Synthesis of FMN Analogues as Probes for the Photocycle of *Avena sativa*

Andrew Wood

Submitted in fulfilment of the requirement for the degree of Doctor of Philosophy

School of Chemistry,

Cardiff University

February, 2014

Declaration

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed …… (candidate) Date ……

STATEMENT 1

This thesis is being submitted in partial fulfilment of the requirements for the degree of PhD.

Signed …… (candidate) Date ……

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated (see below).

Other sources are acknowledged by explicit references. The views expressed are my own.

Signed …… (candidate) Date ……

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed …… (candidate) Date ……

Declaration of Contributions

I confirm that the work presented within this thesis was performed by the author, with exceptions summarised here, and explicitly referenced in the relevant sections of the thesis.

The original construction of several plasmid vectors used for the expression of recombinant proteins was not performed, as they were freely available (from previous work), and donated to me. Their original preparation is summarised below. Furthermore, during a short side-project the dephosphorylation of 5-deazaFAD was performed (on an analytical scale) by a collaborator (Dr. Hannah Collins, University of Kent) using commercially available *Phosphodiesterase I* from Crotalus atrox. This is described in section 7.1.5 in order to support other results independently obtained during this research, and full reference is made to the contributing experimenter.

C. ammoniagenes RfK-FADS (pMAL-C2-RibF-Ca)

The plasmid vector pMAL-C2 containing the gene for Ca-RfK-FADS was originally prepared by others including the project supervisor (Gerald Richter, University of Cardiff), and its sequence and details are reported in a freely available journal[1].

S. pombe RfK (pET14b-*Sp*RfK):

The original cloning of the *S. pombe* riboflavin kinase from the commercially available gene (GenBank®[2,3] reference: NM_001023386.2) into the plasmid expression vector pET-14b was performed by Dr. Hannah Collins (University of Kent), and the construct was gifted to me during a short side-project.

A.sativa PHOT1-LOV2 (pNCO113-HISACT(C49S)-AsLOV2)

The creation of the plasmid pNCO113-HISACT(C49S)-AsLOV2(WT) was performed by others including the project supervisor (Gerald Richter); this is described in freely available literature[1,4], and its sequence added to the GenBank®[2,3] (reference: ef493211.1). This construct allowed expression of the wild-type AsLOV2 domain as a hisactophilin fusion protein, and was also used as the basis for subsequent mutagenesis of the AsLOV2 domain.

Abstract

The effect of light on plants is fundamental to our understanding of the biological world. Photo-activated processes within botanical cells are controlled by two distinct colours; red light provides the energy required for photosynthesis, and blue light stimulates responses of the cell (growth, stomatal opening, etc.) necessary for biological progression. The key step in blue-light controlled processes has recently been determined to be the formation of a covalent bond between a tightly-bound molecule of flavin and a conserved cysteine residue within a photosensitive protein domain, which causes a physical change to the protein structure upon irradiation.

This research was undertaken to examine light-stimulated covalent bond formation in phototropic proteins using two approaches. Firstly, two analogues of riboflavin (5-deazariboflavin and 1-deazariboflavin) were chemically synthesised, converted to the appropriate cofactor form (flavin mononucleotide; FMN), and incorporated into the photosensitive domain. The results of this were startling; while 1-deazaFMN inactivated the photosensitive domain to all wavelengths of light, 5-deazaFMN allowed control of the domain's behaviour using two separate wavelengths of light, creating a biophotonic nanoswitch.

The second approach was based upon mutagenesis of the amino acid responsible for bond formation, introducing several alternative residues in place of the reactive cysteine. All mutants were found to be inactive to irradiation (including those incorporating either of the deazaFMN analogues), providing further evidence to support the current "radical pair" mechanism of photoadduct formation.

During the synthesis of the flavin analogues, several reactions were found to be inefficient, and a large number of alternative reactions were studied. Additionally, a new route to several flavin analogues was devised and tested, and while unsuccessful, provides a foundation upon which to base a future examination of their synthesis. Finally, the synthesis of a novel riboflavin analogue (5-deaza-8-demethyl-riboflavin) was completed, allowing comparison with the synthesis of 5-deazariboflavin.

Acknowledgements

Firstly, I wish to thank my supervisor, Professor Gerald Richter, for his introduction to flavin chemistry. His knowledge in this field and his confidence in my abilities was fundamental in my decision to attempt the research presented in this thesis.

I would like to thank Dr. James Redman and Professor Rudolf Allemann for allowing me space to work within their laboratories, and for their thought-provoking questioning throughout my research. I would like to thank Professor Nick Tomkinson (University of Strathclyde), and Professor Mark Bagley (University of Sussex) for mentoring me during their readerships in Cardiff, as well as the technical and administrative staff (past and present) in the School of Chemistry for their support. In particular I would like to thank Dr. Robert Jenkins and Mr. Robin Hicks for their assistance with all aspects of NMR Spectroscopy and Mass Spectrometry.

I wish to thank Dr. Robert Mart, and a fellow PhD student whom I worked alongside, Dr. Steven Hill, for their assistance in the development of ideas and their numerous suggestions. Their support led to a vast improvement in my knowledge of chemical synthesis, and greatly improved the scope of this work. I also wish to extend thanks to all members of the Chemical Biology group in the School of Chemistry for their immeasurable help with the biological elements of this work. I began this project with no prior practical experience of this field, and have benefitted greatly by their patient instruction in the techniques and methods used for the molecular biology performed during my research.

Furthermore, I would like to thank my friends (Seni Chanapai, William Dawson, Sarah Adams, Daniel Grundy and Dilruba Meah) for their assistance and reassurance, and my partner Sally Pemberton, for her continued encouragement.

Finally, I must give thanks to my parents and friends for their love, encouragement and support throughout the course of my education, which is truly appreciated. Their contributions have been fundamental in my development both as a researcher and as an individual, and without which it would have been impossible to complete this project.

List of Abbreviations

In addition to abbreviations in common occurrence throughout current literature, the following abbreviations are in use throughout this thesis. These are divided for convenience into abbreviations used for chemicals, and for techniques and methods.

Chemical Abbreviations

18-C-6	18-crown-6
4-DMAP	4-dimethylaminopyridine, also known as *N,N*-dimethylaminopyridine
Ac	acetyl; functional group
AcCl	acetyl chloride
Ac_2O	acetic anhydride
AcOH	acetic acid
ADP	adenosine-5'-diphosphate
AMP	adenosine-5'-monophosphate
APS	ammonium persulphate
Arg	arginine (R); amino acid
Ar	aryl; functional group
Asp	aspartic acid (D); amino acid
ATP	adenosine-5'-triphosphate
Bn	benzyl; functional group
Boc	*tert*-butyloxycarbonyl; functional / protecting group
Boc anhydride	di-*tert*-butyl dicarbonate
CAM	cerium ammonium molybdate
CAN	cerium ammonium nitrate
Cys	cysteine (C); amino acid
CTP	cytidine-5'-triphosphate
DCM	dichloromethane
dioxane	1,4-dioxane
DABCO	diazabicyclo[2.2.2]octane
dATP	deoxyadenosine triphosphate, see dNTP

dCTP	deoxycytidine triphosphate, see dNTP
dGTP	deoxyguanosine triphosphate, see dNTP
DIPEA	*N,N*-diisopropylethylamine
DMAP	*N,N*-dimethylaminopyridine
DMF	*N,N*-dimethylformamide
DMPU	1,3-dimethyl-3,4,5,6-tetrahydro-2(1H)-pyrimidinone
DMSO	dimethylsulphoxide
DNA	deoxyribonucleic acid
DNPH	2,4-dinitrophenylhydrazine
dNTP	deoxynucleotide triphosphate (ATP, GTP, CTP, TTP)
DTT	dithiothreitol
dTTP	deoxythymidine triphosphate, see dNTPs
EtBr*	ethidium Bromide (***not*** bromoethane; by convention)
EDTA	ethylenediaminetetraacetic acid
Et_2O	diethyl ether
EtOAc	ethyl acetate
EtOH	ethanol
FAD	flavin adenine dinucleotide
FMN	flavin mononucleotide
GTP	guanosine-5'-triphosphate
GuHCl	guanidinium hydrochloride
GuSCN	guanidinium thiocyanate
HATU	1-[Bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate
HCl	hydrochloric acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IPTG	isopropyl-1-thio-β-D-galactopyranoside
KOH	potassium hydroxide
LB (growth media)	Leuria-Bertani media for bacterial growth
Me	methyl; functional group
MeOH	methanol
MES	2-(*N*-morpholino)ethanesulfonic acid

MOPS	3-(N-morpholino)propanesulfonic acid
NAD^+	nicotinamide adenine dinucleotide
$NADP^+$	nicotinamide adenine dinucleotide phosphate
NaOH	sodium hydroxide
PMSF	phenylmethylsulfonyl fluoride
Ph	phenyl; functional group
PIPES	piperazine-N,N'-bis(2-ethanesulfonic acid)
PLP	pyridoxal 5'-phosphate
$POCl_3$	phosphorous $^{(V)}$ oxychloride
Pyr	pyridine
Rf	riboflavin
riboaniline	3,4-dimethyl-*N*-(ribityl)-aniline
SDS	sodium dodecyl sulphate
TAE (buffer)	tris base, acetic acid and EDTA buffer
TB (growth media)	"terrific broth" media for bacterial growth
TCP	2,4,6-trichloropyrimidine
TCPyr	2,4,6-trichlropyridine
TEA (chemical synthesis)	triethylamine
TEA (buffer)	triethanolamine buffer
TEMED	tetramethylethylenediamine
TFA	trifluoroacetic acid
THF	tetrahydrofuran
Tris (buffer)	tris(hydroxymethyl)aminomethane
TTP	thymidine triphosphate
Vitamin B_2	riboflavin
Vitamin B_{12}	cobalamin

Non-Chemical Abbreviations & Glossary of terms

Å	Angstrom
Alkyl-DOB	alkylimino-de-oxo-bisubstitution
APSCI	Atmospheric Pressure Chemical Ionisation; see Mass Spectrometry
aq.	aqueous
*As*LOV2	recombinant PHOT1-LOV2 domain of *Avena sativa*
atm	atmosphere
b.p.	base pair
Ca-RfK-FADS	recombinant bifunctional RfK-FADS protein from *Corynebacterium ammoniagenes*
CI	Chemical Ionisation; see Mass Spectrometry
CD	Circular Dichroism
c. v.	column volumes
Da	Dalton; unit of mass equal to 1.0 g mol^{-1}
DFT	Density Functional Theory
EI	Electron Impact ionisation; see Mass Spectrometry
ENDOR	Electron Nuclear Double Resonance spectroscopy
EPR	Electron Paramagnetic Resonance spectroscopy
ESI	Electrospray Ionisation; see Mass Spectrometry
FADS	FAD Synthetase; enzymatic domain
FMN-AT	FMN Adenyl Transferase; enzymatic domain
FPLC	Fast Protein Liquid Chromatography
FT-IR	Fourier Transform Infrared spectroscopy
HPLC	High-Performance Liquid Chromatography
IMAC	Immobilised Metal Affinity Chromatography
IR	Infrared spectroscopy, see FT-IR
ISC	Inter-system crossing
k.b.p	kilobase pair
kDa	kilo Dalton; unit of mass equal to 1000 Daltons (1 kg mol^{-1})
LOV	Light-Oxygen-Voltage domain

MB	maltose-binding domain
MCS	Multiple Cloning Site
MS	Mass Spectrometry
m- / meta-	aromatic substituent position
MWt	Molecular Weight, used for protein / DNA markers
MRE	Mean Residue Ellipticity, for normalisation of CD data
MRW	Mean Residue Weight, for calculation of MRE
NMR	Nuclear Magnetic Resonance spectroscopy
NT	Nucleotidyl Transferase; enzymatic domain
o- / ortho-	aromatic substituent position
OD	Optical Density
OOF	Out Of Frame; codon frameshift of DNA
ORF	Open Reading Frame
p- / para-	aromatic substituent position
PCR	Polymerase Chain Reaction
PHOT	phototropin domain
ppm	parts per million
ppb	parts per billion
QM	Quantum Mechanical
RF	Retention Factor; used with TLC
RfK	Riboflavin Kinase; enzymatic domain
RNA	ribonucleic acid
RNAP	RNA polymerase
RPM	Rotations per Minute
RT	Room Temperature OR Retention Time
SDM	Site Directed Mutagenesis
SDS-PAGE	SDS-denaturing Polyacrylamide Gel Electrophoresis
Sp-RfK	recombinant RfK domain from *Schizosaccharomyces pombe*
t-, tert-	structural feature, e.g. *t*-butyl
TLC	Thin Layer Chromatography
UV-Vis	Ultra Violet-visible spectroscopy
w.r.t	with respect to

Contents

Chapter 1

Introduction

Chapter 1 - Introduction

1.0.1 Synopsis

In this chapter, an introduction to the topics discussed within this thesis is made, comprising the photocycle of a flavin-mediated plant phototropin, and the chemical synthesis of riboflavin analogues. A review of the current literature in these fields is presented, in order to give a foundation for the results presented in subsequent chapters of this thesis. Due to the separate strands of research performed, linked by flavin and the analogues 5-deazaflavin and 1-deazaflavin, this chapter has been divided into two sections to give a clear introduction to each topic.

1.0.2 Motivation

The study of phototropism is of significant interest to current research in the field of biophotonics, with potential applications throughout molecular biology. Examination of the mechanism and photo-physical effects of phototropism has previously been based largely on spectroscopic examination and protein mutagenesis studies. Instead, an alternative approach to understanding the mechanism of phototropism using artificial cofactors has the potential to give further insight to this biological processes.

The synthesis of riboflavin analogues has been a topic of some contention in the literature for several decades. No improvement to the synthesis of 5-deazariboflavin has been reported since the seminal work of Carlson and Kiessling[5] (2004); the revisiting of this work is somewhat overdue, and may afford more efficient methods to produce riboflavin analogues for use in future studies. Furthermore, the synthesis of 1-deazariboflavin is inefficient, and may benefit from further study.

Therefore, an examination of the synthesis of several riboflavin analogues, and their incorporation into several phototropin domains (both wild-type and mutant) was performed to address these key questions.

1.1 Introduction to Phototropins

1.1.1 Overview of Phototropism

The growth and behaviour of plants is a complex subject, which has been of significant interest to science since the time of Charles Darwin (1809-1882). One of the earliest discoveries within this field was the influence of light upon intracellular processes (such as respiration and growth) of plants, which was first described by the German botanist Julius von Sachs in 1887. Through carefully controlled experiments, he was able to use separate colours of light to determine that while red-yellow light was responsible for plant respiration[6], "the mechanical actions of light on plants are due chiefly or exclusively to the highly refrangible, blue portion of the spectrum" [7]. Thus, the first description of phototropism of plants under blue light irradiation was made.

The reason for the influence of blue light on plant movement and growth was initially unknown, until the later identification of the plant hormones, comprising auxins, cytokinins, gibberellins, ethylene and abscisic acids[8]. While animal hormones are produced in specific glands and are transported to their site(s) of action (within blood or lymph fluids), all plant cells are capable of producing their own hormones[9], which may therefore be thought of as systemically disperse. The turning movement of leaf, flower and stem direction (tropism) towards light is controlled by differential levels of auxins within the cells (causing asymmetric levels of growth across the growing plant tip (coleoptile) or other tissue, leading to its bending), which is controlled by light – hence the term "phototropism".

The predominant theory of phototropism is based on the parallel work of Cholodny and Went between 1920-1940[10,11]. This suggested that photo-stimulation of the coleoptile induces a "lateral translocation of auxin" from the more heavily irradiated side of the tip to the dark-side of the plant, leading to the observation of tropism. While this theory has come under increased scrutiny within the last 30 years[12,13], its continued use as a model remains valid according to experimental evidence[14–16].

The key step of phototropism is the initial perception of blue light by the coleoptile, which alters auxin levels in the irradiated tissue and causes the tropism effect. This photosensing function is performed by photosensitive proteins, known as phototropins, whose role is to detect light and translate this signal into a physical effect. These phototropins may therefore be thought of as protein-derived "switches". Along with their roles in phototropism, a number of other cellular processes also utilise phototropins to perform other roles, including organelle migration[17–19], stomatal opening[20–22], or other processes necessary for life or evolutionary success.

The phototropin proteins isolated from plants each appear to possess a similar action spectrum, first identified by Johnston[23]. These spectra resemble UV-visible absorption spectra, although a relative rate of physiological activity (such as tropism) induced by a specific wavelength of light is used in place of absorbance; a representative example is given in figure 1 below[24]. When compared to the UV-visible absorption spectra of isolated phototropin domains (*Avena sativa*, *Arabidopsis thaliana*, and *Adiantum capillus-veneris*[25]), a remarkable similarity may be observed.

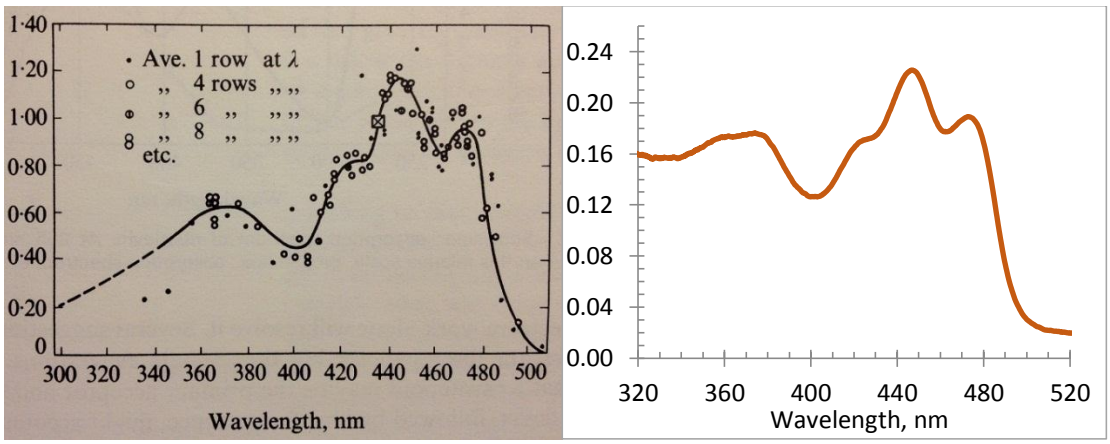


Figure 1: The Action Spectrum [left] of Avena *coleoptiles (describing relative movement per quantum at a given wavelength) shown by Curry* et al. *in "Physiology of Plant Growth and Development" (1969)[24] bears close resemblance to the UV-visible absorbance spectrum of recombinant Avena sativa PHOT1-LOV2 domain, recorded during this research*

As shown in fig. 1, the action spectrum of the *Avena* phototropin is almost identical in wavelength to the absorbance spectrum observed, suggesting a direct link between the UV-visible absorbance of the protein, and its phototropic behaviour.

1.1.2 Phototropin Structure

The gene encoding phototropism in plants has been identified as occurring at the nonphototropic hypocotyl (nph1) locus[26], so-named because it was identified by mutagenic knock-out. The corresponding plant phototropin protein is termed PHOT[27,28].

The phototropin protein comprises several separate domains, each with a discrete function; typically, a photoactive domain is directly coupled to an effector domain, such as a protein kinase[27,29] (fig. 2, *vide infra*). Upon irradiation of the photosensitive domain, a physical change takes place which activates (or deactivates) the function of the coupled effector domain, causing the downstream signalling effect.

The photosensitive agent responsible for the spectra shown in fig. 1 has been identified as a chromophore located within the photosensitive domain, able to absorb incident light of a specific wavelength and undergo an interaction with the protein to initiate the signalling effect. As described above, earlier experimental evidence demonstrated that phototropism is controlled by blue light, and so the chromophore must have a corresponding absorbance in the blue region of the spectrum. The UV-visible and action spectra each show a major absorbance at approximately 450 nm, and the photosensitive agent within phototropins has been identified as based upon a flavin moiety (see section 1.2), which has a strong absorbance in this region of the spectrum.

The absorbance of light by flavin stimulates formation of a covalent bond between itself and a cysteine residue in close proximity to the flavin chromophore (fig. 3(b) below). This cysteine residue is highly conserved throughout all phototropins, with the flavin-cysteine bond formation fundamental to the behaviour of the protein[25,29–31]. Mutant proteins of several photoreceptor families, where the cysteine has been

replaced with an unreactive alternative amino acid (such as glycine or alanine), have each been reported to lack photoactivity upon blue-light irradiation[4,31,32].

Several distinct photoreceptor families are found throughout nature; while PHOT-based photoreceptors are similar across several genera (e.g. *Avena sativa*, *Arabidopsis thaliana*, and *Adiantum capillus-veneris*[25]), using the same flavin cofactor (flavin mononucleotide, FMN), alternative photoreceptors such as cryptochrome[22,33], BLUF[34–36] or VVD[37,38] also exist based around the alternative flavin cofactor, flavin adenine dinucleotide (FAD). The differences in structure between the flavins does not vastly alter the photoreceptor's spectral characteristics (as the point of modification occurs at the terminal end of the ribityl side-chain; see section 1.2). However, significant differences in protein sequence and structure of the FAD-based phototropins makes direct comparison between FMN and FAD-based photoreceptors difficult. Subsequent discussion is therefore dedicated to the PHOT-based photoreceptors.

The complete phototropin protein is comprised of several sub-units, with separate photosensitive and effector domains. PHOT1 from *Avena sativa* comprises two signalling domains (LOV1 and LOV2), which are coupled to a serine / threonine kinase by a C-terminal linking section from the LOV2 domain. This linking section, known as the *J*-α helix, plays a key role in the transduction of the signal from the photoactive domain[39–42], as the photo-stimulated bond formation between flavin and the protein domain causes undocking of this α-helical section from the secondary structure of the LOV2 domain[43], allowing activation of the coupled protein kinase. The photo-controlled docking / undocking process occurs with both high specificity and selectivity; examination by NMR shows a relative ratio of docked to undocked *J*-α helix of 98.4 : 1.6 in the dark state, and a similar ratio of 9 : 91 in the light state[44]. Mutagenesis of the *J*-α helix has shown that when modified to prevent interaction with the LOV2 domain, the coupled kinase is active even without photoirradiation[45], confirming the role of this helical region on phototropism.

As the LOV2 domain of PHOT1 is responsible for the initial translation of light stimulation into a physical response, this domain (along with the coupled *J*-α helix) has become the primary target for phototropin research. The DNA encoding this single domain has been identified, isolated, characterised and expressed using recombinant techniques[1,4,46], allowing the characterisation of the structure and behaviour of the *A. sativa* LOV2 domain[43,47,48].

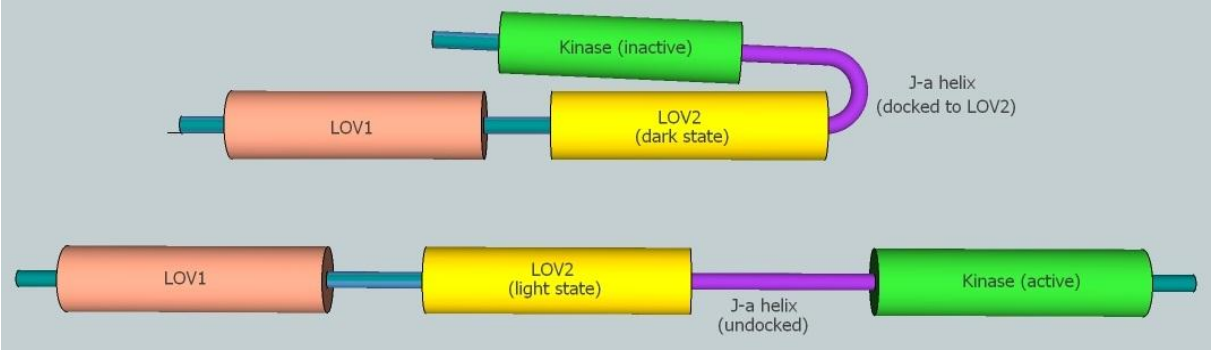


Figure 2: Schematic representations of PHOT1, (not to scale), in the dark [above] and light [below] states. In the dark state, interaction between the J-α helix and the LOV2 domain prevent activation of the coupled protein kinase. Upon irradiation of LOV2, the J-α helix is released from the domain, allowing activation of the coupled kinase domain

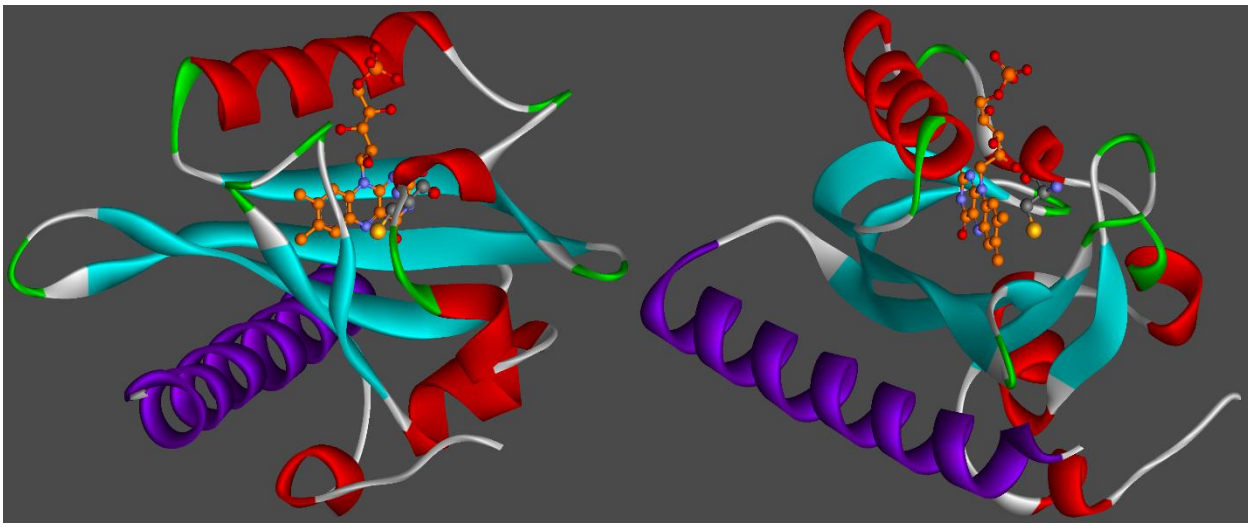


Figure 3(a): Two views of the dark-state crystal structure of the AsLOV2 domain in the dark state, from Protein DataBank sequence 2V0U[48]. The J-α helix is highlighted in purple

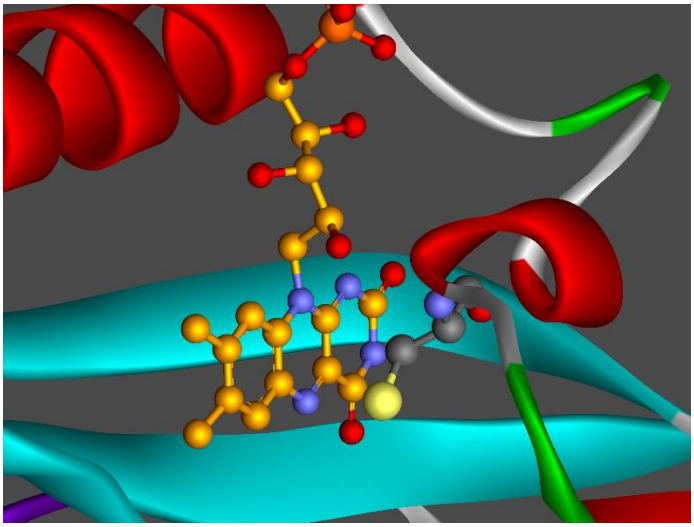


Figure 3(b): Expanded view of flavin within the photoactive site of AsLOV2 in the dark state, from Protein DataBank sequence 2V0U[48]. The reactive cysteine residue (with the sulphur atom highlighted in yellow) is in close proximity to the flavin

1.1.3 Mechanism of Phototropin Photoreaction

Two mechanisms for bond formation have been considered within the literature; originally, flavin-protein bond formation was believed to occur via a nucleophilic attack of the photo-excited flavin (described as the "nucleophilic" or "ionic" mechanisms)[30,49]. However, experimental evidence has led to this mechanism being superseded by an alternative, which more closely matches observed results.

As flavin is as an efficient chromophore (due to a large delocalised system of electrons), the absorption of light allows excitation of an electron from the ground state to a primary (singlet) excited state. As the stability of the excited state in flavin is relatively high, it is possible for an excited electron to undergo intersystem crossing (ISC), giving a triplet excited state for the radical flavin. This photo-induced radical may then combine with a cysteinyl radical to form a covalent bond between the flavin and the protein; this is known as the "radical pair" mechanism)[4,46].

A schematic representation for the proposed nucleophilic mechanism[30] is shown in figure 4 below. This suggests that photoexcitation of flavin allows deprotonation of the cysteinyl group by the N5 position of flavin, with the deprotonated cysteinyl

sulphur attacking flavin at position C4a to form the covalent adduct. This was devised based upon earlier molecular orbital calculations for flavin[50] which suggest an increase in the relative basicity of position N5 of flavin upon photoexcitation to the lowest singlet or triplet states.

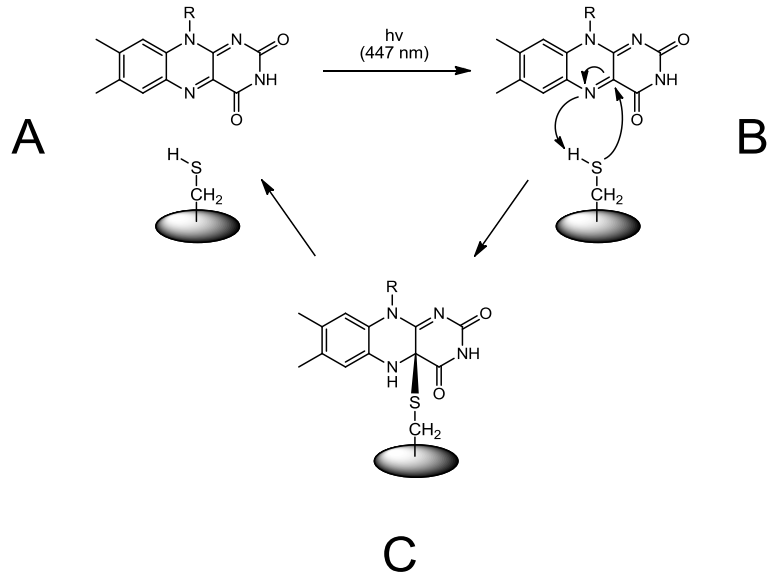


Figure 4: Schematic representation of a nucleophilic mechanism for flavin-cysteine photoadduct formation, proposed by Crosson and Moffat[30]. The interaction between flavin and cysteine may proceed either in a concerted manner (as shown), or as a sequential reaction (also referred to as the "ionic" mechanism), due to stabilisation of the charged intermediates by neighbouring residues within the protein

However, several flaws with the nucleophilic or ionic mechanisms have been found. While initial evidence suggested a partial protonation of the FMN triplet state over a nanosecond timescale[51], more recent experimental evidence suggests that this is not the case, as the lifetime of the flavin triplet state has been found to persist for several nanoseconds without showing evidence of protonation[52]. Furthermore, no clear consensus has been reached about whether the transferred proton originates from cysteine[30,53], or another source[47].

On the other hand, the contrasting radical pair mechanism[46] described above has found significant favour in the recent literature, due to an overwhelming number of supporting results. The stability of the excited flavin radical within phototropin to

protonation over extended timescales[52], density-functional theory (DFT)[54] and quantum-mechanical[55] calculations of the energies and permissible reactions of both ground-state and excited flavin within the phototropin, supplemented by the experimental observation of neutral $FMNH^{\bullet}$ semiquinone radicals in wild-type and mutant phototropins using electron paramagnetic resonance (EPR)[4,46], each support the proposed radical pair mechanism.

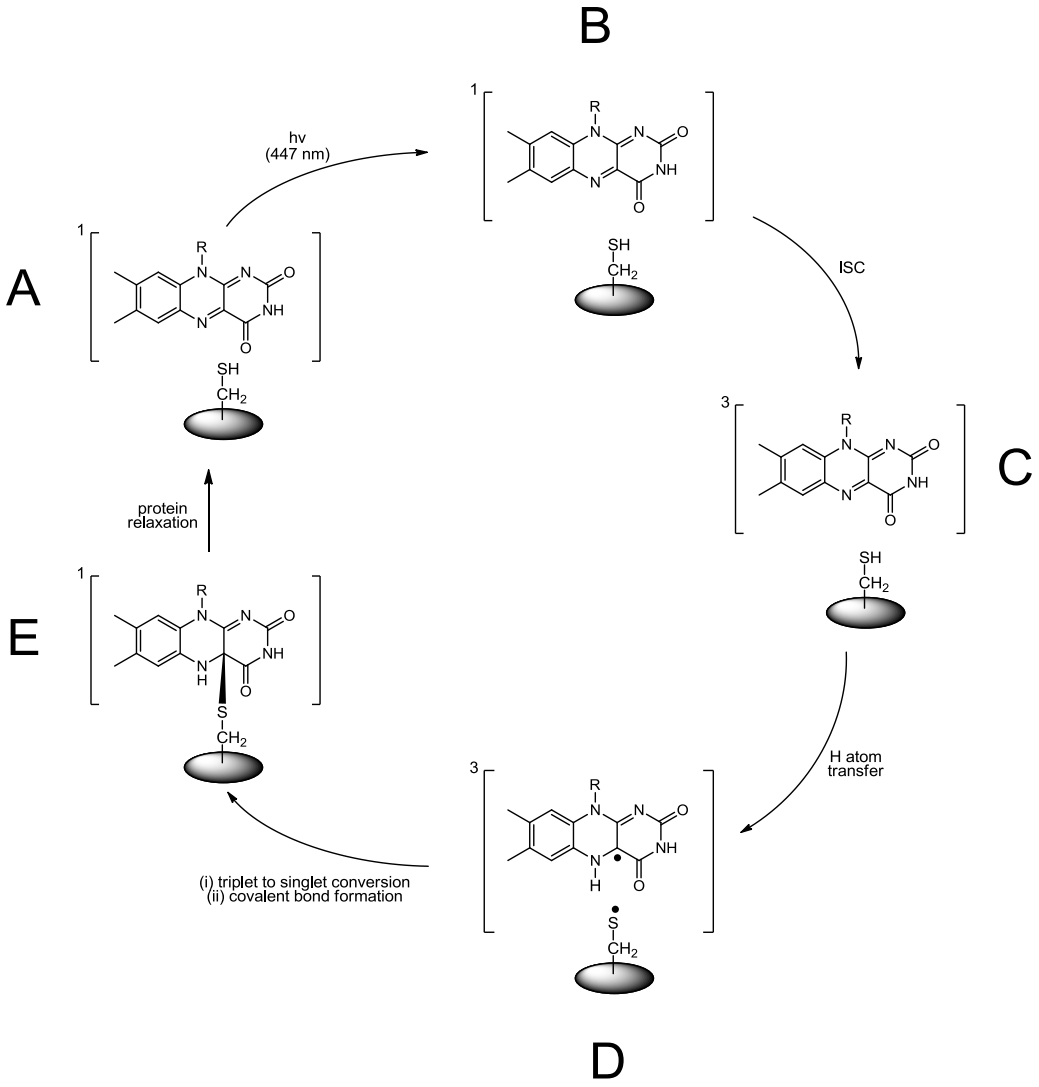


*Figure 5: Proposed mechanism for the photocycle of the AsLOV2(WT) domain, originally described by Schleicher et al.[46]. Several steps may be bidirectional (or lead directly back to dark-state **A**) due to alternative pathways for radiative decay from excited states **B**-**D***

The photoreversion of the light-state occurs spontaneously, with a half-life of 28-32 seconds upon cessation of irradiation[31,56]. This short lifetime ensures that the coupled

protein kinase is only active during illumination, although it has been shown that the lifetime of the undocked state of the *J*-α helix may be modulated by the mutation of various residues within the domain by altering the stability of the docked *J*-α helix[56,57]. Imidazole has also been shown as being particularly effective at increasing the rate of photoreversion of the illuminated species[58], thus careful consideration to address each of these factors is necessary in the design of subsequent experiments.

1.2 Introduction to Flavins

1.2.1 Flavin Structure

The term "flavin" (derived from the Latin '*flavus*', meaning yellow) describes a complex biological molecule containing a tricyclic isoalloxazine structure, as shown in figure 6 below. This unique structure permits the use of flavins for a diverse range of functions; they may act as a redox cofactor[59–63], a light accepting chromophore (as described above), in the detoxification and removal of free-radicals and peroxides within the cell[64], or as a signal transducer during programmed cell death (apoptosis)[65,66].

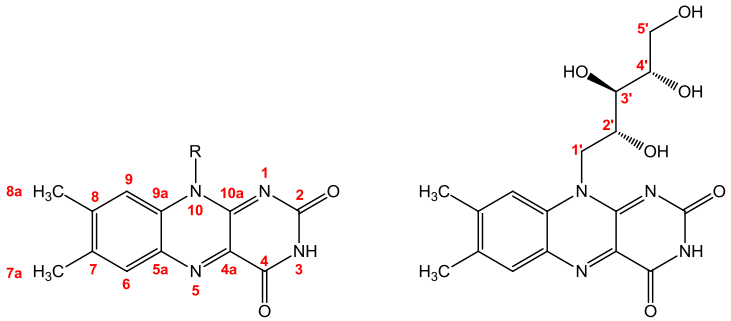


Figure 6: Structure and atomic numbering of flavin [left] and riboflavin [right]. Numbering proceeds clockwise from the ordinarily-substituted N10 position; ring-connecting carbon positions (4a, 5a, 9a and 10a) are numbered according to the preceding position. As only the D-enantiomer of ribose is found in nature, the side-group of riboflavin derived from this retains this stereochemistry, and is assumed in all subsequent figures

The simplest flavin product formed by nature is riboflavin (vitamin B_2; figures 6 and 7). This name is a portmanteau of ribose (the prosthetic group is derived from the

pentose sugar D-ribose) and flavin, which are linked at the N10 position of the isoalloxazine (figure 6 above). As the D-ribityl group is originally derived from the sugar ribose (which is only present in nature as the D-enantiomer), the original stereochemistry of this group is retained throughout all biological forms of flavin, and provides a tool for recognition of the flavin by flavoproteins.

The biosynthesis of riboflavin begins using GTP[67] (figure 7), and occurs via a similar pathway across plant, bacterial and fungal species[68,67,69,70]. However, all flavoenzymes use forms of riboflavin which are further modified, known as flavin cofactors.

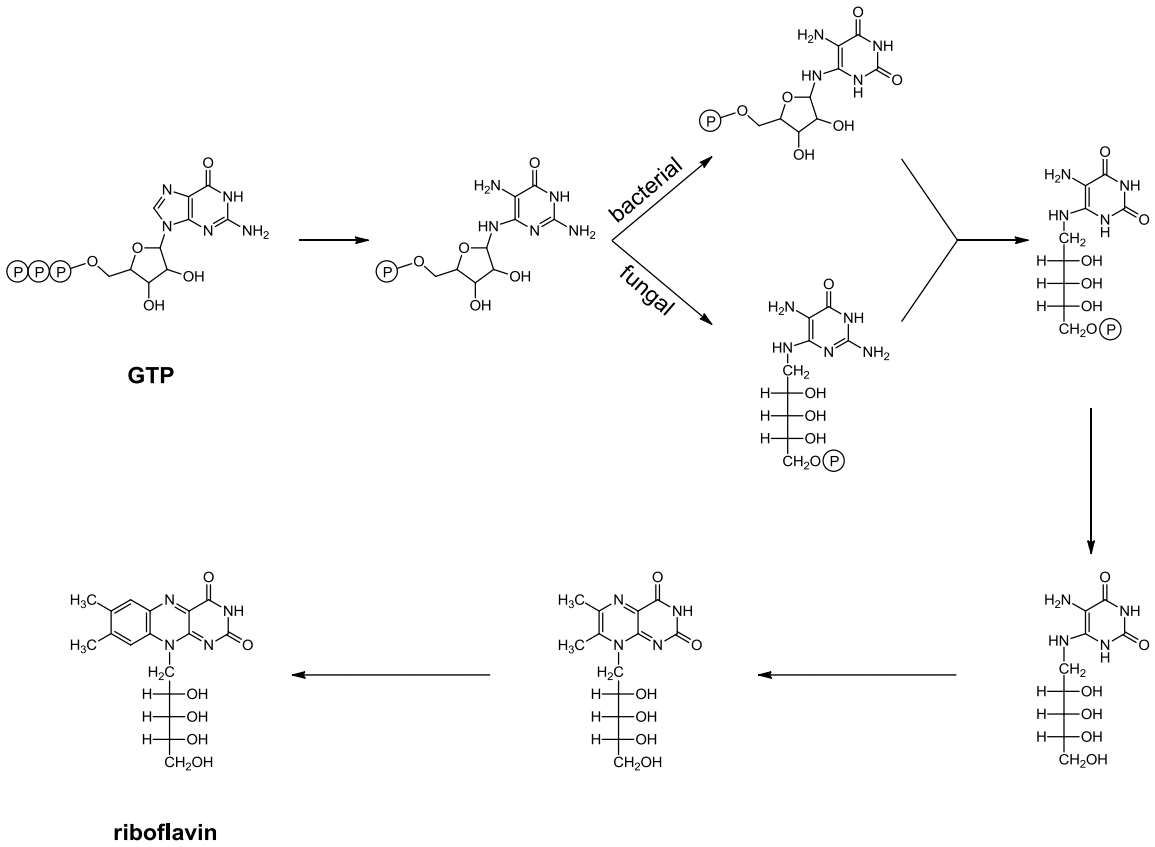


Figure 7: Biosynthetic formation of riboflavin, described by Bacher et al.[67]. GTP is dephosphorylated, before divergence of the bacterial and fungal routes. However, each intermediate leads to the same product, which is further modified by incorporation of additional ribulose-5-phosphate, leading to riboflavin

1.2.2 Flavin Cofactor Forms and Enzymatic Phosphorylation

Two derivatives of riboflavin (shown in figure 8) are used by nature as enzymatic cofactors. These are flavin mononucleotide (FMN; also known as riboflavin 5'-monophosphate) and flavin adenine dinucleotide (FAD; also known as riboflavin 5'-adenosine diphosphate), and are each formed by enzymatic phosphorylation of riboflavin). In addition to significantly improved solubility, incorporation of phosphate or adenosine diphosphate groups acts as a convenient recognition tool for the protein, and allows secure retention of the desired cofactor by providing an ionic interaction between polar amino acids and the charged phosphate group(s).

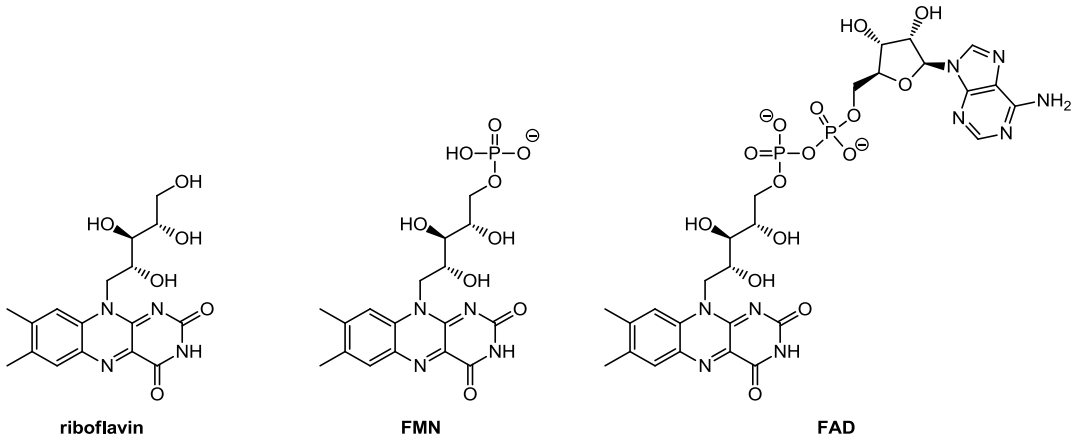


Figure 8: Structures of riboflavin, FMN and FAD

Formation of FAD from riboflavin *in vivo* is performed sequentially; FMN is produced initially by Riboflavin Kinase (RfK), before further modification to FAD by a Nucleotidyl Transferase (NT) or an FMN Adenyltransferase (FMN-AT). Conventionally, the NT and FMN-AT domains are both commonly referred to as FAD synthetase (FADS) domains[71].

The precise route of FAD formation differs between prokaryotic and eukaryotic cells, as eukaryotic cells typically use separately expressed RfK and FADS domains to perform each reaction. In contrast, the majority of prokaryotic cells utilise a combined protein containing both domains. The reason for this remains largely unknown, although recent advances in the statistical analysis of DNA sequences (using methods such as genome mining) may shed light on the reason for this apparent divergence.

In terms of protein structure, the RfK domain of the bifunctional prokaryotic enzyme maintains a strong homology with the RfK domain expressed by eukaryotic cells[72–75]. While the FMN formed within the RfK domain of the bifunctional enzyme must be released and re-bound separately by the NT domain[73] (as there is no shuttling mechanism to transfer the substrate between domains[74,75]), their close proximity assists cofactor transport.

However, the FADS domains are remarkably different. While the structure of the eukaryotic domain is classified as an FMN-AT domain[72], the prokaryotic FADS domain is defined by its structure as an NT domain[74,76,77]. The crucial difference between these distinctions is that while the FMN-AT produces FAD unidirectionally (figure 9), the NT domain catalyses both forward and reverse reactions (figure 10).

The reverse reaction (FAD to FMN) occurs as the NT domain stabilises the transition state of the FMN-FAD reaction[74,75]. It has been postulated that this allows the bifunctional protein to perform a further regulatory role in the maintenance of intracellular stores of the flavin cofactors[78,79]. Thus, if an intracellular deficit of either flavin cofactor occurs in the prokaryotic cell, this protein is able to rapidly use any source of flavin to meet the demand, providing a useful feedback mechanism for the cell and allowing ready supply of both cofactors at all times[73,78,79].

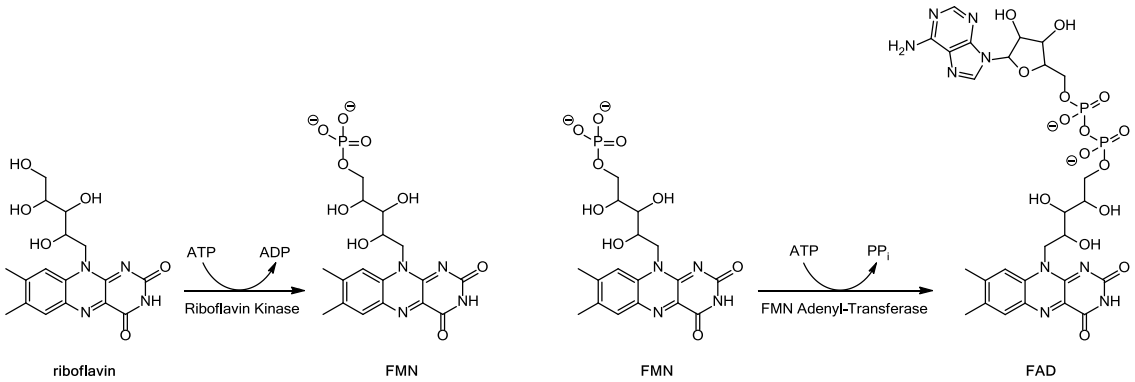


Figure 9: Formation of FMN from riboflavin by a eukaryotic RfK enzyme, and the conversion of FMN to FAD by a separate FMN-AT enzyme. Each enzyme works independently. Note that formation of FAD by an FMN-AT enzyme is unidirectional

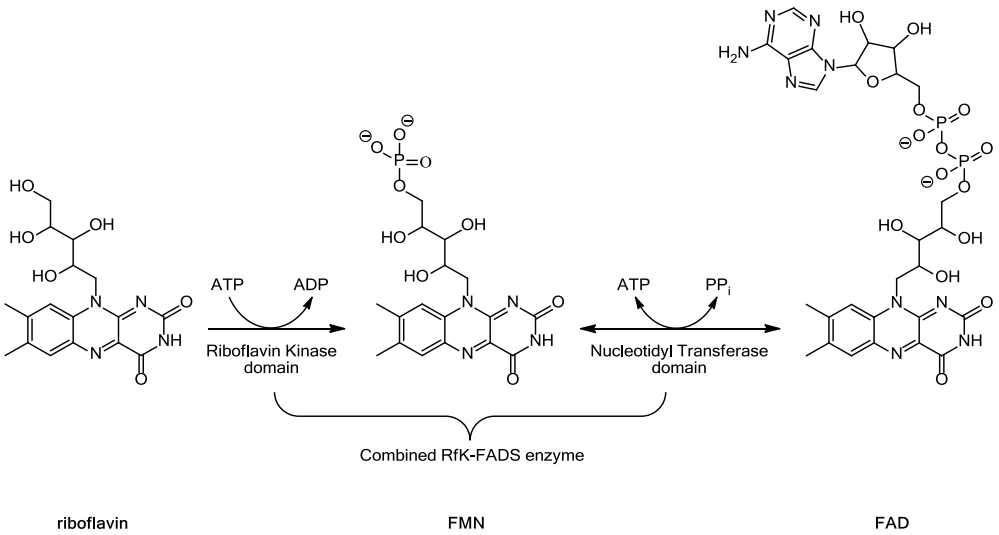
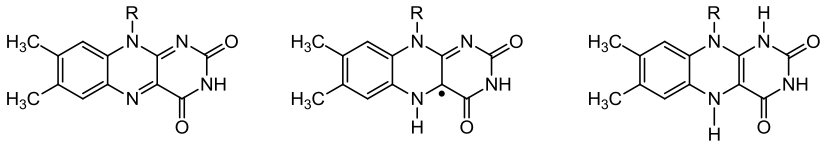


Figure 10: Formation of both FMN and FAD by a bifunctional prokaryotic RfK-FADS protein. The production of both cofactors is linked by the bidirectional functionality of the NT domain, allowing the cell to maintain a ready supply of both cofactors using a single protein. Each catalytic domain is located at separate sites of the protein; FMN formed by the RfK is released and re-bound separately by the NT domain

1.2.3 Flavin Redox Behaviour

The isoalloxazine moiety of each cofactor is a particularly good chromophore, due to the extended delocalisation of electrons throughout its planar, sp^2 hybridised structure. Flavin compounds are able to accept photons of visible wavelengths (445-450 nm) and thus appear yellow in colour. However, not all flavin-mediated processes are light-dependant. Instead, flavin may be easily reduced at positions N1 and N5 (linked via C10a and C4a) to a "hydroquinone" form, by acceptance of 2 protons and 2 electrons, forming $FMNH_2$ or $FADH_2$ (commonly referred to as Fl_{ox} and Fl_{red}, respectively; fig.11(a)). A penalty of this is the loss of aromaticity, demonstrated by the UV-visible absorption of $FMNH_2$ compared to FMN; loss of conjugation leads to greatly reduced extinction coefficient and significant blue-shifting of the spectrum, demonstrated by the migration of the Fl_{ox} absorbance of FMN from 445 nm (ε = 12 200 M^{-1} cm^{-1}) to 395 nm (ε = 2 250 M^{-1} cm^{-1}) for Fl_{red}[80] shown in fig. 11(b). As the loss of aromaticity is energetically unfavourable, this allows the flavin to perform both electron transfer and energy transfer, recognised via its important role in the electron transport chain.



Fl_{ox} $FlH^{\bullet}$ Fl_{red}

Figure 11(a): structure of quinone-form Fl_{ox}, semiquinone form $FlH^{\bullet}$ and hydroquinone Fl_{red}

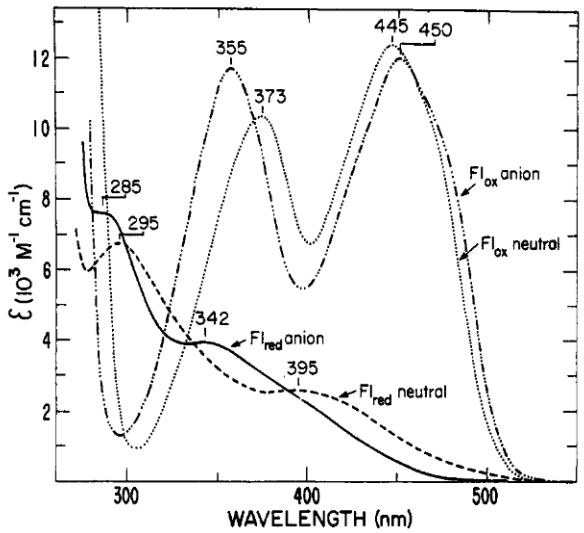


Figure 11(b): UV-visible absorption spectrum of flavin in oxidised and reduced states, at neutral and basic pH, shown by Ghisla et al.[80]

While other redox cofactors (such as NADH / NADPH) perform similar redox reactions to flavin, flavin has been shown to allow a stable radical intermediate state[81–83], known as the flavin semiquinone ($FlH^{\bullet}$). The stability of this intermediate allows the sequential transfer of single electrons in the electron transport chain, which is necessary due to the limitation of iron-sulphur clusters (used to mediate electron transport) to accept only a single electron. This is in contrast to other redox cofactors, which perform only two-electron transfers.

Due to the wide distribution of flavin throughout nature and its importance for a diverse range of enzymatic processes, several chemists of the early 20th century attempted the total synthesis of riboflavin. This was successfully completed by Paul Karrer in 1935[84], who was subsequently awarded the Nobel Prize for Chemistry in 1937 for this and the associated synthesis of carotenoids and vitamin A. The method

described by Karrer, shown in figure 12, has several themes which are revisited in subsequent syntheses of flavin (and its analogues).

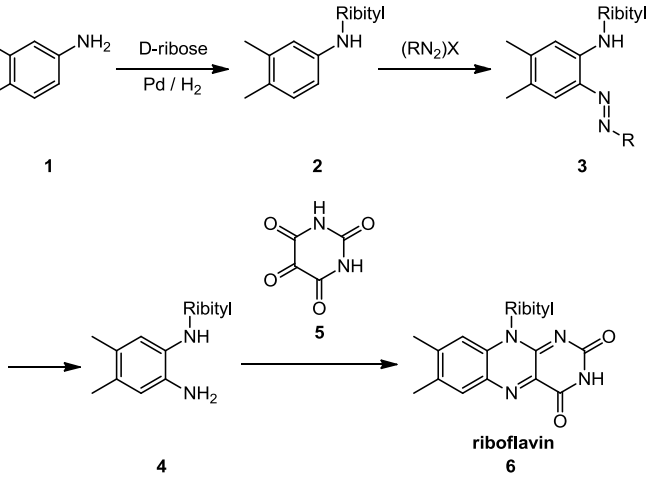


Figure 12: Synthetic route for the total synthesis of riboflavin reported by Karrer[84]. "$(RN_2)X$" refers to one of a number of diazonium salts used by Karrer

1.2.4 Flavin Analogue Structure and Behaviour

After the identification of the structure of flavin and its successful total synthesis, subsequent experimentation was performed to create structural analogues of the compound, to allow the deduction of the mechanisms of flavoproteins.

As the ribityl side-chain of riboflavin was incorporated as a complete unit during Karrer's synthetic method (and it does not play a significant role in the electrochemistry of the flavin), it was possible to easily substitute this for similar alternatives, to examine its role in protein binding[85,86]. However, the results of this were of limited use (as the recognition and incorporation of flavin cofactors was found to be strongly linked to the D-ribityl group), and so direct modification of the flavin structure was used to better assess the behaviour of flavin.

The role of the two nitrogen positions (N1 and N5) was found to be key to the chemistry of flavin, as these positions may each be protonated (forming the semiquinone or hydroquinone form; see above), as well as potentially forming hydrogen bond donor-acceptor bridges. The significant difference in lability between the N-H bond and the C-H bond, and its exploitation by the flavoprotein, made the

examination of these positions using stable analogues crucial for mechanistic investagations[87]. When the nitrogen atoms at positions N1 and N5 were (individually) isoelectrically substituted for CH units to form the corresponding carba-deazaflavins (also known as deazaflavins; figure 13), a demonstrable loss of function was observed in protein samples using the incorporated synthetic analogues.

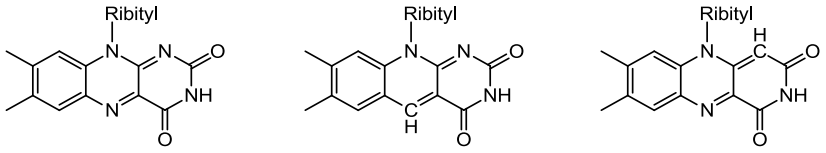


Figure 13: Structures of riboflavin, 5-deazariboflavin and 1-deazariboflavin, showing incorporation of a CH unit in place of a nitrogen atom

The UV-visible spectra of the 1-deaza- and 5-deazaflavins compared to the original spectrum of riboflavin is shown in figure 14. The major absorbance of FMN (at 447 nm) is blue-shifted by approximately 50 nm for 5-deazaFMN, while it is red-shifted by 90 nm for 1-deazaFMN. The extinction coefficient of 5-deazariboflavin (ε_{400} = 12500 M^{-1} cm^{-1})[88] remains the same as riboflavin, whereas the extinction coefficient of 1-deazariboflavin is significantly lower (ε_{535} = 6800 $M^{-1}cm^{-1}$)[88].

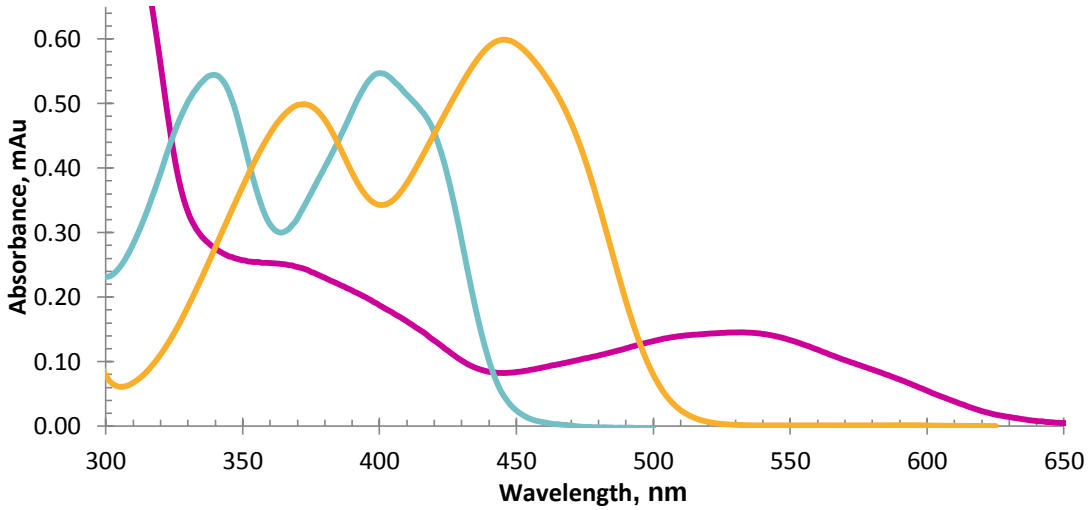


Figure 14: UV-visible absorbance spectra of FMN (yellow), 5-deazaFMN (blue) and 1-deazaFMN (purple). The spectra of the flavin analogues are either blue or red shifted compared to FMN, while displaying similar spectral characteristics

One of the key reasons for the loss of function with 5-deazaflavins, as reported by Edmondson *et al.*[89], is that the N5 position is crucial for the stabilisation of a radical semiquinone state of the flavin. Theoretical calculations for 5-deazariboflavin suggest the semiquinone state is "energetically highly unfavourable" [90,91], a conclusion supported by experimental evidence where no semiquinone state is observed[92]. However, 5-deazaflavin does retain the ability to undergo two-electron transfers, shown by acceptance of hydride from both chemical (e.g. borohydride) and enzymatic (e.g. NADPH reductase) sources[93–96]. However, physical changes to the flavin analogue upon incorporation of an additional substituent at C(H)5 would lead to significant distortion of the (previously planar) isoalloxazine structure, providing a significant energetic barrier to formation of this species.

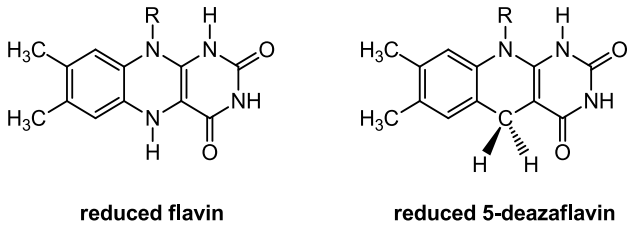


Figure 15: Structure of reduced flavin, compared to reduced 5-deazariboflavin. The incorporation of hydrogen at position C5 of the deazaflavin causes significant distortion of the flavin molecule

While 5-deazaflavins have been successfully utilised to elucidate the mechanism of several flavoprotein-catalysed reactions, 1-deazaflavins have been less well studied. This is largely due to the difficulty of their synthesis (often formed as a side-product during alternative syntheses, with yields below 10 %, see below), along with their limited use as mechanistic probes. 1-deazariboflavin has been found to undergo both one- and two-electron redox processes[97], with the only difference being a slightly more negative redox potential (E_m(1-deazaFAD) = -280 mV[98]; E_m(FAD) = -208 mV)[99]. The significant differences in UV-visible absorption spectra compared to native or 5-deazaflavins (figure 14), along with an observed lack of fluorescence[97,100,101] has generally precluded their use in cases where the flavin acts as a chromophore (although they have recently been successfully used for this purpose[102]).

1.2.5 Flavin Analogue Synthesis

The synthesis of riboflavin devised by Karrer[84] (shown in figure 12, *vide supra*) has provided the foundation for subsequent syntheses of the riboflavin analogues 5-deazariboflavin **7** and 1-deazariboflavin **8**. The original method describes the reductive amination of 3,4-dimethylaniline **1** with D-ribose, followed by incorporation of an amine (via a diazonium species, with the remaining nitrogen eventually becoming position N5 of the flavin) *ortho* to the ribitylated amine **2**. Finally, alloxan **5** was reacted with intermediate species **4**, resulting in formation of riboflavin **6**.

Some steps of this synthesis have been developed, while other aspects of the method have been retained. Improvements to individual reactions or alternative routes have been developed largely due to the advancement knowledge of organic chemistry; for instance, the use of protecting group theory affords an alternative method for formation of intermediate **22** (figure 23).

1.2.6 5-Deazariboflavin Synthesis

The initial reductive amination of D-ribose by 3,4-dimethylaniline **1** has been rationally improved through extensive experimentation[103,104], culminating in the method of Carlson and Kiessling[5] whereby the use of sodium cyanoborohydride ($Na[CNBH_3]$) afforded reliable formation of the ribitylated aniline without the previously required use of high pressure hydrogenation with Pt/Pd catalysis. Alternative hydrides (such as sodium triacetoxyborohydride, $Na[(OAc)_3BH]$) have been successfully reported to improve reductive aminations in some cases[105], and the results of their use as a potential candidate to replace cyanoborohydride for this reaction are discussed in subsequent sections of this thesis.

The key step in the synthesis of 5-deazariboflavin is the successful incorporation of carbon at position 5 of the isoalloxazine. Relatively few methods were known to permit incorporation of a carbon substituent to an aromatic ring; therefore early syntheses utilised an aniline with a protected functionalised carbon at the *ortho*-position (figure 17, intermediate **9**). The initial method for the synthesis of 5-

deazariboflavin **7** described by O'Brien *et al.*[103] reported a somewhat difficult condensation between 4,5-dimethyl-N-ribitylanthraniladehyde **10** and barbituric acid **11**. A similar reaction was also reported for the synthesis of 1,5-dideazariboflavin, using 2,4,6-piperidinetrione[106] in place of barbituric acid. However, several undesired side-products were identified within the reaction mixture, leading to lower yields of 5-deazaflavin overall.

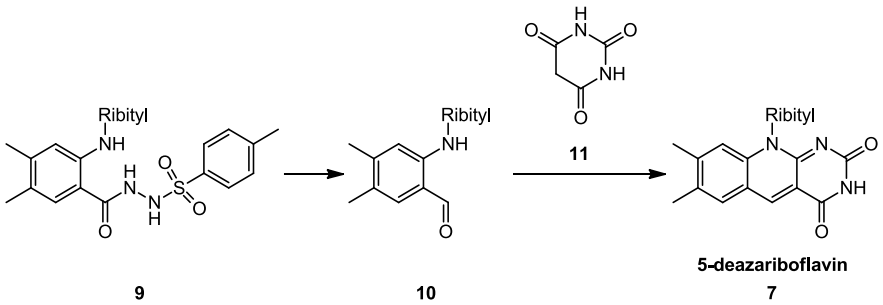


*Figure 17: Synthetic method for formation of 5-deazariboflavin **7** devised by O'Brien et al.[103]*

A breakthrough in the formation of 5-deazariboflavin was made during the preparation of riboflavin and riboflavin-N5-oxide via a bicyclic intermediate **13** (figure 18), reported by Yoneda *et al.*[107,108]. This identified an alternative method for reliably incorporating the uracil-moiety of flavin using 6-chlorouracil **12**, with subsequent introduction of an amine (to give the eventual N5 nitrogen atom) by nitration or nitrosation giving the respective riboflavin or riboflavin-N5-oxide. As 6-chlorouracil **12** is functionalised at position 6, attack by the ribitylated amine is directed towards this position, lowering the potential for undesired side-product formation. This method paved the way for future synthesis of 5-deazariboflavin using an electrophilic carbon species for the incorporation of carbon in bicyclic intermediate **13**, in vastly improved yields over previous reactions.

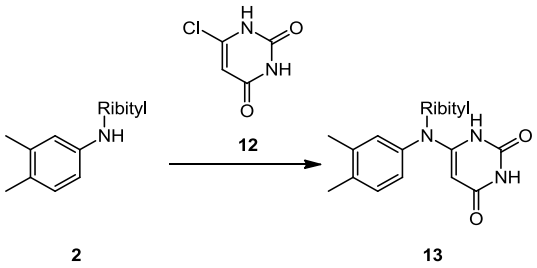


*Figure 18: Formation of bicyclic intermediate **13**, originally reported by Yoneda et al.[107,108]*

A number of methods to perform the closure of the central pyridine ring have been presented in the literature. The treatment of bicyclic intermediate **13** using trimethyl orthoformate (in the presence of a catalytic quantity of *p*-toluenesulfonic acid; figure 19) produced with bis(methoxymethylene)-protected 5-deazariboflavin **15**[109]. Removal of the methoxylmethylene groups afforded 5-deazariboflavin **7** in yields up to 46 % (two steps) from bicyclic intermediate **13**.

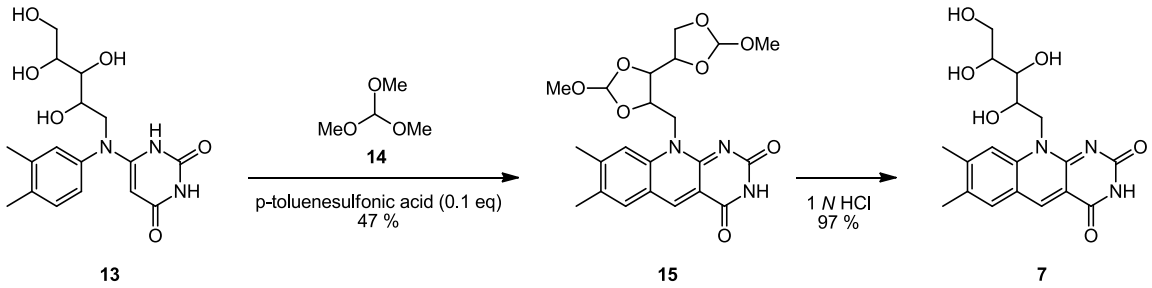


*Figure 19: Closure of the central pyridine ring of 5-deazariboflavin performed using trimethyl-orthoformate **14**[109]. Unintended protection of the ribityl side-chain necessitated subsequent deprotection, although this was performed in high yield, affording 5-deazariboflavin **7** in 46 % yield overall*

An alternative to the use of trimethyl orthoformate was the use of the Vilsmeier-Haack reagents, phosphorus oxychloride ($POCl_3$) and dimethylformamide (DMF)[110], to introduce the formyl group to the bicyclic intermediate **16** (figure 21). This method, reported by Yoneda *et al.*[111,112] (in 57 % yield) uses the reaction between $POCl_3$ and DMF to form a reactive chloriminium intermediate species (figure 20), which is particularly susceptible to nucleophilic attack. Upon reaction with bicyclic intermediate **16** and hydrolysis of the resultant bicyclic iminium species (**16a**) to the

formylated species (**16b**), an intramolecular reaction (followed by elimination of water) may furnish 5-deazaflavin **7**. Thus, this method has been successfully used to prepare 5-deazariboflavin[5,100].

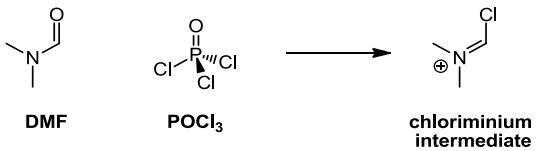


Figure 20: Chloriminium species formed by reaction between $POCl_3$ and DMF

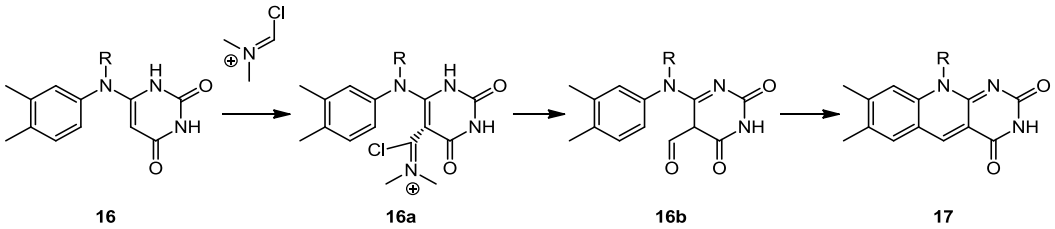


*Figure 21: Formation of 5-deazaflavin from bicyclic intermediate **16** utilising the Vilsmeier-Haack reagents, as reported by Yoneda et al.[111,112]. "R" corresponds to a tetra-acetylated D-ribityl group*

As shown, several stages in the formation of 5-deazariboflavin **7** remain poorly understood, with few advancements made since development of the original synthetic route[103]. Therefore, a rational re-examination of each stage may be of particular interest to the scientific community, and assist in the identification of the most effective route for this compound's synthesis.

1.2.7 1-Deazariboflavin Synthesis

The synthesis of 1-deazariboflavin **8** has been sparsely covered within the literature, (compared to the alternative isoelectric analogue 5-deazariboflavin **7**, *vide supra*). The "modern" method described by Carlson and Kiessling[5] (repeated by Mansurova *et al.*[100]) is essentially identical to the method for the synthesis of 1-deazalumichrome **21** (a flavin lacking a substituent at position N10) originally used by Ashton *et al.*[104] 27 years previously, shown in figure 22 below.

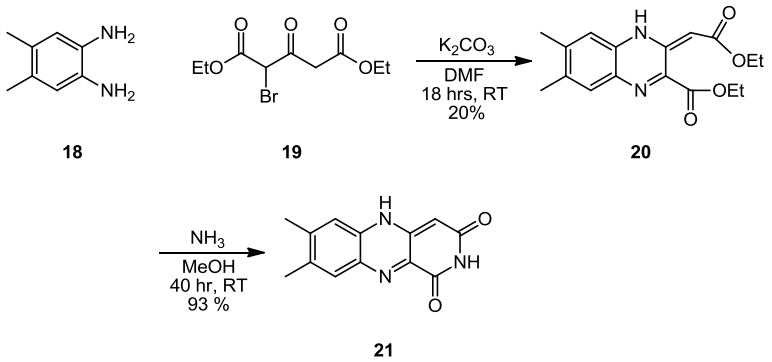


*Figure 22: Synthetic route to 1-deazalumichrome **21** reported by Ashton et al.[104] in 1977, and which was the basis for contemporary syntheses of 1-deazariboflavin*

This method was successfully utilised within the original study to furnish 1-deazariboflavin utilising a ribitylated derivative **22** of the initial diamine, although the yield of 1-deazariboflavin **8** in this case was reported as 14 % over the two steps shown in figure 22 (compared to 19 % for 1-deazalumichrome **21**)[104].

An alternative route for the formation of 1-deazariboflavin **8** was also presented by Ashton *et al.*[104], whereby ribitylated diamine **22** (figure 23) was instead condensed with 2,4,6-piperidinetrione **23** to furnish 1-deazariboflavin tetraacetate **24**, again isolated in 14 % yield. However, the major product of this reaction was in fact found to be 3-deaza-riboflavin tetraacetate **25**, obtained in 36 % yield. This appears somewhat counterintuitive when the relative availability of the carbonyl positions of 2,4,6-piperidinetrione is considered, as two carbonyl groups are available α- to the nitrogen of the piperidine heterocycle, compared to one in the γ- position. However, the relative electrophilicity of the two carbonyl positions is considerably different, with the carbonyl group at the γ- position being a significantly more effective electrophile. It is important to note that although each amine would be expected to be a weak nucleophile, if 2,4,6-piperidinetrione **23** was attacked at either position by the free amine of ribitylated diamine **22**, neither product would be formed, and so the desired product must only be formed by attack of 2,4,6-piperidinetrione **23** by the secondary (ribitylated) amine. As the primary product of this reaction is the formation of 3-deazaflavin **25**, rather than the desired 1-deazaflavin **24**, this method has been largely abandoned, with no further reference made in the literature.

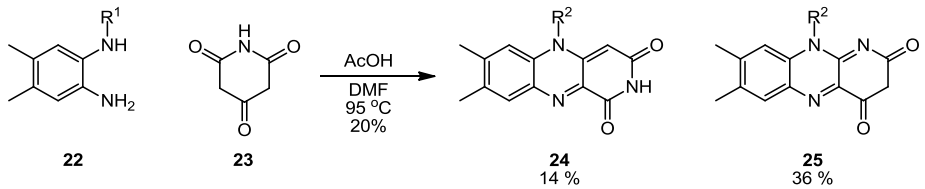


*Figure 23: Alternative route for the synthesis of 1-deazaflavin developed by Ashton et al.[104]; R^1 refers to a D-ribityl group, while R^2 describes the tetra-acetylated derivative of this group. The predominant product of this reaction was 3-deazariboflavin tetraacetate **25**, rather than 1-deazariboflavin tetraacetate **24***

As described above, the recent method[5] for the synthesis of 1-deazariboflavin relies heavily on the original work of Ashton *et al.*[104]. However, a description of the initial ribitylation of the starting material 4,5-dimethylbenzene-1,2-diamine **18** by Carlson and Kiessling[5] has been useful in providing reliable conditions for the initial de-symmetrisation of the compound using Boc-protection prior to ribitylation via a reductive amination, preventing formation of a 1,2-diribitylated species. Additionally, Carlson and Kiessling discovered that substitution of the previous base (K_2CO_3) with an alternative (Cs_2CO_3), and the use of an additional solvent (CH_2Cl_2, alongside DMF used previously), allowed the yield for the subsequent reaction to be improved from 20 % to 55 % (figure 24). However, this yield remains mediocre, and thus improvement to this step may lead to improved efficiency in the synthesis of 1-deazariboflavin **8**.

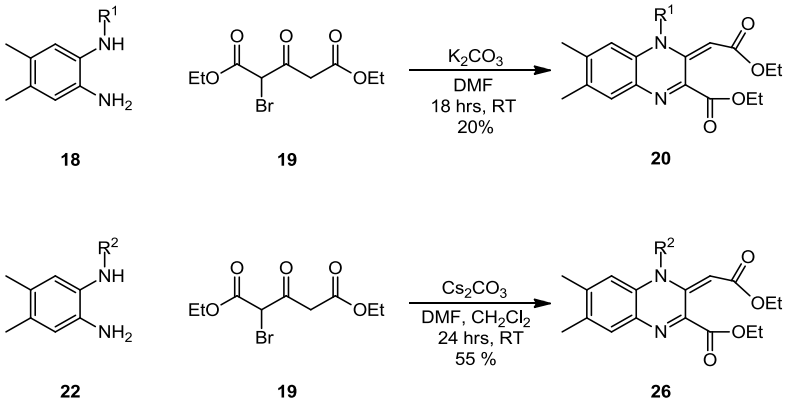


*Figure 24: Improved coupling conditions between ribitylated diamino compound **22** with 2-bromo-3-oxoglutarate **19** affords improved reliability for the formation of bicyclic intermediate **22**, when compared to the previous method of Ashton et al.[104]. R^1 refers to a proton, while R^2 describes a D-ribityl group*

1.3 Aims and Objectives

In order to accomplish the goals described in the motivation for this thesis (section 1.0.2), a number of individual objectives were set, as described below. The outcome of each objective is given and discussed in subsequent sections of this thesis.

(1) Performing and examining potential optimisations for the synthesis of 5-deazariboflavin, 1-deazariboflavin and other related flavin analogues.

(2) Devising a new route for the synthesis of additional analogues of riboflavin.

(3) Expressing, purifying and determining optimal conditions for a riboflavin kinase protein, necessary to perform the phosphorylation of the chemically synthesised riboflavin analogues to the FMN cofactor form.

(4) Expressing, purifying and devising a reliable method to produce quantities of recombinant PHOT1-LOV2 from *A. sativa* containing native FMN, sufficient for later examination.

(5) Determination of conditions for cofactor exchange within the expressed protein, and incorporation of the FMN analogues prepared in earlier stages.

(6) Creating, expressing and purifying a range of C450X mutants of PHOT1-LOV2, including incorporation of FMN analogues within mutant protein samples.

(7) Performing spectroscopic examination of the samples formed in stages 4 to 6, to obtain evidence for the mechanism of the PHOT1-LOV2 protein photocycle.

Chapter 2

Results and Discussion –

Chemical Synthesis

2.0 Results and Discussion - Chemical Synthesis

2.0.1 Introduction

In this chapter, the syntheses of 5-deazariboflavin, 1-deazariboflavin and 8-demethyl-5-deazariboflavin are discussed. The synthetic method used to produce these compounds has been analysed, improvements found are described, and other unsuccessful reactions also noted in order to improve the overall understanding of this area.

Additionally, an alternative pathway with the potential to produce 4 riboflavin analogues (1-deazariboflavin, 1,5-dideazariboflavin, 3-deazariboflavin and 3,5-dideazariboflavin) from a single starting material is described. This was devised using the benefit of similar reactions performed previously, requiring the synthesis of pyridine-based compounds in place of 6-chlorouracil. As several stages of this method were intensively studied during the synthesis of 5-deazariboflavin, their application to this route may have enabled the formation of other riboflavin analogues which have seldom (if ever) been prepared. While the formation of the compounds required was unsuccessful under all conditions examined, the work described here may provide the basis for a future synthesis of these compounds.

2.1 Synthesis of 5-deazariboflavin

2.1.0 Introduction

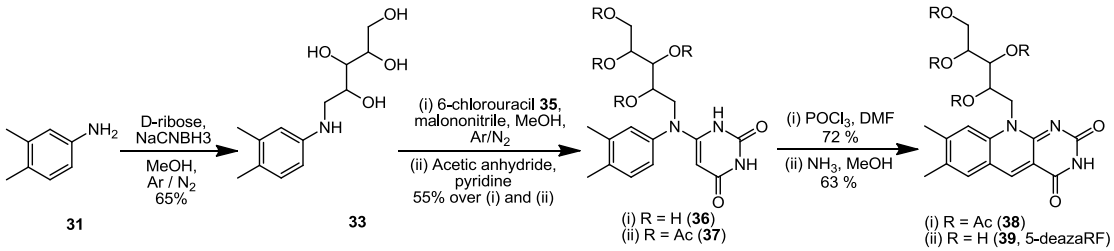


*Figure 25: Overall Synthetic Route to 5–deazariboflavin **39***

The synthesis of this analogue was based upon the route described by Erin E. Carlson and Laura L. Kiessling[5], which has been repeated, studied, and undergone experimental optimisation to better understand the formation of 5-deazariboflavin **39**, achieved with a reported yield of 16.2 % overall. This built upon decades of

comprehensive study by numerous devoted organic chemists, and inevitably few stones have been left unturned in finding improvements to the yield or a simplification of the synthesis. However, significant experimentation has been performed to better-understand each step of this reaction.

2.1.1 Formation of 3,4-dimethyl-N(ribityl)-aniline 33

2.1.1.1 Reaction overview

This step involves the reductive amination of D-ribose facilitated by the attack of nitrogen from 3,4-dimethylaniline **31** (figure 26), proceeding via an alkylimino-de-oxo-bisubstitution[113] (alkyl-DOB) reaction. This is followed by reduction of the Schiff-base intermediate to the corresponding amine using $Na[CNBH_3]$, producing 3,4-dimethylriboaniline **33** with a yield of 65 %. While this is lower than previously reported[5,100], this yield was incredibly reliable, with over 50 repetitions at varying scales found to be consistent within ± 2 %. This poses questions of the yields described elsewhere for this reaction, and so extensive investigation was performed to obtain a better understanding of factors affecting product formation.

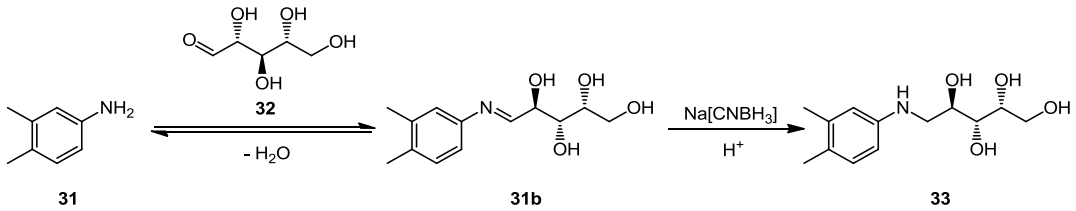


*Figure 26: Reaction between 3,4-dimethylaniline **31** and D-ribose **32** led to Schiff-base intermediate **31b**, which was subsequently reduced to give 3,4-dimethyl-N(ribityl)-aniline **33***

The alkyl-DOB reaction describes the formation of the imine intermediate **31b**, in which the nucleophilic amine attacks the electrophilic carbonyl position of D-ribose followed by a proton transfer. Water is subsequently eliminated across the newly-formed C-N bond, resulting in Schiff-base intermediate **31b**, which is reduced by cyanoborohydride to give 3,4-dimethyl-N(ribityl)-aniline **33**. As originally described, this reaction is performed as a "one-pot" reaction (where all reactants are added together at the beginning of the reaction).

It was noted that upon solvent removal and acidification of the resultant residue prior to extraction, the addition of ethyl acetate triggered spontaneous precipitation of high-purity 3,4-dimethyl-N(ribityl)-aniline **33**. Interestingly, this has not been observed or reported elsewhere[5,100], although similar precipitation of this compound had been observed in aqueous solution during an earlier synthesis of 5-deazariboflavin[109] (albeit during a different stage in the reaction). The majority of the 3,4-dimethyl-N(ribityl)-aniline **33** formed during the reaction was isolated as a precipitate, with the extraction protocol previously described used to ensure the greatest possible level of product recovery. However, in all cases this was far below those yields previously reported[5,100].

Due to the complex mechanism of this reaction, numerous factors may have been responsible for the observed reduction in yield, and so a wide survey of conditions was undertaken, which are summarised in tables 1-3 and described below. Factors examined include examination of Schiff-base intermediate formation (including the use of an organic catalyst), and whether performing the reaction using a "one-pot" method gave a lower yield than a multiple-step method. Also studied was the effect of a range of other solvents and conditions upon the reaction, including those in which reactants were insoluble, and solvents which were not anhydrous. Finally, replacement of the reducing agent was attempted, to determine whether cyanoborohydride was optimal.

As described in table 1 below, reaction scale was found to have little effect on the reaction. Reactions using 0.83 mM (100 mg) to 75 mM (9.0 g) of 3,4-dimethylaniline **31** were performed successfully in yields up to 65 %. Time had a marginal effect on reaction efficiency, with longer times (72 hours) than previously reported (48 hours) making little difference. However, significantly extended reaction times (96 hours) did appear somewhat detrimental to product formation, although no immediate cause was found. Heating at reflux was necessary, with no product formation observed when the reaction was performed at room temperature.

Scale (mM)	Solvent	Temperature	Conditions	Atmosphere	Time (hours)	Yield (%)
12.5	MeOH	reflux		Ar/N_2	48	57
25	MeOH	reflux		Ar/N_2	48	65
25	MeOH	reflux		Ar/N_2	72	64
75	MeOH	reflux		Ar/N_2	72	63
0.83	MeOH	reflux		Ar/N_2	72	65
0.83	MeOH	reflux		Ar/N_2	96	48
0.83	MeOH	RT		Ar/N_2	72	0
0.83	MeOH	reflux	Undried solvent	Air	72	22
0.83	MeOH	reflux	AcOH catalyst, 0.2 eq.	Ar/N_2	24	32
0.83	MeOH	reflux	AcOH catalyst, 0.2 eq.	Ar/N_2	48	19
12.5	MeOH	reflux	DMAP catalyst, 0.2 eq.	Ar/N_2	72	39

*Table 1: Conditions used to examine formation of ribitylated product **33**. "Scale" refers to the quantity of 3,4-dimethylaniline **31** used. "RT" refers to room temperature, approx. 15 °C. Each reaction was performed using 2 equivalents of both D-ribose and the reducing agent $Na[(CN)BH_3]$. All solvents were thoroughly dried before use (unless explicitly stated), as described in section 5.1.3*

As the original reaction conditions were found to give the greatest yields, an appraisal of the reactivity of the reagents was undertaken. The amine group of aniline **31** has an aryl substituent, and so the lone pair of electrons from the nitrogen atom (residing in a p-orbital) is delocalised throughout the aromatic π-bonding system. This lowers the electron density of the amine, thus reducing its nucleophilicity; the pKa of 3,4-dimethylaniline is 5.28[114] to 5.29[115], whereas the pKa of the structurally similar (non-aromatic) amine cyclohexylamine is 10.64[116,117]. Therefore, initial attack at the carbonyl carbon (C1) of D-ribose by 3,4-dimethylaniline **31** would be expected to be poor compared to other nucleophiles, and may be assisted by catalysis.

2.1.1.2 Effects of catalysis

The use of acid catalysis was initially considered, as this is suggested by the mechanism. Acidification was expected to activate the carbonyl carbon (C1) of D-ribose by protonation of the carbonyl-oxygen, assisting attack of the weakly nucleophilic amine. However, results obtained using a catalytic volume of glacial acetic acid (table 1) demonstrated a decreased in the yield of product **33** (32% in 24 hours), and which decreased further with extended reaction times (19 % in 48 hours). This result is interesting, as the reaction appears to reach its maximum efficiency significantly faster than the uncatalysed reaction; however this was at expense of a substantially reduced yield. Shorter reaction times did not show complete consumption of starting materials (by TLC) before 24 hours.

The likely reason for this decrease is that the presence of H^+ in the reaction mixture also caused protonation of the amine, further decreasing its nucleophilicity and preventing its reaction with D-ribose. Even though only a catalytic volume (0.2 equivalents) of acid was used, this appeared sufficient to significantly affect the outcome of the reaction. This result, although somewhat understandable, is somewhat contradictory to both the reaction mechanism and other reported results for similar reductive aminations[105,118,119], where addition of a catalytic volume of acid is necessary for efficient reaction. However, the effect observed experimentally has been described by Abdel-Magid *et al.*[105]. They reported during a thorough review of reductive aminations that whilst the presence of acid in a reaction between an amine and a ketone generally increased the rate of reaction, in the case of an amine reacting with an aldehyde, addition of AcOH to slow reactions such as those with weak nucleophiles tended to lead to a significant (approximately 50%) reduction in product formation. This mirrors the result obtained experimentally; although while Abdel-Magid *et al.* suggest the principal by-product of the reaction is the reduced aldehyde[105], the major products from this reaction were unreacted aniline **31** and a complex mixture of polymerised sugars.

Addition of a base to the reaction was also attempted, in order to deprotonate the reactive amine and thus enhance its nucleophilicity. However, a catalytic amount of

DMAP added to the reaction drastically decreased the formation of compound **33**, similarly to addition of AcOH. Reaction time was related to product formation as with the non-acidified mixture, with no product observed at 24 hours and little at 48 hours; 72 hours gave a yield of only 39 %, which was not improved using longer reaction times.

2.1.1.3 Schiff-base formation

Another structural feature of the starting material **31** was the presence of two methyl groups at positions *meta* and *para* with respect to the amino group. The effect of alkyl substituents of aromatic rings are to activate positions *ortho* and *para* to themselves, while deactivating *meta* positions, by enhancing the electron density of these positions. This is discussed in greater depth in section 2.3.1, where repetition of the reaction using an aniline containing a single methyl-substituent at the *para* position (*p*-toluidine **49**) gave some improvement to the reaction yield (72 %, compared to 65 % previously). This evidence suggested that the presence of the *meta* methyl group in 3,4-dimethylaniline **31** plays some part in reducing the efficiency of the reaction, as it is likely to effectively cancel-out the activating effect from the methyl group *para* to the reactive amine group.

While aromatic delocalisation played a large role in the poor nucleophilicity of the amine, it was hoped that it would assist formation of **33** by enhancing the stability of the Schiff base intermediate **31b** (containing an extended conjugated pi-electron system), thus promoting its formation. However, the Schiff base formed in this reaction was not possible to isolate, suggesting a transient existence. Late addition of the reducing agent (including reactions containing acid catalyst, previously shown to decrease reaction times) also failed to furnish improved yields; 51 % was obtained without acidification, compared to 23 % with acidification (31 % and 7 %, respectively, when repeated using $Na[(AcO)_3BH]$).

Scale (mM)	Reducing Agent	Conditions	Time (hours)	Yield (%)
25	$Na[CNBH_3]$		48	65
0.83	$Na[(AcO)_3BH]$		48	19
0.83	$Na[CNBH_3]$	Reducing agent added after 4 hrs	48	51
0.83	$Na[CNBH_3]$	AcOH added at 0 hrs; reducing agent added after 4 hrs	24	23
0.83	$Na[(AcO)_3BH]$	Reducing agent added after 4 hrs	48	31
0.83	$Na[(AcO)_3BH]$	AcOH added at 0 hrs; reducing agent added after 4 hrs	24	7
0.83	$Na[CNBH_3]$	$MgSO_4$ added, 5 eq	48	64
0.83	$Na[(AcO)_3BH]$	$MgSO_4$ added, 5 eq	24	28
0.83	$Na[(AcO)_3BH]$	$MgSO_4$ added, 5 eq	48	54

*Table 2: Conditions used to examine formation of ribitylated product **33**. "Scale" refers to the quantity of 3,4-dimethylaniline **31** used. Each reaction was performed using 2 equivalents of both D-ribose and the reducing agent described, under an atmosphere of Ar / N_2 in thoroughly dried refluxing MeOH*

2.1.1.4 Effect of water on the reaction

One of the key steps in all reductive aminations is the loss of water to produce an imine (or Schiff base) intermediate. The use of anhydrous conditions is typical[105,118,120], assisting the reaction by forcing the equilibrium to the right (towards product formation). For the reaction between 3,4-dimethylaniline **31** and D-ribose **32**, methanol was originally used as the solvent. It was able to fully dissolve each of the starting materials, and withstand sufficient heating (under reflux conditions) to allow the reaction to proceed. While MeOH is a good solvent for the reaction, its high polarity allows the dissolution of relatively large quantities of water (175 ppm for "wet" methanol, compared to 22.4 ppm for "wet" dichloromethane[121]). Since elimination of water to form the Schiff base is integral to formation of product **33**, the importance of anhydrous conditions was examined.

As originally described[5], there is no mention of whether the reaction should be performed under anhydrous conditions; however, it was found that performing the

reaction using undried solvent reduced the yield to 22 % (table 1, *vide supra*). Subsequent reactions were therefore performed in flame-dried glassware under an Ar or N_2 atmosphere using pre-dried solvent. Methanol was dried by distillation over CaH^{122} and storage over molecular sieves (3 Å) for a minimum of 5 days, as it has been reported that this reduced the residual water content to below 10.5 ppm[121]. Solvent prepared in this manner gave optimal yields of product **33**.

As the use of undried solvent greatly reduced product yield, experiments were performed to examine if the removal of water eliminated upon formation of the Schiff base played any role on reaction yield. Due to the small quantities involved (1 equivalent of water is equivalent to 0.015 mL when performed at scales of 0.83 mM), the use of specialised equipment such as a Dean-Stark apparatus for the constant distillation of the reaction solvent was considered to be of negligible benefit. Instead, the use of desiccants was examined.

$MgSO_4$ was selected due to its particularly fast-acting nature when compared to other drying agents, along with its ability to sequester up to 7 equivalents of water[123]. However, it was found that addition of 5 equivalents of $MgSO_4$ to the reaction mixture (performed with cyanoborohydride) gave a yield of 64 % – identical to the yield obtained without desiccant. This result suggests that $MgSO_4$ does not affect the reaction, and that the small volume of water generated by the reaction was adequately solubilised by the bulk solvent. However, results when using $Na[(AcO]_3BH]$ were more interesting, as yields obtained with addition of a similar quantity of $MgSO_4$ significantly improved the reaction yield (from 19 % to 54 %, table 2), and is discussed further in section 2.1.1.6 below. The presence of water within the $Na[(AcO]_3BH]$ used for experimentation was discounted (as freshly purchased reagent gave the same result). As the use of these conditions was still similar or less efficient than the original conditions using cyanoborohydride, they were not explored further.

2.1.1.5 Solvent effects

Another factor investigated was the effect of the solvent on the reaction. As described above, methanol was used in the original method[5], although a number of alternative

solvents were examined to determine if the yield would be improved when using them. These results are summarised in table 3.

Scale (mM)	Reducing Agent	Solvent	Temperature	Conditions	Time (hours)	Yield (%)
25	$Na[(CN)BH_3]$	MeOH	reflux		48	65
0.83	$Na[(CN)BH_3]$	EtOH	reflux		68	44
0.83	$Na[(CN)BH_3]$	isopropanol	reflux		72	49
0.83	$Na[(CN)BH_3]$	CH_2Cl_2	reflux		48	8
0.83	$Na[(CN)BH_3]$	$CHCl_3$	reflux		72	33
0.83	$Na[(CN)BH_3]$	MeCN	RT		48	0
0.83	$Na[(CN)BH_3]$	MeCN	reflux		48	8
0.83	$Na[(CN)BH_3]$	MeCN	reflux	AcOH catalyst, 0.2 eq.	24	0
0.85	$Na[(CN)BH_3]$	THF	RT		72	0
0.85	$Na[(CN)BH_3]$	THF	reflux		48	12
0.85	$Na[(AcO)_3BH]$	THF	RT		72	0
0.85	$Na[(AcO)_3BH]$	THF	reflux		48	18

*Table 3: Conditions used to examine formation of ribitylated product **33**. "Scale" refers to the quantity of 3,4-dimethylaniline **31** used. "RT" refers to room temperature, approx. 15 °C. All reactions were performed using 2 equivalents of D-ribose and the reducing agent under an Ar/N_2 atmosphere, using thoroughly dried solvents (section 5.1.3)*

THF, dichloromethane and chloroform each failed to dissolve the reactants at reflux, which may have contributed to the generally poor yields observed. Although acetonitrile solubilised the compounds at reflux, yields obtained were lower than those when using both THF and chloroform. EtOH was in turn slightly less effective than isopropanol, with both alcohols affording reasonable yields of product **33** – although MeOH remained optimal.

These results indicated the importance of a protic, polar solvent for this reaction. Aprotic solvents (both polar and non-polar) were generally unable to afford reasonable yields for the reaction, although the yield obtained using chloroform was somewhat surprising.

2.1.1.6 Reducing agent

In several cases, the reducing agent (sodium cyanoborohydride, $Na[CNBH_3]$) was replaced the similar compound sodium triacetoxyborohydride ($Na[(AcO)_3BH]$). This is a similar mild reducing agent, and has been widely used in other reductive aminations as a direct replacement for cyanoborohydride[105]. This is due to its significantly reduced toxicity (as there is no risk up hydrogen cyanide release upon hydration or acidification), and it is reportedly more stable to common reaction conditions (such as extended periods at reflux)[105], making it more appropriate for a "one-pot" synthesis where the intermediate is not isolated before elimination of water.

However, the yields obtained using this reducing agent were consistently lower than those achieved using cyanoborohydride, with the only exception to this found in the presence of $MgSO_4$. The reason for the general decrease observed using triacetoxyborohydride may be found in the extensive study of reductive aminations reported by Abdel-Magid *et al.* [105] using this reducing agent, where over 120 reductive aminations are described. While almost all were improved by replacement of cyanoborohydride for triacetoxyborohydride, they report that for reactions using unreactive or aryl amines, cyanoborohydride does give improved yields over triacetoxyborohydride (although the reasons for this remain unexplored). This is generally in line with the experimental results obtained.

2.1.1.7 Summary – formation of 3,4-dimethyl-N(ribityl)-aniline 33

A number of potential modifications to the original reaction method were attempted, although none has given any improvement to reaction yield. A range of factors was examined, including the use of catalysis, alternative solvents, different reducing agents, and the benefits of a "one-pot" method over separate additions of reagents. The main effect observed was the influence of water, although the use of dried

solvents appeared sufficient to prevent this. Therefore, as no yield for the formation of 3,4-dimethyl-N(ribityl)-aniline **33** was obtained above 65 %, questions must be raised regarding the yields reported previously[5,100].

2.1.2 Formation of 6-chlorouracil 35

The synthesis of 6-chlorouracil **35** was originally performed according to the procedure described by M. Mansurova *et al.*[100]. This was a simple hydrolysis of 2,4,6-trichloropyridine **34** performed using aqueous NaOH, producing 6-chlorouracil **35** in moderate yield (isolated by precipitation and purified by recrystallisation). However, yields for this reaction were often poor, with significant potential for degradation via tautomerisation or other routes. Therefore, a full survey of reaction conditions was performed. A number of trends were observed, summarised in table 4 (*vide infra*), and which are discussed below.

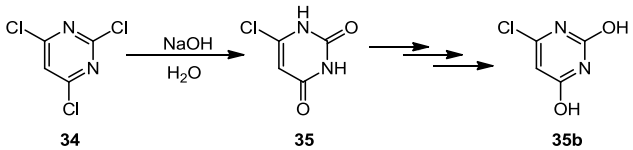


*Figure 27: Hydrolysis of 2,4,6-trichloropyrimidine **34** to form 6-chlorouracil **35**, which degraded quickly in aqueous solution to a range of products, including tautomer **35b***

Extension of the reaction times (from 12 to 72 hours) under similar conditions gave a demonstrable increase in reaction yield. However, with greater reaction times (96 hours), a drastic fall in the percentage of isolated product was observed, which is attributed to the re-aromatisation of the pyrimidine ring by tautomerisation, or degradation of 6-chlorouracil in the highly basic environment; numerous unidentifiable products were recovered, but not analysed further due to instability and difficulty in their separation. At high concentrations of base (discussed separately later), the effect of durations longer than 48 hours was negligible, with little enhancement to reaction yields obtained at 72 hours, due to faster initial reaction.

	Scale (mM)	Solvent	Base	Concentration (M)	Time (hours)	Reaction Temperature	Yield (%)
1	2.5	H_2O / HO^-	NaOH	2.0	12	100 °C	35
2	35	H_2O / HO^-	NaOH	2.2	48	100 °C	65
3	82	H_2O / HO^-	NaOH	2.5	48	100 °C	52
4	10	H_2O / HO^-	NaOH	2.0	72	100 °C	56
5	22	H_2O / HO^-	NaOH	5.0	72	100 °C	71
6	8.5	H_2O / HO^-	NaOH	2.6	96	100 °C	12
7	1.7	H_2O / HO^-	NaOH	1.0	48	100 °C	4
8	1.7	H_2O / HO^-	NaOH	3.0	48	100 °C	39
9	1.7	H_2O / HO^-	NaOH	5.0	48	100 °C	68
10	1.7	H_2O / HO^-	NaOH	5.0	48	RT	0
11	1.7	H_2O / HO^-	KOH	1.0	48	100 °C	20
12	1.7	H_2O / HO^-	KOH	3.0	48	100 °C	37
13	1.7	H_2O / HO^-	KOH	5.0	48	100 °C	52
14	1.7	H_2O / HO^-	KOH	5.0	48	RT	0
15	1.7	H_2O / HO^-	$Mg(OH)_2$	3.0	72	100 °C	0
16	3	toluene; H_2O / HO^-	NaOH	2.5	72	100 °C	21
17	8.2	DCM; H_2O / HO^-	NaOH	2.5	72	40 °C	14
18	10	H_2O / HO^-	NaOH	5.0	48	100 °C	*64,* ***81****

*Table 4: Conditions examined for the formation of 6-chlorouracil **35**. "Scale" refers to the quantity of 2,4,6-trichloropyrimidine **34** used. "RT" describes room temperature, approx. 15 °C. Reaction 18 gave a yield of 64 % when the product was isolated by precipitation alone; further extraction using EtOAc (described below) allowed recovery of additional product, increasing the yield to 81 % (combined)*

Temperature was found to have a broadly linear influence on the reaction. At ambient temperatures (typically 15-20 °C), no product formation was observed, while at reflux (100 °C) the reaction successfully produced 6-chlorouracil **35**. Higher temperatures were not examined due to the need for a hydroxylic base for the reaction; while it was

found that from the biphasic reaction using dichloromethane and NaOH (where a temperature of 40 °C was used due to the volatility of DCM), a modest yield of 6-chlorouracil **35** was obtained. If solvent effects are discounted, this suggests that increasing temperature lead to an improvement to reaction yield, with elevated temperatures allowing most efficient reaction.

The reaction solvent in most cases was aqueous NaOH, which acted as both solvent and reactive base for the hydrolysis. However, as 2,4,6-trichloropyrimidine **34** is only sparingly soluble in refluxing hydroxide, performing the reaction in organic solvents (toluene and dichloromethane) was attempted. Each of these reactions was performed as a biphasic mixture with the aqueous base; however even with vigorous stirring low yields were obtained. Examination of these mixtures using TLC indicated that 2,4,6-trichloropyrimidine **35** was generally localised within the organic solvent (particularly in the case of toluene), potentially reducing product yield because of poor mixing between the two phases. It was hoped that as 6-chlorouracil **35** may be more soluble in polar solvents, it would be retained within the aqueous layer after formation; however it was observed to re-dissolve in both organic layers, indicating a previously unreported affinity for these solvents. The use of a phase-transfer catalyst[124–126] to improve mixing was considered but discounted, as 6-chlorouracil **35** often formed large aggregated structures within the mixture (*vide infra*), which physically prevented effective mixing of the layers.

A further trend was noted, with an increase in hydroxylic base concentration (up to 5 M) facilitating a greater yield of 6-chlorouracil **35**. Several alternative hydroxylic bases were considered, including KOH and $Mg(OH)_2$. However, the poor solubility of $Mg(OH)_2$ in water led to no product formation after 72 hours. KOH was successful at furnishing 6-chlorouracil **35**, especially at lower concentrations (1 M) – where a yield of 20 % was obtained, compared to the yield produced using NaOH under the same conditions being 4 % – although at high concentrations (5 M), the effectiveness of KOH was significantly below that of NaOH.

It was noted during this reaction was that in almost all cases, an insoluble plaque was observed to precipitate out of solution above the reaction solvent after 12-18 hours. Mass spectrometry and NMR (^{1}H and ^{13}C) data suggests that this was comprised primarily of 6-chlorouracil **35**, although this was highly contaminated with other materials (such as a mass of 163.95 and characteristic isotope pattern indicating an intermediate dichlorinated compound). Manual disruption of this plaque allowed greater contact with the solvent, and it slowly re-dissolved over approximately 8 hours. The formation of this plaque, and its slow re-dissolution suggests high stability of this aggregated structure.

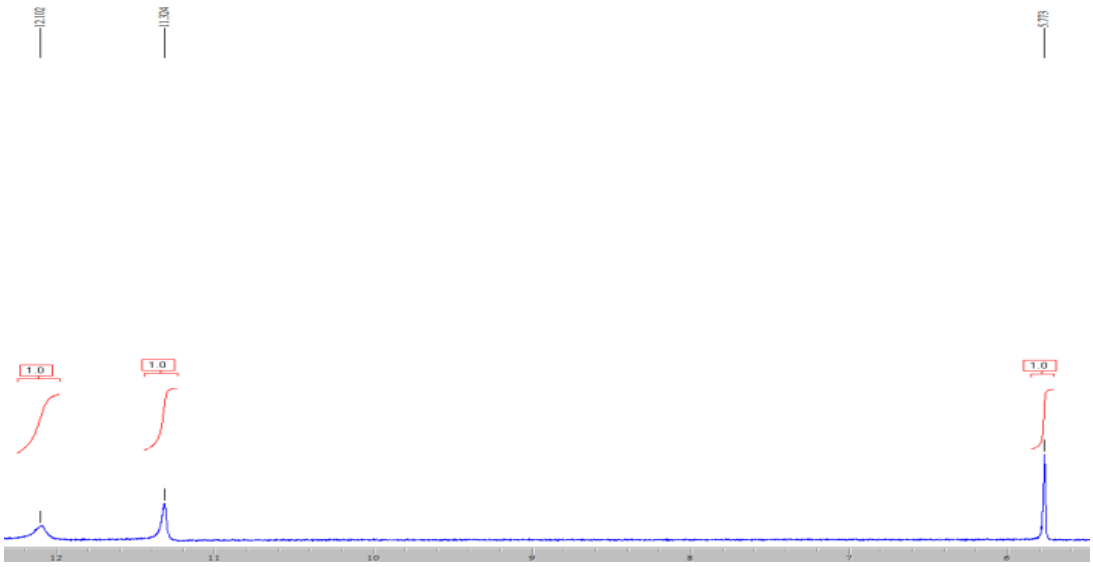


*Figure 28: ^{1}H NMR of 6-chlorouracil **35**, visualised using Mestrelab MESTREC[127]. Resonances for each N-H are visible (12.10, 11.32 ppm), with the single C-H resonance at 5.77 ppm. No lower resonances were observed for this compound*

One probable influence on the formation of 6-chlorouracil **35** is the strong hydrogen bond donor-acceptor motif of the uracil ring. Uracil is a nucleobase present in RNA (similar to thymine found in DNA), forming a hydrogen-bond pair with adenine. However, as uracil has both hydrogen bond donor and acceptor moieties, uracil-uracil mismatches are possible, with the stability of this inter-molecular effect utilised to by nature to stabilise elements of RNA secondary structure[128–130] (as shown in fig. 29). This intra-molecular effect is also likely to occur with free uracil in solution, minimising entropy by forming aggregated structures.

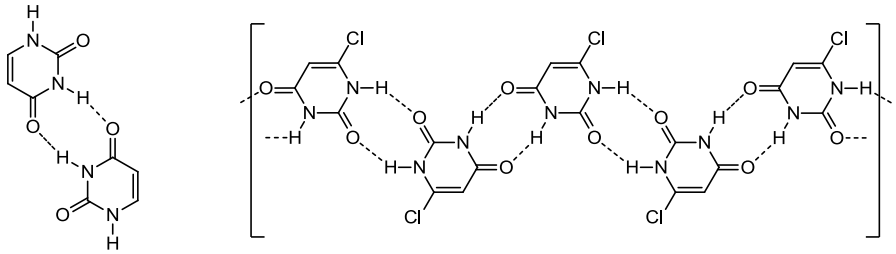


Figure 29: Mismatch hydrogen-bonding of uracil (left), as found in secondary structures of RNA [128–130]*. This may potentially extend between several molecules, leading to an aggregated "polymeric" structure (right) for 6-chlorouracil* ***35****.*

The strong influence of uracil's hydrogen bond donor-acceptor motif is further enhanced by a number of potential resonance structures, with each having a differing effect on the overall aromaticity of uracil[131]. The ionic effect of these resonance structures would be expected to improve the strength of bonding between molecules of uracil, supporting potential aggregation of the compound. When the effect of the high electronegativity of the chlorine withdrawing electron density from the uracil ring is considered for 6-chlorouracil **35**, it may stabilise some charged resonance forms, further enhancing the aggregation of the compound.

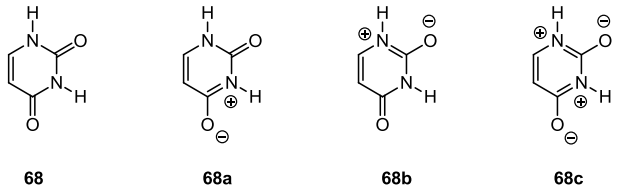


Figure 30: Uracil ***68****, with resonance-contributing structures as described by Galvão et al.*[131]*. Note the increasing level of aromaticity of structures* ***68a****-****c***

Upon isolation, 6-chlorouracil **35** was found to be particularly unstable, with near-total degradation to a complex mixture of products observed after 7 days at RT. To prevent degradation before reaction with 3,4-dimethyl-N(ribityl)-aniline **33**, 6-chlorouracil **35** was originally used immediately upon precipitation. However, it became clear that the compound retained quantities of water, with a detrimental effect on the subsequent

reaction. Drying 6-chlorouracil **35** under high vacuum for 8 hours improved the yield of the subsequent reaction from between 0-30% to 55%.

In order to find a swifter method to remove water from 6-chlorouracil **35**, azeotropic distillation of the product (from EtOH or toluene, approximately 100 mL per 1 g crude product) gave a similar improvement to the subsequent reaction as drying under high vacuum, although the use of elevated temperature was not ideal as this may promote tautomerisation.

An alternative method to isolate 6-chlorouracil **35** from the reaction mixture was discovered by TLC examination of the mother liquor from the precipitation, revealing significant quantities of the product remaining in solution. Trituration using hexane, diethyl ether, toluene or other solvents was ineffective, and the lyophilised product was found to be particularly impure. However, extraction of the aqueous solution with EtOAc allowed isolation of 6-chlorouracil **35** from the mixture, enhancing the overall yield by up to 17 %. As this previously unreported method was both convenient and efficient at furnishing pure 6-chlorouracil **35**, the reaction method was adjusted to reflect this. The use of extraction using EtOAc (following filtration if precipitation upon acidification was observed), gave 6-chorouracil in yields up to 81 %.

2.1.3 Formation of bicyclic intermediate 36

The coupling of 3,4-dimethyl-N(ribityl)-aniline **33** and 6-chlorouracil **35** is one of the steps of this synthesis which has been particularly contentious in the literature. At least three distinct routes (with numerous variations) have been described; each stating that their method is an improvement over previous routes. It is difficult to imagine that the weakly nucleophilic ribitylated nitrogen will readily attack 6-chlorouracil, and thus the yields reported for this reaction have remained moderate. Isolation of the intermediate bicyclic product **36** is difficult due to the similar polarities of the product, intermediates and reagents; however, acetylation of the bicyclic intermediate **37** immediately after reaction allows purification by FC.

Due to the competing reactions described in the literature, each of the three methods to perform this coupling was performed in order to determine the optimal route.

2.1.3.1 Catalysis using malononitrile

Initially, the method of Carlson and Kiessling (72 hours, refluxing MeOH, malononitrile catalyst, anhydrous conditions) was repeated, which furnished product **36** in low yield (35 %, after acetylation). This reaction was found to be particularly sensitive to the presence of water; when non-dried solvents were used, only starting materials were observed (by TLC) within the reaction mixture after 72 hours, and bicyclic intermediate **36** was not obtained. As described above, ensuring that 6-chlorouracil **35** was thoroughly dried gave a marked improvement to the yield of this reaction, reaching 55 % (after acetylation).

The addition of malononitrile appears necessary for effective reaction; very low yields (10-15 %) of product **36** were obtained when the reaction was performed without this compound. Addition of malononitrile was first reported by Carlson and Kiessling[5], who justify its use by its influence on a previous reaction[132], where malononitrile assists the coupling of conjugated vinyl halides. The mechanism of this malononitrile-catalysis has not been discussed in the literature, however similar reactions of malononitrile may shed light on its role in this case.

Some evidence is available that malononitrile may act as a reasonable nucleophile (assisted by palladium catalysis), especially in the displacement of aryl halogens[133,134] (fig. 31). As cyanide is an effective leaving group during nucleophilic substitutions (often preferentially to halogens)[134], it may be reasonably expected that di-nitriles such as malononitrile will be an even more effective leaving group. Thus in formation of **36**, malononitrile may activate 6-chlorouracil, allowing reaction with poor nucleophile of 3,4-dimethyl-N(ribityl)-aniline **33**.

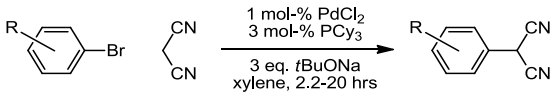


Figure 31: Scheme for reaction between a substituted aryl bromide and malononitrile, described by Schnyder et al.[133]*. "R" describes several aromatic substituents (electron withdrawing and donating) at various positions of the aromatic ring. The resultant aryl malononitriles were synthesised in yields up to 92 %*

One difficulty with the use of malononitrile is an apparent instability to the reaction conditions; optimal results were obtained when portions were added regularly throughout the course of the reaction. When monitoring by TLC, the reaction was found to be incomplete after the described 72 hours, with each of the starting materials visible up to 120 hours. Typically, the reaction was allowed to continue for 144 hours before work-up to maximise product formation. Notably, a characteristic blue fluorescence was observed for the bicyclic intermediate **36** when examined by TLC at 365 nm, due to significant extension of the chromophore.

An attempted modification to this reaction was the addition of 1.2 eq. of a base (either triethylamine (TEA) or 4-dimethylaminopyridine (DMAP)), to aid deprotonation the aryl amine of 3,4-dimethyl-N(ribityl)aniline **33**. This was intended to enhance the amine's nucleophilicity, and to remove acid generated by the reaction. However, in both cases no enhancement to the yield was made; yields decreased to 34 % and 21 % respectively (isolated after acetylation; principal side-products observed were acetylated starting materials).

Thus, the principal conclusion from the study of this method is that the reaction was particularly susceptible to both the presence of water or base, and should be performed with exclusion of both.

2.1.3.2 Solvent-free coupling

An alternative route used to form bicyclic intermediate **36** utilised a high-temperature solvent-free heating method to achieve formation of the bicyclic product[108]. This simple method required heating of reactants **33** and **35** (with vigorous mixing) at

185 °C for 15 minutes, followed by washing with diethyl ether and water to remove unreacted starting materials, furnishing the bicyclic compound in reported yields up to 75 % [108].

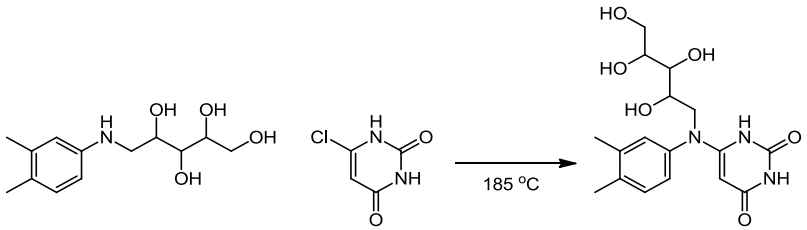


*Figure 32: Solvent-free heating of 3,4-dimethyl-N(ribityl)aniline **33** and 6-chlorouracil **35** afforded formation of intermediate **36***

Evolution of acidic vapours indicated a reaction between the two compounds. It was found that when performed at scales below 0.75 mmol, quantitative yields of bicyclic compound **36** were obtained. However, when performed using greater quantities of reactants, product formation was inefficient and frequently failed to give any isolatable product. This may be due to the instability of the ribityl group of 3,4-dimethyl-N(ribityl)aniline **33** at high temperatures, indicated by formation of a dark-brown caramel-like residue, and shifts in the ^{1}H NMR of peaks assigned to this functional group from compound recovered from the reaction. Alternatively, the reduction in yield may be due simply to the difficulty of efficiently mixing large quantities of the reagents without solvent. Therefore, while this method appears to show some promise, its application to preparative-scale production of the bicyclic intermediate **36** was not possible in its current form.

2.1.3.3 Reaction in aqueous solution

The third method which was attempted in order to facilitate the coupling between 3,4-dimethyl-N(ribityl)aniline **33** and 6-chlorouracil **35** was first reported in 1978 by W. T. Ashton *et al.*[109], with a variation reported in contemporary literature in 1997 by Frier *et al.*[135]. The original method describes performing the coupling reaction in refluxing water, producing **36** with a yield of 57 %; while the modern interpretation (using a similar 3,4-dimethyl-N(hexyl)-aniline, and using a 50:50 mixture of water and dioxane) reports a yield of 63 %. During the course of this project these reactions were

repeated using 3,4-dimethyl-N(ribityl)-aniline **33**, however in all cases no product formation was identified, with starting materials (or degradation products) recovered after 72 hours.

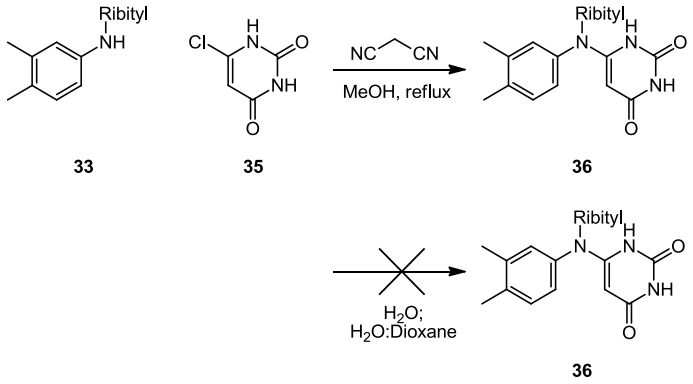


Figure 33: While formation of intermediate 36 was possible in anhydrous methanol, it was not possible to form the product in aqueous solution or solvent mixtures

The apparent failure of the reaction under these conditions was somewhat expected, as it was previously found in the case of the malononitrile-catalysed coupling described above that anhydrous conditions are crucial for effective reaction between the two reagents. Even the small quantities of water retained within 6-chlorouracil **35** from its prior precipitation had been found to have a dramatic effect on the yield of bicyclic compound **36**. Further work may assist in identifying the reasons for this.

2.1.3.4 Summary

The study of these three methods to perform the coupling between 6-chlorouracil **35** and 3,4-dimethyl-N(ribityl)-aniline **33** confirmed the use of the solvent-free method to perform the reaction in quantitative yield on an analytical scale, although it was not possible to use this reaction to furnish bicyclic intermediate **36** on a preparative scale, possibly due to insufficient mixing of the reagents. The method reported by Carlson *et al.*[5] performed in MeOH using malononitrile catalyst was found to be optimal, and was therefore used to prepare the bicyclic intermediate, in yields up to 55 % (including the subsequent acetylation).

After solvent removal from the coupling reaction, all material was used for the subsequent protection, typically without prior purification. This was originally attempted by recrystallisation (as described by Carlson and Kiessling[5]), or using FC over silica. However, product isolation was not possible by recrystallisation (as unreacted starting materials continually co-crystallised); although FC was possible on small scales, for spectroscopic purposes.

2.1.4 Protection of bicyclic intermediate 36

The protection of the D-ribityl group was performed to prevent reaction during the subsequent ring-closure using $POCl_3$. Almost all previous syntheses of riboflavin analogues use the same protection method described by Carlson *et al.*[5], where the ribityl hydroxyl groups were protected by acetylation.

2.1.4.1 Potential alternative protection strategies

The use of other protecting groups was considered, with the structural arrangement of the hydroxyl groups (1,2,3,4-relative to each other) leading to the interesting possibility of using cyclic protecting groups to protect the two pairs of diols, either 1,2 or 1,3 to each other. However, using these would have been difficult for several reasons. Firstly, the presence of both 1,2- and 1,3-diols within the ribityl group would cause complications during the reaction, leading to a mixture of different products with many not fully protecting all four hydroxyl groups of the ribityl chain. Secondly, while some cyclic protecting groups are resistant to harsh acidic conditions, due to their stability strongly basic conditions are often required for their removal – but riboflavin is particularly unstable in basic conditions[136–138].

Furthermore, other cyclic protecting groups such as dioxolanes alter the stereochemistry of the original hydroxyl group upon formation[139]. While in most cases stereochemistry is recovered upon deprotection, due to the importance of the stereochemistry of the D-enantiomer of ribose for recognition and binding by flavoproteins, damage to this functionality was particularly unattractive.

A previous method to perform the ring closure of the bicyclic compound which also resulted in the protection of the ribityl chain was using trimethylorthoformate[109], shown in figure 34 and discussed previously in chapter 1.2.6. This was unsuitable, as the bis(methoxymethylene) protecting groups are easily removed under acidic conditions, and the reported yield for the ring-closure reaction using trimethylorthoformate was 47 % [109]**;** significantly lower than the 64 % obtained by Carlson *et al.*[5] using $POCl_3$ / DMF.

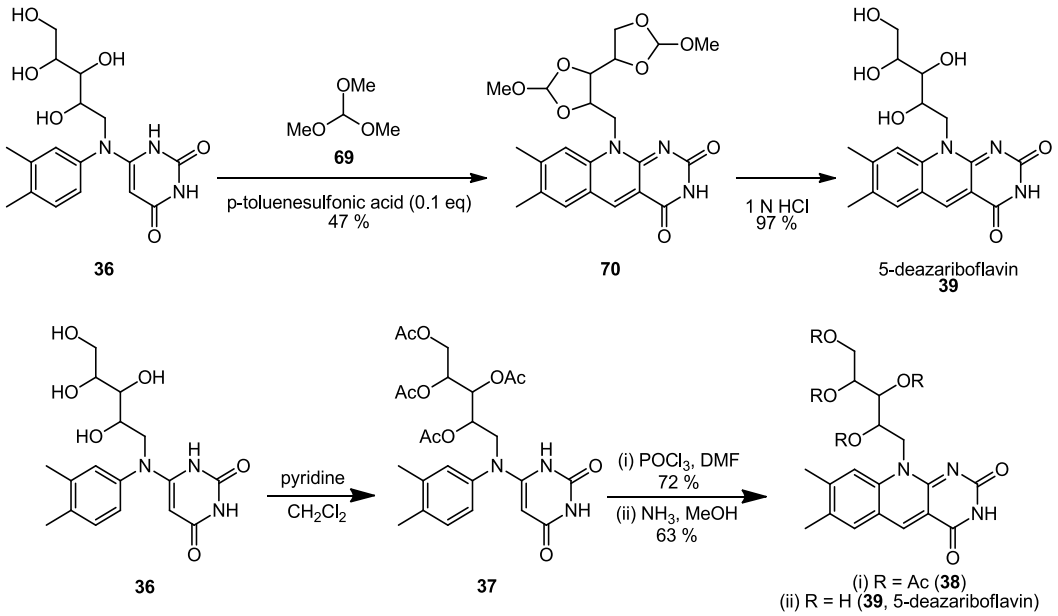


*Figure 34: Triethylorthoformate **69** successfully incorporated bis(methoxymethylene) protection for the ribityl side-chain of intermediate **36** while also allowing closure of the central pyridine ring of 5-deazariboflavin **39** (above), however this was less efficient than acetylation followed by the Vilsmeier-Haack method (below)*

After acetylation, purification of **37** by FC was much more convenient, using the slightly polar mixture of $CHCl_3$ and MeOH in a ratio of 20 : 1 ($^{v}/_{v}$). This was significantly less polar than that required to separate the starting materials for the reaction, as the polarity of the ribityl chain is vastly decreased by acetylation. Therefore, while protection of the ribityl chain was useful to prevent unwanted reaction in subsequent steps, it also allowed convenient purification of product **37**.

2.1.4.2 Examination of protection conditions

Protection of bicyclic compound **36** was originally performed using acetic anhydride in pyridine, acting as both solvent and catalyst[139,140]. As each of the four hydroxyl groups required protection, a 1.25 molar excess (5 equivalents) of acetic anhydride was used, which is not greater than that reported for numerous other acetylations[139]. However, even after FC and numerous rounds of extraction into dichloromethane, it was found that pyridine often contaminated the product, making purification and assignment (particularly by NMR) difficult. Therefore, numerous alternative conditions were examined, as described in table 5.

	Scale (mmol)	Bulk Solvent	Catalyst	Acetate source	Temp.	Reaction Time (hrs)	Yield (%)
1	6.0	Pyridine	Pyridine	acetic anhydride	RT	16	47
2	0.5	CH_2Cl_2	Pyridine	acetic anhydride	RT	16	55
3	0.5	CH_2Cl_2	-	acetic anhydride	RT	24 (incomplete)	15
4	0.5	CH_2Cl_2	TEA	acetic anhydride	RT	16	24
5	0.5	CH_2Cl_2	DMAP	acetic anhydride	RT	16	31
6	0.5	CH_2Cl_2	Imidazole	acetic anhydride	RT	16	41
7	4.5	CH_2Cl_2	Pyridine	acetic anhydride	RT	16	54
8	0.5	THF	Pyridine	acetic anhydride	RT	16	45
9	0.4	THF	TEA	acetic anhydride	RT	2	36
10	0.4	THF	TEA	acetic anhydride	RT	16	41
11	0.5	CH_2Cl_2	Pyridine	acetyl chloride	RT	4	*0 **
12	0.5	CH_2Cl_2	Pyridine	acetyl chloride	0 °C	4	*0 **
13	0.5	CH_2Cl_2	Pyridine	acetyl chloride	-78 °C	4	*0 **

*Table 5: Reactions performed to establish the optimal protocol for acetylation of intermediate **36**. "Scale" refers to the number of moles of riboaniline **33** used for the initial formation of bicyclic **36**, formed according to the method described in section 6.1.3.1; all yields reported were obtained over two steps. 0.2 equivalents of catalyst were used in all cases. 5 equivalents (1.25 molar excess) of acetic anhydride was used, while 4.1 equivalents (1.03 molar excess) of acetyl chloride were used*

Without the catalytic effect of pyridine, the acetylation of bicyclic intermediate **36** by acetic anhydride (in dichloromethane) was incomplete after 24 hours, while the

presence of 0.2 equivalents of pyridine (again using dichloromethane as the bulk solvent) gave optimal yields after 16 hours. Reactions were not performed for greater than 16 hours, due to the formation of undesired side-products (principal among these was N-acetylated 3,4-dimethyl-N(ribityl)-aniline, indicating rupture of the crucial xylidine-uracil bond, which has been reported to be acid-sensitive[5]).

Using dichloromethane as the reaction solvent was particularly beneficial, as the acetylated product was readily soluble in the organic solvent. While dissolution of the starting material required a solvent with higher polarity than dichloromethane, the catalytic volume of pyridine required for efficient reaction was sufficient to allowing dissolution of the starting material within the bulk solvent. Yields obtained using dilute pyridine in dichloromethane were in fact greater than those obtained when using pyridine as a lone solvent (55 % vs. 47 %), and this effect applied on all scales (from 0.4 mmol to 4.5 mmol), with little difference.

Furthermore, this method allowed convenient purification by extraction against aqueous solutions, due to the greatly reduced volume of base to be removed. It was further found that this extraction process enabled isolation of the desired product in good purity (as unreacted materials from previous steps were also removed to the aqueous layer), without requiring chromatographic procedures such as FC, thus saving a considerable amount of time.

As described above, it was found that a small volume of base (0.2 molar equivalents) was sufficient to catalyse the reaction and allow effective acetylation. However, a number of other organic bases (triethylamine, DMAP and imidazole) were evaluated against pyridine in order to determine if a more effective replacement was possible, as each has been widely reported to assist acetylation reactions[141]. Each base was used as a direct replacement for pyridine in the reaction, although some problems with solubility occurred; both pyridine and triethylamine each allowed the starting material to dissolve in the reaction solvent, while the solid bases DMAP and imidazole had no effect. In these cases, the solvent was gently warmed to allow dissolution of the starting material, and the reaction was started as a slight suspension if necessary.

A further complication was that the use of both imidazole and DMAP required lengthy purification by FC after the reaction to remove the catalyst.

The use of stronger bases has been reported to reduce the time required for acetylations; in an example described by Klemenc[142], DMAP was found to be able to catalyse acetylation of morphine (to heroin) in under 10 minutes at room temperature, in contrast to 2 hours at reflux using pyridine. However, a similar effect was not observed in this case. In contrast, a general trend inversely related to the strength of the base was found: while pyridine is the weakest base (pKa 5.2), the yield obtained using this was greatest; as pKa of the base increased, the yield of product decreased. This may be due to the stability of the active intermediate species (in the case of reaction with pyridine, intermediate **71a**, figure 35), which dissociates upon attack by the group undergoing protection. Although stronger bases would be expected to react more readily with acetic anhydride to form the activated species, the increased stability of the intermediate makes it more difficult for the ribityl group to react. This result therefore suggests that the nucleophilicity of the group to be protected must be balanced against the strength of base used, which may form a more stable intermediate.

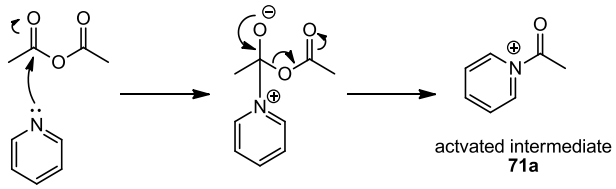


*Figure 35: Reaction between pyridine and acetic anhydride forms the intermediate species **71a**, which assists acetylation of the ribityl-hydroxyl groups*

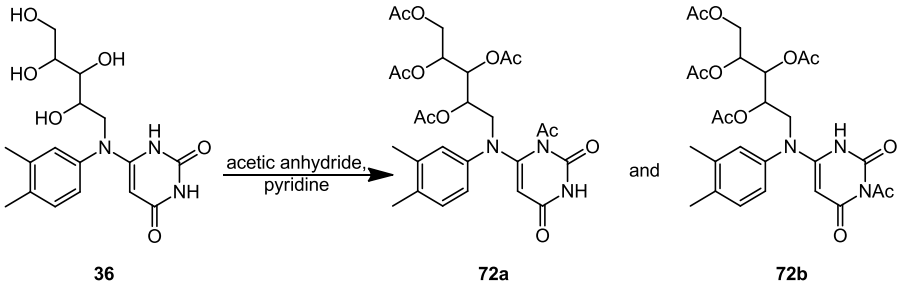
As an alternative to the use of dichloromethane, THF was examined in its place, where it was found to give yields only slightly lower those obtained previously. Interestingly, while triethylamine appeared to be less effective than pyridine in dichloromethane (24 % after 16 hours in dichloromethane, vs. 55% using pyridine); using THF, triethylamine gave 41 % after 16 hours, and 36 % after only two, which is very close to the yield obtained using pyridine in THF (45 %). A probable reason for the differing

yields when using dichloromethane and THF is the increase in polarity of THF relative to dichloromethane, which more adequately solubilised the reactants and facilitated a fast initial reaction. However, the overall reaction yield remains inferior to the yield obtained using dichloromethane/pyridine, supporting the supposition that the stronger base is less effective as a catalyst.

After determination of the optimal conditions for this reaction, it was repeated using the (significantly more reactive) reagent acetyl chloride. However, even using the minimum quantity (4.1 equivalents) at reduced temperatures (-78 °C), and with careful monitoring throughout the course of the reaction, the formation of a wide variety of unwanted side-products was observed without formation of protected product **37**. Principal among these was evidence of acetylation of the amine group at the expense of either the ribityl side-chain or the uracil moiety, with a high proportion of N-acetylated derivative of 3,4-dimethyl-N(ribityl)-aniline again recovered. Therefore, the use of acetyl chloride was not explored further, and acetic anhydride in a mixture of dichloromethane and pyridine was used to perform this reaction.

2.1.4.3 Undesired Acetylation

As described above, five equivalents of acetic anhydride were used to protect the four hydroxyl groups of the D-ribityl side chain. However, during careful examination of the spectroscopic data, it was apparent that on some occasions, a previously unreported side-product was formed, with potentially damaging consequences for the synthetic route. In a number of cases, the slight excess of acetic anhydride led to additional acetylation of the uracil ring. Initial examination by NMR did not show a preference for acetylation of either amino group of the uracil moiety, with a mixture of both generally observed. This led to a more complex NMR structure of the acetylated compound **37**, which was incredibly difficult to separate from the undesired side-products by FC due to their very similar polarities.



36 **72a** **72b**

*Figure 36: Acetylation of the ribityl side-chain of bicyclic compound **36** often led to acetylation of the uracil moiety. No preference for either amine position was noted*

The removal of acetyl protecting groups from amines is notoriously difficult[139], due to a strong conjugation effect. However, (as discussed in greater detail later) the conditions used for the final step of the reaction were found to be sufficient to overcome this, and thus the presence of undesired acetylation did not affect the outcome of subsequent reactions.

2.1.5 Ring closure at position C5 of the isoalloxazine

2.1.5.1 Previously reported methods

The insertion of a carbon atom at position C5 of the flavin ring is key to the final structure of 5-deazariboflavin, but is one of the most complex steps of the synthesis. Previous syntheses of 5-deazariboflavin utilised alternative strategies to afford formation of 5-deazariboflavin; either using triethylorthoformate[109] (discussed above), or with a starting material already containing a functionalised carbon at this position (2-amino-N(ribityl)-4,5-dimethylbenzaldehyde **73b**, fig. 37)[103,106]. The presence of a reactive carbonyl group required protection to prevent unwanted reaction during the course of the synthesis, which was performed using a tosylated hydrazide group, removed immediately prior to coupling with 6-chlorouracil[103] (or an analogue[106]), to form a 5-deaza-flavin product.

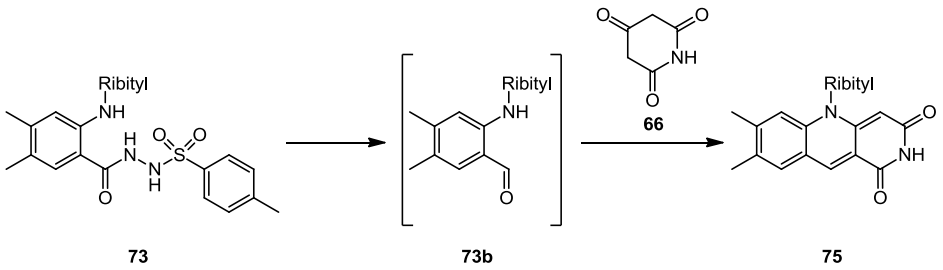


*Figure 37: Scheme for reported formation of 1,5-dideazarboflavin **75** from protected 2-amino-N(ribityl)-4,5-dimethylbenzaldehyde **73**[106]. This is deprotected in situ, before addition of 2,4,6-piperidine trione **66**; condensation furnished the flavin analogue*

The insertion of the eventual C5-carbon of 5-deazariboflavin used for this reaction was originally devised by F. Yoneda *et al.* in 1976 [111]. This reaction performed the ring-closure of the central ring of isoalloxazine using the Vilsmeier-Haack reagents $POCl_3$ and DMF[110]. $POCl_3$ activates DMF, which is then attacked by an electrophilic substrate (generally an arene[143–145]). DMF acts as both a reactant and the reaction solvent.

2.1.5.2 Position of Formylation

The formyl group formed by this reaction undergoes a secondary attack by either the neighbouring xylidine or uracil ring (*vide infra*), which closes the central pyridine ring of the isoalloxazine analogue; final elimination of water allows aromaticity across the flavin system.

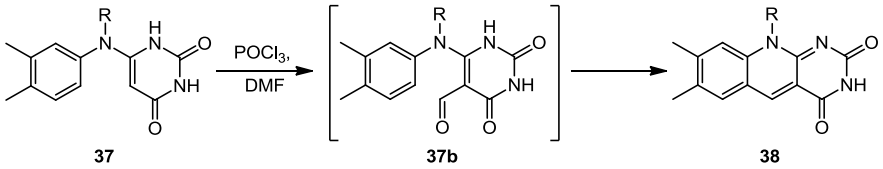


*Figure 38: Formylation of intermediate **37** (R = D-ribose tetraacetate) using the Vilsmeier reagents ($POCl_3$ and DMF); attack by the xylidine ring allows formation of the 5-hydroxyl-deazaflavin intermediate (not shown), leading to 5-deazaflavin **38***

Interestingly, a strong argument is made by Yoneda *et al.*[111] for the initial formylation of the uracil moiety of bicyclic compound **37**, rather than formylation of the xylidine ring. This is somewhat contrary to initial expectations, as the principal reported use of the Vilsmeier-Haack reaction is the formylation of aromatic molecules[143–145] (although

this is disputed by some sources which report that this reaction fails with benzenes[146]). On the other hand, the formylation of non-aromatic alkenes using these reagents is widely reported, particularly when activated by neighbouring (or conjugated) electron donating groups[147–149].

The mechanism of attack by an activated alkene is the same as that performed by an arene ring, especially when enhanced by an electron donor. Key amongst the evidence presented by Yoneda *et al.*[111] is that the formylation of 6-chlorouracil (to 6-chloro-5-formyluracil) prior to reaction with 3,4-dimethyl-N(ribityl)-aniline **33** successfully produced 5-deazariboflavin **39** in moderate yield. Hence, upon bond formation between the amine and chlorinated position of uracil, the formylated uracil moiety must be susceptible to an intramolecular attack by the xylidine ring, to close the central ring of the flavin analogue. The pre-formylation of uracil was not considered during this study, as the yields obtained by Yoneda *et al.* for this reaction are substantially lower than those obtained without prior formylation of 6-chlorouracil (from 57 % to 37 %[111]). However, it may be useful to re-examine this method to confirm this result, as it may significantly simplify this synthesis.

2.1.5.3 Relative Positions of the Methyl Groups of the Xylidine Moiety

In order to furnish 5-deazariboflavin **39**, an important consideration of the relative geometry of the xylidine moiety of bicyclic intermediate **37** is necessary. Rotation around the C-N bond originating from 3,4-dimethyl-N(ribityl)-aniline **33** changes the position of methyl groups relative to position of C5 in the final 5-deazaflavin structure. The two positions of the aniline *ortho* to the amino bond are non-equivalent due to the neighbouring methyl groups, and so two potential products may have been formed by this reaction – the desired 5-deazariboflavin **39** (i.e. 7,8-dimethyl), and 5-deaza-6-methyl-8-demethylriboflavin **76** (fig. 39).

However, the only reported product from this reaction was 5-deazariboflavin **39**, which is confirmed by spectroscopic evidence show in figure 40.

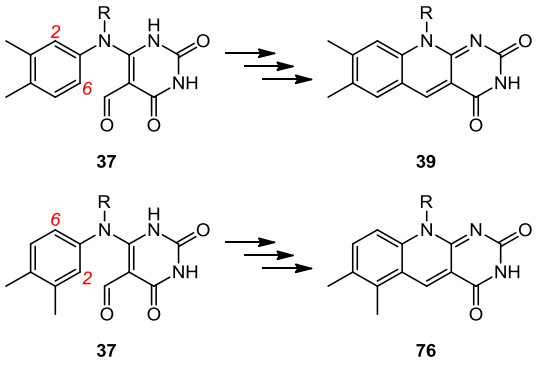


*Figure 39: Rotation around the C-N bond of bicyclic intermediate **37** offers two positions (2 or 6) of the xylidine ring for attack at the formylated uracil; the pro-5-deazariboflavin structure (above) and the pro-5-deaza-6-methyl-8-demethylriboflavin structure (below), however only 5-deazariboflavin **39** is observed experimentally*

Evidence confirming the formation of 5-deazariboflavin was obtained from the splitting pattern and coupling constants in the aromatic region of the proton NMR spectrum, shown in figure 40. The spectrum of 5-deazariboflavin **39** (fig. 40, black) shows no splitting of the proton resonances, with some possible peak broadening of the resonances at 8.05 ppm and 7.95 ppm indicating relative *para* coupling. Compared to the proton NMR spectrum of 3,4-dimethyl-N(ribityl)-aniline **33** (fig.40, blue), which shows relative *ortho*, *para* and *ortho+para* coupling (from left to right), it is clear that if compound **76** was formed, the aromatic resonances observed would be expected to each have doublet multiplicity, with a coupling constant of approximately 7 Hz.

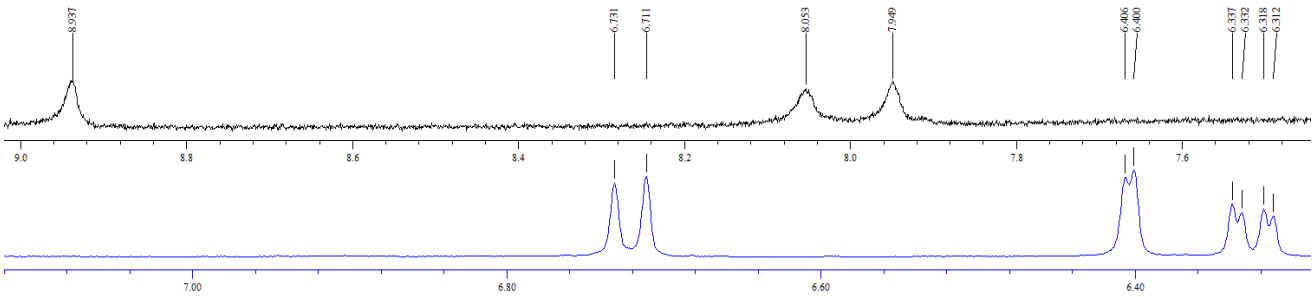


*Figure 40: Proton NMR spectra of 5-deazariboflavin **39** (black, above), and 3,4-dimethyl-N(ribityl)-aniline **33** (blue, below)*

There are several possible reasons for why the 5-deazaflavin form is preferred by the reaction, although these were not examined closely during this research. One likely

reason is the activation of position 6 of the xylidine ring (figure 39) by a donating effect from the methyl substituent at position 3. Similar activation has already been shown to play some effect amino group activation in the reductive amination of aromatic amines 3,4-dimethylaniline **31** and 4-methylamine **49**, discussed in sections 2.1.1 and 2.3.1. Additionally, the effect of delocalisation of electrons from the ribitylated amine across the aromatic ring gives this C-N bond considerable π-character, providing a high energetic barrier for rotation around this bond. Thus, the arrangement of methyl groups leading to 5-deazariboflavin formation may be favoured, and lead to the formation of the single observed product.

2.1.5.4 Isolation of 5-deazariboflavin tetraacetate 38

A key improvement to this reaction was made using an organic extraction procedure to isolate further product **38** (fig. 38, *vide supra*) from the mother liquor after precipitation. However, the first extraction also isolated unreacted staring materials; 5-deazariboflavin tetraacetate **38** was isolated by acidification (protonating the flavin analogue), which was separated from other contaminants by extraction into water. Deprotonation allowed re-extraction using $CHCl_3$, and isolation of further 5-deazariboflavin tetraacetate **38**. The yield obtained using this method was 72 %, which is greater than the yield obtained previously (64 % [5]).

2.1.6 Deprotection of 5-deaza(tetraacetyl)riboflavin 38

Deprotection of the acetylated flavin was conveniently performed using methanol saturated with ammonia. This was performed with care taken to exclude all light from the sample, generally using foil-lined glassware in the dark. It was found that the reaction proceeded quickly and efficiently at room temperature, with a typical reaction time of 6 hours. As both NH_3 and MeOH are volatile, these were easily removed at reduced pressures, along with ammonium acetate formed during the course of the deprotection.

While similar methods are reported elsewhere for the removal of acetyl protection from hydroxyl groups[139,150], one key question of this reaction is whether it was possible to remove the unintended amino- acetylation of the uracil moiety observed

from protection of the ribityl chain, described above. The stability of the acetamide group is generally exceptionally high, and stable to hydrolysis under acidic and basic conditions[139,151–154]. However, careful scrutiny of the spectroscopic data from this reaction showed the complete removal of all acetyl groups present before the reaction, forming 5-deazariboflavin **39**; therefore these conditions were suitable for hydrolysis of this bond.

2.1.7 Conclusion – Synthesis of 5-deazariboflavin

During this section, the optimal route for the synthesis of 5-deazariboflavin **39** was identified, and used to furnish the product in 16.2 % yield overall. Several reactions were found to be particularly difficult, and where necessary thorough surveys of experimental conditions and reagents were performed in order to better understand the reactions. One of the more intriguing results obtained has been the formation of a previously unreported acetylated intermediate during acetyl protection of the ribityl chain; however, it has been determined that this has no effect on subsequent reactions, and is readily removed using the standard deprotection procedure described in section 6.1.6.

Future work which may be of benefit to this synthesis would be the examination of the outcome for potential for the acetylation of 6-chlorouracil **35** prior to coupling with 3,4-dimethyl-N(ribityl)-aniline **33**, as there is some evidence in the literature[108,111] that this may give improved yields for the coupling reaction. Likewise, formylation of **35** prior to the coupling reaction may further improve the reaction efficiency, by potentially removing the need for protection, deprotection, or ring closure steps[111].

2.2 Synthesis of 1-deazariboflavin

2.2.0 Introduction

This section describes the method used to synthesise 1-deazriboflavin, following the original route of Carlson and Kiessling[5]. Few approaches for the synthesis of this analogue exist in the literature, and so this study has been performed to bolster the further the understanding of this reaction. In this section, the improvements made

and the rationale behind various methods is described, which has enabled overall yield to be improved from 11.0 % reported elsewhere[5] to 61.3 %.

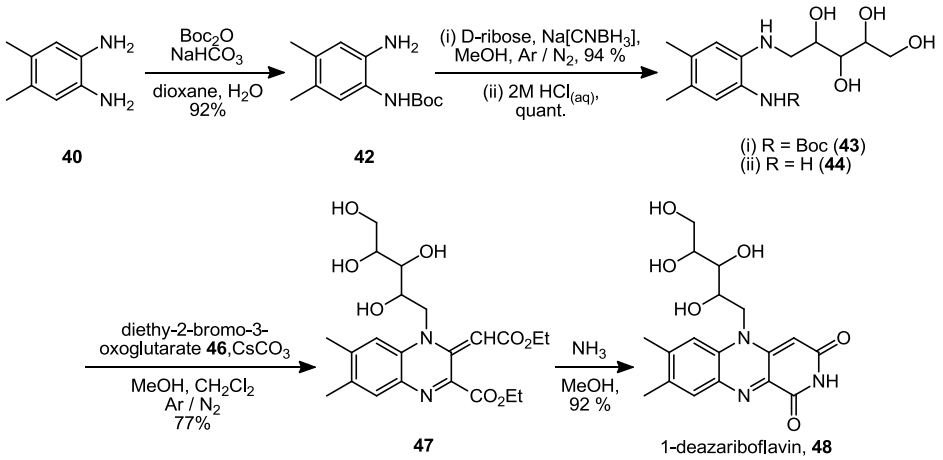


*Figure 41: Overall scheme for the synthesis of 1-deazariboflavin **48***

2.2.1 Protection of 2-amino-4,5-dimethylphenylamine 40

Mono-protected product **42** was produced in very high yields (92 %), in excess of those quoted previously (66 % [5], 90 % [100]). A side-product commonly reported[5] for this reaction is the protection of both amino groups, however this was not observed using the conditions reported in section 6.2.1. A strict stoichiometric amount of ^{t}-Boc anhydride was used, along with careful monitoring of the reaction using TLC, which showed completion at shorter times than previously reported (3 hours).

Commonly, protection of amine functionality using the Boc- group is performed with heating at reflux[139]; however the diamine used for this reaction reacts readily at room temperature. Cooling of the reaction mixture was considered in the event that di-protection was observed as reported[5], however as this was not the case it was not necessary. Purification of protected product **42** was performed using flash chromatography, with very good separation between the desired product and other reaction contents enabling swift purification.

2.2.2 Ribitylation of Boc-protected intermediate 42

The asymmetric protected diamine **42** was ribitylated using the most effective conditions found for synthesis of 3,4-dimethyl-N(ribityl)-aniline **33**, described in

section 2.1.1. However, the yield for this reaction was significantly improved, from 65 % previously to 94 % using the protected diamine, providing further evidence that the weak nucleophilicity of the nitrogen atom in 3,4-dimethylaniline **31** plays a significant role in poor yields obtained when undergoing reductive amination, as the donating effect of the protected *ortho*- amine of diamine **42** appears to considerably enhance the efficiency of the ribitylation.

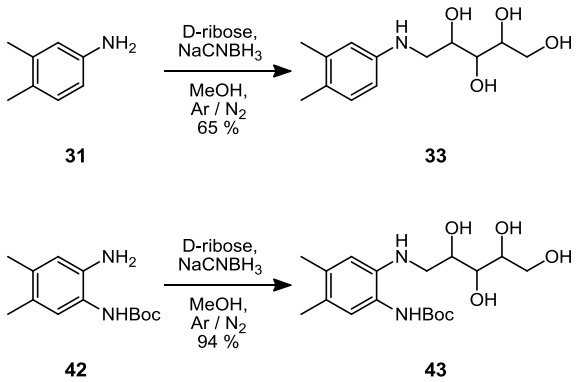


*Figure 42: Comparison of reductive aminations for 3,4-dimethylaniline **31** and Boc-protected diamine **42**. When performed* ceteris paribus*, vast improvement in the yield of **43** was found, suggesting an ortho-donating effect from the protected amine*

The greatest difficulty with this reaction was the lability of the Boc- protecting group of product **43** to the acidic conditions used in the reaction work-up. The use of acid is necessary to ensure complete reduction of the Schiff-base intermediate, however (as was found later), this group was successfully cleaved using 1 M hydrochloric acid. Therefore, this stage of the reaction was performed as swiftly as possible, in order to maximise product yield.

Furthermore, while the purification of 3,4-dimethyl-N(ribityl)-aniline **33** caused spontaneous precipitation of the pure product, no precipitation of the similar Boc-protected product **43** was observed during this reaction, believed to be due to the increase in solubility afforded by the protecting group. Final yields of the purified compound reached 94 %, which was reliably reproduced at all scales between 0.8 mmol and 50 mmol.

2.2.3 Deprotection of 1-(N-ribityl),2-(N-Boc)-diamino-4,5-dimethylbenzene 43

Removal of the Boc- protecting group was originally performed as previously reported, using 4 M HCl in dioxane. While other protocols for this deprotection typically use less concentrated conditions (2 M aqueous HCl)[139], the suggested reagent is generally used to only perform this reaction under oxygen-free or anhydrous conditions[139,155]. Interestingly, the synthetic method of Carlson makes no mention or justification for the use of air-free conditions.

A plausible reason for performing this reaction with the exclusion of oxygen was found in the results of Barry *et al.*[156] and Maggio-Hall *et al.*[157], describing the formation of 5,6-dimethylbenzimidazole (DMB) **44b** as a ligand for vitamin B_{12} biosynthesis. They report that this compound may be formed by a complex "oxidative cascade" from riboflavin, via the deprotected ribitylated diamine **44**, simply by aerobic oxidation[156]. This was also reportedly possible using catalytic amount of a mild oxidising agent, such as potassium ferricyanide[157]; providing a plausible explanation for performing the deprotection under air-free conditions.

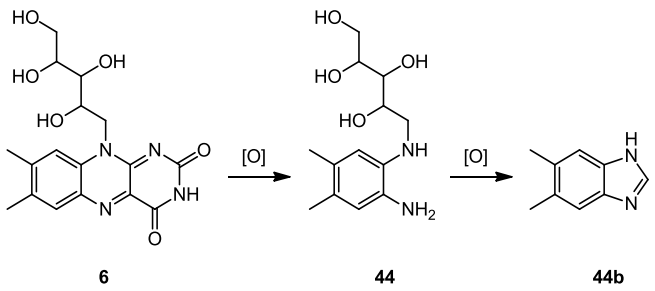


*Figure 43: Proposed formation of 5,6-dimethylbenzimidazole **44b** via an oxidative cascade[156,157] from riboflavin **6** and intermediate **44** (formed during this synthesis)*

The use of 4 M HCl in dioxane was reasonably efficient (giving yields up to 78 %), although slow, with reaction times of 16 hours necessary. As an oxygen-free environment was originally considered necessary, an alternative method used HCl (gas) generated alongside the reaction vessel, and bubbled through an anhydrous solution of the compound (in THF, MeOH or dioxane). This proved significantly more effective (with reaction complete after 30 minutes, forming the compound as the hydrochloride salt), although the use of strong acid was undesirable due to the

potential risk of undesired dehydration of the ribityl group. While this was not observed, the high risk of this led to examination of other deprotection strategies.

As the use of oxygen-free conditions limited the availability of methods to perform the deprotection, an examination into the degradation of deprotected diamine **44** to dimethylbenzimidazole **44b** was performed to assess to influence of atmospheric quantities of O_2. Deprotected amine **44** previously formed under anhydrous conditions was stirred at RT in a continually-aerated aqueous solution, and formation of DMB **44b** was indeed observed; however, this reaction was found to be particularly slow and ineffective (less than 10 % conversion to DMB **44b** at 96 hours). Therefore, as the degradation of intermediate **44** was found to be very slow, the use of simplified deprotection conditions was possible.

Principal among these was the use of 2 M aqueous HCl; using this reagent, it was possible to furnish a quantitative yield of deprotected amine **44** within two hours. Importantly, no formation of DMB **44b** was observed, and so this reaction step was used for subsequent reactions. Lower concentrations (1 M, 0.5 M) of the acid were also successful at facilitating deprotection, requiring an increase in reaction duration.

Purification of the product using this method was not performed, as product purity (assessed by NMR) was greater than 98 %. No justification for the previously described extraction (using Et_2O) could be found, as the products from Boc-deprotection are CO_2 and isobutylene – both volatile gasses, readily removed by lyophilisation. Upon evaporation of the ethereal layer, nothing was isolated; thus extraction was unnecessary, and no longer performed.

As discussed above, although the risk of oxidative degradation was found to be low, the subsequent coupling of intermediate **44** to diethyl-2-bromo-3-oxoglutarate **46** was performed immediately in order to mitigate any possible degradation.

2.2.4 Formation of diethyl-2-bromo-3-oxoglutarate 46

It was found that performing this reaction according to the original method of Carlson *et al.*[5] was ineffective. As with reaction 3.2.3 above, no explicit mention of anhydrous conditions was made – although the reaction was reportedly performed under a stream of CO_2. The reaction was initially performed as described, assuming anhydrous conditions, at 65° C.

Mass Spectrometry data for the product of this reaction indicated that a significant proportion of both mono- and di- brominated products was formed, which were immediately identifiable due to the characteristic isotope pattern of bromine (^{79}Br and ^{81}Br). The presence of di-brominated oxoglutarate **46c** (figure 44) may have caused difficulties during subsequent reaction steps, which is discussed in further detail below. Evidence suggesting formation of the di-brominated species alongside the desired mono-brominated species was also found for the original reaction conditions at room temperature and at 0 °C; therefore, the use of elevated temperatures did not assist the formation of the desired product.

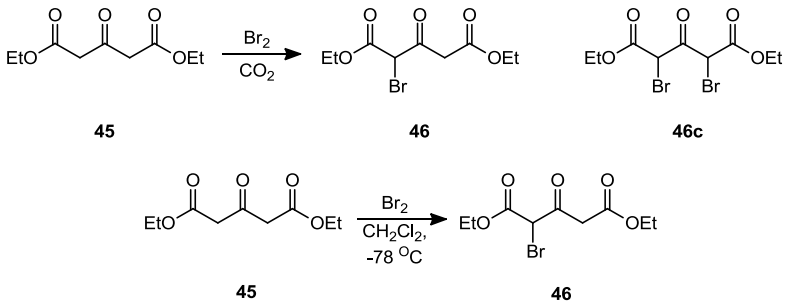


*Figure 44: Methods for the bromination of diethyl-3-oxoglutarate **45**. The original method (above) led to formation of a mixture containing mono and dibrominated compounds, with a potential detrimental effect on subsequent reaction steps*

An alternative method to produce brominated product **46** reported the use of strict anhydrous conditions in dried dichloromethane at room temperature[158]. Bromine was used as the determinant reactant (a large excess was used previously[5]), added dropwise over 10 minutes. Using this procedure, it was possible to produce brominated product **46** effectively; an acidic vapour (believed to be HBr) was observed when the flask was opened to the air, and allowed to freely disperse before

subsequent purification. This was performed by FC using a short plug of silica, allowing fast and reliable isolation of mono-brominated compound **46** in high yield (assessed by NMR due to short product lifetime). The reaction was further improved by cooling the mixture to -78 °C, before allowing gradual warming (to RT) of the mixture after 45 minutes. Using this method, no evidence of di-bromination was observed, allowing subsequent steps to be performed without the risk of undesired side-reactions.

2.2.5 Coupling of diethyl-2-bromo-3-oxoglutarate 46 and deprotected aniline 44

The efficient coupling of deprotected amine **44** and diethyl-2-bromo-3-oxoglutarate **46** was crucial to the success of this synthesis, but had previously been particularly difficult to achieve. The reaction proceeds by attack of the deprotected amino group of **44** at the bromine-functionalised position of glutarate **46**, which then undergoes intramolecular attack from the secondary (ribitylated) amine.

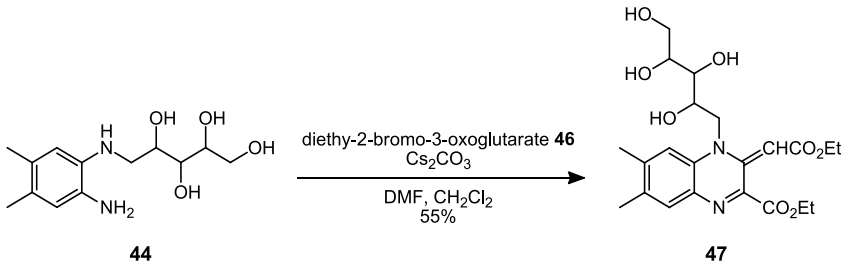


*Figure 45: Original method to produce intermediate **47** reported by Carlson and Kiessling[5], using intermediate **44** and diethyl-2-bromo-3-oxo-glutarate **46***

When this reaction was performed using brominated glutarate **46** formed using the original synthetic method (which also produced the di-brominated species **46b**), product **47** was obtained in poor yields (35 %). The presence of the dibrominated species allows several undesired reactions to occur, drastically affecting the formation of the intended product, and forming a number of difficult to separate side-products. However, when the reaction was repeated using purified bromoglutarate **46** formed by the improved technique described above (section 2.2.4) yields up to 55 % were obtained, in line with the original method[5].

Isolation of the final product **47** by extraction against an organic solvent was found to be particularly difficult due to the amphoteric nature of the product, facilitating

formation of an emulsion. The use of alternative solvents for extraction (including Et_2O, petroleum ether, CH_2Cl_2 or $CHCl_3$) made little difference; this problem was eventually overcame by addition of large quantities of saturated NaCl solution to the mixture, which assisted phase separation.

The effect of the base added to the reaction was examined using a range of alternatives, including several other carbonates ($NaHCO_3$, Na_2CO_3, K_2CO_3 and Cs_2CO_3), hydroxylic bases (NaOH, KOH), and organic bases (pyridine, imidazole and DMAP). Cs_2CO_3 remained the most effective base, with neither of the hydroxylic bases, imidazole or DMAP facilitating product formation. The most likely reason for this is that each of these is a more effective nucleophile than the relatively weak aromatic amine, with the stable intermediate formed preventing progress of the reaction.

Base	**Yield (%)**
Cs_2CO_3	55
Cs_2CO_3 + DMAP (0.2 eq)	22
K_2CO_3	34
Na_2CO_3	14
$NaHCO_3$	11
Pyridine	4
Imidazole	0
DMAP	0
NaOH	0
KOH	0

*Table 6: Yields obtained for the formation of **47** using alternative bases for the reaction between deprotected aniline **44** and diethyl-2-bromo-3-oxoglutarate **46***

While DMAP alone was ineffective for this reaction, it was also used alongside Cs_2CO_3 to assess any phase-transfer effect[124–126] between the immiscible solvent layers (DMF and CH_2Cl_2). Even in low concentrations (0.2 equivalents), the reaction yield was reduced to 22 %, from 55 % previously, demonstrating that this was not suitable for phase transfer. As each of the reactants appeared to reside preferentially in the separate solvent layers (**44** in DMF; **46** in CH_2Cl_2), the use of a solvent combination which allowed greater mixing was sought.

It was found that deprotected compound **44** was soluble in several polar solvents (DMF, H_2O, MeOH), while it was almost completely insoluble in dichloromethane or other organic solvents. Conversely, diethyl-2-bromo-3-oxoglutarate **46** was soluble in several organic solvents, but insoluble in polar solvents. A survey of numerous combinations of these solvents was performed, described in table 7 (below). It was found that an equivolume mixture of MeOH and dichloromethane was optimal, allowing isolation of compound **47** in 77 % yield; a significant improvement over the previous reported yield of 55 %. This is attributed to the slight miscibility of the two solvents, allowing improved contact between the two reactants and facilitating improved yields of product **47**.

Organic Solvent	Secondary Solvent	Yield (%)
CH_2Cl_2	DMF	55
CH_2Cl_2	MeOH	77
CH_2Cl_2	EtOH	44
CH_2Cl_2	H_2O	8
CH_2Cl_2	DMSO	0
CH_2Cl_2	Hexane	0
$CHCl_3$	MeOH	65
$CHCl_3$	DMF	39
Et_2O	MeOH	16
Et_2O	H_2O	11

*Table 7: Solvent mixtures used to examine the coupling of deprotected ribitylated aniline **44** and diethyl-2-bromo-3-oxoglutarate **46**. All reactions were carried out as described in section 6.2.5.2, using 250 mg (925 µM) of the determinant reactant. All solvents were used in equivolume ratios; the mixture of CH_2Cl_2 and hexane did not allow dissolution of starting materials, and was performed as a suspension.*

2.2.6 Ring closure of bicyclic intermediate 47, forming 1-deazariboflavin 48

This reaction was performed to close the isoalloxazine moiety of the flavin analogue at the 3- position, using NH_3 dissolved in MeOH. As the reaction required the formation of two amide bonds, it was expected to be particularly slow and ineffective,

with the method reported by Carlson and Kiessling reporting a yield of 33 % for their small-scale reaction (56 µmol of the starting material)[5].

While the initial bond formation using the original method was expected to be ineffective (*vide infra*), once one amide had been formed, the second (intramolecular) reaction to close the ring would be expected to occur much more readily. This is likely due to a much greater effective concentration of the amine, and is supported by evidence from other syntheses which report that intramolecular formation of amides is particularly favoured (with some reactions proceeding spontaneously[159]) when formation of a cyclic product occurs[140].

The creation of an amide bond is typically a difficult reaction to perform by synthetic chemistry approaches. Low reactivity between the reactant carbonyl and amine groups prevents reaction even under reasonably harsh conditions, which is ordinarily overcome by either (or both) the use of coupling agents to assist the reaction, or the use of reactants with significantly higher reactivity (such as an acyl chloride in place of a carboxylic acid)[140,160]. In biological systems, several enzymes exist to assist and promote amide formation (for instance during protein biosynthesis), which utilise carbonyl group activation to promote reaction.

Intermediate **47** contains two ethyl esters; upon amide formation, the ethanolate group of these is replaced by an amine. However, this leaving group is relatively poor compared to other functionalities; for amide formation, alkyl esters generally "cannot be considered as activated species" (stated by Montalbettli and Falque[140]). Of the few reliable conditions found in the literature to both remove the ester and facilitate effective amide formation, the use of Lewis Acid ($TiCl_4$) catalysis[161] allow both reactions to occur in a limited number of cases.

Based on the poor potential for this reaction, in concert with the previously reported low yields[5], it was suspected that in this case the reaction would remain ineffective. Therefore, the potential use of coupling agents was explored for utilisation if required. The formation of amide bonds (as peptides) in Solid-phase Peptide Synthesis (SPPS)

utilises coupling agents to allow reaction between a carboxylic acid and an amine of two amino acids, which undergo sequential reaction (with functionality protected by various protecting groups) with other amino acids to form peptides. These reactions are ordinarily performed using mild reaction conditions (so as not to interfere with side-groups or structural features), and so the use of a coupling agent such as HATU[162] is required to catalyse the reaction, giving efficient reactions at room temperature (even with sterically-hindered reagents)[140].

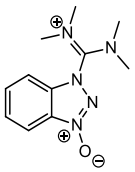
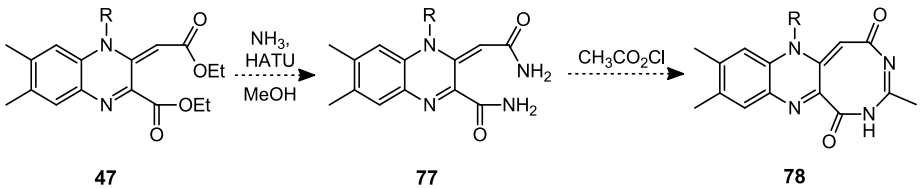


Figure 46: HATU structure, supplied as a salt (typically PF_6). This reagent allows the formation of a reactive ester with a (deprotonated) carboxylic acid, acting as a good leaving group and thus enabling attack of an amine to form an amide bond

While the use of a coupling agent such as HATU may have enhanced the formation of 1-deazariboflavin **48**, a potential problem was the presence of two ethyl esters in compound **47**. If each was to react to form the relevant amide, the isoalloxazine moiety would remain unclosed at position 3 (**77**, fig. 47). Ring closure would be unlikely to occur spontaneously, due to the high stability of the newly-formed amide bonds; although this method may allow formation of 1-deaza- (or other-functionalised) riboflavin analogues with an enlarged uracil moiety, such as compound **78** suggested in figure 47.



47 **77** **78**

*Figure 47: The use of a coupling agent to enhance formation of **48** may lead to the formation of diamide **77**, which (due to the stability of the newly-formed amide bonds) may not close at the 3- position to form 1-deazariboflavin **48**. However, this may instead allow the formation of ring-expanded isoalloxazine analogues by bridging the two amides, e.g. product **78**.*

It was found experimentally that it was possible to form 1-deazariboflavin **48** using the original reaction conditions, with extended reaction times (72 hours, vs. 48 previously) furnishing 1-deazariboflavin **48** in yields up to 51 %. While already a significant improvement over previous reported yield, observations made during the course of the reaction indicated that further improvement may be possible.

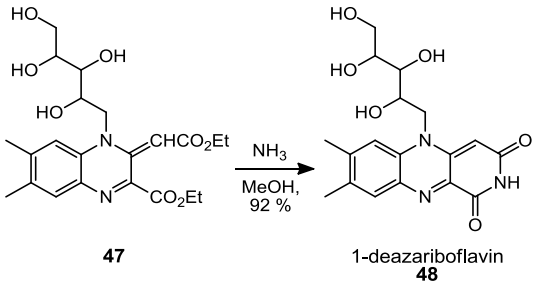


*Figure 48: Optimal reaction conditions for the formation of 1-deazariboflavin **48***

Originally, gaseous ammonia was dissolved in MeOH at -5 °C. However, it was found during formation of 1-deazariboflavin **48** that a significant proportion of the ammonia swiftly evaporated from the mixture, leaving little remaining by the end of the reaction. Instead, addition of further portions of solvent saturated with ammonia through the course of reaction improved 1-deazariboflavin yields up to 63 %, suggesting that concentration of ammonia may play a key part in the success of the reaction. The use of a dry ice / acetone bath (-78 °C) to reduce the temperature of the methanol further during dissolution of ammonia allowed condensation of the vapours (ammonia boils at -33 °C). A 1:1 ratio of ammonia in methanol was found to be possible, and as the mixture was gently warmed to RT, no significant immediate evaporation was observed. Naturally this did occur over time, and was replaced by addition of a further aliquot of solvent after 36 hours of reaction. In this manner, formation of 1-deazariboflavin **48** in yields up to 92 % was possible; a vast improvement over 33 % previously reported[5].

2.2.7 Conclusion – Synthesis of 1-deazariboflavin

The successful formation of 1-deazariboflavin **48** was performed, with an overall yield of 61.3%. This is far superior to other syntheses (11.0 % [5]; 21 % [100]), and was possible

due to the exhaustive study of each reaction, allowing significant optimisation, with two exceptional improvements identified.

Future work which may be performed based on these results is the examination of the final ring-closure reaction using several alternative nucleophiles in place of ammonia, such as alkyl- or other functionalised amines. If successful, this would allow the construction of a library of 1-deaza-3(N-functionalised)-flavin analogues, which may be of significant interest.

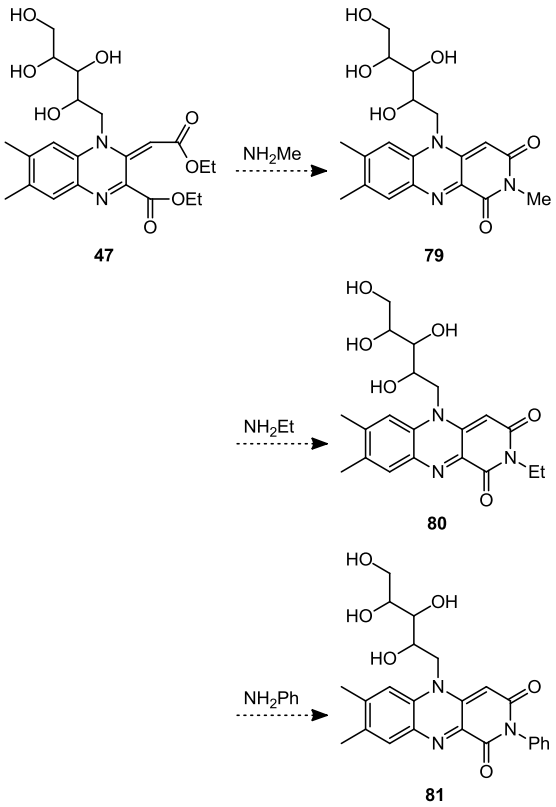


Figure 49: Potential derivatives of 1-deazariboflavin with substituted N(3) positions, formed using N-functionalised reagents

2.3 Synthesis of 8-demethyl-5-deazariboflavin

This analogue was synthesised using the same route described for 5-deazariboflavin above, using 4-methylaniline **49** (commonly referred to as *p*-toluidine) as the initial reactant in place of 3,4-dimethylaniline **31**. This had an interesting effect on several reaction steps, providing further evidence that the poor nucleophilicity of the aromatic amine of the anilines is largely responsible for the low yields observed.

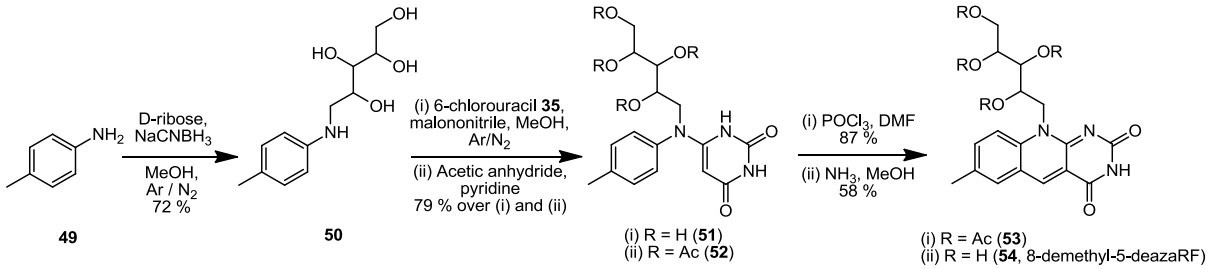


*Figure 50: Overall synthetic route to 8-demethyl-5–deazariboflavin **54***

2.3.1 Reductive amination of 4-methylaniline 49

This step was performed using the optimal conditions found during the similar reaction described in section 2.1.1. However, the yield for this reaction using the same method was 72 %; a slight (7 %) enhancement over the analogous reaction using 3,4-dimethylaniline **31**.

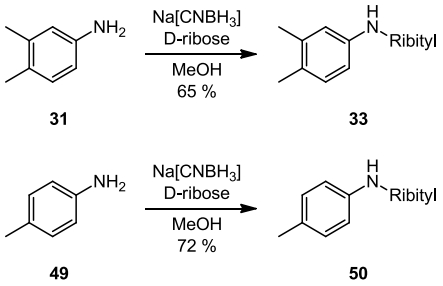


*Figure 51: Comparison in reductive amination of 4-methylaniline **49** and 3,4-dimethylaniline **31** indicated the influence of the meta methyl group of **31** reduces the yield of the reaction*

This result was expected, due to the effect of the methyl substituents of the aromatic ring affecting the nucleophilicity of the amino group of the aniline. For 3,4-dimethylaniline **31**, the methyl substituents are *meta* and *para* with respect to the reactive amine, whereas 4-methylaniline **49** only contains the *para* methyl group. As alkyl substituents of aromatic systems behave as electron donors, their presence activates positions of the aromatic ring which are *ortho* and *para* to themselves, making these positions more nucleophilic. On the other hand, *meta* positions are deactivated by this effect.

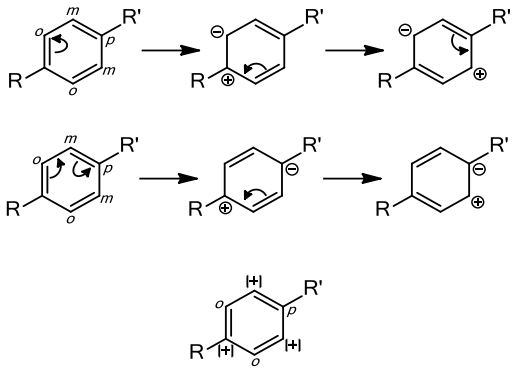


Figure 52: Schematic representation of the donating effect of an alkyl group (R) on benzene. Two resonance forms are possible, allowing localisation of electron density at positions ortho or para with respect to the donating group, and thus increasing the electron density available to a substituent (R') at these positions. Furthermore, positions meta to the donating group obtain a partial positive charge, reducing the nucleophilicity of substituents at these positions

For 3,4-dimethylaniline **31** the methyl groups present at positions *meta* and *para* with respect to the amino group leads to a conflict between activation and deactivation effects. While the methyl group at position 3 (*meta*) is deactivating to the amine, the methyl at position 4 (*para*) is activating. However, 4-methylaniline **49** contains only a *para*-methyl group, which confers the activating effect of the (electron donating) methyl group without the *meta*-deactivating effect; the nucleophilicity of the amine of 4-methylaniline **49** is thus slightly increased compared to 3,4-dimethylaniline **31**, allowing an improved yield for the reductive amination.

This may be further tested in future work using several aromatic molecules with substituents at alternative positions with respect to the reactive amine, such as 2-methylaniline, 3-methylaniline or phenylamine. If the above hypothesis is true, it would be expected that in each case a reduced yield of the ribitylated product would be obtained.

2.3.2 Coupling of ribitylated aniline 50 and 6-chlorouracil 35

As with the ribitylation described above, this reaction was performed using the optimal conditions found previously, in refluxing anhydrous MeOH (section 6.1.3.1;

discussed in section 2.1.3). 6-chlorouracil **35** used for this reaction was formed using the optimal conditions found previously, and isolated using the novel organic extraction method described in section 6.1.2.3.

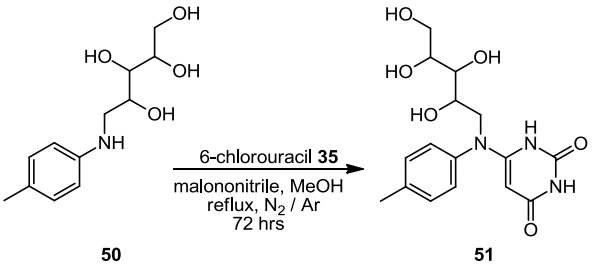


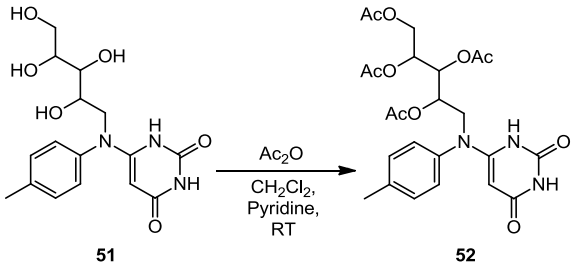
*Figure 53: Overview of the coupling of 4-methyl(N-ribityl)-aniline **50** and 6-chlorouracil **35***

This reaction proceeded in a very similar manner to the analogous reaction performed earlier, with formation of a fluorescent product (under irradiation at 365 nm), and completion identified by TLC at a similar reaction time. As before, bicyclic product **51** from this reaction was used without isolation or purification, as the amphoteric nature of the product caused considerable difficulty with purification by methods including organic / aqueous extraction and FC. As the subsequent step (protection of the ribityl functionality using acetyl protecting groups) reduced this effect, reliable purification was performed at that stage, allowing accurate assessment of the reaction yield over two steps.

2.3.3 Protection of ribityl side-chain of intermediate 51

The acetylation of intermediate product **51** was performed using a mixture of pyridine and dichloromethane, as described in section 2.1.4.2. However, a slightly greater excess of acetic anhydride was used (1.5 equiv. per hydroxyl group; 1.25 equiv. previously), to prophylactically ensure complete protection of the ribityl side-chain even in the case of acetylation of the uracil moiety (as had been previously observed). Although this additional acetylation would originally have been considered a problem (due to the notorious stability of acetamides), previous results (section 6.1.5) demonstrated that this would not affect the outcome of the subsequent ring closure,

and that it was possible to remove unwanted acetylation at the same time as deprotection of the ribityl chain (section 6.1.6).



51 52

*Figure 54: Acetylation of the ribityl side-chain of **51** using acetic anhydride*

The spectroscopic data for this reaction clearly showed no undesired additional acetylation of the uracil moiety, and so the yield of product **52** obtained after purification was 79 % (over two steps). This is a significant improvement over that obtained from the analogous reactions performed using 3,4-dimethyl-N(ribityl)-aniline **33**, which gave a yield of 55 %. As both reactions were performed *ceteris paribus* (with scale adjusted proportionally), this difference in yield is again likely to be due to the effect of the methyl substituents of the N(ribityl)anilines on the nucleophilicity of the reactive amine.

2.3.4 Formation of 8-demethyl-5-deaza-(tetraacetyl)riboflavin 53

The use of the Vilsmeier-Haack method was again used to close the central ring of the isoalloxazine analogue, using a mixture of $POCl_3$ and DMF to first formylate intermediate **52**, which underwent an intramolecular reaction to close the central ring of the flavin analogue. As discussed in section 2.1.5 above, the position of formylation is a source of debate within the literature; isolation or further study of the formylated intermediate species was not deemed to be possible due to a potentially transient existence (due to the reactivity of this group), and the hazards associated with $POCl_3$ necessitating neutralisation of this before chromatographic separation and analysis. Alternatively, formylation of the xylidine or uracil moieties prior to coupling may assist the understanding of this reaction (as a similar protected formylated intermediate was utilised during earlier syntheses of 5-deazariboflavin[111,112]); this is discussed alongside other potential future work in chapter 7.1.3.

The yield for this reaction was found to be 87 %, which is a significant improvement over the yield of 72 % obtained when using bicyclic intermediate **37** (fig. 55). This provides evidence to assist the understanding of the formylation reaction, discussed below.

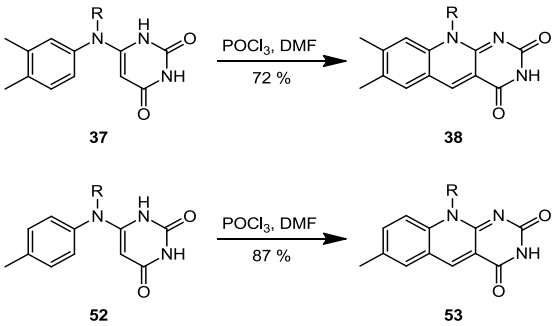


Figure 55: Ring closure of the isoalloxazine moiety of the 5-deazaflavin compounds. "R" denotes a D-ribityl(tetraacetate) group

As discussed for several reactions above, the relative arrangement of the methyl substituents of 3,4-dimethylaniline appear to play a significant role on the outcome of the reactions during this synthesis. While generally appearing to be detrimental to other steps of the reaction, one benefit of the 3-methyl group of intermediate **37** is that with respect to the desired position of formylation (position 6 of the aromatic ring), it is *para*, and thus activating. The increase in electron density at this position would be expected to assist reaction at this position with the formylation agent (or formylated intermediate), and may be key to ensuring the correct relative arrangement of the methyl groups in the isoalloxazine structure of 5-deazariboflavin, discussed in section 2.1.5.3.

In the case of the reaction performed to produce 8-demethyl-5-deaza-riboflavin, no group is present on the xylidine ring to direct formylation. If formylation of the aromatic moiety does occur (as opposed to the uracil moiety, *vide supra*), this may take place at either position 2 or position 3 of the symmetrical aniline. Attack by position 2 would lead to the desired product, while position 3 would lead to formation of an alternative undesired product. As the lone methyl group present (which is *ortho*-

, para- activating) would be expected to favour attack by position 3 of the aromatic ring to react with the formylating agent, this would lead to a low yield of the target compound.

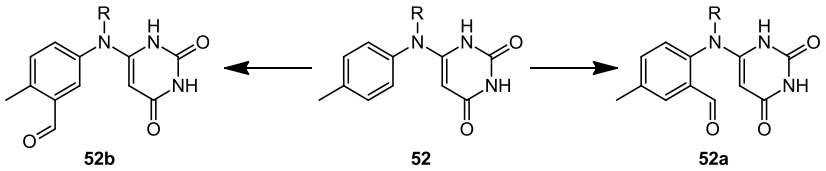


*Figure 56: Two structural isomers for the formylation of bicyclic intermediate **52** at the xylidine ring exist; only **52a** will lead to the formation of 8-demethyl-5-deazariboflavin **54**. Without the directing effect previously present when using the analogous compound **37** (*vide supra*), **52b** would be expected to predominate if formylation of the xylidine moiety occurred (rather than formylation of the uracil)*

As demonstrated by the spectroscopic data, the target isoalloxazine analogue **53** was indeed formed by this reaction. This result therefore implies that the formylation stage of the reaction initially takes place at the uracil ring, before ring closure via an intramolecular attack. While uncommon, some sources do report the formylation of non-aromatic alkenes using the Vilsmeier-Haack reagents, particularly when activated by neighbouring electron donating groups[147–149] (discussed in depth in section 2.1.5). Thus, the early results of Yoneda *et al.*[111] appear to be accurate, and potentially applicable to future syntheses of 5-deazaflavins.

Furthermore, one of the greatest enhancements to the method used to isolate 5-deazariboflavin tetraacetate **38** was the augmentation of spontaneously-precipitated product with further product isolated from the solution. This method was adopted for the 8-demethyl compound **54**, which was isolated using organic extraction before separation from starting materials and other contaminants by FC.

2.3.5 Formation of 8-demethyl-5-deazariboflavin 54

The deprotection of the acetylated intermediate **53** to form 8-demethyl-5-deazariboflavin **54** was performed using the same solvent mixture (NH_3 in MeOH) as

before. Again, extra care was taken to exclude light from the sample, due to the potential for flavin to be degraded under basic conditions when exposed to light.

As with the analogous reaction (section 2.1.6), this reaction proceeded quickly and efficiently at room temperature, with a typical reaction time of 6 hours furnishing the target compound in 58 % yield. This is in comparison to a yield of 63 % obtained during the synthesis of 5-deazariboflavin **39**, although it may be possible to improve slightly upon this, bringing the yield of **54** more closely in line with that of **39**.

2.3.6 Conclusions – Synthesis of 8-demethyl-5-deazariboflavin

The total synthesis of the novel riboflavin analogue 8-demethyl-5-deazariboflavin has been reported, based upon the thoroughly-studied synthesis of 5-deazariboflavin. The target compound was successfully synthesised in 28.7 % yield overall, which is significantly greater than that obtained using the optimal method to form 5-deazariboflavin (16.2 %), using the same synthetic route and reaction conditions. This direct comparison of the syntheses provides evidence that the methyl groups of 3,4-dimethylaniline **31** play a key role in the nucleophilicity of the reactive amine during several stages of the synthesis of 5-deazariboflavin. The initial reaction steps (described in sections 2.3.1 and 2.3.2) were both significantly improved using 4-methylaniline **49**, indicating that the *meta*-methyl group (with respect to the amine) plays a direct role in reduction of reaction yield. To further support this conclusion, it would be interesting to repeat the synthesis using 3-methylaniline and 3,5-dimethyl-aniline, in order to provide further evidence of this effect.

2.4 Attempted formation of multiple riboflavin analogues using a single synthetic method

2.4.0 Introduction

In this section, a potential route for the synthesis of four riboflavin analogues (1-deazariboflavin, 1,5-dideazariboflavin, 3-deazariboflavin and 3,5-dideazariboflavin) is described. This was attempted (unfortunately without success), to devise a novel

route to these four analogues using minor adjustments to reactions previously examined during the synthesis of 5-deazariboflavin.

Two reactions crucial for the success of this method were the synthesis of pyridine-based analogues of uracil, and their efficient coupling to 3,4-dimethyl-N(ribityl)-aniline **33**. As formation of these pyridine compounds was expected to produce two structural isomers (**58a** and **58b**, fig. 57) from the same starting material, it was anticipated that this would enable formation of the four riboflavin analogues with little wastage. As 1-deazariboflavin had been synthesised earlier (5.2), it was intended to use for direct comparison of the effectiveness between both two routes.

Numerous methods were attempted (including the use of several alternative functionalized uracil analogues, and the use of adjusted reaction conditions) to furnish the products; however in all cases the key coupling reaction was a failure, preventing examination of later synthetic steps. Although this method therefore proved unsuccessful, the results obtained are reported in order to provide a foundation for future work into the formation of these four riboflavin analogues, where it may be of significant use to others.

Reactions **A** and **D** of figure 57 (below) describe the coupling between the uracil analogues and ribitylated aniline **33**. Upon formation of the bicyclic intermediates, two methods are available to close the central ring of the isoalloxazine: (i) using the Vilsmeier-Haack method (successfully used to form 5-deazariboflavin and 8-demethyl-5-deazariboflavin) via reactions **C** and **F**; or (ii) using $NaNO_2$ (previously reported by Yoneda et al. during their study of isoalloxazine formation[107,108]) via reactions **B** and **E**. Therefore, from the two intermediates **82** and **85**, four riboflavin analogues may be produced.

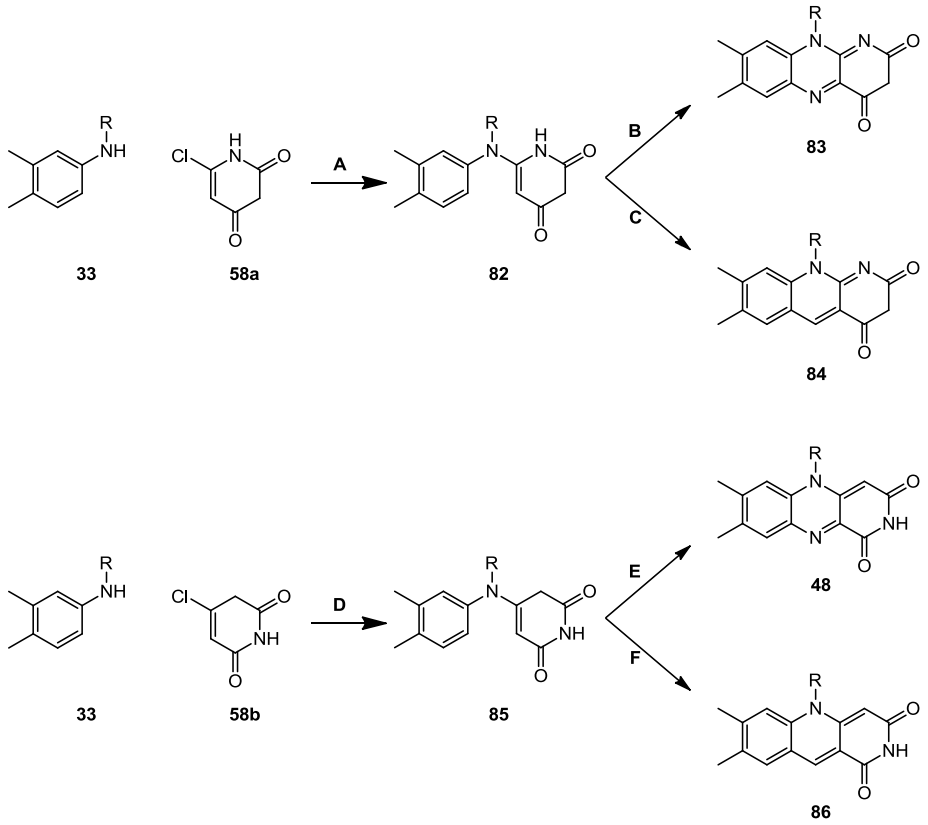


*Figure 57: Proposed synthetic route to four riboflavin analogues (3-deazariboflavin **83**, 3,5-dideazariboflavin **84**, 1-deazariboflavin **48** and 1,5-dideazariboflavin **86**) from two pyridine-based analogues (**58a** and **58b**) of uracil.*

2.4.1 Reaction overview

The crucial reaction for the success of this method was the coupling between the functionalised uracil analogue and the ribitylated analogue. As conditions for the successful analogous coupling between 3,4-dimethyl-N(ribityl)-aniline **33** and 6-chlorouracil **35** had been found previously (section 2.1.3), this method was used as the basis for the novel reactions.

As described above, experimental results suggests that the primary reason for the moderate yield obtained during the original reaction (55 %, including subsequent acetylation) was the low reactivity of the ribitylated amine, lowered further by the presence of a methyl group *meta* to this position. When 4-methyl-N(ribityl)-aniline **50** (section 2.3.2) was utilised for the coupling with 6-chlorouracil **35**, the yield improved to 79 % (including acetylation); while 4-methyl-N(ribityl)-aniline **50** was clearly more

successful at furnishing the desired product, its use in the novel synthetic route under consideration (fig. 57) would lead to the formation of 8-demethyl- flavin analogues. Although the potential formation of a suite of novel 8-demethyl-X-deaza- (or 8-demethyl-X,Y-dideaza-) flavin compounds was particularly interesting, these were less desirable for the initial evaluation of the synthetic route, as the spectroscopic characteristics of the products are unknown. It was therefore decided that 3,4-dimethyl-N(ribityl)-aniline **33** would be sufficiently reactive to give an initial representation of whether the proposed synthetic method would be successful, and which could be performed using alternative regents if desired.

2.4.2 Formation of 2,4,6-trichloropyrimidine 57

As each of the desired products lacked a nitrogen atom within the uracil moiety of the isoalloxazine structure, the uracil analogues necessary were derived from single-nitrogen heterocyclic compounds (pyridines or piperidines). In order to closely follow the well-studied previous synthesis of 6-chlorouracil **35** from 2,4,6-trichloropyrimidine **34**, 2,4,6-trichloropyridine **57** was required for the initial reactions, which was synthesised in good yield (78 % over two steps) from the inexpensive starting material 2,6-dichloropyridine **55**.

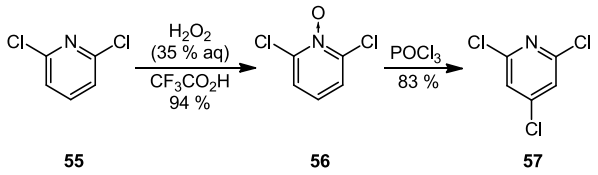


*Figure 58: Formation of 2,4,6-trichloropyridine **57** proposed by Kaneko et al.*[163]

Following the method devised by Kaneko *et al.*[163], synthesis of 2,4,6-trichloropyridine **57** entailed initial activation of the 2,6-dichloropyridine **55** via formation of an N-oxide intermediate **56** using hydrogen peroxide in TFA, before chlorination of the intermediate using refluxing $POCl_3$ to furnish product **57**. Each reaction was swift and efficient, with simple purification; although FC was originally performed, due to the crystalline nature of all products it was possible to successfully purify these using recrystallisation.

This sequence of reactions was necessary, as it was found that direct treatment of 2,6-dichloropyridine **55** using the chlorinating agent without prior activation failed to produce the desired product (section 6.4.1.3). The formation of the N-oxide permits chlorination of the pyridine ring by selectively reducing the electron density of the pyridine. As a dative bond is formed between nitrogen oxygen to form the oxide, electron density is withdrawn from the aromatic system at positions *ortho* and *para* to the nitrogen. As both *ortho* positions (2 and 6) are already occupied by chlorine, the *para* position is the only possible site for chlorination, leading to the single isomer product **57** obtained in very high yield. Similar N-oxide activation of pyridines prior to halogenation has been reported elsewhere[164–166] with an interesting adjunct; the converse reaction (chlorination of 4-substituted pyridines, such as 4-methylpyridine, at the *ortho* position) is also only possible using the analogous N-oxide intermediate[166].

An alternative route to prepare 2,4,6-trichloropyridine **57** was devised by Schlosser *et al.*[167]. 2,6-dichloropyridine **55** was treated with n-butyllithium at -75 °C and iodine to form 2,6-dichloro-4-iodopyridine **87** (figure 59) in 64 % yield. The moderate yield was attributed to the undesired formation of the alternative 3-iodo product in place of the 4-iodo compound, in a relative ratio of 5:2. The iodine group of intermediate **87** was subsequently replaced by chlorine using a mixture of isobutylmagnesium chloride and *N*-chlorosuccinimide, forming 2,4,6-trichloropyridine **57**.

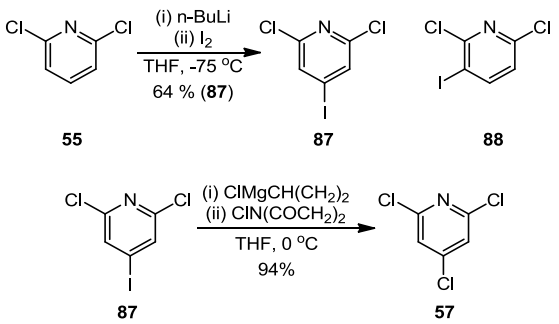


*Figure 59: Alternative synthetic route to **57** devised by Schlosser et al.[167]. Formation of 4-iodo-2,6-dichloropyridine **87** was hampered by the undesired formation of 3-iodo isomer **88***

The choice of butyllithium by Schlosser *et al.*[167] to perform the lithium-halogen exchange is interesting, as the reaction of pyridine **55** with n-butyllithium would be expected to form a number of intermediates leading to several undesired products. For unsubstituted pyridine, although the relative acidity of the proton position *ortho* to the nitrogen atom is the weakest of the three proton positions, the potential for coordination of lithium by nitrogen favours lithiation at this position. However, as both *ortho* positions of 2,6-dichloropyridine are already substituted, attack must occur at either of the remaining positions (*meta* or *para*). These are favoured either kinetically or thermodynamically (respectively)[168], with relative proportions of each product controlled by the reaction conditions; this was observed by Schlosser *et al*.[167], with a slight excess of the *para*- species formed (in a ratio of 5:2).

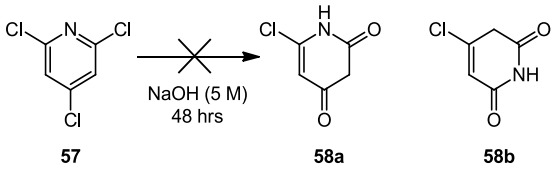
However, as the method of Schlosser *et al.* was much less efficient overall than the method of Kaneko *et al.* (60 %, compared to 78 % (respectively), from starting material **55**), and may have formed undesired 3-substiuted side-products, the method of Kaneko *et al.* was used to form 2,4,6-trichloropyridine **57**.

2.4.3 Attempted conversion of 2,4,6-trichloropyridine 57 to 58a and 58b

The formation of 6-chlorouracil **35** from 2,4,6-trichloropyrimidine **34** described earlier (section 2.1.2) used a high concentration of hydroxylic base to facilitate hydrolysis of the pyrimidine ring, forming 6-chlorouracil **35** in yields up to 81 %. Chlorination served to activation the pyrimidine for attack by hydroxide, with the remaining chlorine being used to direct formation of the desired bicyclic product **36** upon reaction with 3,4--dimethyl-N(ribityl)-aniline **33**. Overreaction (hydrolysis of all three chlorine groups) was not observed under any condition examined, or with unrealistically extended reaction times (up to 120 hours).

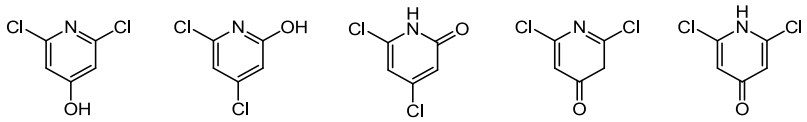
The most successful conditions for production of 6-chlorouracil **35** (5M NaOH, 48 hours at reflux) were found to be unsuccessful at furnishing the desired products **58a** and **58b** when applied to 2,4,6-trichloropyridine **57**. Instead, a large proportion of unreacted **57** was recovered (66 %), along with an unidentified white crystalline solid (138 mg, 31 % based on later MS evidence, *vide infra*). While it was originally hoped

that this would be one (or a mixture) of the intended products, mass spectrometry presented an m/z value of 162.93. This strongly indicated a dichlorinated intermediate with the molecular formula $C_5H_3Cl_2NO$ (162.96), further supported by relative isotope patterns for ^{35}Cl and ^{37}Cl observed. However, even with reaction times up to 144 hours, this remained the dominant product, with no mono-chlorinated product (determined by MS) obtained.



57 **58a** **58b**

*Figure 60(a): Conditions previously used to form 6-chlorouracil from 2,4,6-trichloropyrimidine were unsuccessful at furnishing either **58a** or **58b***



89a **89b** **89c** **89d** **89e**

*Figure 60(b): Several possible structural isomers of $C_5H_3Cl_2NO$ (**89a-e**) exist, formed from 2,4,6-trichloropyridine **57***

The ^1H NMR structure of the isolated solid (shown in figure 61 below) showed two aromatic resonances at 6.76 and 6.69 ppm, indicating two differing proton environments and therefore a lack of symmetry for the molecule. Relative intensity of each resonance suggested a single proton, which agrees with the structure for compound **89b**. The splitting pattern of each proton signal (with coupling constants of 1.6 Hz for both peaks), suggests relative *meta* potitioning of the protons, which is the conformation expected for the majority of possible structural isomers (**89a-89e**).

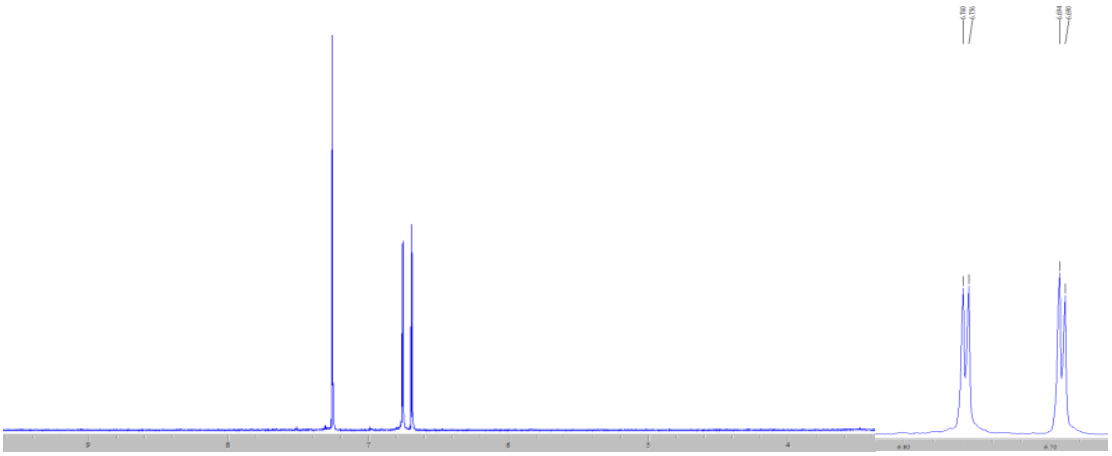


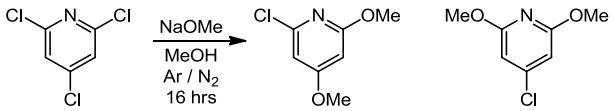
Figure 61: ^{1}H NMR spectrum of the product from reaction 3.4.2, showing the relative positioning of the proton resonances. From the left, these are the solvent residual peak ($CDCl_3$) at 7.26 ppm, a pair of doublets (6.76, 6.69 ppm; each 1H, with J = 1.6 Hz), and a broad singlet (1.66 ppm, 1H; not shown). An expansion of the region 6.80-6.60 ppm is provided, demonstrating a small coupling constants for each resonance

As **89b** was the only product formed by this reaction, even using vastly extended reaction times (144 hours), this method appeared unsuitable for formation of **58a** or **58b**. When the reaction was repeated using 2,6-dichloropyridine **55** (section 6.4.3), no product was observed after 48 hours, with near-quantitative recovery of the starting material. Due to the failure of this reaction to produce the desired products, it was decided to abandon this method and instead investigate the synthesis of other nitrogen-based heterocycles.

2.4.4 Formation of 2-chloro-4,6-dimethoxypyridine

The conversion of 2,4,6-trichloropyridine **57** to the dimethoxy-derivatives **60** or **61** was performed using the method previously reported by Kaneko *et al.*[163]. The product again retained a single chlorine, which was intended to direct attack of 3,4-dimethyl-N(ribityl)-aniline **33**, for formation of the corresponding bicyclic intermediates (figure 63, *vide infra*). However, a complication reported with this method was the formation of a secondary product, 4-chloro-2,6-dimethoxypyridine **61**[163] (a structural isomer of the intended product, fig. 62). The mixture of products **60** and **61** were found in

relative ratios of 7:3; no evidence of 2,4,6- trimethoxypyridine was found using these conditions.



57 60 61

*Figure 62: Formation of 2-chloro-4,6-dimethoxypyridine **60** and 4-chloro-2,6-dimethoxy-pyridine **61** from 2,4,6-trichloropyridine **57**, in a ratio of 7:3 (respectively)*

Interestingly, results described by Schlosser *et al.*[167] report that reaction of sodium ethoxide with 2,4,6-trifluoropyridine under similar conditions to those used to form **60** and **61** led to the regioselective formation of a single product, 4-ethoxy-2,6-difluoropyridine. The use of an alternative nucleophile, dimethylamine, in reaction with both 2,4,6-trifluoropyridine and 2,4,6-trichloropyridine also reportedly led to the sole formation of the 4-dimethyl-2,6-dihalogen compounds[167]. This result is unexplained, and is at odds with the results reported by Kaneko *et al.*[163] and confirmed during this work, which clearly demonstrate the successful formation of dimethoxypyridines **60** and **61**, with no mono-methoxypyridine intermediate observed upon completion of the reaction.

2.4.5 Attempted coupling between 3,4-dimethyl-N(ribityl)-aniline 33, and dimethoxypyridine compounds 60 or 61.

The coupling between 3,4-dimethyl-N(ribityl)-aniline **33**, and either 2-chloro-4,6-dimethoxypyridine **60** or 4-chloro-2,6-dimethoxypyridine **61** as shown in figure 63 below, was attempted in order to facilitate the formation of the bicyclic intermediates **62** or **63**. Removal of the methyl groups was to be performed later upon successful formation of the final riboflavin analogues, to prevent unintended reaction at these positions.

Each of the three coupling methods examined previously for the formation of **36** (section 2.1.3) were examined to facilitate these reactions, however none was able to permit product formation. In all cases, only unreacted starting materials were

recovered after reaction times significantly longer than necessary for the equivalent coupling using 6-chlorouracil **35**, indicating that these reagents were not sufficiently reactive to allow product formation.

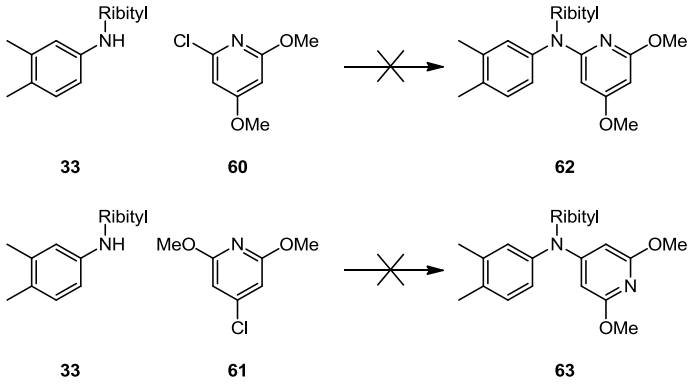


*Figure 63: Attempted reaction between **33** and **60** or **61**, to give bicyclic intermediates **62** or **63**. All conditions examined failed to give the desired products*

Failure to form either product **62** or **63** suggests that the nucleophile (3,4-dimethyl-N(ribityl)-aniline **33**) is not sufficiently reactive to attack reactants **60** or **61**, which may be predicted based upon the aromaticity of compared to 6-chlorouracil **35**. While pyridines typically contain less electron density throughout the aromatic ring compared to benzenes (due to the electrophilic nature of the nitrogen heteroatom), it is still a much greater density than for the non-aromatic uracil compound used previously, potentially preventing reaction with 3,4-dimethyl-N(ribityl)-aniline **33**.

Furthermore, the effect of conjugation possible with 6-chlorouracil **35** may also play a key role in this reaction. The carbonyl oxygens of uracil are able to act as an 'electron sink' (via enolate formation) upon nucleophilic attack, and thus stabilise the reactive intermediate; whereas no such facility is possible with intermediates **60** and **61** due to the methyl substituents. Thus, the uracil form is significantly more susceptible to nucleophilic attack than the aromatic methoxy-pyridines.

As described in section 2.1.3.1, the malononitrile-catalysed reaction between 3,4-dimethyl-N(ribityl)-aniline **33** and 6-chlorouracil **35** was expected to proceed via displacement of the chlorine of 6-chlorouracil **35** by malononitrile, before attack of **33**

leading to formation of the bicyclic product. This was based upon the reported reaction of malononitrile with vinyl halides[132], or with aryl halogens in a similar displacement-type reaction[133,134] (fig. 64).

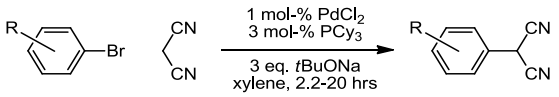


Figure 64: Scheme for reaction between a substituted aryl bromide and malononitrile, described by Schnyder et al.[133]*. "R" describes several aromatic substituents (electron withdrawing and donating) at various positions of the aromatic ring. The resultant aryl malononitriles were synthesised in yields up to 92 %*

However, while the reaction of malononitrile with vinyl halides or 6-chlorouracil **35** proceeded without assistance[132], the method of Schnyder *et al*.[133] for displacement of aryl halogens described the use of Pd catalysis. This was not attempted, due to the risk of damage to the alcohol groups of the ribityl side-chain[169–172].

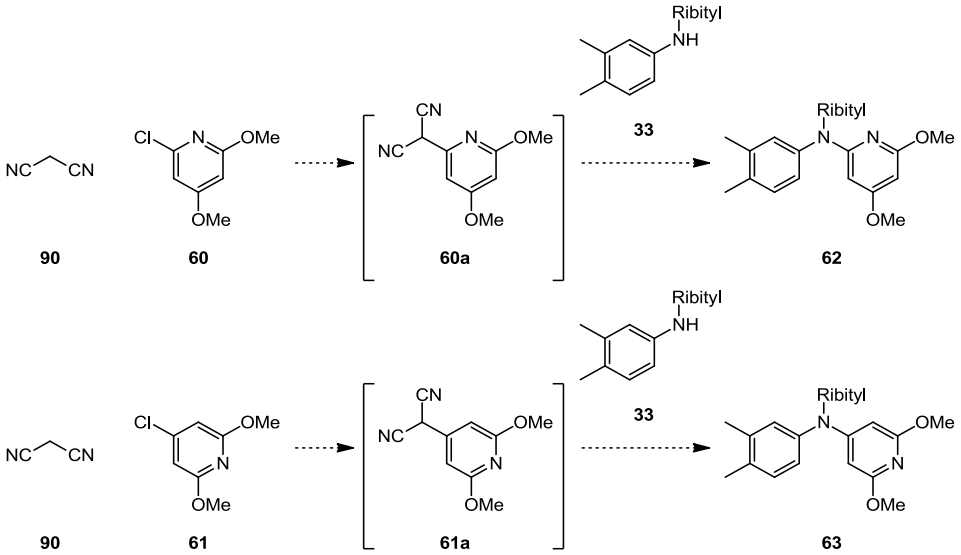


*Figure 65: Intended formation of intermediates **62** and **63** via malononitrile displacement of chlorine from **60** or **61**; formation of **62** and **63** was not observed*

As neither product **62** or **63** was formed by this reaction, it suggests that malononitrile **90** was not able to form the activated intermediate species **60a** or **61a** (fig. 65). This was supported by spectroscopic examination of the products from the reaction, where both **60** and **61** were recovered in high yield. However, this route may be successful if

the incorporation of malononitrile **90** could be performed (with additional catalysis), and the activated intermediates isolated without degradation. These compounds may be more sensitive to attack by 3,4-dimethyl-N(ribityl)-aniline **33**, potentially giving a viable route to **62** and **63**.

2.4.6 Formation of 2,4,6-piperidinetrione 66

As two likely reasons for the failure of the reaction between 3,4-dimethyl-*N*(ribityl)-aniline **33** and dimethoxychloropyridines **60** or **61** were the difficultly in nucleophilic attack (due to the high electron density of the aromatic ring), and the lack of a conjugate form for the bicyclic intermediate, the use of an unsaturated compound such as 2,4,6-piperidine trione **66** may have allowed successful formation of bicyclic products **67a** or **67b** (figure 66). As before, nucleophilic attack at either position α or γ to the nitrogen of the piperidine-trione would furnish two useful bicyclic compounds, potentially providing a method to form several riboflavin analogues. Therefore, the synthesis of **66** was performed, as shown in figure 67.

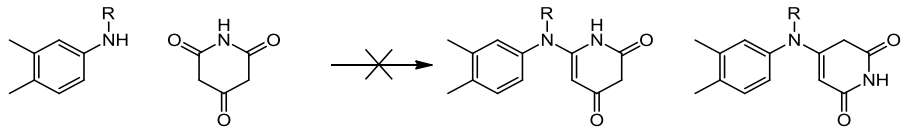


*Figure 66: Coupling between 3,4-dimethyl-N(ribityl)-aniline **33** and 2,4,6-piperidine trione **66** may have produced both bicyclic products **67a** and **67b**. However, no conditions were found to permit this reaction. "R" describes a D-ribityl group*

The formation of piperidine-trione **66** was described in great detail in work by Weinstock *et al.*[106], alongside their reported synthesis of 1,5-dideazariboflavin **75**. This was performed by cyclisation of diethyl-2-cyano-3-iminoglutarate **65** (referred to synonymously as ethyl-β-imino-α-cyanoglutarate in earlier literature), synthesised according to the procedure of Baron *et al.*[173] via a bimolecular condensation of ethylcyanoacetate **64** in refluxing NaOEt (fig. 67). Cyclisation was achieved using relatively harsh conditions (refluxing 12 M HCl, followed by boiling in NaOH), furnishing the sodium salt of piperidine-trione **66** in a yield of 73 %.

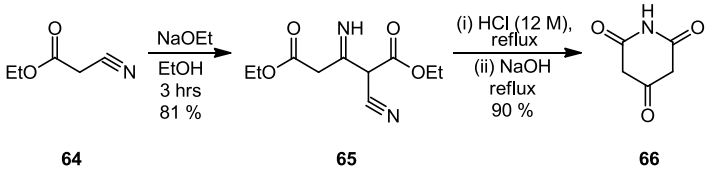


*Figure 67: Synthesis of 2,4,6-piperidine-trione **66**, via ethyl-β-imino-α-cyanoglutarate **65** formed from elthylcyanoacetate **64***

The formation of product **66** as the sodium salt was reported to be crucial to the stability, as the pure isolated compound appeared highly susceptible to oxidation by air[174], and readily underwent intermolecular attack with other molecules to form a bimolecular compound which was reportedly inactive[106]; isolation of **66** as the sodium salt prevented this.

2.4.7 Attempted coupling of ribitylated aniline 33 and piperidine trione 66

The reported formation of 1,5-dideazariboflavin **75** devised by Weinstock *et al.*[106] using 2,4,6-piperidinetrione **66** was used as the basis for examination of an analogous route to intermediates **67a** and **67b**. The original method (summarised in figure 68) describes the use of a tosylated hydrazido-N(ribityl)-aniline analogue **73**, with the tosyl hydrazide protecting the functionality of the aldehyde. This was necessary as at this time, the Vilsmeier-Haack method of formylation had not yet been applied to this synthesis. The protecting group was removed immediately prior to the reaction with 2,4,6-piperidinetrione **66**.

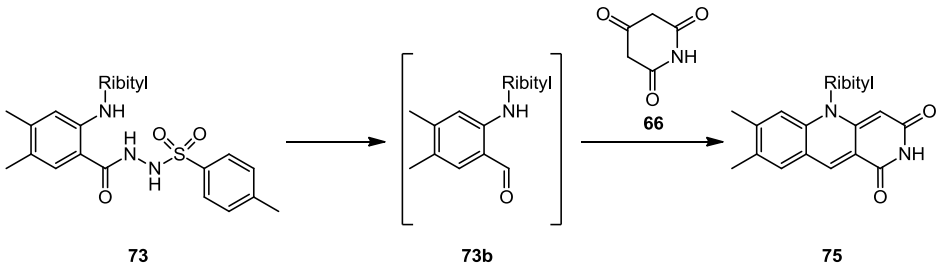


*Figure 68: Scheme for reported formation of 1,5-dideazarboflavin **75** from protected 2-amino-N(ribityl)-4,5-dimethylbenzaldehyde **73**[106], which was deprotected in situ before addition of 2,4,6-piperidine trione **66**; condensation furnished 1,5-didezaflavin*

It was hoped that the use of 3,4-dimethyl-N(ribityl)-aniline **33** would react similarly to tosylated hydrazido-N(ribityl)-aniline analogue **73**, with attack at either α or γ positions of 2,4,6-piperidinetrione **66** forming one (or both) of the bicyclic intermediates **67a** or **67b** described in figure 66 above.

While **66** was previously isolated as the sodium salt, it was readily reconstituted in mildly acidic solution immediately prior to reaction[106]; the reaction between **73b** and **66** was also performed in acidified aqueous solution. However, aqueous solution was previously found during this research to be unsuitable (section 2.1.3.3) for reaction between 3,4-dimethyl-N(ribityl)-aniline **33** and 6-chlorouracil **35**, and so when these conditions were applied to the reaction between **33** and **66** (which required a small proportion of methanol to aid dissolution of **33**), again no reaction was observed.

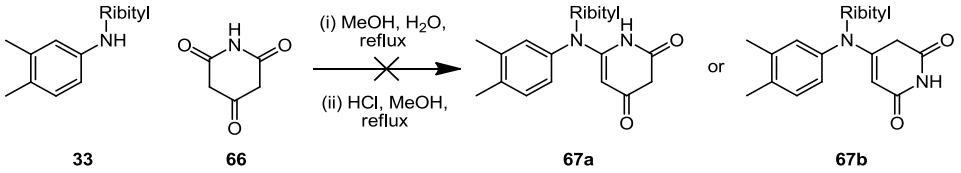


*Figure 69: Attempted reaction between **33** and **66** as described by Weinstock et al.[106] for their formation of 1,5-dideazariboflavin **75** failed to furnish either **67a** or **67b***

Several additional routes the form the bicyclic compounds were also examined, including the use of the previously successful solvent-free heating method, or performing the reaction in an anhydrous organic solvent, using a crown ether to allow dissolution of the salt of **66**. However, each reaction again failed to produce the desired products. In all cases, starting materials (or degraded side-products) were isolated in near-quantitative yield.

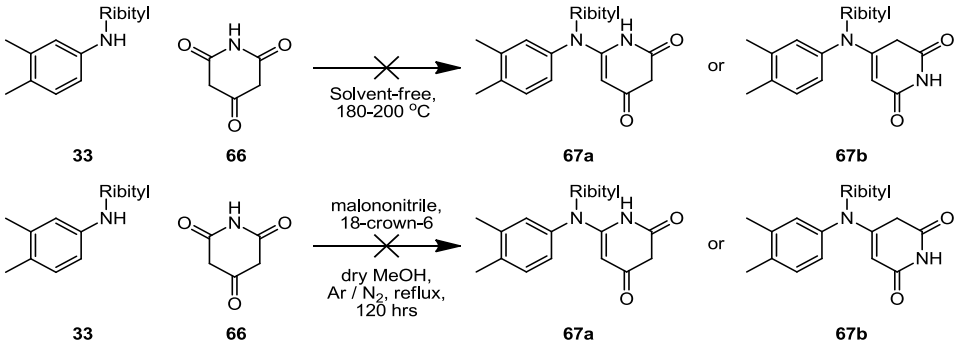


*Figure 70: Alternative routes attempted to form **67a** and **67b***

This result was somewhat difficult to understand, due to the precedent for the similar reaction set by Weinstock *et al*.[106] (figure 68). Furthermore, 2,4,6-piperidinetrione **66** was expected to be significantly more susceptible to nucleophilic attack by 3,4-dimethyl-N(ribityl)-aniline **33** than either of the dimethoxypyridine species discussed earlier (section 2.4.5). Thus, the lack of reaction between **33** and **66** indicates that this must be due to another reason, with the poor nucleophilicity of **33** initially suspected. However, while reactions performed using 4-methyl-N(ribityl)aniline **49** were generally significantly more efficient due to the enhanced nucleophilicity of the amine, even using this reagent it was not possible to facilitate coupling with **66**. Therefore, while the low nucleophilicity of the aniline is undoubtedly some hindrance to the reaction, it is likely that an additional factor (or factors) are also to blame, and thus this reaction is therefore worthy of further study. The activation of 2,4,6-piperidinetrione **66** (for example, via tosylation) may allow the reaction to proceed by providing an enhanced leaving group, potentially allow formation of the intermediate compounds **67a** or **67b** and leading to the riboflavin analogues described in fig. 57.

2.4.8 Conclusion – attempted formation of multiple riboflavin analogues

In this section, the formation of several heterocyclic molecules was discussed, for use as analogues of 6-chlorouracil **35**. While these were typically produced in moderate yield, no enhancements to the reactions were sought, as their primary purpose was to test a novel synthetic pathway potentially leading to several riboflavin analogues.

Reaction of these uracil analogues with 3,4-dimethyl-N(ribityl)-aniline **33** was intended to provide two bicyclic intermediates **67a** and **67b**, which may have each been used to prepare two analogues of riboflavin as described in section 2.4.0. However, no heterocycle was successfully coupled to 3,4-dimethyl-N(ribityl)-aniline **33**. While frustrating, useful information regarding conditions and outcomes was recovered for each reaction, which may allow future work in this area to proceed more rapidly. Key to this may be the reliable formation and functionalization of 2,4,6-piperidinetrione **66** to assist coupling with **33**, finally allowing successful formation of the bicyclic intermediates as a precursor for flavin analogue synthesis.

Chapter 3

Results and Discussion –

Molecular Biology

Chapter 3 - Results and Discussion - Molecular Biology

3.0.1 Introduction

In this chapter the results obtained within the field of Molecular Biology are reported and discussed. These results include detail of the expression and purification of a bifunctional (prokaryotic) riboflavin kinase-FAD synthetase from *Corynebacterium ammoniagenes*, and a monofunctional (eukaryotic) riboflavin kinase protein from *Schizosaccharomyces pombe*. The formation of the phosphorylated riboflavin analogues (FMN and FAD) by these enzymes is compared and discussed, as while FMN was the desired cofactor form for future studies, the FAD form was the predominant product of the bifunctional enzyme, thus requiring additional treatment before use.

Additionally, the expression, purification and photo-reaction of the PHOT1-LOV2 domain from *Avena sativa* is detailed. This domain has been studied using both the wild type protein, and a series of mutants in which the active cysteine residue (C450) was modified.

Each AsLOV2 sample was studied containing their native cofactor (FMN), and reconstituted with both the 1-deaza-FMN and 5-deaza-FMN analogues, allowing collection of evidence to support the understanding of the LOV domain photocycle.

3.1 *C. ammoniagenes* Bifunctional Rf Kinase – FAD-Synthetase

3.1.1 Protein expression and purification

This protein, first isolated by Hagihara *et al.*[175,176], was expressed using the plasmid expression vector pMAL-C2, provided as a ready-to-use construct (prepared previously[177]) by the project supervisor. This vector allowed expression of the *C. ammoniagenes* enzyme as a fusion protein with an N-terminal maltose-binding (MB) domain, with a weight of approx. 66 kDa for the fusion protein (~28 kDa for the MB domain; 38 kDa for the RfK-FADS). The Ca-RfK-FADS gene was initially expressed in *E. coli* XL1-blue (see fig. 71 below) in moderate yield, although later expression in *E. coli* strain BL21 significantly improved protein yields.

The MB domain allowed isolation of the recombinant protein by affinity chromatography over amylose resin. Elution was performed using a maltose-containing buffer, which competitively bound to the MB domain, releasing the protein from the resin. Due to the efficiency of the overexpression, several binding cycles were typically necessary, due to the limited capacity of the amylose resin.

Each fraction of the flow-through from the column was scrutinised immediately using UV spectroscopy (280nm) and SDS-PAGE. When the column capacity had been reached, it was possible to visualise unbound protein passing through the column; those fractions found to contain residual protein were reserved and re-bound later. Protein fractions were dialysed into buffer RfK 3 (section 5.2.7) and concentrated by ultrafiltration over a 30 kDa membrane. The purified protein was found to be stable in solution at 4°C for 1 week, after which a loss of activity and precipitation was observed. Attempts were made to flash-freeze the protein before storage at -80°C, but the thawed protein exhibited a total loss of activity.

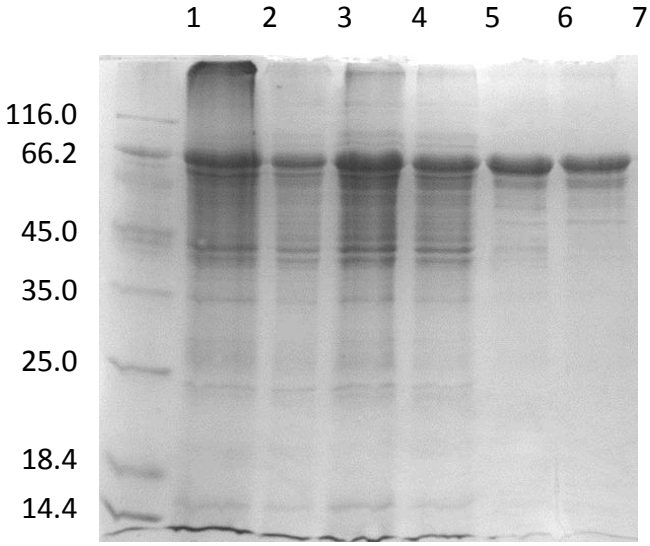


Figure 71: SDS-PAGE Gel of recombinant C. ammoniagenes RfK-FADS.
Lane 1 - molecular weight marker (SM0431); lane 2 - whole cell lysate; lanes 4 and 5 - flow-through from column (showing unbound protein, due to column capacity being reached); lanes 6 and 7 – purified bifunctional Ca-RfK-FADS protein (66 kDa)

3.1.2 Phosphorylation of riboflavin analogues using bifunctional RfK-FADS

Phosphorylation of riboflavin (and the synthesised 5-deazariboflavin analogue) by the *Corynebacterium ammoniagenes* RfK-FADS produced a mixture of both FMN and FAD cofactor forms. This was expected, as this enzyme is used within the cell to regulate and maintain intracellular stores of both flavin cofactors[78,79]. However, while the FMN form was required for incorporation into the AsLOV2 domain, this appeared to be the minor product of the enzyme. An in-depth examination of this phosphorylation was therefore performed to determine the optimal to form the FMN analogue.

During the course of the enzymatic reaction, it was determined (assayed using TLC) that upon initial formation of FMN, this compound rapidly underwent further reaction to form FAD. Thus, FMN concentration appeared to remain lower than FAD concentration throughout the reaction. As the reaction proceeded, the relative proportion of FAD increased while RF concentration decreased, eventually leading to a sample containing FAD as the predominant constituent. Upon consumption of all riboflavin (or deaza-analogue), the NT domain of the protein appeared to remain active for some time, catalysing the interconversion between FAD and FMN to reach an eventual equilibrium of the phosphorylated products.

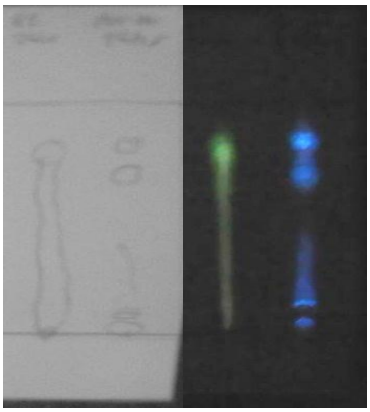


Figure 72: TLC of true riboflavin (green fluorescence; RF = 0.81) and 5-deaza-products from enzymatic reaction with Ca-RfK-FADS (blue fluorescence). 5-deazariboflavin (RF = 0.83), 5-deazaFMN (Rf = 0.70) and 5-deazaFAD (Rf = 0.16) were identified; 5-deazaflavin bound by the enzyme is also observed (Rf = 0.0)

Due to the complex nature of the reaction, the rate of flavin turnover by this enzyme was not examined during this project. However, the kinetics reported by Efimov *et al.*[73] for the reactions catalysed by *Ca*-RfK-FADS describe that for the conversion of riboflavin to FMN, a k_{cat} of 0.41 ± 0.9 s^{-1} (with an indeterminate rate for the reverse reaction) is reached; for the interconversion of FMN to FAD, a k_{cat} of 0.300 ± 0.023 s^{-1} (with a slightly decreased k_{cat} for the reverse reaction of 0.270 ± 0.023 s^{-1}) is possible. These reported rates suggest a more rapid formation of FMN than FAD (as the rate of FMN to FAD conversion is lower than the rate of FMN formation from riboflavin, and is decreased further by the rate of the reverse reaction), rather than the experimentally observed greater titres of FAD than FMN. However, this difference may be explained by the presence of a large excess of ATP within the experimental reaction mixture, which would adversely affect the kinetics of several steps (according to Le Chatelier's Principle) and thus increase the overall rate of FAD formation.

3.1.3 Cofactor purification

Preparative-scale purification of the phosphorylated deaza-analogues was somewhat difficult. Whilst it was possible to assay the reaction using TLC (silica stationary phase), the mobile phase (3:1:1 mixture of isopropanol, H_2O and AcOH; respectively) necessary for adequate separation rendered this process unsuitable for purification on a preparative scale. The high polarity of the flavin cofactors (due to the incorporated phosphate or adenosine diphosphate groups) led to a strong interaction with the silica stationary phase using a less polar mobile phase. However, the use of "reversed phase" (i.e. hydrophobic-covered silica matrix) stationary phases was successfully used to separate the products of the enzymatic mixture on a preparative scale, using HPLC. Whereas before, the most polar component of the reaction mixture (FAD) was retained for the longest by the silica matrix (resulting in the lowest retention factor value, RF), reversed-phase chromatography allowed the more highly polar FAD to be eluted before less polar cofactor forms.

HPLC was performed using similar methods to those previously described[178,179], after removal of the enzyme by passing the mixture over an amylose column (and confirmed by SDS-PAGE). The conditions were adjusted to allow good separation of

the reaction components, with FAD, FMN and riboflavin forms eluting with retention times of 7.35-9.45 minutes, 20.05-20.90 minutes, and 28.70-29.70 minutes respectively, shown in figure 73. The peak area of the HPLC indicated formation of FAD, FMN and riboflavin in relative ratios of 25 : 1 : 3, respectively.

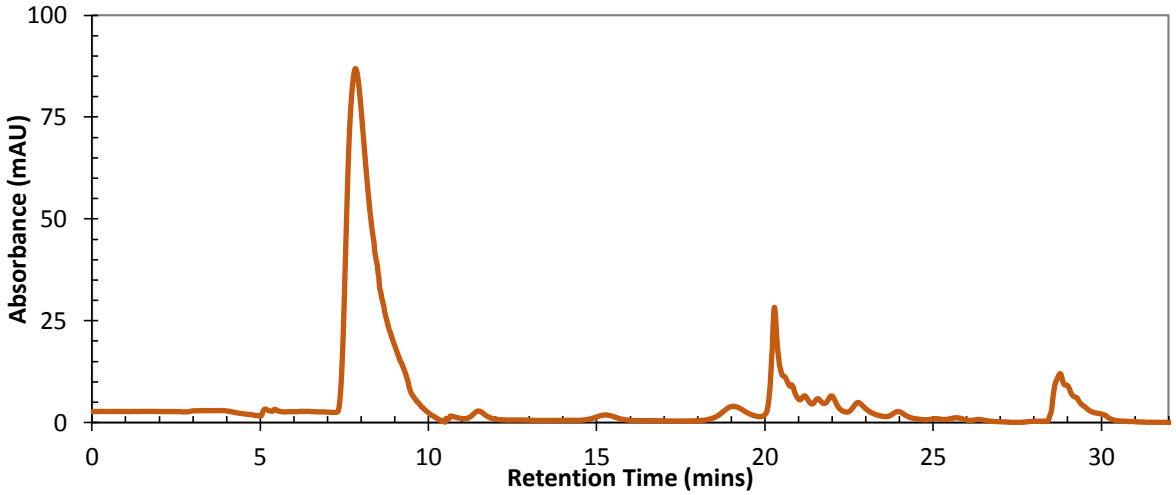


Figure 73: HPLC trace of 5-deazaFAD formed by the Ca-RfK-FADS, observed at 400nm. Conditions and retention times are described fully in section 7.1.5

After lyophilisation, product purity was very high, although a typical HPLC run only isolated between 0.5-1.5 mg of purified products. Therefore, repeated rounds of HPLC were necessary to isolate the compound in sufficient quantities for experimental use.

3.1.4 Conversion of 5-deazaFAD to 5-deazaFMN

As the FAD-cofactor form was isolated as the major enzymatic component, conversion to the FMN form was required for use in the AsLOV2 domain. Eukaryotic cells typically utilise an alkaline phosphatase, which strips the adenosenyl-monophosphate moiety from FAD to give FMN, and was successfully employed for isolation of 5-deazaFMN by a colleague during a short side-project (Dr. Hannah Collins, University of Kent), using commercially available Phosphodiesterase I from *Crotalus atrox* (Western Diamondback Rattlesnake).

Hydrolysis of the phosphoanhydride was followed by UV-visible observation of the sample, with a blue-shift of approx. 5 nm for FMN compared to FAD, coupled to an increased absorbance (due to an increase in extinction coefficient[88]). For FAD, the

adenine ring (held in close proximity to the flavin by the flexible ribityl chain) undergoes a degree of π-π stacking with isoalloxazine, which has a small effect on the electrochemistry and thus induces a spectral shift; cleavage of the AMP moiety removes this effect. Dephosphorylation was confirmed by HPLC and MS.

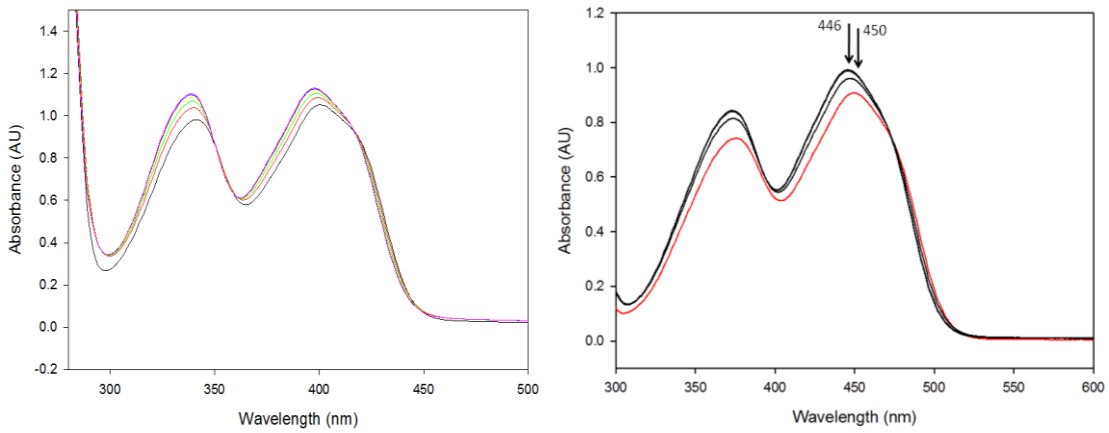


Figure 74: UV-visible spectra showing dephosphorylation of 5-deazaFAD (left), and FAD (right), recorded by Dr. Hannah Collins (University of Kent) using Phosphodiesterase I from C. atrox. An approximate 5nm blue-shift in flavin absorbance (previously 404 nm or 445 nm respectively) indicated hydrolysis of the phosphoanhydride. Spectra were recorded at 10 min. intervals with reaction complete after 30 mins using 4 units of enzyme per flavin equivalent

The use of large-scale enzymatic hydrolysis of the FAD cofactor analogues was considered, but discounted due to the discovery of alternative conditions to hydrolyse the phosphoanhydride bond. An acidified buffer (containing a strong acid, such as trichloroacetic acid (TClA) or trifluoroacetic acid (TFA)) had been previously reported to hydrolyse FAD to FMN within 1 hour[180,181]. This was successfully used to quantitatively hydrolyse 5-deazaFAD (within 1 hour, using 1 % (v/v) TFA), although subsequent purification by HPLC was again necessary to ensure isolation of the desired cofactor.

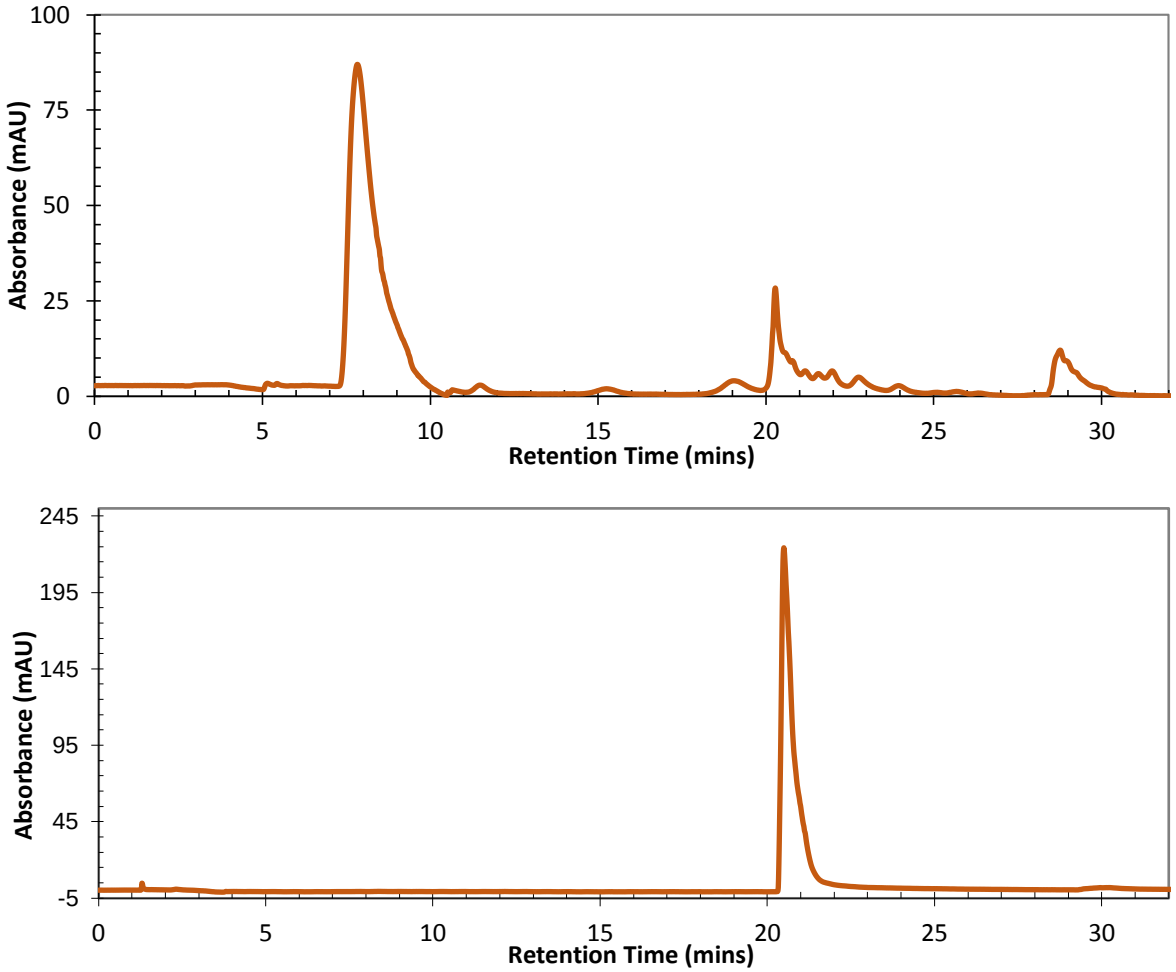


Figure 75: HPLC chromatographs of 5-deazaFAD, 5-deazaFMN and 5-deazariboflavin, observed at 400nm (above), and 5-deazaFMN formed by treatment of 5-deazaFAD with 1 % TFA (below). Conditions and retention times described in section 7.1.5

Direct production of the FMN analogue by inactivation of the FADS domain of the bifunctional RfK-FADS was also considered. However, as only the key residues for catalysis within the riboflavin kinase domain have so far been reported[182] (which are relatively easily determined due the homology between other prokaryotic and eukaryotic riboflavin kinases), this was deemed to be a divergence from the aim of this project, and was thus not attempted.

3.1.5 Conclusion - bifunctional Ca-RfK-FADS

Conditions were found for the successful expression and purification of the *Ca*-RfK-FADS protein. However, several difficulties were encountered, associated with the

bifunctional nature of the enzyme (producing FAD as major product, rather than FMN). An effective method to hydrolyse the FAD produced to FMN was found; however the time-consuming purification procedures to isolate the desired cofactor analogue rendered this route inefficient. Thus, an alternative method was sought, to perform the phosphorylation of the riboflavin analogue to the FMN form more swiftly and effectively.

3.2 *S. pombe* Riboflavin Kinase

During a collaboration with Professor Martin Warren and Dr. Hannah Collins (University of Kent, Canterbury) on a side-project, I was gifted a plasmid construct containing the riboflavin kinase gene from *Schizosaccharomyces pombe*. This is a eukaryotic riboflavin kinase of 166 residues[183], which had been inserted into a pET-14b vector by Dr. Hannah Collins. The Sp-RfK protein catalyses the formation of FMN from riboflavin, and was successfully used with 1-deazariboflavin and 5-deazariboflavin to produce the respective FMN analogues.

3.2.1 Protein expression and purification

Expression of a gene contained within the multiple cloning site (MCS) of pET-14b is controlled by a T7 promoter sequence, for which the *E. coli* expression strain BL21 is optimised, resulting in efficient transcription of the target gene. Immediately upstream of the MCS is a sequence for the expression of an N-terminal 6 x His polyhistidine tag, which is expressed as a fusion partner at the N-terminus of the target protein. The polyhistidine tag allowed convenient protein isolation (for purification, and removal after use) using an IMAC procedure.

In order to maintain the plasmid, it was immediately transformed into *E. Coli* XL1-blue (for plasmid storage, with an efficiency of 1.20×10^5 transformants per µg DNA), and BL21(DE3) (for protein expression, with an efficiency of 4.81×10^4 transformants per µg DNA). Expression of the protein was performed with high yields in both LB and TB media, with the polyhistidine tag allowing separation of the target protein from the cell lysate using a high-throughput method over nickel-chelating resin (Sepharose Fast

Flow), discussed below. Enzyme concentration was estimated using an extinction coefficient (ε) value of 21,890 M^{-1} cm^{-1} at 280 nm (calculated using the ExPASy ProtParam tool[184] using calculations defined elsewhere[185]).

Swift purification of the expressed protein was performed using a column of Ni-Sepharose Fast Flow resin. This resin is optimised for use in high-pressure systems (such as FPLC), and so it was possible to use a peristaltic pump to pass a large volume (50 mL) of filtered cell lysate over the column swiftly, without damaging the system. After binding and washing, retained protein was eluted from the column using a buffer containing a high concentration of imidazole, which competed with the histidine residues of the protein and weakened its interaction with the resin, causing elution. Fractions containing the desired protein were identified by SDS-PAGE, and concentrated (if required) by ultrafiltration, followed by dialysis or treatment using a desalting column to remove all traces of imidazole.

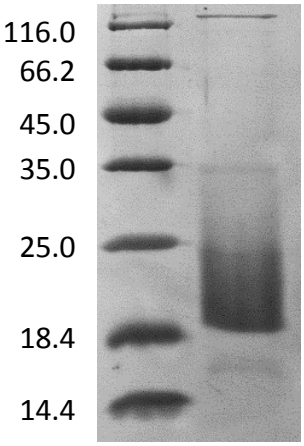


Figure 76: SDS-PAGE of purified Sp-RfK (18.9 kDa), compared to gel marker SM0431

3.2.2 Formation of 1-deaza and 5-deaza FMN using Sp-RfK

It was found that the expressed *S. pombe* RfK (Sp-RfK) successfully phosphorylated riboflavin and the synthesised 1-deaza and 5-deaza analogues much more efficiently than the previous bifunctional *Ca*-RfK-FADS. When assayed by the original creator of the plasmid construct (Dr. Hannah Collins, University of Kent), it was found that 1 nmol of purified protein was able to turn over a minimum of 130 nmol of riboflavin within 1 hour. No literature studies of this enzyme were available with which to compare this

result at the time of writing; however a similar result was observed for both 1-deazariboflavin and 5-deazariboflavin (130 nm of the flavin analogue per nmol of Sp-RfK was incubated in the presence of the enzyme and ATP; after 1 hour no riboflavin analogue remained, assayed by TLC).

The purified Sp-RfK was found to be most active at 37 °C, and was stable to storage in solution at 4 °C for 7 days or to lyophilisation. The reconstituted enzyme from lyophilisation had only slight loss of activity, and in this case a turnover of 100 nmol per nmol of enzyme was assumed for the subsequent enzymatic step. For large-scale preparation of the deaza-FMN analogues, very high (> 99%) yields were obtained using only 1.2 equivalents of ATP, with reaction times under 4 hours.

Removal of Sp-RfK from the mixture after completion of the reaction was straightforward, as the polyhistidine tag again allowed retention of the protein when the mixture was applied to a Ni-chelating resin. After protein removal, the solution was lyophilised and re-suspended in the minimum volume of water to allow purification by HPLC, as before. Alternative methods to isolate the purified product (such as "cofactor trapping" [186], based on *in situ* product removal (ISPR)[187,188]) were considered, but were deemed to have an unacceptably low specificity to be used without further intensive study.

3.2.3 Conclusion – formation of FMN analogues using Sp-RfK

The Sp-RfK gene (supplied by a collaborator, Dr. Hannah Collins, University of Canterbury), was expressed and purified in very high yield, and used to successfully phosphorylate the 1-deaza- and 5-deaza- analogues of riboflavin formed previously. Cofactor formation was successfully achieved using as little as 1 nmol of protein per 130 nmol of riboflavin analogue, with 1.2 equivalents of ATP, rendering this enzyme significantly more efficient than the Ca-RfK-FADS described in section 3.1. Purity of the FMN analogue formed by this reaction was very high (greater than 99 % by HPLC).

3.3 AsLOV2 Sample Preparation

3.3.1 Introduction to the AsLOV2 construct

The *Avena sativa* PHOT1-LOV2 gene was previously incorporated into a modified pNCO113 vector by others prior to this work. Modification of the pNCO113 plasmid vector[189] was first performed by Kay *et al.*[4] by insertion of the gene encoding hisactophilin from *Dictyostelium discoideum*[190] in the MCS of the plasmid, between *EcoRI* and *BamHI* restriction sites. The pNCO113-HISACT plasmid was further modified by Kay *et al.* to incorporate a thrombin cleavage site, followed by modification of cysteine 49 of hisactophilin to serine (pNCO113-HISACT(C49S)). Modification of hisactophilin was necessary to remove the only cysteine residue of the protein, ensuring that the results of future studies using AsLOV2 were due only to the effects of the conserved cysteine of the LOV domain. Finally, the *Avena sativa* PHOT1-LOV2 gene was incorporated between *BamHI* and *HindIII* sites within the MCS of the modified vector by M. Fisher, to give the final plasmid pNCO113-HISACT(C49S)-AsLOV2[1,4], which was provided by the project supervisor to enable further study of this flavoprotein.

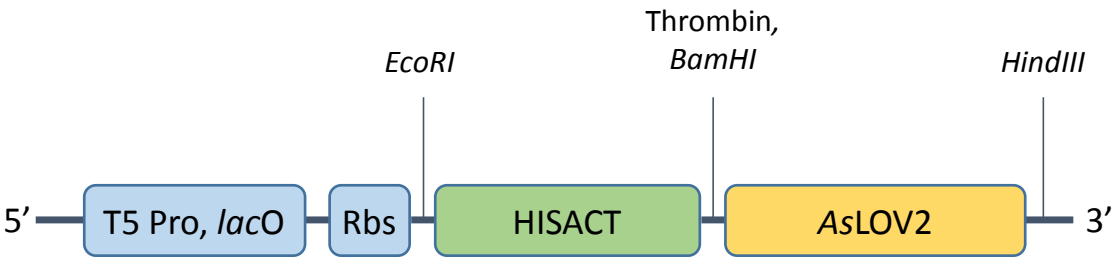


Figure 77: Schematic representation of the hisactophilin-AsLOV2 construct (not to scale). "T5 Pro, lacO" refers to the T5 promoter and lac operator, followed by "Rbs", the ribosomal binding site. HISACT and AsLOV2 domains are linked by 6 residues

Expression of this plasmid results in a fusion construct of hisactophilin and AsLOV2, comprising a total of 297 amino acids with a mass of 34.3 kDa; 133 residues for hisactophilin (15.5 kDa)[190], the 6-residue linker sequence Leu-Val-Pro-Arg-Gly-Ser (0.6 kDa), and 158 residues for AsLOV2 (18.2 kDa)[4]. This plasmid is subsequently referred to as "pNCO-AsLOV2", and when expressed, the resulting fusion protein is referred to as AsLOV2 for brevity, as (in almost all of cases) the hisactophilin tag was retained (see section 3.3.3). In later sections, mutated variants of the wild-type protein are referred

to as AsLOV2(C450X), where X corresponds to an amino acid substituted for the original reactive cysteine.

As previously described in section 1.1, the LOV2 domain of *A. sativa* undergoes a conformational change upon irradiation with blue light (440 - 450 nm). The absorbance of this wavelength by the flavin chromophore stimulates bond formation between the cofactor and a neighbouring cysteine residue within the active site of the protein. This may be conveniently observed photometrically, as the photoadduct lacks the flavin absorbance at 445 nm, although this spontaneously recovers upon cessation of irradiation. Covalent bond formation initiates the physical undocking of an α-helix (known as the *J*-α helix) from the central protein β-sheet, which allows activation of the coupled protein kinase in the full-length phototropin.

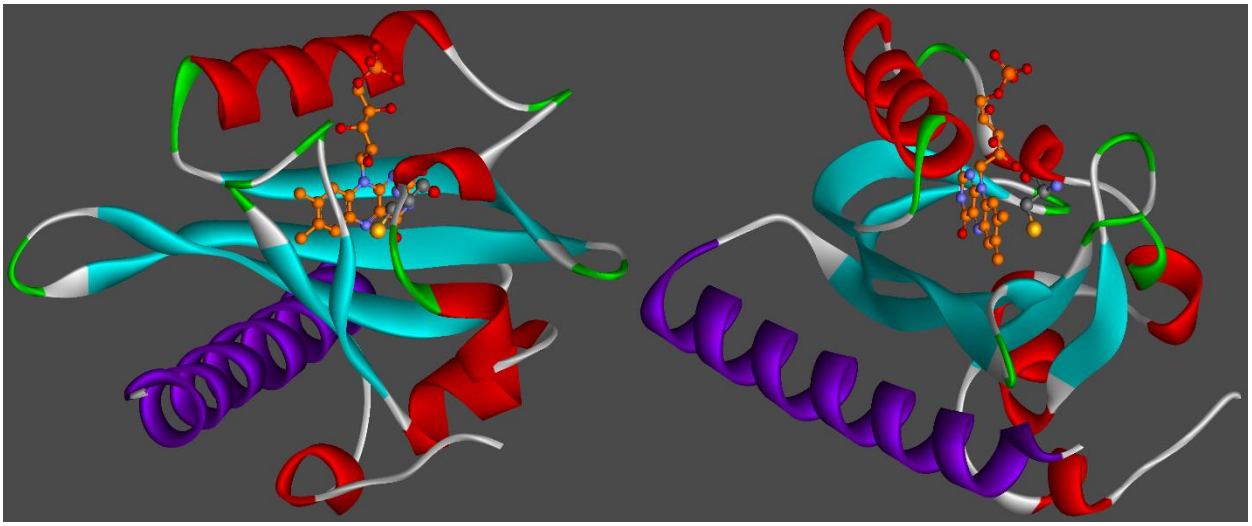


Figure 78: Two views of the dark-state crystal structure of the AsLOV2 domain, from Protein DataBank sequence 2V0U[48]. The J-α helix is highlighted in purple; FMN (orange) is located within the central binding site of the domain, in close proximity to the sulphur atom (gold) of the reactive cysteine residue

As recombinant protein expression terminates at the C-terminus of the *J*-α helix, irradiation of the sample causes this helix to become free within solution, leading to a corresponding decrease in α-helicity upon successful photoadduct formation. Thus, the photochemical and physical effects of the LOV2 domain's reaction with flavin may be measured using UV-visible and CD spectroscopy; both are convenient, fast and non-destructive methods, which do not affect the protein sample under analysis.

3.3.2 Expression strain selection

Prior to use, the pNCO-AsLOV2 construct was sequenced, including regulatory regions. For initial protein expression, the plasmid was transformed into *E. coli* strain BL21(DE3) cells, although (as fully described in section 3.3.5) expression of AsLOV2(WT) using this strain was found to be unreliable (AsLOV2(C450X) mutants were also not reliably expressed in this strain). While several approaches were taken to understand and address this problem (*vide infra*), the use of alternative *E. coli* strains was examined to improve protein expression.

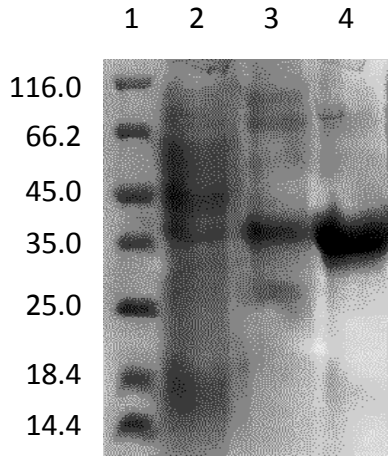


Figure 79: SDS-PAGE gel of AsLOV2(WT).

Lane 1 - weight marker; lane 2 - cell lysate; lane 4 - purified AsLOV2 (34.3 kDa)

Initially, two derivatives of BL21(DE3) were examined (BL21(DE3) STAR, and BL21(DE3) [pLysS], where co-plasmid pLysS assists in regulation of transcription of other plasmids within the cell), however no improvement in protein expression was found. Instead, strain M15[191] (a derivative of *E. coli* K12), was examined for use with both the wild-type and mutant plasmids. The name is derived from the deletion of a single amino acid (M15) in the *lacZ* gene, which is part of the larger lac operon; this deletion renders the *lac* operon ineffective for metabolism of lactose, assisting in the regulation of genes under control of a *lac* promoter. This is of particular use for plasmids under control of a T5 expression system, as a *lac* promoter sequence is incorporated into the sequence of T5 (see section 3.3.2.1). Furthermore, strain M15 used for this project also contained the repressor-plasmid pRep4, which assisted in the regulation of

protein expression by causing expression of the *lac* repressor protein (encoded by the *lacI* gene[192]), until addition of IPTG to the cell culture caused inactivation of this gene and simultaneously permitted transcription of the (previously blocked) plasmid gene.

A second alternative strain used in place of BL21 was strain JM101[193]. As for M15, this is also a derivative of *E. coli* K12, but is a less popular choice for plasmid expression due to poor regulation of lactose production within the cell, often leading to unregulated induction with consequent low levels of recombinant protein production. However, it is often successful at expression of difficult genes, due to the retention of a variety of non-optimised transcriptional machinery[193–195].

Expression of wild-type AsLOV2 using M15[pRep4] was found to be generally as effective as strain BL21(DE3) used previously, while expression of the mutant plasmids remained low in this strain. However, the use of strain JM101 was able to successfully allow reliable expression of each mutant formed as described in section 3.3.4, with the expression methodology itself described in greater detail in section 3.3.5.

Two reasons for the experimentally-observed poor levels of expression were identified following a detailed analysis of the plasmid vector and the DNA sequence of the AsLOV2 gene, and are described below.

3.3.2.1 Plasmid promoter sequence

The promoter sequence of the pNCO-AsLOV2 plasmid, originally derived from bacteriophage T5[196], was not ideally suited for use in *E. coli* strains optimised for alternative promoter sequences. Strain BL21(DE3) was originally developed by incorporation of the T7 RNA polymerase (RNAP) from bacteriophage λ[197] (see section 5.7.1 for genotype), and is thus optimised for expression of genes under control of a T7 promoter, such as pET-14b used for the successful expression of Sp-RfK (section 3.2) previously.

The T5 promoter sequence is significantly longer than other promoter sequences commonly used for recombinant gene expression as it contains an embedded *lac*

operator, which is more usually sited immediately downstream of the promoter. Little homogeneity is found between promoter sequences, as they are specific to a sigma-factor (σ-factor) responsible for identification of the promoter, and initial binding of the RNA polymerase (RNAP). For effective expression of recombinant genes within bacterial systems, the host cell requires both an appropriate σ-factor and RNAP, and so development of strains which express high levels of a defined RNAP has been used to optimise expression for a specific corresponding plasmid.

5'-TCATAAAAAATTTATTTGCTTTGTGAGCGGATAACAATTATAATA-3'

5'-TAATACGACTCACTATAGGGAGA-3'

Figure 80: T5 promoter sequence, with Lac operator *sequence the promoter underlined (above), and T7 bacteriophage promoter sequence(below)*

A promoter sequence may occasionally be recognised by an RNAP which is not optimised for the sequence for a number of reasons, although transcription from these is significantly less effective[198]. The high specificity of the T7 bacteriophage RNAP is well reported[199–201], with single point mutations to the polymerase completely preventing the expression of genes under control of the T7 promoter[202,203]. Therefore, as the specificity of the T7 RNAP is very high, it is unlikely for reliable expression of genes under control of a T5 promoter to proceed in strains optimised for the T7 promoter sequence.

This would explain the inconsistency observed with the expression of AsLOV2, and was addressed using two alternative *E. coli* strains for expression (M15 and JM101; *vide supra*), which are not optimised for expression using a single promoter sequence.

3.3.2.2 Codon optimisation

Another potential reason for the inconsistent expression of AsLOV2 was due to an apparent oversight in the DNA sequence of the AsLOV2 gene. When originally isolated and incorporated within the plasmid[1,4], no consideration was made for the codon-optimisation of the AsLOV2 gene for expression in *E. coli*; instead, the original codons were retained from the host species. Codon utilisation by plants is significantly

different to the codons utilised by prokaryotes (typically at the third base pair of the codon[204]) with a preference for guanine and cytosine rather than adenine and thymine in this position[205]. Although the standard genetic code remains the same (leading to the correct amino acid sequence in the translated protein), the utilisation of each codon combination is specific to the host (with full details available for *Avena sativa*[204,205] and *E. coli*[206]). Thus, the use of rarely-utilised codons will lead to retarded or failed expression, furnishing lower levels of the recombinant protein.

Similar difficulties have been encountered for other flavin-mediated phototropin domains, including *Arabidopsis thaliana* PLP[207] and PAS/LOV domains[208], where mutation of each sequence to optimise codon usage vastly improved protein yields. As the effect of codon optimisation was not considered during this project (due to the previously reported successful expression[1,4]), this discussion is intended to assist future work.

Utilising the web-based OPTIMIZER tool[209] (based upon methodology described elsewhere[210]), the optimal codon usage of the AsLOV2 gene was compared to the codon-usage of *E. coli* K12. This generated an alignment report (shown in full in appendix 2) which suggests that of the 477 nucleotides (159 codons) of the AsLOV2 gene, 97 mutations (20 % of all nucleotides) would be necessary to fully optimise the gene for expression in *E. coli*. This startling result demonstrates the difference in codon usage, with significant mutagenesis of the AsLOV2 gene necessary to improve the efficiency of expression in bacterial hosts.

3.3.3 Influence of Hisactophilin on AsLOV2 photocycle

The primary role of the N-terminal Hisactophilin protein was to act as a convenient tag for IMAC purification, as the high number of histidine residues (31 out of 118 [190]) ensures this protein has an exceptionally strong affinity for metal ligands. This allowed isolation of the target protein using IMAC (also used to purify Sp-RfK, described earlier in section 3.2.1). A discussion of the IMAC protocol is given in section 3.3.6 below.

Cleavage of hisactophilin from AsLOV2 was possible using the protease thrombin, which selectively cleaves the amino acid sequence 5'-Leu-Val-Pro-Arg-Gly-Ser-3', between arginine and glycine, with very high specificity. This sequence was used to link the proteins during original construction of the hisactophilin-AsLOV2 construct[4], and gave an effective method for the selective removal of the protein tag, if desired.

An additional benefit of using hisactophilin instead of a shorter affinity protein (such as 6xHis) was a reported enhancement to both the stability and the photochemistry of the AsLOV2 domain. During the work detailing the original plasmid formation[4], a calmodulin-binding protein (CBP)[25,31] construct was also produced, and used to compare the properties of the AsLOV2 photocycle when coupled to either protein. During these studies, an inactive mutant of the AsLOV2 domain (C450A) was used to prevent formation of the photoproduct.

Spectral measurements of the ground (i.e. dark) state of AsLOV2(C450A) made using both CBP and hisactophilin-fusion constructs, were reported to be essentially unaffected by the presence of either tagging protein[4]. However, hisactophilin-tagged AsLOV2(C450A) gave a significantly higher radical yield upon illumination, compared to either CBP-fused, or free- AsLOV2(C450A) (shown in table 8, below). As radical formation is key to the photoreaction of the AsLOV2 domain, the retention of hisactophilin therefore gave two beneficial effects: (i) improved protein stability, enabling storage and reliable observation over extended periods; and (ii) an enhanced radical yield for the protein, improving the sensitivity of photometric measurements.

While the results presented by Kay *et al.*[4] (table 8, below) support the retention of hisactophilin during spectroscopic examination, the photoquenching they report using the C450A mutant domain is not normally observed[56,139,211,212]; furthermore, this was also not observed during this project using an AsLOV2(T418I, C450A) double mutant. However, it was assumed that the reported effect of the hisactophilin on the C450A mutant would be identical for the wild-type protein, and that along with the improved protein stability, its retention would be generally beneficial. Therefore, the

majority of the experiments reported in this thesis were performed using the hisactophilin-AsLOV2 fusion protein.

Protein Sample	Condition	Radical Yield (%)
LOV2 C450A-HISACT	Aerobic	90
	Anaerobic	92
	Aerobic + EDTA	97
	Anaerobic + EDTA	100
LOV2 C450A-CBP	Aerobic	29
	Anaerobic	57
LOV2C450A	Aerobic	20
	Anaerobic	56

Table 8: Reproduced from "Table II" of the paper by Kay et al.[4], listing the relative radical yields of AsLOV2(C450A) as the lone domain and with hisactophilin (HISACT) or calmodulin binding protein (CBP) fusion partners. Radical yields were normalised to the greatest radical yield (HISACT, in anaerobic solution with additional EDTA)

3.3.4 Site Directed Mutagenesis (SDM)

Several mutant proteins were designed and expressed during this project, each with mutation to the single cysteine residue of the AsLOV2 domain which is responsible for formation of a covalent bond with flavin during the phototropin photocycle. These C450X mutants (where X indicates the alternative amino acid incorporated) used alanine, serine, methionine, glycine and aspartic acid in place of cysteine, and were formed by mutation of the codon within the original plasmid via whole-plasmid mutagenesis using the polymerase chain reaction (PCR) method of Weiner *et al.*[213]. This method was used due to much shorter turnaround times compared to mutagenesis of the isolated gene, however, this method has two associated risks. Firstly, there is potential for undesired mutation at a random position of the plasmid by incorrect incorporation of a nucleotide by the DNA polymerase. Secondly, PCR cycle times may be required to be significantly extended (to 15 or 20 minutes per cycle) to allow replication of many kilo-bases of DNA. To mitigate these risks, all mutant products were sequenced to confirm the identity of the target gene, and a range of times for the elongation phase of the PCR was used to find the optimal duration.

Mutagenesis was performed using purified DNA (previously maintained in *E. coli* strain XL1-blue before isolation). Prior to mutagenesis, the wild-type (also referred to as template) DNA was sequenced to confirm its fidelity, along with a phenotypical examination (via expression) to confirm that the protein expressed was fully functional. As *E. coli* strain XL-1 blue contains both *dam+* and *dcm+* genes, all DNA within the cell (including plasmid DNA) is methylated at both adenosine and cytosine nucleobases (respectively) by DNA methylases. However, the "daughter" DNA formed by the *Pfu* polymerase during the PCR process remained unmethylated, giving a useful differential to allow selection of only the mutant DNA from the PCR mixture. The parental (methylated) DNA was successfully removed by digestion using *DpnI* restriction endonuclease, which cuts methylated DNA (recognising adenosenyl methylation in the sequence 5'-GA*TC-3'). Therefore, only daughter DNA remained after digestion, which was expected to have incorporated the desired mutation.

To perform mutation using the PCR, two primers were required, each complimentary to the sense or antisense template DNA strands and containing the desired mutation as a mismatch in the codon encoding the amino acid to be modified. In all cases, primer length was 40 b.p., in order to give a sufficient overlap on both sides of the mismatch to allow strong binding of the primer during the annealing phase of the PCR cycle. Sixteen complete cycles of the PCR were used to give optimal yields of DNA.

As described above, one of the greatest risks of whole-plasmid mutagenesis is the potential for undesired additional mutation of the daughter DNA by the polymerase enzyme, which becomes more likely with longer DNA sequences or additional cycles of the PCR. Therefore, selection of an appropriate DNA polymerase (DNAP) was key. Numerous DNAPs are available, with *Taq* (from *Thermus aquaticus*[214]) and *Pfu* (from *Pyrococcus furiosus*[215]) being most common. Although *Taq* performs elongation more rapidly than *Pfu* (1,000 nucleobases in 10 seconds using *Taq*[216], versus a recommended allowance of 1 minute per 1,000 nucleobases with *Pfu*), *Taq* has a significantly lower reported fidelity for replicated DNA (1.1×10^{-4} spontaneous mutations per base pair per duplication for *Taq*[217], compared to 1.3×10^{-6} for *Pfu*[218]).

This is due to the presence of 3' → 5' exonuclease activity for *Pfu*, removing incorrectly incorporated nucleobases during replication; *Pfu* was thus used for mutagenesis.

Following digestion with *DpnI*, portions of each PCR mixture were transformed into *E. coli* XL1-blue, with successful transformants identified by growth on LB-agar plates containing ampicillin. Initially, no transformants were obtained, which was determined to be due to the DNA polymerase working less efficiently than expected and thus requiring longer times to complete DNA replication. An elongation time of 8 minutes was originally used, which was gradually increased to fifteen minutes per PCR cycle, giving optimal yields of the full-length DNA product. Agarose gel electrophoresis (figure 81) confirmed formation of correctly-sized daughter DNA, which was subsequently successfully transformed into XL1-blue with reasonable efficiency (3.3 x 10^3 to 4.7 x 10^4 transformants per µg of DNA). Small liquid cultures were used to provide sufficient quantities of DNA for sequencing.

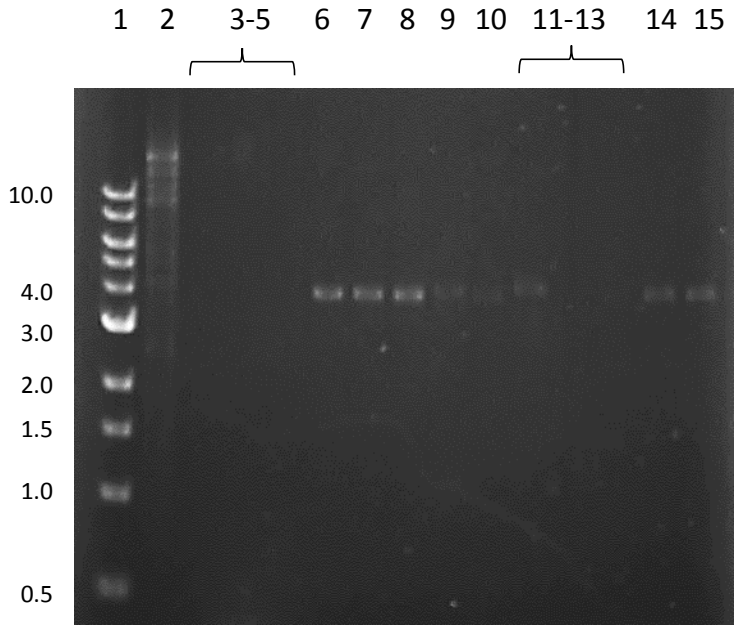


Figure 81: Agarose gel of PCR products from AsLOV2 mutagenesis. All DNA (except lane 2) nicked using BamHI restriction endonuclease.

Lane 1 - 1 k.b.p. DNA ladder; lane 2 - undigested plasmid DNA; lanes 3-5 - DNA products with PCR elongation times of 8, 10 and 12 minutes (respectively); lane 6 - C450A mutant (4.2 kBp.); lane 7 - C450S mutant; lane 8-C450M mutant; lane 9 - C450G mutant; lane 10 - C450D mutant; lanes 14,15 - AsLOV2(WT) (4.2 k.b.p.)

Sequencing of C450S, C450G, C450D and C450M mutants confirmed successful mutagenesis, with no other mutation found throughout in the gene compared to the template DNA. However, sequencing data for the C450A mutant showed a spontaneous mutation present in all transformant colonies sequenced. This was a single nucleobase (C to T) mutation in codon 418, which previously encoded threonine (sequence ACT). The effect of the mutation changed this codon to A<u>T</u>T, which instead encodes isoleucine, leading to an undesired double-mutant. However, as the effect of this mutation was expected to be minimal, and the effects of the AsLOV2(C450A) mutant are particularly well characterised[4,31,54,208,219], the AsLOV2(T418I, C450A) double mutant was used in its place.

5'- GTC ATT ACT GAC CCA-3' 5'-Val Ile Thr Asp Pro-3'

5'- GTC ATT A<u>T</u>T GAC CCA-3' 5'-Val Ile <u>Ile</u> Asp Pro-3'

Figure 82: DNA and translated protein sequences of AsLOV2(WT) (above), and undesired double mutant AsLOV2(T418I, C450A) (below). Mutation site is underlined

3.3.5 Expression of wild-type and mutant AsLOV2 proteins

After sequencing of the PCR products, the mutant plasmids were transformed into a number of *E. coli* strains (BL21(DE3), BL21(DE3) STAR, BL21(DE3)[pLysS], M15[pRep4] and JM101) to evaluate relative rates of expression, as discussed in section 3.3.2. While *E. coli* strains BL21(DE3) and M15[pRep4] were found to be reasonably effective at expression of AsLOV2(WT), the expression of the mutant derivatives was particularly unreliable in these strains. This was not believed to be due to the effect of mutagenesis (as the codon selected for mutation was optimised to be the most-used codon for *E. coli*), but for other reasons discussed earlier in sections 3.3.2.1 and 3.3.2.2. Protein yields obtained using strain JM101 were acceptable, and so this was used to express all mutant plasmids.

Culture growth for over-expression was performed in a liquid medium. In all cases, protein yields were greatest using a highly enriched medium, with typical wet cell masses (for BL21(DE3) pNCO113-AsLOV2(WT)) of 18-26 g L^{-1} in TB medium, compared

to 6-8 g L^{-1} using LB medium; similar results were obtained for mutant plasmids within strain JM101. Successful over-expression was indicated visually (and observable photometrically at 445-450 nm) by the vivid yellow colour of the cells, as a result of FMN incorporation within the overexpressed protein. The temperature (post-induction) appeared to play a role in the yield of protein obtained, with 16-18 °C found to be optimal. At elevated temperatures, a significant amount of misfolded or otherwise unstable protein was found in the cell lysate, which spontaneously precipitated upon lysis or caused rapid culture death during expression. As incorporation of FMN within AsLOV2 was integral to protein stability (the apoprotein spontaneously precipitated, see section 3.3.9.2), the use of lower temperatures aided AsLOV2 expression by allowing the cell sufficient opportunity to form the FMN required for incorporation within the recombinant protein. Notably however, supplementation of the growth medium with either riboflavin or FMN was not found to assist protein expression.

Other attempts were made to improve levels of protein expression, including performing all stages of cell growth with the exclusion of light (due to the photoreaction of the expressed AsLOV2 protein), using increased concentrations of IPTG (up to 5 mM), or using freshly-transformed competent cells. However, no conditions were found to significantly improve protein yields; while it was possible to produce sufficient levels for use during this project, protein expression remained unreliable. This supports the earlier argument (section 3.3.2.1 and 3.3.2.2) that sub-optimal promoter sequence or codon usage within the plasmid dramatically affects expression of the protein.

3.3.6 Purification of expressed AsLOV2 samples

All expressed AsLOV2 proteins (WT and C450 mutants) were lysed using sonication, although a more gentle procedure than previously used was necessary, as vigorous lysis led to fragmentation of the fusion protein.

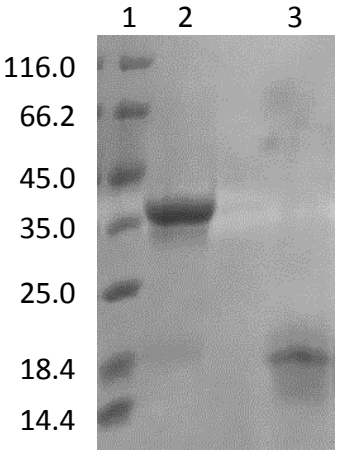


Figure 83: Comparison of AsLOV2(WT) lysed and purified gently (lane 2; see section 7.3.2), and AsLOV2(WT) lysed using the same protocol as Sp-RfK purification and collected in the flow-through from IMAC purification (lane 3)

Lysis and purification was performed at pH 8.0, which was found to be optimal to minimise the effects of pH-dependent protein aggregation. At this pH, the excited state of the photoreaction is also reported to be optimally stabilised (at physiological temperature)[219]. The crude lysate solution (which exhibited a strong yellow colour; figure 84) was purified using a high-throughput IMAC method (section 5.6.7.2). Due to the strong affinity of hisactophilin for metals, the protein became immobilised on nickel-chelating resin, while other non-tagged proteins passed through the column unimpeded. Two types of resin were evaluated: Ni-NTA® and Ni-Sepharose® fast flow. NTA resin was found to be greatly inferior to Sepharose resin, as it appeared unstable under the pressure applied to the column using a peristaltic pump.

As the target protein exhibited a strong yellow colour (due to the presence of FMN within the LOV domain), binding of this protein was observed visually as the vivid yellow lysate (figure 84, left) was rapidly decolourised (figure 84, right) by passing over the resin column. The binding of the yellow protein to the blue column produced a characteristic green colour, which aided assessment of the column capacity.

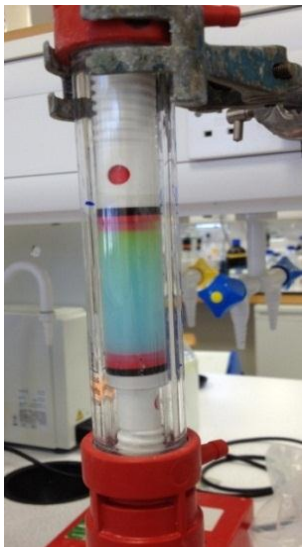



Figure 84 [left image]: Vivid yellow AsLOV2(WT) crude lysate solution (left), and almost colourless flow-through from the column after protein binding (right)
Figure 85 [right image]: AsLOV2(WT) immobilised on Ni-chelating Sepharose resin. The green colour localised at the top of the resin bed indicated retention of AsLOV2

After binding and washing of the retained protein samples, the AsLOV2 domains (wild-type or mutant) containing the native cofactor FMN were eluted from the column using a high concentrations of imidazole (0.5-1.0 M). After elution, protein-containing fractions were dialysed or applied to a desalting column, to remove imidazole.

3.3.7 Removal of imidazole

The effect of imidazole on the photoreaction of LOV domains is remarkable. It is able to dramatically increase the rate of reversion for the photoadduct, with this effect observed in a broad range of LOV domains including *Avena sativa* LOV1 and LOV2[58], *Neurospora crassa* VIVID (VVD) LOV[39], and *Solano lycopersicum* LOV/LOV protein (LLP)[207]. In *S. lycopersicum* LLP, an imidazole concentration of 200 mM decreased the half-life for the photoproduct approximately 64 fold (from 65.4 minutes to 61.4 seconds)[207]. However, for *A.sativa* LOV2, a considerably lower imidazole concentration (20 mM) caused a much greater increase (100-fold) in the rate of photoadduct reversion[58].

Several potential mechanisms have been proposed in the literature, including the direct action of a base (either intra- or extra- protein)[58], interruption of the hydrogen bonding network[58], or a steric interaction stabilising the ground state of the complex[220]. The predominant mechanism, postulated by Alexandre *et al.*[58] (shown in figure 86), is supported by their analysis of the protein crystal structure, where a potential tunnel and cavity was found to within 4-5 Å of flavin (under certain kinetic models), of a sufficient size to permit theoretical incorporation of imidazole in the reactive site.[58]

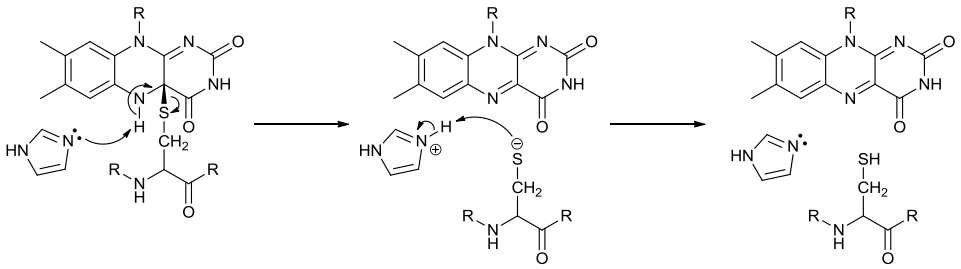


Figure 86: Ionic mechanism proposed by Alexandre et al.[58] for the imidazole-catalysed reversion of the FMN-LOV photoadduct

As the presence of these reported voids may be dynamic (expanding and contracting due to structural changes of the protein in “dark” and “light” states), the incorporation of imidazole via these is expected to be temporary, related to its concentration in solution. Thus, removal of imidazole from the protein solution was crucial for the accuracy of further studies.

While dialysis was appropriate for large volumes of protein solutions (up to 250 mL), the time required to perform the requisite number of repetitions (typically 8-10) made this method particularly time-consuming. Alternatively, the use of size-exclusion or desalting columns was both swift and effective, but limited by the volume of protein sample which could be accommodated by the column (typically 5-10 mL per loading). A combination of both techniques was used to remove imidazole from protein samples, with dialysis remaining the preferred method for large-scale purification.

3.3.8 Spectroscopic determination of protein concentration

Estimation of AsLOV2 concentration within a sample was performed photometrically using the absorbance of the chromophore within the domain. While free-FMN in solution has an extinction coefficient (ε) of 12500 M^{-1} cm^{-1} at 445 nm[88], the effects of binding within the LOV domain cause a slight red-shift of the flavin absorbance to 447 nm, and an enhancement of the extinction coefficient to 13800 M^{-1} cm^{-1}, as determined by Salomon *et al.*[31]. Cofactor binding is indicated by resolution of the spectrum between 400-500 nm, showing two distinct vibronic side-bands at 422 nm and 473 nm (figure 87). These each correspond to vibronic energy levels for low energy singlet (excited) states of FMN[219], and appear highly resolved as only a single environment exists for the bound cofactor.

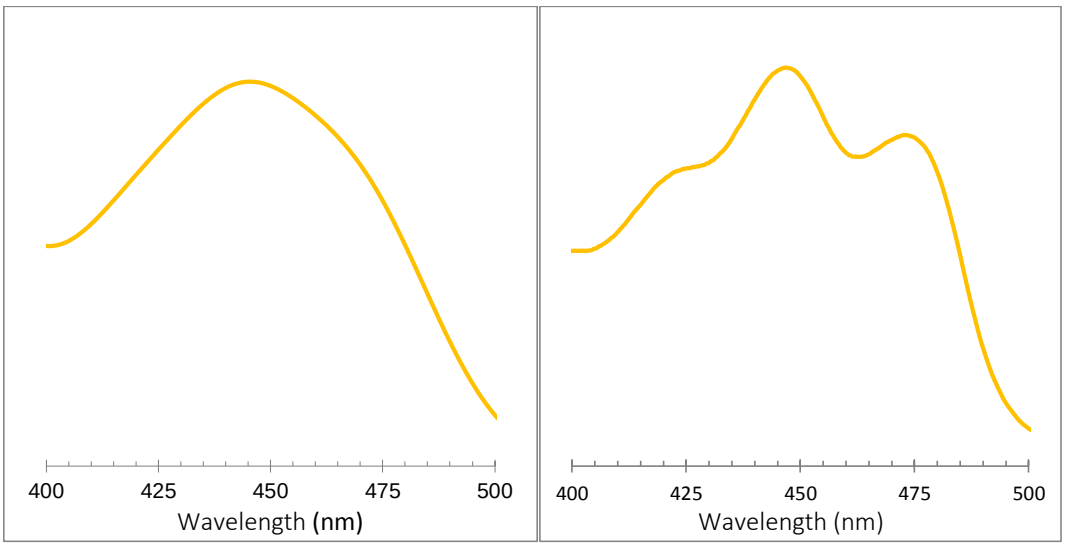


Figure 87: UV-visible absorption spectra of free FMN in solution (left), and FMN bound within AsLOV2(WT) (right). While FMN exhibits a single broad absorbance at 445 nm, the spectrum of the protein-bound cofactor resolves to reveal three absorbances in the same region, with the major resonance remaining at 447 nm

It is interesting that the LOV1 domain from the same PHOT protein (which is somewhat similar in structure to the LOV2 domain) gives an extinction coefficient of 12200 M^{-1} cm^{-1} at 449 nm[31]; much closer to the value for the free cofactor in solution. This is reportedly due to differences between each domain in the flavin binding site, with LOV1 appearing to influence the cofactor less than LOV2[31]. A value of

13800 M^{-1} cm^{-1} at 447 nm was thus used (without correction) to estimate the concentration of AsLOV2 samples containing FMN.

For the evaluation of protein samples containing either 5-deazaFMN or 1-deazaFMN, protein concentration was estimated based upon the 'enhancement ratio' between the extinction coefficient value for protein-bound FMN (13800 M^{-1} cm^{-1} [31]) and free-solution FMN (12500 M^{-1} cm^{-1} [88]) of 1.104. Therefore, using 5-deazaFMN, the protein concentration was estimated using an extinction coefficient of 13800 M^{-1} cm^{-1} at 404 nm (from 12500 M^{-1} cm^{-1} at 404 nm unbound [88]), and for 1-deazaFMN, protein concentrations were derived using an extinction coefficient of 7507 M^{-1} cm^{-1} at 535 nm (from 6800 M^{-1} cm^{-1} at 535 nm unbound [88]).

3.3.9 Cofactor exchange within the AsLOV2 domain

3.3.9.1 Protein denaturation

This process was performed after the protein was immobilised on Ni-sepharose resin (used previously to purify the protein sample). Denaturing was originally attempted using solutions of three denaturing agents, at varying concentrations: urea (4.0-8.0 M), guanidinium hydrochloride (0.5-6.0 M) and guanidinium thiocyanate (0.5-4.0 M). Urea did not cause elution of FMN from the bound protein, and guanidinium hydrochloride was found to be unsuitable at all concentrations (due to the high concentration of chloride present in solution causing instability of the resin). However, guanidinium thiocyanate was found to be successful at denaturing the bound protein, causing immediate release of the native cofactor. This compound has been previously used to successfully denature recombinant proteins[221–223], with an optimal concentration between 1.8 and 4 M reported. During this research, it was found that thiocyanate concentrations as low as 0.5 M were able to cause displacement of the native cofactor, typically within 3-5 column volumes.

In order to ensure complete removal of the native flavin (along with traces of the denaturing agent), the bound apoprotein was washed with 2 column volumes of the washing buffer, or a saline solution (up to 2 %). This was key to the effectiveness of

the subsequent cofactor exchange, as the presence of remaining thiocyanate may have prevented successful renaturing of the protein.

3.3.9.2 Incorporation of artificial cofactors

The washed denatured protein was immediately treated using a solution of the purified FMN analogue (typically 20 mM), dissolved in the same buffer or saline solution used to wash the protein sample. The solution was continuously re-circulated over the column (with exclusion of light) at a low flow rate for a minimum of 4 hours, although times up to 16 hours gave greater analogue incorporation yields.

Successful incorporation of the synthetic cofactor allowed the previously denatured protein to re-form around the introduced cofactor. Subsequently, the column was washed to remove any unbound flavin analogue, and the reconstituted protein eluted as before (section 3.3.6). Each eluted protein adopted the characteristic colour and UV-visible characteristics of the cofactor introduced, allowing convenient identification of protein-containing fractions.

Removal of imidazole from samples containing synthetic cofactors was not performed immediately, due to the rapid precipitation of any apoprotein present within the mixture. This was found to be incredibly unstable due to the lack of structural integrity in the absence of the cofactor, and thus precipitated rapidly from solution. This allowed conveniently separation of the holo- and apo- protein forms, ensuring that subsequent examination of the protein samples was not influenced by this. Typical incorporation efficiencies above 80 % (estimated photometrically) were obtained using this method.

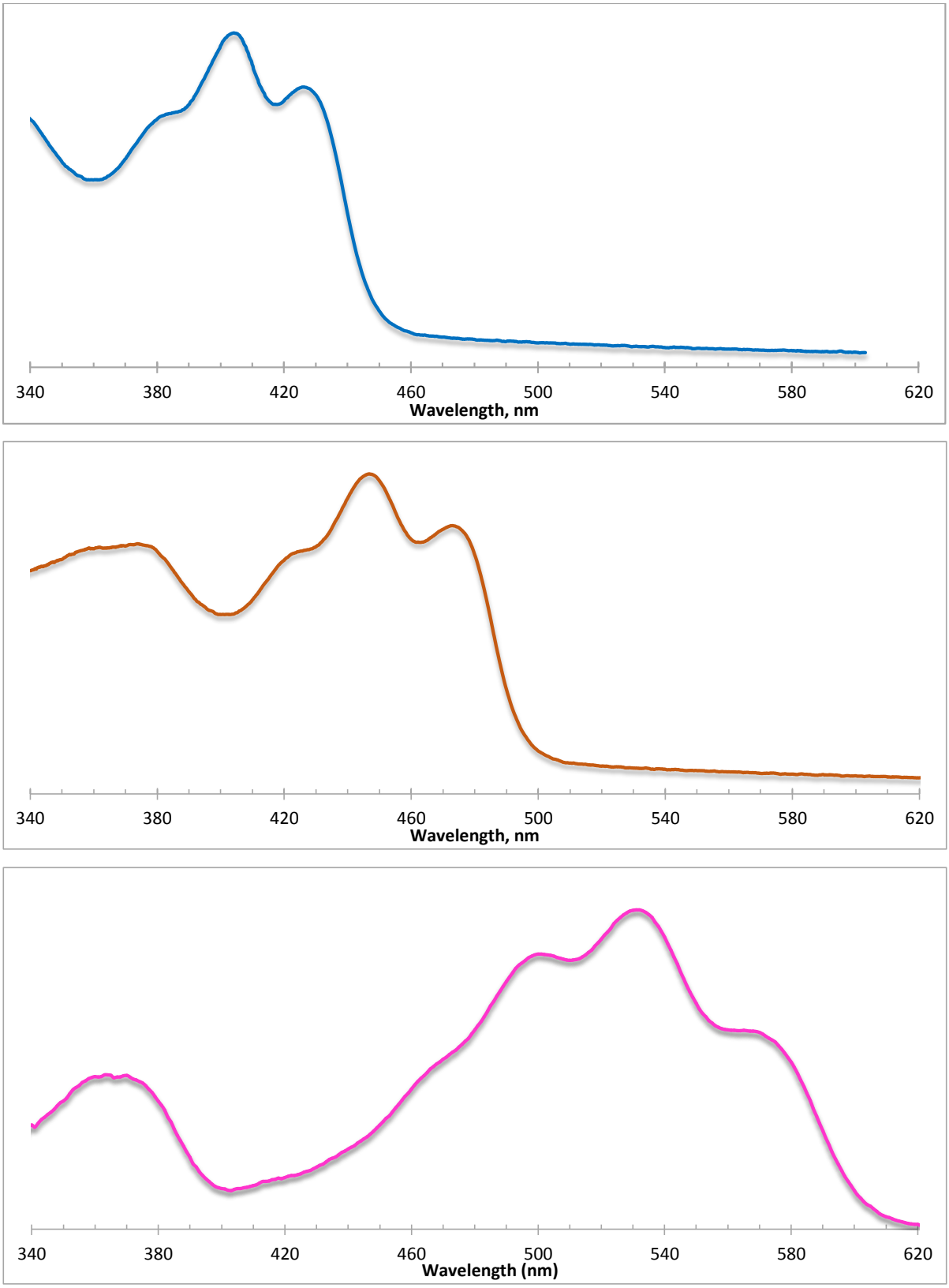


Figure 88: UV-visible absorbance spectra of AsLOV2(WT) containing 5-deazaFMN (top, blue), native FMN (centre, orange) and 1-deazaFMN (bottom, purple)

3.3.10 Conclusions – expression, preparation and purification of AsLOV2 samples

The expression and purification of the AsLOV2(WT) domain was performed, with some difficulty. Expression was found to be inconsistent, and was examined using a number of alternative *E. coli* host strains. However, little enhancement was made for the expression of the wild-type protein, with evidence found to suggest that the non-optimised codon usage of the gene was responsible. However, sufficient quantities of the protein for study were obtained by expression using nutrient-rich mixtures, and purified using the IMAC protocol described.

Additionally, the site directed mutagenesis of the original plasmid was performed, producing four desired mutants (AsLOV2(C450S, M, G and D)), and a single double-mutant (AsLOV2(T418I,C450A)). These were each expressed in strain JM101 (which was found to be most efficient), and also purified using the IMAC method used for the wild-type protein.

Furthermore, the artificial cofactors 5-deazaFMN and 1-deazaFMN were successfully incorporated into samples of each expressed protein, enabling subsequent photometric examination of the samples.

3.4 Photometric Examination of AsLOV2 (Wild-Type and Mutant), using Native FMN, 5-deazaFMN and 1-deazaFMN

3.4.0 Introduction

Protein samples prepared as described in section 3.3 were used to examine the photoreaction of the AsLOV2 domain. Two approaches were made: firstly, synthetically produced cofactors 5-deazaFMN and 1-deazaFMN were incorporated within AsLOV2(WT), to examine their effect on the photocycle of the domain. Secondly, the effects of mutating the reactive cysteine residue (C450) of AsLOV2 were examined, to determine if photoadduct formation was possible with similar amino acids. Each mutant was further examined with the artificially synthesised cofactors, to provide additional evidence.

Samples were examined using UV-visible and CD spectroscopy as appropriate. The strong UV-visible absorbance of the flavin chromophore allowed swift non-destructive analysis of photoadduct formation, with any stimulated physical effects observed by CD. Irradiation of samples was performed as described in section 5.6.8.2.

A range of protein concentrations was used, calculated using the methodology described in section 3.3.8, with the spectral effects of irradiation appearing relative to protein concentration. Due to the strong absorbance of FMN, typical protein concentrations between 15-20 µM were used for UV-visible spectroscopy of samples containing FMN or 5-deazaFMN, while a concentration of 25-35 µM was used for those containing 1-deazaFMN. All CD values were normalised using Mean Residue Ellipticity (MRE, $[\theta]_{mrw}$), based upon a Mean Residue Weight (MRW) of 115.9, and calculated according to the formula given below, to account for the differences in sample concentration.

$$[\theta]_{mrw} = (MRW \times \theta_{obs}) / 10 \times d \times c$$

Figure 89: Formula for calculation of MRE for normalisation of CD data, where MRW corresponds to Mean Residue Weight of the protein, θ_{obs} is the observed ellipticity, d is the path length (in centimetres) and c is the concentration of the solution, in g mL^{-1}

3.4.1 Photoreaction of AsLOV2(WT) with native FMN

3.4.1.1 UV-visible spectroscopy results

AsLOV2(WT) containing the native cofactor (FMN), expressed and purified as above (with retention of hisactophilin) was examined using UV-visible spectroscopy and CD, to obtain a basis for comparison with other results.

As shown in figure 90 below, the behaviour of the expressed AsLOV2-hisactophilin construct was found to be almost identical to that of the original plant domain, with irradiation of the sample at 447 nm leading to quenching of the flavin absorbance at 447 nm (and the associated vibronic side-bands). When irradiation ceased, the “dark-

state" (pre-irradiation) spectrum recovered spontaneously, with an approximate half-life of 30 seconds. This was in line with other similar results (28-32 seconds)[31,56].

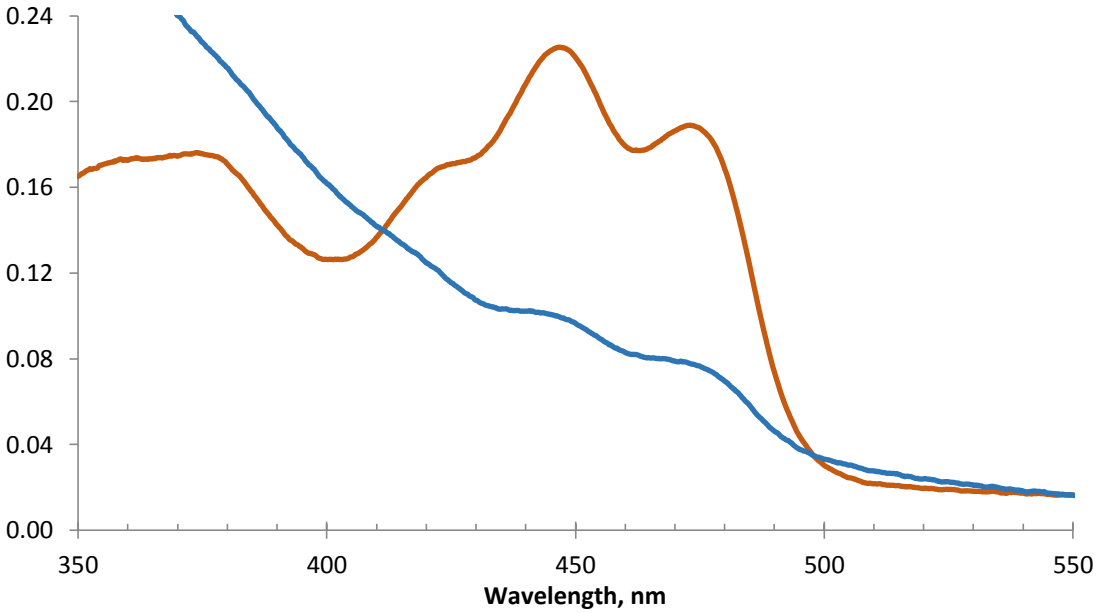


Figure 90: Photocycle of AsLOV2(WT) containing FMN. Irradiation of the sample at 447 nm caused quenching of the flavin absorbance from the "dark" state (orange) to the "light" state (blue). Recovery to the pre-irradiation spectra occurred with a half-life of ~30 seconds

3.4.1.2 CD spectroscopy results

The physical effect of flavin-protein bond formation in LOV domains has been thoroughly described in the literature[40,207,219,224,]. A partial unwinding of the *J*-α helix has been previously demonstrated using FT-IR and polarimetry[225–227], which is stimulated by photoirradiation of the domain. This effect was observed using CD spectroscopy of AsLOV2(WT) containing native FMN (fig. 91). Upon irradiation, the mean residue ellipticity in the region 240-205 nm decreased, indicating an overall decrease in α-helicity of the protein sample, which recovered to pre-irradiation levels within 600s (20 half-lives).

Based upon an analysis of the crystal structure of the AsLOV2 domain (PDB structure 2V0U[48]), 22 residues (of a total of 52 helical residues; thus 42.3 %) are within the *J*-α helix. The reduction in MRE observed was 30.9 % (at 208 nm), which although lower than predicted, is reasonable considering the short half-life of the light-state sample. Overall, the expected trend for a loss of ellipticity is broadly in line with expectations.

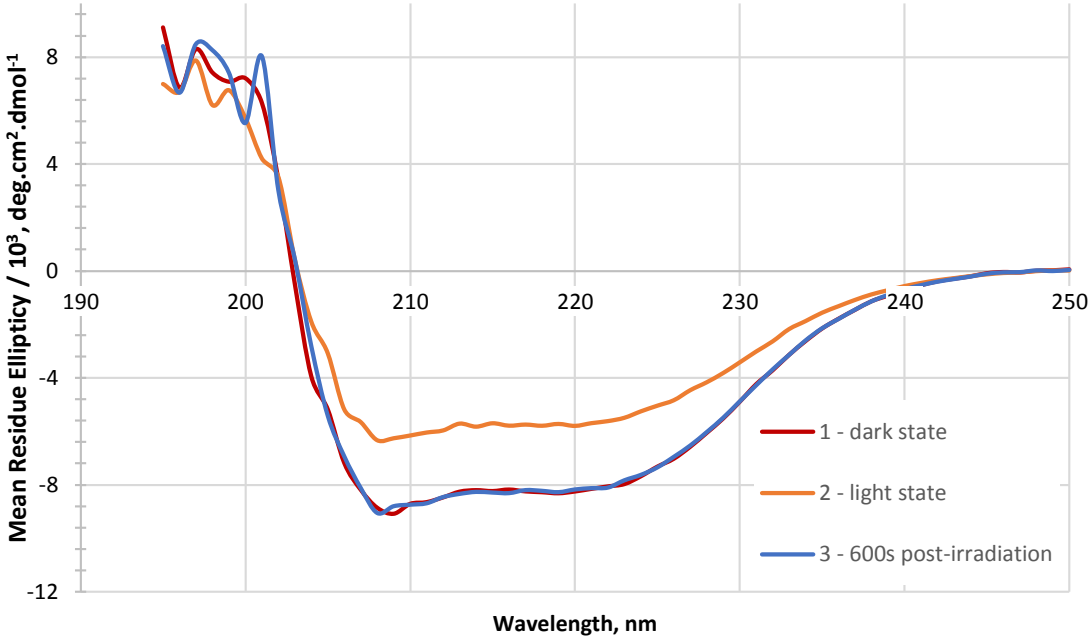


Figure 91: CD Spectrum of native FMN within the AsLOV2(WT)-hisactophilin protein

3.4.2 Photoreaction of AsLOV2(WT) with 5-deazaFMN

3.4.2.1 UV-visible spectroscopy results

The UV-visible spectrum of the AsLOV2(WT) domain was affected by the incorporation of 5-deazaFMN. The dark-state spectrum was blue-shifted by approximately 50 nm, in line with the shift observed between free FMN and 5-deazaFMN in solution.

Upon irradiation at 404 nm (the wavelength of 5-deazaflavin corresponding to 447 nm of FMN), a decrease in absorbance was observed in an analogous manner to AsLOV2(WT) containing the native cofactor (figure 90). However, the photoproduct formed was found to be stable, and did not revert spontaneously; both the fully-irradiated "light state" and partially irradiated intermediate states were stable for over 72 hours. Furthermore, this result was also observed in the presence of up to 200 mM imidazole.

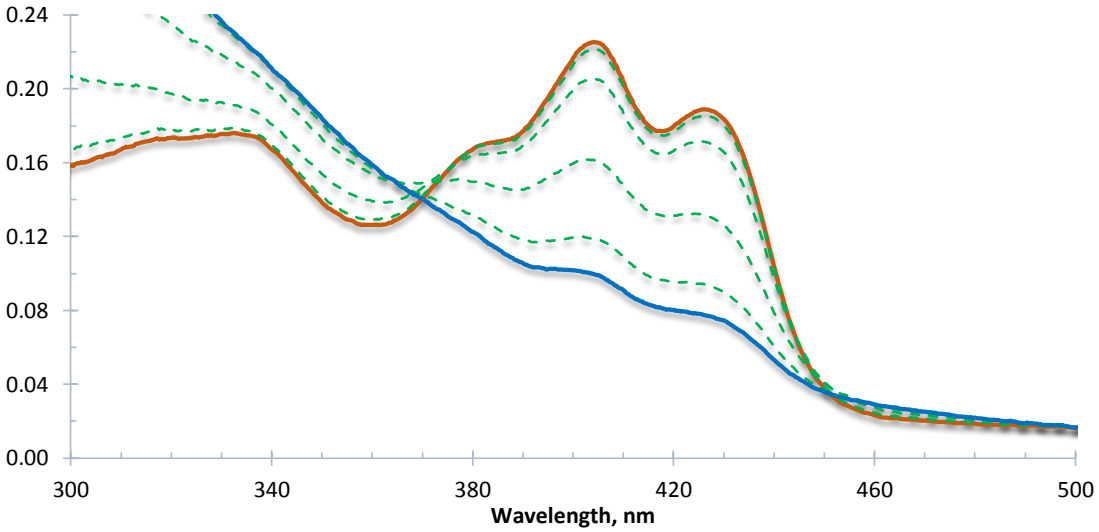


Figure 92: UV-visible spectrum of light-state formation of 5-deazaFMN within AsLOV2(WT); the primary absorbance is at 404nm, with two side-bands at 428 nm and 384 nm. The dark-state sample (orange) was irradiated at varying intervals using monochromatic light (404 nm), which caused a sequential decrease in the absorbance spectra observed (green, dashed), until the sample became fully photobleached (blue)

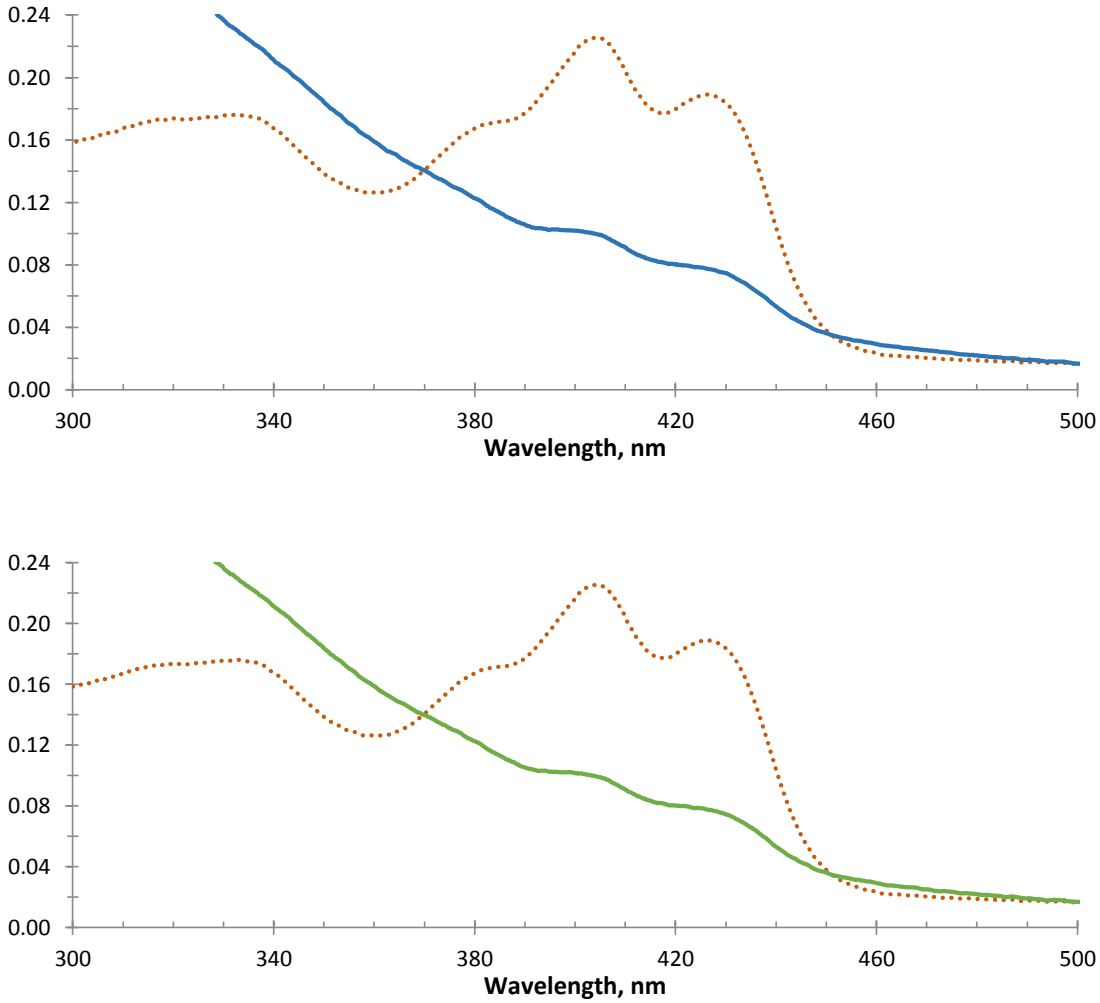


Figure 93: UV-visible absorption spectra of the light-state photoadduct of 5-deazaFMN-AsLOV2(WT). After irradiation at 404 nm from the dark-state (orange, dotted), the light-state spectrum obtained immediately following irradiation (blue, above) remained unchanged after 72 hours (green, below) in the dark at 4 °C

While the photoadduct was found to be stable, it was not immediately apparent if the dark-state spectrum would be recoverable. Numerous wavelengths were examined in order to determine if this may be achieved photometrically, with light in the ultraviolet region of the spectrum found to be effective at inducing reversion of the photoadduct. After somewhat exhaustive study, it was found that 313 nm was the optimal

wavelength to facilitate recovery of the spectral and physical properties of the dark-state sample.

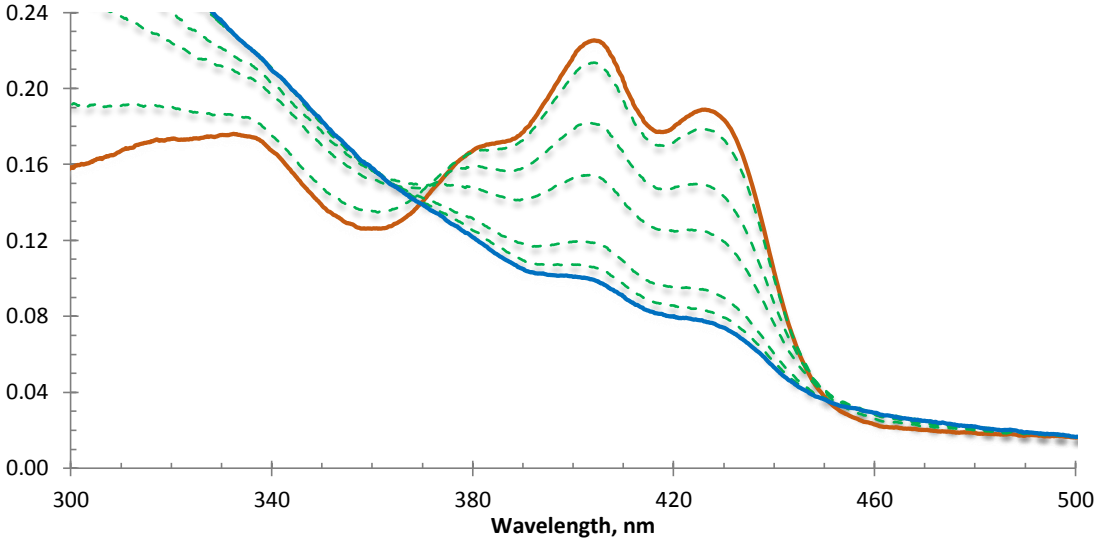


Figure 94: UV-visible spectrum of the dark-state recovery of 5-deazaFMN within AsLOV2(WT). The stable light-state sample (shown in blue) was irradiated periodically using monochromatic light (313 nm), facilitating reversion of the sample (sequentially; green, dashed) until the pre-irradiation spectrum (orange) was recovered. Note that the absorbance recovered is identical to that of the pre-irradiation dark state (figure 92), indicating that the photoreaction was both temporary and fully reversible

3.4.2.2 CD spectroscopy results

Examination of the 5-deazaFMN-AsLOV2(WT) sample using CD spectroscopy demonstrated a similar loss of ellipticity as found with native-FMN-AsLOV2(WT)[225–227] (section 3.4.1.2). However, the magnitude of the change was markedly different. While the reduction in ellipticity observed with the native cofactor was 30.9 % (at 208 nm), the 5-deazaFMN sample gave a 42.2% reduction at the same wavelength. As discussed above, the *J*-α helix is estimated to contain 42.3 % of the helical residues of the protein, and thus the observed loss of ellipticity very closely matches the predicted value. As found using UV-visible spectroscopy, the light-state sample was found to have no change in ellipticity over extended periods (600 seconds; 20 half-lives for

AsLOV2(WT) containing native FMN). As before, irradiation at 313 nm allowed recovery of the pre-irradiation spectrum.

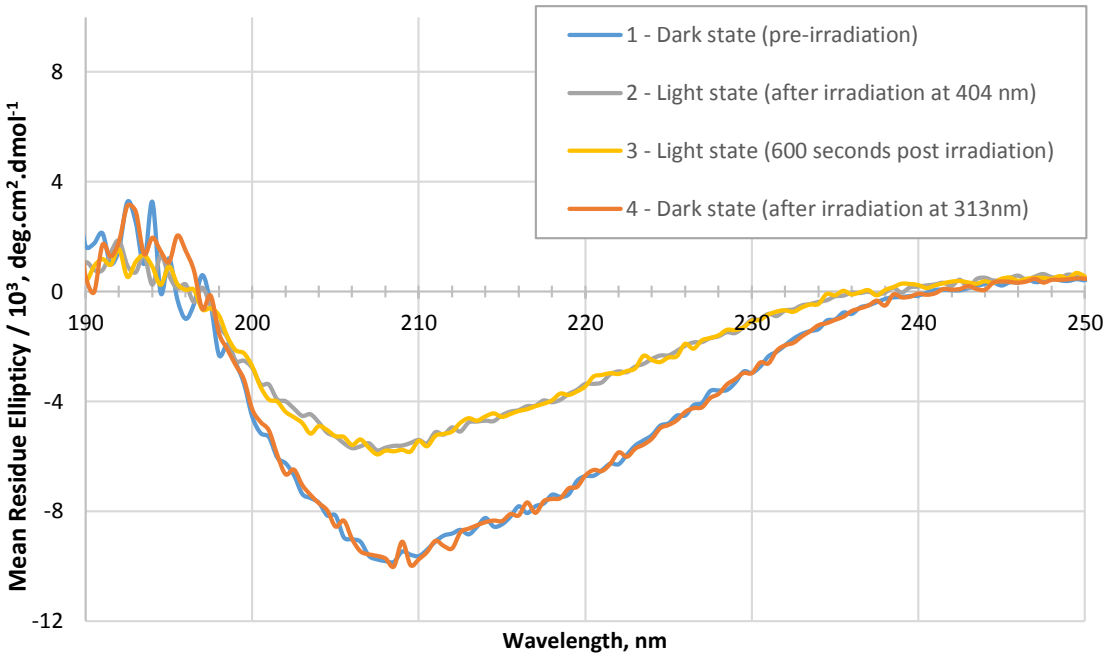


Figure 95: CD Spectrum of 5-deazaFMN within the AsLOV2(WT) domain. Upon irradiation, ellipticity in the region 240-200 nm decreased, indicating a decrease in α-helicity of the protein sample. This was found to be stable, with no change to the CD spectrum after 600 seconds. Irradiation of the sample at 313 nm allowed recovery of the dark state spectrum

These results indicate that the photo-reaction with 5-deazaFMN was successful at forming a covalent photoadduct with the AsLOV2(WT) domain, as it resulted in the release and subsequent loss of helicity for the *J*-α helix as reported for the native-FMN containing protein[225–227]. As observed using UV-visible spectroscopy, the resultant photoadduct remained stable, with no spontaneous recovery of helicity. Recovery of the dark state was again facilitated by irradiation at 313 nm.

3.4.2.3 5-deazaFMN – AsLOV(WT) discussion

Control of the photoreaction between 5-deazaFMN and AsLOV2(WT) using two wavelengths of light (404 nm and 313 nm) is due to the energy provided by photons of these wavelengths providing the energy to allow specific electronic transitions to take place. While the commonly accepted mechanism for the LOV domain photocycle suggests an intermediate triplet-state (section 1.1.3), direct transition between a ground-state singlet (S_0) and excited triplet (T_i) state is forbidden (due to opposing electronic spins), and are thus ordinarily reached via an excited singlet state (S_i). An appreciation of the photophysics of flavin was therefore undertaken.

The lowest-energy $S_0 \rightarrow S_i$ transition ($^1[\pi,\pi^*]$) calculated for riboflavin has a wavenumber of 22.5 x 10^{-3} cm^{-1}, equating to a wavelength of 444.4 nm[228]. Assuming a negligible effect of the C5'-phophate group of FMN on the photophysical properties of the flavin ring (as suggested by their near-identical UV-visible absorption spectra[88], *vide supra*), this value is remarkably close to the wavelength required to stimulate bond formation between AsLOV2(WT) and FMN (447 nm). As expected, no direct $S_0 \rightarrow T_i$ transition is reported for riboflavin within reasonable limits of this wavelength – $^3[\pi,\pi^*]$ calculated from 21.7 x 10^{-3} cm^{-1} as 460.8 nm; $^3[n,\pi^*]$ calculated from 22.8 x 10^{-3} cm^{-1} as 438.6 nm[228].

It is likely that the small difference between the observed wavelength and that calculated from the energy of the $S_0 \rightarrow S_i$ transition may be due to the local conditions within the protein binding site, suggesting that this transition is responsible for the photoreaction of the domain. If true, formation of the flavin triplet state within the protein (to allow cofactor-protein bond formation) must occur as a result of intersystem crossing (ISC), and not a direct (spin-forbidden) photo-initiated $S_0 \rightarrow T_i$ transition.

The use of EPR to identify a flavin radical formed upon irradiation[46], combined with other evidence, has allowed the deduction of the mechanism shown in figure 96 below. The theoretical calculations described above appear to compliment this, supporting the currently accepted mechanism for the AsLOV2-domain photocycle.

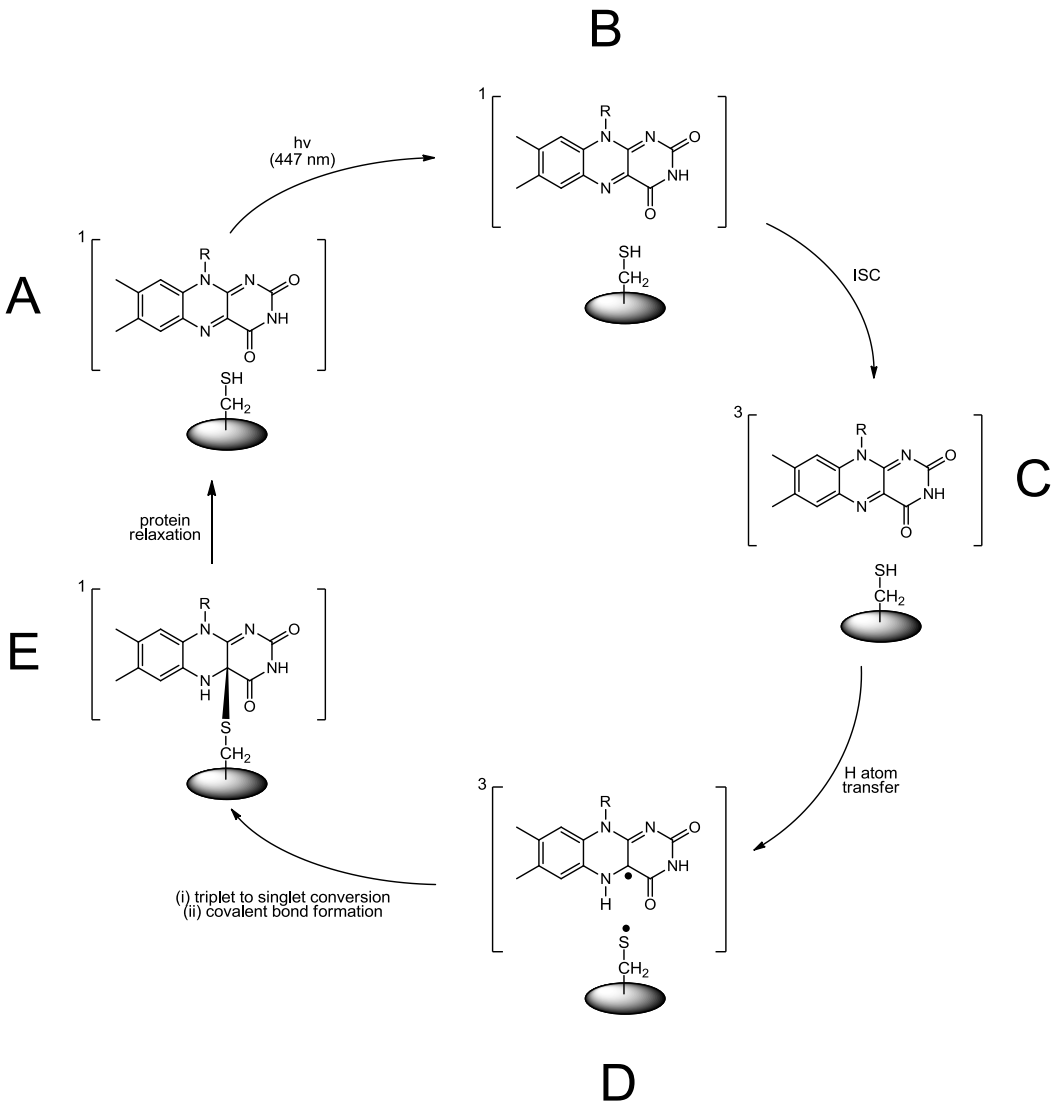


*Figure 96: Mechanism for the photocycle of the AsLOV2(WT) domain (containing native FMN), proposed by Schleicher et al.[46], from dark-state "A" to light-state "E". Irradiation of the AsLOV2 domain does not cause direct formation of the FMN triplet state; which is instead formed by an $S_i \rightarrow T_i$ transition via intersystem crossing (ISC). Although reactions are shown as unidirectional, some may be bidirectional or lead directly to dark-state **A** via alternative pathways for radiative decay*

Assuming similar behaviour of 5-deazariboflavin compared to riboflavin (and thus also the phosphorylated cofactors), the lowest-energy calculated $S_0 \rightarrow S_i$ transition ($^1[\pi,\pi^*]$) is reported at 25.0 x 10^{-3} cm^{-1}, equivalent to a wavelength of 400.0 nm[91]. As before, this value is remarkably close to the wavelength required to initiate reaction between 5-deazaFMN and the phototropin domain (404 nm), with the difference potentially accounted for by local effects within the protein. Again, no $S_0 \rightarrow T_i$ transition for 5-deazariboflavin has been calculated with a similar energy to this wavelength (between 430 nm and 360 nm), implying that a triplet state intermediate for 5-deazaFMN may only be formed via intersystem crossing rather than direct photo stimulation. This is particularly interesting, since a $S_0 \rightarrow T_i$ transition (usually spin-forbidden) may have explained the exceptional stability of the 5-deazaFMN-AsLOV2(WT) photoadduct.

As the 5-deazaFMN-AsLOV2 photoadduct only reverts to the dark state with irradiation at 313 nm, this wavelength must be implicated in providing the requisite energy to facilitate cleavage of this bond, and which may also require an intermediate triplet-state. This would provide a sufficient barrier to the photoreversion to "lock" the conformation of the light-state photoadduct, and thus may play more of a role in the stability of the intermediate than the photo-formation step.

As before, two potential transitions for 5-deazariboflavin were found at similar energies to this wavelength – a $S_0 \rightarrow S_i$ transition ($^1[\pi,\pi^*]$) with a wavenumber of 31.9 x 10^{-3} cm^{-1} (313.5 nm), or a $S_0 \rightarrow T_i$ transition ($^3[\pi,\pi^*]$) with a wavenumber of 31.1 x 10^{-3} cm^{-1} (321.5 nm)[91]. While the ($^1[\pi,\pi^*]$) transition is very close to the expected value, the presence of a $S_0 \rightarrow T_i$ transition of similar energy is highly significant. As a ground-state singlet to excited-state triplet transition is ordinarily a spin-disallowed process, the process is unlikely to occur spontaneously, and thus the potential for photoreversion would be incredibly low. This may be a key reason for the stability of the 5-deazaFMN-AsLOV2(WT) photoadduct, however further work is necessary to confirm this suggestion.

During the preparation of this thesis, similar work has been reported by Silva-Junior *et al.*[102], describing the incorporation of 5-deazaFMN into the LOV domain of YtvA,

reaching similar conclusions. While the 5-deazaFMN-AsLOV2(WT) construct has been shown to be most effectively controlled at 404 nm and 313 nm, Silva-Junior *et al.* report that photocontrol of 5-deazaFMN-YtvA was performed at 396 nm and 334 nm. A mechanism based upon various photophysical calculations performed is shown in figure 97 below.

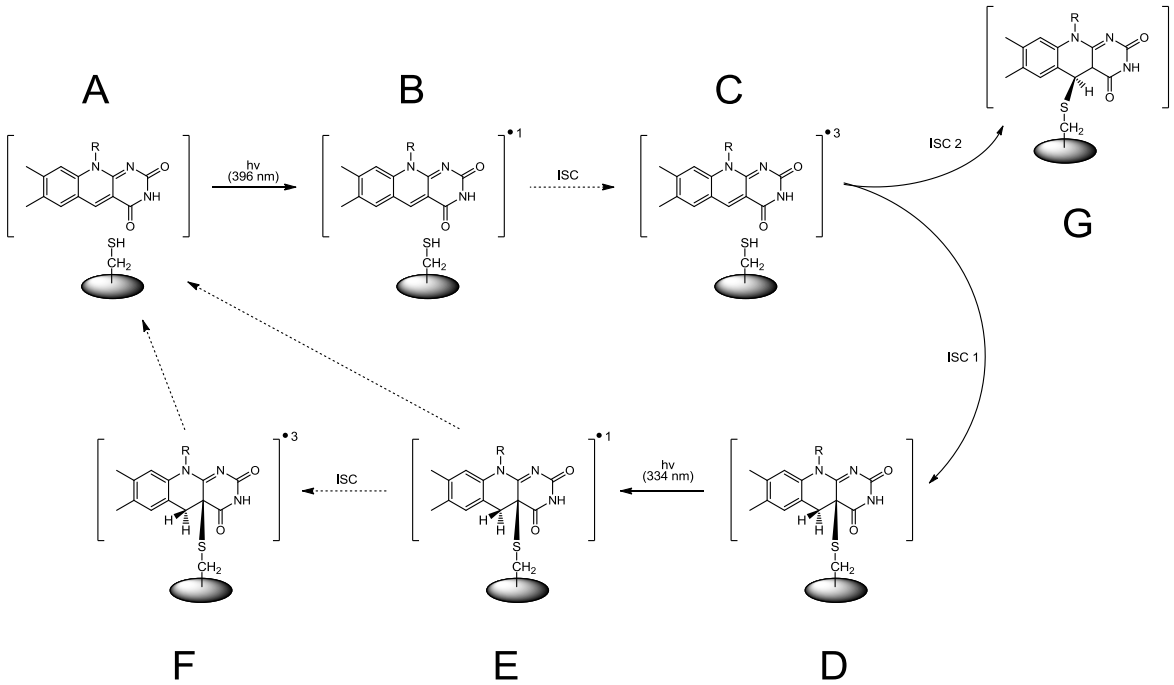


Figure 97: Proposed mechanism for 5-deazaFMN-YtvA-LOV shown as reported by Silva-Junior et al.[102] (letters added to aid discussion), based largely upon that of the native-FMN AsLOV2(WT) photocycle (figure 96), although the wavelengths reported are somewhat different to those found during this project

That rationale for the wavelengths reported by Silva-Junior *et al.* is unclear, as the $S_0 \rightarrow S_1$ transition wavelengths of 5-deazariboflavin are well characterised[91]. Furthermore, their mechanism reports two separate potential modes of photoreversion, from a singlet-state (**E**; figure 97), or from a triplet state (**F**) formed by a second ISC after irradiation of the light-state protein. As described above, two transitions are possible for 5-deazariboflavin in this region ($S_0 \rightarrow S_i$ transition ($^1[\pi,\pi^*]$) at 313.5 nm, or $S_0 \rightarrow T_i$ transition ($^3[\pi,\pi^*]$) at 321.5 nm[91]), although the wavelength reported by Silva-Junior *et al.* is closer to the wavelength required for direct formation of an excited triplet state, and thus may support this theory. However, further work is necessary to conclusively confirm this mechanism.

One further interesting alternative was also suggested within this mechanism. Intermediate **G** (figure 97) instead shows covalent bond formation between cysteine and 5-deazaFMN at position C5 of the 5-deazaflavin. As has been shown (by both NMR and crystallographic studies[43,48,54]), it is the C4a position of FMN which reacts with cysteine in a LOV domain, but in the case of 5-deazaflavin , this requires the C5 position to act both as a base (deprotonating the reactive cysteine), and as an acid (losing this proton upon reversion to the dark state). This appears somewhat unlikely, based upon an appreciation of the behaviour and strength of a C-H bond compared to an N-H bond; consequently, the mechanism proposed may not be the correct pathway for photoadduct formation. As an alternative, it is possible that position C5 of the deazaflavin may instead behave as a conjugate acceptor, and thus itself be labile to nucleophilic attack from cysteine. This is assisted by the relative positioning of the uracil moiety of the deazaflavin, which may to act as an electron sink for the reaction. Several potential conjugate partners are shown in figure 98.

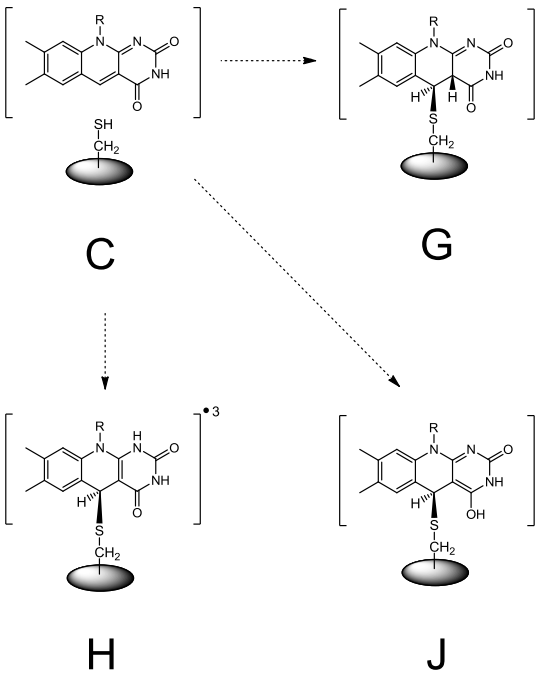


*Figure 98: Several potential conjugate partners exist for the reaction between 5-deazaFMN and LOV. It is possible that the consequential transfer of hydrogen (from cysteine) may occur at positions C4a (**G**), N1 (**H**), or C4(O) (**J**); forms **H** and **J** allow retention of the sp^2 hybridisation of position C4a*

3.4.3 5-deazaFMN reactivity with glutathione and cysteine

In order to provide further evidence to support or refute the potential conjugate addition theory, a series of experiments were performed in order to examine whether FMN or 5-deazaFMN would react spontaneously, or with photoirradiation, to form an adduct with a short cysteine-containing peptide (glutathione), or free cysteine (fully reduced) in solution. The peptide glutathione[229] was chosen as although it is small and has a simple sequence (5'-Gly-Cys-Glu-3'), it behaves as a strong thiolic nucleophile. Its role in nature is to react preferentially with conjugate acceptors, reducing the risk of inactivation of labile groups within cells[230–232].

UV-visible spectroscopy was again used as a convenient and fast method of determining whether a reaction had taken place, due to an expected photoquenching and / or band shift of the flavin absorbance upon successful bond formation.

Solutions of FMN, 5-deazaFMN and 5-deazariboflavin were each treated with glutathione or cysteine solutions (both in their fully reduced state, at concentrations of 50 mM or 150 mM, respectively), however no spontaneous reaction with any flavin was observed photometrically. Irradiation using the wavelength of light appropriate to cause covalent bond formation within the AsLOV2 domain (447.5 nm for FMN; 404 nm for 5-deazaflavins) also failed to give any change in the UV-visible spectra recorded, even with extended irradiation times. Other wavelengths were also examined, but no changes in the UV-visible spectra were observed.

Failure to form an adduct therefore suggests that a more complex reaction may be occurring than a simple conjugate addition of the reactive cysteine of AsLOV2 to 5-deazaFMN. However, these experiments do not conclusively rule this out, as other factors (such as the surrounding hydrogen-bonding network of the domain) may be contributory.

3.4.4 Photoreaction of AsLOV2(WT) with 1-deazaFMN

3.4.4.1 Spectroscopic results

As with the native cofactor FMN and 5-deazaFMN shown above, the successful incorporation of 1-deazaFMN within the AsLOV2(WT) domain was indicated by a resolution of the flavin absorbance spectrum, to reveal three resonances at 535, 503 and 570 nm, as demonstrated in figures 99(a) and 99(b) below.

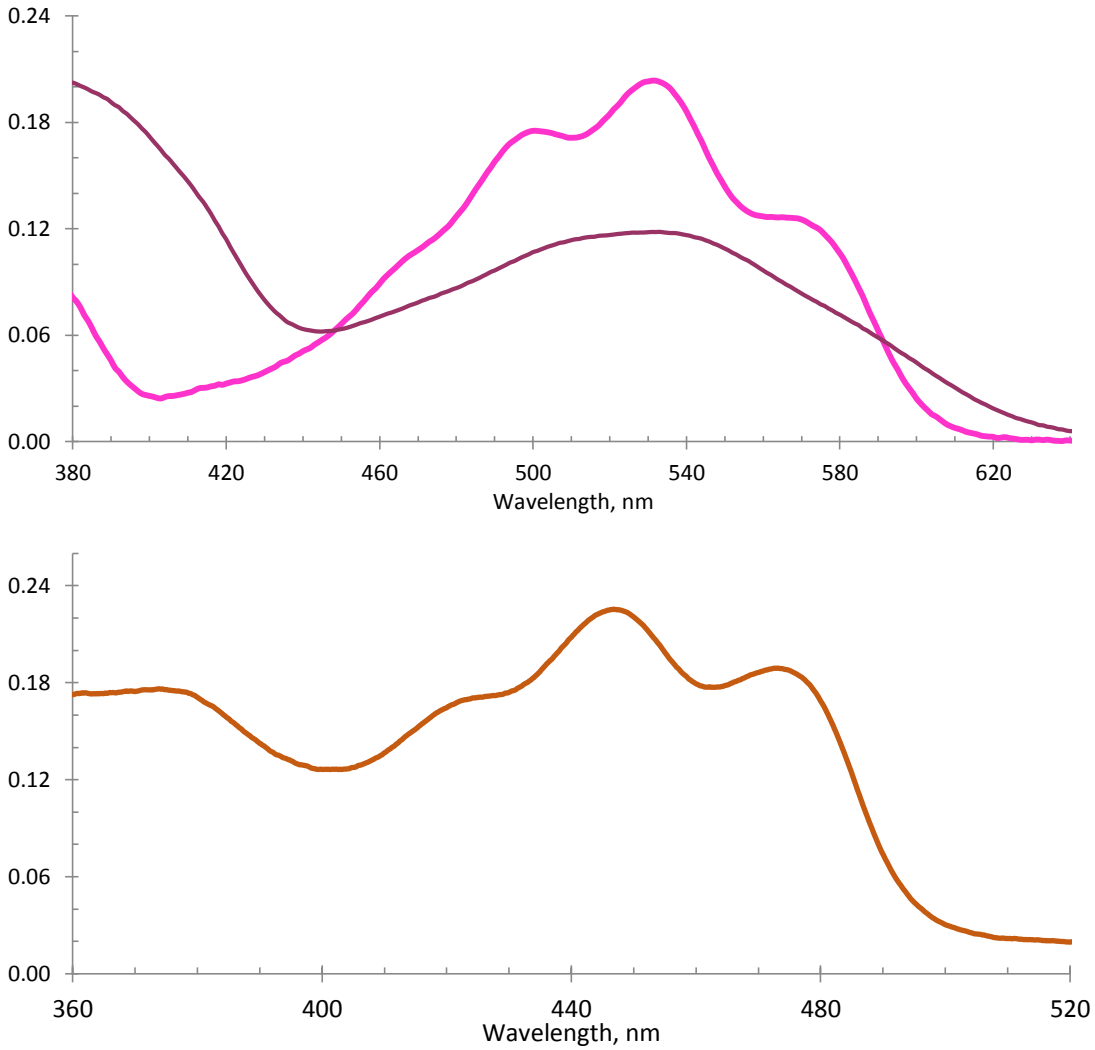


Figure 99(a) [above]: UV-visible absorbance spectrum of 1-deazaFMN free in solution (dark purple) and 1-deaza-FMN bound within the AsLOV2(WT) domain (pink)

Figure 99(b) [below]: UV-visible absorbance spectrum of native FMN within the AsLOV2(WT) domain (orange)

Interestingly, while the second-most predominant side-band of both native FMN and 5-deazaFMN–AsLOV2 has a longer wavelength than the central absorbance (see figures 90 and 92, respectively), this is reversed for 1-deazaFMN-LOV2 (figure 99(a)). No reason for this could be found, although it was believed to be unrelated to the function of the cofactor within the domain, and may also be noted in the spectra presented by Silva-Junior *et al.*[102].

However, whereas both native FMN and 5-deazaFMN have been shown to form a photoadduct with the protein under irradiation with distinct wavelengths of light, no such adduct was found when 1-deazaFMN–AsLOV2(WT) samples were irradiated using any wavelength of light between 260 nm and 750 nm. No changes to either the UV-visible spectrum or CD spectrum (which was identical to the dark-state FMN–AsLOV2(WT) spectrum shown in fig. 91) were found. Therefore, 1-deazaFMN appeared to prevent photoswitching of the AsLOV2(WT) domain; a similar result was also reported by Silva-Junior *et al.*[102] using the YtvA protein.

3.4.4.2 1-deazaFMN–AsLOV(WT) discussion

Very little discussion of the photophysics of 1-deazaflavins has been made in the literature, almost certainly due in part to the previous difficulty of their synthesis, and their limited application to flavin-catalysed processes (*vide supra*).

However, since its early syntheses, 1-deazaflavin has been reported to prevent generation of a stable triplet-state[97]. A discussion of this has been performed by Slavov *et al.*[101] and Salzmann *et al.*[233], with numerous calculations performed to predict the energies and relative positions of the singlet and triplet transition bands possible for the molecule.

Calculations by Slavov *et al.*[101] predict that as for both riboflavin and 5-deazariboflavin, both the lowest-lying triplet state (T_1; due to a $^3[\pi\rightarrow\pi^*]$ transition), and the second triplet state (T_2; $^3[n\rightarrow\pi^*]$) lie below the energy of the S_1 transition, and so each may be populated via intersystem crossing from an excited singlet state. On the other hand, Salzmann *et al.*[233] have determined that only the T_1 state of 1-deazaflavin lies

below the S_1 level, and thus the $^3[\pi\rightarrow\pi^*]$ is the only possible route for formation of a triplet state in 1-deazaflavin.

Furthermore, both authors report that the rate of ISC is too slow to permit a transition from $S_1\rightarrow T_1$ levels, with the energy of the transition therefore lost via internal conversion or radiative decay[233]. The slow rate of intersystem crossing was determined by Slavov *et al.*[101] by measurement of the fluorescence quantum yields of 1-deazariboflavin, which was found to be incredibly low (0.002, compared to 0.27 for riboflavin[233]), demonstrating that for 1-deazaflavins this is not the predominant route for energy loss from an excited state.

When this evidence is applied to 1-deazaFMN within AsLOV2, it may be inferred that as a maximum of two available triplet states of 1-deazaflavin ($^3[\pi\rightarrow\pi^*]$) lie lower than the energy of the S_1 state, and that the rate of intersystem crossing between the singlet excited state and triplet excited state is too slow to permit the transition, no reaction should be expected. The absence of an observed 1-deazaFMN triplet state is particularly useful in this case, as it provides further evidence to support the triplet-state mechanism of the AsLOV2 photoreaction. If the alternative nucleophilic mechanism (postulated by Crosson and Moffat[30]; see section 1.1.3) was true, it would be expected that 1-deazaFMN would also form an adduct with the AsLOV2 domain, which was not observed under any condition attempted.

Finally, the results obtained have also been confirmed by the work of Silva-Junior *et al.*[102], within the LOV domain of YtvA. They report similar findings for 1-deazaFMN within the YtvA LOV domain as found in this study using the LOV2 domain of *A.sativa*, which further supports the radical-pair mechanism of photoadduct formation (*vide supra*).

3.4.5 AsLOV2(C450X) mutants, incorporating 5-deazaFMN and 1-deazaFMN

3.4.5.1 Spectroscopic results

Examination of each mutant protein formed as described in section 3.3.4 with each of the three cofactors (native FMN, 5-deazaFMN and 1-deazaFMN) was performed, with

the UV-visible spectral results summarised in table 9 below, for brevity. A representative spectrum of native-FMN-AsLOV2(C450M) is shown in figure 100, below. While the dark-state spectra of the mutant proteins were similar to those of the wild-type protein (with similar vibronic side-bands observed), under no condition examined was it possible to observe any photoswitching behaviour from any mutant protein containing either of the three cofactors examined.

Protein Sample	**Cofactor**	**Primary Absorbance Wavelength, nm**	**Vibronic Side-Band Wavelength, nm**	
			Most Intense	**Least Intense**
Wild-Type	native FMN	447	473	422
	5-deazaFMN	404	428	384
	1-deazaFMN	535	503	570
C450M	native FMN	447	472	421
	5-deazaFMN	404	428	383
	1-deazaFMN	535	502	570
C450S	native FMN	447	473	422
	5-deazaFMN	404	428	384
	1-deazaFMN	535	503	571
C450D	native FMN	448	474	422
	5-deazaFMN	404	428	384
	1-deazaFMN	535	503	570
C450G	native FMN	447	473	422
	5-deazaFMN	404	427	383
	1-deazaFMN	535	502	571
C450A(T418I)	native FMN	447	473	422
	5-deazaFMN	404	428	384
	1-deazaFMN	533	502	570

Table 9: Key spectral features of all AsLOV2 samples examined. The primary absorbance (related to the original absorbance of the cofactor) is reported, along with the wavelengths observed for the two characteristic vibronic side-bands. All samples (except wild-type AsLOV2 containing native FMN or 5-deazaFMN) were inactive to photoirradiation

While all mutant proteins were found to be inactive to photoirradiation for durations previously used to cause formation of the photoadduct, it was found that when continually irradiated for exceptionally extended periods (up to 1 hour), the

absorbance observed steadily decreased, giving a “pseudo-bleached” state (via stable intermediate stages), similar to the effect seen with 5-deazaFMN in AsLOV2(WT). However, this was found to be due to destruction of the flavin by continual irradiation at high intensity; simultaneous observation of significant protein precipitation was also observed, with no recovery of the dark-state spectrum possible. This decrease in absorbance was not due to formation of a stable photo-adduct.

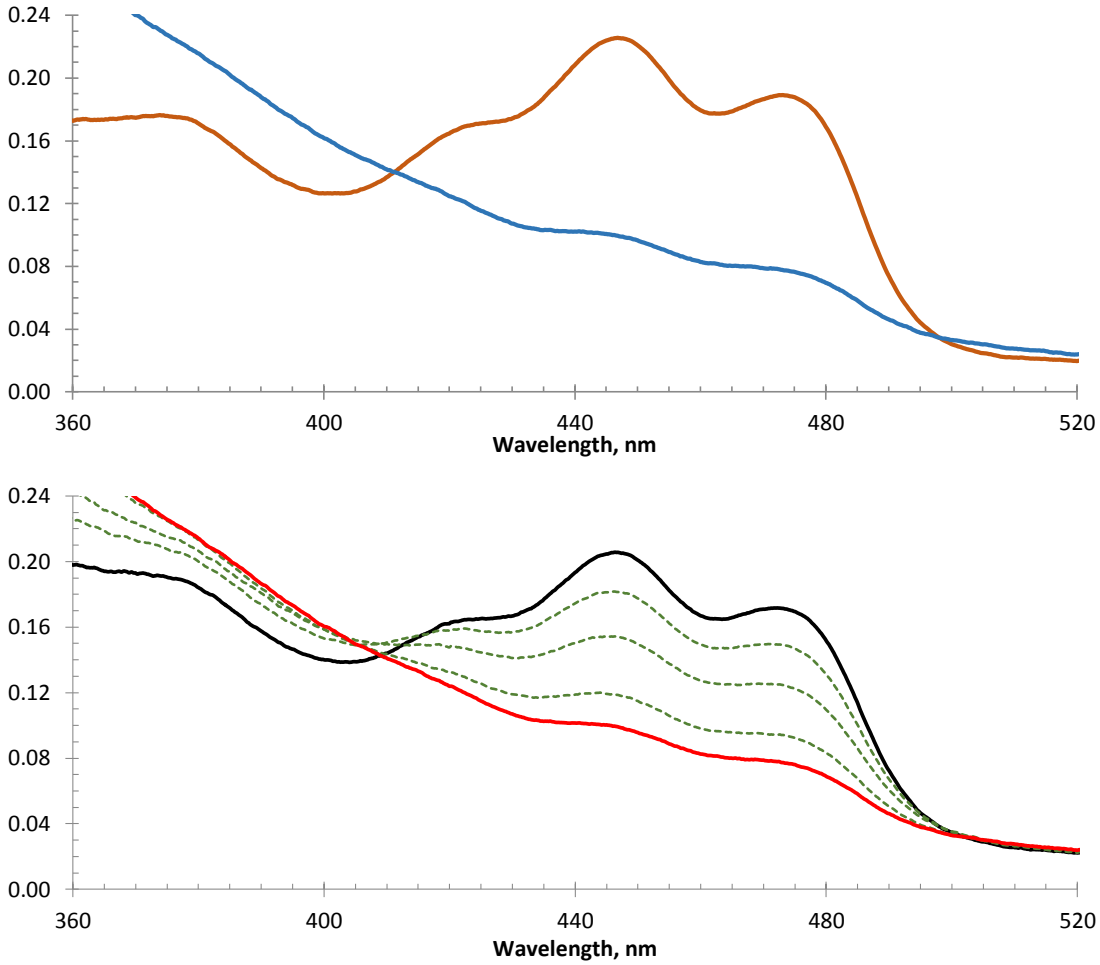


Figure 100: UV-visible absorption spectra of native FMN–AsLOV2(WT)] in the dark state (orange), and after irradiation at 447 nm (blue) [above], and native FMN–AsLOV2(C450M) [below]. The mutant dark state (black) initially remained after irradiation at 447 nm; vastly extended irradiation times led to some decrease in absorbance via stable intermediate stages (green) to a pseudo-photobleached state (red), which did not revert to pre-irradiation levels; examination of the sample indicated that loss of absorbance was due to destruction of the flavin and consequent precipitation of the apoprotein

3.4.5.2 AsLOV2(C450X) mutant discussion

The failure of each mutant to give a photoadduct with any cofactor demonstrates the importance of the cysteine group for formation of the photoadduct. Interestingly, the behaviour of the sulphur atom of cysteine appears key, since methionine and serine mutants both failed to allow photoadduct formation. While methionine's methylated sulphur group may prevent formation of the adduct for steric reasons, serine is almost identical in structure to cysteine, with an oxygen atom substituted for sulphur.

The ability of the thiolic amino acid to undergo the photoreaction while the hydroxylic does not may be due to the difference in spin-orbit coupling between sulphur and oxygen. This is due to the electronic arrangement of the heavier sulphur atom, with the effect of reducing the unfavourability in the unpairing of electron spins, as the additional energy levels required for electrons become degenerate[234,235]. Furthermore, the increased mass of sulphur reduces the orbital angular momentum necessary for singlet to triplet state conversion[236] (via intersystem crossing), allowing sulphur to more easily adopt an excited triplet state. This is an important consideration for the photocycle of LOV domains, as a radical pair mechanism requires each reactant to be of the same excitation state. As photoirradiation of flavin stimulates triplet state formation, the reduced energetic barrier to the adoption of an excited triplet state by cysteine permits bond formation between the flavin and protein.

3.4.6 Conclusions – Spectroscopic investigation of AsLOV2 domain

The results shown in section 3.4.2 demonstrate that incorporation of the artificial cofactor 5-deazaFMN within the AsLOV2(WT) domain allows full control of the phototropin using two wavelengths of light. Stability of both the light and dark states was found, and remained even in the presence of high concentrations of imidazole (previously reported to catalyse the photoreversion of the photoadduct[58]). It may be possible to further develop this for practical uses in molecular biology, as LOV domains are rapidly becoming a common functional tool to examine and control both *in vivo*[32,211] and *in vitro*[237] processes. As the rapid reversion of the light-state of AsLOV2

containing the native cofactor necessitates constant illumination or significant mutation of the protein to prolong the functional effect[39,56,57], this may be potentially damaging for the cell or system under analysis. Therefore, the incorporation of 5-deazaFMN may allow a controlled regulation of the protein without constant irradiation.

Furthermore, the artificial cofactor within the AsLOV2 domain may facilitate the preparation of protein crystals in the light-state suitable for examination using X-ray crystallography, potentially revealing important structural information. While a high quality dark-state crystal structure has been obtained[48], and a low resolution dark and light structures have been partially elucidated by NMR[43], the collection of high resolution structural detail for the light-state has not yet been possible due to the requirement for constant irradiation during growth. The use of the stable light-state afforded by incorporation of 5-deazaFMN would allow collection of this data, which may give fresh insight for the behaviour of the domain.

On the other hand, 1-deazaFMN prevented formation of a photoadduct with AsLOV2(WT) upon irradiation at any wavelength examined (260-750 nm), which appears to be largely due to the inability of this analogue to adopt triplet state upon irradiation. However, this supports the radical pair mechanism of flavin-LOV photoreaction, as 1-deazaFMN would be expected to form the photoadduct via the ionic mechanism proposed by Crosson and Moffat[30].

Examination of each mutant protein (section 3.3.4) with each cofactor showed no photoadduct formation under any condition, providing some evidence for the unique role that sulphur plays within the domain. These results may be built upon using the inactive cofactor 1-dezaFMN within the AsLOV2(WT) domain; as 1-deazaFMN does not readily form a triplet-state under irradiation, it would be possible to isolate any triplet-state formation of cysteine using techniques such as Electron Paramagnetic Resonance (EPR), and allow deduction of the point at which triplet state formation occurs (prior to or during irradiation), which may further the understanding of this part of the LOV domain photocycle mechanism.

Chapter 4

Conclusions and

Future Work

Chapter 4 - Conclusions and Future Work

Overall, the conditions set in the motivation and aims for this project (sections 1.0.2 and 1.3) have been met. The syntheses of 5-deazariboflavin, 1-deazariboflavin and 8-demethyl-5-deazariboflavin have been reported and optimised (particularly in the case of 1-deazariboflavin), along with foundations laid for the synthesis of several alternative analogues of riboflavin. After successful enzymatic phosphorylation, 5-deazaFMN and 1-deazaFMN were incorporated into the AsLOV2 domain, and used to examine the photocycle of this protein. The results of these studies have enabled the confirmation of the triplet-state mechanism of photoadduct formation, and demonstrated the potential for development of this as a tool for use in molecular biology and optogenetics using a synthetic cofactor.

4.1 Conclusions - Organic Synthesis

4.1.1 5-deazariboflavin synthesis

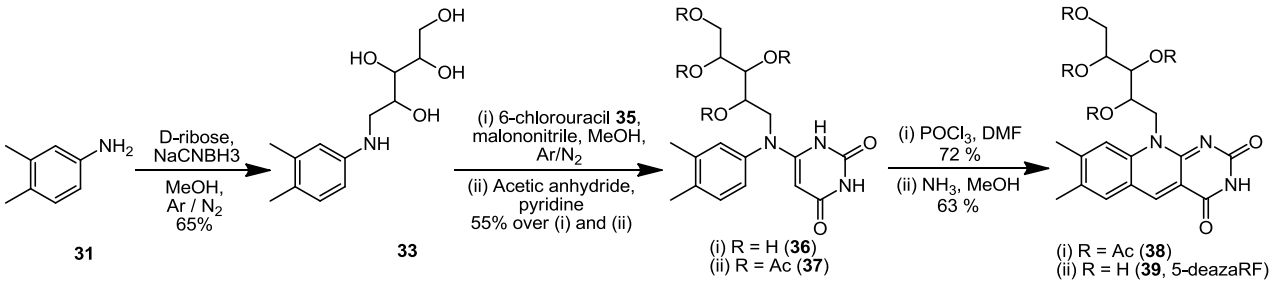


*Figure 101: Overall synthetic route to 5–deazariboflavin **39***

The formation of 5-deazariboflavin **39** is among the most popular of all flavin analogue syntheses described in the literature, with decades of synthetic work leading to the optimal route described by Carlson and Kiessling[5]. While subsequent syntheses[100] based upon this method have reported inflated yields without any change or enhancement to the original reaction conditions, the exhaustive study of conditions described in this thesis confirm the yields of Carlson and Kiessling, forming the target compound in 16.2 % yield overall.

The most significant result obtained during the synthesis of 5-deazariboflavin **39** was obtained during the protection of the ribityl chain of intermediate **36** by acetylation,

where the previously unreported acetylation of the uracil moiety of bicyclic compound **37** was occasionally observed. This was found to have no effect on subsequent reactions, and the N-acetyl groups were removed using the same deprotection conditions used to remove the hydroxyl protecting groups. It does however indicate that future syntheses of riboflavin analogues may be able to utilise a fully-protected uracil moiety, which is supported by some evidence in the literature[108,111], and may enable more efficient coupling between 3,4-dimethyl-N(ribityl)-aniline **33** and 6-chlorouracil **35**.

4.1.2 1-deazariboflavin synthesis

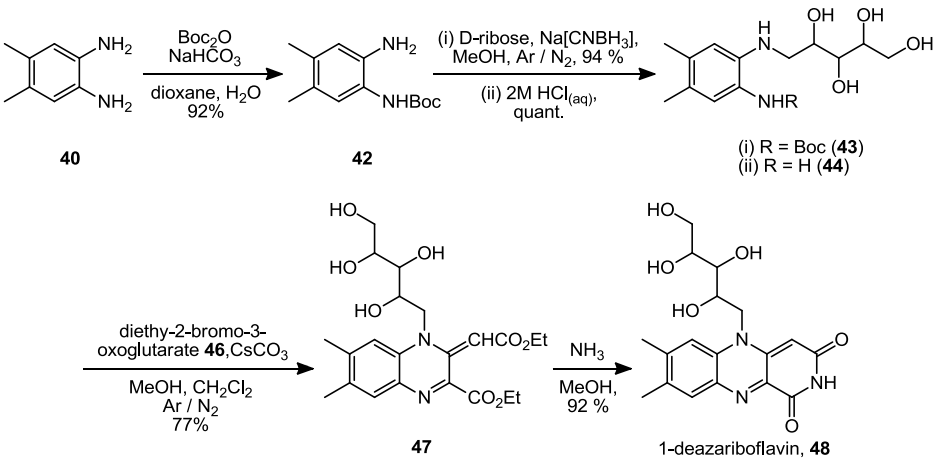


*Figure 102: Overall scheme for the synthesis of 1-deazariboflavin **48***

The synthesis of 1-deazariboflavin **48** was performed in 61.3 % yield, which is a vast improvement over previous methods (11.0 % [5]; 21 % [100]). This exceptional improvement was due to the enhancements made during two key reactions: firstly, the coupling of the deprotected intermediate **44** and brominated oxoglutarate **46**, and secondly, closure of the final ring of intermediate **47** at the N(3) position. For both reactions, adjustments to the reaction solvent led to the improvements, by allowing more ready solubilisation of reactants in the first case, or using a greater concentration of ammonia in the second.

It may be of particular interest to repeat this method using several alternative nucleophiles for the final step of the reaction. This would allow the construction of a

suite of novel 1-deaza-3(N-functionalised) analogues, which may be of interest both as structural analogues of flavin, and by using electron donating or withdrawing groups to tune the electrochemical properties of the flavin.

4.1.3 8-demethyl-5-deazariboflavin synthesis

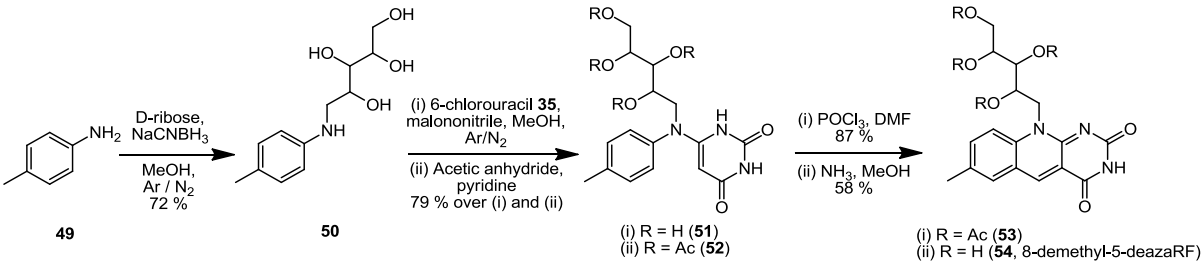


*Figure 103: Scheme for the synthesis of 8-demethyl-5–deazariboflavin **54***

The novel synthesis of 8-demethyl-5-deazariboflavin **54** has been described, based upon the optimal method for the synthesis of 5-deazariboflavin **39**, using 4-methylaniline **49** in place of 3,4-dimethylaniline **33**. Use of 4-methylaniline **49** led to a significant improvement in overall yield (28.7 %, versus 16.2% for 5-deazariboflavin), and assisted the discussion of the efficiency of 5-deazariboflavin synthesis by allowing consideration of the relative activating or deactivating positions of the aromatic ring.

Furthermore, this also showed that the synthetic route described may tolerate alternative starting materials, and it would therefore be particularly interesting to apply a number of alternatives (such as 3-methylaniline or 3,5-dimethylaniline) to form a range of other riboflavin analogues.

4.2 Conclusions - Molecular Biology

4.2.1 *C. ammoniagenes* Ca-RfK-FADS

The expression and purification of the bifunctional Ca-RfK-FADS protein was performed successfully, in high yield. However, the use of this enzyme to phosphorylate riboflavin analogues was less successful.

While the enzyme worked efficiently, the predominant product of the enzymatic reaction was the FAD cofactor form, rather than the desired FMN form. This required

lengthy purification using preparative-scale HPLC, followed by chemical hydrolysis of the phosphoanhydride bond, and re-purification. Therefore, an alternative method to isolate the FMN form of the deaza analogues was sought.

4.2.2 *S. pombe* RfK

The expression of the Sp-RfK gene, and its purification, were again performed successfully, in very high yield. Phosphorylation of 5-deazariboflavin and 1-deazariboflavin was swift and efficient; 1 equivalent of protein was able to turn over 130 equivalents of the flavin cofactor within 4 hours.

The purity of the FMN analogues formed by this enzyme were assessed by HPLC, and found to typically be over 99 % of the total flavin content of the reaction mixture. Purification was ordinarily performed via preparative-scale HPLC, although due to the high purity of the product this may not have been strictly necessary.

4.2.3 *A. sativa* LOV2 preparation

Expression of the AsLOV2(WT) protein was generally unreliable. Several factors may have been responsible, although two likely reasons were that (i) the plasmid vector containing the gene used a different promoter sequence to that which the expression cell strain was optimised for; or (ii) the AsLOV2(WT) gene had not been optimised for expression in *E. coli*.

To address these factors, expression was attempted in *E. coli* strains BL21(DE3), BL21(DE3)-STAR, BL21(DE3) [pLysS], M15 [pRep4] and JM101. Each strain generally performed poorly, although strains BL21(DE3) and M15 [pRep4] were found to be the most reliable strains for expression of the wild-type protein. Furthermore, an analysis of the AsLOV2(WT) gene was performed using the web-based OPTIMIZER tool[209,210], which showed poor codon optimisation for *E. coli*. Of the 477 nucleotides of the gene, 97 mutations would have been required to fully-optimise the gene for *E. coli*, suggesting that this may play a large role in the poor expression of the gene.

Several mutants of the AsLOV2 gene were formed, replacing the cysteine residue at position 450 (of the full-length phototropin) with either alanine, serine, methionine, glycine or aspartic acid (abbreviated to C450A, S, M, G or D). In the case of the alanine mutant, an undesired mutation occurred during the PCR process, leading to an unintended double-mutant (AsLOV2(T418I, C450A)). However, as the behaviour of the C450A mutant has been previously well characterised, this was not considered to be detrimental.

Expression of each of these mutants was performed in *E. coli* strain JM101 (which gave optimal yields), although this was also unreliable due to poor regulatory control within the cell often resulting in early induction of the plasmid.

Conditions were found to allow incorporation of the synthetic FMN analogues within each of the AsLOV2 protein samples. This was performed effectively, resulting in protein samples containing either native FMN, 5-deazaFMN or 1-deazaFMN. These each adopted the colour of the cofactor they contained, which was reflected by the UV-visible spectra. The spectra obtained were found to be resolved (compared to the spectrum of the free cofactor), to reveal two vibronic side-bands either side of the central absorbance, which was indicative of protein binding. Crucially, the apoprotein form was found to be unstable in solution, and rapidly precipitated; thus, this did not influence subsequent spectroscopic examination.

4.2.4 Examination of the AsLOV2 domain

Samples of the AsLOV2(WT) domain containing each cofactor were examined using UV-visible and CD spectroscopy. The results for native-FMN samples showed formation of a light-state upon irradiation at 447 nm (the wavelength of the FMN chromophore), with photo-bleaching of the sample observed by UV-visible spectroscopy, and a partial loss (30.9 %) of ellipticity observed in the CD spectrum commensurate with unwinding of the *J*-α helix of the protein (which accounted for 42.3 % of the helical residues). The irradiated “light” state of the sample was unstable, and decayed to pre-irradiation levels (with recovery of pre-irradiation spectra) with an approximate half-life of 30 seconds.

When 5-deazaFMN was incorporated within the AsLOV2(WT) domain, irradiation at the wavelength of the flavin chromophore (404 nm) led to photo-bleaching of the sample as with the native-FMN protein. However, the irradiated state was found to be fully stable for periods over 72 hours post-irradiation, and in the presence of a catalyst for the reversion. Furthermore, the loss of ellipticity observed in the CD spectrum was found to be 42.2 %, which closely matches the percentage of residues which make up the *J*-α helix of the protein (42.3 %). Reversion of the 5-deazaFMN photoadduct was found to be possible using irradiation at 313 nm, which allowed recovery of the dark-state spectra.

Incorporation of 1-deazaFMN fully inactivated the AsLOV2(WT) domain to all wavelengths of light. This analogue, which is unable to form a triplet-state, therefore acts as an inhibitor to the photoreaction, and thus supports the currently-accepted radical pair mechanism of photoadduct formation.

Each mutant protein, containing either cofactor, was found to be inactive to photoirradiation at all wavelengths. While this was expected for mutants lacking a reactive residue (C450A, T418I and C450G), it was more surprising that methionine, serine and aspartic acid also failed to react with the bound cofactor. The replacement of cysteine with serine indicates that the sulphur atom is key to the photoreaction, as sulphur possesses several energy levels of similar energy and has a relatively low barrier for intersystem crossing (necessary to form a triplet state, to bond with triplet-state flavin) via spin-orbit coupling.

Suggested future work in this area would be the coupling of 5-deazaFMN-containing AsLOV2(WT) to a functional domain, in order to examine whether the observed photoreaction allows control of a physical process. This may result in the development of a key tool for optogenetics and molecular biology, allowing manipulation of complex systems using two wavelengths of light.

Chapter 5

Materials and Methods

5.1 Chemical suppliers, solvent and reagent purification

Several methods for the purification of the reagents and solvents used for this research were obtained from the widely-used reference text "Purification of Laboratory Chemicals" by W. F. Armarego and C. Chai[122], which have been supplemented by experimentally-derived methods where possible.

5.1.1 Chemical suppliers

Solid compounds and solvents were of high purity (98 % or above), purchased from Acros Organics, Sigma-Aldrich, Fischer Scientific, Fluka, Alfa Aesar and Novabiochem; generally used without further purification (exceptions detailed below or in relevant experimental section). Isotopically labelled compounds were obtained as follows: ^{13}C-carbonyl-dimethyformamide was supplied by Sigma-Aldrich; ^{13}C-C_6 glucose and $^{15}NH_4Cl$ were supplied by Cambridge Isotope Laboratories; deuterated solvents for NMR were supplied by Sigma Aldrich ($CDCl_3$, acetone-D_6, MeOD) and Fischer (DMSO-D_6, MeOD, D_2O).

5.1.2 Desiccants

Molecular sieves were prepared either by (i) heating in an oven at 300 °C for a minimum of 48 hours, followed by storage under high vacuum (< 10^{-2} Torr) for a minimum of 4 hours[238,239]; alternatively, by (ii) successive rounds of heating in a domestic microwave oven, followed again by storage under high vacuum for a minimum of 6 hours. Dryness was confirmed by addition of water to a small number of dried sieves; exothermic reaction indicated successful drying. Other desiccants used include silica, as used for Flash Chromatography, and alumina.

5.1.3 Solvent Purification

Methanol was distilled over CaH_2[122] and further dried using 3 Å molecular sieves[121]. Other methods are available, but are reportedly less efficient[121], and may potentially lead to formation of undesirable methoxides[240,241]. Pyridine was dried by storage over 4 Å molecular sieves in the dark for 24 hours. Triethylamine was distilled over CaH_2 at atmospheric pressure. DMSO was dried using 4 Å molecular sieves or standing over

alumina, followed by reflux over CaH_2 and distillation[122,242]. Anhydrous THF, MeCN, CH_2Cl_2, toluene and hexane were collected from an MBraun MB SPS-800 solvent purification system. THF was further dried by storage over 3 Å molecular sieves.

Sodium methoxide and ethoxide were freshly prepared using sodium which had been stored under mineral oil. The protective oil was removed by washing in a bath of anhydrous hexane and the oxidised surface material was scraped away, before cautious portionwise addition of the cleaned metal to the pre-dried stirring solvent.

5.1.4 Recrystallisation

3,4-dimethylaniline was recrystallised from ligroin[122], or from warmed 60 % ($^{v}/_{v}$) aqueous EtOH. Special care was necessary due to the low melting point (51 °C) of the compound. 6-chlorouracil was initially recrystallised from a 40 % solution of EtOH in water. However, due to the instability of this compound in aqueous conditions, purification by recrystallisation is not recommended. 2,6-dichloropyridine was purified by recrystallisation from absolute EtOH[122]. 2,4,6-trichloropyridine was recrystallised from 75 % or 50% EtOH in water. Extreme care are was necessary due to the low melting point of the compound (33-36 °C).

4-dimethylaminopyridine (DMAP) was recrystallised from toluene[122] or EtOAc. Imidazole was recrystallised from CH_2Cl_2 or petroleum ether (40-60). *p*-Tosyl chloride was recrystallised from an equivolume mixture of petroleum ether (40-60) and toluene. 18-crown-6 was recrystallised from acetonitrile; several rounds of recrystallisation were necessary to ensure complete removal of water.

5.2 Buffers and Solutions for Molecular Biology

5.2.1 Media for bacterial growth

LB (Leuria-Bertani)

10 gL^{-1}	tryptone
5 gL^{-1}	yeast extract
5 gL^{-1}	NaCl
H_2O, to final volume	

TB (Terrific Broth)

This is made up in two parts. The masses given are sufficient for 1 L (total volume) of growth medium once combined. Each component is sterilised separately, and mixed immediately prior to use.

Component A (per litre final volume)

12 g	tryptone
24 g	yeast extract
4 mL	glycerol
900 mL	H_2O

Component B (per litre final volume)

2.31 g	KH_2PO_4
12.54 g	K_2HPO_4
100 mL	H_2O

M9 minimal media

Starred solutions were sterile-filtered using MillexGP syringe filters (0.22 μ metre). Compounds in squared-brackets were added after sterilisation – by autoclave or sterile filter as appropriate.

M9 minimal medium (complete) (1 L)

100 mL	M9 Salts (10 x)
1 mL	1 M $MgSO_4$*
0.1 mL	1 M $CaCl_2$
1 mL	trace element solution
[10 mL	20 % $^{w}/_{v}$ glucose (aqueous)]

M9 salts (10 x, 1 L)*

150 g	Na_2HPO_4
30 g	KH_2PO_4
5 g	NaCl
[10 g	NH_4Cl]

In the case of isotopically enriched mixtures, the ^{13}C-glucose and / or $^{15}NH_4Cl$ were omitted from the initial mixtures, and added aseptically immediately before introduction of the starter-culture.

Trace elements (1000 x, 100 mL)*

1.6 g	$MnCl_2.4H_2O$
1.5 g	$CuCl_2.2H2O$
2.7 g	CoCl2.6H2O
3.0 g	$FeCl_2$
4.0g	Na-citrate.$2H_2O$
10 mg	$Zn(OAc)_2.2H_2O$
5 mg	disodium edetate (EDTA.Na_2)

Enriched Minimal Media mixtures

M9 + supplemental vitamins

M9 minimal media made up as above, plus

2.5 mL L^{-1}	vitamin mixture *

Vitamin Mixture (250 x, 1 L)*

50 mg	pyridoxal 5'-phosphate (monohydrate)
10 mg	thiamine hydrochloride
20 mg	4-aminobenzoic acid
20 mg	D-pantothenic acid (hemicalcium salt)
10 mg	biotin
10 mg	folic acid
5 mg	cyanocobalamin
20 mg	*riboflavin*
50 mg	niacin
270 mg	$ZnCl_2$

When expression of labelled riboflavin was desired, the supplemental riboflavin (underlined) was omitted from the vitamin mixture.

M9Y Media *

M9 minimal media made up as above, plus

5 g L^{-1} yeast extract
10 mg L^{-1} $FeCl_2.6H_2O$
Mixture was filtered prior to use.

M9 + Amino acid(s)

M9 minimal media made up as above, plus

50 mg L^{-1} desired amino acid, dissolved in minimal volume H_2O*

Up to three amino acids were supplemented in this manner.

5.2.2 Buffers for Chemically Competent Cells

Super-competent Cells

Buffer Tfb I

100 mM RbCl
50 mM *$MnCl_2$*
30 mM KOAc
10 mM $CaCl_2$
15 % ($^w/_v$) glycerol
H_2O, to final volume
adjusted to pH 5.8 using dilute (0.2 % $^v/_v$) acetic acid

N.B. manganese chloride was added *after pH adjustment* to prevent precipitation.

Solution was sterile-filtered and stored at RT, but chilled to 4 °C before use.

Buffer Tfb II

10 mM MOPS
75 mM $CaCl_2$
10 mM RbCl
15 % ($^w/_v$) glycerol
H_2O, to final volume
adjusted to pH 6.5 using KOH *with caution*

Solution was sterile-filtered and stored at RT, but chilled to 4 °C before use.

Simple ($CaCl_2$) Method for Competent Cells

Buffer Ca for calcium-only competent cells

10 mM PIPES
60 mM $CaCl_2$
15 % ($^w/_v$) glycerin
H_2O, to final volume
adjusted to pH 7.0

Solution was sterile-filtered and stored at RT, but chilled to 4 °C before use.

N.B. cells produced using this method were most competent after 12-24 hours stored in this buffer, on ice

5.2.3 Buffers for Protein and DNA Separation

SDS-PAGE Resolving Gel Buffer (1L)

1.5 M Tris base
100 µL 10% $^w/_v$ aqueous solution of SDS
adjusted to pH 8.0

H_2O, to final volume

SDS-PAGE Stacking Gel Buffer (1L)

0.5 M Tris base
100 µL 10% $^w/_v$ aqueous solution of SDS
adjusted to pH 6.8

H_2O, to final volume

SDS-PAGE Running Buffer

250 mM Tris base
1.5 M glycine
35 mM SDS
adjusted to pH 8.3

H_2O, to final volume

SDS-PAGE Sample Buffer

60 mM	Tris HCl
5 % ($^{w}/_{v}$)	SDS
10 % ($^{w}/_{v}$)	sucrose
30 % ($^{v}/_{v}$)	glycerol
3 % ($^{v}/_{v}$)	β-mercaptoethanol
0.02 % ($^{w}/_{v}$)	Bromophenol blue

Ammonium Persulfate Solution (1 mL, 10% $^{w}/_{v}$)

100 mg	APS
1 mL	H_2O

This solution was not stable for storage of greater than 4 days at -80 °C.

SDS-Resolving Gel (12 %)

4.0 mL	Bis-acrylamide, 30% solution
3.4 mL	H_2O
2.5 mL	Resolving gel buffer
75 µL	10 % APS solution
5 µL	TEMED

SDS-Resolving Gel (15 %)

5.0 mL	Bis-acrylamide, 30% solution
2.4 mL	H_2O
2.5 mL	Resolving gel buffer
75 µL	10 % APS solution
5 µL	TEMED

SDS- Stacking Gel (5 %)

1.7 mL	Bis-acrylamide, 30% solution
5.7 mL	H_2O
2.5 mL	Stacking gel buffer
75 µL	10 % APS solution
10 µL	TEMED

Coomassie (AKA Brilliant) Blue Staining Solution for SDS gels

0.25 % ($^{w}/_{v}$) Coomassie Blue R-250
45% ($^{v}/_{v}$) MeOH
9 % ($^{v}/_{v}$) AcOH
H_2O, to final volume

Destaining Solution (overnight)

15 % ($^{v}/_{v}$) MeOH
7.5 % ($^{v}/_{v}$) AcOH
H_2O, to final volume

Destaining Solution (heating method)

50 % ($^{v}/_{v}$) EtOH
10 % ($^{v}/_{v}$) AcOH
H_2O, to final volume

TAE (50x concentration)

2 M Tris base
1 M glacial acetic acid
100 mM EDTA
adjusted to pH 8.2

H_2O, to final volume

Agarose Gel (1 % $^{w}/_{v}$, for plasmids up to 8 kBp)

0.5 g Agarose
50 mL TAE buffer (diluted)

Sample Dye Buffer (100 mL)

30 % ($^{v}/_{v}$) glycerol
100 µL 0.6 % ($^{w}/_{v}$) aqueous Bromophenol blue solution
H_2O, to final volume

Ethidium Bromide Solution (100 mL)

100 µg Ethidium Bromide
H_2O, to final volume

5.2.4 Buffers for Alkaline Lysis of Bacterial Cells, and Precipitation of DNA

Solution I

50 mM glucose
25 mM Tris base
10 mM EDTA
adjusted to pH 8.0

Solution II

0.2 M NaOH
1% (w/v) SDS

Solution III

3 M KOAc
adjusted to pH 5.2

TE Buffer (10 x)

100 mM Tris HCl
10 mM Na_2EDTA
H_2O to 100 mL
adjusted to pH 8.2

Ethanol Precipitation of intracellular DNA

7.5 M NH_4OAc
adjusted to pH 7.0
EtOH, as required

5.2.5 Buffers for Bacterial Lysis and Purification using Spin Columns

These buffers are used following the protocol of the QIAGEN 'miniprep' plasmid purification kit.

Buffer P1 (suspension buffer)

50 mM Tris HCl
10 mM EDTA
100 µg / mL RNase A
adjusted to pH 8.0

stored at 2-5°C

Buffer P2 (lysis buffer)

0.2 M NaOH
1% (w/v) SDS

Buffer N3 (neutralisation buffer)

4 M guanidinium HCl
0.5 M KOAc
adjusted to pH 4.2

Buffer PB (wash buffer I)

5 M guanidinium HCl
20 mM Tris HCl
38% (v/v) EtOH
adjusted to pH 6.6

Buffer PE (wash buffer II)

20 mM NaCl
2 mM Tris HCl
80% (v/v) EtOH
adjusted to pH 7.5

Buffer EB (elution buffer)

10 mM Tris HCl
adjusted to pH 8.5

5.2.6 Restriction Endonuclease Buffers

NEBuffer 1 (10x)

100 mM Tris HCl
100 mM $MgCl_2$
10 mM DTT
pH adjusted to pH 7.0

NEBuffer 2 (10x)

100 mM Tris HCl
500 mM NaCl
100 mM $MgCl_2$
10 mM DTT
pH adjusted to pH 7.9

NEBuffer 3 (10x)

500 mM Tris HCl
1 M NaCl
100 mM $MgCl_2$
10 mM DTT
pH adjusted to pH 7.9

NEBuffer 4 (10x)

200 mM Tris-OAc
500 mM K-OAc
100 mM $Mg\text{-}(OAc)_2$
10 mM DTT
pH adjusted to pH 7.9

BSA Storage Buffer (for storage of 1x dilution from BSA (100x) solution)

20 mM KPO_4
50 mM NaCl
100 µM EDTA
5% (v/v) glycerol
adjusted to pH 7.0

5.2.7 Buffers for Purification of *C. ammoniagenes* RfK-FADS

Cell Lysis Buffer

Cell Lysis was performed using the binding / washing buffer RfK 1.

Buffer RfK 1 (binding / washing)

50 mM KH_2PO_4
5 mM EDTA
5 mM Na_2SO_4
adjusted to pH 8.0

Buffer RfK 2 (elution)

Buffer RfK1, plus

45 mM maltose (to a final concentration of 50 mM
0.02% (w/v) NaN_3
adjusted to pH 8.0

Buffer RfK 3 (dialysis)

50 mM Tris Base
50 mL NaCl
0.02% (w/v) NaN_3
adjusted to pH 7.5

Buffer RfK 4 (for enzymatic conversion of riboflavin to FMN / FAD)

50 mM Tris Base
50 mM NaCl
30 mM $MgCl_2$
100 µM Purified Ca-RfK-FADS (in buffer RfK 3).
[n] mM riboflavin / riboflavin analogue – up to 1 mg / mL
[n x 5] mM Adenosine triphosphate
adjusted to pH 7.5

The minimum volume of buffer RfK 4 was used to dissolve riboflavin (or analogue).

5.2.8 Buffers for IMAC purification of 6xHIS tagged *S. pombe* Rf kinase

Buffer $SpRfK_{LYS}$ (cell lysis buffer)

50 mM Tris HCl
100 mM NaCl
adjusted to pH 8.0
H_2O, to final volume

Buffer SpRfK 1 (binding buffer)

20 mM Tris HCl
500 mM NaCl
10 mM imidazole
adjusted to pH 8.0
H_2O, to final volume

Buffer SpRfK 2 (washing buffer)

20 mM Tris HCl
500 mM NaCl
50 mM Imidazole
adjusted to pH 8.0
H_2O, to final volume

Buffer SpRfK 3 (elution buffer)

20 mM Tris HCl
500 mM NaCl
400 mM imidazole
adjusted to pH 8.0
H_2O, to final volume

Buffer SpRfK 4 (storage buffer)

20 mM Tris HCl
100 mM NaCl
adjusted to pH 8.0
H_2O, to final volume

Buffer SpRfK 5 (FMN formation buffer)

Buffer SpRfK 1, plus

10 mM	$MgSO_4$
5 mM	ATP
1 mM	riboflavin / riboflavin analogue
100 µM	SpRfK protein
H_2O to V_{tot}	Minimum volume used to dissolve riboflavin / riboflavin analogue

5.2.9 Buffers for IMAC Purification of Hisactophilin-tagged *A. sativa* LOV2

Buffer LOV_R (resuspension buffer)

10 mM	NaH_2PO_4
0.9% (v/v)	NaCl

adjusted to pH 7.5

Buffer LOV_{LYS} (cell lysis buffer)

50 mM	NaH_2PO_4
300 mM	NaCl
10 mM	Imidazole

adjusted to pH 8.0

Buffer LOV_A (binding & washing buffer)

50 mM	NaH_2PO_4
300 mM	NaCl
10 mM	Imidazole
0.02% (w/v)	NaN_3

adjusted to pH 8.0

Buffer LOV_B (denaturing buffer)

Buffer LOV_A, plus

4.5 M	Guanidine thiocyanate

adjusted to pH 8.0

Buffer LOV_C (elution buffer)

Buffer LOV_A, plus

490 mM imidazole (total solution concentration of 1M)
adjusted to pH 8.0

Buffer LOV_D 20x (dialysis buffer, to remove imidazole and perform photometric analysis)

100 mM NaH_2PO_4
200m M NaCl
0.4% (w/v) NaN_3 (added after spectroscopic analysis)
adjusted to pH 8.0

Buffer LOV_E (cleavage buffer for removal of Hisactophilin)

50 mM NaH_2PO_4 OR Tris HCl (dependant on current sample buffer)
200 mM NaCl
2.5 mM $CaCl_2$
[n] mM HISACT-AsLOV2 fusion protein – up to 0.5 mg / mL
[n x 5] units thrombin
adjusted to pH 7.5

5.3 TLC Systems

All TLCs were performed on Merck TLC Silica Gel (60 Å) plates with aluminium foil backing. TLCs were observed using a Spectroline ENF-240C/FE Dual Wavelength lamp providing UV light (254nm & 365 nm). TLC plates were also developed for visual examination using appropriate chemical staining: ninhydrin (for compounds with amine functionality), dinitro-phenylhydrazine (DNPH, for aldehydes and ketones), cerium molybdate (CAM, also known as or Hanessian's Stain, for general visualisation), or iodine (for other functionalities, in particular aromatics), as appropriate. RF values of each product determined is given in the relevant experimental section.

5.4 Instruments Used

5.4.1 Instruments Used - Synthetic Chemistry

Balance (Analytical)	KERN ALS 220-4N (accurate to 0.05 mg)
Centrifuge (up to 4 000 rpm)	Thermo Electron Corporation IEC CL31R Multispeed Centrifuge (up to 4000 rpm)
Centrifuge (up to 11 000 rpm)	Hettich Zentrifugen Rotina 38 R (up to 11 000 rpm)
Fluorimeter	Perkin Elmer LS 55 Luminescence Spectrometer
Freeze Dryer	Thermo Electron Corporation ModulyoD coupled to ThemoSavant RV8 belt-driven high vacuum (< 0.1 mBar) oil pump
HPLC	Dionex Ultimate 3000 system (complete)
HPLC	Dionex P680 HPLC Pump with Dionex UVD170U Variable Wavelength Detector
IR Spectrometer	Jasco FT/IR-660 Plus
Mass Spectrometer	Waters LCT Premier XE TOF
Melting Point	Gallekamp Variable Heater
NMR Spectrometer	Brüker AVANCE DPX 250 MHz
NMR Spectrometer	Brüker Fourier 300 MHz
NMR Spectrometer	Brüker AVANCE DPX 400 MHz (Shielded)
NMR Spectrometer	Brüker AVANCE DPX 400 MHz (ULTRA-Shielded)
NMR Spectrometer	Brüker AVANCE DPX 500 MHz (ULTRA-Shielded)
pH Meter	HANNA Instruments pH 209 meter
Polarimeter	Schmidt + Haensch UniPol L-Series polarimeter
Press (for KBr disk formation)	GRASEBY SPECAC (15 tonne)

Rotary Evaporator (water cooled)	Heidolph Laborota 4000 Efficient (complete system), coupled to a VAKUUBRAND GmbH & Co. MZ-2C Pump, and controlled by a Heidolph feedback pressure sensor
Rotary Evaporator (dry ice / acetone cooled)	Bücki Rotavapor RE 120, with Köttermann waterbath, coupled to VAKUUBRAND GmbH & Co. Membrane and Vacuum pumps
Solvent Drying Machine	MBraun MB SPS-800
Ultrasonic Bath	Grant XB3 Ultrasonic Bath
Stirrer/ Heating Plate	IKA-WERKE RCT basic heating / stirring block, with removable IKA-WERKE ETS-D4 fuzzy temperature feedback probe
TLC Visualisation Lamp	Spectroline ENF-240C/FE Dual Wavelength (254nm & 365 nm)
UV Spectrometer	Jasco UV-660 Spectrophotometer
Vacuum Pump	Oerlikon TRIVAC mechanical high vacuum (< 10^-2 Torr) oil pump
Water Purification System	Elga Purelab Ultra systems (Ω < 15.0)
Water Purification System	Elga Purelab Ultra-pure (Ω < 5.0)

5.4.2 Instruments Used - Molecular Biology

Agarose Gel Electrophoresis	Bio-Rad mini-PROTEAN® Tetra system
Autoclave	Astell Scientific ASA-260
Autoclave	Astell Scientific AMA-260BT with SPL-400 USB
Balances	Kern ALJ 120-4 (to 0.1 mg)
Balances	Mettler Toledo AB204-5 (to 0.1 g)
Cell Counter (UV-visible)	SHIMADZU BioSpec-mini DNA/RNA/Protein Analyzer
Centrifuge (Large Volume)	Sorvall RC-5 Plus
Centrifuge (Large Volume)	Sorvall RC-6 Plus
Rotors	Sorvall SLA-3000
Rotors	Sorvall SS-34
Rotors	Piramoon Technologies FIBERLite F21-8x50y
Rotors	Thermo Scientific FIBERLite F10-4x1000 LEX
Centrifuge (Bench-top)	JENCONS-PLS Spectrafuge 24D (ambient temperature)
Centrifuge (Bench-top)	Eppendorf Centrifuge 5424 (with cooling)
Circular Dichroism	Applied Photophysics Chirascan instrument, coupled to Applied Physics Circular Dichroism Photomultiplier and Scanning Emission Monochromator
Fermenter	LAMBDA Minifor Laboratory Fermenter
Fluorimeter	*see section* 5.4.1
FPLC	Amersham Biosciences ÄKTA FPLC System, fitted with UPC-900, Inv-907, P-920, M-925 and Frac-950 (autosampler) units
Gel Visualisation	Syngene Bio Imaging "Geneflash" lightbox, Syngene UP-895MD Video Graphic Printer
Heating Block	JENCONS-PLS SB1 mono digital block heater; 2 block model

IMAC Column Housing	GE Healthcare XK 16/20 Column, fitted with Quicklock AK 16 adaptors
Incubators	New Brunswick Scientific Innova 43
Incubators	New Brunswick Scientific Innova 44 (variable temperature)
Ion Exchange Chromatography	Dionex ICS-3000 system
Nanodrop Spectrometer	Thermo Scientific NANODROP 1000 Spectrophotometer
Peristaltic Pump	Watson-Marlow 101U/P Pump
pH Meter	HANNA Instruments pH 209 meter
Plate Oven	Heraeus Instruments Variable Temperature
SDS-PAGE Gel Casting	Hoefer Scientific Instruments SE 250 MIGHTY Small II
See-Saw Rocker	Labortechnik KS501 digital
Sonnicator	Vibra-cell SONICS
Thermocycler (PCR)	JENCONS-PLS TECHNE Touchgene Gradient
Thermocycler (PCR)	Biometra TGradient 96
Thermomixer Block	Eppendorf Thermomixer "comfort"
UV-Visible Spectroscopy	*see section* 5.4.1
Vacuum Pump (filtration & degassing)	Fischer-Biotech Roto-Vac 3
Vortex Mixer	Vortex Genie
Warm Water Bath	Grant SUB28 Unstirred Water Bath
Water Purification System	*see section* 5.4.1

5.5 Techniques and Methods – Chemical Synthesis

5.5.1 General experimental techniques

Chemical reactions were performed in glassware fitted with ground "quick-fit" joints; those in close proximity to reaction mixtures were not sealed using grease as is common practice but instead using a Teflon or PTFE gasket to prevent contamination of the reaction mixture.

Reaction vessels were heated using solid aluminium heating blocks turned specifically to fit the size of flask used, or within a bath of silicon oil for complex-shaped glassware. Heating was monitored using an IKA-WERKE ETS-D4 *fuzzy* temperature probe mounted within a port of the heater block / bath, coupled electronically to the IKA-WERKE RCT *basic* heater / magnetic stirrer apparatus.

In the case of reactions performed to the exclusion of air and / or moisture, flame-dried glassware was used, which had been heated under high vacuum for a minimum of 4 hours. Reagents were powdered and pre-dried, and added immediately after flame-drying, under a stream of N_2 followed by further cycles of high-vacuum and flushing with Ar and N_2 to ensure removal of air and moisture. Dry reactions were performed under Ar, with pressure release afforded by a N_2 bubbler. Solvents (section 5.1.3) used for air-free reactions were added to the sealed flask using syringes or canulae through rubber septa.

"High Vacuum" refers to ultra-low vacuum pressure ($< 10^{-2}$ Torr), provided by an Oerlikon TRIVAC mechanical oil-pump. This was connected via a custom-built Schlenk-type manifold (including a N_2-cooled trap) to the reaction vessels using double-oblique connections, with the alternative position supplying N_2 gas and protected by a pressure-release bubbler.

5.5.2 Purification methods (chemical synthesis)

5.5.2.1 Thin Layer Chromatography (TLC)

Analysis by TLC is discussed in detail in section 5.3 above.

5.5.2.2 Flash Chromatography (FC)

Purification by FC was performed using a gel of silica formed using an appropriate solvent mixture, and technical grade silica gel (pore size 60 Å, 234-400 mesh size, 40-63 µm particle size). This was retained between a layers of sand and glass sinter within a glass column, and pressurised using a hand pump. Fractions of an appropriate volume were collected before analysis by TLC to indicate fraction composition.

5.5.2.3 High Performance Liquid Chromatography (HPLC)

HPLC Analysis was performed using a complete Dionex UltiMate 3000 system, including a variable wavelength UV-visible detector, or using a Dionex P680 HPLC pump system coupled to a Dionex UVD170U variable wavelength detector. Solvents for HPLC were of HPLC or analytical grade, supplied by Fischer or (in the case of buffers) freshly prepared using water obtained from an Elga Purelab Ultra-pure system. All buffers were thoroughly degassed using a vacuum pump prior to use. Conditions and gradients are described in the appropriate experimental chapters (sections 6.5 and 7.1.5). Gradients are shown graphically in appendix 1. Columns used include Phenomenex Gemini® 10 µm RP C18 110 Å 250 x 10 mm (part no. 00G-4436-N0), Phenomenex Luna® 10 µm RP C18(2) 100 Å 250 x 10 mm (part no. 00G-4252-N0) and YMC-Actus Triart Prep C18-S 250 x 20 mm (part TA12S05-2520WX). Fractions from HPLC were lyophilised using a Thermo Electron Corporation MOLULYOD freeze dryer, with high vacuum (< 0.1 mbar) provided by a ThermoSavant RV8 belt-driven oil pump.

5.5.3 Methods for spectral analysis

Solid samples for NMR spectroscopy (10 mg for ^{1}H; 20-50 mg for ^{13}C) were prepared by dissolution in a suitable deuterated solvent, and added to an NMR tube to a depth of 45 mm. Liquid samples were added neat to a depth of 10-15 mm, topped up to a final depth of 45 mm using a suitable deuterated solvent. This was necessary to provide sufficient deuteration to allow locking of the spectrometer to the deuterium

signal, as an internal standard. Spectra were measured using an NMR spectrometer (section 5.4.1) of an appropriate field strength.

Solid samples for Mass Spectrometry (MS) were submitted to the internal MS service at Cardiff University (5-20 mg). Samples were dissolved in solvent immediately prior to examination except samples which were only soluble in DMSO, which were analysed as a solid; liquid samples were submitted neat. Photosensitive samples were stored in the dark until immediately prior to examination. Samples were analysed using a Waters LCT Premier XE Time-of-Flight Spectrometer. Solid samples for IR analysis (1-5 mg) were milled together with KBr (spectroscopic grade). The resultant fine powder was formed into a disk using a GRASEBY SPECAC press (15 tonne) before analysis using a Jasco FT/IR-660 Plus spectrophotometer. Melting point (MP) measurements were collected using Gallenkamp Variable Heater apparatus fitted with an internal temperature probe and confirmed using a mercury thermometer, and are uncorrected. Chemical samples for UV-visible spectroscopy were prepared using a suitable solvent of HPLC or spectroscopic grade, measured in a capped Hellma semimicro CXA-145-307L far-UV quartz cuvette. A further description of UV spectroscopy is given in section 5.6.8.

5.6 Techniques and Methods – Molecular Biology

5.6.1 General experimental techniques & suppliers

The pH of buffers and solutions were measured using a HANNA Instruments 'pH 209' device accurate to 2 decimal places, and were adjusted using the appropriate conjugate acid or base. All pH values given are at 20 °C unless otherwise stated. Liquid cultures were incubated with shaking in New Brunswick Scientific Innova 43 or Innova 44 shakers, with variable temperature control between 10 °C and 40 °C. Temperatures of 37 °C were typically used for cell growth, and between 16-24 °C for protein expression. Enzymes obtained from commercial sources (thrombin, lysozyme) were purchased from Sigma-Aldrich. *Pfu* and *Taq* DNA polymerase and restriction endonucleases were obtained from New England Biolabs. RNase A was purchased

from Quiagen. Amylose resin, Ni-NTA and Chelating Metal Sepharose (fast flow) were obtained from GE Healthcare. Columns for FPLC were obtained from the same source.

Sequencing of DNA samples was organised by the internal DNA sequencing service at Cardiff University's School of Biosciences, operated by Dr Steven Turner. The sequencing operation was sub-contracted to Eurofins, and performed using ABI Genetic Analysers 3739xl machines (96 capillary), using the ddNTP chain termination method. Read-lengths of 900 – 1000 b.p. (after initial non-readable data) were typically obtained.

5.6.2 Bacteriological culturing and protein expression

5.6.2.1 Culture preservation

Bacterial cultures were stored on LB-agar or TB-agar plates (supplemented with appropriate antibiotic(s), sealed using parafilm and stored at 4 °C in the dark. At periods not longer than 4 weeks, new cultures were prepared from a single colony on fresh medium, with periodic sequencing of the plasmid from the daughter cells to confirm fidelity.

Storage for extended periods was performed by preparation of a 20 mL starter culture (as described in section 5.6.2.2) containing an appropriate antibiotic, incubated with shaking at 37 °C overnight. The resultant mixture was centrifuged, the cells washed with 0.9 % saline, and resuspended using a mixture of 0.5 mL LB media, 0.25 mL sterilised H_2O and 0.25 mL glycerol. The resuspended mixture was flash frozen in 2 mL Eppendorf tubes using liquid N_2, followed by storage at -80 °C.

5.6.2.2 Microbial culturing

An overnight 'starter culture' was made of the bacterium containing the desired construct by selecting a single colony from its storage plate, or by collecting a small number of cells using a red-hot sterile culture loop from a flash-frozen liquid sample prepared as described above. This was used to inoculate a small volume of LB media (100 mL was sufficient to allow subsequent large-scale expression in 3 L of growth media), which was treated with an appropriate antibiotic and incubated with shaking

at 37 °C overnight. The resultant mixture was centrifuged, cells were washed with 0.9 % ($^{w}/_{v}$) saline and resuspended in a small volume of an appropriate growth media.

Media for bacterial growth was made up as described in section 5.2.1. Prepared growth media was divided into baffled 2 L Erlenmeyer flasks (500 mL per flask) capped with air-permeable foam bungs and foil, and autoclaved. Flasks were sealed immediately upon sterilisation, and cooled media was stored at 4 °C in the dark until required. Immediately prior to use, cooled media was supplemented with appropriate antibiotic(s) under aseptic conditions and warmed to incubation temperature. To this, a portion of the starter-culture prepared as above was added, and the flasks were incubated at 37 °C until an OD_{600} between 0.6-0.8 (LB media), 1.2-1.6 (TB media), or 0.5-0.7 (M9 minimal media) was reached. At this point, isopropyl-1-thio-β-D-galactopyranoside (IPTG, 1-5 mM) was added, along with supplemental antibiotic(s) if required. The mixture was incubated as determined by expression tests (*vide infra*) and described for each construct in chapter 4; generally at 37 °C for 4 hours, or 20 °C overnight. Cells were harvested by centrifugation (30 minutes, 4500 RPM, 4 °C,), resuspended and washed with 0.9 % saline, recentrifuged, and stored at -20 °C until required.

5.6.2.3 Expression efficiency testing

The efficiency of bacterial expression was examined during the microbial culturing phase by collection of samples (2-5 mL) at timed intervals (generally 30 or 60 minutes) after supplementation of the media with IPTG (in varying concentrations). Cells within the sample was harvested as before, and the samples lysed using the alkaline lysis method (section 5.6.3.1). Whole-cell lysate was analysed using SDS-PAGE (discussed in section 5.6.4.1) to indicate levels of protein expression.

5.6.2.4 Protein expression methodology

The gene(s) of interest were located within a plasmid expression vector. Several vectors were used during this project (section 5.7.2) Those genes within a pMal-C2 vector were under the control of a *tac* promoter, pET14b and pLysS vectors were under the control of a *T7 promoter* and *lac operator*, while pNCO113 and pRep4

vectors were similarly under control of *T5 promoter* containing a *lac operator*. Expression of genes contained within the expression vectors (pMal-C2, pEt-14b, pNCO113) was initiated by addition of IPTG (1-5 mM) to the bacterial growth medium, while the control vectors (pLysS and pRep4) were used to suppress basal levels of gene expression in specific strains, and expression of their genes was thus inhibited by addition of IPTG.

5.6.3 Methods for cell lysis

Buffers and solutions used in this section are given in sections 5.2.4 and 5.2.5 above.

5.6.3.1 Alkaline lysis method

This method was used for general-purpose periodical analysis of bacterial samples, for example during the expression test described in section 5.6.2.3. This method was also occasionally used to isolate intracellular DNA, as an alternative to the QIAprep kit (section 5.6.3.3). This method uses solutions I, II and III to lyse the cells and neutralise the alkaline mixture, before analysis using SDS-PAGE.

A portion of cells were centrifuged (4000 RPM, 30 minutes, 4 °C). Solution I (150 µL) was used to resuspend the cells. Solution II (150 µL) was added, and the mixture gently mixed by pipetting or inversion. Solution III (150 µL) was added, which occasionally caused the precipitation of the potassium salt of SDS. In this case, addition of 1-2 drops of $CHCl_3$ was added to dissolve this. After thorough mixing, the mixture was centrifuged (15000 RPM, 5 minutes, 4°C), and the supernatant collected by decanting. If analysis of the protein content of the cell lysate was required, the supernatant liquid was analysed by SDS-PAGE at this stage.

To collect DNA, EtOH (1 mL) was added, and the mixture recentrifuged (15000 RPM, 15 minutes, 4 °C). The cell pellet was retained, and dried at 37 °C for 30 minutes. Once dry, buffer TE (100 µL) and DNase-free RNase A (2 µL, 10 mg mL^{-1}) was added to the solution. The mixture was incubated at 37 °C for 30 minutes, and the purified DNA was isolated by precipitation using the method described in section 5.6.3.2 below.

5.6.3.2 Purification of DNA by precipitation using EtOH

Precipitation of plasmid DNA using EtOH was performed in order to provide purified (salt and RNA-free) DNA for further uses. To a solution (50 µL) of DNA obtained as above, or collected from a spin-column, NH4OAc (7.5 M, 50 µL) was added, along with 2.5 volumes of absolute EtOH. The mixture was gently mixed, and allowed to stand at RT for 10-15 minutes. After centrifugation (13000 RPM, 15 minutes, RT), the supernatant was discarded, and the pellet resuspended in a solution of EtOH (80 % ($^{v}/_{v}$) aqueous, 250 µL), before further centrifugation (as before). The supernatant was again discarded, and the pellet of purified DNA dried at 37 °C in an oven before storage at -20 °C. Resuspension was generally performed using 50 µL of water, although lower volumes gave relative increases in DNA concentration.

5.6.3.3 Lysis and purification using QIAprep 'mini-prep' plasmid isolation kit

This method provided a fast and convenient method to collect both whole-cell lysates and to isolate intracellular DNA (which is its main purpose). It was performed using a mixture of buffers (P1, P2 and N3) to lyse the cells and neutralise the lysate, followed by purification of the intracellular DNA using the supplied anion-exchange (spin) columns. Retention of the DNA allowed purification by washing, before a reduction in the ionic strength of the washing solution enabled elution of the purified DNA. Purification was often supplemented by precipitation using EtOH (*vide supra*).

To an overnight sample of cells which had been harvested by centrifugation, the cell pellet was resuspended in buffer P1 (250 µL) on ice. To this, buffer P2 (250 µL) was added, with gentle inversion of the tube to facilitate mixing. This was allowed to stand for 5 minutes, before addition of buffer N3 (350 µL), with immediate mixing. The mixture was centrifuged (13000 RPM, 10 minutes, 4 °C), and the supernatant liquid (cell lysate) collected. For purification of intracellular DNA, the supernatant was applied to a QIAprep spin column, and centrifuged (13000 RPM, 1 minute). The flow-through was discarded, and buffer PB (500 µL) added before centrifugation (as before). Buffer PE (500 µL) was added, and centrifugation (as before) performed in duplicate, in order to fully remove all of the ethanolic buffer. Finally, DNA was eluted

from the column using sterilised H_2O, followed by standing at RT for 5 minutes and centrifugation as before. Purified DNA was isolated from the flow-through.

5.6.3.4 Sonication lysis of cells from large scale microbial culture

Cell pellets obtained from large-scale microbial culture (section 5.6.2.2) were lysed using sonication. Rupture of the cell walls was necessary to release proteins over-expressed by the inserted plasmid, which were expressed within the cytoplasm of the cell. This method was selected due to the fast, reliable disruption of large quantities of cells without addition of reagents to the mixture, with the main risk being from heating induced by the high-powered sonic probe, or excessive duration leading to degradation of the recombinant protein. All lysis using this method was performed while cooling the sample tube in a bath of crushed ice.

Cell pellets were thoroughly resuspended in the relevant lysis buffer, before subjection of the mixture to sonication for an intermittent cycle to minimise inductive heating. SDS-PAGE was used to identify the optimal sonication pattern. PMSF was added pre-sonication to the lysis mixtures, except for those containing AsLOV2. Once lysed, the mixture was centrifuged (11000 RPM, 1 hour, 4 °C or 15000 RPM, 45 minutes, 4 °C), and the supernatant decanted. This was filtered using a MillexGP syringe filter (0.22 µm) in order to remove any further suspended particles, before purification as necessary.

5.6.4 Electrophoresis of protein and DNA samples

All buffers and solutions used in this section are described above in section 5.2.3.

5.6.4.1 Sodium Dodecyl Sulphate Polyacrylamide Electrophoresis (SDS-PAGE)

SDS-PAGE was performed according to the discontinuous procedure of Laemmli[243], in order to analyse the protein content of bacterial lysates and solutions. SDS is an anionic surfactant, capable of denaturing and solubilizing a protein sample. SDS also confers a charge to the protein proportional to its size, stabilising the protein in a rod-like form by electrostatic repulsion. This negative charge allows electrophoretic

separation of different sized proteins through a gel of polymerised acrylamide, with motility of a given protein inversely proportional to its size.

The gel was prepared using the discontinuous buffer system, using a resolving gel (where protein analysis occurs) topped by a low-strength stacking gel (confining separate protein samples). The resolving gel was prepared first; after mixing of the components, initiation of polymerisation was performed by a small volume of ammonium persulfate solution. The initiated solution was cast between two glass plates, with a recess (3 cm) at the top filled with isopropanol to give a defined level surface between layers.

After completion of polymerisation, isopropanol was removed and the top of the gel rinsed with water before being thoroughly dried. After preparation of the stacking gel, initiation was again performed by APS, and the initiated solution added above the resolving gel in the unfilled recess between plates. Once added, a 10 or 15 well comb was inserted to leave defined areas in which to add the protein sample. Once polymerisation was complete, SDS-PAGE electrophoresis was performed immediately. The casting plate was fitted within the electrophoresis system, and submerged in SDS running buffer. The gel comb was removed, and checks for leaks were made. Protein samples were mixed in equivolume with SDS-sample buffer (section 5.2.3). The mixture was heated at 95 °C for 10 minutes, to ensure protein denaturation by the SDS. Prepared samples were loaded into the gel, alongside a standard marker containing a mixture of protein molecular weights (SM0431, section 5.7.4.1) to allow determination of relative protein masses.

Electrophoresis was performed by passing a current of 50 mA (180 V) across the gel for approximately 1 hour, causing resolution of the protein samples. Larger proteins were more strongly retarded by the polyacrylamide gel, and thus did not move as far as smaller proteins. Upon completion, the gel was removed from the apparatus and separated from its glass plates, before staining with the SDS-PAGE staining solution (section 5.2.3). This was generally performed at room temperature for 1 hour. The gel

was destained with the destaining solution overnight, to leave visible bands of protein for characterisation (section 5.6.4.3).

5.6.4.2 Agarose gel electrophoresis

For analytical and preparative scale purification and separation of DNA fragments, agarose gel electrophoresis was used. Due to the innate negative charge of DNA, it can be separated electrophoretically through a gel of agarose, with motility inversely proportional to the DNA fragment size. Using this method, it was possible to separate both small fragments of DNA, and large plasmid vectors containing several genes.

Due to the secondary structural effects of plasmids (such as supercoiling), circular DNA was nicked using an appropriate restriction enzyme. Typically, gels of 1 % ($^{w}/_{v}$) agarose were used for complete vectors (up to 8 k.b.p.), or 2 % for smaller DNA fragments (up to 2 k.b.p.). Electrophoresis was performed using a Bio-Rad mini-PROTEAN® Tetra system.

Powdered agarose was suspended in diluted TAE buffer, and heated using a domestic microwave oven (with extreme caution) until the powder was fully dissolved. The solution was allowed to cool to 50 °C, and poured into the gel casting / running case, with a gel comb added to provide wells to accept the electrophoresis samples. The gel was cooled to 4 °C for 30 minutes. Samples for electrophoresis (2-5 µL for analytical; 10-20 µL for preparative-scale uses) were mixed with DNA sample dye (2 µL of 6x concentration, section 5.2.3), along with a small volume of glycerol (for low DNA concentrations) if required. The solidified gel was placed into the electrophoresis tray filled with diluted TAE buffer, and completely submerged. The gel comb was removed and DNA samples loaded into the formed wells, with one well containing a DNA marker (0.1 k.b.p. or 1 k.b.p. ladder, section 5.7.4.2). The gel was subjected to electrophoresis at 80 mA for 120 minutes (with care taken to ensure the gel did not heat during the procedure), with performance monitored by observing the progression of the sample dye.

Following completion of electrophoresis, the gel was rinsed with water and submerged in a bath of freshly-prepared ethidium bromide (EtBr*, section 5.2.3) for 30 minutes, and rinsed with water. Due to the strong intercalation effect of EtBr with DNA, bands containing DNA became visible under UV light, at concentrations as low as 10 ng. These were characterised as described below.

* - By convention, ethidium bromide is abbreviated to EtBr.

5.6.4.3 Gel visualisation

Both SDS-PAGE and Agarose Gels were visualised using a Syngene Bio Imaging "Geneflash" light box, coupled to a Syngene UP-895MD video-graphic printer. This provided both a white and UV light source, and allowed photography of the gel to provide documentary evidence of the results.

Polyacrylamide gels from SDS-PAGE were previously stained using Coomassie (Brilliant) blue, which leaves visible blue bands indicating the presence of protein. SDS-PAGE gels were run alongside a sample of a molecular weight marker (SM0431; section 5.7.4.1) comprised of several proteins of known sizes. This was used to provide an accurate assessment of sample composition.

Agarose gels stained using ethidium bromide were compared to a DNA ladder (1 k.b.p.; section 5.7.4.2) containing DNA fragments of known size, which was subjected to electrophoresis alongside DNA samples. Comparison of bands (visible under UV light after EtBr staining) within the sample allowed assessment of DNA size.

5.6.5 Competent cells and transformation

Chemically competent cells were made using two methods, to give "super competent" cells for use with low concentrations of DNA and products from PCR, and ordinarily competent cells which were used for procedures which were known to be successful, when a large excess of DNA was available or if several repetitions were to be performed. Both methods make use of positive cations which are able to aid the incorporation of DNA into the cell by preventing the unfavourable electrostatic charge between the negatively-charged DNA molecule and anions within the cell wall. The

use of rubidium is often reported[244] to be superior to calcium for this purpose, along with the inclusion of other multivalent cations in the buffers used for super-competent cells. Buffers for these procedures are found in section 5.2.2.

5.6.5.1 Chemically (ordinary) competent cells

This method made use of buffers containing a high concentration of Ca^{2+} ions to facilitate formation of small pores in the cell wall of a bacterium, allowing exogenous DNA to be incorporated into the cell via a disruptive process such as heat-shocking.

A small-scale culture of the desired strain of bacterium was grown overnight, of which a portion was used to inoculate a sterilised flask of LB growth medium (250 mL). This culture was inoculated with appropriate antibiotic (if available), and incubated at 37 °C with shaking until an OD_{600} of 0.5-0.6 was reached. The culture was immediately cooled on ice, and centrifuged when cool (4500 RPM, 30 minutes, 4°C). The cell pellet was resuspended in Buffer Ca (section 5.2.2) and centrifuged as before, which was repeated. Finally, the cells were resuspended in 10 mL of Buffer Ca. The cells were allowed to stand on ice overnight, with a maximum reported competency obtained after 24 hours[245]. At this point, aliquots (100 µL) were selected under aseptic conditions, and flash-frozen in liquid N_2 before storage at -80 °C. Transformation efficiencies up to 5×10^3 transformants per µg DNA were obtained using these cells.

5.6.5.2 Chemically (super) competent cells

Using this method, a culture was grown as before to the point at which the cells were harvested by centrifugation. However, the cells were resuspended in buffer TFB I (150 mL), and stood on ice for a further 15 minutes. After centrifugation (4500 RPM, 30 minutes, 4°C), the cell pellet was resuspended in buffer TFB II (20 mL), and flash frozen in aliquots (100 mL) before storage at -80 °C. Transformation efficiencies up to 5×10^6 transformants per µg DNA were obtained using these cells.

5.6.5.3 Transformation of plasmid DNA into competent cells

Transformation describes the process of inserting plasmid DNA into an appropriate host bacterium. A number of bacterial strains were used during this project, with

several subtle nuances required for each. In all cases, each step was performed on ice unless otherwise stated. A sample of competent cells (100 µL) was thawed on ice, which were then added to the DNA to be transformed, containing approximately 1 ng of DNA. The mixture was thoroughly mixed and allowed to stand for 30 minutes, followed by heat-shocking at 42 °C in a water bath for 60 seconds. Cells were immediately cooled on ice for 15 minutes, after which LB media (1.5 mL) at 4 °C was added, and the mixture allowed to gradually warm to room temperature. After 45 minutes from heat-shocking, cells were incubated at 37 °C for 1 hour, followed by centrifugation. The cell pellet was resuspended in LB media (150 µL), and plated on LB-agar plates prepared with appropriate antibiotics in order to provide selection for successful transformants.

5.6.6 Site Directed Mutagenesis (SDM) using Polymerase Chain Reaction (PCR)

5.6.6.1 Introduction to the Polymerase Chain Reaction (PCR)

The Polymerase Chain Reaction[246,247] is one of the key techniques performed in modern-day molecular biology, typically used to duplicate small quantities of DNA. When performed over a number of cycles, DNA is replicated almost exponentially, greatly increasing the amount present for other uses. The four key constituents of the PCR are (i) the template DNA, (ii) a set of short (20-30 b.p.) oligonucleotide “primers”, specific to known regions of the template DNA, (iii) deoxynucleotide triphosphate (dNTP) forms of the four DNA nucleobases, and (iv) a thermostable DNA polymerase, such as *Taq* or *Pfu*.

The reaction takes place in a cycle of three steps (denaturing, annealing and extension or elongation), each occurring at different temperatures. During the denaturing step (94-96 °C), the double-stranded DNA template separates, to give single DNA strands. The temperature is then reduced to the melting temperature of the specific primers used (usually 45-60 °C), allowing the oligonucleotide to anneal to each strand of the DNA. The oligonucleotide anneals preferentially to the complimentary DNA strand, as it is both much smaller, and present in great excess.

Once annealing is complete, the temperature of the reaction is increased to the optimal temperature for the DNA polymerase (72 °C for both *Taq* and *Pfu*), allowing elongation to proceed. The DNA polymerase binds at the annealed primer, and adds dNTPs complimentary to the template strand of DNA, forming a new strand of DNA. An average rate of incorporation of 1000 nucleotides per minute is typically assumed, however this is reduced for DNA polymerases which possess 3'→5' exonuclease (proof-reading) activity, such as *Pfu*. After elongation, the cycle is then repeated, with the newly-formed DNA after each cycle also able to act as a template for further reactions, increasing the proportion of DNA in a near-exponential manner.

The PCR is commonly performed using short (500 b.p. to 1.5 k.b.p.) sequences of DNA, although it is possible to perform the reaction with longer DNA strands (such as an entire plasmid) acting as the template. While larger lengths of DNA require extended elongation times, and their use carries an increased risk in undesired mutation due to processional errors by the DNA polymerase, the oligonucleotide primer may be located at any point of the DNA sequence, allowing the DNA polymerase to process in the usual 5'→3' direction until it has reproduced the entire strand of DNA. This vastly reduce the time required to prepare the reaction, as the gene of interest does not have to be isolated from the vector before the PCR.

5.6.6.2 Site Directed Mutagenesis using the PCR

A recent application of the PCR process has been to perform site directed mutagenesis (SDM) of an amino acid within a recombinant protein by mutation of an individual base pair(s) of the codon responsible for expression of the amino acid. Using a DNA primer complimentary to the template DNA containing a mismatch at the position where mutagenesis is desired allows incorporation of the template's sequence in all daughter DNA created by the DNA polymerase. Further reactions using either strand of DNA (parental or daughter), each give rise to a mutant strand, which proportionally increases the quantity of mutant DNA with each successive round of PCR.

Key to the success of this method is separation of the parental (original sequence) and daughter (mutant) DNA. The template DNA used for the PCR was obtained from XL1-

blue cells, which possesses the dam^+ genotype, leading to methylation of every adenine group in DNA within the cell. This allows differentiation between parental (original) and daughter (mutant) DNA created by the PCR, as adenine present within the daughter DNA is unmethylated. Treating the mixture using the restriction enzyme *DpnI* (see section 5.6.6.4) cleaves DNA between adenine and thymine in the sequence 5'-GA(Me)*TC-3' when adenine is methylated. This effectively removes the methylated parental DNA, preventing its subsequent transformation – while newly-formed daughter (mutant) DNA is undamaged by *DpnI*, and free to be transformed. DNA created by the PCR was transformed into super-competent cells (2.6.5.2), with initial transformation efficiencies up to 1×10^2 transformants per µg DNA.

5.6.6.3 Construction of AsLOV2(C450X) mutants

Construction of several mutants of AsLOV(C450X) (where X = alanine, glycine, serine, methionine or aspartic acid) was performed using the methodology given in sections 5.6.6.2 and 5.6.6.3 above. Oligonucleotides used to create the mutants using the PCR are described in section 5.7.3.

5.6.6.4 Restriction digestion of DNA samples

The ability to manipulate DNA by cutting is performed using enzymes known as restriction endonucleases. These enzymes recognise specific sequences of DNA, which are generally palindromic (enabling digestion of both the sense and anti-sense strands). The enzymes cut at a specific point within the restriction site, with two possible outcomes: cuts with an overhanging (also known as "sticky") DNA section (either 5' or 3'), or a cut with no overhanging section (known as a "blunt" cut). Other uses of restriction endonucleases include the differentiation between methylated and non-methylated DNA (described in section 5.6.6.2) and the nicking of cyclic DNA from plasmid vectors (cutting at a single site, to allow the ordinarily super-coiled cyclic DNA to relax, for accurate results using agarose electrophoresis).

In cases where isolation of a DNA fragment was desired (for example, removal of a gene from a plasmid vector), digestion was necessary at both ends of the DNA fragment. This could be performed in two ways: (i) using two sequential digestions,

requiring inactivation of the first endonuclease before addition of the second, but allowing purification of the DNA after each reaction, or (ii) using a "double digestion". The use of double digestions (with the mixture containing both endonucleases) was significantly faster to perform, but extra care was necessary to check the compatibility of the enzymes and buffers used.

For analytical digestions, between 100-500 ng of DNA was used, treated with 1 unit (U) of the desired restriction endonuclease (1 U is defined by the manufacturer (New England Biolabs) as "the amount of enzyme required to completely digest 1 µg of substrate DNA in a reaction volume of 50 µL in 1 hour") and buffer (appropriate to the enzyme) added to the required volume. The reaction was heated at 37 °C for 1 hour, and analysed using agarose gel electrophoresis (section 5.6.4.2). Preparative-scale digestions were performed using 1-5 µg of DNA, in a total reaction volume of 100-250 µL, with appropriate volumes of buffer and 5-10 U of endonuclease. Incubation times up to 4 hours were used for single-enzyme reactions, and 6 hours for double-digestions. Single-enzyme reactions were heat-inactivated in a hot water bath (temperatures specific to each endonuclease given in table 10 below) for 20 minutes after completion of the reaction. Conditions for double-digests were obtained from New England Biolabs.

Enzyme	Recognition Sequence	Optimal NeBuffer	Inactivation Temperature	Comments
BamHI	5'-G*GATCC-3'	3 (+ BSA)	-	
BamHI-HF	5'-G*GATCC-3'	4	-	
DpnI	5'-GA(Me)*TC-3'	4	80 °C	Only cuts where adenine is methylated
EcoRI	5'-G*AATTC-3'	1-4	65 °C	
EcoRI-HF	5'-G*AATTC-3'	4	65 °C	
HindIII	5'-A*AGCTT-3'	2	65 °C	
HindIII-HF	5'-A*AGCTT-3'	4	80 °C	
KpnI	5'-GGTAC*C-3'	1 (+ BSA)	-	
NdeI	5'CA*TATG-3'	4	65 °C	

Table 10: Restriction endonucleases used during this research. All enzymes were provided by New England Biolabs. Cutting site within the recognition sequence is indicated by a star. All endonucleases (except DpnI) caused sticky-ended fragments

5.6.7 Purification of expressed proteins

5.6.7.1 Amylose Affinity Chromatography

Buffers used for amylose affinity chromatography of the *C. ammoniagenes* bifunctional RfK-FADS protein are described fully in section 5.2.7. Cells were thawed in buffer RfK 1, before addition of PMSF (5 mM). Cells were ultrasonically lysed and centrifuged (section 5.6.3.4), with the filtered lysate loaded onto a column (1 cm^2 x 6 cm) of amylose resin, which had been prepared according to the manufacturer's instructions (GE Healthcare). After loading, the column was washed using 5 column volumes (c.v.) of buffer RfK 1. UV absorbance of the flow-through was monitored at 280 nm, and once this decreased to zero, the RfK-FADS protein was eluted using buffer RfK 2. Fractions were collected and analysed by SDS-PAGE.

5.6.7.2 Immobilised Metal Affinity Chromatography (IMAC)

IMAC was performed to purify several protein samples, which are individually described in chapter 4. Buffers required are listed in sections 5.2.8 and 5.2.9.

IMAC relies on the principle of a target protein coupled to residues or a further protein, which contains several surface amino acids with a high affinity for metal ions. Typically, histidine residues are most commonly used for this purpose, due to their high affinity for transition metals (particularly Ni, Cu, Zn and Fe). This can be utilised to purify a particular protein from a complex mixture, using a metal-chelating resin. Protein retained by the resin may be released using a buffer containing imidazole, which competes with the histidine residues and releases the protein. Nickel was most commonly used as it is strongly bound by the resin, while allowing efficient release of isolated proteins using gentle gradients of imidazole.

Two metal-chelating resins were utilised – Ni-NTA and Chelating Sepharose Fast Flow (GE Healthcare) – retained within an XK 16/20 Column (GE Healthcare), equipped with an external jacket for cooling of the sample, if required. The column was capped with two QuickLock AK 16 adaptors, to enable visualisation of the entire resin bed to allow accurate monitoring of protein binding. The resin bed (5-20 mL) was cast with continual pouring, and the column packed by hand with headspace over the resin

minimised. Solutions were loaded onto the column and forced through the resin using a Watson-Marlow 101U/P peristaltic pump. Purified proteins were released with imidazole-containing buffers, with imidazole removed from the eluted protein by dialysis, FPLC (size-exclusion) or desalting (section 5.6.7.4)

5.6.7.3 Fast Protein Liquid Chromatography (FPLC)

FPLC was used to purify protein samples. This technique is similar to HPLC (section 5.5.2.3), where a sample containing several analytes is forced through a chromatographic column (stationary phase) containing a solid medium which retards the progress of the analytes based upon their physical characteristics (size, polarity etc.), allowing separation. The buffer(s) used for the FPLC act as the mobile phase, with either a single buffer or a gradient of ionic strength used to facilitate separation.

The apparatus for this included an Amersham Biosciences ÄKTA FPLC system, fitted with UPC-900, P-920, M-925 and Inv-907 modules and a Frac-950 autosampler unit. Columns used were Superdex® 75 or 200 size exclusion columns, or Q-Sepharose® anion exchange columns (Amersham Biosciences). Each was equilibrated with the desired mobile phase (5 c.v.), before samples were applied. The sample was eluted with absorbance of the eluent monitored at an appropriate wavelength. Fractions containing the target protein were examined by SDS-PAGE to confirm the presence of the desired protein, and concentrated by ultrafiltration if necessary.

5.6.7.4 Desalting using FPLC

A pre-packed HiPrep™ 26/10 desalting column was used to exchange the buffer of protein samples. The column was equilibrated with a minimum of 5 c.v. of the desired buffer. Pre-concentrated protein samples (up to 10 mL) were applied to the column at a flow rate of 2 mL min^{-1}, followed by elution with a flow rate of 5 mL min^{-1}. Fractions were examined by UV-visible spectroscopy (for chromophore-containing solutions) and SDS-PAGE, and similar fractions were combined. Electrical conductivity was also monitored, and used to indicate elution of the original buffer.

5.6.8 Spectroscopy of protein samples

5.6.8.1 UV-visible spectroscopy

UV spectra were measured using a Jasco UV-660 UV-visible spectrophotometer coupled to a variable-temperature control unit, able to maintain housing temperatures at +/- 0.5°C from 5 °C to above 60 °C. Flushing the chamber with N_2 was used to prevent the formation of condensation on the surface of the sample cuvette at non-ambient temperatures. Spectra were recorded at 20 °C unless otherwise stated. Samples were examined in suitable solvents and measured within a capped Hellma semimicro CXA-145-307L far-UV quartz 1400 µL cuvette.

Analysis of DNA concentrations was performed using a Thermo Scientific NANODROP 1000 UV-visible Nanodrop Spectrophotometer, to determine sample concentrations and purity.

5.6.8.2 Irradiation of AsLOV2 Samples

Irradiation of samples was performed using a variable-power LED lighting apparatus created specifically for this project. This comprised a 12 V DC power supply powering an LED, controlled through a potentiometer. SMD LEDs were fitted to a removable base, allowing replacement of the LED to give alternative wavelengths. Light was focussed to a narrow beam using an encapsulating lens (14 ° beam angle, part no. OPC1-1). SMD LEDs (Philips LUMILEDS, Rebel range) of varying colours were adapted for use with this equipment; Royal Blue (440-460 nm, part number LXML-PR01-0425) and Clear White (400-800 nm part number LXML-PWC1-0900) were commonly used.

Irradiation at lower wavelengths (254 and 365 nm) was provided by a Spectroline ENF-240C/FE Dual Wavelength lamp, used previously for analysis of TLCs (section 5.3). Other wavelengths between 260 nm and 600 nm were obtained using a Perkin Elmer LS 55 Luminescence Spectrometer, with typical bandwidths of 2-10 nm.

5.6.8.3 Circular Dichroism (CD)

Circular Dichroism was performed using achiral buffers within an Applied Photophysics Chirascan spectrometer, coupled to a CD Photomultiplier (with a

detection range of 160-850 nm) and a Scanning Emission Monochromator (4.65 nm/mm band pass). All spectra were recorded at 20 °C, and have been normalised by calculation of the Mean Residue Ellipticity (MRE), as fully described in section 3.4.0.

5.7 Bacterial strains & plasmid vectors

5.7.1 Bacterial strains and genotypes

***E. coli* Strain**	**Genotype**	**Source**
XL1-blue	recA1, endA1, gyrA96(Nal^r), thi-1, hsdR17, supE44, relA1, lac, dam [F', proAB, lacI^q ZΔM15, Tn10 (tet^r)]	Bullock et al., 1987 obtained from Stratagene
BL21 (DE3)	F-, ompT, hsdS$_B$(r$_B$-, m$_B$-), gal, dcm, λ(DE3)	Promega
BL21(DE3) [pLysS]	F-, ompT, hsdS$_B$(r$_B$-, m$_B$-), gal, dcm, λ(DE3), pLysS (Cam^r)	Invitrogen
BL21(DE3) STAR	F-, ompT, hsdS$_B$(r$_B$-, m$_B$-), gal, dcm, rne131, λ(DE3)	Invitrogen
M15 [pRep4]	Nal^r, Str^s, Lac$^-$, Ara$^+$, Gal$^+$, Mtl$^-$, F$^-$, RecA$^+$, Uvr$^+$, Lon$^+$	D. Stüeber, F. Hoffmann-La Roche Ltd.[191]; obtained from QIAgen
JM101	supE, thi-1 Δ(lac-proAB) endA$^+$ [F´, traD36, proAB, lacI^q ZΔM15	J. Messing *et al.*[193], 1981 obtained from Stratagene

Table 11: Bacterial strains, genotypes and original sources from each cell line

E. coli strains BL21(DE3)[pLysS] and M15[pRep4] each contain a plasmid (pLysS and pRep4, respectively) to regulate expression of other plasmids within the cell. Strains XL1-blue and JM101 require supplemental thiamine, as their ability to produce this was removed to prevent unwanted propagation outside of controlled conditions.

5.7.2 Plasmid vectors

Plasmid Vector	Description	Source	Resistance
pMal-C2-RibfK-Ca	Expression vector for *E. coli*, containing the gene from *C. ammoniagenes* encoding the bifunctional riboflavin kinase-FAD synthetase enzyme with an N-terminal maltose binding domain	K. Kemter *et al.*,[177] 2004	Amp
pET14b-*Sp*RfK	Expression vector for *E. coli*, containing the gene for expression of the riboflavin kinase from *Schizosaccharomyces pombe* as an N-terminal 6xHis-tagged protein	H. Collins, University of Canterbury, 2010	Amp
pNCO113-HISACT(C49S)	Expression vector based on pNCO113 (Stüeber *et al.*[189]), modified by Kay *et al.*[4] to contain an N-terminal gene encoding the protein Hisactophilin from *Dictyostelium discoideum*[190] with C-terminal thrombin site, and further modified (Fischer *et al.*[1]) to mutate the single cysteine residue of hisactophilin (to ensure no effect with flavin)	M. Fischer, G. Richter[1], 2007	Amp
pNCO113-HISACT(C49S)-AsLOV2 (WT)	Expression vector as above, including the gene for expression of Avena sativa PHOT1-LOV2 protein (wild type)	M. Fischer, G. Richter[1], 2007	Amp
pNCO113-HISACT(C49S)-AsLOV2(T418I, C450A)	As above, with mutation of AsLOV2 gene to replace cysteine 450 with alanine. A second mutation (threonine 418 to isoleucine) was also present	A. Wood, 2012	Amp
pNCO113-HISACT(C49S)-AsLOV2(C450G)	As above, with mutation of AsLOV2 gene to replace cysteine 450 with glycine	A. Wood, 2012	Amp
pNCO113-HISACT(C49S)-AsLOV2(C450M)	As above, with mutation of AsLOV2 gene to replace cysteine 450 with methionine	A. Wood, 2012	Amp
pNCO113-HISACT(C49S)-AsLOV2(C450S)	As above, with mutation to AsLOV2 gene to replace cysteine 450 with serine	A. Wood, 2012	Amp

pNCO113-HISACT(C49S)-AsLOV2(C450D)	As above, with mutation of AsLOV2 gene to replace cysteine 450 with aspartic acid	A. Wood, 2012	Amp
pRep4	Plasmid used to regulate expression by the T5 promoter. This plasmid expresses the *lac repressor* protein encoded by the *lacI* gene[192]; IPTG inactivates this gene allowing transcription	QIAgen (within *E. coli* strain M15 [pRep4])	Kan
pLysS	Plasmid used to supress basal expression from the *T7 promoter* by expression of *T7 lysozyme*, which acts as an inhibitor to *T7 RNA polymerase*. IPTG blocks expression of this gene, allowing transcription from the *T7 promoter*	Novagen (within *E. coli* strain BL21(DE3) [pLysS])	Cam

Table 12: Description of Plasmid Vectors. "Amp" refers to ampicillin, "Kan" refers to kanamycin, and "Cam" refers to chloramphenicol

5.7.3 Oligonucleotide primers

Primers were designed using the PrimerX tool[248], and supplied by Eurofins MWG Operon.

Name	**Sequence (5' --> 3')**
T5 Promoter (Fw)	TCATAAAAAATTTATTTGCTTTGTGAGCGGATAACAATTATAATA
T5 Promoter(Re)	TATTATAATTGTTATCCGCTCACAAAGCAAATAAATTTTTTATGA
AsLOV2-C217A (Fw)	GAAATTCTGGGTCGTAAC**GCG**CGTTTTCTTCAAGGTCCTG
AsLOV2-C217A (Re)	CAGGACCTTGAAGAAAACG**CGC**GTTACGACCCAGAATTTC
AsLOV2-C217G (Fw)	GAAATTCTGGGTCGTAAC**GGT**CGTTTTCTTCAAGGTCCTG
AsLOV2-C217G (Re)	CAGGACCTTGAAGAAAACG**ACC**GTTACGACCCAGAATTTC
AsLOV2-C217M (Fw)	GAAATTCTGGGTCGTAAC**ATG**CGTTTTCTTCAAGGTCCTG
AsLOV2-C217M (Re)	CAGGACCTTGAAGAAAACG**CAT**GTTACGACCCAGAATTTC
AsLOV2-C217S (Fw)	GAAATTCTGGGTCGTAACT**CT**CGTTTTCTTCAAGGTCCTG
AsLOV2-C217S (Re)	CAGGACCTTGAAGAAAACG**AG**AGTTACGACCCAGAATTTC
AsLOV2-C217D (Fw)	GAAATTCTGGGTCGTAAC**GAT**CGTTTTCTTCAAGGTCCTG
AsLOV2-C217D (Re)	CAGGACCTTGAAGAAAACG**ATC**GTTACGACCCAGAATTTC

Table 13: Primers used for sequencing, or creation of site-specific mutants of Avena sativa PHOT1-LOV2 domains. The codon describing the desired mutation is underlined, and nucleobases altered (from the wild-type gene) are shown in bold type

5.7.4 Gel Ladders

5.7.4.1 SDS-PAGE

SDS-PAGE was analysed by comparison to a molecular weight marker (SM0431), supplied by Fermentas. This contained proteins of the following masses:

MWt, kDa	Protein	Source
116.0	β-galactosidase	*E. coli*
66.2	Bovine serum albumin (BSA)	bovine plasma
45.0	Ovalbumin	avian (chicken) albumen
35.0	Lactate dehydrogenase	porcine muscle
25.0	Rease Bsp98l	*E. coli*
18.4	β-lactoglobulin	bovine milk
14.4	Lysozyme	avian (chicken) albumen

Table 14: Components of protein molecular weight marker SM 0431

5.7.4.2 Agarose Gel Electrophoresis

Ladders for analysis of DNA products separated by agarose gel electrophoresis were supplied by New England Biolabs. Composition is described in table 15 below.

1 k.b.p. ladder

Kilobase pairs	Mass (ng) DNA present
10.0	42
8.0	42
6.0	50
5.0	42
4.0	33
3.0	**125**
2.0	48
1.5	36
1.0	42
0.5	42

Table 15: Components of 1 k.b.p. ladder DNA ladder, visible with staining using EtBr. Bold type indicates strongly-absorbing bands, for additional reference

Chapter 6

Experimental Detail –

Chemical Synthesis

Chapter 6 - Experimental - Chemical Synthesis

6.0.1 Synopsis

In this chapter, the experimental detail necessary to produce 5-deazariboflavin (section 6.1) and 1-deazariboflavin (section 6.2) is reported. In addition, the synthesis of 8-demethyl-5-deazariboflavin (section 6.3) is described, which has not been previously reported in the scientific literature, and serves as a good comparison for the efficiency of the formation of 5-deazariboflavin. The formation of several small molecules is also reported (section 6.4), which were used in an attempt to provide novel routes to prepare 1-deazariboflavin, 1,5-dideazariboflavin, 3-deazariboflavin and 3,5-dideazariboflavin (3.6). The rationale for this is described in chapter 2.

Each successful reaction described is supported by appropriate experimental detail and appropriate spectroscopic information, for reference. Several reactions have been studied exhaustively using alternative conditions, reagents, or solvents; a general method has been given for these reactions, with alternative reagents, conditions and solvents (along with yields obtained) described in tabular form for brevity. In these cases, spectroscopy was compared to the product formed by the original reaction and is described only as necessary to assist later discussion. Each reaction was performed a minimum of three times, with average yields quoted.

6.1 Synthesis of 5-Deazariboflavin

This section describes the synthesis of 5-deazariboflavin **39**.

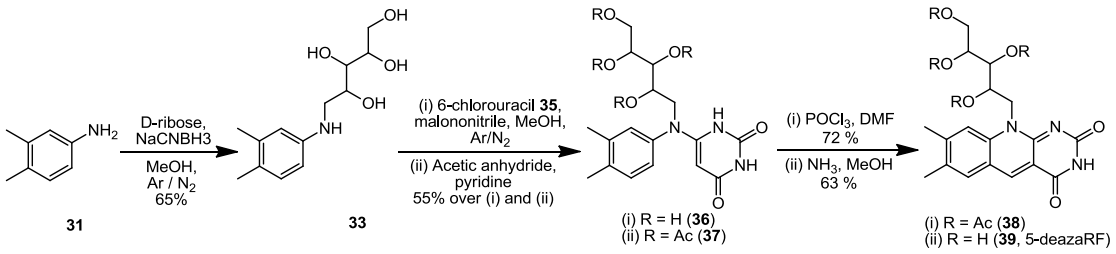
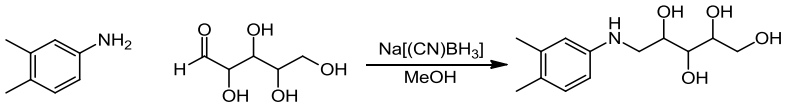


*Figure104: Overall Synthetic Route to 5–deazariboflavin **39***

6.1.1.1 Formation of 3,4-dimethyl-N(ribityl)-aniline 33



31 32 33

3,4-Dimethylaniline **31** (3.00 g, 25 mmol), $Na[(CN)BH_3]$ (3.15 g, 50 mmol, 2 equiv.) and D-ribose (7.50 g, 50 mmol, 2 equiv.) were combined in a flame-dried 3-neck RBF and dried under high vacuum for 4 hours. Thoroughly dried MeOH (120mL) was added, and the mixture was stirred at reflux under an Ar / N_2 atmosphere until completion (48 hours, determined by TLC). The solvent was removed under reduced pressure to furnish a dark brown viscous oil, which was dissolved in $HCl_{(aq)}$ (20mL, 1M) with shaking until gas evolution ceased. The pH of the mixture was adjusted to pH 7.0 using a saturated $NaHCO_3$ solution (8mL) and washed using an aliquot of EtOAc (40 mL). Addition of EtOAc stimulated formation of a solid precipitate within the mixture, which was isolated by filtration to give a grey-white solid. This was dried under high vacuum for 4 hours to give a white powder, with spectroscopic analysis confirming formation of 3,4-dimethyl-N(ribityl)-aniline **33** (5.97 g, 23.4 mmol, 47 %). The aqueous layer was extracted using further EtOAc (4 x 40 mL), washed with brine (2 x 40mL) and dried over $MgSO_4$. Solvent was removed under reduced pressure, and the light-brown crystals were recrystallised from 10 % $EtOH_{(aq)}$ to furnish purified 3,4-dimethyl-N(ribityl)-aniline **33** as a white powder (2.36 g, 9.24 mmol, 18 %). The combined yield of 3,4-dimethyl-N(ribityl)-aniline **33** (8.33 g, 32.64 mmol) was therefore 65 %.

TLC 10:1 $CHCl_3$:MeOH **RF**: 0.82 (**31**), 0.18 (**33**), 0.00 (**32**); NinH staining; **MP** 142-144 °C (*lit. 142-143 °C* [249]); **^{1}H NMR** (400 MHz, MeOD-d_3) δ (ppm) 6.90 (d, 1H, *J* = 8.2 Hz), 6.57 (d, 1H, *J* = 2.4 Hz), 6.49 (dd, 1H, J_{AM} = 2.4 Hz, J_{AX} = 8.2 Hz), 4.91 (br s (solvent, residual)), 3.93 (m, 1H), 3.78 (m, 2H), 3.65 (m, 2H), 3.45 (dd, 1H, J_{AM} = 3.6 Hz, J_{AX} = 12.8 Hz), 3.33 (m, 2H), 3.10 (q, *J* = 8.0 Hz), 2.19 (s, 3H), 2.14 (s, 3H); **^{13}C NMR** (100 MHz) δ (ppm) 146.2 (C), 136.7 (C), 129.6 (CH), 125.8 (C), 115.6 (CH), 111.3 (CH), 73.2 (CH), 72.5 (CH), 70.2 (CH), 62.8 (CH_2), 46.5 (CH_2), 18.6 (CH_3), 17.3 (CH_3); **EI MS (+)** (m/z) 255 ([M], 100%); **HR MS (APCI +)** found 256.1539, calc. 256.1549 for [M+1]

6.1.1.2 Optimisation of conditions - 3,4-dimethyl-N(ribityl)-aniline 33

Numerous solvents, reducing agents, reaction times and other conditions for this reaction were examined following the general method described above, with results shown in table 16 below. Full spectrometric data was obtained for each reaction (including TLC RF values, ^{1}H and ^{13}C NMR, and EI or ESI MS) and compared to that obtained for the original reaction. In all cases, each reaction was performed using two molar equivalents of D-ribose and the reducing agent selected. Solvent volumes required were adjusted for each reaction, with optimal volumes allowing solubility of the reagents at temperatures slightly below the reflux temperature of the solvent. Some solvents (THF, DCM, $CHCl_3$) were unable to fully solubilise the reagents even at reflux; in these cases, a similar volume to those used in other reactions performed at similar scales was used.

Alterations to the original reaction were carried out as direct substitutions (if possible), with changes made to the general reaction method only made when necessary due to solubility problems or with the use of additional reagents. Spontaneous product precipitation did not occur in all cases. All reaction mixtures were extracted using an appropriate volume of EtOAc (based upon reaction scale), until no further product remained in the aqueous layer (determined by TLC). Reactions performed using additional desiccants were performed with the desiccant added alongside all other solid reagents, and were removed by filtration immediately before initial concentration of the reaction mixture upon completion. Any product isolated was purified as necessary by recrystallisation.

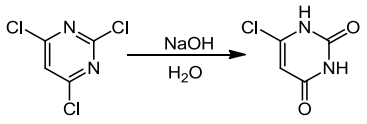
In the case of reactions performed using acidic catalysis, glacial acetic acid was pre-dried using molecular sieves, and added to the reaction mixture dropwise after all components had fully dissolved. Reactions performed using DMAP were subjected to further purification by recrystallisation or flash chromatography to remove DMAP from the product after extraction.

	Scale (aniline, mM)	Reducing Agent	Solvent	Temperature (° C)	Conditions	Atmosphere	Reaction Time (hours)	Yield (%)
1	12.5	$Na[(CN)BH_3]$	MeOH	reflux	-	Ar/N_2	48	57
2	25.0	$Na[(CN)BH_3]$	MeOH	reflux	-	Ar/N_2	48	65
3	25.0	$Na[(CN)BH_3]$	MeOH	reflux	-	Ar/N_2	72	64
4	75.0	$Na[(CN)BH_3]$	MeOH	reflux	-	Ar/N_2	72	63
5	0.83	$Na[(CN)BH_3]$	MeOH	RT	Undried solvent	Air	72	-
6	0.83	$Na[(CN)BH_3]$	MeOH	reflux	Undried solvent	Air	72	22
7	0.83	$Na[(CN)BH_3]$	MeOH	reflux	AcOH catalyst, 0.2 eq.	Ar/N_2	24	32
8	0.83	$Na[(CN)BH_3]$	MeOH	reflux	AcOH catalyst, 0.2 eq.	Ar/N_2	48	19
9	12.5	$Na[(CN)BH_3]$	MeOH	reflux	DMAP catalyst, 0.2 eq	Ar/N_2	72	39
10	12.5	$Na[(CN)BH_3]$	MeOH	reflux	-	Ar/N_2	48	57
11	0.83	$Na[(AcO)_3BH]$	MeOH	reflux	-	Ar/N_2	48	19
12	0.83	$Na[(CN)BH_3]$	MeOH	reflux	Reducing agent added after 4 hrs	Ar/N_2	48	51
13	0.83	$Na[(CN)BH_3]$	MeOH	reflux	AcOH added at 0 hours, reducing agent added after 4 hours	Ar/N_2	24	23
14	0.83	$Na[(AcO)_3BH]$	MeOH	reflux	Reducing agent added after 4 hours	Ar/N_2	48	31

	Scale (aniline, mM)	Reducing Agent	Solvent	Temperature (° C)	Conditions	Atmosphere	Reaction Time (hours)	Yield (%)
15	0.83	$Na[(AcO)_3BH]$	MeOH	reflux	AcOH added at 0 hrs, reducing agent added after 4 hrs	Ar/N_2	24	7
16	0.83	$Na[(CN)BH_3]$	MeOH	reflux	MgSO4 added, 5 eq.	Ar/N_2	48	64
17	0.83	$Na[(AcO)_3BH]$	MeOH	reflux	MgSO4 added, 5 eq.	Ar/N_2	24	28
18	0.83	$Na[(AcO)_3BH]$	MeOH	reflux	MgSO4 added, 5 eq.	Ar/N_2	48	54
19	0.85	$Na[(CN)BH_3]$	THF	RT	-	Ar/N_2	72	44
20	0.85	$Na[(CN)BH_3]$	THF	reflux	-	Ar/N_2	48	49
21	0.85	$Na[(AcO)_3BH]$	THF	RT	-	Ar/N_2	72	8
22	0.85	$Na[(AcO)_3BH]$	THF	reflux	-	Ar/N_2	48	33
23	0.83	$Na[(CN)BH_3]$	MeCN	RT	-	Ar/N_2	48	-
24	0.83	$Na[(CN)BH_3]$	MeCN	reflux	-	Ar/N_2	48	8
25	0.83	$Na[(CN)BH_3]$	MeCN	reflux	AcOH catalyst, 0.2 eq.	Ar/N_2	24	-
26	0.83	$Na[(CN)BH_3]$	EtOH	reflux	-	Ar/N_2	68	-
27	0.83	$Na[(CN)BH_3]$	isopropanol	reflux	-	Ar/N_2	72	12
28	0.83	$Na[(CN)BH_3]$	DCM	reflux	-	Ar/N_2	48	-
29	0.83	$Na[(CN)BH_3]$	CHCl3	reflux	-	Ar/N_2	72	18

Table 8: Varying reaction conditions used for the formation of 3,4-dimethyl-N(ribityl)-aniline

6.1.2.1 Formation of 6-chlorouracil 35 (original method)



34 35

To a freshly-prepared solution of aqueous NaOH (40 mL, 5 M, 192 mmol), 2,4,6-trichloropyrimidine **34** (2.50 mL, 3.99 g, 21.7 mmol) was added dropwise, and the mixture heated to reflux. A solid white plaque was commonly observed above the reaction mixture after 16 hours, which was manually disrupted and slowly re-dissolved over 3 hours. After 72 hours, the reaction was cooled to room temperature, and the acidity of the solution adjusted to pH 1-2 using concentrated HCl (6 M, 10 mL). A white precipitate formed upon addition of the acid, and the mixture was maintained in an ice bath for 30 minutes to facilitate further precipitation. The fine white precipitate was collected by filtration under suction, and washed with warm deionised water (5 mL), acetone (5 mL), and dried under high vacuum, yielding 6-chlorouracil **35** (2.26 g, 15.4 mmol, 71 %) as a fine white powder.

TLC 10:1 $CHCl_3$:MeOH **RF**: 0.26 (**35**); 5:1 $CHCl_3$:MeOH **RF**: 0.59 (**35**); NinH / I_2 staining; **MP** 312-315°C (deg; *lit: deg >300°C*[250]); **^{1}H NMR** (400 MHz, DMSO-d_6) δ (ppm) 12.10 (br s, 1H,), 11.32 (br s, 1H), 5.77 (s, 1H), 3.36 (br s, H_2O); **^{13}C NMR** (100 MHz, DMSO-d_6) δ (ppm) 162.7, 150.5, 145.1, 99.7; **EI MS (+)** (m/z) 146 ([M+1], 100%)

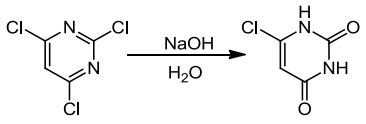
6.1.2.2 Formation of 6-chlorouracil 35 (survey of conditions)

Alternative conditions for the formation of 6-chlorouracil **35** were examined leading to the enhanced method described in section 6.1.2.3. Each of these reactions was performed following the general method given in section 6.1.2.1 above, and is described in table 17 below. The product of each reaction was isolated as before, by precipitation upon pH adjustment to pH 1-2, with full spectral data obtained for each product. However, TLC analysis revealed significant levels of 6-chlorouracil remaining in the aqueous layer, which was previously unrecovered. Extraction of the aqueous layer using EtOAc (table 17, reaction 18) significant improved isolation of **35**, from 64 % by precipitation alone to 81 %.

	Scale (mM)	Solvent	Base	Base conc. (M)	Time (hours)	Reaction Temperature	Yield (%)
1	2.5	H_2O / HO^-	NaOH	2.0	12	100 °C	35
2	35	H_2O / HO^-	NaOH	2.2	48	100 °C	65
3	82	H_2O / HO^-	NaOH	2.5	48	100 °C	52
4	10	H_2O / HO^-	NaOH	2.0	72	100 °C	56
5	22	H_2O / HO^-	NaOH	5.0	72	100 °C	71
6	8.5	H_2O / HO^-	NaOH	2.6	96	100 °C	12
7	1.7	H_2O / HO^-	NaOH	1.0	48	100 °C	4
8	1.7	H_2O / HO^-	NaOH	3.0	48	100 °C	39
9	1.7	H_2O / HO^-	NaOH	5.0	48	100 °C	68
10	1.7	H_2O / HO^-	NaOH	5.0	48	RT	0
11	1.7	H_2O / HO^-	KOH	1.0	48	100 °C	20
12	1.7	H_2O / HO^-	KOH	3.0	48	100 °C	37
13	1.7	H_2O / HO^-	KOH	5.0	48	100 °C	52
14	1.7	H_2O / HO^-	KOH	5.0	48	RT	0
15	1.7	H_2O / HO^-	$Mg(OH)_2$	3.0	72	100 °C	0
16	3	toluene; H_2O / HO^-	NaOH	2.5	72	100 °C	21
17	8.2	DCM; H_2O / HO^-	NaOH	2.5	72	40 °C	14
18	10	H_2O / HO^-	NaOH	5.0	48	100 °C	*64*
							81*

*Table 17: Conditions used for the formation of 6-chlorouracil **35**. Reaction 18 gave a yield of 64 % when the product was isolated by precipitation; however, extraction of the aqueous mother liquor using EtOAc allowed isolation of further 6-chlorouracil with combined (filtrate and extract) yields of 81 %. "RT" refers to an assumed room temperature between 15 and 20 °C, with each reaction repeated in duplicate or more.*

6.1.2.3 Formation of 6-chlorouracil (enhanced method)

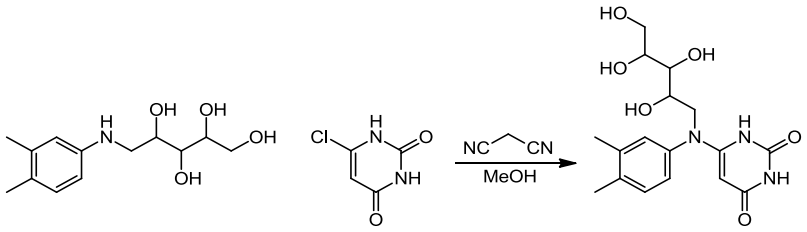


34 **35**

This reaction was initiated as described in section 6.1.2.1, using 1.15 mL (1.83 g, 10.0 mmol) of 2,4,6-trichloropyrimidine **34**. Upon completion (48 hours, determined by TLC), the solution was cooled to room temperature and the pH adjusted to 6.5 using aqueous HCl (2 M, 6 mL); precipitation of 6-chlorouracil did not routinely occur under these conditions. The aqueous solution was extracted thoroughly using EtOAc (5 x 30 mL), the organic layer was washed using brine (2 x 30 mL) and dried over $MgSO_4$. Removal of solvent under reduced pressure furnished 6-chlorouracil **35** (1.18 g, 8.05 mmol, 81 %) as a fine white solid.

TLC 10:1 $CHCl_3$:MeOH **RF**: 0.25 (**35**); 5:1 $CHCl_3$:MeOH **RF**: 0.61 (**35**); NinH / I_2 staining; **MP** 312-315°C (deg; *lit: deg >300°C*[250]); **^{1}H NMR** (250 MHz, DMSO-d_6) δ (ppm) 12.12 (br s, 1H,), 11.28 (br s, 1H), 5.75 (s, 1H), 3.37 (br s, H_2O), **^{13}C NMR** (62.5 MHz, DMSO-d_6) δ (ppm) 162.4, 150.1, 144.7, 99.6; **EI MS (+)** (m/z) 146 ([M], 100%)

6.1.3.1 Formation of bicyclic intermediate 36 (original method)



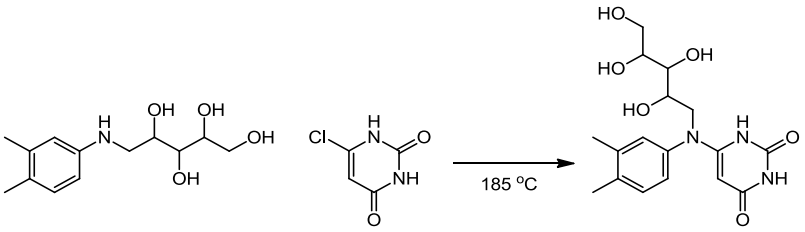
33 **35** **36**

3,4-Dimethyl-N(ribityl)-aniline **33** (1.53 g, 6.0 mmol), 6-chlorouracil **35** (1.06 g, 7.2 mmol, 1.2 equiv.) and malononitrile (160 mg, 2.4 mmol, 0.4 molar equiv.) were combined and dried under vacuum in flame-dried glassware. Thoroughly dried methanol (50 mL) was added, and the resulting suspension heated under N_2 to reflux. Further malononitrile (80 mg, 1.2 mmol, 0.2 molar equiv.) was added after 54 hours, with completion observed 94 hours. The reaction was cooled to room temperature, and solvent removed under reduced pressure. The product was dried under high vacuum for 4 hours, yielding crude bicyclic intermediate **36** as a viscous brown oil which was typically used for the subsequent reaction without further purification.

This reaction was repeated at several scales using a number of additives (described in section 2.1.3) added at appropriate points of the reaction; the effect of these is detailed in the relevant discussion, although the method described above gave optimal results. All products were subjected to acetylation immediately following this reaction, after which purification of the acetylated product allowed accurate assessment of the yield over two steps.

TLC 10:1 $CHCl_3$:MeOH **RF**: 0.24 (**35**), 0.20 (**36**), 0.18 (**33**); 20:1 $CHCl_3$:MeOH **RF**: 0.66 (**35**), 0.62 (**36**), 0.12 (**33**); NinH / I_2 staining; **36** exhibits blue-white fluorescence; **^{1}H NMR** (400 MHz, MeOD-d_3) δ (ppm) 6.79 (d, 1H, J = 8.0 Hz), 6.47 (d, 1H, J = 2.2 Hz), 6.39 (dd, 1H, J_{AM} = 2.2 Hz, J_{AX} = 8.0 Hz), 5.63 (s, 1H), 4.51 (br s, 1H), 3.82 (m, 1H), 3.67 (m, 2H), 3.54 (m, 2H), 3.33 (dd, 1H, J_{AM} = 3.6 Hz, J_{AX} = 12.8 Hz), 3.00 (m, 2H), 2.90 (dd, 1H, J_{AM} = 8.0 Hz, J_{AX} = 12.8 Hz), 2.08 (s, 3H), 2.03 (s, 3H); **EI MS (+)** (m/z) 388 ([M + Na^+], 65 %); 366 ([M+1], 100 %); **ESI MS (-)** 364 ([M-1], 100 %)

6.1.3.2 Formation of bicyclic intermediate 36 (solvent-free heating)



33 **35** **36**

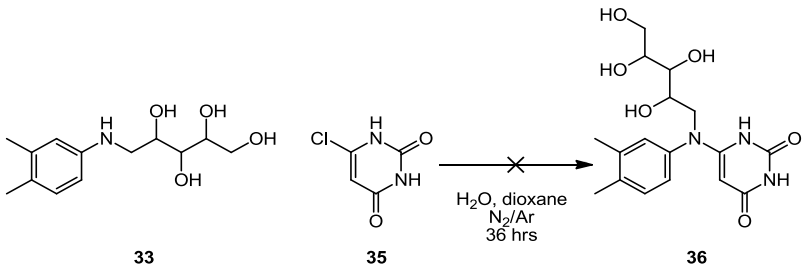
3,4-Dimethyl-N(ribityl)-aniline **33** (150 mg, 0.59 mmol) and 6-chlorouracil **35** (95 mg, 0.65 mmol, 1.1 equiv.) were powdered together and dried under high vacuum for 4 hours. The solid mixture was heated under ambient pressure to 185°C for 15 minutes, with evolution of acidic vapours observed. The mixture was allowed to cool to room temperature, and washed with diethyl ether (5 mL) and filtered. The retentate was washed with further ether (5 mL) and dissolved in MeOH (10 mL). This was filtered again, and the filtrate was collected. Solvent was removed under vacuum to give bicyclic intermediate **36** as an orange-yellow solid (215 mg, 0.59 mmol, quantitative yield).

TLC 20:1 $CHCl_3$:MeOH **RF**: 0.65 (**35**), 0.61 (**36**), 0.11 (**33**); NinH / I_2 staining; **36** exhibits blue-white fluorescence; **^{1}H NMR** (400MH_z, MeOD-d_3) δ (ppm) 7.08 (d, 1H, J = 8.1 Hz), 6.88 (d, 1H, J = 2.0 Hz), 6.58 (s, 1H), 6.49 (dd, 1H, J_{AM} = 2.4 Hz, J_{AX} = 7.9 Hz), 5.78 (s, 2H), 4.00 (m, 1H), 3.77

(m, 2H), 3.64 (m, 2H), 3.49 (dd, 1H, J_{AM} = 3.6 Hz, J_{AX} = 12.8 Hz), 3.34 (m, 1H), 3.11 (dd, 1H, J_{AM} = 8.0 Hz, J_{AX} = 12.8 Hz), 2.26 (s, 3H), 2.22 (s, 3H); **^{13}C NMR** δ (ppm) 136.8, 130.9, 129.9, 128.4, 124.8, 115.9, 111.7, 99.5, 78.1, 77.8, 77.5, 73.3, 72.9, 70.6, 63.2, 19.0, 17.8; **ESI MS (-)** (m/z) 364 ([M-1], 100%); **HR MS (ESI -)** found 364.1510, calc. 364.1509 for [M-1]

6.1.3.3 Attempted formation of intermediate 36 (in aqueous solution)

This method was originally described by W.T. Ashton *et al.*[109] in 1978, using a 1:1 mixture of deionised water and dioxane. This was repeated, and additionally the reaction was also performed using acetonitrile in place of dioxane. In both cases, this method was found to be unsuccessful in facilitating the coupling between 3,4-dimethyl-N(ribityl)-aniline **33** and 6-chlorouracil **35**, with a high proportion of starting materials recovered.

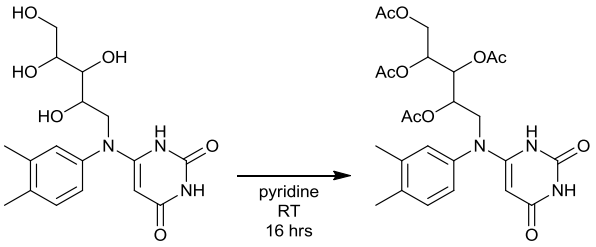


33 **35** **36**

A suspension of 3,4-dimethyl-N(ribityl)-aniline **33** (439 mg, 1.70 mmol, 2 equiv.) and 6-chlorouracil **35** (125 mg, 0.85 mmol) in an equivolume mixture of water and dioxane (40 mL) was heated at reflux for 36 hours, followed by stirring at room temperature for 12 hours. The solution was adjusted to pH 11 using aqueous NaOH (30 mL, 2M), and extracted using DCM (3 x 50 mL). Organic solvent was washed with brine (50 mL) and dried over $MgSO_4$, before rotary evaporation revealed no product isolation. The aqueous mixture was neutralised using HCl (15 mL, 1 M), before extraction using EtOAc (3 x 60 mL). Evaporation of the organic solvent and separation by FC allowed recovery of 3,4-dimethyl-N(ribityl)-aniline **33** in quantitative yield, and isolation of unreacted 6-chlorouracil **35** in yields up to 20 %.

TLC 10:1 $CHCl_3$:MeOH **RF**: 0.24 (**35**), 0.18 (**33**); NinH / I_2 staining; **FC stationary phase** Silica gel using 10 : 1 $CHCl_3$:MeOH, column bed 2.0 x 10 cm; **FC mobile phase** 10 : 1 $CHCl_3$:MeOH

6.1.4.1 Protection of intermediate 36 (original method)



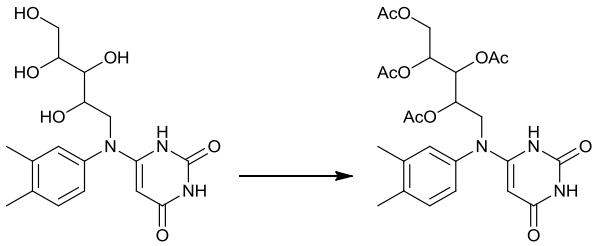
36 **37**

The crude solid from reaction 3.1.3.1 (2.98 g, impure) was dissolved in pyridine (25 mL), and acetic anhydride (2.80 mL, 29.4 mmol, 5-fold excess calculated from prior reaction) was added dropwise. The solution was stirred vigorously at room temperature overnight. The solvent was removed under reduced pressure, and the orange-brown residue dissolved in dichloromethane (120 mL). This was washed with water (3 x 20 mL) and brine (20 mL), and the organic layer was dried over $MgSO_4$. Solvent was removed using rotary evaporation to yield an orange oil, which was purified by FC. Product-containing fractions from FC were evaporated and the orange oil re-dissolved in CH_2Cl_2 (100 mL), washed with water (5 x 100 mL), saturated aqueous $NaHCO_3$ (2 x 50 mL) and brine (3 x 50 mL), before removal of the organic solvent under reduced pressure. The resultant orange-brown oil was dried under high vacuum for 2 days to yield bicyclic intermediate **37** as an orange-brown solid (1.51 g, 2.83 mmol, 47 % yield over 2 steps).

TLC 10:1 $CHCl_3$:MeOH **RF**: 0.37-0.28 (**37**), 0.19 (**36**); 20:1 $CHCl_3$:MeOH **RF**: 0.62 (**36**), 0.44-0.39 (**37**); NinH / I_2 staining; **36** and **37** exhibit blue-white fluorescence; **FC stationary phase** Silica gel using 50:1 $CHCl_3$:MeOH, column bed 2.5 x 15 cm; **FC mobile phase** Initially 50:1 $CHCl_3$:MeOH, before increasing gradient of MeOH (to 20:1); **^{1}H NMR** (400MH_z, $CDCl_3$) δ (ppm) 7.07 (d, 1H, *J* = 8.0 Hz), 6.85 (d, 1H, *J* = 2.4 Hz), 6.81 (dd, 1H, J_{AM} = 2.4 Hz, J_{AX} = 8.0 Hz), 5.24 (d, 2H, *J* = 6.4 Hz), 5.14 (m, 1H), 4.22 (dd, 1H, J_{AM} = 3.0 Hz, J_{AX} = 12.2 Hz), 4.04 (dd, 1H, J_{AM} = 6.2 Hz, J_{AX} = 12.2 Hz), 3.93 (m, 1H), 3.82 (dd, 1H, J_{AM} = 1.8 Hz, J_{AX} = 14.2 Hz), 3.36 (s, 1H), 2.19 (s, 6H), 1.96 (s, 3H), 1.95 (s, 3H), 1.84 (s, 3H), 1.73 (s, 3H); **^{13}C NMR** (62.9 MHz, DMSO-d_6) δ (ppm) 170.3, 169.8, 169.4, 169.0, 149.2, 140.6, 138.3, 136.3, 130.4, 128.3, 125.2, 70.3, 69.4, 69.1, 61.7, 48.2, 20.6, 20.4, 20.1, 20.0, 19.8, 19.1; **ESI MS (-)** (m/z) 531 ([M-2], 100 %), 490 ([M-1-Ac], 40 %)

6.1.4.2 Protection of intermediate 36 (alternative methods)

Numerous alternative methods for the effective formation of protected intermediate **37** were attempted, and are summarised in table 18 (*vide infra*). Each reaction was performed according to the general method given below.



36 **37**

Crude product from the previous reaction (performed at scales of 0.50 mmol) was dissolved in a mixture of CH_2Cl_2 (10 mL) and pyridine (8 µL, 0.10 mmol, 0.2 equiv.). Acetic anhydride (240 µL, 2.54 mmol; 5 equiv.) was added dropwise, and the mixture stirred at room temperature for 16 hours. The solution diluted using CH_2Cl_2 (25 mL), washed with water (3 x 20 mL), $NaHCO_3$ (20 mL) and brine (20 mL), before drying over $MgSO_4$ and removal of the solvent under reduced pressure. The resultant brown oil was dried under high vacuum to yield an orange-brown solid.

TLC 10:1 $CHCl_3$:MeOH **RF**: 0.36-0.26 (**37**), 0.17 (**36**); 20:1 $CHCl_3$:MeOH **RF**: 0.65 (**36**), 0.48-0.42 (**37**); NinH / I_2 staining; **36** and **37** exhibit blue-white fluorescence; **^{1}H NMR** (400MH_z, $CDCl_3$) δ (ppm) 7.07 (d, 1H, *J* = 8.0 Hz), 6.85 (d, 1H, *J* = 2.4 Hz), 6.81 (dd, 1H, J_{AM} = 2.4 Hz, J_{AX} = 8.0 Hz), 5.24 (d, 2H, *J* = 6.4 Hz), 5.14 (m, 1H), 4.22 (dd, 1H, J_{AM} = 3.0 Hz, J_{AX} = 12.2 Hz), 4.04 (dd, 1H, J_{AM} = 6.2 Hz, J_{AX} = 12.2 Hz), 3.93 (m, 1H), 3.82 (dd, 1H, J_{AM} = 1.8 Hz, J_{AX} = 14.2 Hz), 3.36 (s, 1H), 2.19 (s, 6H), 1.96 (s, 3H), 1.95 (s, 3H), 1.84 (s, 3H), 1.73 (s, 3H); **^{13}C NMR** (62.9 MHz, DMSO-d_6) δ (ppm) 170.5, 170.1, 169.8, 169.5, 149.6, 140.8, 138.1, 136.4, 130.6, 128.6, 124.9, 70.5, 69.8, 69.2, 61.8, 48.3, 20.7, 20.7, 20.6, 20.5, 19.7, 19.3; **ESI MS (+)** (m/z) 532 ([M], 100 %), 474 ([M - OAc], 40 %).

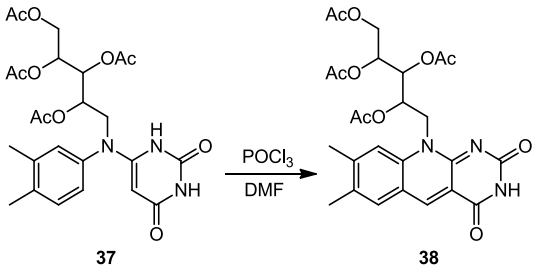
In cases where acetylation of the uracil moiety was observed, NMR spectra included the following additional resonances: ^{1}H NMR at 1.99 ppm (794.1 Hz), as a singlet with a relative integration of 3H, and ^{13}C NMR at 171.0 ppm and 22.5 ppm. MS gave a m/z value of 574 ([M], 25 %).

	Scale (mmol)	Bulk Solvent	Catalyst	Acetate source	Temp.	Reaction Time (hrs)	Yield (%)
1	6.0	Pyridine	-	Acetic Anhydride	RT	16	47
2	0.5	CH_2Cl_2	Pyridine	Acetic Anhydride	RT	16	55
3	0.5	CH_2Cl_2	-	Acetic Anhydride	RT	24 (incomplete)	15
4	0.5	CH_2Cl_2	TEA	Acetic Anhydride	RT	16	24
5	0.5	CH_2Cl_2	DMAP	Acetic Anhydride	RT	16	31
6	0.5	CH_2Cl_2	Imidazole	Acetic Anhydride	RT	16	41
7	4.5	CH_2Cl_2	Pyridine	Acetic Anhydride	RT	16	54
8	0.5	THF	Pyridine	Acetic Anhydride	RT	16	45
9	0.4	THF	TEA	Acetic Anhydride	RT	2	36
10	0.4	THF	TEA	Acetic Anhydride	RT	16	41
11	0.5	CH_2Cl_2	Pyridine	Acetyl chloride	RT	4	*0 **
12	0.5	CH_2Cl_2	Pyridine	Acetyl chloride	0 °C	4	*0 **
13	0.5	CH_2Cl_2	Pyridine	Acetyl chloride	-78 °C	4	*0 **

*Table 18: Reactions performed to establish the optimal method for acetylation of bicyclic intermediate **36**. "Scale" refers to the number of moles of riboaniline used for the initial formation of bicyclic **36**; all yields reported were obtained over two steps. When used, 0.2 equivalents of catalyst (calculated relative to riboaniline quantity) were used in all cases. 5 equivalents (1.25 molar excess) of acetic anhydride was used, while 4.1 equivalents (1.03 molar excess) of acetyl chloride were used.*

**- starting materials (or degradation products) only were isolated from these reactions*

6.1.5 Formation of 5-deazariboflavin tetraacetate 38

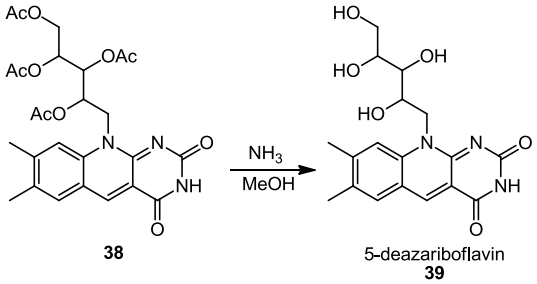


Acetylated intermediate **37** (900 mg, 1.69 mmol) was dissolved in DMF (6 mL), and $POCl_3$ (0.6 mL, 990 mg, 6.35 mmol) was cautiously added dropwise over 10 minutes. The solution was allowed to stir at room temperature for 1 hour, before heating to 100°C for 15 minutes. The mixture was allowed to cool to room temperature, and ice (30 cm^3) was added directly to the solution. A yellow-orange precipitate developed within the solution as it cooled. The pH of the mixture was adjusted to pH 7 using NH_4OH (5% aqueous solution, 10mL), and the precipitate isolated by filtration under suction, before drying under high vacuum. Further product was isolated from the mother liquor by extraction using $CHCl_3$ (3 x 20 mL). The product was isolated from the mixture by washing the organic solvent using a small volume of HCl (0.5 M, 5 mL) and washing with water (3 x 10 mL). The pH of the combined aqueous layers was adjusted to pH 9-10 using a saturated solution of $NaHCO_3$, and re-extracted using $CHCl_3$ (3 x 30 mL); combined organic layers were dried and concentrated under reduced pressure to yield additional 5-deazariboflavin tetraacetate **38**.

Spectroscopic data for both precipitated and extracted products were compared, and found to be identical. Thus, 5-deazariboflavin tetraacetate **38** was isolated in 72 % yield (665 mg, 1.22 mmol) as a bright yellow solid, with a characteristic blue fluorescence.

TLC 20:1 $CHCl_3$:MeOH **RF**: 0.58 (**38**), 0.38-0.32 (**37**), 0.19 (**36**); I_2 / CAM staining; **37** and **38** exhibit blue fluorescence; **^{1}H NMR** (400 MHz, $CDCl_3$) δ (ppm) 8.81 (s, 1 H), 7.61 (s, 1 H), 7.53 (s, 2 H), 5.39 (m, 2H), 5.30 (m, 1H), 4.40 (m, 1H), 4.28 (dd, 1H, J_{AM} = 5.8 Hz, J_{AX} = 12.0 Hz), 3.87 (dd, 1H, J_{AM} = 1.9 Hz, J_{AX} = 14.2 Hz), 2.51 (s, 3H), 2.41 (s, 3H), 2.26 (s, 3H), 2.16 (s, 3H), 2.05 (s, 3H), 2.02 (s, 3H); **^{13}C NMR** (62.9 MHz, $CDCl_3$) δ (ppm) 170.5, 170.1, 169.8, 169.5, 161.1, 158.1, 155.0, 138.1, 136.4, 130.6, 119.6, 116.1, 113.5, 70.3, 69.8, 69.1, 61.7, 21.4, 21.0, 20.6, 20.5, 19.6, 19.0; **ESI MS (+)** (m/z) 544 ([M+1], 100 %), 502 ([M+1, -OAc], 25 %)

6.1.6 Deprotection of 5-deazariboflavin tetraacetate 38



Crude 5-deazariboflavin tetraacetate **38** (250 mg, 0.460 mmol) was dissolved in methanolic ammonia (20 mL) and stirred at room temperature for 6 hours. The mixture was filtered (to remove a colourless precipitate), and the solvent was removed by rotary evaporation, yielding 5-deazariboflavin **39** (109 mg, 0.290 mmol, 63 %) as a brilliant yellow powder.

TLC 20:1 $CHCl_3$:MeOH **RF**: 0.58 (**38**), 0.0 (**39**); 3:1:1 isopropanol:AcOH:H_2O **RF**: 0.90 (**39**), 0.0 (**38**); I_2 / CAM staining; **38** and **39** exhibit blue fluorescence; **^{1}H NMR** (400 MHz, D_2O) δ (ppm) 8.80 (s, 1H), 7.92 (s, 1 H), 7.80 (s, 1H), 4.89 (m, 1 H), 4.63 (d, 1 H, J = 13.9 Hz), 4.23 (dd, 1 H, J_{AM} = 5.8 Hz, J_{AX} = 12.0 Hz), 3.70 (m, 4 H), 2.44 (s, 3 H), 2.34 (s, 3 H); **^{13}C NMR** (62.9 MHz, $CDCl_3$) δ (ppm) 161.9, 157.4, 156.3, 145.9, 141.0 139.9, 133.5, 130.4, 119.7, 117.8, 113.8, 73.4, 72.3, 69.7, 63.4, 46.1, 21.0, 19.0; **ESI MS (+)** (m/z) 398 ([M+Na$^+$], 55 %), 376 ([M+1], 100 %); **HR MS (ESI +)** (m/z) found 398.1325, calc. for the sodium adduct, $C_{18}H_{21}N_3O_6Na^+$ as 398.1323

6.2 Synthesis of 1-Deazariboflavin

This section describes the synthesis of 1-deazariboflavin from 1,2-diamino-4,5-dimethylbenzene. An overview of the method used is shown in figure 105 below.

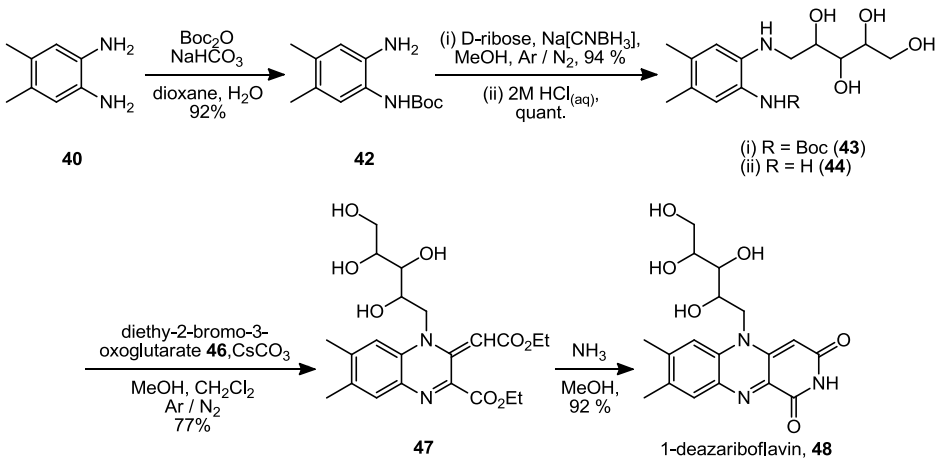
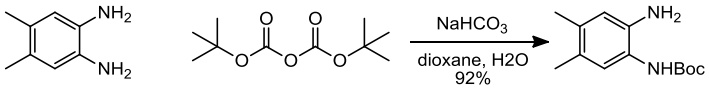


*Figure 105: Overall scheme for the synthesis of 1-deazariboflavin **48**.*

6.2.1 Boc- protection of 2-amino-4,5-dimethyphenylamine 40

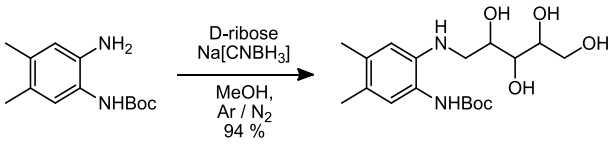


40 **41** **42**

2-Amino-4,5-dimethyphenylamine **40** (1.00 g, 7.30 mmol), di-*tert*-butyldicarbonate **41** (1.60 g, 7.30 mmol, 1.0 equiv.) and $NaHCO_3$ (0.62 g, 8.34 mmol, 1.15 equiv.) were dissolved in an equivolume mixture of dioxane and H_2O (total volume 100 mL). The mixture was stirred for three hours at room temperature, before dilution with additional H_2O (150 mL) and extraction into CH_2Cl_2 (3 x 50 mL). Combined organic layers were washed with saturated $NaHCO_3$ (50 mL), brine (50 mL), and dried over $MgSO_4$. Solvent was removed under reduced pressure and the product dried under high vacuum for 4 hours, yielding 2-amino(N-Boc)-4,5-dimethyphenylamine **42** as a red-orange solid (1.59 g, 6.73 mmol, 92%).

TLC 4:1 hexane:EtOAc **RF**: 0.60 (**41**), 0.21 (**42**), 0.0 (**40**); 10:1 $CHCl_3$:MeOH **RF**: 0.97 (**41**), 0.36 (**42**), 0.0 (**40**); NinH / CAM / I_2 staining; **MP** 153-155 °C; **^{1}H NMR** (400 MHz, $CDCl_3$) δ (ppm) 7.19 (s, 1H), 6.95 (s, 1H), 6.56(s, 1H), 6.12 (br s, 1H), 2.08 (s, 3H), 2.07 (s, 3H), 1.44 (s, 9H); **^{13}C NMR** (125.7 MHz) δ (ppm) 154.1, 134.4, 134.3, 132.3, 127.9, 122.3, 118.6, 28.4, 19.3, 18.9; **EI MS (+)** (m/z) 236 ([M], 100%)

6.2.2 Ribitylation of Boc-protected intermediate 42

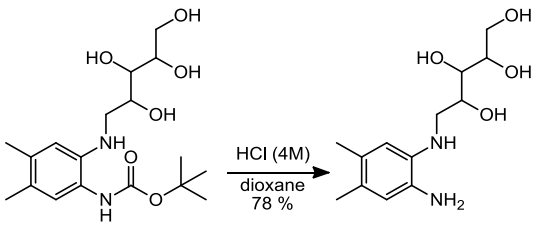


42 **43**

2-Amino(N-Boc)-4,5-dimethylphenylamine **42** (3.30 g, 14.0 mmol), D-ribose (5.25 g, 35 mmol, 2.5 equiv.) and $NaCNBH_3$ (1.57 g, 25 mmol, 1.8 equiv.) were dissolved in MeOH (150 mL) in a 2-neck RBF. The mixture was heated at reflux for 48 hours, and following cooling the solvent was removed under reduced pressure. The brown residue was swiftly dissolved in 1 M $HCl_{(aq)}$ (50 mL) and swirled until gas evolution ceased. This was immediately neutralised using a saturated solution of $NaHCO_3$ (75 mL) and extracted into ethyl acetate (5 x 75 mL). The combined organic layers were washed with brine (2 x 100 mL) and dried using Na_2SO_4. Solvent was removed, and the product was dried under high vacuum for 8 hours, to give ribitylated intermediate **43** as an orange-coloured solid (4.89 g, 13.20 mmol, 94%).

TLC 4:1 hexane :EtOAc **RF**: 0.14 (**43**), 0.21 (**42**); 10:1 $CHCl_3$:MeOH **RF**: 0.68 (**43**), 0.36 (**42**); NinH / CAM / I_2 staining; **MP** 110-112 °C; **^{1}H NMR** (500MH_z, MeOD-d_3) δ (ppm) 6.77 (s, 1H), 6.51 (s, 1H), 4.73, (H_2O), 3.85-3.83 (m, 1H), 3.69 (dd, 1H, J_{AM} = 3.5 Hz, J_{AX} = 11.0 Hz), 3.64 (m, 1H), 3.54 (dd, 2H, J_{AM} = 6.5 Hz, J_{AX} = 12.8 Hz), 3.33 (dd, 1H, J_{AM} = 3.0 Hz, J_{AX} = 12.8 Hz), 3.21 (m, 1H), 3.05 (dd, 1H, J_{AM} = 3.1 Hz, J_{AX} = 8.0 Hz), 2.09 (s, 3H), 2.03 (s, 3H), 1.40 (s, 9H); **^{13}C NMR** (100 MHz) δ (ppm) 153.1, 132.9, 132.3, 131.8, 126.4, 122.0, 118.6, 74.4, 73.1, 70.7, 63.4, 46.5, 28.2, 19.3, 18.9; **EI MS (+)** (m/z) 394 ([M+1+Na^+], 65%), 370 ([M+1], 100 %); **ESI MS (+)** (m/z) 394 ([M+1+Na^+], 80 %), 370 ([M+1], 100 %)

6.2.3.1 Deprotection of intermediate 43 (original method)

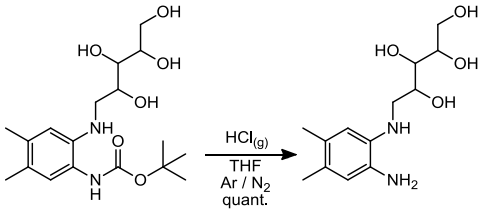


43 **44**

1-N(ribityl),2-N(Boc)-diamino-4,5-dimethylbenzene **43** (1.00 g, 2.70 mmol) was dissolved under anaerobic conditions in a solution of HCl (4 M) in dioxane (35 mL), and stirred at RT for 16 hours. The mixture was washed with Et_2O (2 x 30 mL), and the aqueous solvent removed under reduced pressure. The solid residue was dried under high vacuum for 8 hours, to yield deprotected intermediate **44** as a green solid (570 mg, 2.11 mmol, 78 %).

TLC 10:1 $CHCl_3$:MeOH **RF**: 0.68 (**43**), 0.00 (**44**); NinH / I_2 staining; **^{1}H NMR** (400MH_z, MeOD-d_3) δ (ppm) 6.79 (s, 1H), 6.52 (s, 1H), 4.76, (H_2O), 3.88-3.86 (m, 1H), 3.71 (dd, 1H, J_{AM} = 3.5 Hz, J_{AX} = 11.0 Hz), 3.67 (m, 1H), 3.55 (dd, 2H, J_{AM} = 6.5 Hz, J_{AX} = 12.8 Hz), 3.34 (dd, 1H, J_{AM} = 3.0 Hz, J_{AX} = 12.8 Hz), 3.24 (m, 1H), 3.06 (dd, 1H, J_{AM} = 3.1 Hz, J_{AX} = 8.0 Hz), 2.10 (s, 3H), 2.05 (s, 3H); **^{13}C NMR** (100 MHz, D_2O) δ (ppm) 131.9, 131.2, 126.4, 123.0, 115.6, 113.2, 74.8, 73.9, 72.4, 64.0, 46.7, 19.2, 18.7; **EI MS (+)** (m/z) 294 ([M+1+Na^+] 60%), 270 ([M], 100%)

6.2.3.2 Deprotection of intermediate 43 (using HCl gas)

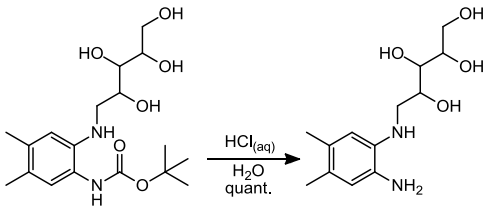


43 **44**

1-N(ribityl),2-N(Boc)-diamino-4,5-dimethylbenzene **43** (800 mg, 2.16 mmol) was dissolved in anhydrous THF (50 mL), and stirred at RT. A stream of HCl gas was bubbled directly through the reaction mixture for 30 minutes. Solvent was removed under reduced pressure, and the resultant green solid dried under high vacuum for 4 hours to yield the hydrochloride salt of deprotected intermediate **44** as a green solid (662 mg, 2.16 mmol, quantitative yield).

TLC 10:1 $CHCl_3$:MeOH **RF**: 0.66 (**43**), 0.00 (**44**); NinH / I_2 staining; **^{1}H NMR** (250MH_z, D_2O) δ (ppm) 7.36 (s, 1H), 7.27 (s, 1H), 4.09 (m, 1H), 3.79-3.68 (m, 4H), 3.35 (s, 1H), 3.31 (s, 1H), 2.34 (s, 3H), 2.32 (s, 3H); **^{13}C NMR** (D_2O) δ (ppm) 128.7, 124.4, 120.4, 96.5, 93.6, 79.2, 78.7, 78.2, 63.9, 61.1, 45.2, 27.4, 26.0 **ESI MS (+)** (m/z) 294 ([M+1+Na^+] 85 %), 270 ([M], 100 %); **ESI MS (-)** (m/z) 306 ([M+1+^{35}Cl], 100 %), 308 ([M+1+^{37}Cl], 30 %), 269 ([M-1], 40 %)

6.2.3.3 Deprotection of intermediate 43 (using aqueous HCl)

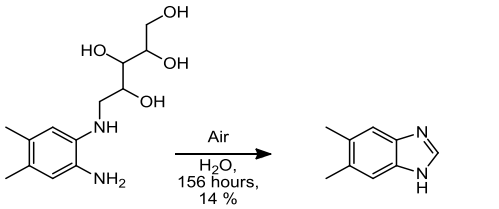


43 **44**

1-N(ribityl),2-N(Boc)-diamino-4,5-dimethylbenzene **43** (1.20 g, 3.24 mmol) was dissolved in aqueous HCl (2M) and stirred at RT for 3 hours. The mixture was lyophilised, furnishing deprotected intermediate **44** as green crystals (876 mg, 3.24 mmol, quantitative yield).

TLC 10:1 $CHCl_3$:MeOH **RF**: 0.67 (**43**), 0.00 (**44**); NinH / I_2 staining; **^{1}H NMR** (400MH_z, D_2O) δ (ppm) 7.07 (s, 1H), 6.76 (s, 1H), 3.92 (m, 1H), 3.83 (dd, 1H, J_{AM} = 3.5 Hz, J_{AX} = 11.2 Hz), 3.66 (m, 1H), 3.62 (dd, 2H, J_{AM} = 6.5 Hz, J_{AX} = 12.8 Hz), 3.47 (dd, 1H, J_{AM} = 3.0 Hz, J_{AX} = 12.8 Hz), 3.31 (m, 1H), 3.16 (dd, 1H, J_{AM} = 3.1 Hz, J_{AX} = 8.0 Hz), 2.14 (s, 3H), 2.08 (s, 3H); **^{13}C NMR** (100 MHz, D_2O) δ (ppm) 131.1, 130.4, 125.8, 124.0, 115.8, 113.2, 73.8, 72.9, 72.0, 64.7, 47.2, 19.0, 18.5; **EI MS (+)** (m/z) 294 ([M+1+Na^+] 30 %), 271 ([M+1], 100 %)

6.2.3.4 Formation of 5,6-dimethylbenzimidazole

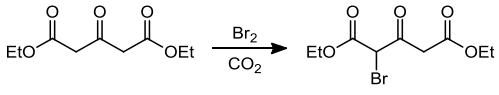


44 5,6-dimethylbenzimidazole, **44b**

Deprotected ribitylated diamine **44** (200 mg, 0.740 mmol) formed under anaerobic conditions was dissolved in H_2O (100 mL), and stirred vigorously at RT, with continuous aeration of the solution using compressed air. TLC at 156 hours indicated no reaction had occurred over the previous 24 hours (and was thus assumed to be complete). The mixture was extracted using EtOAc (3 x 30 mL), and the organic layer washed with saturated solutions of $NaHCO_3$ (50 mL) and brine (2 x 50 mL), before drying over $MgSO_4$. Solvent was removed by rotary evaporation to give the impure product as a crude brown solid. Recrystallization from Et_2O furnished 5,6-dimethylbenzimidazole **44b** (15 mg, 0.103 mmol, 14 %) as a pale yellow crystalline solid.

MP 203-205 °C (*lit 202-206 °C*); **^{1}H NMR** (400 MH_z, $CDCl_3$) δ (ppm) 8.00 (s, 1H), 7.44 (s, 2 H), 2.38 (s, 6H); **^{13}C NMR** (100 MHz), δ (ppm) 139.9, 136.3, 132.0, 115.6, 20.4 **ESI MS (+)** (m/z) 146 ([M], 100 %)

6.2.4.1 Formation of diethyl-2-bromo-3-oxoglutarate 46 (original method)

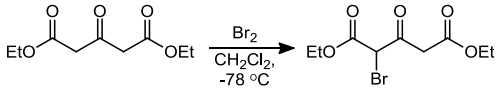


45 **46**

Diethyl-3-oxoglutarate **45** (1.60 mL, 1.79 g, 8.86 mmol) was stirred in a two-necked flask at 65 °C. CO_2 was bubbled through the solution, and bromine (650 µL, 2.00 g, 12.5 mmol, 1.4 equiv.) was added dropwise. The solution stirred at 65 °C for 30 minutes, before being cooled to RT, whereupon CH_2Cl_2 (60 mL) was added. The organic solvent was washed using a 10 % ($^w/_v$) solution of Na_2SO_3, followed by saturated solutions of Na_2SO_4 (25 mL), $NaHCO_3$ (25 mL) and brine (2 x 25 mL). The organic solvent was dried using $MgSO_4$ and removed under reduced pressure to yield a crude red oil. Yield was not calculated due to fast degradation of the product; diethyl-2-bromo-3-oxoglutarate **46** was synthesised immediately prior to use, with an assumption of 50 % efficiency.

^{1}H NMR (400MH$_z$, $CDCl_3$; impure due to degradation) δ (ppm) 5.05 (s, 1H), 4.21 (dq, 2H, J_{AM} = 3.1 Hz, J_{AX} = 7.2 Hz), 4.15-4.08, (m, 3H), 4.02 (q, 1H, *J* = 7.0 Hz), 3.77 (q, 1H, *J* = 15.6 Hz), 3.54 (s, 1H), 1.95 (s, 2H), 1.20 (m, 9H); **^{13}C NMR** (100MHz; impure due to degradation) δ (ppm) 191.9, 171.1, 167.0, 78.0, 77.4, 63.5, 61.9, 61.6, 60.5, 48.9, 45.8, 21.6, 14.4, 14.0; **APCI MS (+)** (m/z) 362 ([M($^{81}Br_2$)], 50%), 360 ([M($^{79}Br^{81}Br$)], 100%), 358 ([M($^{79}Br_2$)], 50%), 283 ($C_9H_{13}{}^{81}BrO_5$ [M(^{81}Br)+1], 100%), 281 ([M(^{79}Br)+1], 100%)

6.2.4.2 Formation of diethyl-2-bromo-3-oxoglutarate 46 (improved method)

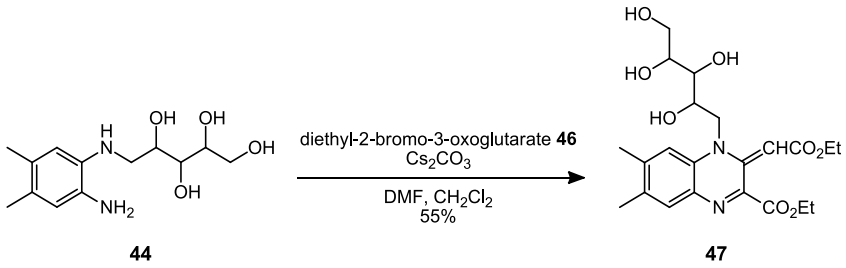


45 **46**

In a flame-dried three-neck-round bottom flask, diethyl-3-oxoglutarate **45** (2.90 mL, 3.50 g, 20 mmol) was dissolved in dried CH_2Cl_2 (60 mL) under N_2 at -78°C. Bromine (1.03 mL, 3.20 g, 20 mmol, 1 equiv.) was added dropwise over 10 minutes. After 45 minutes, the flask was allowed to warm to room temperature and the stoppers removed to release an acidic vapour (HBr). The solution was filtered under suction, and solvent removed under reduced pressure to furnish a red oil. This was purified by passing it over a short column of silica, allowing isolation of diethyl-2-bromo-3-oxoglutarate **46** as a rose-coloured oil (4.0 mL). This was synthesised immediately prior to use; due to fast degradation of the compound, yield calculation was found to be unreliable. Instead, sufficient reagent was used to supply the subsequent reaction with between 1.2 to 1.5 equivalents of the brominated oxoglutarate **46**, based on calculation from the quantity of **45** initially used.

TLC 3:2 hexane:EtOAc **RF**: 0.99 (**45**), 0.55 (**46**); NinH / I_2 staining; **FC stationary phase** Silica gel using 3:2 hexane:EtOAc, column bed 2.5 x 2.5 cm; **FC mobile phase** 3:2 hexane:EtOAc; **^{1}H NMR** (400MH$_z$, $CDCl_3$) δ (ppm) 5.05 (s, 1H), 4.14 (dq, 4H, *J* = 7.2 Hz), 3.54, (s, 1H), 1.95 (s, 1H), 1.20 (m, 6H); **^{13}C NMR** (100MHz) δ (ppm) 198.6, 168.6, 165.0, 63.1, 61.4, 60.5, 45.8, 14.4, 14.2, 14.0; **ESI MS (-)** (m/z) 280 ([M(^{81}Br)-2], 100%), 278 ([M(^{79}Br)-2], 100%)

6.2.5.1 Coupling of aniline 44 and diethyl-2-bromo-3-oxoglutarate 46 (original method)

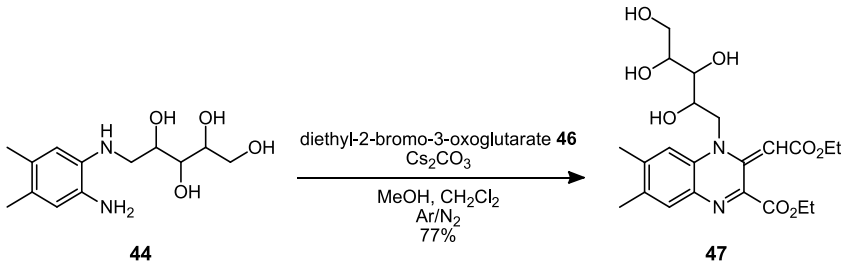


1-(N(ribityl)),2-diamino-4,5-dimethylbenzene **44** (280 mg, 1.04 mmol) dissolved in DMF (10 mL) was mixed with a solution of diethyl-2-bromo-3-oxoglutarate **46** (600 mg, 2.1 mmol, 2 equiv.) in CH_2Cl_2 (15 mL). Cs_2CO_3 (1.00 g, 3.1 mmol, 3 equiv.) was added, and the biphasic mixture stirred at RT for 36 hours, and filtered. The solution was diluted with H_2O (60 mL) and extracted into ethyl acetate (5 x 50 mL), with addition of brine as necessary to facilitate separation of the layers. The combined organic layers were washed with further brine (2 x 50 mL), and dried over $MgSO_4$. Solvent was removed under reduced pressure to yield a dark-red viscous oil. After purification by FC over silica, the product was dried under high vacuum for 8 hours to give the bicyclic intermediate **47** (257 mg, 0.57 mmol, 55 %) as a dark red oil.

The reaction was repeated using several alternative bases (in appropriate relative quantities), including K_2CO_3 (giving 34 % yield), Na_2CO_3 (14 % yield), $NaHCO_3$ (11 % yield), pyridine (4 % yield), imidazole, DMAP, NaOH and KOH. Product purity was assessed compared to the spectroscopic data given below. No product was isolated using imidazole, DMAP or the hydroxylic bases, and no base examined was more effective than Cs_2CO_3.

TLC 20 : 1 $CHCl_3$:MeOH **RF**: 0.49 (**46**), 0.07 (**47**), 0.00 (**44**); NinH / CAM / I_2 staining; **FC stationary phase** Silica gel using 20 : 1 $CHCl_3$:MeOH, column bed 2.5 x 10 cm; **FC mobile phase** 20 : 1 $CHCl_3$:MeOH; **^{1}H NMR** (250MH_z, $CDCl_3$) δ (ppm) 7.94 (s, 1H), 7.26, (s, 1H) 7.10 (br s, 1H), 5.24 (s, 1H), 4.60 (m, 3H), 4.40 (q, 1H, *J* = 12Hz), 4.30 (m, 1H), 3.80 (m, 1H), 2.89 (s, 3H), 2.81 (s, 3H), 1.25 (m, 3H); **^{13}C NMR** (62.5 MHz) δ (ppm) 162.5, 162.3, 148.3, 132.8, 132.5, 131.6, 130.7, 123.1, 106.5, 77.7, 77.1, 76.6, 73.2, 62.9, 60.8, 40.7, 36.5, 31.4, 29.6, 15.0, 14.4; **ESI MS (+)** (m/z) 473 ([M+Na^+], 65 %) 451 ([M+1], 100 %); **HR MS (ESI +)** (m/z) found 451.2336, calc. 451.2080 for [M]

6.2.5.2 Coupling of aniline 44 and diethyl-2-bromo-3-oxoglutarate 46 (enhanced method, and survey of solvents)



44 **47**

1-(N(ribityl)),2-diamino-4,5-dimethylbenzene **44** (3.10 g, 11.5 mmol) was added to a stirring biphasic mixture of MeOH (30 mL) and CH_2Cl_2 (30 mL), containing diethyl-2-bromo-3-oxoglutarate **47** (4.0 mL) and Cs_2CO_3 (2.40 g, 7.4 mmol, 0.64 equiv.). The mixture was stirred at RT for 48 hours, followed by filtration under suction, before solvent was removed under reduced pressure. The black residue was dissolved in H_2O (100 mL) and extracted into ethyl acetate (5 x 50 mL), with brine added to facilitate separation of the layers. The combined organic layers were washed with further brine (2 x 50 mL), and dried over $MgSO_4$. Solvent was removed under reduced pressure to furnish a dark-red viscous oil, which was purified by FC. Solvent was removed under reduced pressure, and the final product dried under high vacuum for 8 hours to give the final product **47** (4.00 g, 8.89 mmol, 77%) as a dark red oil.

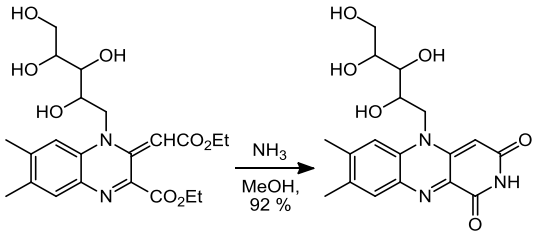
TLC 20 : 1 $CHCl_3$:MeOH **RF**: 0.50 (**46**), 0.08 (**47**), 0.00 (**44**); NinH / CAM / I_2 staining; **FC stationary phase** Silica gel using 20 : 1 $CHCl_3$:MeOH, column bed 2.5 x 10 cm; **FC mobile phase** 20 : 1 $CHCl_3$:MeOH; **^{1}H NMR** (400 MH_z, $CDCl_3$) δ (ppm) 7.36, (s, 1H) 6.49 (s, 1H), 6.39 (s, 1H), 5.24 (s, 1H), 4.65-4.52 (m, 4H), 4.37 (q, 1H, *J* = 7.0 Hz), 4.22 (m, 1H), 3.78 (m, 1H), 2.33 (s, 3H), 2.26 (s, 3H), 1.30 (m, 3H); **^{13}C NMR** (62.5 MHz) δ (ppm) 168.3, 162.5, 162.3, 148.3, 132.8, 132.5, 131.6, 130.7, 123.1, 106.5, 77.7, 77.1, 76.6, 73.2, 62.9, 60.8, 40.7, 36.5, 31.4, 29.6, 15.0, 14.4; **ESI MS (+)** (m/z) 473 ([M+Na^+], 45 %), 451 ([M+1], 100 %); **HR MS (ESI +)** (m/z) found 451.2106, calc. 451.2080 for [M+1]

This reaction was repeated using the alternative solvent mixtures given in table 19 below. Each reaction was performed using 250 mg (925 μM) of the determinant reactant. All products were purified by FC, and spectral data compared to that given above.

Primary Solvent	Secondary Solvent	Yield		
		mass (mg)	µmol	%
CH_2Cl_2	DMF			55
CH_2Cl_2	Hexane	-	-	0
CH_2Cl_2	EtOH	184	408	44
CH_2Cl_2	MeOH	322	715	77
CH_2Cl_2	H_2O	34	75	8
CH_2Cl_2	DMSO	-	-	0
$CHCl_3$	MeOH	270	599	65
$CHCl_3$	DMF	162	360	39
Et_2O	MeOH	68	151	16
Et_2O	H_2O	45	100	11

*Table 19: Solvent mixtures used to examine the coupling between intermediates **44** and **46***

6.2.6 Formation of 1-deazariboflavin 48



47 1-deazariboflavin **48**

Bicyclic intermediate **47** (120 mg, 267 µmol) was dissolved at room temperature in methanolic ammonia (prepared at -78 °C. The solution was stirred at RT for 72 hours, with additional solvent (10 mL) added after 36 hours. Solvent was removed under reduced pressure and the resultant purple-brown solid dried under high vacuum for 8 hours to yield 1-deazariboflavin **48** as a purple-black solid (92 mg, 245 µmol, 92%).

TLC 3:1:1 isopropanol:AcOH:H_2O **RF**: 0.87 (**48**), 0.00 (**47**); **47** exhibits some white fluorescence, while **48** exhibits a total lack of fluorescence; **^{1}H NMR** (400 MH_z, D_2O) δ (ppm) 7.48, (s, 1H) 7.15 (s, 1H), 5.56 (s, 1H), 3.94 (m, 1H) 3.56 (m, 2H), 3.38 (m, 1H), 2.88 (q, 1H, *J* = 7.6 Hz), 2.79 (m, 2H), 2.11 (s, 3H), 2.02 (s, 3H); **^{13}C NMR** (62.5 MHz, MeOD) δ (ppm) 167.3, 161.1, 148.0, 146.7, 141.1, 136.0, 132.8, 131.3, 130.1, 122.9, 107.3, 76.0, 71.4, 70.2, 65.3, 53.8, 21.5, 18.1; **ESI MS (-)** (m/z) 412 ([M+^{37}Cl], 15 %), 410 ([M+^{35}Cl], 50 %), 374 ([M-1], 100 %)

6.3 Synthesis of 8-Demethyl-5-deazariboflavin 54

This section describes the synthesis of the 5-deazariboflavin derivative, 8-demethyl-5-deazariboflavin **54**. This was performed using the optimal route found for the synthesis of 5-deazariboflavin **39**, using 4-methylaniline **49** in place of 3,4-dimethylaniline **31**.

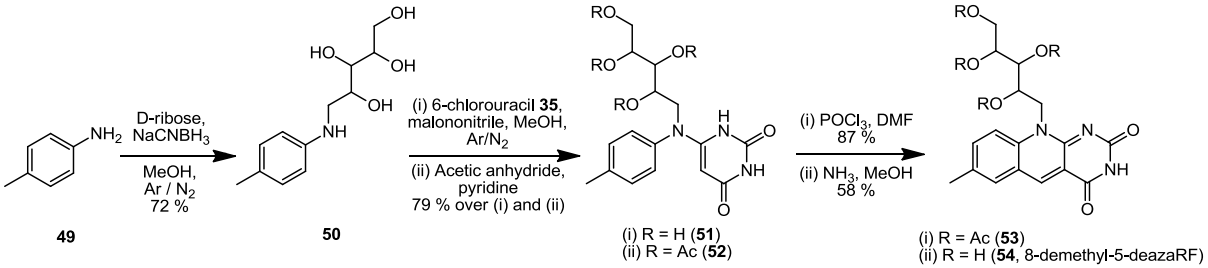
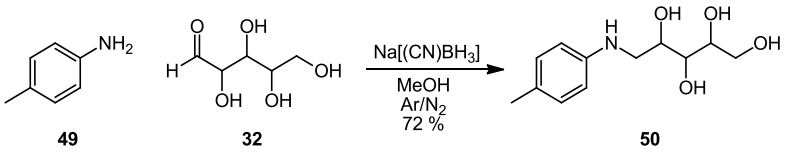


*Figure 106: Overview of the synthesis of 8-demethyl-5-deazariboflavin **54***

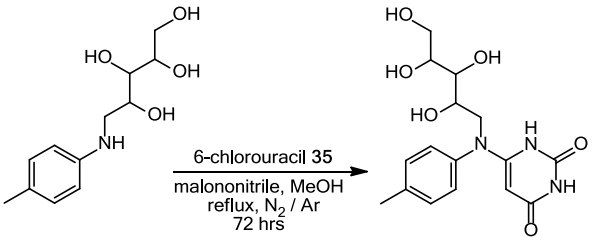
6.3.1 Reductive amination of 4-methylaniline 49 using D-ribose



p-Toluidine **49** (5.00 g, 47 mmol), Na[(CN)BH_3] (5.84 g, 93 mmol, 2 equiv.) and D-ribose **32** (17.5 g, 117 mmol, 2.5 equiv.) were combined in a flame-dried 3-neck RBF and dried under high vacuum for 4 hours. Anhydrous MeOH (250mL) was added, and the mixture heated at reflux under Ar / N_2 until completion. The solvent was removed under reduced pressure to give a dark brown solid, which was dissolved in $HCl_{(aq)}$ (50mL, 1M) and swirled until gas evolution ceased. The resulting suspension was carefully neutralised using saturated $NaHCO_3$ (100 mL) and washed with an aliquot of EtOAc (50 mL). As with the earlier formation of 3,4-dimethyl-N-(ribityl)-aniline (*vide supra*), EtOAc stimulated formation of a white precipitate which was collected by filtration and dried under high vacuum to furnish (N-ribityl)-*p*-toluidine (2.00 g, 8.29 mmol) as a white solid. The aqueous layer was extracted with further EtOAc (4 x 50 mL), and combined organic fractions washed with brine (2 x 50mL) and dried over $MgSO_4$. Solvent was removed under reduced pressure to give an impure pale yellow solid, which was purified by recrystallisation to give product **50** (6.11 g, 25.32 mmol) as a white solid. The combined yield of purified (N-ribityl)-*p*-toluidine from both precipitate and organic extraction (8.11 g, 33.6 mmol) was therefore 72%.

TLC 10:1 $CHCl_3$:MeOH **RF**: 0.78 (**49**), 0.57 (**50**), 0.00 (**32**); NinH staining; **MP** 138-141 °C; **^{1}H NMR** (400 MHz, MeOD-d_3) δ (ppm) 6.83 (d, 2H, *J* = 7.6 Hz), 6.52 (d, 2H, *J* = 8.5 Hz), 4.81 (br s (hydroxylic), 4H), 3.81 (m, 1H), 3.67 (m, 2H), 3.54 (dd, 2H, J_{AM} = 6.2 Hz, J_{AX} = 9.8 Hz), 3.33 (dd, 1H, J_{AM} = 3.2 Hz, J_{AX} = 12.8 Hz), 3.21 (m, 2H), 2.99 (dd, 1H, J_{AM} = 8.0 Hz, J_{AX} = 12.8 Hz), 2.09 (s, 3H); **^{13}C NMR** (62.5 MHz, MeOD-d_3) δ (ppm) 147.8, 130.5, 127.8, 115.0, 74.9, 74.4, 72.5, 64.7, 42.4, 20.6; **ESI MS (+)** (m/z) 264 ([M+Na^+], 100 %) 242 ([M+1], 100 %); **HR MS (EI +)** found 241.1313, calculated as 241.1314 [M]

6.3.2 Coupling of ribitylated aniline 50 and 6-chlorouracil 35

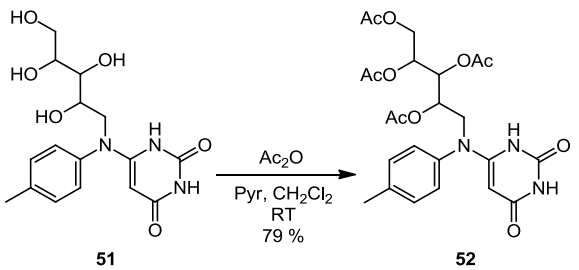


50 **51**

Ribitylated aniline **50** (280 mg, 1.16 mmol), 6-chlorouracil **35** (203 mg, 1.39 mmol, 1.2 equiv.) and malononitrile (50 mg, 0.76 mmol, 0.66 equiv.) were combined in flame-dried glassware. Thoroughly dried MeOH (30 mL) was added, and the suspension heated under N_2 to reflux. After 48 hours, further malononitrile (50mg, 0.76 mmol, 0.66 equiv.) was added. TLC at 96 hours indicated complete consumption of starting materials, and the reaction was cooled to room temperature. Solvent was removed under reduced pressure, and the resultant brown solid was used for the subsequent reaction without further purification.

TLC 20:1 $CHCl_3$:MeOH **RF**: 0.65 (**35**), 0.12 (**51**), 0.21 (**50**); NinH / I_2 staining; **51** exhibits blue fluorescence; **ESI MS (-)** (m/z) 351 ([M], 100%)

6.3.3 Protection of ribityl side-chain of intermediate 51



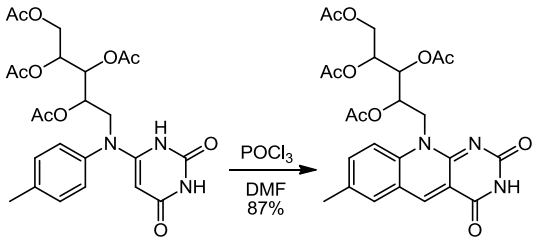
51 **52**

Crude intermediate **51** from the previous reaction was dissolved in a solution of pyridine (2.5 mL) and CH_2Cl_2 (10 mL). Acetic anhydride (0.65 mL, 6.89 mmol, 6-fold excess) was added

dropwise, and the mixture was stirred at room temperature for 16 hours. Solvent was removed under reduced pressure, and the orange-brown residue dissolved in water (20 mL) before extraction into CH_2Cl_2 (3 x 50 mL). Combined organic layers were washed with a saturated solution of $NaHCO_3$ (50 mL), water (3 x 50 mL) and brine (2 x 50 mL), and dried over $MgSO_4$. Solvent was removed under reduced pressure, and the resultant oil thoroughly dried under reduced pressure to furnished acetylated product **52** as a yellow-orange oil in 79 % yield (477 mg, 0.92 mmol) over two steps.

TLC 20:1 $CHCl_3$:MeOH **RF**: 0.34-0.42 [streak] (**52**), 0.12 (**51**); NinH / I_2 staining; **51** and **52** each exhibit blue fluorescence; **^{1}H NMR** (400 MHz, $CDCl_3$) δ (ppm), 7.19 (d, 2H, *J* = 8.2 Hz), 7.03 (d, 2H *J* = 8.2 Hz), 5.20 (m, 1H), 4.28 (dd, 1H, J_{AM} = 3.2 Hz, J_{AX} = 12.4 Hz), 4.23-4.15 (m, 2H), 4.12-4.03 (m, 2H), 3.85 (dd, 1H, J_{AM} = 2.2 Hz, J_{AX} = 14.6 Hz), 3.65 (m, 1H), 2.31 (s, 3H), 2.05 (s, 3H), 2.02 (s, 3H), 2.01 (s, 3H), 1.90 (s, 3H); **^{13}C NMR** (62.5 MHz, $CDCl_3$) δ (ppm) 177.2 (C), 171.8 (C), 169.9 (C), 166.6 (C), 164.8 (C), 160.3 (C), 144.0 (C), 130.4 (CH), 126.9 (CH), 126.4 (CH), 73.3 (CH), 72.5 (CH), 68.9 (CH), 63.8 (CH), 61.4 (CH_2), 42.8 (CH_2), 36.3 (CH_3), 19.9 (CH_3); **EI MS (+)** (m/z) molecular ion ($C_{24}H_{29}N_3O_{10}$, 519.19) was unobserved; characteristic fragmentation indicated product formation: 475 ([M+1-Ac], 100 %), 432 ([M-(Ac)$_2$], 35 %)

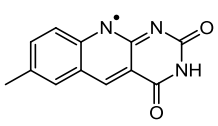
6.3.4 Formation of 8-demethyl-5-deazariboflavin tetraacetate 53



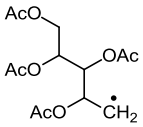
Bicyclic acetylated intermediate **52** (400mg, 0.77 mmol) was dissolved in DMF (5 mL) with stirring. $POCl_3$ (86 µL, 142 mg, 0.92 mmol, 1.2 equiv.) was added dropwise over 10 minutes, and the mixture stirred at room temperature for 30 minutes, before heating to 100°C for 60 minutes. Once cooled to room temperature, ice (20 cm^3) was added directly to the mixture, facilitating precipitation of a yellow-orange precipitate. The pH of the mixture was adjusted to pH 7 using NH_4OH (5% aqueous solution, 15 mL), and the precipitate isolated by filtration under suction, before drying under high vacuum. The aqueous layer was extracted using CH_2Cl_2 (3 x 20 mL), and the combined organic fractions washed using HCl (0.5 M, 10 mL)

followed by water (3 x 10 mL). The pH of the aqueous solution was adjusted to pH 9-10 using a saturated solution of $NaHCO_3$, and re-extracted using $CHCl_3$ (3 x 20 mL). Combined organic layers were dried and solvent removed under reduced pressure to yield additional product **53** as a vibrant yellow solid, with a combined yield of 87 % (356 mg, 0.67 mmol).

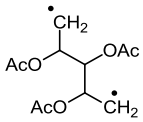
TLC 20:1 $CHCl_3$:MeOH **RF**: 0.38-0.42 [streak] (**52**), 0.31 (**53**); NinH / I_2 staining; **52** and **53** each exhibit blue fluorescence, **53** is particularly vivid **^{1}H NMR** (500 MHz, $CDCl_3$) δ (ppm) 8.01 (s, 1H), 7.19 (d, 1H, *J* = 6.4 Hz), 7.13 (dd, 1H, J_{AM} = 1.6 Hz, J_{AX} = 6.8 Hz), 5.30 (m, 1H), 5.20 (m, 1H), 4.28 (dd, 1H, J_{AM} = 2.4 Hz, J_{AX} = 10.0 Hz), 4.17-4.05 (m, 2H), 3.75 (dd, 1H, J_{AM} = 11.4 Hz, J_{AX} = 80.4 Hz), 2.95 (DMSO, residual), 2.36 (s, 3H), 2.11-2.02 (m, 12H) **^{13}C NMR** (125 MHz, $CDCl_3$) δ (ppm), 177.3 (C), 172.2 (C), 169.2 (C), 164.2 (C), 162.5 (C), 157.2 (C), 151.0 (C), 138.0 (C), 130.4 (CH), 128.0 (CH), 127.7, 122.8 (CH), 119.6 (CH), 114.0, 70.7 (CH), 69.8 (CH), 69.2 (CH), 61.9 (CH_2), 48.4 (CH_2), 36.6 (CH_3), 21.1 (CH_3), 20.9 (CH_3), 20.8 (CH_3), 20.8 (CH_3); **ESI MS (-)** (m/z) molecular ion ($C_{24}H_{29}N_3O_{10}$, 519.19) unobserved; fragmentation indicated product formation: 485 ([M-1, -Ac], 100 %); **EI MS (+)** (m/z) molecular ion **53** was unobserved; fragmentation indicated product formation (predicted structures below): 303 (**53c** [M], 55 %), 243 (**53d** [M], 15 %) 229 (**53b** [M], 60 %),



53b

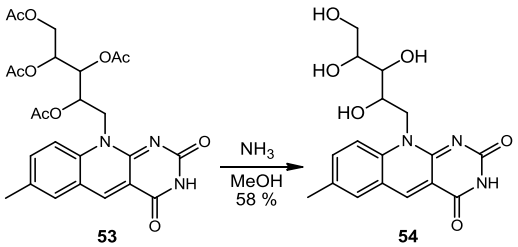


53c



53d

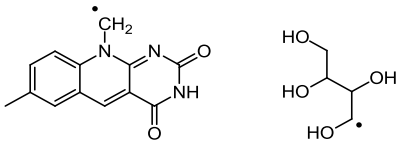
6.3.5 Formation of 8-demethyl-5-deazariboflavin 54



53 **54**

Crude acetylated intermediate **53** (300 mg, 0.567 mmol) was dissolved in methanolic ammonia (20 mL) and stirred at room temperature for 6 hours. The mixture was filtered, and the clarified solution concentrated by rotary evaporation, yielding crude 8-demethyl-5-deazariboflavin as a yellow oil. Purified 8-demethyl-5-deazariboflavin **54** was isolated as a brilliant yellow powder in 58 % yield (119 mg, 0.329 mmol) by HPLC.

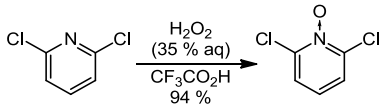
TLC 20:1 $CHCl_3$:MeOH **RF**: 0.31 (**53**), 0.00 (**54**); 3:1:1 isopropanol:AcOH:H_2O **RF**: 0.86 (**54**), 0.00 (**53**); I_2 / CAM staining; **53** and **54** each exhibit vivid blue fluorescence; **^{1}H NMR** (500 MHz, DMSO-d_6) δ (ppm) 8.01 (s, 1H), 7.25 (d, 1H, *J* = 7.2 Hz), 6.88 (d, 1H, *J* = 6.4 Hz), 6.71 (br s, 1H), 6.51 (d, 1H, *J* = 6.8 Hz), 3.80-3.75 (m, 1H), 3.63-3.56 (m, 2H), 3.49-3.39 (m, 3H), 3.29 (d, 1H), 2.36 (s, 3H); **^{13}C NMR (**125 MHz, DMSO-d_6) δ (ppm), 162.4, 155.8, 153.3, 148.5, 130.6 (CH), 129.7 (CH), 128.5, 123.1, 121.8, 115.2 (CH), 112.8, 107.4 (CH), 73.9 (CH), 71.8 (CH), 70.6 (CH), 63.6 (CH_2), 46.4 (CH_2), 23.0 (CH_3); **ESI MS (-)** (m/z) 356 ([M-4], 100 %); **ESI MS (+)** (m/z) fragments indicated product formation; 240 (**54b** [M], 100 %), 121 (**54c** [M], 55 %)



54b **54c**

6.4 Attempted Formation of Multiple Riboflavin Analogues Using a Single Synthetic Method

6.4.1.1 Synthesis of 2,6-dichloropyridine-1-oxide 56

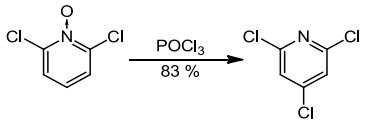


55 **56**

2,6-Dichloropyridine **55** (11.2 g, 75.7 mmol) was dissolved in a mixture of CF_3CO_2H (50 mL) and 35 % aqueous [v/v] H_2O_2 (14 mL), and heated at reflux for 16 hours. After cooling to RT, the mixture was basified (pH 14) using saturated aqueous $NaHCO_3$ (50 mL), followed by NaOH (4.5 g), extracted using EtOAc (3 x 60 mL), and the combined organic extracts washed with saturated $NaHCO_3$ (50 mL), brine (50 mL) and dried over Na_2SO_4. Solvent was removed under reduced pressure to give crude yellow crystals (97 %), which were purified by recrystallisation from EtOAc to give pure product **56** as colourless crystals (11.67 g, 71.2 mmol, 94 %).

TLC 20:1 $CHCl_3$:MeOH **RF:** 0.65-0.71 (**56**)**;** 10:1 $CHCl_3$:MeOH **RF:** 0.74-0.85 (**56**); **MP** 136-138 °C (*lit 135-139 °C*); **^{1}H NMR** (400MH_z, $CDCl_3$) δ (ppm) 7.57 (t, 1H, *J* = 7.8 Hz), 7.20 (d, 2H, *J* = 7.9 Hz); **^{13}C NMR** (62.5 MH_z, $CDCl_3$) δ (ppm) 149.5, 139.9, 121.9; **ESI MS (+)** (m/z) 186 ([M + Na^+], 100 %), 164 ([M+1], 65 %); **HR MS (ESI +)** (m/z) found 185.9486, calc. 185.9489 for [M+Na^+]; found 162.9675, calc. 162.9670 for [M]

6.4.1.2 Synthesis of 2,4,6-trichloropyridine 57

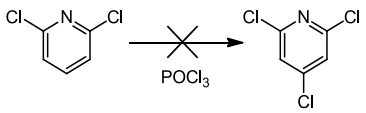


56 **57**

2,6-Dichloropyridine-N(oxide) **56** (9.00 g, 54.9 mmol) was dissolved in undiluted $POCl_3$ (10 mL) and heated at reflux for 16 hours. The mixture was allowed to cool to 0 °C, and ice (35 cm^3) was cautiously added over the course of 2 hours. The pH of the solution was adjusted using a solution of Na_2CO_3 (0.5 M) until pH 7, followed by NaOH (2M) to pH 14. The solution was extracted into Et_2O (3 x 80 mL), and the combined organic layers washed with $NaHCO_3$ (50 mL) and brine (50 mL) before drying over $MgSO_4$. Solvent was removed under reduced pressure (at 40 °C) to give a green-brown oil, which crystallised upon cooling to RT to furnish green-grey needles (10.01 g, 54.9 mmol, quant.). 2,4,6-trichloropyridine **57** was recrystallised from pentane and ice-cold ethanol as large grey-white needles (9.02 g, 49.4 mmol, 83 %).

TLC 20:1 $CHCl_3$:MeOH **RF:** 0.98 (**57**) 0.66-0.71 (**56**)**;** 10:1 $CHCl_3$:MeOH **RF:** 0.94 (**57**), 0.74-0.85 (**56**); **MP** 33-34 °C (*lit 33-37 °C*); **^{1}H NMR** (400MH_z, $CDCl_3$) δ (ppm) 8.07 (s, 1H); **^{13}C NMR** (62.5 MH_z, $CDCl_3$) δ (ppm) 150.6, 147.2, 123.1; **EI MS (+)** (m/z) 185 ([M($^{37}Cl_2$)], 25 %), 183 ([M($^{37}Cl^{35}Cl$)], 95 %), 181 ([M($^{35}Cl_2$)], 100 %)

6.4.1.3 Attempted direct formation of 2,4,6-trichloropyridine 57 (unsuccessful)

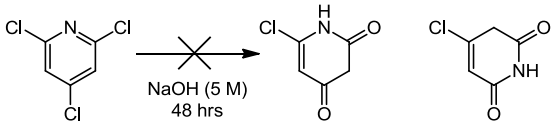


55 **57**

2,6-dichloropyridine **55** (1.48 g, 10.0 mmol) was dissolved in undiluted $POCl_3$ (4 mL), and the solution heated at reflux for 16 hours. The mixture was allowed to cool to 0 °C, and ice (10 cm^3) cautiously added over the course of 1 hour. The acidity of the solution was adjusted to pH 7 using Na_2CO_3 (10 mL, 0.5 M), followed by NaOH (2M) to adjust the pH to 14, before extraction using diethyl ether (3 x 50 mL); no product was isolated in this solvent. The aqueous layer was re-extracted using EtOAc (3 x 50 mL), and the combined organic layers washed with $NaHCO_3$ (50 mL) and brine before drying over $MgSO_4$. Solvent was removed under reduced pressure to yield a white crystalline solid. Spectroscopy indicated recovery of 2,6-dichloropyridine **55** (1.46 g, 9.86 mmol, 99%).

TLC 20:1 $CHCl_3$:MeOH **RF:** 0.82 (**55**); **MP** 84-86 °C (*lit 83-86* °C); **^{1}H NMR** (400MH_z, $CDCl_3$) δ (ppm) 7.63 (t, 1H, *J* = 7.9 Hz), 7.28 (d, 2H, *J* = 7.6 Hz); **EI MS (-)** (m/z) 146 ([M-1], 100 %)

6.4.2 Attempted conversion of 2,4,6-trichloropyridine 57 to pyridione 58 (unsuccessful)



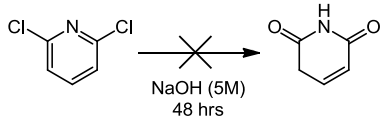
57 **58a** **58b**

2,4,6-trichlropyridine **57** (500 mg, 2.74 mmol) was stirred in a refluxing aqueous solution of NaOH (10 mL, 5 M, 50 mmol) for 48 hours. The reaction was allowed to cool to RT, and the pH of the mixture adjusted to pH 2 using aqueous HCl (2 M). Acidification of the mixture facilitated precipitation of an off-white solid, which was collected by filtration under suction, washed with ice-cold water (5 mL), and recrystallised from acetone to furnish a white crystalline solid (138 mg, *0.85 mmol*, 31 %* [*based on later MS evidence]). The aqueous filtrate was extracted thoroughly using EtOAc (5 x 30 mL), dried using brine (30 mL) and dried over $MgSO_4$. Removal of solvent under reduced pressure afforded recovery of 2,4,6-trichloropyridine **57** (305 mg, 1.81 mmol, 66 %).

Spectroscopy was performed for the unknown (minor) product:

TLC 20:1 $CHCl_3$:MeOH **RF:** 0.98 (**57**) 0.14, 0.17 (unidentified products); 10:1 $CHCl_3$:MeOH **RF:** 0.94 (**57**), 0.34, 0.36 (unidentified products); **MP** 224-226 °C; **^{1}H NMR** (400 MHz, $CDCl_3$) δ (ppm) 6.76 (d, 1H, *J* = 1.6 Hz), 6.69 (d, 1H, *J* = 1.6 Hz), 1.66 (br s); **^{13}C NMR** δ (ppm) 205.0, 145.4, 135.2, 114.0; **EI MS (+)** (m/z) isotope pattern strongly indicated a dichlorinated product: 167 ([M($^{37}Cl_2$)], 10 %), 165 ([M($^{37}Cl^{35}Cl$), 65 %], 163 ([M($^{35}Cl_2$)], 100 %)

6.4.3 Attempted conversion of 2,4-dichloropyridine 55 to 2,6-pyridione 59 (unsuccessful)



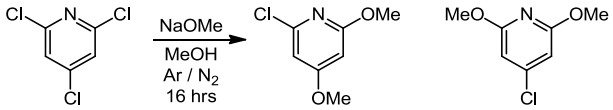
55 **59**

A solution of 2,6-dichloropyridine **55** (1.00 g, 6.76 mmol) in refluxing aqueous NaOH (6 mL, 5 M, 4.5 equiv.) was stirred for 48 hours. The reaction was allowed to cool to RT, the pH of the mixture was adjusted to between pH 1-2 using aqueous HCl (25 mL, 1 M), and the mixture

cooled at 4°C overnight, with no precipitation observed. The aqueous mixture was extracted using EtOAc (3 x 20 mL). The organic layer was washed with a saturated solution of $NaHCO_3$ (30 mL), brine (30 mL), and dried over $MgSO_4$. Removal of solvent under reduced pressure furnished a white crystalline solid, identified as starting material **55** (982 mg, 98%).

TLC 20:1 $CHCl_3$:MeOH **RF:** 0.82 (**55**) **MP** 84-86 °C *(lit 83-86 °C)* **^{1}H NMR** (400MH_z, $CDCl_3$) δ (ppm) 7.63 (t, 1H, *J* = 7.8 Hz), 7.28 (d, 2H, *J* = 7.8 Hz) **EI MS(-)** (m/z) 146 ([M-1], 100 %)

6.4.4 2-chloro-4,6-dimethoxypyridine 60 and 4-chloro-2,6-dimethoxy-pyridine 61



57 **60** **61**

To a freshly-prepared solution of NaOMe (20 mL) under Ar / N_2, 2,4,6-trichloropyridine **57** (850 mg, 4.70 mmol) was added, and the solution heated under reflux for 16 hours. Solvent was removed under reduced pressure, and H_2O (20 mL) cautiously added. The product was extracted into CH_2Cl_2 (3 x 20 mL); combined organic layers were washed with brine, and dried over $MgSO_4$. Solvent was removed under reduced pressure, and the yellow residue was purified by FC. Elution with $CHCl_3$:MeOH (10:1, $^v/_v$) furnished 2-chloro-4,6-dimethoxylpyridine **60** (major product) as a colourless crystalline solid, which was recrystallised from pentane (203 mg, 1.17 mmol, 25 %); 4-chloro-2,6-dimethoxypyridine **61** (minor product) was isolated as a colourless crystalline solid (31 mg, 0.18 mmol, 4 %).

FC stationary phase Silica gel using 10 : 1 $CHCl_3$:MeOH, column bed 2.0 x 15 cm; **FC mobile phase** 10 : 1 $CHCl_3$:MeOH

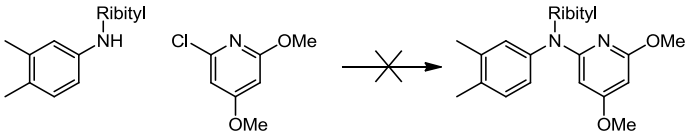
<u>2-chloro-4,6-dimethoxylpyridine **60**</u>

TLC 10:1 $CHCl_3$:MeOH **RF:** 0.94 (**60**); 4 : 1 hexane : EtOAc **RF:** 0.81 (**60**) **MP** 82-84 °C *(lit 81-85 °C)* **^{1}H NMR** (400 MHz, $CDCL_3$) δ (ppm) 6.49 (d, 1H, *J* = 2.0 Hz), 6.10 (d, 1H, *J* = 1.6 Hz), 3.89 (s, 3H), 3.79 (s, 3H) **EI MS (+)** (m/z) 172 ([M], 100 %) **HR MS (EI +)** found 172.0159, calc. 172.0165 for [M (^{35}Cl)]

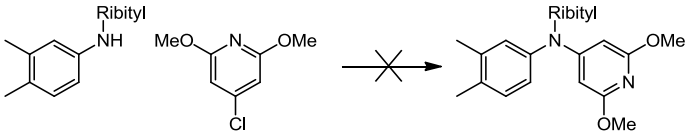
4-chloro-2,6-dimethoxypyridine **61**

TLC 10:1 $CHCl_3$:MeOH **RF:** 0.80 (**61**); 4 : 1 hexane : EtOAc **RF:** 0.69 (**61**) **MP** 63-65 °C *(lit 63-66 °C)* **^{1}H NMR** (400 MHz, $CDCL_3$) δ (ppm) 6.38 (s, 2H), 3.74 (s, 6H) **EI MS (+)** (m/z) 173 ([M+1], 100 %) **HR MS (EI +)** found 172.0159, calc. 172.0165 for [M (^{35}Cl)

6.4.5 Attempted coupling of ribitylated aniline 33 and chloro dimethoxypyridines 60 or 61



33 **60** **62**



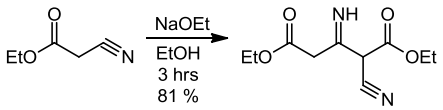
33 **61** **63**

Several methods to perform this reaction were attempted, based on the methods described in sections 6.1.3.1 - 6.1.3.3; however in all cases, the reaction failed to produce any products. Residues isolated from these reactions after purification (FC or recrystallisation) were determined from the spectroscopic evidence as initial reactants which had not undergone reaction. No products or side-products indicating degradation were visible.

6.4.6 Formation of 2,4,6-piperidinetrione 66

The synthesis of diethyl-2-cyano-3-iminoglutarate **65** was originally described by Baron *et al.*[173], under the synonym ethyl-β-imino-α-cyanoglutarate; the cyclisation of this to form 2,4,6-piperidinetrione **66** was originally described by Weinstock *et al.*[106].

6.4.6.1 Preparation of diethyl-2-cyano-3-iminoglutarate 65



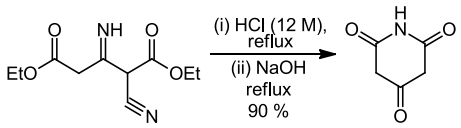
64 **65**

Ethyl cyanoacetate (11.3 g, 10.6 mL, 100 mmol) was dissolved in freshly-prepared sodium ethoxide (15 mL), and the solution heated under reflux for 3 hours. Once cool, H_2O (10 mL) was added cautiously, followed by aqueous HCl (5 mL, 2 M), to yield a purple solution which was extracted using Et_2O (3 x 80 mL). Combined organic layers were washed with a saturated

solution of $NaHCO_3$ (50 mL) and brine (50 mL), before drying over $MgSO_4$; concentration under reduced pressure furnished impure product **65** as a red oil. This was allowed to stand overnight, forming red needles, which were collected under suction and recrystallised from aqueous EtOH to furnish diethyl-2-cyano-3-iminoglutarate **65** (18.24g, 81 mmol, 81 %) as large white needles**.**

MP 52-53 °C *(lit 52-54 °C;)* **^{1}H NMR** (400 MHz, DMSO-d_6) δ (ppm) 9.10 (br s, 1H), 8.85 (br s, 1H), 4.17 (dq, 4H, J_{AM} = 3.6 Hz*,* J_{AX} = 4.4 Hz), 3.56 (s, 2H), 3.34 (s, 1H), 1.21 (dt, J_{AM} = 2.3 Hz*,* J_{AX} = 7.1 Hz); **APCI MS (+)** (m/z) 227 ([M+1], 100 %) **HR MS (APCI +)** found 227.1022, calc. 227.1032 for [M+1]

6.4.6.2 Cyclisation of diethyl-2-cyano-3-iminoglutarate to 2,4,6-piperidine-trione 66



65 **66**

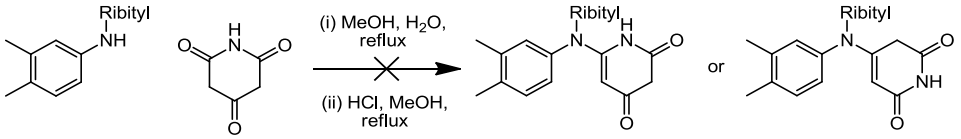
Diethyl-2-cyano-3-iminoglutarate **65** (15.4 g, 68 mmol) was suspended in concentrated (12 M) HCl (10 mL), and stirred at reflux for 3 hours forming a yellow solution. Solvent was removed under reduced pressure to furnish a pale-yellow solid, which was washed with Et_2O (20 mL) and collected by filtration. The solid retentate was dissolved in H_2O (50 mL), and the pH rapidly adjusted to 13-14 using solid NaOH (0.5 g), before boiling for 2 hours (with evolution of an alkaline gas smelling strongly of NH_3). Once cool, the pH was carefully adjusted to pH 7 using HCl (1M); care was taken not to over-acidify the mixture. The pale yellow solution was allowed to stand undisturbed overnight at RT, with the colour becoming significantly more vibrant. Solvent was removed under reduced pressure to yield the sodium adduct of piperidinetrione **66** as a luminous yellow powder (9.20 g, 61 mmol, 90 %).

MP: >360 °C (deg; *lit: deg > 330 °C);* **^{1}H NMR** (400 MHz, D_2O) δ (ppm) 3.16 (s, 2 H) [No resonance for N*H* was visible due solvent exchange; **^{13}C NMR** (100 MHz, D_2O) δ (ppm) 40.34; **^{13}C NMR DEPT 45** δ (ppm) 40.34 (CH_2); **^{13}C NMR DEPT 90** [no resonance]; **^{13}C NMR DEPT 135** (100 MHz) δ (ppm) **Positive Polarisation:** 40.34 **Negative Polarisation:** [no resonance]; **EI MS(-)** (m/z) 126 ([M-1], 100%); **HR MS (EI+)** found 127.0267, calc. 127.0269 for [M]

6.4.7 Coupling of 3,4-dimethyl-N(ribityl)-aniline 33 and 2,4,6-piperidinetrione 66

Each of these reactions was unsuccessful at furnishing the desired products. Of the isolated products, in all cases 3,4-dimethyl-N(ribityl)-aniline **33** was typically recovered in high yield; 2,4,6-piperidinetrione **66** was recovered in lower yield due to degradation, giving several unassignable low molecular weight compounds. Reaction 3.4.7.1 was repeated using 1-(N-ribityl),2-(N-Boc)-diamino-4,5-dimethyl-benzene **43** and 4-methyl-N(ribityl)-aniline **50**; no reaction was again observed, with staring materials isolated as before.

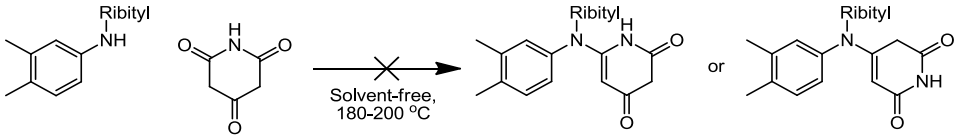
6.4.7.1 Aqueous coupling method (unsuccessful)



33 **66** **67a** **67b**

3,4-dimethyl-N(ribityl)-aniline **33** (500 mg, 1.96 mmol) and 2,4,6-piperidinetrione **66** (300 mg, 2.00 mmol) were suspended in a mixture of MeOH (10 mL) and H_2O (60 mL). The mixture was heated at reflux for 24 hours, with no reaction was observed. Aqueous HCl (10 mL) was added, and reflux maintained for 16 hours. The mixture was cooled to RT, with no precipitation observed; the solution was extracted using EtOAc (3 x 50 mL), and the organic layer washed using $NaHCO_3$ (2 x 50 mL), brine, and dried over $MgSO_4$. Solvent was removed under reduced pressure to furnish 3,4-dimethyl-N(ribityl)-aniline **33** (455 mg, 1.78 mmol, 91 %). 2,4,6-piperidinetrione **66** (and degradation products) were identified within the aqueous layer.

6.4.7.2 Solvent free heating method (unsuccessful)

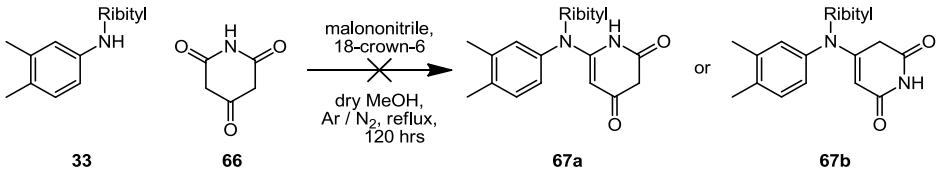


33 **66** **67a** **67b**

3,4-dimethyl-N(ribityl)-aniline **33** (500 mg, 1.96 mmol) and 2,4,6-piperidinetrione **66** (300 mg, 2.00 mmol) were powdered together in a RBF and dried under high vacuum for 4 hours. The solid mixture was heated to between 180–200 °C for 20 minutes, before being allowed to cool to RT. The brown solid was washed with Et_2O (5 mL) [although TLC indicated no dissolution of any compound into the organic solvent] and filtered. The solid retentate was dissolved in

water and extracted using EtOAc (3 x 40 mL). The organic layer was washed with $NaHCO_3$ (50 mL) and brine, dried over $MgSO_4$, before solvent removal under reduced pressure furnished 3,4-dimethyl-N(ribityl)-aniline **33** (quant.). As seen previously (reaction 3.4.7.1), 2,4,6-piperidinetrione **66** (and degradation products) were identified in the aqueous layer.

6.4.7.3 Malononitrile catalysed coupling method (unsuccessful)



33 **66** **67a** **67b**

3,4-dimethyl-N(ribityl)-aniline **33** (500 mg, 1.96 mmol), 2,4,6-piperidinetrione **66** (300 mg, 2.00 mmol), malononitrile (50 mg, 0.76 mmol) and 18-crown-6 (530 mg, 2.01 mmol) were suspended in thoroughly dried MeOH (50 mL) and heated under an Ar / N_2 atmosphere at reflux for 48 hours; no reaction was observed, with starting materials noted using TLC. A further portion of malononitrile (50 mg, 0.76 mmol) was added, and reflux maintained for a further 72 hours. The mixture was filtered and the solvent removed under reduced pressure; the viscous solid was dissolved in water (50 mL) and extracted using EtOAc (5 x 50 mL). Combined organic layers were washed using $NaHCO_3$ (2 x 50 mL) and brine, before drying over $MgSO_4$. Solvent was removed under reduced pressure to furnish 3,4-dimethyl-N(ribityl)-aniline **33** (495 mg, 99 %). 2,4,6-piperidinetrione **66** (and degradation products) were identified by TLC and MS within the aqueous layer.

6.5 HPLC purification of Riboflavin Analogues

Preparative-scale purification of the riboflavin analogues formed was performed using HPLC (section 5.5.2.3), following established protocols[251,178,179]. The gradient used is shown graphically in appendix 1..Phenomenex Luna, or YMC-Actus Triart columns (section 5.5.2.3) were used, with injection volumes of 2-10 mL. Mobile phase composition was acetonitrile ("A") and H_2O (containing 0.01 % $^{v}/_{v}$ TFA; "B") with an initial linear gradient (table 20), and a flow rate of 5 mL minute^{-1}. Flavin products were detected by monitoring the absorbance of the eluent from the column at a fixed wavelength, relative to the cofactor eluted.

Time (minutes)	Gradient (% A)
0	0
25	50
30	50
35	0

Retention Time (minutes)	Product Eluted	Wavelength monitored (nm)
14.10-16.05	Riboflavin	450
14.26-16.63	5-deazaRf	400
14.36-15.36	1-deazaRf	535
14.95-17.20	8-demethyl-5-deazaRf	450, 400

Table 20: Gradient used for separation of deazariboflavin products, and retention times of each riboflavin analogue

Chapter 7

Experimental Detail –

Molecular Biology

Chapter 7 - Experimental Detail - Molecular Biology

7.0.1 Introduction

In this chapter, the unique experimental conditions required for the effective expression and purification of several enzymes within various strains of *E. coli* are reported, in addition to the general description of methods presented in section 5.6. Where applicable, additional experiments utilising the recombinant proteins are also described. Instruments and apparatus required are listed in section 5.4.2; all buffers and solutions are listed in section 5.2.

7.1 *C.ammoniagenes* RfK-FADS

7.1.1 Expression of *C.ammoniagenes* RfK-FADS

Expression of the *C.ammoniagenes* RfK-FADS protein was performed using the general method given in section 5.6.2.2. The pre-prepared[1] pMAL-C2-RibfK-Ca construct containing the necessary gene was transformed into *E. coli* strains XL1-blue and BL21(DE3); expression was performed in both, with BL21(DE3) giving optimal yields. Typical liquid culture volumes were 3 L, with IPTG concentrations of 1-2 mmol used to induce protein expression at optimal cell density. Expression was allowed to continue for 4 hours at 37 °C, after which the cells were harvested using the general method (2.6.2.2). Typical wet cell masses of 10 gL^{-1} were obtained using strain XL1-blue, while strain BL21(DE3) gave 14 gL^{-1}. Harvested cells were stored at -20°C until purification.

7.1.2 Cell lysis – Ca-RfK-FADS

Cell lysis was performed in RfK 1, following sonication as described in section 5.6.3.4.

7.1.3 Purification of Ca-RfK-FADS

Amylose affinity chromatography was used to purify the Ca-RfK-FADS protein, as described in section 5.6.7.1. The eluted protein solution was concentrated by ultrafiltration followed by dialysis into buffer RfK 3. Activity remained for 7 days (maintained at pH 7.5, 4°C).

7.1.4 Formation of FMN and FAD analogues using Ca-RfK-FADS

Phosphorylation of riboflavin and 5-deazariboflavin using Ca-RfK-FADS was performed in buffer RfK 4. Non-phosphorylated flavin or analogue concentrations up to 20 mg L^{-1} were

used, with 5 equivalents of ATP. Concentrations of Ca-RfK-FADS up to 100 µM were added immediately prior to reaction. The enzymatic mixture was incubated at 37 °C with shaking overnight and assayed using TLC, with a complex mixture of products observed. Relative intensities of each flavin form were used to estimate reaction progress; upon reaching the estimated point of equilibrium between FMN and FAD, the mixture was lyophilised and stored at -20 °C until purification.

7.1.5 Purification of flavin cofactors; conversion of 5-deazaFAD to 5-deazaFMN

Preparative-scale purification of the phosphorylated products 5-deazaFMN and 5-deazaFAD was performed using HPLC (section 5.5.2.3) following established protocols[251,178,179]. The gradient used is shown graphically in appendix 1. A Phenomenex Luna column (5.5.2.3) was used, with injection volumes between 2-10 mL. Mobile phase composition was acetonitrile ("A") and 0.01 M potassium phosphate, pH 7 ("B"). A gradient of these solvents was used to give separation, using a stepwise gradient (table 21), with a flow rate of 5 mL minute^{-1}.

Time (minutes)	Gradient (% A)
0	15
14	15
15	35
24	35
25	65
29	65
30	90
35	90

Retention Time (minutes)	Product Eluted
7.35-9.45	5-deazaFAD
20.05-20.90	5-deazaFMN
28.70-29.70	5-deazaRf

Table 21: Gradient used for separation of 5-deazaFAD, FMN and Rf, and retention times of each. Chromatogram and gradient plot is shown graphically in appendix 1

The major product of the enzymatic phosphorylation was 5-deazaFAD (typically in ratios greater than 20:1 compared to 5-deazaFMN, estimated by HPLC). Cleavage of the phosphoanhydride bond of 5-deazaFAD was achieved using 1.2 equivalents of TFA added to an aqueous solution of purified 5-deazaFAD, resulting in complete cleavage of the phosphoanhydride after 2 hours. HPLC of the neutralised lyophilised mixture allowed isolation of 5-deazaFMN.

7.2 *S. pombe* RfK

7.2.1 Expression of Sp-RfK

During a collaborative side project with Professor Martin Warren and Dr. Hannah Collins (University of Kent, Canterbury), I was gifted a plasmid construct containing the Riboflavin Kinase (RfK) gene from *Schizosaccharomyces pombe*[183] (GenBank®[2,3] ref: NM_001023386.2), within a pET-14b plasmid vector. This plasmid was transformed into *E. Coli* strains XL1-blue and BL21(DE3); expression within BL21(DE3) was performed according to the general method given in section 5.6.2.2. Culture volumes between 1.5–3 L were used, with IPTG concentrations of 1-2 mmol used to induce protein expression. Once induced, expression was allowed to continue for 2 hours at 37 °C, or 4 hours at 20 °C, after which the cells were harvested (5.6.2.2). Typical wet cell masses of 12 gL^{-1} were obtained using LB media, and 19 gL^{-1} using TB media; harvested products were stored at -20°C until purification.

7.2.2 Cell lysis – Sp-RfK

Cell lysis was performed in buffer $SpRfK_{LYS}$, according to the general method given in section 5.6.3.4. Additional dilution of the cell lysate (1:1, using buffer SpRfK 1) after initial centrifugation prevented precipitation of the recombinant protein during purification.

7.2.3 Purification of Sp-RfK

IMAC purification of the Sp-RfK protein was performed using a column of Ni-sepharose, following the general method given in section 5.6.7.2. An XK 16/20 Column (GE Healthcare), equipped with an external jacket for cooling and capped with two QuickLock AK 16 adaptors was packed with metal-chelating Sepharose fast flow resin (GE Healthcare). The resin bed (5-20 mL) was cast with continual pouring, with headspace over the resin minimised. Solutions were loaded onto the column using a Watson-Marlow 101U/P peristaltic pump (flow rate 0.5-5.0 mL $minute^{-1}$), and the column pre-equilibrated with 10 column volumes (c.v.) of buffer Sp-RfK 1. Filtered lysate was loaded, and the bound sample washed using buffer Sp-RfK 1 (2 c.v.), followed by Sp-RfK 2 (5 c.v.). Elution was achieved using buffer Sp-RfK 3, after which the purified protein solution was concentrated by ultrafiltration and dialysed into buffer Sp-RfK 4 to remove imidazole. Long-term storage of the protein was possible after flash-freezing.

7.2.4 Formation of FMN analogues using Sp-RfK

Phosphorylation of 5-deazariboflavin and 1-deazariboflavin using Sp-RfK utilised a portion of the prepared protein sample within buffer Sp-RfK 5. Reactions were performed at 37 °C with shaking, typically utilising 5 equivalents of ATP to give completion in 1 hour, although lower ATP concentrations (1.2 equivalents) still gave complete reaction in 4 hours. Once complete, the enzymatic mixture was lyophilised and stored at -20 °C until purification.

7.2.5 Purification of FMN analogues formed using Sp-RfK

Purification of the phosphorylated products (5-deazaFMN and 1-deazaFMN) was performed on a preparative scale using HPLC following the same protocol used for the purification of the Ca-RfK-FADS mixtures[251,178,179], with the same gradient (appendix 1) and conditions (section 7.1.5) giving similar retention times.

Since using HPLC for preparative isolation of the product was particularly time consuming, an alternative method of purification was attempted in cases where subsequent reaction or use of the cofactor would not be hindered or otherwise affected by the presence of low-concentration contaminants (such as ADP) within the reaction. Upon completion of the enzymatic reaction, the crude mixture was passed over the IMAC column previously used to separate the Sp-RfK from other intracellular proteins. As the polyhistidine tag was not removed from the protein, it was possible to remove the kinase from the solution of FMN. The mixture was subsequently lyophilised prior to storage at -20 °C.

7.3 *A. sativa* LOV2

7.3.1 Expression of AsLOV2(WT) and mutant proteins

The modified pNCO113-HISACT(C49S)-AsLOV2(WT) vector prepared by Fisher, Kay *et al.*[1,4] was used for the expression of the wild-type *Avena sativa* PHOT1-LOV2 domain. This was maintained within *E. coli* strain XL1-blue, and expressed within BL21(DE3) cells. Alternative strains of *E. coli* examined for expression of this plasmid include M15[pRep4] and JM101, although JM101 furnished lower protein yields than BL21(DE3) or M15[pRep4]. Mutant derivatives of the wild-type AsLOV2 gene (referred to subsequently as AsLOV2(C450X)) were expressed using a mutated vector within strain JM101.

Culture growth was performed in line with the general method given in section 5.6.2.2, in TB media (LB was found to be ineffective). Culture volumes between 0.5-3 L were used, with IPTG concentrations of 1-5 mmol used to induce protein expression. However, while an OD_{600} of 0.6-0.8 was used for expression of the C450X mutants, an OD_{600} between 1.2-1.4 was possible for expression of the wild-type LOV2 domain, as the richness of the TB growth media allowed significantly extended bacterial growth[245]. After induction, expression was performed overnight at 16-20 °C, after which the cells were harvested as before. Typical wet cell masses between 18 - 26 g L^{-1} were obtained; harvested products were stored at -20°C until purification.

7.3.2 Cell Lysis – AsLOV2(WT) and AsLOV2(C450X)

Frozen cells were resuspended and thawed in an equivolume quantity of buffer LOV_R. Immediately prior to cell lysis, this was diluted using two volumes of buffer LOV_{LYS}. Cells were sonnicated using an intermittent cycle of 5 seconds sonication, followed by 15 seconds of no disruption, for 10 minutes, as more vigorous sonication led to extensive fragmentation of the hisactophilin-AsLOV2 fusion protein. After lysis, the mixture was centrifuged (15000 RPM, 30 minutes, 4 °C); the supernatant was decanted, and centrifugation twice repeated. The centrifuged lysate was finally filtered using a MillexGP syringe filter (0.22 µm).

7.3.3 Purification of AsLOV2(WT) and AsLOV2(C450X)

Purification of the cell lysate was performed using an IMAC procedure similar to that used for the purification of Sp-RfK (section 7.2.3, *vide supra*). The vivid yellow lysate was passed over a column of Ni-Sepharose (fast-flow) resin using a peristaltic pump (flow rate 0.5 – 2.0 mL $minute^{-1}$), which had been packed and pre-equilibrated using buffer LOV_A. Binding of the hisactophilin-AsLOV2 protein was indicated visually, as the yellow protein appeared green when immobilised on the blue resin (figure 87, section 3.3.6). The column was washed with a minimum of 5 column volumes of buffer LOV_A, and protein samples for analysis containing the native FMN cofactor were eluted using buffer LOV_C.

Eluted fractions were analysed using SDS-PAGE, and concentrated via ultrafiltration if necessary. Fractions containing the purified protein were combined, and imidazole was

removed by dialysis using buffer LOV_D, or by passing the solution through a pre-packed HiPrep™ 26/10 desalting column equilibrated with buffer LOV_D.

7.3.4 Cofactor Exchange within the AsLOV2(WT) and AsLOV2(C450X) Domains

Replacement of the native FMN cofactor with either of the enzymatically modified analogues (5-deazaFMN or 1-deazaFMN) was performed prior to elution of the protein sample from the IMAC resin. After washing of the retained protein, buffer LOV_B (containing 1.0-4.5 M guanidine thiocyanate; 3 column volumes) was applied to the column, causing denaturing and elution of the native FMN cofactor. Thiocyanate was removed by washing with 2 column volumes of buffer LOV_A or a 2 % saline solution. A solution of the desired FMN analogue to be incorporated (typically 20 mM, although 10 mM was also effective) was cycled over the retained denatured sample continuously for 16 hours at 4 ° C. Care was taken to ensure both the column and deazaFMN solution remained in the dark, to prevent photo-degradation of the cofactor. After 16 hours, the synthetic FMN analogue solution was removed and the column thoroughly washed with buffer LOV_A; successful incorporation of the analogue was immediately observable by the retained protein adopting the colour of the synthetic cofactor. Protein was eluted from the column using the method described in section 7.3.3 above.

Samples prepared using this method were allowed to stand overnight at 4 °C in the dark, as any misfolded or apoprotein-form proteins readily precipitated from the mixture, which were removed by centrifugation (4000 RPM, 30 minutes, 4°C) prior to removal of imidazole as before. Typical incorporation efficiencies above 80 % (determined upon the concentration of FMN eluted) were obtained using this method.

7.3.5 Site Directed Mutagenesis (SDM) of AsLOV2(WT) to form AsLOV2(C450X) mutants

Mutagenesis of the AsLOV2 gene was performed based upon the PCR-based whole plasmid method of Weiner *et al.*[213], using the original vector pNCO113-HISACT(C49S)-AsLOV2(WT) isolated from *E. coli* XL1-blue and thoroughly purified to remove any salts. An introduction to the technique is given in section 5.6.6.

Primers (listed in table 13, section 5.7.3) incorporating the desired mutation(s) to alter the gene transcribed by the codon originally expressing cysteine 450 of AsLOV2 (C450) were designed using the PrimerX tool[248], and supplied by Eurofins MWG Operon.

A typical PCR mixture for the SDM process was comprised of:

Volume	Mixture Component
39 µL	ddH_2O
5 µL	*Pfu* DNA polymerase buffer (10 x)
2 µL	DNA primer, forward (2 nanomoles)
2 µL	DNA primer, reverse (2 nanomoles)
1 µL	dNTP mixture (100 µM)
1 µL	parental DNA (20 ng)

Table 22: Typical composition of a PCR mixture for SDM of the AsLOV2 gene

All reactant mixtures were kept on ice until initiation of the PCR. *Pfu* DNA polymerase (1 µL) was added immediately before initiation of the reaction. 16 cycles of the PCR were performed using an elongation time of between 10-15 minutes per cycle, to allow sufficient time for the polymerase to complete its synthesis. Upon completion of the PCR cycle, all products were digested using *DpnI*, which cut any methylated DNA present. As the parental DNA was methylated (the newly-formed daughter DNA was not), this allowed selection of only the daughter (mutant) DNA.

Purified DNA was transformed into *E. coli* strain XL1-blue, and plated on LB-Agar plates containing ampicillin to allow selection. Transformants were grown in liquid cultures, before the DNA was extracted and sequenced to confirm desired mutation and to ensure no other undesired alterations (including frame-shifting and silent mutations). Mutant DNA was subsequently transformed into *E. coli* strain JM101 for expression (as described in sections 7.3.1 to 7.3.4, above).

7.3.6 Spectroscopic examination of AsLOV2(WT) and AsLOV2(C450X)

Spectroscopic examination of all AsLOV2 – derived protein samples was performed as described in section 5.6.8, in buffer LOV_D, at pH 8.0. Spectra of the wild-type protein containing the native cofactor, 5-deaza FMN and 1-deazaFMN are shown in the appropriate

experimental sections, along with selected spectra of the AsLOV2-mutant proteins. UV-visible spectra for each mutant sample were largely identical to the wild-type (dark state) spectrum, with the primary absorbance wavelength (and accompanying vibronic side-bands) determined by the cofactor bound by the protein. Key spectral characteristics of each sample are summarised in table 23 below.

Protein Sample	**Cofactor**	**Primary Absorbance Wavelength, nm**	**Vibronic Side-Band Wavelength, nm**	
			Most Intense	**Least Intense**
Wild-Type	native FMN	447	473	422
	5-deazaFMN	404	428	384
	1-deazaFMN	535	503	570
C450M	native FMN	447	472	421
	5-deazaFMN	404	428	383
	1-deazaFMN	535	502	570
C450S	native FMN	447	473	422
	5-deazaFMN	404	428	384
	1-deazaFMN	535	503	571
C450D	native FMN	448	474	422
	5-deazaFMN	404	428	384
	1-deazaFMN	535	503	570
C450G	native FMN	447	473	422
	5-deazaFMN	404	427	383
	1-deazaFMN	535	502	571
C450A(T418I)	native FMN	447	473	422
	5-deazaFMN	404	428	384
	1-deazaFMN	533	502	570

Table 23: Summary of UV-visible spectral results for the incorporation of synthetic flavins into several mutant derivatives of AsLOV2. The primary absorbance (determined by absorbance of the cofactor) is shown, along with the wavelength observed for the vibronic side-bands surrounding this. All of those listed (except for the wild-type protein containing native FMN or 5-deazaFMN) were inactive to photoirradiation. Note that for all samples containing 1-deazaFMN, the relative intensity of the side bands is the reverse of the intensity found for native FMN and 5-deazaFMN

7.4 Additional Experiments

7.4.1 Examination of possible photo-stimulated reaction between FMN, 5-deazaFMN (and 5-deazaRf), and two cysteine based Michael acceptors

A range of solutions containing 5-deazaRf and 5-deazaFMN were prepared, at concentrations between 50 μM and 250 μM. Freshly prepared solutions of free-(*L*)-cysteine and fully-reduced glutathione in H_2O (100 mM or 300 mM) were mixed with an equivolume quantity of each flavin solution (giving minimum solution concentrations of 0.25 μM flavin and 50 mM cysteine or glutathione).

UV-visible spectra were immediately recorded, and compared to the spectrum of a similarly-diluted flavin solution containing no thiol. No difference between the two spectra was observed. Re-examination after several hours (stored in the dark) showed no decrease in flavin absorbance in the cysteine-containing mixture.

The cysteine and glutathione-containing samples were irradiated at 400 nm for periods from 30 seconds to 300 seconds, and the UV spectra immediately recorded. In every case, no change was found to any UV spectrum compared to pre-irradiation levels, suggesting no formation of a flavin-cysteine bond.

These experiments were repeated using FMN (with irradiation at 450 nm), with the same result observed.

Chapter 8 - References

1. W. Eisenreich, M. Joshi, B. Illarionov, G. Richter, W. Römisch-Margl, F. Müller, A. Bacher, and M. Fischer, *FEBS Journal*, 2007, **274**, 5876–5890.
2. D. A. Benson, I. Karsch-Mizrachi, D. J. Lipman, J. Ostell, and E. W. Sayers, *Nucleic Acids Res*, 2011, **39**, D32–D37.
3. GenBank and http://www.ncbi.nlm.nih.gov/genbank/, .
4. C. W. M. Kay, E. Schleicher, A. Kuppig, H. Hofner, W. Rüdiger, M. Schleicher, M. Fischer, A. Bacher, S. Weber, and G. Richter, *J. Biol. Chem.*, 2003, **278**, 10973–10982.
5. E. E. Carlson and L. L. Kiessling, *J. Org. Chem.*, 2004, **69**, 2614–2617.
6. J. Sachs, *Lectures on the Physiology of Plants; Lexture XVIII (translated from "Vorlesgungen uber Pflanzen-physiologie" by H. Marshall Ward).*, Oxford, Clarendon Press, Page 304, 1887.
7. J. Sachs, *Lectures on the Physiology of Plants; Lexture XXXV (translated from "Vorlesgungen uber Pflanzen-physiologie" by H. Marshall Ward).*, Oxford, Clarendon Press, Page 612, 1887.
8. H. Kende and J. Zeevaart, *Plant Cell*, 1997, **9**, 1197–1210.
9. P. F. Wareing and I. D. J. Phillips, *The control of growth and differentiation in plants,*, Pergamon Press, Pages 134-140, 1st ed., 1970.
10. N. Cholodny, *Science*, 1937, **86**, 468.
11. F. W. Went, *Bot. Rev*, 1935, **1**, 162–182.
12. I. R. Macdonald and J. W. Hart, *Plant Physiol.*, 1987, **84**, 568–570.
13. A. J. Trewavas, *Plant Cell Environ.*, 1992, **15**, 761.
14. A. C. Leopold, *Plant Cell Environ.*, 1992, **15**, 777–778.
15. M. Michniewicz, P. B. Brewer, and J. Friml, *Arabidopsis Book*, 2007, **5**.
16. R. J. Pattison and C. Catalá, *The Plant Journal*, 2012, **70**, 585–598.
17. N. A. Eckardt, *Plant Cell*, 2003, **15**, 2755–2757.
18. M. Kasahara, T. Kagawa, Y. Sato, T. Kiyosue, and M. Wada, *Plant Physiol*, 2004, **135**, 1388–1397.
19. C.-H. Goh, *Plant Signal Behav*, 2009, **4**, 693–695.
20. T. Kinoshita, M. Doi, N. Suetsugu, T. Kagawa, M. Wada, and K. Shimazaki, *Nature*, 2001, **414**, 656–660.
21. S. Inoue, A. Takemiya, and K. Shimazaki, *Curr. Opin. Plant Biol.*, 2010, **13**, 587–593.
22. H. E. Boccalandro, C. V. Giordano, E. L. Ploschuk, P. N. Piccoli, R. Bottini, and J. J. Casal, *Plant Physiol.*, 2011, pp.111.187237.
23. E. S. Johnston, *Smithsonian Miscellaneous Collections*, 1934, **12**, 1–17.
24. M. B. Wilkins, Ed., in *Physiology of Plant Growth and Development*, McGraw-Hill, Page 249, 1st edn., 1969.
25. J. M. Christie, M. Salomon, K. Nozue, M. Wada, and W. R. Briggs, *PNAS*, 1999, **96**, 8779–8783.
26. E. Liscum and W. R. Briggs, *Plant Cell*, 1995, **7**, 473–485.
27. E. Huala, P. W. Oeller, E. Liscum, I.-S. Han, E. Larsen, and W. R. Briggs, *Science*, 1997, **278**, 2120–2123.
28. C. Murray, J. T. Christeller, L. N. Gatehouse, and W. A. Laing, *Plant Physiology*, 1998, **116**, 869–869.
29. S. Crosson, S. Rajagopal, and K. Moffat, *Biochemistry*, 2003, **42**, 2–10.
30. S. Crosson and K. Moffat, *PNAS*, 2001, **98**, 2995–3000.
31. M. Salomon, J. M. Christie, E. Knieb, U. Lempert, and W. R. Briggs, *Biochemistry*, 2000, **39**, 9401–9410.

32. Y. I. Wu, D. Frey, O. I. Lungu, A. Jaehrig, I. Schlichting, B. Kuhlman, and K. M. Hahn, *Nature*, 2009, **461**, 104–108.
33. A. T. Vaidya, D. Top, C. C. Manahan, J. M. Tokuda, S. Zhang, L. Pollack, M. W. Young, and B. R. Crane, *Proc. Natl. Acad. Sci. U.S.A.*, 2013.
34. M. Gomelsky and G. Klug, *Trends Biochem. Sci.*, 2002, **27**, 497–500.
35. W. Laan, M. A. van der Horst, I. H. van Stokkum, and K. J. Hellingwerf, *Photochem. Photobiol.*, 2003, **78**, 290–297.
36. S. Masuda, K. Hasegawa, A. Ishii, and T. Ono, *Biochemistry*, 2004, **43**, 5304–5313.
37. C. Schwerdtfeger and H. Linden, *EMBO J*, 2003, **22**, 4846–4855.
38. B. D. Zoltowski and B. R. Crane, *Biochemistry*, 2008, **47**, 7012–7019.
39. B. D. Zoltowski, B. Vaccaro, and B. R. Crane, *Nature Chemical Biology*, 2009, **5**, 827–834.
40. E. Peter, B. Dick, and S. A. Baeurle, *Nature Communications*, 2011, **1**, Article No. 122.
41. J. P. Zayner, C. Antoniou, and T. R. Sosnick, *J. Mol. Biol.*, 2012, **419**, 61–74.
42. E. Herman, M. Sachse, P. G. Kroth, and T. Kottke, *Biochemistry*, 2013, **52**, 3094–3101.
43. S. M. Harper, L. C. Neil, and K. H. Gardner, *Science*, 2003, **301**, 1541–1544.
44. X. Yao, M. K. Rosen, and K. H. Gardner, *Nat Chem Biol*, 2008, **4**, 491–497.
45. S. M. Harper, J. M. Christie, and K. H. Gardner, *Biochemistry*, 2004, **43**, 16184–16192.
46. E. Schleicher, R. M. Kowalczyk, C. W. M. Kay, P. Hegemann, A. Bacher, M. Fischer, R. Bittl, G. Richter, and S. Weber, *J. Am. Chem. Soc.*, 2004, **126**, 11067–11076.
47. T. E. Swartz, S. B. Corchnoy, J. M. Christie, J. W. Lewis, I. Szundi, W. R. Briggs, and R. A. Bogomolni, *J. Biol. Chem.*, 2001, **276**, 36493–36500.
48. A. S. Halavaty and K. Moffat, *Biochemistry*, 2007, **47**, 14001–14009.
49. M. Salomon, W. Eisenreich, H. Dürr, E. Schleicher, E. Knieb, V. Massey, W. Rüdiger, F. Müller, A. Bacher, and G. Richter, *Proc. Natl. Acad. Sci. U.S.A.*, 2001, **98**, 12357–12361.
50. P.-S. Song, *Photochemistry and Photobiology*, 1968, **7**, 311–313.
51. J. T. M. Kennis, S. Crosson, M. Gauden, I. H. M. van Stokkum, K. Moffat, and R. van Grondelle, *Biochemistry*, 2003, **42**, 3385–3392.
52. M. T. A. Alexandre, T. Domratcheva, C. Bonetti, L. J. G. W. van Wilderen, R. van Grondelle, M.-L. Groot, K. J. Hellingwerf, and J. T. M. Kennis, *Biophysical Journal*, 2009, **97**, 227–237.
53. R. Fedorov, I. Schlichting, E. Hartmann, T. Domratcheva, M. Fuhrmann, and P. Hegemann, *Biophys J*, 2003, **84**, 2474–2482.
54. G. Richter, S. Weber, W. Romisch, A. Bacher, M. Fischer, and W. Eisenreich, *J. Am. Chem. Soc.*, 2005, **127**, 17245–17252.
55. C. Neiß and P. Saalfrank, *Photochemistry and Photobiology*, 2003, **77**, 101–109.
56. O. I. Lungu, R. A. Hallett, E. J. Choi, M. J. Aiken, K. M. Hahn, and B. Kuhlman, *Chem Biol*, 2012, **19**, 507–517.
57. D. Strickland, X. Yao, G. Gawlak, M. K. Rosen, K. H. Gardner, and T. R. Sosnick, *Nat Meth*, 2010, **7**, 623–626.
58. M. T. A. Alexandre, J. C. Arents, R. van Grondelle, K. J. Hellingwerf, and J. T. M. Kennis, *Biochemistry*, 2007, **46**, 3129–3137.
59. T. Iyanagi and I. Yamazaki, *Biochim. Biophys. Acta*, 1969, **172**, 370–381.
60. S. Zenno, K. Saigo, H. Kanoh, and S. Inouye, *J. Bacteriol.*, 1994, **176**, 3536–3543.
61. H. Nishida, K. Inaka, and K. Miki, *FEBS Lett.*, 1995, **361**, 97–100.
62. S. C. Tu, *Antioxid. Redox Signal.*, 2001, **3**, 881–897.
63. G. W. Snider, C. M. Dustin, E. L. Ruggles, and R. J. Hondal, *Biochemistry*, 2014.
64. A. Meister, *J. Biol. Chem.*, 1988, **263**, 17205–17208.
65. S. A. Susin, H. K. Lorenzo, N. Zamzami, I. Marzo, B. E. Snow, G. M. Brothers, J. Mangion, E. Jacotot, P. Costantini, M. Loeffler, N. Larochette, D. R. Goodlett, R. Aebersold, D. P. Siderovski, J. M. Penninger, and G. Kroemer, *Nature*, 1999, **397**, 441–6.

66. V. Massey, *Biochem. Soc. Trans.*, 2000, **28**, 283–296.
67. A. Bacher, S. Eberhardt, W. Eisenreich, M. Fischer, S. Herz, B. Illarionov, K. Kis, and G. Richter, *Vitam. Horm.*, 2001, **61**, 1–49.
68. S. Mörtl, M. Fischer, G. Richter, J. Tack, S. Weinkauf, and A. Bacher, *J. Biol. Chem.*, 1996, **271**, 33201–33207.
69. A. G. Vitreschak, D. A. Rodionov, A. A. Mironov, and M. S. Gelfand, *Nucl. Acids Res.*, 2002, **30**, 3141–3151.
70. H. Marx, D. Mattanovich, and M. Sauer, *Microbial Cell Factories*, 2008, **7**, 23.
71. P. Macheroux, B. Kappes, and S. E. Ealick, *FEBS Journal*, 2011, **278**, 2625–2634.
72. H. Nakano and D. B. McCormick, in *Flavins and Flavoproteins 1990*, 1990, pp. pp. 89–92.
73. I. Efimov, V. Kuusk, X. Zhang, and W. S. McIntire, *Biochemistry*, 1998, **37**, 9716–9723.
74. S. Frago, M. Martínez-Júlvez, A. Serrano, and M. Medina, *BMC Microbiology*, 2008, **8**, 160.
75. S. Frago, A. Velázquez-Campoy, and M. Medina, *J. Biol. Chem.*, 2009, **284**, 6610–6619.
76. A. Krupa, K. Sandhya, N. Srinivasan, and S. Jonnalagadda, *Trends in Biochemical Sciences*, 2003, **28**, 9–12.
77. C. Huerta, D. Borek, M. Machius, N. V. Grishin, and H. Zhang, *J. Mol. Biol.*, 2009, **389**, 388–400.
78. D. J. Manstein and E. F. Pai, *J. Biol. Chem*, 1986, **261**, 16169–16173.
79. B. Herguedas, M. Martínez-Júlvez, S. Frago, M. Medina, and J. A. Hermoso, *J. Mol. Biol*, 2010, **400**, 218–230.
80. S. Ghisla, V. Massey, J.-M. Lhoste, and S. G. Mayhew, *Biochemistry*, 1974, **13**, 589–597.
81. V. Massey, *J. Biol. Chem.*, 1994, **269**, 22459–22462.
82. C. W. Kay, R. Feicht, K. Schulz, P. Sadewater, A. Sancar, A. Bacher, K. Möbius, G. Richter, and S. Weber, *Biochemistry*, 1999, **38**, 16740–16748.
83. E. Schleicher and S. Webber, *EPR Spectroscopy: Applications in Chemistry and Biology*, Springer, “Radicals In Flavoproteins”, Page 43, 2012.
84. P. Karrer, *Monatshefte für Chemie*, 1935, **66**, 367–392.
85. J. C. M. Tsibris, D. B. McCormick, and L. D. Wright, *J. Biol. Chem.*, 1966, **241**, 1138–1143.
86. S. Ghisla and V. Massey, *Biochem J.*, 1986, **239**, 1–12.
87. D. Bender and S. Ghisla, in *Flavoprotein Protocols*, eds. S. K. Chapman and G. A. Reid, Humana Press, 1999, pp. 157–179.
88. C. Walsh, J. Fisher, R. Spencer, D. W. Graham, W. T. Ashton, J. E. Brown, R. D. Brown, and E. F. Rogers, *Biochemistry*, 1978, **17**, 1942–1951.
89. D. E. Edmondson, B. Barman, and G. Tollin, *Biochem.*, 1972, **11**, 1133–1138.
90. G. Blankenhorn, *Eur. J. Biochem.*, 1976, **67**, 67–80.
91. M. Insinska-Rak, E. Sikorska, J. L. Bourdelande, I. V. Khmelinskii, W. Prukala, and K. Dobek, *J. Mol. Struc.*, 2006, **783**, 184–190.
92. G. Eberlein and T. C. Bruice, *J. Am. Chem. Soc.*, 1982, **104**, 1449–1452.
93. L. B. Hersh and M. S. Jorns, *J. Biol. Chem.*, 1975, **250**, 8728–8734.
94. B. A. Averill, A. Schonbrunn, and R. H. Abeles, *J. Biol. Chem.*, 1975, **250**, 1603–1605.
95. J. Fisher, R. Spencer, and C. Walsh, *Biochemistry*, 1976, **15**, 1054–1064.
96. S. Ghisla, C. Thorpe, and V. Massey, *Biochemistry*, 1984, **23**, 3154–3161.
97. R. Spencer, J. Fisher, and C. Walsh, *Biochemistry*, 1977, **16**, 3586–3594.
98. R. Spencer, J. Fisher, and C. Walsh, *Biochemistry*, 1977, **16**, 3594–3602.
99. R. D. Draper and L. L. Ingraham, *Archives of Biochemistry and Biophysics*, 1968, **125**, 802–808.
100. M. Mansurova, M. S. Koay, and W. Gartner, *Eur. J. Org. Chem.*, 2008, **2008**, 5401–5406.
101. C. Slavov, M. Mansurova, A. R. Holzwarth, and W. Gartner, *Photochemistry and Photobiology*, 2010, **86**, 31–38.

102. M. R. Silva-Junior, M. Mansurova, W. Gärtner, and W. Thiel, *ChemBioChem*, 2013, **14**, 1648–1661.
103. D. E. O'Brien, L. T. Weinstock, and C. C. Cheng, *J. Het. Chem.*, 1970, **7**, 99–105.
104. W. T. Ashton, D. W. Graham, R. D. Brown, and E. F. Rogers, *Tetrahed. Lett.*, 1977, **18**, 2551–2554.
105. A. F. Abdel-Magid, K. G. Carson, B. D. Harris, C. A. Maryanoff, and R. D. Shah, *The Journal of Organic Chemistry*, 1996, **61**, 3849–3862.
106. L. T. Weinstock, C. J. W. Wiegand, and C. C. Cheng, *Journal of Heterocyclic Chemistry*, 1977, **14**, 1261–1262.
107. F. Yoneda, Y. Sakuma, M. Ichiba, and K. Shinomura, *Chemical & Pharmaceutical Bulletin*, 1972, **20**, 1832–1834.
108. F. Yoneda, Y. Sakuma, M. Ichiba, and K. Shinomura, *J. Am. Chem. Soc.*, 1976, **98**, 830–835.
109. W. T. Ashton, R. D. Brown, and R. L. Tolman, *J. Het. Chem.*, 1978, **15**, 489–491.
110. A. Vilsmeier and A. Haack, *Berichte der deutschen chemischen Gesellschaft (A and B Series)*, 1927, **60**, 119–122.
111. F. Yoneda, Y. Sakuma, S. Mizumoto, and R. Ito, *J. Chem. Soc. Perkin 1*, 1976, 1805–8.
112. F. Yoneda and Y. Sakuma, *J. Chem. Soc., Chem. Commun.*, 1976, 203–204.
113. R. A. Y. Jones and J. F. Bunnett, *Pure and Applied Chemistry*, 1989, **61**, 725–768.
114. Schwarzenbach, R. P., P. M. Gschwend, and Imboden, D. M., *Environmental Organic Chemistry*, John Wiley and Sons, Inc., Page 267, 2nd edn., 2003.
115. J. C. Westall and C. A. Johnson, *Environ. Sci. and Technol.*, 1990, **24**, 1803–1810.
116. N. F. Hall and M. R. Sprinkle, *J. Am. Chem. Soc.*, 1932, **54**, 3469–3485.
117. H. K. Hall, *J. Am. Chem. Soc.*, 1957, **79**, 5441–5444.
118. M. McLaughlin, M. Palucki, and I. W. Davies, *Org. Lett.*, 2006, **8**, 3307–3310.
119. M. Horn, H. Mayr, E. Lacôte, E. Merling, J. Deaner, S. Wells, T. McFadden, and D. P. Curran, *Org. Lett.*, 2012, **14**, 82–85.
120. J. Zhou and B. List, *J. Am. Chem. Soc.*, 2007, **129**, 7498–7499.
121. D. B. G. Williams and M. Lawton, *J. Org. Chem.*, 2010, **75**, 8351–8354.
122. W. L. F. Armarego and D. D. Perrin, *Purification of Laboratory Chemicals*, Butterworth-Heinemann Publishers, 5th edn., 1997.
123. J. Leonard, B. Lygo, and G. Proctor, *Advanced Practical Organic Chemistry, Third Edition*, CRC Press, 3rd edn., 2013.
124. E. V. Dehmlow, *Angewandte Chemie International Edition in English*, 1974, **13**, 170–179.
125. Starks Charles M., in *Phase-Transfer Catalysis*, American Chemical Society, 1987, vol. 326, pp. 1–7.
126. A. Sakakura, K. Kawajiri, T. Ohkubo, Y. Kosugi, and K. Ishihara, *J. Am. Chem. Soc.*, 2007, **129**, 14775–14779.
127. Mestrelab MESTREC and http://mestrelab.com/, .
128. K. J. Baeyens, H. L. De Bondt, and S. R. Holbrook, *Nat Struct Mol Biol*, 1995, **2**, 56–62.
129. S. Schroeder, J. Kim, and D. H. Turner, *Biochemistry*, 1996, **35**, 16105–16109.
130. B. N. Bourdélat-Parks and R. M. Wartell, *Biochemistry*, 2005, **44**, 16710–16717.
131. T. L. P. Galvão, I. M. Rocha, M. D. M. C. Ribeiro da Silva, and M. A. V. Ribeiro da Silva, *J. Phys. Chem. A*, 2013, **117**, 5826–5836.
132. P. Safar, F. Povazanec, P. Cepec, and N. Pronayova, *Collect. Czech. Chem. Commun.*, 1997, **62**, 1105–1113.
133. A. Schnyder, A. Indolese, T. Maetzke, J. Wenger, and H.-U. Blaser, *Synlett*, 2006, **2006**, 3167–3169.
134. A. D. Thompson and M. P. Huestis, *J. Org. Chem.*, 2013, **78**, 762–769.
135. C. Frier, J.-L. Decout, and M. Fontecave, *J. Org. Chem.*, 1997, **62**, 3520–3528.

136. C. Mielczarek, *Farmacja Polska*, 2003, **59**.
137. P. L. López-de-Alba, L. López-Martínez, V. Cerdá, and J. Amador-Hernández, *Journal of the Brazilian Chemical Society*, 2006, **17**, 715–722.
138. E. Brzezińska, C. Mielczarek, and W. Pajak, *Acta Pol Pharm*, 2008, **65**, 59–63.
139. P. G. M. Wuts and T. W. Greene, *Protective Groups in Organic Synthesis*, Wiley, 4th edn., 2006.
140. C. A. G. N. Montalbetti and V. Falque, *Tetrahedron*, 2005, **61**, 10827–10852.
141. P. Tiwari, R. Kumar, P. R. Maulik, and A. K. Misra, *European Journal of Organic Chemistry*, 2005, **2005**, 4265–4270.
142. S. Klemenc, *Forensic Sci. Int.*, 2002, **129**, 194–199.
143. G. Jones and S. P. Stanforth, in *Organic Reactions*, John Wiley & Sons, Inc., 2004.
144. S. Ushijima and H. Togo, *Synlett*, 2010, **2010**, 1067–1070.
145. A. P. Rajput and P. D. Girase, *IJPCBS*, 2013, **3**, 25–43.
146. R. M. NAIK and V. M. THAKOR, *J. Org. Chem.*, 1957, **22**, 1630–1633.
147. R. A. Aungst, C. Chan, and R. L. Funk, *Org. Lett.*, 2001, **3**, 2611–2613.
148. D. L. Comins and A. L. Williams, *Org. Lett.*, 2001, **3**, 3217–3220.
149. G. Jones and S. P. Stanforth, in *Organic Reactions*, John Wiley & Sons, Inc., 2004.
150. T. Neilson and E. S. Werstiuk, *Can. J. Chem.*, 1971, **49**, 493–499.
151. V. M. Csizmadia, K. M. Koshy, K. C. M. Lau, R. A. McClelland, V. J. Nowlan, and T. T. Tidwell, *J. Am. Chem. Soc.*, 1979, **101**, 974–979.
152. I. Lee, C. K. Kim, and H. S. Seo, *Tetrahedron*, 1986, **42**, 6627–6633.
153. E. R. Koft, P. Dorff, and R. Kullnig, *J. Org. Chem.*, 1989, **54**, 2936–2940.
154. T. H. Fife, R. Bembi, and R. Natarajan, *J. Am. Chem. Soc.*, 1996, **118**, 12956–12963.
155. E. A. Englund, H. N. Gopi, and D. H. Appella, *Org. Lett.*, 2004, **6**, 213–215.
156. C. E. Barry, P. G. Nayar, and T. P. Begley, *Biochemistry*, 1989, **28**, 6323–6333.
157. L. A. Maggio-Hall, P. C. Dorrestein, J. C. Escalante-Semerena, and T. P. Begley, *Org. Lett.*, 2003, **5**, 2211–2213.
158. J. J. Reddick, R. Nicewonger, and T. P. Begley, *Biochemistry*, 2001, **40**, 10095–10102.
159. R. Chicharro, S. de Castro, J. L. Reino, and V. J. Arán, *European Journal of Organic Chemistry*, 2003, **2003**, 2314–2326.
160. E. Valeur and M. Bradley, *Chem. Soc. Rev.*, 2009, **38**, 606–631.
161. J. K. Chakrabarti, T. M. Hotten, I. A. Pullar, and N. C. Tye, *J. Med. Chem.*, 1989, **32**, 2573–2582.
162. L. A. Carpino, *J. Am. Chem. Soc.*, 1993, **115**, 4397–4398.
163. C. Kaneko, K. Uchiyama, M. Sato, and N. Katagiti, *Chem. Pharm. Bull.*, 1986, **34**, 3658–3671.
164. R. J. Rousseau and R. K. Robins, *Journal of Heterocyclic Chemistry*, 1965, **2**, 196–201.
165. G. E. Chivers and H. Suschitzky, *J. Chem. Soc. C*, 1971, 2867–2871.
166. R. A. Abramovitch, J. Campbell, E. E. Knaus, and A. Silhankova, *Journal of Heterocyclic Chemistry*, 1972, **9**, 1367–1371.
167. M. Schlosser, C. Bobbio, and T. Rausis, *J. Org. Chem.*, 2005, **70**, 2494–2502.
168. C. J. Woltermann and J. A. Schwindeman, *ChemInform*, 2004, **35**, no–no.
169. K. P. Peterson and R. C. Larock, *J. Org. Chem.*, 1998, **63**, 3185–3189.
170. T. Nishimura, T. Onoue, K. Ohe, and S. Uemura, *J. Org. Chem.*, 1999, **64**, 6750–6755.
171. E. M. Ferreira and B. M. Stoltz, *J. Am. Chem. Soc.*, 2001, **123**, 7725–7726.
172. S. Gowrisankar, H. Neumann, D. Gördes, K. Thurow, H. Jiao, and M. Beller, *Chemistry*, 2013, **19**, 15979–15984.
173. H. Baron, F. G. P. Remfry, and J. F. Thorpe, *J. Chem. Soc., Trans.*, 1904, **85**, 1726–1761.
174. A. Albert and J. N. Phillips, *J. Chem. Soc.*, 1956, 1294–1304.

175. T. Hagihara, T. Fujio, and K. Aisaka, *Appl. Microbiol. Biotechnol.*, 1995, **42**, 724–729.
176. S. Nakagawa, A. Igarashi, T. Ohta, T. Hagihara, T. Fujio, and K. Aisaka, *Bioscience, Biotechnology, and Biochemistry*, 1995, **59**, 694–702.
177. W. Eisenreich, K. Kemter, A. Bacher, S. B. Mulrooney, C. H. W. Willams Jr., and F. Muller, *Eur. J. Biochem.*, 2004, **271**, 1437–1452.
178. P. Nielsen, P. Rauschenbach, and A. Bacher, *Analytical Biochemistry*, 1983, **130**, 359–368.
179. P. Viñas, N. Balsalobre, C. López-Erroz, and M. Hernández-Córdoba, *Journal of Agricultural and Food Chemistry*, 2004, **52**, 1789–1794.
180. O. A. Bessey, O. H. Lowry, and R. H. Love, *Journal of Biological Chemistry*, 1949, **180**, 755–769.
181. S. Hustad, P. M. Ueland, and J. Schneede, *Clin. Chem.*, 1999, **45**, 862–868.
182. A. Serrano, S. Frago, B. Herguedas, M. Martínez-Júlvez, A. Velázquez-Campoy, and M. Medina, *Cell Biochem. Biophys.*, 2013, **65**, 57–68.
183. S. Bauer, K. Kemter, A. Bacher, R. Huber, M. Fischer, and S. Steinbacher, *J. Mol. Biol.*, 2003, **326**, 1463–1473.
184. ExPASy - http://web.expasy.org/protparam/, .
185. S. C. Gill and P. H. von Hippel, *Anal. Biochem.*, 1989, **182**, 319–326.
186. U. Krauss, V. Svensson, A. Wirtz, E. Knieps-Grünhagen, and K.-E. Jaeger, *Appl. Environ. Microbiol.*, 2011, **77**, 1097–1100.
187. A. Freeman, J. M. Woodley, and M. D. Lilly, *Nat Biotech*, 1993, **11**, 1007–1012.
188. G. J. Lye and J. M. Woodley, *Trends in Biotechnology*, 1999, **17**, 395–402.
189. D. Stüeber, H. Matile, and G. Garotta, in *Immunological Methods IV (Lefkovits, I. & Pernis, P. eds.)*, Academic Press, USA, 1990, pp. 121–152.
190. J. Scheel, K. Ziegelbauer, T. Kupke, B. M. Humbel, A. A. Noegel, G. Gerisch, and M. Schleicher, *J. Biol. Chem.*, 1989, **264**, 2832–2839.
191. M. R. Villarejo and I. Zabin, *J Bacteriol*, 1974, **120**, 466–474.
192. P. J. Farabaugh, *Nature*, 1978, **274**, 765–769.
193. J. Messing, R. Crea, and P. H. Seeburg, *Nucleic Acids Res.*, 1981, **9**, 309–321.
194. C. Yanisch-Perron, J. Vieira, and J. Messing, *Gene*, 1985, **33**, 103–119.
195. L. Olvera, A. Mendoza-Vargas, N. Flores, M. Olvera, J. C. Sigala, G. Gosset, E. Morett, and F. Bolívar, *PLoS ONE*, 2009, **4**, e7466.
196. R. Gentz and H. Bujard, *J Bacteriol*, 1985, **164**, 70–77.
197. F. W. Studier and B. A. Moffatt, *Journal of Molecular Biology*, 1986, **189**, 113–130.
198. J. M. Fernandez and J. P. Hoeffler, Eds., *Gene Expression Systems: Using Nature for the Art of Expression*, Academic Press, USA, Pages 14-29, 1st edn., 1999.
199. M. Rong, B. He, W. T. McAllister, and R. K. Durbin, *PNAS*, 1998, **95**, 515–519.
200. R. W. Siegel, S. Adkins, and C. C. Kao, *PNAS*, 1997, **94**, 11238–11243.
201. D. Imburgio, M. Rong, K. Ma, and W. T. McAllister, *Biochemistry*, 2000, **39**, 10419–10430.
202. K. A. Chapman and R. R. Burgess, *Nucleic Acids Res.*, 1987, **15**, 5413–5432.
203. K. A. Chapman, S. I. Gunderson, M. Anello, R. D. Wells, and R. R. Burgess, *Nucleic Acids Res*, 1988, **16**, 4511–4524.
204. T. Ikemura, *Mol Biol Evol*, 1985, **2**, 13–34.
205. E. E. Murray, J. Lotzer, and M. Eberle, *Nucleic Acids Res*, 1989, **17**, 477–498.
206. D. Chen and D. E. Texada, *Gene Therapy & Molecular Biology*, 2006, **10**, 1–12.
207. M. Kasahara, M. Torii, A. Fujita, and K. Tainaka, *J. Biol. Chem.*, 2010, **285**, 34765–34772.
208. Y. Ogura, A. Komatsu, K. Zikihara, T. Nanjo, S. Tokutomi, M. Wada, and T. Kiyosue, *J. Plant Res.*, 2008, **121**, 97–105.
209. OPTIMIZER, http://genomes.urv.es/OPTIMIZER/Form.php, .

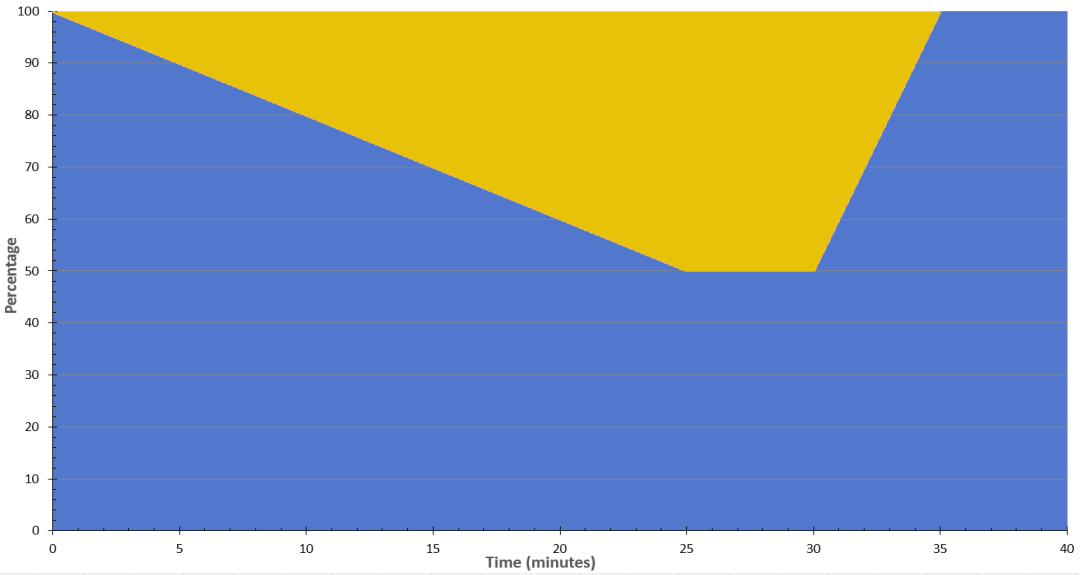
210. P. Puigbo, E. Guzman, A. Romeu, and S. Garcia-Vallve, *Nucleic Acids Res*, 2007, **35**, W126–W131.
211. D. Strickland, Y. Lin, E. Wagner, C. M. Hope, J. Zayner, C. Antoniou, T. R. Sosnick, E. L. Weiss, and M. Glotzer, *Nat Methods*, 2012, **9**, 379–384.
212. D. M. Dietz, H. Sun, M. K. Lobo, M. E. Cahill, B. Chadwick, V. Gao, J. W. Koo, M. S. Mazei-Robison, C. Dias, I. Maze, D. Damez-Werno, K. C. Dietz, K. N. Scobie, D. Ferguson, D. Christoffel, Y. Ohnishi, G. E. Hodes, Y. Zheng, R. L. Neve, K. M. Hahn, S. J. Russo, and E. J. Nestler, *Nat. Neurosci.*, 2012, **15**, 891–896.
213. M. P. Weiner, G. L. Costa, W. Schoettlin, J. Cline, E. Mathur, and J. C. Bauer, *Gene*, 1994, **151**, 119–123.
214. A. Chien, D. B. Edgar, and J. M. Trela, *J Bacteriol*, 1976, **127**, 1550–1557.
215. K. S. Lundberg, D. D. Shoemaker, M. W. Adams, J. M. Short, J. A. Sorge, and E. J. Mathur, *Gene*, 1991, **108**, 1–6.
216. F. C. Lawyer, S. Stoffel, R. K. Saiki, S. Y. Chang, P. A. Landre, R. D. Abramson, and D. H. Gelfand, *PCR Methods Appl.*, 1993, **2**, 275–287.
217. K. R. Tindall and T. A. Kunkel, *Biochemistry*, 1988, **27**, 6008–6013.
218. J. Cline, J. C. Braman, and H. H. Hogrefe, *Nucleic Acids Res.*, 1996, **24**, 3546–3551.
219. I. H. M. van Stokkum, M. Gauden, S. Crosson, R. van Grondelle, K. Moffat, and J. T. M. Kennis, *Photochemistry and Photobiology*, 2011, **87**, 534–541.
220. J. M. Christie, S. B. Corchnoy, T. E. Swartz, M. Hokenson, I.-S. Han, W. R. Briggs, and R. A. Bogomolni, *Biochemistry*, 2007, **46**, 9310–9319.
221. F. J. Castellino and R. Barker, *Biochemistry*, 1968, **7**, 4135–4138.
222. S. Lapanje, *Biochimica et Biophysica Acta (BBA) - Protein Structure*, 1971, **243**, 349–356.
223. J. A. Gordon, *Biochemistry*, 1972, **11**, 1862–1870.
224. B. D. Zoltowski, A. I. Nash, and K. H. Gardner, *Biochemistry*, 2011, **50**, 8771–8779.
225. Y. Nakasone, T. Eitoku, D. Matsuoka, S. Tokutomi, and M. Terazima, *J. Mol. Biol.*, 2007, **367**, 432–442.
226. M. T. A. Alexandre, R. van Grondelle, K. J. Hellingwerf, and J. T. M. Kennis, *Biophys J*, 2009, **97**, 238–247.
227. A. Pfeifer, T. Mathes, Y. Lu, P. Hegemann, and T. Kottke, *Biochemistry*, 2010, **49**, 1024–1032.
228. E. Sikorska, I. Khmelinskii, A. Komasa, J. Koput, L. F. V. Ferreira, J. R. Herance, J. L. Bourdelande, S. L. Williams, D. R. Worrall, M. Insińska-Rak, and M. Sikorski, *Chemical Physics*, 2005, **314**, 239–247.
229. B. Ketterer, B. Coles, and D. J. Meyer, *Environ Health Perspect*, 1983, **49**, 59–69.
230. B. Coles, B. Ketterer, F. A. Beland, and F. F. Kadlubar, *Carcinogenesis*, 1984, **5**, 917–920.
231. Clayden, Greeves, Warren, and Wothers, *Organic Chemistry*, Oxford University Press, Page 584, 6th edn., 2001.
232. B. F. Coles and F. F. Kadlubar, *Biofactors*, 2003, **17**, 115–130.
233. S. Salzmann, V. Martinez-Junza, S. E. Braslavsky, C. M. Marian, and W. Gartner, *J. Phys. Chem. A*, 2009, **113**, 9365–9375.
234. A. M. Ellis, M. Feher, and T. G. Wright, *Electronic and Photoelectron Spectroscopy: Fundamentals and Case Studies*, Cambridge University Press, Page 125, 2005.
235. P. Atkins and J. de Paula, *Elements of Physical Chemistry*, W. H. Freeman, Page 339, Fourth Edition., 2009.
236. P. Atkins and J. de Paula, *Elements of Physical Chemistry*, W. H. Freeman, Page 519, Fourth Edition., 2009.
237. B. Schierling and A. Pingoud, *Bioconjugate Chem.*, 2012, **23**, 1105–1109.

238. A. Dyer, *An introduction to zeolite molecular sieves*, John Wiley & Sons Australia, Limited, 1988.
239. R. Szostak, *Molecular Sieves: Principles of Synthesis and Identification*, Springer, 1998.
240. H. Lund and J. Bjerrum, *Berichte der deutschen chemischen Gesellschaft (A and B Series)*, 1931, **64**, 210–213.
241. E. Friedman, T. J. Gill, and P. Doty, *J. Am. Chem. Soc.*, 1961, **83**, 4050–4053.
242. D. R. Burfield and R. H. Smithers, *J. Org. Chem.*, 1978, **43**, 3966–3968.
243. U. K. Laemmli, *Nature*, 1970, **227**, 680–685.
244. C. Mulhardt, *Molecular Biology and Genomics*, AP Press / Elsevier, Page 120, 1st edn., 2007.
245. C. Mulhardt, *Molecular Biology and Genomics*, AP Press / Elsevier, Page 125, 1st edn., 2007.
246. K. Mullis, F. Faloona, S. Scharf, R. Saiki, G. Horn, and H. Erlich, *Cold Spring Harb. Symp. Quant. Biol.*, 1986, **51 Pt 1**, 263–273.
247. R. K. Saiki, D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich, *Science*, 1988, **239**, 487–491.
248. (first) PrimerX, http://www.bioinformatics.org/primerx/, .
249. T. Sato and K. Fujiwara, .
250. H. Yan and J.-Y. Liu, *Journal of Chinese Pharmaceutical Sciences*, 2011, **20**.
251. D. R. Light, C. Walsh, and M. A. Marletta, *Analytical Biochemistry*, 1980, **109**, 87–93.

Appendix 1 – Graphical Representation of HPLC gradients

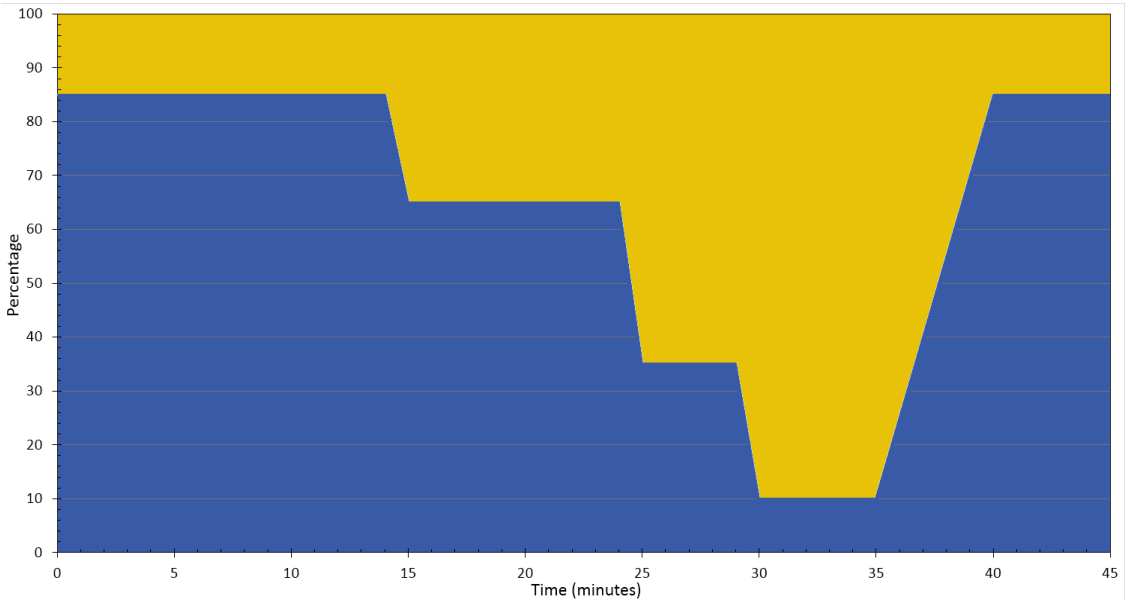
Formation of Riboflavin Analogues

The gradient used for the purification of riboflavin analogues (as described in section 6.5) is shown below. The aqueous mobile phase is shown in blue; acetonitrile is shown in yellow.



Formation of FMN Analogues

The gradient used for the purification of the enzymatically formed flavin cofactors (as described in section 7.1.5) is shown below. The aqueous mobile phase is shown in blue; acetonitrile is shown in yellow.



Appendix 2 – Codon Usage Report, Generated by OPTIMIZER

The gene encoding AsLOV2(WT) was compared to *Escherichia coli* K12, to identify relative codon optimisation. The AsLOV2(WT) gene was referred to as the "Query".

The OPTIMIZER tool generated an Optimized [sic] sequence, an Alignment Report of the Query and Optimized DNA sequence, and Individual Codon Optimization Summary.

	DNA Sequence	CAI	% GC	% AT
Query	GAATTTCTTGCTACTACACTTGAACGTATTGAGAAGAAC TTTGTCATTACTGACCCACGTTTGCCAGATAATCCCATT ATCTTCGCGTCCGATAGTTTCTTGCAGTTGACAGAATAT TCGCGAGAAGAAATTCTGGGTCGTAACTGCCGTTTTCTT CAAGGTCCTGAAACCGATCGCGCGACAGTGCGCAAAA TTCGTGATGCCATCGATAACCAAACAGAGGTCACTGTA CAGCTGATTAATTATACAAAGAGTGGTAAAAAGTTCTG GAACCTCTTTCACTTGCAGCCTATGCGTGATCAGAAGG GTGATGTCCAGTACTTTATTGGTGTCCAGTTGGATGGT ACCGAACATGTCCGTGATGCGGCCGAGCGTGAGGGTG TCATGCTGATTAAGAAAACTGCAGAAAATATTGATGAG GCGGCAAAAGAACTTCCAGATGCTAATCTGCGTCCAGA GGATTTGTGGGCTAACTAA	0.446	44.7	55.3
Optimized	GAATTCCTGGCGACCACCCTGGAACGTATCGAAAAAAA CTTCGTTATCACCGACCCGCGTCTGCCGGACAACCCGAT CATCTTCGCGTCTGACTCTTTCCTGCAGCTGACCGAATA CTCTCGTGAAGAAATCCTGGGTCGTAACTGCCGTTTCCT GCAGGGTCCGGAAACCGACCGTGCGACCGTTCGTAAA ATCCGTGACGCGATCGACAACCAGACCGAAGTTACCGT TCAGCTGATCAACTACACCAAATCTGGTAAAAAATTCTG GAACCTGTTCCACCTGCAGCCGATGCGTGACCAGAAAG GTGACGTTCAGTACTTCATCGGTGTTCAGCTGGACGGT ACCGAACACGTTCGTGACGCGGCGGAACGTGAAGGTG TTATGCTGATCAAAAAAACCGCGGAAAACATCGACGAA GCGGCGAAAGAACTGCCGGACGCGAACCTGCGTCCGG AAGACCTGTGGGCGAACTAA	1.000	53.7	46.3

Abbreviations:

CAI	Codon Adaptation Index
% GC	Percentage of G+C nucleobases
% AT	Percentage of A+T nucleobases
|	Unchanged Nucleotide
*	Transversion Change (Purine <-> Pyrimidine)
#	Transitional Change (Purine <-> Purine or Pyrimidine <-> Pyrimidine)

Alignment Report

```
Query  GAA TTT  CTT  GCT  ACT  ACA CTT GAA CGT ATT GAG AAG AAC TTT GTC ATT ACT GAC CCA CGT
       ||| ||#  ||*  ||*  ||#  ||* ||* ||| ||| ||# ||# ||#||| ||#||# ||# ||# ||| ||# |||
Optim. GAA TTC  CTG GCG ACC ACC CTG GAA CGT ATC GAA AAA AAC TTC GTT ATC ACC GAC CCG CGT

Query  TTG CCA GAT  AAT CCC ATT ATC TTC GCG TCC GAT AGT TTC TTG CAG  TTG ACA GAA TAT TCG
       #|| ||#  ||#  ||#  ||*  ||# ||| ||| ||| ||# ||#  **| ||| #||  |||  #|| ||* ||| ||#  ||*
Optim. CTG CCG GAC AAC CCG ATC ATC TTC GCG TCT GAC TCT TTC CTG CAG CTG ACC GAA TAC TCT

Query  CGA GAA GAA ATT CTG  GGT CGT AAC TGC CGT TTT  CTT CAA  GGT CCT GAA ACC GAT CGC GCG
       ||*  ||| |||  ||# |||  ||| |||  ||| ||| ||# ||* ||#  |||  ||* ||| ||| ||#  ||#  |||
Optim. CGT GAA GAA ATC CTG GGT CGT AAC TGC CGT TTC CTG CAG GGT CCG GAA ACC GAC CGT GCG

Query  ACA GTG CGC AAA ATT CGT GAT GCC ATC GAT  AAC CAA ACA GAG GTC ACT GTA CAG CTG ATT
       ||* ||* ||#  |||  ||# ||| ||#  ||* ||| ||# |||  ||#  ||* ||#  ||#  ||#  ||* |||  |||  ||#
Optim. ACC GTT CGT  AAA ATC CGT GAC GCG ATC GAC AAC CAG ACC GAA GTT ACC GTT CAG CTG ATC

Query  AAT TAT  ACA AAG AGT GGT AAA AAG TTC TGG AAC CTC TTT CAC TTG  CAG CCT ATG CGT GAT
       ||#  ||#  ||*  ||#  **| |||  ||| ||# ||| ||| |||  ||* ||# ||| #||  |||  ||* |||  |||  ||#
Optim. AAC TAC ACC AAA TCT  GGT AAA AAA TTC TGG AAC CTG TTC CAC CTG CAG CCG ATG CGT GAC

Query  CAG AAG GGT GAT GTC CAG TAC TTT ATT GGT GTC CAG TTG GAT  GGT ACC GAA CAT GTC CGT
       |||  ||#  |||  ||# ||#  |||  |||  ||# ||# |||  ||#  |||  #||  ||#  |||  |||  |||  ||#  ||#  |||
Optim. CAG AAA GGT GAC GTT CAG TAC TTC ATC GGT GTT CAG CTG GAC GGT ACC GAA CAC GTT CGT

Query  GAT GCG GCC GAG CGT GAG GGT GTC ATG CTG ATT AAG AAA ACT GCA GAA AAT ATT  GAT GAG
       ||#  |||  ||*  ||#  |||  ||#  |||  ||#  |||  |||  ||#  ||#  |||  ||#  ||#  |||  ||#  ||#  ||#  ||#
Optim. GAC GCG GCG GAA CGT GAA GGT GTT ATG CTG ATC AAA AAA ACC GCG GAA AAC ATC GAC GAA

Query  GCG GCA AAA GAA CTT  CCA  GAT  GCT  AAT  CTG CGT CCA GAG GAT TTG  TGG GCT AAC TAA
       |||  ||#  |||  |||  ||*  ||#  ||#  ||*  ||#  |||  |||  ||#  ||#  ||#  #||  |||  ||*  |||  |||
Optim. GCG GCG AAA GAA CTG CCG GAC  GCG AAC  CTG CGT CCG GAA GAC CTG TGG GCG AAC TAA
```

Individual Codon Optimization Summary

Each codon of the standard genetic code is listed, with the amino acid encoded indicated in brackets. Usage of each codon in the query sequence and optimized sequence is shown.

Codons	Query	Optimized	Codons	Query	Optimized	Codons	Query	Optimized	Codons	Query	Optimized
GCA (A)	2	0	GCC (A)	2	0	GCG (A)	4	11	GCT (A)	3	0
TGC (C)	1	1	TGT (C)	0	0	GAC (D)	1	13	GAT (D)	12	0
GAA (E)	9	15	GAG (E)	6	0	TTC (F)	3	8	TTT (F)	5	0
GGA (G)	0	0	GGC (G)	0	0	GGG (G)	0	0	GGT (G)	7	7
CAC (H)	1	2	CAT (H)	1	0	ATA (I)	0	0	ATC (I)	2	11
ATT (I)	9	0	AAA (K)	4	9	AAG (K)	5	0	TTA (L)	0	0
TTG (L)	6	0	CTA (L)	0	0	CTC (L)	1	0	CTG (L)	4	15
CTT (L)	4	0	ATG (M)	2	2	AAC (N)	5	9	AAT (N)	4	0
CCA (P)	4	0	CCC (P)	1	0	CCG (P)	0	7	CCT (P)	2	0
CAA (Q)	2	0	CAG (Q)	6	8	AGA (R)	0	0	AGG (R)	0	0
CGA (R)	1	0	CGC (R)	2	0	CGG (R)	0	0	CGT (R)	9	12
AGC (S)	0	0	AGT (S)	2	0	TCA (S)	0	0	TCC (S)	1	0
TCG (S)	1	0	TCT (S)	0	4	ACA (T)	5	0	ACC (T)	2	11
ACG (T)	0	0	ACT (T)	4	0	GTA (V)	1	0	GTC (V)	6	0
GTG (V)	1	0	GTT (V)	0	8	TGG (W)	2	2	TAC (Y)	1	3
TAT (Y)	2	0	TAA (.)	1	1	TGA (.)	0	0	TAG (.)	0	0