNLKSINLAISARGIDALKSIDPDACEHILQDMIPMKGRMIHDLKGRQ-ESQLYGLHG 101
PfKMO (All Bact)  ARG--RSINLALAERGAHALRLAG--LEREVLAEAVMMRGRMVH-VPGTPPNLQPYGRDD 102
PfKMO (Close Hs)  ARG--RSINLALAERGAHALRLAG--LEREVLAEAVMMRGRMVH-VPGTPPNLQPYGRDD 102

h.sapiens KMO     -QYILSVSRENLNKDLLTAAEKYPNVKMHFNHR----------LLKCNPEEGM-ITVLGS 151
C.elegans KMO     -EHIVSINRRHLNEVMITQAEKSPNVKFFFEHKVKNVDYDKKQLVVQCTSQPSKIPTFGN 156
S.cerevisiae KMO  -EAINSINRSVLNNSLLDELEKS-TTELKFGHK----------LVKIEWTDDKQICHFAI 149
PfKMO (All Bact)  SEVIWSINRDRLNRILLDGAEAA-GASIHFNLG----------LDSVDFARQRLTLSNVS 151
PfKMO (Close Hs)  SEVIWSINRDRLNRILLDGAEAA-GASIHFNLG----------LDSVDFARQRLTLSNVS 151

h.sapiens KMO     DKVPK-----DVTCDLIVGCDGAYSTVRSHLMKKPRFDYSQQYIPHGYMELTIPPKN--- 203
C.elegans KMO     KSPPQEHAEFHVEADLILACDGAYSAVRRSLMTIPRFDFSQEYIEHGYVELNIMANN--- 213
S.cerevisiae KMO  GEDLKTP--HTEKYDFVIGCDGAYSATRSQMQRKVEMDFSQEYMNLRYIELYIPPTEEFK 207
PfKMO (All Bact)  GERLEKR------FHLLIGADGCNSAVRQAMASVVDLGEHLETQPHGYKELQITPEAS-- 203
PfKMO (Close Hs)  GERLEKR------FHLLIGADGCNSAVRQAMASVVDLGEHLETQPHGYKELQITPEAS-- 203

h.sapiens KMO     ----GDYAMEPNYLHIWPRNTFMMIALPNMNKSFTCTLFM--------PFEEFEKLL-TS 250
C.elegans KMO     ----NEFAFEENVFHLWPRGHFTLIALANRDKTFTVTIFA--------PFSEFEKHMSTS 261
S.cerevisiae KMO  PNYGGNFAIAPDHLHIWPRHKFMLIALANSDGSFTSTFFG--------SKDQISDLITSK 259
PfKMO (All Bact)  ----AQFNLEPNALHIWPHGDYMCIALPNLDRSFTVTLFLHHQSPAAQPASPCFAQLVDG 259
PfKMO (Close Hs)  ----AQFNLEPNALHIWPHGDYMCIALPNLDRSFTVTLFLHHQSPAAQPASPCFAQLVDG 259

h.sapiens KMO     NDVVDFFQKYFPDAIPLIGEKLLVQDFFLLPAQPMISVKCSSFHFKS-HCVLLGDAAHAI 309
C.elegans KMO     EDVLSFFEENFPDAFLLLGKEHIADTFNRVKPQPLVSIKCSPHSFFD-NLVLMGDAAHAM 320
S.cerevisiae KMO  SRVREFLIENFPDIINIMDLDDAVKRFITYPKESLVCVNCKPYDVPGGKAILLGDAAHAM 319
PfKMO (All Bact)  HAARRFFQRQFPDLSPMLDS--LEQDFEHHPTGKLATLRLTTWHVGG-QAVLLGDAAHPM 316
PfKMO (Close Hs)  HAARRFFQRQFPDLSPMLDS--LEQDFEHHPTGKLATLRLTTWHVGG-QAVLLGDAAHPM 316

h.sapiens KMO     VPFFGQGMNAGFEDCLVFD-ELMDKFSNDLSLCLPVFSRLRIPDDHAISDLSMYNYIEMR 368
Caen.ele_KMO      VPFYGQGMNCGFEDCLVFS-ETLEEYGNDIAKAVKVYSDGRVNDAHSINDLAMYNYEELK 379
S.cerevisiae KMO  VPFYGQGMNCGFEDVRILM-ALLKKHSGDRSRAFTEYTQTRHKDLVSITELAKRNYKEMS 378
PfKMO (All Bact)  VPFHGQGMNCALEDAVALA-EHLQSAA-DNASALAAFTAQRQPDALAIQAMALENYVEMS 374
PfKMO (Close Hs)  VPFHGQGMNCALEDAVALA-EHLQSAA-DNASALAAFTAQRQPDALAIQAMALENYVEMS 374

h.sapiens KMO     AHVNSSWFIFQKNMERFLHAIMPSTFIPLYTMVTF-SRIRYHEAVQRWHWQKKVINKGLF 427
C.elegans KMO     DLVNKSSYKLRKKFDTIMNSIFPKSWIPLYSMVTF-SRIPYSEVIERRKRQDKILSR--- 435
S.cerevisiae KMO  HDVTSKRFLLRKKLDALFSIIMKDKWIPLYTMISFRSDISYSRALERAGKQTRILKF--- 435
PfKMO (All Bact)  SKVASPTYLLERELGQIMAQRQPTRFIPRYSMVTF-SRLPYAQAMARGQIQEQLLKFA-- 431
PfKMO (Close Hs)  SKVASPTYLLERELGQIMAQRQPTRFIPRYSMVTF-SRLPYAQAMARGQIQEQLLKFA-- 431

h.sapiens KMO     FLGSLIAISSTYLLIHYMSPRSFLRLRRPWNWIAHFRNTTCFPAKAVDSLEQISNLISR- 486
C.elegans KMO     ----IMTTTSTLALIGAAAG------------IYVNRGKLGL------------------ 461
S.cerevisiae KMO  -------LESLTLGMLSIGG--------------YKLFKFLTRERS-------------- 460
PfKMO (All Bact)  -----VANHSDLTSINLDA--------------VEHEVTRCLPPLSHLC----------- 461
PfKMO (Close Hs)  -----VANHSDLTSINLDA--------------VEHEVTRCLPPLSHLC----------- 461
```

1.4 Alignment to FAH sequences - secondary structure prediction

CLUSTAL W2 (2.0.10) multiple sequence alignment with secondary structure information labelled with β-sheets in yellow and α-helices in red. For PfKMO the data is based on secondary structure predictions as described in section **3.2.3**.




2qa1 -------TRIIVCERGTPPQ----RRETPPSWHEVADAWK-RLTGDDIAH---------A 256
1foh -------VRFYVQLQARAEKGGRVDRTKFTPEVVIANAKK-IFHPYTFDV---------Q 328
2x3n -------ARLVVSFPREEARELMADTRGESLRRRLQRFVGDESAEAIAAV---------T 264
3c96 -------GKSLVNWVCMVPSAAVGQLDNEADWNRDGRLEDVLPFFADWDLGWFDIRDLLT 281
2y6r -------ALHFGISFKTPDEWKNQTQVDFQNRNSVVDFLLKEFSDWDERYK-----ELIH 269
3gmc -------LYLGLMAPAADPRGSSVPIDLEVWVEMFPFLEPCLIEAAKLKT---------- 264
2vou ESPRLNFQWYWNVAEGPDLDELMTDVRGIRLPTSVHNNSLNPHNLRQFHSKGESLFKPFR 274
PfKMO ----SFTVTLFLHHQSPAAQPASPCFAQLVDGHAARRFFQRQFPDLSPMLD--------S 279
1dod ----------SRYYVQVPLSEKVEDWSDERFWTELKARLPSEVAEKLVTG---------- 258
3ihg ELVDIQGWEMAARIAERWREG----RVFLAGDAAKVTPPTGGMSGNAAVADGFDLAWKLA 336
2qa1 EPVWVSAFGNATRQVTEYRRG----RVILAGDSAHIHLPAGGQGMNTSIQDAVNLGWKLG 312
1foh QLDWFTAYHIGQRVTEKFSKDE---RVFIAGDACHTHSPKAGQGMNTSMMDTYNLGWKLG 385
2x3n GTSRFKGIPIGYLNLDRYWAD----NVAMLGDAIHNVHPITGQGMNLAIEDASALADALD 320
3c96 RNQLILQYPMVDRDPLPHWGRG---RITLLGDAAHLMYPMGANGASQAILDGIELAAALA 338
2y6r TTLSFVGLATRIFPLEKPWKSKRPLPITMIGDAAHIMPPFAGQGVNSGLVDALILSDNLA 329
3gmc -----ARYDKYETTKLDSWTRG---KVALVGDAAHAMCPALAQGAGCAMVNAFSLSQDLE 316
2vou DLVLNASSPFVTVVADATVDRMVHGRVLLIGDAAVTPRPHAAAGGAKASDDARTLAEVFT 334
PfKMO LEQDFEHHPTGKLATLRLTTWHVGGQAVLLGDAAHPMVPFHGQGMNCALEDAVALAEHLQ 339
1dod PSLEKSIAPLRSFVVEPMQHG----RLFLAGDAAHIVPPTGAKGLNLAASDVSTLYRLLL 314
3ihg AVLQG-QAGAGLLDTYEDERKVAAELVVAEA--LAIYAQRMAPHMAEVWDKSVGYPETLL 393
2qa1 AVVNG-TATEELLDSYHSERHAVGKRLLMNT--QAQGLLFLSGPEVQPLRDVLTELIQYG 369
1foh LVLTG-RAKRDILKTYEEERHAFAQALIDFD--HQFSRLFSGRPAKDVADEMGVSMDVFK 442
2x3n LALRDACALEDALAGYQAERFPVNQAIVSYG--HALATSLEDR------QRFAGVFDTAL 373
3c96 RNADV----AAALREYEEARRPTANKIILAN--RREKEEWAAASRPKTEKSAALEAITG 392
2y6r DGKFN--SIEEAVKNYEQQMFIYGKEAQEES--TQNEIEMFKPDFTFQQLLNV - 378
3gmc EGSSV----EDALVAWETRIRPITDRCQALSGDYAANRSLSKGNMFTPAALEAARYDPLR 372
2vou KNHDI----RGSLQSWETRQLQQGHAYLNKVK-KMASRLQHGGSFEPGNPAFAFGLPKVD 389
PfKMO SAADN----ASALAAFTAQRQPDALAIQAMALENYVEMSSKVASPTYLLERELGQIMAQR 395
1dod KAYRE--GRGELLERYSAICLRRIWKAERFSWWMTSVLHRFPDTDAFSQRIQQTELEYYL 372
3ihg GFRYRSSAVLATDDDPARVENPLTPSGRPGFRGPHVLVSRHGERLSTVDLFGDGWTLLAG 453
2qa1 EVARHLAGMVSGLEITYDVGTGSHP--LLGKRMPALELTTATRETSSTELLHTARGVLLD 427
1foh EAFVKGNEFASGTAINYDENLVTDKKSSKQELAKNCVVGTRFKSQPVVRHSEGLWMHFGD 502
2x3n QGSSRTPEALGGERSYQPVRSPAPLG - 399
3c96 SYRNQVERPRLE - 404
2y6r
3gmc RVYSWPQ - 379
2vou EPSVVTNS - 397
PfKMO QPTRFIPRYSMVTFSRLPYAQAMARGQIQEQLLKFAVANHSDLTSINLDAVEHEVTRCLP 455
1dod GSEAGLATIAENYVGLPYEEIE - 394
3ihg ELGAD----------WVAAAEAVSAELGVPVRAYRVGAGLTDPE 487
2qa1 LADNP----------RLRARAAAWSDR---------VDIVTAVP 452
1foh RLVTDGRFRIIVFAGKATDATQMSRIKKFSAYLDSENSVISLYTPKVSDRNSRIDVITIH 562
PfKMO PLSHLC - 461
3ihg SAVSERYGIGK----------AGASLVRPDG 508
2qa1 GEVSATSGLRD----------TTAVLIRPDG 473
1foh SCHRDDIEMHDFPAPALHPKWQYDFIYADCDSWHHPHPKSYQAWGVDETKGAVVVVRPDG 622
3ihg IVAWRTDEAAADAAQTLEGVLRRVLDR - 535
2qa1 HVAWAAPGSHHDLPMALERWFGAPLTG - 500
1foh YTSLVTDLEGTAEIDRYFSGILVEPKEKSGAQTEADWTKSTA - 664

Appendix 2 – Structural inferences for human KMO

2.1 Alignment of human KMO with PfKMO

An alignment of the amino acid sequences of human KMO and *Pf*KMO with residues within 5, 6 and 8 Å of the substrate in the dm2 structure highlighted. The alignment shows strong conservation of the residues in the vicinity of the substrate-binding pocket.

5A substrate
6A substrate
8A substrate

```
PfKMO           -MTATDNARQVTIIGAGLAGTLVARLLARNGWQVNLFERRPDPRIETGARGRSINLALAE 59
humanKMO        MDSSVIQRKKVAVIGGGLVGSLQACFLAKRNFQIDVYEAREDTRVATFTRGRSINLALSH 60
                  ::. : ::*::**.**.*:* * :**:..:*::::* * *.*: * :*********:.

PfKMO           RGAHALRLAGLEREVLAEAVMMRGRMVHVPGTPPNLQPYGRDDSEVIWSINRDRLNRILL 119
humanKMO        RGRQALKAVGLEDQIVSQGIPMRARMIHSLSGKKSAIPYG-TKSQYILSVSRENLNKDLL 119
                ** :**: .*** ::::::.: **.**:*  .   .  ***  .*: * *:.*:.**: **

PfKMO           DGAEAAG-ASIHFNLGLDSVDFARQRLTLSNVSGERLEKRFHLLIGADGCNSAVRQAMAS 178
humanKMO        TAAEKYPNVKMHFNHRLLKCNPEEGMITVLGSDKVPKDVTCDLIVGCDGAYSTVRSHLMK 179
                 .**    ..:***  * . :  .  :*: . .    :   .*::*.**. *:**. : .

PfKMO           VVDLGEHLETQPHGYKELQITPEASAQFNLEPNALHIWPHGDYMCIALPNLDRSFTVTLF 238
humanKMO        KPRFDYSQQYIPHGYMELTIPPKN-GDYAMEPNYLHIWPRNTFMMIALPNMNKSFTCTLF 238
                   :.   :  **** ** *.*:  .:: :*** *****:. :* *****:::*** ***

PfKMO           LHHQSPAAQPASPCFAQLVDGHAARRFFQRQFPDLSPMLDS--LEQDFEHHPTGKLATLR 296
humanKMO        MPFEE---------FEKLLTSNDVVDFFQKYFPDAIPLIGEKLLVQDFFLLPAQPMISVK 289
                : .:.         * :*: .: .  ***: ***  *::..  * ***   *:  : :::

PfKMO           LTTWHVGGQAVLLGDAAHPMVPFHGQGMNCALEDAVALAEHLQSAADNAS-ALAAFTAQR 355
humanKMO        CSSFHFKSHCVLLGDAAHAIVPFFGQGMNAGFEDCLVFDELMDKFSNDLSLCLPVFSRLR 349
                 :::*. .:.********.:***.*****..:**.:.: * ::. ::: * .*..*:  *

PfKMO           QPDALAIQAMALENYVEMSSKVASPTYLLERELGQIMAQRQPTRFIPRYSMVTFSRLPYA 415
humanKMO        IPDDHAISDLSMYNYIEMRAHVNSSWFIFQKNMERFLHAIMPSTFIPLYTMVTFSRIRYH 409
                 **  **. ::: **:** ::* *. ::::::: :::    *: *** *:******: *

PfKMO           QAMARGQIQEQLLKFAVANHSDLTSIN----------------------------LDAV 446
humanKMO        EAVQRWHWQKKVINKGLFFLGSLIAISSTYLLIHYMSPRSFLRLRRPWNWIAHFRNTTCF 469
                :*: * : *::::: .:   ..* :*.                              ..

PfKMO           EHEVTRCLPPLSHLC-- 461
humanKMO        PAKAVDSLEQISNLISR 486
                  :.. .*  :*:*
```

2.2 Homology model of human KMO tertiary structure

Tertiary protein structure models for human KMO (*yellow*) were generated with the i.model online server (IBI Biosolutions pvt ltd) using the: i) substrate-bound (*green*) and ii) substrate-free (*mauve*) dm2 *Pf*KMO structures as templates. The resulting models are shown in figures A2.1 and A2.2 prepared using PyMol. The predicted trans-membrane region is highlighted in red and caused problems in the models. It is likely that the trans-membrane region actually forms a helix that extends out away from the rest of the protein. It could then span the outer mitochondrial membrane with the remaining C-terminal forty residues on the other side of the membrane, as depicted in figure A2.3.

Figure A2.1 – human KMO model based on substrate-bound *Pf*KMO structure

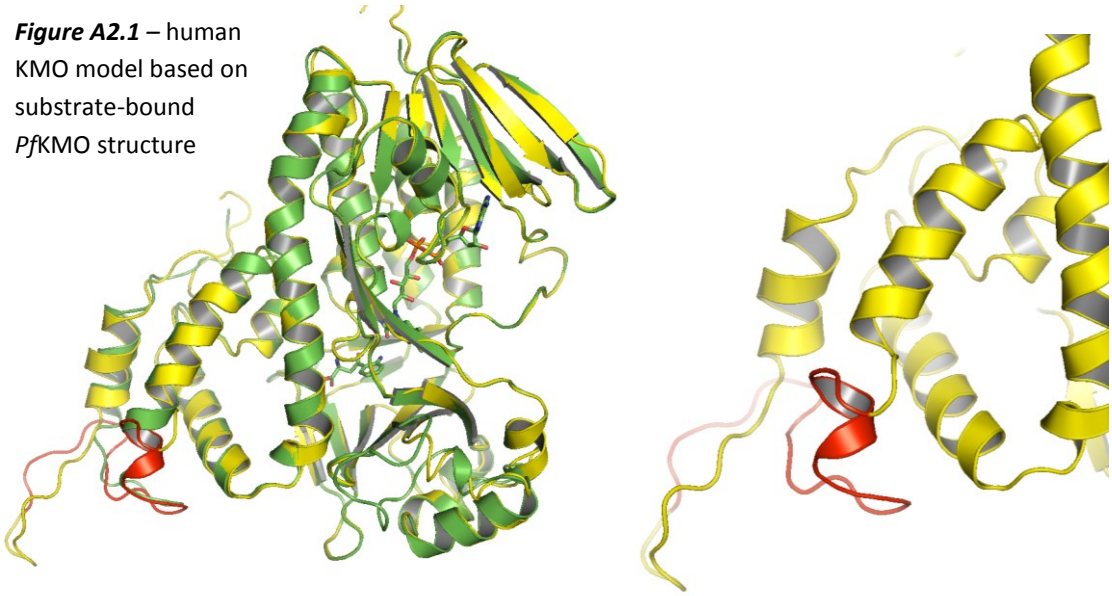


Figure A2.2 – human KMO model based on substrate-free *Pf*KMO structure

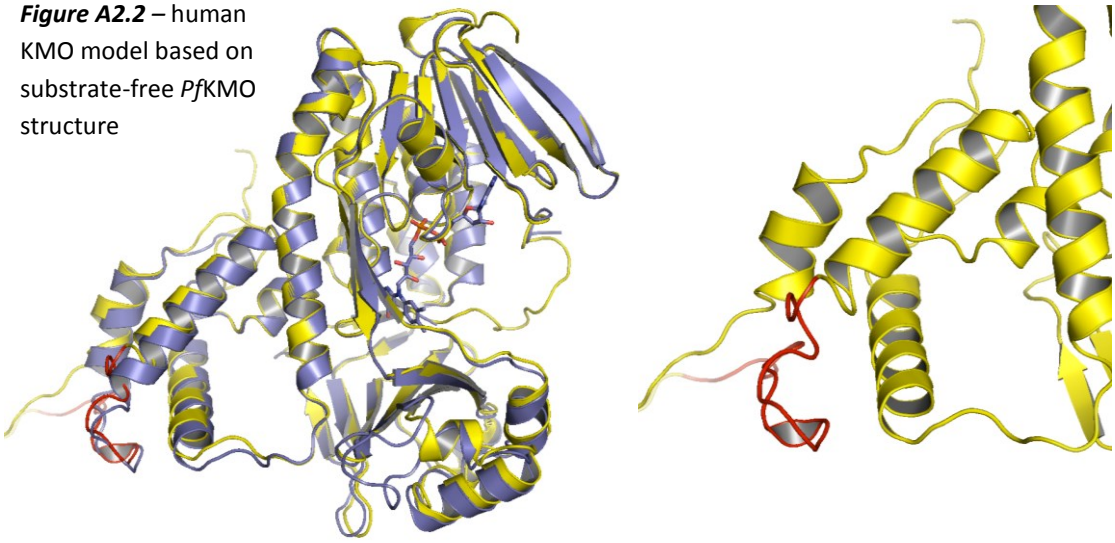
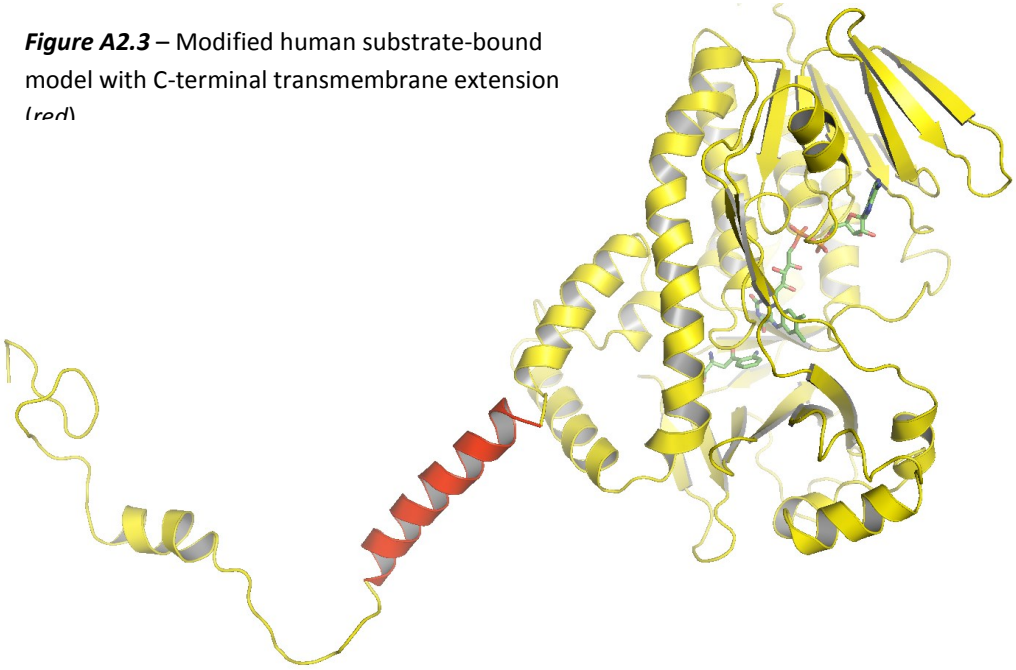
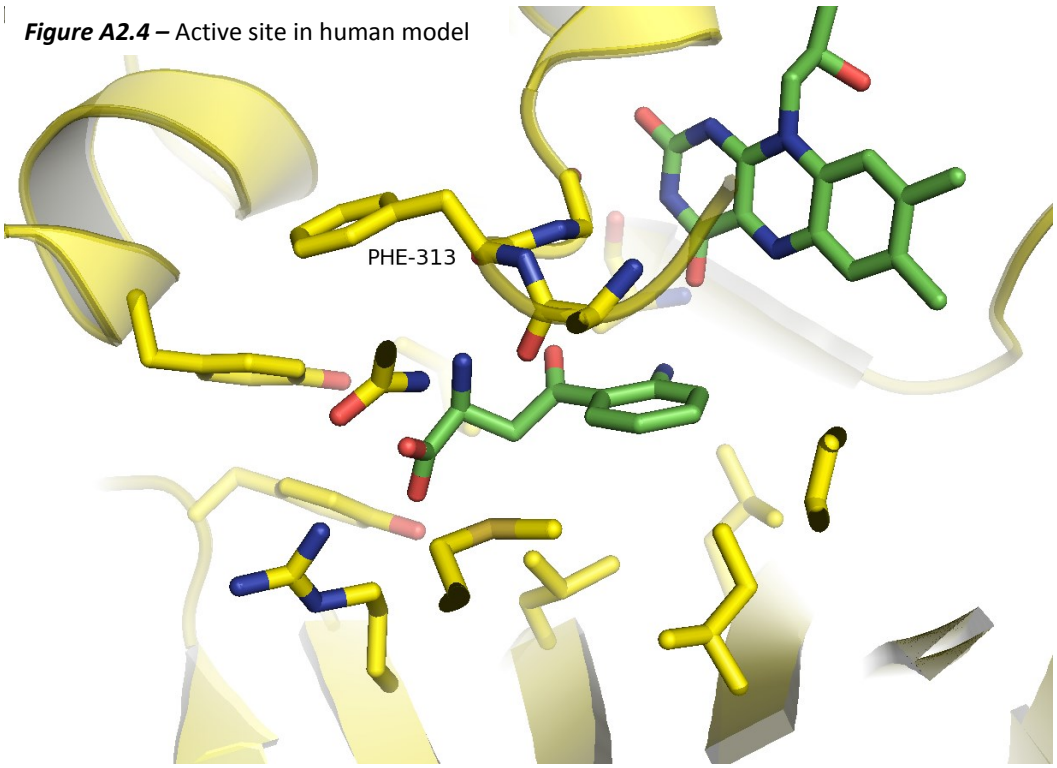


Figure A2.3 – Modified human substrate-bound model with C-terminal transmembrane extension (*red*)



The active site residues in the human model shown as sticks (see figure A2.4) are highly similar to the substrate-binding residues identified in the structures of *Pf*KMO. The only difference is the labelled residue Phe-313, which relates to His-320 of the bacterial enzyme. FAD and L-Kyn are modelled in as green sticks.

Figure A2.4 – Active site in human model



Appendix 3 – *Pf*KMO X-ray data tables

3.1 Selenomethionine-labelled substrate-bound dm2 PfKMO

Table A3.1 – Data collection and refinement statistics for single SeMet dm2 L-Kyn-bound crystal

	Peak	Inflection	HER	LER
Data collection				
Beamline	I04	I04	I04	I04
X-ray wavelength (Å)	0.9795	0.9797	0.9686	0.9873
Space group	$I4_122$	$I4_122$	$I4_122$	$I4_122$
Average cell dimensions				
$a = b$, c (Å)	149.9, 273.4	150.0, 273.4	149.9, 273.4	150.0, 273.5
$\alpha = \beta = \gamma$ (°)	90.0	90.0	90.0	90.0
Resolution (Å)	77.86-3.40 (3.58-3.40)*	77.86-3.40 (3.58-3.40)*	77.86-3.40 (3.58-3.40)*	77.86-3.40 (3.58-3.40)*
R_{merge}	0.140 (0.654)	0.137 (0.673)	0.137 (0.599)	0.114 (0.581)
$I / \sigma I$	8.9 (2.7)	10.3 (2.8)	9.7 (3.1)	11.7 (3.3)
Completeness (%)	99.9 (100.0)	98.8 (98.7)	99.7 (99.8)	98.8 (99.2)
Redundancy	7.5 (7.7)	8.8 (9.0)	8.7 (9.0)	8.8 (9.0)
Anomalous Redundancy	4.0 (4.1)	4.6 (4.6)	4.6 (4.6)	4.6 (4.6)
Refinement				
Resolution (Å)	77.86-3.40			
No. reflections	20381			
R_{work} / R_{free}	0.249/0.281			
No. atoms				
Protein				
Ligand/ion				
Water				
B-factors				
Protein				
Ligand/ion				
Water				
rmsd				
Bond lengths (Å)	0.005			
Bond angles (°)	0.936			

*Values in parentheses are for highest-resolution shell.

3.2 Substrate-free dm2 PfKMO

Table A3.2 – Data collection and refinement statistics for unbound and chloride-bound crystals

	Unbound	Chloride
Data collection		
Beamline	I24	I04
X-ray wavelength (Å)	0.9778	0.9797
Space group	$P2_122_1$	$P2_122_1$
Cell dimensions		
a, *b*, *c* (Å)	104.9, 134.2, 188.3	105.0, 133.9, 189.1
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution (Å)	75.59-2.26 (2.38-2.26)*	94.54-2.45 (2.58-2.45)
R_{merge}	0.055 (0.652)	0.063 (0.597)
$I / \sigma I$	13.5 (2.0)	12.4 (2.0)
Completeness (%)	99.4 (99.6)	97.8 (99.0)
Redundancy	3.8 (3.7)	3.8 (3.6)
Refinement		
Resolution (Å)	50.00-2.26	35.00-2.45
No. reflections	117656	91114
R_{work} / R_{free}	0.219/0.232	0.218/0.244
No. atoms		
Protein		
Ligand/ion		
Water		
B-factors		
Protein		
Ligand/ion		
Water		
rmsd		
Bond lengths (Å)	0.005	0.007
Bond angles (°)	0.940	1.113

*Values in parentheses are for highest-resolution shell.

3.3 Substrate-bound C-terminal PfKMO mutants

Table A3.3 – Data collection and refinement statistics for substrate-bound mutants.

	Y382F dm2	R386K dm2	R386T dm2	E372T [(†)] dm2
Data collection				
Beamline	I04-1	I02	I03	I03
X-ray wavelength (Å)	0.9200	0.9790	0.9763	0.9763
Space group	$I4_122$	$I4_122$	$I4_122$	$I4_122$
Cell dimensions				
$a = b$, c (Å)	150.2, 273.6	149.8, 273.5	149.7, 272.1	149.6, 270.7
$\alpha = \beta = \gamma$ (°)	90.0	90.0	90.0	90.0
Resolution (Å)	68.36-3.36 (3.54-3.36)*	53.96-3.30 (3.50-3.30)	65.65-3.40 (3.58-3.40)	44.66-3.24 (3.42-3.24)
R_{merge}	0.095 (0.599)	0.070 (0.545)	0.088 (0.644)	0.088 (0.691)
$I / \sigma I$	14.4 (3.6)	20.1 (4.9)	9.9 (2.3)	14.2 (2.8)
Completeness (%)	99.7 (99.8)	100.0 (100.0)	99.9 (99.9)	99.9 (100.0)
Redundancy	9.0 (9.1)	14.3 (14.8)	5.7 (5.7)	8.5 (8.8)
Refinement				
Resolution (Å)	65.93-3.36	53.96-3.30	65.68-3.40	44.70-3.24
No. reflections	21461	22568	20512	23527
R_{work} / R_{free}	0.247/0.277	0.247/0.259	0.260/0.270	0.230/0.256
No. atoms				
Protein				
Ligand/ion				
Water				
B-factors				
Protein				
Ligand/ion				
Water				
rmsd				
Bond lengths (Å)	0.005	0.005	0.005	0.005
Bond angles (°)	0.900	0.953	0.981	0.966

*Values in parentheses are for highest-resolution shell.
[(†)] E372T dm2 mutant contained an additional L367R mutation.

3.4 Substrate-free C-terminal PfKMO mutants

Table A3.4 – Data collection and refinement statistics for substrate-free mutant crystals.

	Unbound R386K dm2	Cl-bound R386T dm2	Cl-bound E372T (†) dm2
Data collection			
Beamline	I04	I03	I03
X-ray wavelength (Å)	0.9800	0.9763	0.9763
Space group	$P2_122_1$	$P2_122_1$	$P2_122_1$
Cell dimensions			
a, *b*, *c* (Å)	105.2, 133.9, 188.5	105.4, 134.0, 189.5	104.7, 133.9, 189.1
$\alpha = \beta = \gamma$ (°)	90.0	90.0	90.0
Resolution (Å)	45.00-2.63 (2.77-2.63)	45.00-2.52 (2.66-2.52)	50.51-2.32 (2.45-2.32)
R_{merge}	0.115 (0.619)	0.106 (0.599)	0.072 (0.599)
$I / \sigma I$	7.8 (2.0)	8.9 (2.6)	12.1 (2.5)
Completeness (%)	98.3 (99.7)	97.6 (99.6)	99.8 (99.8)
Redundancy	3.6 (3.6)	5.3 (5.1)	4.8 (4.9)
Refinement			
Resolution (Å)	45.00-2.63	44.71-2.52	50.51-2.32
No. reflections	74239	84072	109394
R_{work} / R_{free}	0.225/0.241	0.223/0.241	0.221/0.242
No. atoms			
Protein			
Ligand/ion			
Water			
B-factors			
Protein			
Ligand/ion			
Water			
rmsd			
Bond lengths (Å)	0.005	0.005	0.005
Bond angles (°)	0.953	0.977	0.990

*Values in parentheses are for highest-resolution shell.
(†) E372T dm2 mutant contained an additional L367R mutation.

3.5 Substrate-analogue bound dm2 PfKMO

Table A3.5 – Data collection and refinement statistics for substrate-analogue bound structures.

	m-NBA bound dm2 *Pf*KMO	3,4-CBA bound dm2 *Pf*KMO
Data collection		
Beamline	I02	I24
X-ray wavelength (Å)	0.9790	0.9778
Space group	$I4_122$	$I4_122$
Cell dimensions		
$a = b$, c (Å)	149.7, 273.0	149.4, 273.8
$\alpha = \beta = \gamma$ (°)	90.0	90.0
Resolution (Å)	45.00-3.55 (3.74-3.55)*	35.00-3.37 (3.55-3.37)
R_{merge}	0.096 (0.799)	0.099 (0.651)
$I / \sigma I$	7.9 (1.9)	11.2 (2.8)
Completeness (%)	96.1 (97.9)	96.4 (97.3)
Redundancy	4.5 (4.6)	8.0 (8.1)
Refinement		
Resolution (Å)	44.76-3.55	34.56-3.37
No. reflections	17259	20256
R_{work} / R_{free}	0.247/0.254	0.237/0.239
No. atoms		
Protein		
Ligand/ion		
Water		
B-factors		
Protein		
Ligand/ion		
Water		
rmsd		
Bond lengths (Å)	0.005	0.004
Bond angles (°)	0.919	0.882

*Values in parentheses are for highest-resolution shell.

3.6 Inhibitor-bound dm2 PfKMO

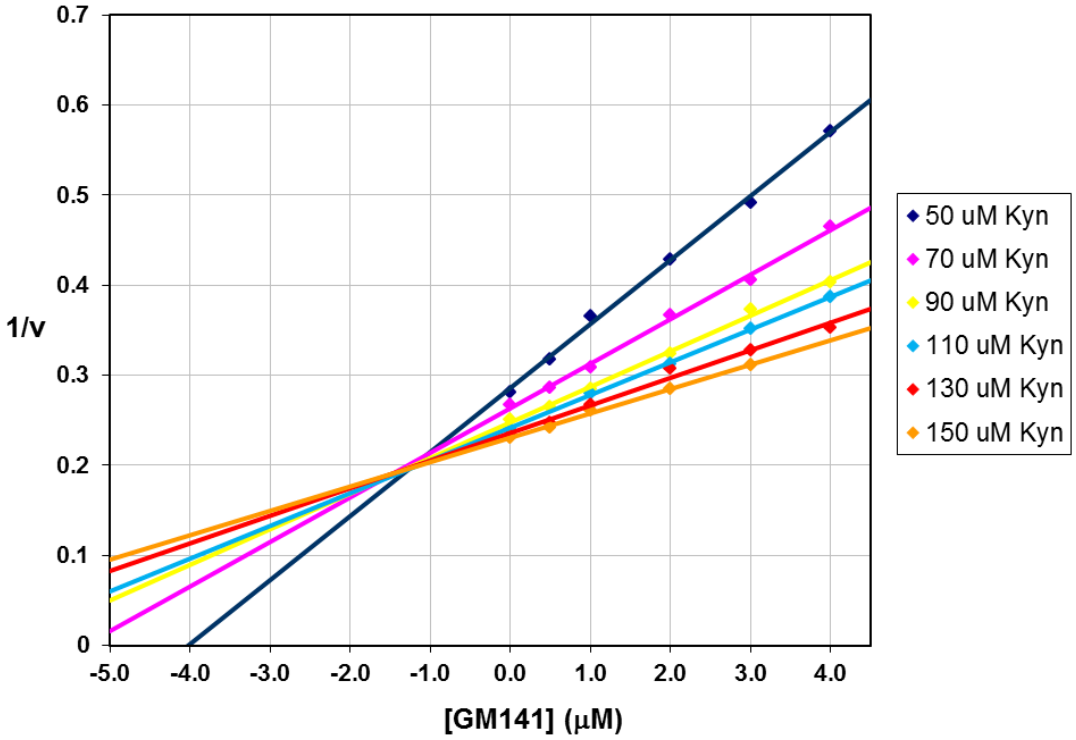
Table A3.6 – Data collection and refinement statistics for inhibitor bound structures.

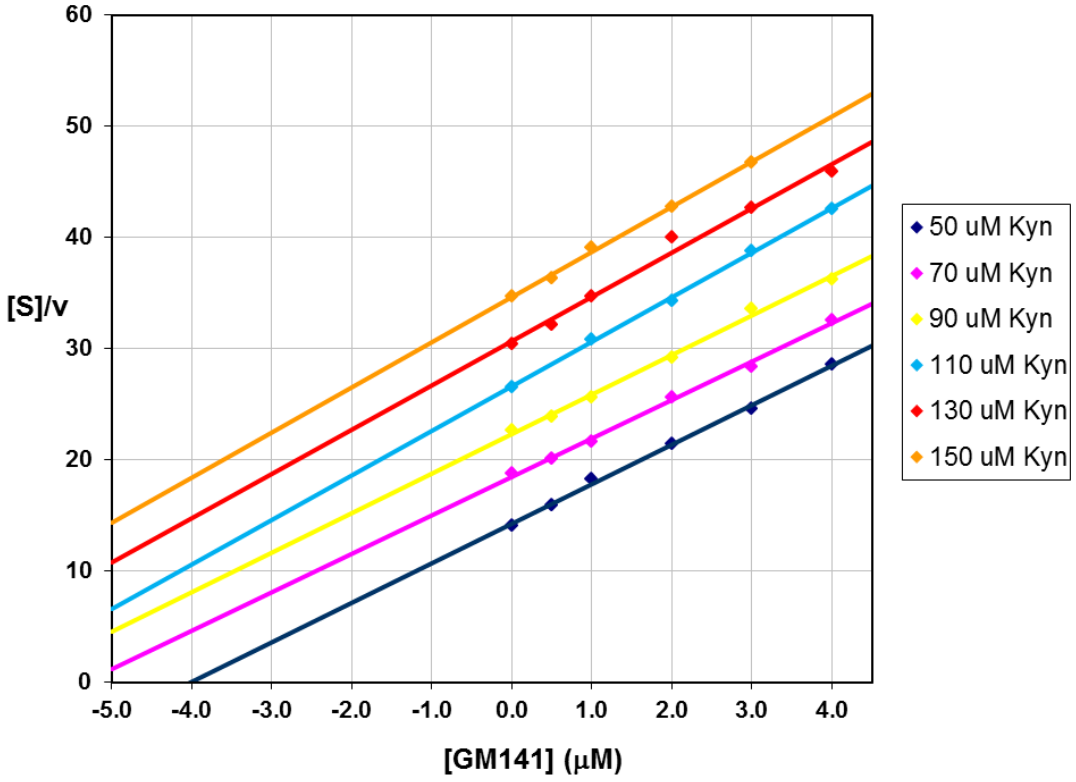
	GM303-bound dm2 *Pf*KMO	GM347-bound dm2 *Pf*KMO	GM760-bound dm2 *Pf*KMO	GM769-bound dm2 *Pf*KMO
Data collection				
Beamline	I03	I04-1	I02	I02
X-ray wavelength (Å)	0.9763	0.9200	0.9790	0.9790
Space group	$P6_122$	$P6_122$	$P6_122$	$P6_122$
Cell dimensions				
$a = b$, c (Å)	117.2, 418.2	117.7, 420.1	117.3, 418.1	117.5, 417.8
$\alpha = \beta, \gamma$ (°)	90.0, 120.0	90.0, 120.0	90.0, 120.0	90.0, 120.0
Resolution (Å)	46.47-2.18 (2.30-2.18)*	59.98-2.15 (2.27-2.15)	58.12-2.40 (2.52-2.40)	39.03-2.30 (2.42-2.30)
R_{merge}	0.136 (0.547)	0.137 (0.549)	0.156 (0.682)	0.157 (0.675)
$I / \sigma I$	12.7 (4.2)	11.0 (2.7)	11.0 (2.6)	12.5 (3.2)
Completeness (%)	100.0 (99.9)	99.9 (99.2)	100.0 (99.9)	99.9 (99.6)
Redundancy	10.7 (9.8)	10.1 (5.2)	11.1 (8.0)	14.9 (10.5)
Refinement				
Resolution (Å)	46.52-2.18	45.00-2.15	58.12-2.40	36.00-2.30
No. reflections	85024	89610	64181	72900
R_{work} / R_{free}	0.197/0.238	0.208/0.234	0.203/0.248	0.202/0.229
No. atoms				
Protein				
Ligand/ion				
Water				
B-factors				
Protein				
Ligand/ion				
Water				
rmsd				
Bond lengths (Å)	0.005	0.006	0.005	0.005
Bond angles (°)	1.096	1.117	1.081	1.052

*Values in parentheses are for highest-resolution shell.
(†) E372T dm2 mutant contained an additional L367R mutation.

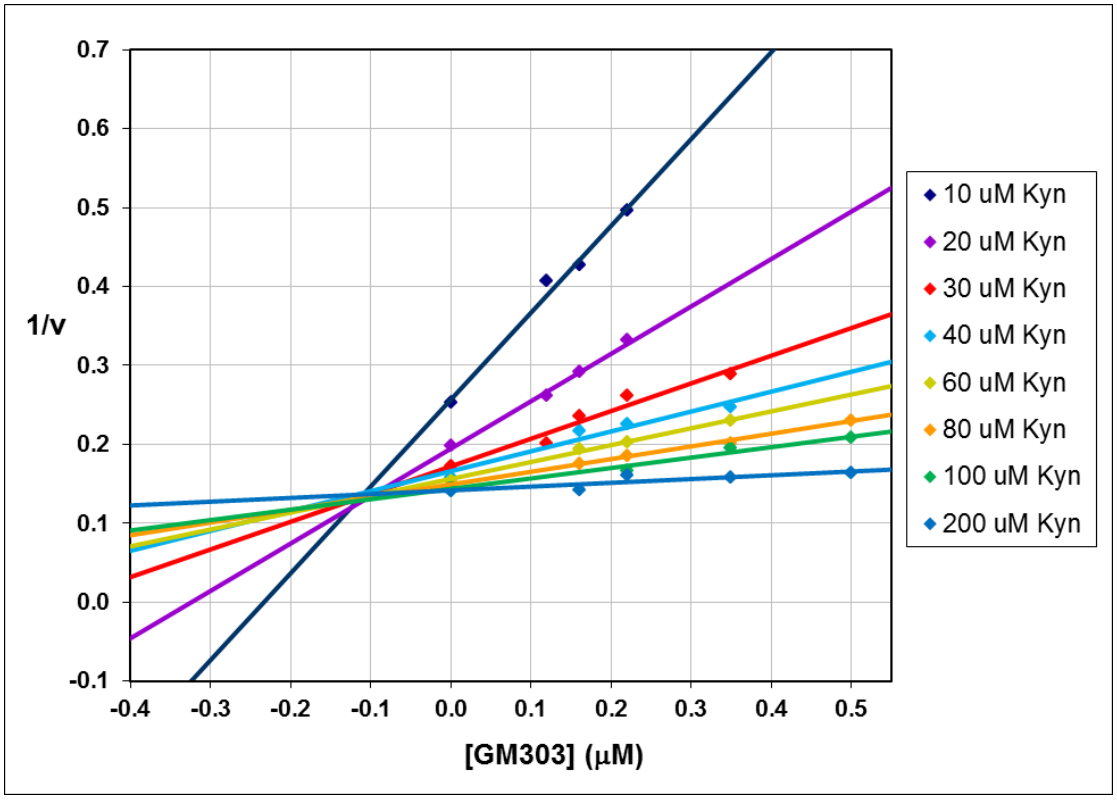
Appendix 4 – *Pf*KMO Inhibition data

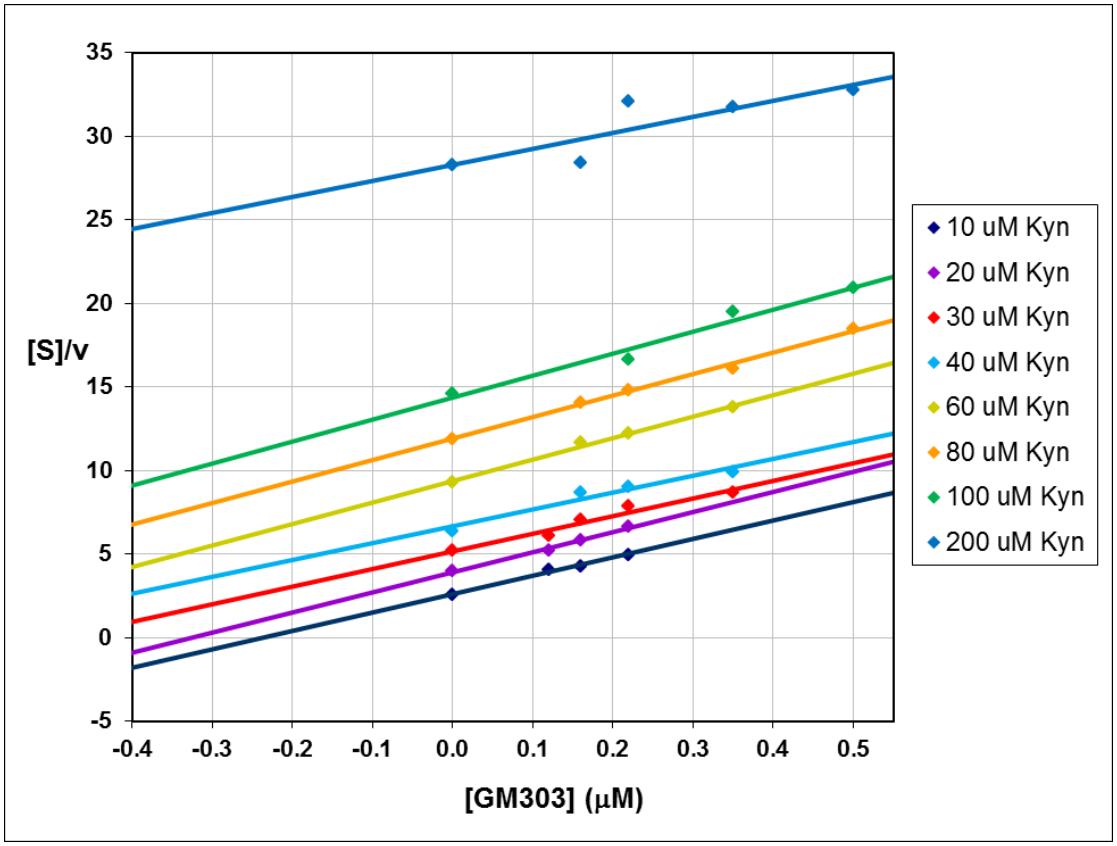
4.1 Inhibition plots for GM141



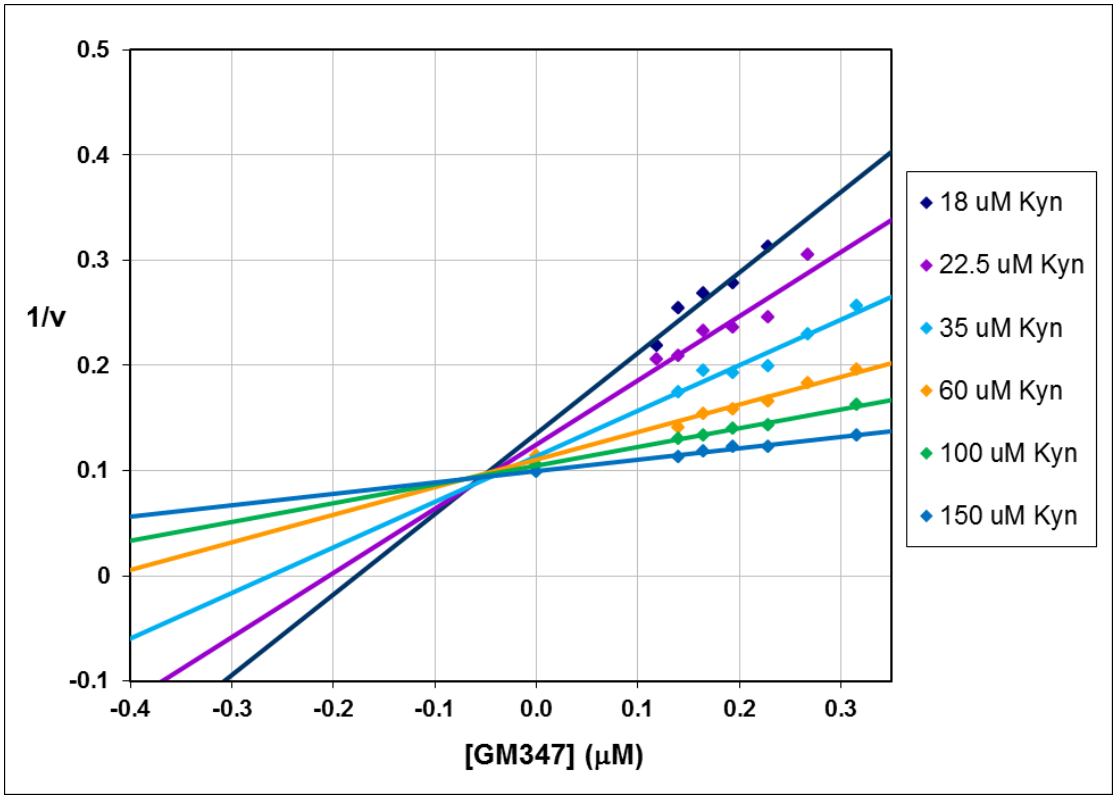


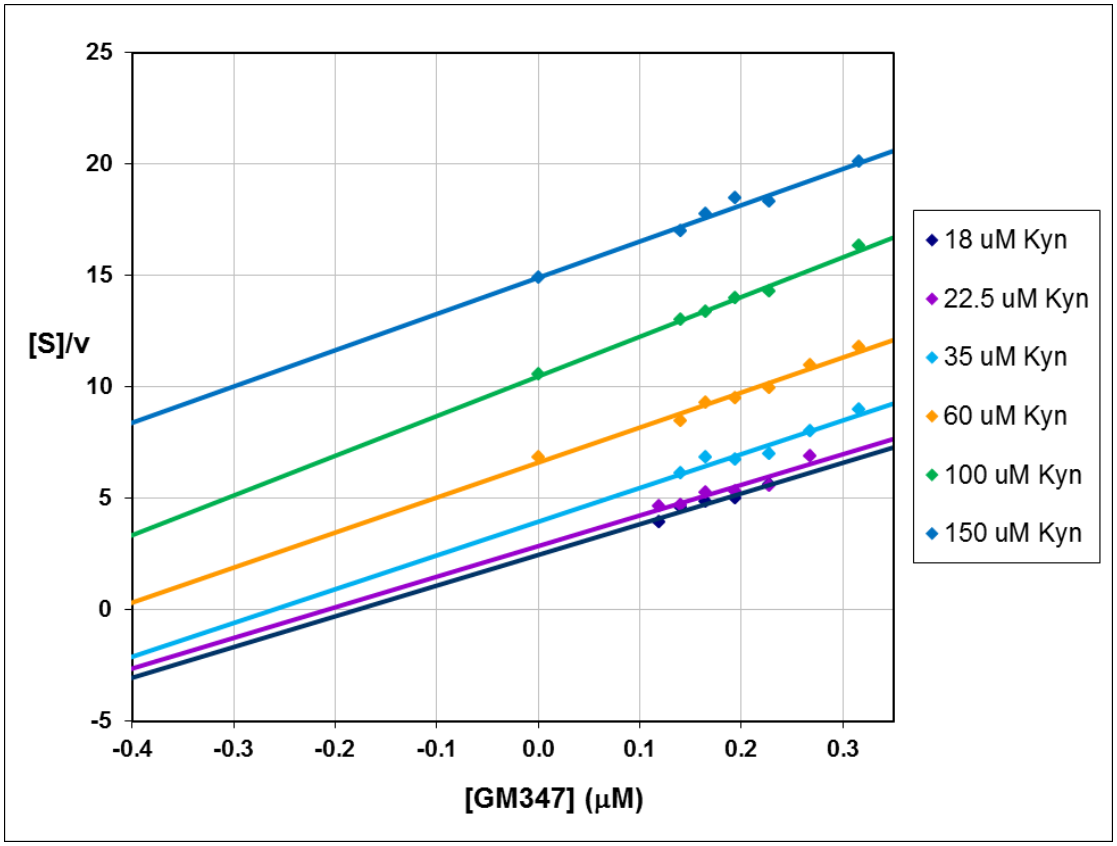
4.2 Inhibition plots for GM303



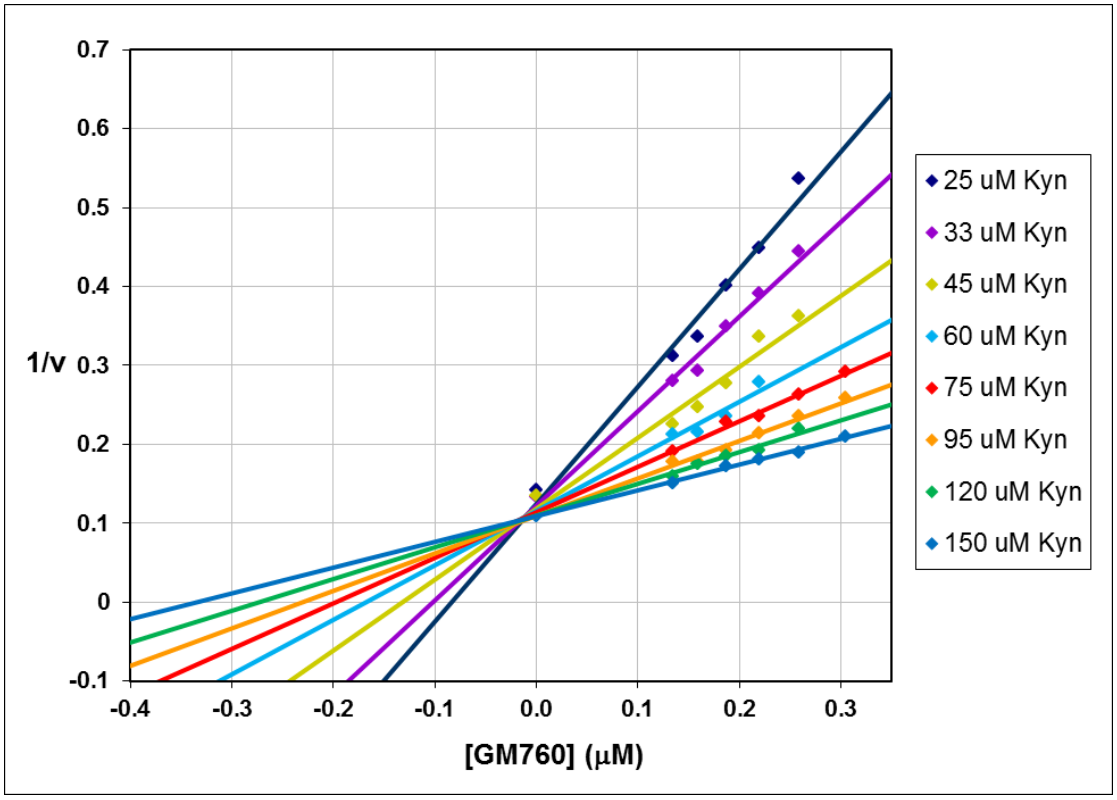


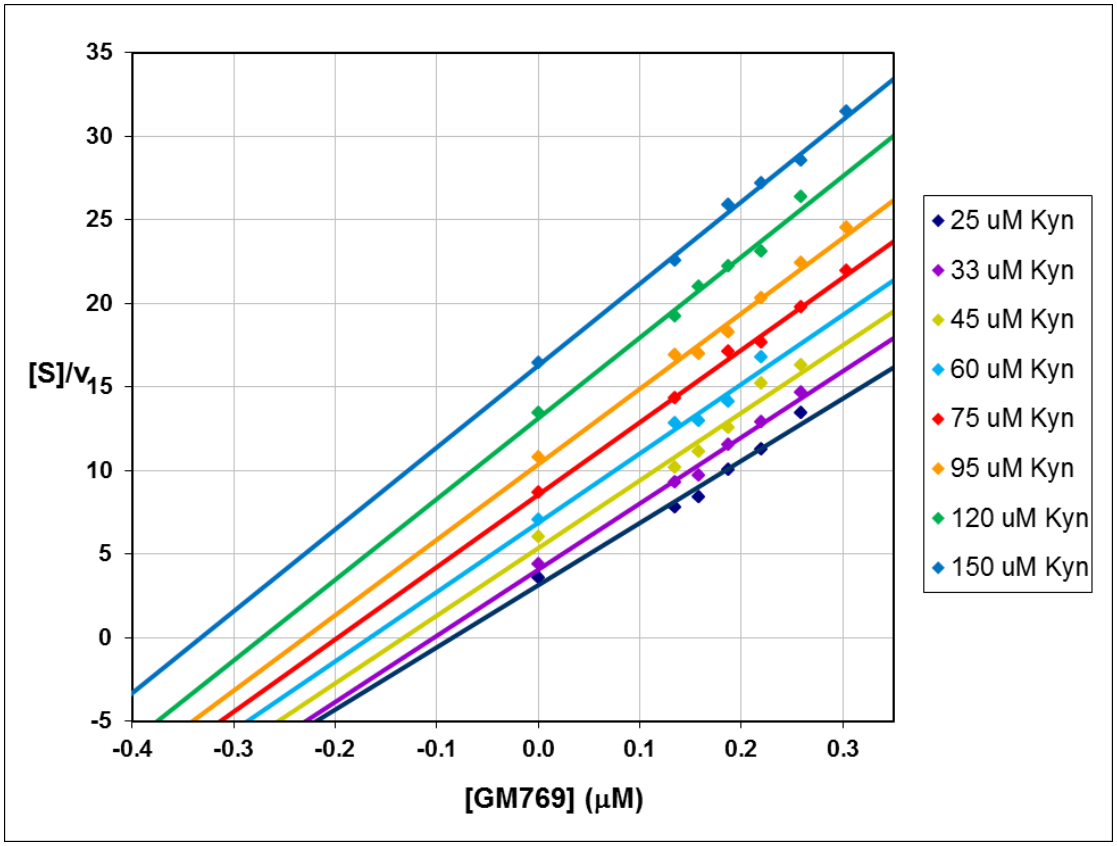
4.3 Inhibition plots for GM347





4.4 Inhibition plots for GM760





4.5 Inhibition plots for GM769