

# **Mechanistic and Structural Studies of Light-Oxygen- Voltage (LOV) Domain Proteins**

A thesis submitted to Cardiff University  
for the degree of Doctor of Philosophy by:

**Mindaugas Edvardas Kalvaitis**

Supervisor: Professor Rudolf K. Allemann



Cardiff University  
School of Chemistry  
September 2019





## Abstract

In this work, structural and mechanistic understanding of LOV domain proteins has been provided by applying biophysical techniques including protein X-ray crystallography, nuclear magnetic resonance (NMR) and circular dichroism (CD). LOV domains are photosensory modules where exposure to blue light triggers the formation of a reversible cysteinyl-flavin covalent photoadduct. The formation of this chemical bond has been studied extensively, but how this covalent adduct governs residue rearrangements within the flavin binding pocket and the allosteric regions of the LOV domain core remained elusive. By capturing *Ochromonas danica* Aureochrome1a LOV (*OdAu1a<sub>LOV</sub>*) domain in three distinct states, structural evidence for a novel residue rearrangement has been provided. X-ray crystal structures of dark, illuminated (where dark grown crystals were photoactivated prior crystal harvesting) and light grown *OdAu1a<sub>LOV</sub>* revealed three distinct conformations of a highly conserved glutamine 293 residue. Reflecting on the three distinct states, a mechanism proposing how this residue links the chromophore binding pocket to the allosteric A' $\alpha$  helix has been proposed. This is the first record of glutamine 293 swinging away from the flavin binding pocket being captured by X-ray crystallography. This transient swing changes the hydrogen bond network formed upon photoadduct formation and could be crucial to facilitate large scale structural changes. This key glutamine residue is highly conserved in other LOV domain proteins, hence understanding the transition between these three distinct conformations provides novel insights to assist the rational design of novel optogenetic tools.

Furthermore, native flavin mononucleotide (FMN) cofactor was replaced with 5-deazaflavin mononucleotide (5dFMN) analogue to understand *OdAu1a<sub>LOV</sub>* structural changes further. The current standing hypothesis suggests that the formation of the covalent cysteinyl-FMN photoadduct upon illumination results in the protonation of the FMN-N5. It has been suggested that this protonation state change, from unprotonated in the dark state and protonated in the lit state, triggers structural changes *via* a conserved glutamine residue. Considering that 5dFMN is a structural FMN analogue where N5 is replaced with a carbon (C5), no protonation is possible in this position and 5dFMN should be inactive if N5 protonation is crucial for the activation. Upon photoactivation of 5dFMN *OdAu1a<sub>LOV</sub>*, a thermally stable cysteinyl-5dFMN covalent adduct was generated and high-resolution crystal structures provide conclusive evidence that, as for FMN, this adduct forms at the C4a position also triggering some residue rearrangements within the flavin binding pocket. Nonetheless, CD and NMR spectra revealed only partial light

responsiveness, yet no major structural changes were evident for 5dFMN reconstituted *OdAu1a<sub>LOV</sub>* post photoactivation. These results therefore suggest that N5 protonation is important to lock the protein in the lit state conformation.

To probe functionality of 5dFMN further, the cofactor was introduced into *Avena sativa* LOV2 (*AsLOV2*), another LOV domain protein. Similarly to *OdAu1a<sub>LOV</sub>*, partial light responsiveness was evident by CD studies suggesting that 5dFMN displayed light dependent structural changes. This was also studied further by <sup>19</sup>F NMR through labelling *AsLOV2* with 5-fluoro-L-tryptophan, a fluorine (<sup>19</sup>F) probe, allowing direct observation of J $\alpha$  unfolding. NMR results suggested that 5dFMN *AsLOV2* experienced changes in the cofactor chemistry including interruption of conjugated isoalloxazine ring system, which from these observations was hypothesised to contribute in achieving the active lit state conformation.

Finally, conformational changes of *OdAu1a<sub>LOV</sub>* were studied by <sup>19</sup>F NMR in order to investigate the importance of allosteric A' $\alpha$  and J $\alpha$  helices in achieving the lit state conformation. Results highlighted that the removal of A' $\alpha$  results in insoluble protein, indicating that this auxiliary helix is crucial, whilst the truncation of the J $\alpha$  helix resulted in a functional protein, indicating that this extension is not required.

## Acknowledgments

Most importantly, I would like to express my gratitude to my supervisor, Professor Rudolf Konrad Allemann (Ruedi), for allowing me to join his research group as a postgraduate student. Throughout my studies, his passion and immense knowledge of mechanistic enzymology has inspired me to proceed with a scientific career to understand enzyme catalysis further and apply this understanding in development of novel therapeutics. Furthermore, I am also extremely grateful for his patience and understanding throughout my studies, thank you very much, I really do appreciate it.

I would also like to thank BBSRC for funding research grant BB/M006158/1, Light-responsive building blocks for synthetic biology. This funding, applied by Ruedi, allowed me to proceed with a career in structural molecular biology allowing me tackle many different scientific problems.

Besides my supervisor, I would also like to thank Drs Luke Johnson (Luke) and Robert Mart (Rob) for all of the scientific discussions and help throughout my PhD. I would like to thank Luke for his help with the NMR experiments including fancy blue light NMR experiments and useful scientific discussions we had in the mornings. I would also like to thank Rob for helping me with the scientific writing. Additionally, I would also like to thank Dr Pierre Rizkallah (Pierre) for all the trips to DLS and all the help with X-ray crystallography, scientific discussions and training me in the field I am currently working in.

I would also like to thank Cardiff Chemical Biology team including academics Drs Yi Jin, Louis Luk and Yu-Hsuan Tsai, and Professor Nigel Richards for all the scientific discussions. Furthermore I would like to thank Postdocs and PhD students working in the Chemical Biology team for all the discussions and chats. Special thank you goes to Thomas Williams for all the help with LC-MS and Dr Robert Jenkins for all the help with NMR.

My sincere thanks also goes to everybody working in the Allemann's group including Postdocs and PhD students, past and present. I would like to thank everybody in the 'Photobiochem' group including Luke, Rob, Martin and Raquel. All those coffees we had made my PhD journey a lot easier. 'DHFR Boys' including Antonio, Adura, RobH, Alan and Gareth for all the chats and discussions, science is amazing is it not? 'The Terpene People' including 'The Boys' (Victor, Mateusz and Chris) for all the jokes in the labs, 'The Single Cell Organism' (Florence, Alice, Jenneh and Gwawr) for all the support and Prabhakar for scientific discussions. I really do appreciate all of these terpene talks. To all

the people in the Allemann's group I really do hope that my silly jokes and comments did not annoy you too much!

My special thanks goes to amazing people I had a pleasure to meet in Cardiff and amazing friendships I made. A very big and special thank you goes to my best friend Lady Raquel Cruz-Samperio (Raquelita), thank you so much for always being there for me. Two years we lived together will be remembered forever! Also a big thank you goes to Magda, Mateusz and Florence, three amazing people that I had a pleasure to become close friends with.

Lastly, I would like to thank my amazing family for all the support and encouragement. You really have helped me a lot especially my mum Violeta.

## Table of Content

<b>ABSTRACT .....</b>	<b>III</b>
<b>ACKNOWLEDGMENTS .....</b>	<b>V</b>
<b>TABLE OF CONTENT .....</b>	<b>VII</b>
<b>LIST OF ABBREVIATIONS.....</b>	<b>XII</b>
<b>LIST OF AMINO ACIDS .....</b>	<b>XV</b>
<b>LIST OF FIGURES .....</b>	<b>XVI</b>
<b>LIST OF TABLES.....</b>	<b>XXII</b>
<b>1 INTRODUCTION.....</b>	<b>1</b>
1.1 LIGHT SENSING .....	2
1.2 BLUE LIGHT SENSING USING FLAVINS AS A CHROMOPHORE.....	2
1.3 CRYPTOCHROMES AS LIGHT SENSORS .....	4
1.4 BLUE LIGHT SENSING UTILISING FLAVIN ADENINE DINUCLEOTIDE (BLUF) DOMAIN.....	6
1.5 LIGHT-OXYGEN-VOLTAGE (LOV) DOMAIN PROTEINS.....	9
1.6 ENGINEERED LOV PHOTORECEPTORS .....	14
1.7 LOV DOMAIN LIGHT SENSING, STRUCTURAL AND MECHANISTIC INSIGHTS.....	19
1.8 PROJECT AIMS AND PURPOSE .....	22
1.9 REFERENCES.....	24
<b>2 GENERAL MATERIALS AND METHODS .....</b>	<b>39</b>
2.1 PREPARATION OF LIQUID AND SOLID CULTURE MEDIA .....	39
2.1.1 <i>Lysogeny Broth (LB) Liquid Media Preparation .....</i>	<i>39</i>
2.1.2 <i>Defined Minimal Media (M9) Preparation.....</i>	<i>40</i>
2.1.3 <i>Preparation of Super Optimal Broth (SOB) and Super Optimal Broth with Catabolite Repression (SOC) Media .....</i>	<i>42</i>
2.1.4 <i>Preparation of Terrific Broth (TB) Media.....</i>	<i>42</i>
2.1.5 <i>Preparation of Autoinduction Media – LB Based .....</i>	<i>43</i>
2.1.6 <i>Preparation of Solid Luria Broth (LB) and Super Optimal Broth with Catabolite Repression (SOC) Agar Plates.....</i>	<i>44</i>
2.2 PREPARATION OF COMPETENT CELLS, TRANSFORMATION AND RECOMBINANT PROTEIN EXPRESSION	45
2.2.1 <i>Escherichia coli (E. coli) DNA Work Strains.....</i>	<i>45</i>
2.2.2 <i>Escherichia coli (E. coli) Protein Overexpression Strains .....</i>	<i>46</i>

2.2.3	<i>Preparation of Competent Cells.....</i>	46
2.2.4	<i>Preparation of Supercompetent Cells.....</i>	47
2.2.5	<i>Transformation, Preparation of Cell Culture Inoculate and Glycerol Stocks.....</i>	48
2.2.6	<i>Preparation of BL21-AI pEVOL-F<sub>n</sub>YR-E3 Cells.....</i>	49
2.2.7	<i>Transformation of Cloning Constructs, Preparation of Overnight Growths and Sequencing of Constructs.....</i>	49
2.2.8	<i>Expression of Recombinant Proteins in M9, LB, TB and Autoinduction Medias, Cell Culture Harvesting and Cell Paste Storage.....</i>	50
2.3	PREPARATION, MAINTENANCE AND REGENERATION OF GRAVITY AND FAST PROTEIN LIQUID CHROMATOGRAPHY (FPLC) COLUMNS.....	50
2.3.1	<i>Immobilised Metal Affinity Chromatography (IMAC) Gravity Drip Column and Prepacked Column Preparation and Regeneration.....</i>	51
2.3.2	<i>Resource Q (ResQ) Column Preparation and Regeneration.....</i>	51
2.3.3	<i>Gel Filtration Chromatography Purification.....</i>	52
2.4	SODIUM DODECYL SULPHATE POLYACRYLAMIDE GEL ELECTROPHORESIS (SDS-PAGE) RUNNING BUFFER, SAMPLE LOADING BUFFER AND POLYACRYLAMIDE GEL PREPARATION.....	53
2.4.1	<i>Preparation of Stock Solutions and Reducing SDS-PAGE Denaturing Loading Dye.....</i>	53
2.4.2	<i>Preparation of SDS-PAGE Loading Dye.....</i>	53
2.4.3	<i>Preparation of Protein Samples for SDS-PAGE Analysis.....</i>	54
2.4.4	<i>Preparation of Acrylamide SDS-PAGE Stacking and Resolving Gels.....</i>	54
2.4.5	<i>Tricine SDS-PAGE Preparation.....</i>	55
2.4.6	<i>Preparation of Coomassie Blue Stain – G250 and Staining of Protein SDS-PAGE Gels.....</i>	56
2.5	CHAPTER 3 MATERIALS AND METHODS.....	57
2.5.1	<i>OdAu1a construct generation.....</i>	57
2.5.2	<i>Protein expression and purification.....</i>	58
2.5.3	<i>FMN to 5dFMN chromophore exchange.....</i>	60
2.5.4	<i>Liquid chromatography-mass spectrometry.....</i>	61
2.5.5	<i>NMR spectroscopy.....</i>	63
2.5.6	<i>Circular dichroism spectroscopy.....</i>	63
2.5.7	<i>Analytical gel filtration chromatography.....</i>	63
2.5.8	<i>Electrophoretic mobility shift assays.....</i>	64
2.5.9	<i>Crystallisation of OdAu1a<sub>LOV</sub>.....</i>	65
2.6	CHAPTER 4 MATERIALS AND METHODS.....	66

2.6.1	<i>Protein Constructs</i> .....	66
2.6.2	<i>Protein Construct Generation</i> .....	67
2.6.3	<i>Protein Expression and Purification</i> .....	67
2.6.4	<i>Cofactor exchange</i> .....	68
2.6.5	<i>UV-Vis Characterization</i> .....	69
2.6.6	<i>Circular dichroism experiments</i> .....	70
2.6.7	<i>Nuclear Magnetic Resonance Experiments</i> .....	70
2.6.8	<i>Protein analysis by Liquid Chromatography Mass Spectroscopy (LC-MS)</i> .....	71
2.7	CHAPTER 5 MATERIALS AND METHODS .....	71
2.7.1	<i>Protein construct generation and plasmid isolation</i> .....	71
2.7.2	<i>Protein expression and storage of culture pellets</i> .....	74
2.7.3	<i>Protein purification</i> .....	76
2.7.4	<i>UV-Visible light experiments</i> .....	77
2.7.5	<i>Protein production and purification for X-ray crystallography</i> .....	77
2.7.6	<i>NMR experiments</i> .....	78
2.7.7	<i>Liquid Chromatography Mass Spectroscopy (LC-MS) experiment</i> .....	79
2.7.8	<i>Analytical gel filtration chromatography experiments</i> .....	79
2.7.9	<i>Circular dichroism experiments</i> .....	80
<b>3</b>	<b>STRUCTURAL STUDIES OF <i>OCHROMONAS DANICA</i> AUREOCHROME1A RECONSTITUTED WITH 5DFMN</b> .....	<b>82</b>
3.1	INTRODUCTION.....	82
3.2	RECONSTITUTION OF <i>OCHROMONAS DANICA</i> AUREOCHROME1A WITH 5-DEAZA FLAVIN MONONUCLEOTIDE (5DFMN) .....	86
3.3	CD AND NUCLEAR MAGNETIC RESONANCE INVESTIGATIONS OF FMN AND 5DFMN AUREOCHROME1A LOV DOMAIN STRUCTURAL CHANGES .....	90
3.4	SOLUTION CHARACTERISATION OF AUREOCHROME1A DIMERISATION AND DNA BINDING .....	94
3.5	STRUCTURAL STUDIES OF AU1ALOVS WITH .....	99
3.6	DISCUSSION AND CONCLUSIONS .....	106
3.7	REFERENCES.....	111
3.8	APPENDIX.....	116
<b>4</b>	<b>UNDERSTANDING OF LOV DOMAIN SIGNALLING – TESTING PROPOSED HYPOTHESIS IN AVENA SATIVA PHOTOTROPIN LOV2 DOMAIN (ASLOV2)</b> .....	<b>125</b>
4.1	INTRODUCTION.....	125

4.2	UV-VISIBLE, CIRCULAR DICHROISM AND NUCLEAR MAGNETIC RESONANCE STUDIES OF NATIVE AND <sup>19</sup> F ASLOV2.....	128
4.3	INVESTIGATIONS OF ASLOV2 STRUCTURAL CHANGES WITH ELONGATED J $\alpha$ HELIX .....	137
4.4	CONCLUSION AND DISCUSSION .....	141
4.5	REFERENCES .....	144
4.6	APPENDIX .....	148
<b>5</b>	<b>IN SOLUTION STRUCTURAL STUDIES OF <i>OCHROMONAS DANICA</i> FULL LENGTH</b>	
	<b>AUREOCHROME1A AND ISOLATED LOV DOMAIN .....</b>	<b>161</b>
5.1	INTRODUCTION .....	161
5.2	ASSESSMENT OF <i>OCHROMONAS DANICA</i> AUREOCHROME1A CRYSTAL STRUCTURES TO GUIDE CONSTRUCT DESIGN .....	163
5.3	STRUCTURAL STUDIES OF 5-FLUORO-L-TRYPTOPHAN LABELLED <i>OCHROMONAS DANICA</i> AUREOCHROME1A AND AUREOCHROME1A LOV DOMAIN CONSTRUCTS .....	167
5.3.1	<i>Labelling of OdAu1a and OdAu1a<sub>LOV</sub> Constructs with 5-fluoro-L-tryptophan.....</i>	167
5.3.2	<i>Structural Investigations of <math>\Delta</math>J<math>\alpha</math>OdAu1a<sub>LOV</sub> Construct .....</i>	170
5.4	STRUCTURAL STUDIES OF AUREOCHROME1A LOV DOMAIN LABELLED WITH 3,5-DIFLUORO-L- TYROSINE.....	175
5.5	DISCUSSION AND CONCLUSIONS.....	184
	IN THIS WORK, <sup>19</sup> F LABELLING AND CONSEQUENT NMR CHEMICAL SHIFT INVESTIGATIONS WERE USED TO PINPOINT STRUCTURAL CHANGES IN THE SOLUTION FOR <i>OdAu1a<sub>LOV</sub></i> . AS DISCUSSED FOR ASLOV2 IN CHAPTER 4, <sup>19</sup> F NMR PROVIDES AN ALTERNATIVE METHODOLOGY TO STUDY OF LIGHT DEPENDENT LOV DOMAIN CONFORMATIONAL CHANGES.....	184
5.6	REFERENCES .....	188
5.7	APPENDIX .....	191
<b>6</b>	<b>FUTURE WORK .....</b>	<b>208</b>
6.1	CHAPTER 3 FUTURE WORK .....	208
6.2	CHAPTER 4 FUTURE WORK .....	209
6.3	CHAPTER 5 FUTURE WORK .....	210





## List of Abbreviations

5dFMN	5-deaza-flavin mononucleotide
AC	Adenylyl cyclase
ADP	Adenosine diphosphate
AMP	Adenosine monophosphate
AMU	Atomic mass unit
ATP	Adenosine triphosphate
bp	Base pair
BLUF	Blue Light Using FAD (flavin adenine dinucleotide)
bZIP	Basic leucine zipper
°C	Degrees Celsius
cAMP	Cyclic adenosine monophosphate
cGMP	Cyclic guanosine monophosphate
CRY	Cryptochrome
CV	Column volumes
Da	Dalton, atomic mass unit
diH <sub>2</sub> O	Deionised water
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNApol	Deoxyribonucleic acid template dependent deoxyribonucleic acid polymerase
dNTP	Deoxyribose nucleotide triphosphate
DTT	Dithiothreitol
E. coli	Escherichia coli
EDTA	Ethylenediaminetetraacetic acid
EDO	1,2-ethanediol/ ethylene glycol
EPR	Electron paramagnetic resonance
EtOH	Ethanol
FAD	Flavin adenine dinucleotide
FMN	Flavin mononucleotide
FRET	Förster-resonance energy transfer
FT	Fourier transform
g	Mass unit gram
GDP	Guanosine diphosphate
GMP	Guanosine monophosphate
GTP	Guanosine triphosphate
h	Hour
HLH	Helix-loop-helix

HTH	Helix-turn-helix
Hz	Hertz
IMAC	Immobilised metal affinity chromatography such as Ni <sup>2+</sup> -NTA
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IR	Infrared
kDa	Kilodalton
L	Litre
λ	Wavelength
LB	Lysogeny broth
LC	Liquid chromatography
LC-MS	Liquid chromatography mass spectrometry
LED	Light emitting diode
LOV	Light-Oxygen-Voltage
M	Molar
mM	Millimolar
μM	Micromolar
mg	Milligram
μg	Microgram
mL	Millilitre
μL	Microlitre
min	Minute
MS	Mass spectrometry
NaCl	Sodium chloride
NaOH	Sodium hydroxide
nm	Nanometre
NMR	Nuclear magnetic resonance
PAGE	Polyacrylamide gel electrophoresis
PAS	Per-ARNT-Sim (Period-Aryl hydrocarbon receptor nuclear translocator-Single minded)
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
ppm	Parts per million
RNA	Ribonucleic acid
RNApol	Deoxyribonucleic acid template dependent ribonucleic acid polymerase
s	Second
SAXS	Small angle X-ray scattering
SDM	Site directed mutagenesis
SDS	Sodium dodecyl sulphate

SEC	Size exclusion chromatography
TAE	Tris.Acetate EDTA
TAG	Amber stop codon
TB	Terrific broth
TCEP	Tris(2-carboxyethyl) phosphine hydrochloride
TEV	Tobacco etch virus
TFA	Trifluoro acetate
tRNA	Transfer ribonucleic acid
UV-Vis	Ultraviolet-visible
v/v	Volume by volume
w/v	Weight by volume

## List of Amino Acids

L-Alanine	Ala	A
L-Cysteine	Cys	C
L-Aspartic acid	Asp	D
L-Glutamic acid	Glu	E
L-Phenylalanine	Phe	F
L-Glycine	Gly	G
L-Histidine	His	H
L-Isoleucine	Ile	I
L-Lysine	Lys	K
L-Leucine	Leu	L
L-Methionine	Met	M
L-Asparagine	Asn	N
L-Proline	Pro	P
L-Glutamine	Gln	Q
L-Arginine	Arg	R
L-Serine	Ser	S
L-Threonine	Thr	T
L-Valine	Val	V
L-Tryptophan	Trp	W
L-Tyrosine	Tyr	Y
5-Fluoro-L-Tryptophan	5fW	
3,5-Difluoro-L-Tyrosine	2fY	

## List of Figures

Figure 1.1 Structure of riboflavin, flavin mononucleotide and flavin adenine dinucleotide.....	1
Figure 1.2 Explanation of two component systems. ....	3
Figure 1.3 Structure of DNA photolyase. ....	4
Figure 1.4 Structure and photochemical mechanism of cryptochrome proteins....	5
Figure 1.5 Structure and Mechanism of BLUF Domain Proteins.....	7
Figure 1.6 PAS domain structure and function. ....	10
Figure 1.7 Structural mechanism of LOV domain activation by light. ....	13
Figure 1.8: Early engineering of light responsive proteins.....	16
Figure 1.9: Engineering of VVD magnets. ....	18
Figure 1.10 Proposed mechanism of LOV domain activation by light. ....	21
Figure 1.11 Photochemical reaction of LOV domain proteins incorporated with flavin analogue.....	23
Figure 2.3 SDS-PAGE analysis of OdAu1aLOV purification. ....	62
Figure 3.1 Photochemical reaction of OdAu1a.....	85
Figure 3.2 Circular dichroism (CD) spectra of FMN and 5dFMN OdAu1aLOV and OdAu1abZIPLOV.....	87
Figure 3.3 <sup>15</sup> N- <sup>1</sup> H TROSY-HSQC NMR characterisation of dark FMN and 5dFMN OdAu1aLOV.....	88
Figure 3.4 UV-Vis spectra of FMN and 5dFMN containing OdAu1aLOV. ....	89
Figure 3.5 Circular dichroism studies of OdAu1aLOV.....	91
Figure 3.6 Circular dichroism studies of OdAu1abZIPLOV. ....	92
Figure 3.7 <sup>15</sup> N- <sup>1</sup> H TROSY-HSQC NMR characterisation of lit FMN and 5dFMN OdAu1aLOV. ....	93
Figure 3.8 Analytical size exclusion chromatography characterisation of OdAu1aLOV. ....	95
Figure 3.9 Size exclusion chromatography characterisation of DNA binding of FMN or 5dFMN-containing OdAu1abZIPLOV constructs.....	97
Figure 3.10 DNA binding studies of FMN and 5dFMN containing OdAu1abZIPLOV construct. ....	98
Figure 3.11 Images of OdAu1aLOV crystals. ....	100

Figure 3.12 X-ray crystal structures of dark state FMN and 5dFMN OdAu1aLOV.	101
Figure 3.13 X-ray crystal structures of illuminated state FMN and 5dFMN OdAu1aLOV.	104
Figure 3.14 X-ray crystal structures of lit state FMN and 5dFMN OdAu1aLOV.	105
Figure 3.15 Modelling of Gln293 swing and flip states from early refinement....	107
Figure 3.16 Proposed structural mechanism of Au1aLOV activation by light. ...	109
Figure 3.17 UV-visible spectra of FMN and 5dFMN OdAu1a constructs.....	118
Figure 3.18 Mass spectra of FMN OdAu1aLOV and OdAu1abZIPLOV. ....	120
Figure 3.19 Mass spectra of 5dFMN incorporated OdAu1aLOV and OdAu1abZIPLOV. ....	121
Figure 3.20 Mass spectrum of harvested lit state 5dFMN containing OdAu1aLOV crystals. ....	122
Figure 3.21 Analytical size exclusion chromatography calibration curve of Superdex 75 column. ....	123
Figure 3.22 Analytical size exclusion chromatography calibration curve of Superdex 200 column. ....	123
Figure 4.1 Avena sativa phototropin LOV2 (AsLOV2) activation by light.....	126
Figure 4.2 5-fluorotryptophan labelling of AsLOV2. ....	129
Figure 4.3 UV-Vis and CD characterization of FMN AsLOV2. ....	131
Figure 4.4 UV-Vis and CD characterization of 5dFMN AsLOV2. ....	132
Figure 4.5 CD of AsLOV2 containing FMN and 5dFMN as a cofactor.....	133
Figure 4.6 CD following FMN and 5dFMN containing AsLOV2 relaxation.....	133
Figure 4.7 <sup>19</sup> F 1D NMR characterisation of 5fW labelled FMN AsLOV2.....	134
Figure 4.8 <sup>19</sup> F 1D NMR relaxation kinetics of 5fW labelled AsLOV2. ....	135
Figure 4.9 <sup>19</sup> F 1D NMR characterisation of 5fW labelled 5dFMN AsLOV2.....	136
Figure 4.10 UV-Vis characterization of AsLOV2FL. ....	138
Figure 4.11 CD characterization of AsLOV2FL containing FMN and 5dFMN as a cofactor.....	139
Figure 4.12 <sup>19</sup> F 1D NMR characterisation of 5fW labelled FMN AsLOV2FL.....	140
Figure 4.13 <sup>19</sup> F 1D NMR characterisation of 5fW labelled 5dFMN AsLOV2FL.	141

Figure 4.14 Room temperature X-ray structures of AsLOV2 in dark and lit states indicating changes in hydrogen bonds extending from the FMN-O4.....	143
Figure 4.15 Liquid chromatography mass spectroscopy (LC-MS) analysis of FMN AsLOV2. ....	148
Figure 4.16 Liquid chromatography mass spectroscopy (LC-MS) analysis of 5dFMN AsLOV2.....	149
Figure 4.17 Liquid chromatography mass spectroscopy (LC-MS) analysis of FMN containing AsLOV2 protein labelled with 5fW.....	150
Figure 4.18 Liquid chromatography mass spectroscopy (LC-MS) analysis of 5dFMN containing AsLOV2 labelled with 5fW. ....	151
Figure 4.19 Liquid chromatography mass spectroscopy (LC-MS) analysis of FMN and 5dFMN containing AsLOV2FL proteins labelled with 5fW. ....	152
Figure 4.20 UV-Vis characterization of FMN and 5dFMN containing AsLOV2. ....	153
Figure 4.21 CD studies of AsLOV2 containing FMN or 5dFMN as a cofactor... ..	154
Figure 4.22 19F 1D NMR characterisation of refolded 5fW labelled FMN AsLOV2. ....	155
Figure 4.23 UV-Vis characterisation of AsLOV2FL. ....	156
Figure 4.24 CD studies of AsLOV2FL containing FMN or 5dFMN as a cofactor. ....	157
Figure 4.25 19F 1D NMR characterisation of refolded 5fW labelled FMN AsLOV2FL. ....	158
Figure 4.26 SDS-PAGE analysis of purified AsLOV2 and AsLOV2FL protein samples. ....	159
Figure 5.1 Photoactivation and structure of <i>Ochromonas danica</i> Aureochrome1a LOV domain.....	162
Figure 5.2 Structural B-factor analysis of <i>Ochromonas danica</i> and <i>Phaeodactylum tricornutum</i> Aureochrome1a LOV (Au1aLOV) domains. ....	164
Figure 5.3 <i>Ochromonas danica</i> Aureochrome1a LOV domain (OdAu1aLOV) construct design.....	166
Figure 5.4 Size exclusion chromatography (SEC) investigations of OdAu1aLOV oligomerisation change upon photoactivation under NMR like conditions. ....	167
Figure 5.5 NMR investigations of 5fW271 labelled OdAu1aLOV. ....	169



Figure 5.6 NMR studies of full length OdAu1a.....	170
Figure 5.7 Characterisation of OdAu1aLOV and $\Delta$ J $\Delta$ OdAu1aLOV by UV-Vis and CD. ....	171
Figure 5.8 NMR investigations of $\Delta$ J $\Delta$ OdAu1aLOV. ....	172
Figure 5.9 Size exclusion chromatography (SEC) investigations of $\Delta$ J $\Delta$ OdAu1aLOV oligomerisation change upon photoactivation under NMR like conditions. ....	173
Figure 5.10 X-ray crystallography studies of $\Delta$ J $\Delta$ OdAu1aLOV construct.....	175
Figure 5.11 Characterisation of 2fY182 and 2fY300 labelled OdAu1aLOV by UV-Vis. ....	176
Figure 5.12 Characterisation 2fY labelled OdAu1aLOV and OdAu1aLOV300TAG by CD. ....	177
Figure 5.13 SEC investigations of 2fY labelled OdAu1aLOV and OdAu1aLOV300TAG.....	179
Figure 5.14 Light dependent multiple dimer formation for 2fY300 OdAu1aLOV sample.....	180
Figure 5.15 <sup>19</sup> F NMR investigations of 2fY182 OdAu1aLOV. ....	180
Figure 5.16 SIIS <sup>19</sup> F NMR investigations of dark state 2fY182 OdAu1aLOV. ..	181
Figure 5.17 SIIS <sup>19</sup> F NMR investigations of lit state 2fY182 OdAu1aLOV.....	183
Figure 5.18 Mapping 2fY182 <sup>19</sup> F NMR results onto crystal structure of OdAu1aLOV and OdAu1aLOV.....	186
Figure 5.19 Dark state <sup>15</sup> N- <sup>1</sup> H TROSY-HSQC spectra of OdAu1aLOV and 5fW271 labelled OdAu1aLOV. ....	191
Figure 5.20 Lit state <sup>15</sup> N- <sup>1</sup> H TROSY-HSQC spectra of OdAu1aLOV and 5fW271 labelled OdAu1aLOV.....	192
Figure 5.21 LC-MS analysis of OdAu1aLOV and OdAu1a. ....	193
Figure 5.22 Dark state <sup>15</sup> N- <sup>1</sup> H TROSY-HSQC spectra of $\Delta$ J $\Delta$ OdAu1aLOV and 5fW271 labelled $\Delta$ J $\Delta$ OdAu1aLOV.....	194
Figure 5.23 Lit state <sup>15</sup> N- <sup>1</sup> H TROSY-HSQC spectra of $\Delta$ J $\Delta$ OdAu1aLOV and 5fW271 labelled $\Delta$ J $\Delta$ OdAu1aLOV.....	195
Figure 5.24 LC-MS analysis of $\Delta$ J $\Delta$ OdAu1aLOV. ....	196

Figure 5.25 LC-MS analysis of 2fY182 and 2fY300 labelled OdAu1aLOV, and OdAu1aLOV300TAG. ....	197
Figure 5.26 SEC75 characterisation of OdAu1aLOV. ....	198
Figure 5.27 SEC75 characterisation of 2fY182 labelled OdAu1aLOV. ....	199
Figure 5.28 SEC75 characterisation of 2fY300 OdAu1aLOV. ....	200
Figure 5.29 SEC75 characterisation of OdAu1aLOV300TAG. ....	201
Figure 5.30 SEC75 characterisation of $\Delta J\alpha$ OdAu1aLOV. ....	202
Figure 5.31 $^{19}\text{F}$ NMR spectra of 3,5-difluoro-L-tyrosine (2fY) in 10 and 100% D <sub>2</sub> O buffer. ....	203
Figure 5.32 Calibration of SEC75 column and investigation of OdAu1aLOV light dependent change in elution maxima. ....	204



## List of Tables

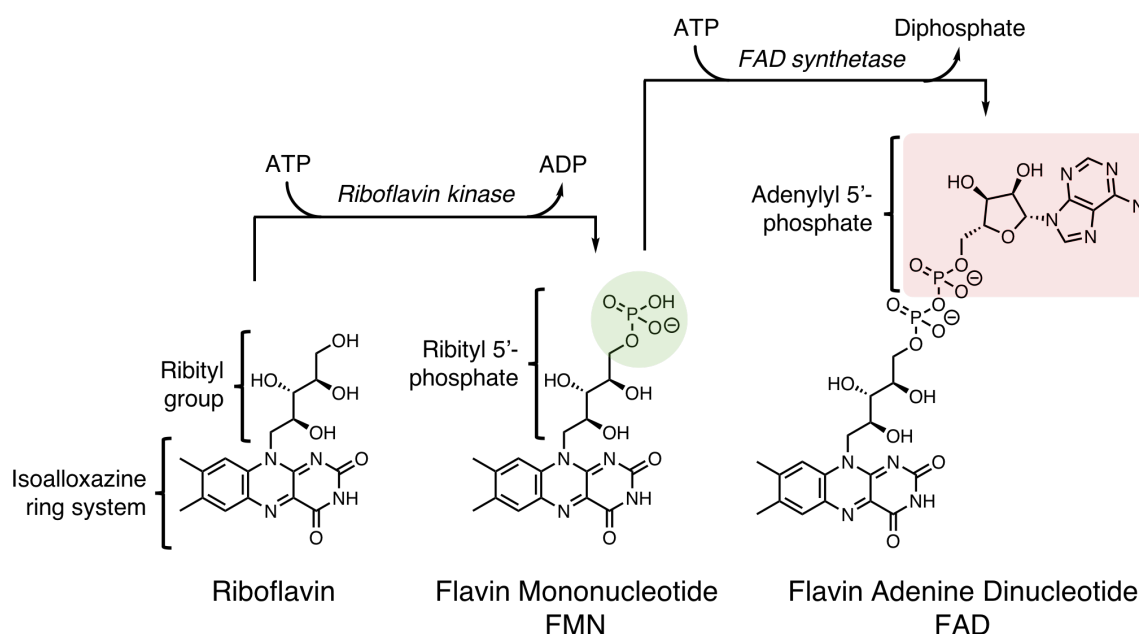
Table 2.1: Commonly used additives in growth media including antibiotics. ....	39
Table 2.2: <i>Escherichia coli</i> ( <i>E. coli</i> ) strains used for plasmid amplification. ....	46
Table 2.3: <i>Escherichia coli</i> ( <i>E. coli</i> ) strains used for overproduction of proteins. ....	46
Table 2.4 $\Delta A'\alpha$ OdAu1a <sub>LOV</sub> and $\Delta J\alpha$ OdAu1a <sub>LOV</sub> construct generation primers. ....	73
Table 3.1 Data sets collected for FMN-containing <i>OdAu1a</i> <sub>LOV</sub> . ....	116
Table 3.2 Data sets for 5dFMN-containing <i>OdAu1a</i> <sub>LOV</sub> . ....	117
Table 5.1 Dataset for $\Delta J\alpha$ OdAu1a <sub>LOV</sub> . ....	205





## 1 Introduction

Flavoproteins are ubiquitous proteins found in all living organisms<sup>1,2</sup>. They contain at least one of the flavin-based cofactors detailed in Figure 1.1, which are riboflavin, flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD)<sup>3–6</sup>. The structure of isoalloxazine ring system allows flavoproteins to catalyse a vast array of different biochemical transformations including metabolic reactions<sup>7–10</sup>, activation of molecular oxygen<sup>11–13</sup> that can lead to luminescence<sup>14–16</sup> and light sensing<sup>17</sup>. FMN and FAD cofactors are usually tightly ( $K_D$  in the nanomolar range) and non-covalently bound to the protein core, where the ribityl chain and the corresponding 5'-hydroxy modifications provide a scaffold for cofactor binding and selectivity<sup>18</sup>. Flavoproteins are probably the most versatile proteins known where nature has also used flavin as a chromophore to sense blue, a focus in this work.



**Figure 1.1 Structure of riboflavin, flavin mononucleotide and flavin adenine dinucleotide.**

Riboflavin is a biological precursor to flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD). The transformation of riboflavin to FMN is catalysed by a riboflavin kinase using adenosine triphosphate (ATP) as a substrate, where the  $\gamma$ -phosphate of ATP is transferred to the 5'-hydroxy of the ribityl chain<sup>19</sup>. Biotransformation of FMN to FAD is catalysed as FMN adenylyltransferase, using ATP as a substrate<sup>20</sup>. Both enzymes require  $Mg^{2+}$  for optimal enzymatic activity where bacteria usually contain a bifunctional single enzyme<sup>20</sup>.

## 1.1 Light Sensing

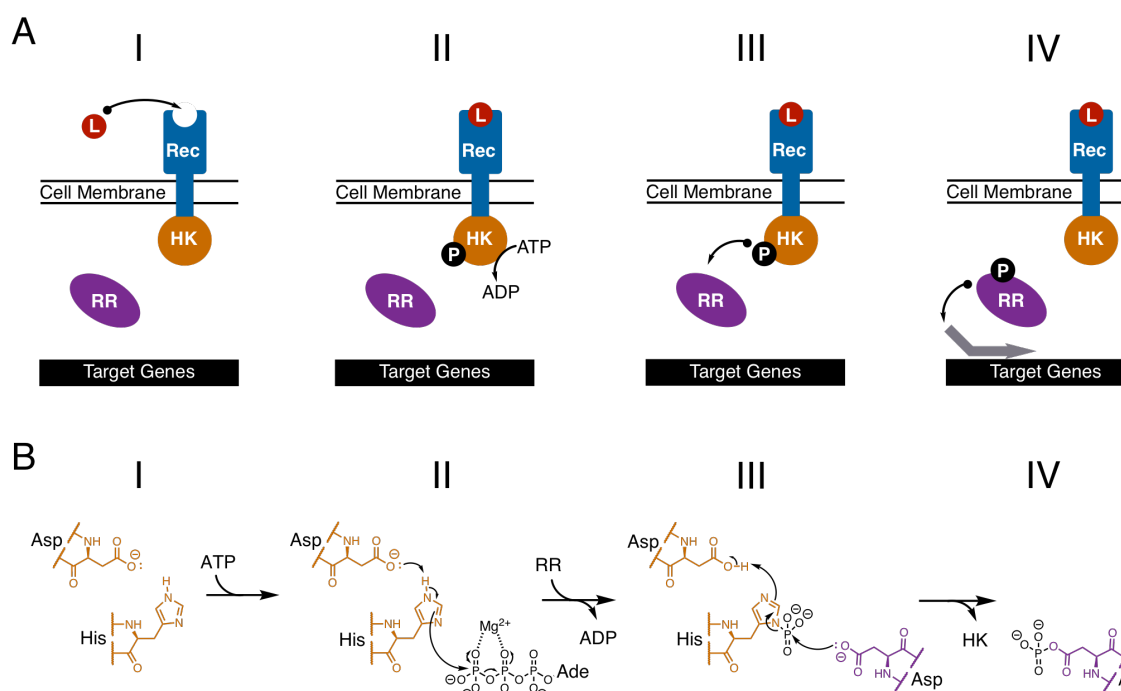
The ability to sense light is crucial for the big majority of living organisms in order to harness light for photosynthesis, protection from photodamage caused to the DNA and RNA *via* regulation of genes involved in damage repair mechanisms and pigment biosynthesis, protection from reactive oxygen species (ROS), sensory purposes such as vision and phototaxis, entrainment of circadian rhythms (biological clocks), seasonal acclimatisation and play a role in other physiological mechanisms<sup>21–24</sup>. Light sensing is achieved through a number of versatile photoreceptor proteins capable of detecting light in the range of UV-B (280 nm) to far-red (710 to 800 nm)<sup>25–36</sup>.

Photoreceptors are usually two-component systems (TCS), where a receptor module senses a physical stimuli and through a secondary module converts this information into the biochemical response<sup>37,38</sup>. The best example of TCS includes bacterial metabolite transmembrane receptors depicted in Figure 1.2<sup>39</sup>. In such system, binding of a ligand to the receptor located in the extracellular environment causes activation of cytosol facing histidine kinase (HK) domain, one of the most common signal transduction domains in bacteria. Allosterically activated HK autophosphorylates at a conserved histidine consuming a molecule of ATP<sup>40</sup>. Generated phosphohistidine then phosphorylates an aspartate residue located on a response regulator (RR)<sup>39,40</sup>. Phosphorylated RR, in this example a transcription factor, undergoes a conformation change facilitating tighter DNA binding and triggering expression of genes in a response to an external stimuli. In many bacteria, such systems allow rapid acclimatisation to changes in the environment such as oxygen<sup>41</sup>, availability of nutrients<sup>42,43</sup> and metal ions<sup>44</sup>.

## 1.2 Blue Light Sensing Using Flavins as a Chromophore

Historically, blue light was known to dictate multiple biological processes in plants including phototropism, stomatal opening, flowering, gene expression, direction of growth and regulation of the circadian rhythms<sup>22–24</sup>. The importance of blue light was known, but a protein responsible for the blue light sensing eluded identification. Initially, it was hypothesised that a chromophore with UV-A and blue absorbance such as pterin, flavin or carotenoid was essential, but low natural abundance and difficult biochemical assays made isolation challenging. The breakthrough appeared upon isolation of *Arabidopsis thaliana* mutant with elongated hypocotyls (HY4 mutant), where the sequential work allowed isolation of HY4 gene<sup>45</sup>. Upon sequencing, it appeared that HY4 gene shared



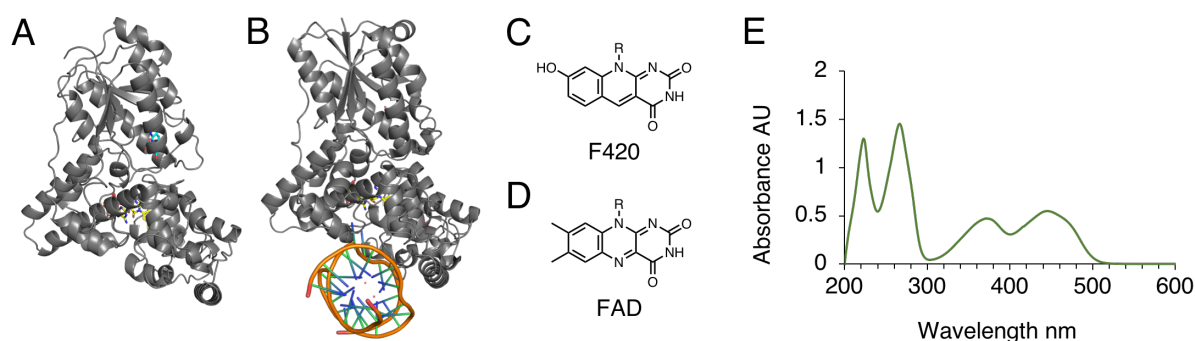


**Figure 1.2 Explanation of two component systems.**

A) Indicates the mechanism of two component systems (TCS). In the resting state I, membrane bound sensor (blue squares) inhibits the activity of cytosol located histidine kinase (HK) domain (orange circle). Ligand binding event (red circle labelled L), state II, results in structural changes leading to activation of the HK domain that autophosphorylates at a key histidine residue consuming a molecule of ATP or another nucleotide triphosphate (NTP). In a sequential step III, a phosphate group (black circle labelled P) is transferred from a histidine onto an aspartate residue located on a response regulator (RR, purple oval). In this example, a RR is a transcription factor that upon phosphorylation is activated and binds onto DNA to induce gene expression, step IV. B) indicates individual chemical steps taking place in corresponding steps I, II, III and IV.

homology with DNA-photolyase enzyme, an enzyme known to use FAD as a cofactor for the activity<sup>46,47</sup>, Figure 1.3 A. HY4 did not indicate any DNA binding domain suggesting an identification of a novel blue light sensor, appropriately called cryptochrome (CRY)<sup>45</sup>.

Upon biochemical characterisation, it was shown that CRY non-covalently bound FAD, a cofactor well suited for blue light sensing<sup>48–50</sup>. Flavins have an absorbance maxima within the blue light and UV-A regions, Figure 1.3 D and E. Photochemical reactivity of flavins has been known in biochemistry, for instance photolyase<sup>46,47</sup>. Although absorption properties of flavins is well suited for a photosensory mechanisms, the actual understanding of the photochemical reactions remained understood for a decade. After the identification of CRY, other light sensing flavoproteins were isolated from multiple organisms including plants, algae, cyanobacteria, bacteria, archaea, fungi, flagellates, diatoms and animals.



**Figure 1.3 Structure of DNA photolyase.**

A) Crystal structure of *Anacystis nidulands* DNA photolyase, (PDB 1QNF)<sup>51</sup>. B) Structure of *Drosophila melanogaster* DNA photolyase-DNA complex, (PDB 3CVY)<sup>47</sup>. C) Structure of cofactor F420, also known as 8-hydroxy-7-desmethyl-5-deazaflavin, used as Förster resonance energy transfer (FRET) antenna for activation of FADH<sup>-</sup> centre in DNA photolyase found in microorganisms and insects. D) Structure of isoalloxazine ring system of FAD. E) UV-Vis absorbance spectra of 50  $\mu$ M FMN in 10 mM potassium phosphate buffer pH 7.0 indicating absorbance maxima at 280, 390 (UV-A region) and 450 nm (blue light in the visible light region).

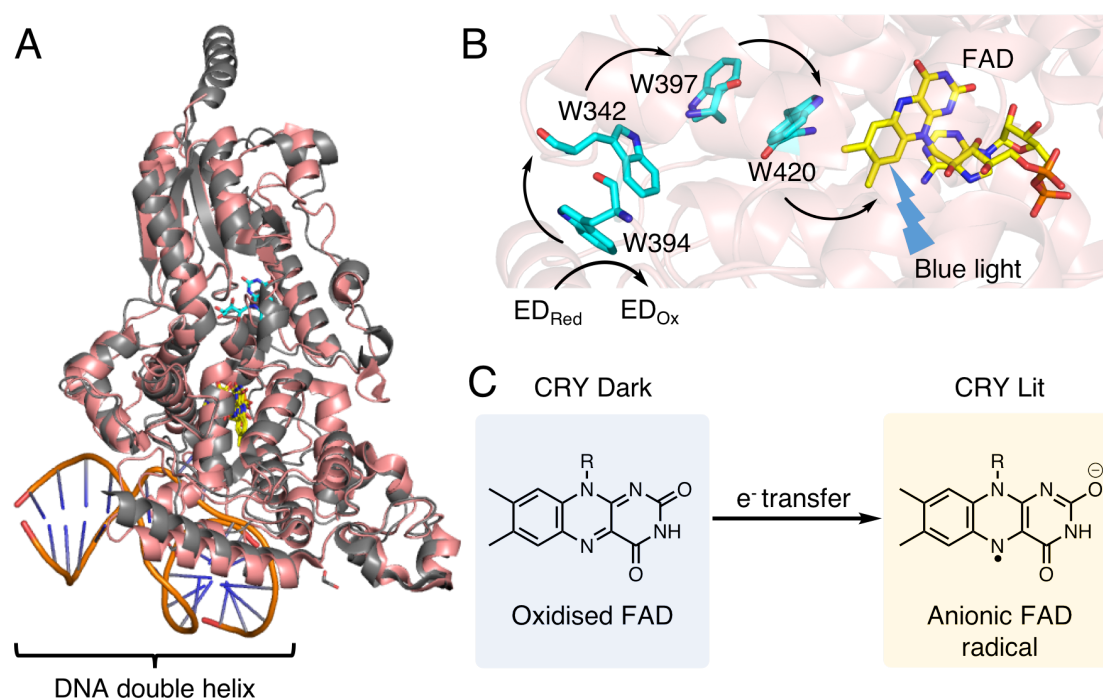
Flavin cofactors are well suited to act as a photosensors due to photochemical properties<sup>52,53</sup>. Excitation with blue light allows excitation of a  $\pi$ -bonding low-energy electron into high energy  $\pi$ -antibonding orbital (lowest unoccupied molecular orbital, LUMO) generating singlet excited flavin species (S1 species)<sup>54</sup>. Most notably, flavins can also undergo an intersystem crossing, where a spin of an excited electron is inverted, generating triplet excited flavin species (T species)<sup>55,56</sup>. Both S1 and T species act as a strong single electron oxidising agents allowing photochemically excited flavins to readily react with a number of redox active amino acid residues including cysteine, tyrosine and tryptophan<sup>23,57–59</sup>. Throughout the course of evolution, flavoproteins have exploited this unique reactivity appropriately applying it in different photoreceptors as discussed later in this chapter. Additionally, flavins can also act as excellent photosensitisers allowing excited isoalloxazine ring system to readily transfer energy to a molecular oxygen generating a very reactive singlet oxygen species<sup>60–62</sup>.

### 1.3 Cryptochromes as Light Sensors

As mentioned earlier, CRY proteins are ubiquitous proteins controlling many physiological functions including entrainment of circadian rhythms and is the only blue light sensing flavoprotein found in animals<sup>63</sup>. Structurally, cryptochromes are similar to FAD dependent photolyases, Figure 1.4 A, that catalyse light dependent photodamaged DNA repair. This structural similarities suggest an evolutionary link furthermore supported by an observation

that some CRY proteins have retained an ability to act as a photolyase as well as transcription repressor in diatoms, an activity not present in plant or animal CRY<sup>64,65</sup>.

In contrast to the photolyases, the actual photochemical mechanism of CRY activation by light is still debated. For the photolyases, the redox state of a biologically active cofactor has been suggested to be FADH<sup>•</sup> considering that the protein purifies in a reduced state<sup>47</sup>. Conversely, CRY purify with FAD in an oxidised state suggesting biologically relevant oxidation state<sup>66–68</sup>. When exposed to blue light CRY accumulate flavin semiquinone radical species *in vitro* where external reducing agent, such as glutathione, acts as a sacrificial electron donor. For the animal CRY it has been shown to be an anionic<sup>63,70,71</sup> and for plant counterpart a neutral flavin semiquinone species<sup>58,72</sup>, respectively.



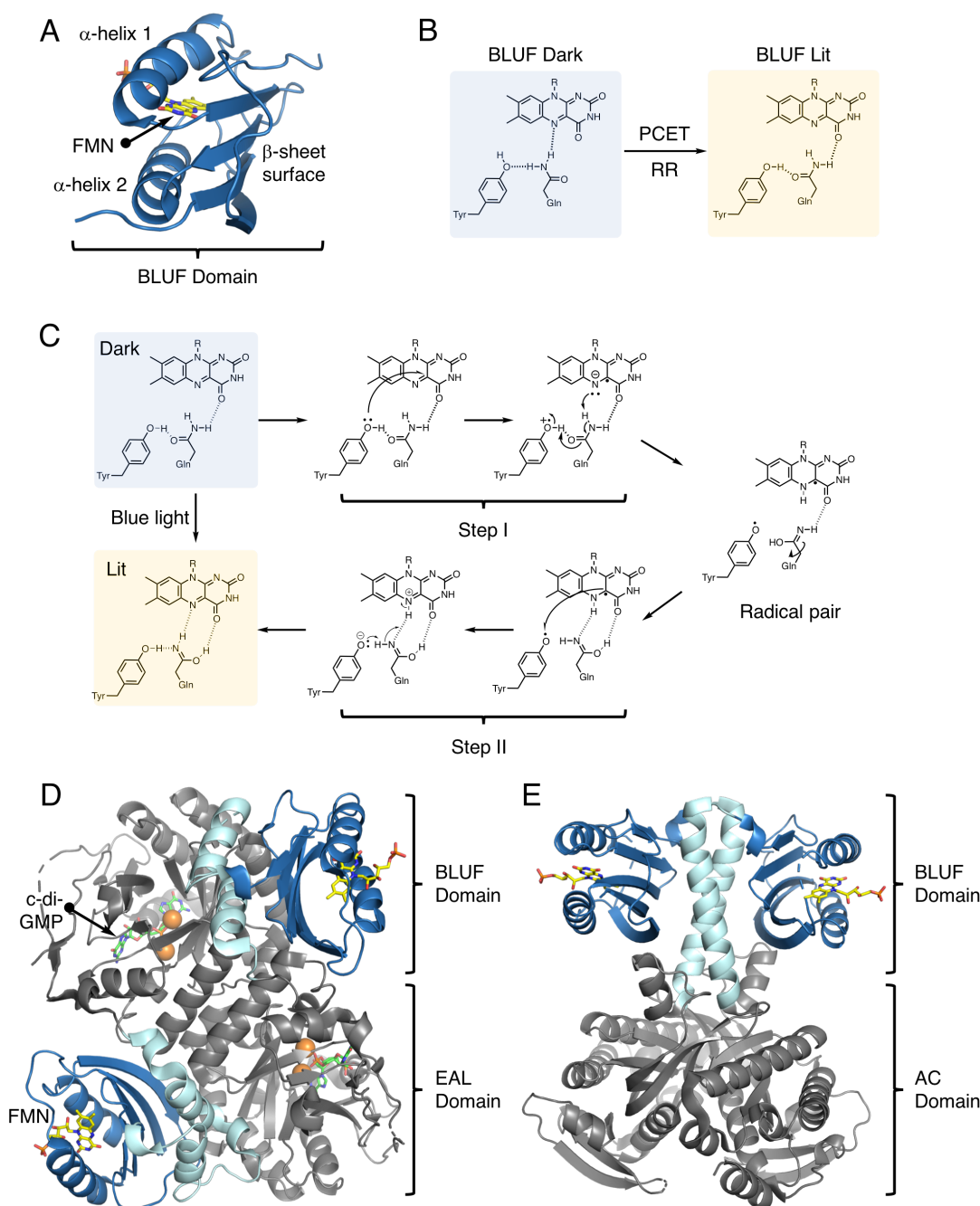
**Figure 1.4 Structure and photochemical mechanism of cryptochrome proteins.**

A) Overlaid structures of *Drosophila melanogaster* DNA photolyase-DNA complex in grey with DNA in orange and blue (PDB 3CVY) and cryptochrome (*DmCRY*) in pink (PDB 4JZY) indicating an identical structural fold. FAD cofactors are depicted as yellow and F420 cofactor as blue sticks, respectively. B) Proposed electron transfer pathway allowing photoreduction of FAD centre in a light dependent manner through tryptophan (Trp/W) tetrad<sup>69</sup>, where ED stands electron donor and ED<sub>Red</sub> corresponds to a reduced and ED<sub>Ox</sub> oxidised states, respectively. FAD is depicted as yellow sticks and Trp tetrad as blue sticks, respectively. Trp residues important for the electron relay pathway are appropriately numbered and labelled. For image generation PDB 4JZY was used (*DmCRY*). C) Photochemical reduction of FAD centre triggered by light, where for *DmCRY* light facilitates accumulation of an anionic semiquinone FAD. Photoreduction then leads to proposed conformational changes triggering biological response.

The actual photochemical mechanism of photoreduction remains unclear and appears to vary amongst different CRY. Several different hypotheses have been put forward reflecting on the *in vitro* studies. Using fruit fly *Drosophila melanogaster* CRY (DmCRY) as a model the proposed photochemical mechanism involves excitation of FAD active centre followed by an electron relay from an external electron donor such as a physiological reducing agent glutathione<sup>69</sup>, Figure 1.4 B. The electron relay is achieved through four highly conserved tryptophan (Trp) residues connecting FAD binding pocket to the surface of the protein. The actual mechanism involves photochemical oxidation of Trp residues by the FAD centre, Figure 1.4 B. Generated Trp<sup>•+</sup> (or Trp<sup>•</sup>) species can furthermore oxidise other proximal redox active Trp/Tyr residues, where surface exposed Trp<sup>•+</sup> can readily oxidise redox active compounds<sup>69</sup>. The generation of a radical intermediate has been shown to trigger a net structural change. Depending on CRY, this then leads to changes in oligomerisation state or interaction with other proteins facilitating translation of light stimuli into a biological response.

#### 1.4 Blue Light Sensing Utilising Flavin Adenine Dinucleotide (BLUF) Domain

Yet another different structural fold is observed in Blue Light Sensing Utilising Flavin Adenine Dinucleotide (BLUF) domain proteins adopting a ferredoxin-like  $\alpha/\beta$  fold, five antiparallel  $\beta$ -sheets flanked by two  $\alpha$ -helices on one side, Figure 1.5 A, with the FAD chromophore 'sandwiched' in-between the two  $\alpha$ -helices and the isoalloxazine positioned perpendicular to the  $\beta$ -sheet surface<sup>73–75</sup>. Unlike CRY proteins, BLUF proteins are mostly found in bacteria, some flagellates and some fungi, and do not show a net change in the cofactor redox state upon photoexcitation<sup>54,76</sup>. Exposure to blue light allows generation of signalling active BLUF domain, lit state, *via* subtle changes in hydrogen bond networks<sup>54</sup>. Unlike the resting dark state, a photoexcited FAD spectrum is redshifted, a shift in maxima from approximately 445 to 455 nm, dark to a lit state, respectively<sup>77</sup>. This redshift of around 10 nm is indicative of changes in hydrogen bonding interactions between FAD and the protein<sup>78–80</sup>. Site directed mutagenesis and structural studies have identified highly conserved glutamine (Gln) and tyrosine (Tyr) residues participating in achieving the lit state conformation<sup>27,54,76,78,81</sup>, Figure 1.5 B and C. The mutation of Tyr to isoleucine (Ile) or Gln to alanine (Ala) abolished generation of redshifted lit intermediate indicating the importance of these residues. Nonetheless, the actual change in the hydrogen bond network is still highly debated and two hypotheses have been put forward to explain it.



**Figure 1.5 Structure and Mechanism of BLUF Domain Proteins.**

A) Canonical structure of BLUF domain consisting of  $\alpha/\beta$  ferredoxin like fold (PDB 2KB2, where C-terminal helices were removed for clarity). B) From X-ray structural studies proposed change in hydrogen bond networks after photoexcitation of BLUF domain. In the proposed mechanism it was suggested that prior photoexcitation a conserved glutamine (Gln) residue hydrogen bonds to FAD-N5. Upon light triggered proton coupled electron transfer (PCET) from a conserved tyrosine (Tyr), hydrogen bond network rearrangement occurs where Gln residue now hydrogen bonds to FAD-O4. C) Proposed photochemically driven hydrogen bond network rearrangement proposed from Fourier transform infrared spectroscopy (FT-IR) studies. In the mechanism, light triggered PCET and possible proton shuffling in Step I allows formation of tyrosyl radical (Tyr $\cdot$ ) and FAD neutral semiquinone radical pair that allows reorientation of a conserved glutamine residue tautomer. This is then followed by a second PCET where Tyr $\cdot$  oxidises FAD neutral semiquinone radical in step II triggering the formation of lit state hydrogen bond network.

D) Crystal structure of *Klebsiella pneumoniae* BlrP1 protein (KpBlrP1, PDB 3GFZ) where BLUF domain is depicted in blue, C-terminal helices important for signalling in cyan and cyclic diguanosine monophosphate (c-di-GMP) hydrolase (EAL) domain in grey. Please note FMN is depicted as yellow sticks, c-di-GMP as green sticks and Mn<sup>2+</sup> ions as orange spheres, respectively. E) Crystal structure of *Oscillatoria acuminata* photoactivatable adenylate cyclase (OaPAC, PDB 5X4T) where N-terminus BLUF domain is depicted in blue, C-terminus coiled-coil shown in cyan and C-terminus adenylate cyclase in grey (PDB 5X4T). FMN is depicted as yellow sticks.

The first hypothesis has been put forward when the structure of BLUF domain was determined in dark and lit states suggesting light induced Gln residue flip and reorientation<sup>73</sup> dictating hydrogen bond rearrangements extending from the O4 of FAD, Figure 1.5 B. Considering that the electron densities of the oxygen in C=O bond and the nitrogen of C-NH<sub>2</sub> are very similar, the orientation of Gln residue sidechain positioning in the reported crystal structures (2.3 Å) was argued to be impossible to assign with certainty. To test this mechanism, studies employing molecular dynamics (MD) were undertaken and failed to support Gln flip suggesting a second mechanism of photoactivation<sup>82–85</sup>, Figure 1.5 C. It was proposed that Gln residue undergoes keto-enol tautomerisation leading to the changes in hydrogen bond networks, although no structural information supported these findings. Further insights came from femtosecond infrared (IR) and Fourier transform IR (FT-IR) studies, which identified a previously observed neutral FADH· radical intermediate in the photocycle as well as providing evidence for the keto-enol tautomerisation of Gln sidechain<sup>86</sup>. FT-IR experiments furthermore identified one unique observation, a change in the hydroxy (OH) stretch of the conserved Tyr residue only in the lit state suggesting a very strong hydrogen bond. FT-IR using <sup>15</sup>N labelled Gln also provided evidence for the tautomerisation taking place. Using this experimental data in the computational studies, it was suggested that such change would only be evident if a new hydrogen bond would be formed between Tyr-OH to the imine of Gln sidechain tautomer and Gln-OH tautomer forming a hydrogen bond with the FAD-O4, Figure 1.5 C. Currently, this is the leading hypothesis to explain BLUF domain signalling, but no structural information has been observed to suggest required flip of a keto-enol tautomer.

To date very few structural studies have described multidomain BLUF domain proteins. A notable BLUF domain protein, BlrP1, was identified in  $\gamma$ -proteobacteria *Klebsiella pneumoniae*, an important nosocomial pathogen<sup>87</sup>. It was shown that BlrP1 contained an N-terminus BLUF domain and C-terminus EAL domain, a cyclic dimeric guanosine monophosphate (c-di-GMP) hydrolase domain, Figure 1.5 D. c-di-GMP is a secondary messenger in many different bacteria and has been shown to regulate motility,



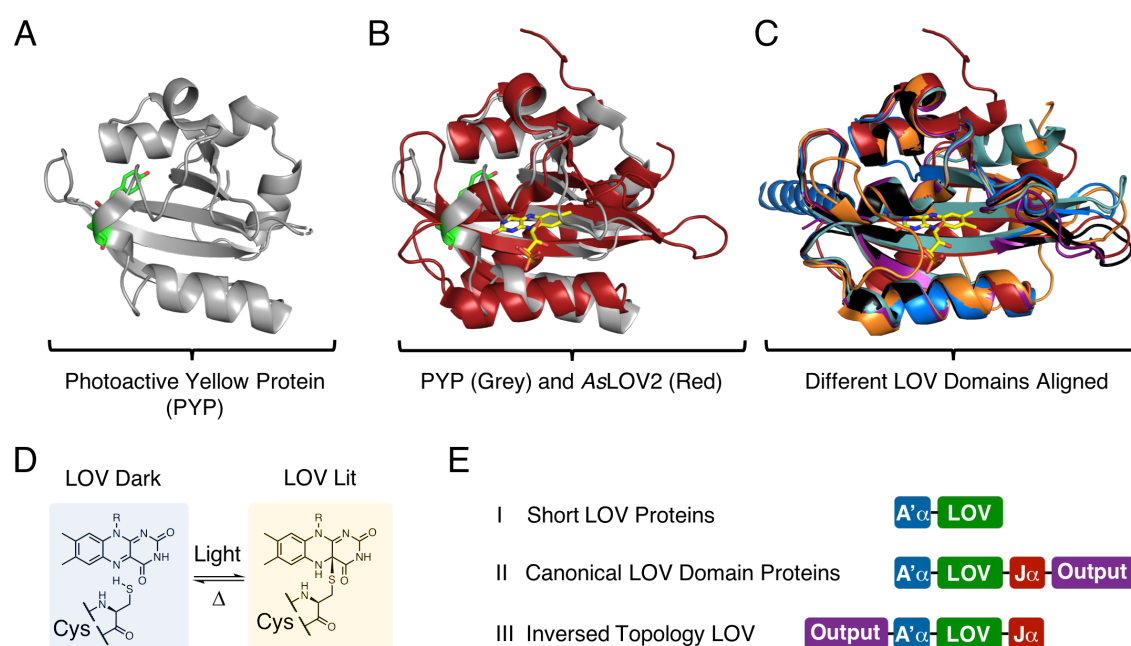
biofilm formation, virulence and antibiotic resistance operon expression<sup>88</sup>. Excitation with light indicated four-fold increase in the catalytic activity, hydrolysis of c-di-GMP, highlighting the fact that BLUF domain structural changes modulated EAL domain activity<sup>87</sup>. Another BLUF domain protein, PAC (photoactivatable adenylate cyclase), was identified in *Oscillatoria acuminata*, a cyanobacterium, and was shown to contain N-terminus BLUF domain and C-terminus adenylate cyclase (AC) domain<sup>89</sup>. Like c-di-GMP, cyclic adenosine monophosphate (cAMP) acts as a secondary messenger in many lower and higher organisms<sup>90</sup>. Strikingly, *OaPAC* indicated almost 20 fold increase in adenylate cyclase activity in response to light, where structure of the photoexcited protein indicated structural changes in the coiled-coil  $\alpha$ -helices suggesting a mechanism where BLUF domains control rearrangements<sup>89</sup>.

Although significant amount of information is known regarding BLUF domain proteins, very few structural studies have described the lit state, especially of full length proteins. Although the general mechanism of photoactivation is becoming more clear, existing inconsistencies amongst different BLUF regarding the mechanistic aspects such as proton coupled electron transfer (PCET) raise questions if a signalling mechanism can be universally described<sup>91</sup>. Additionally, questions still remain how such subtle changes in the hydrogen bond networks from FAD and the protein result in large scale structural changes dictating the activity of intra- and interprotein partners. This control has been suggested to be govern by C-terminal helices such as Ccap or coiled-coils allowing BLUF domains to dictate structural changes of partner proteins. Regardless of the mechanistic questions, BLUF domains prove a unique scaffold in engineered optogenetic tools including already described ability to control pools of secondary messengers such as c-di-GMP and cAMP<sup>88,90</sup> in response to blue light.

### 1.5 Light-Oxygen-Voltage (LOV) Domain Proteins

Like CRY and BLUF proteins, Light-Oxygen-Voltage (LOV) domain proteins are photosensory flavoproteins. Structurally, LOV domain belongs to PAS (Per-ARNT-Sim) domain superfamily containing a conserved  $\alpha/\beta$ -fold with five antiparallel  $\beta$ -sheets flanked by several  $\alpha$ -helices and loops, Figure 1.6 A, B and C<sup>55,92–95</sup>. A flavin chromophore is tightly bound within the hydrophobic core with a highly conserved cysteine residue located on top of the isoalloxazine ring system<sup>96</sup>. Blue light triggers the formation of a thermally reversible cysteinyl-FMN covalent photoadduct<sup>97</sup>, Figure 1.6 D, where this change is coupled to the conformational changes of the LOV domain tertiary structure and residue dynamics<sup>98</sup>. This structural change allows LOV domains to directly control the activity of

intra or interprotein signal output domains, Figure 1.6 E, through auxiliary N- or C-terminal helices  $A'\alpha$  and  $J\alpha$ <sup>17,99–101</sup>, respectively. In a canonical LOV domain architecture, signal output domains are usually fused to the C-terminus, Figure 1.6 E. Exceptions to this rule exist, where blue light responsive transcription factors such as Aureochrome 1a (Au1a) found in stramenopiles<sup>102–105</sup> and so far uncharacterised LOV domain from *Sulfurimonas denitrificans* (SdLOV, UniProt Q30NS0) have an inversed topology. Both proteins have DNA binding domains located in the N-terminus, with Au1a harbouring basic leucine zipper (bZIP)<sup>103</sup> and SdLOV putative helix-turn-helix (HTH) domain.



**Figure 1.6 PAS domain structure and function.**

A) Structure of a prototypical PAS domain protein photoactive yellow protein (PYP), PDB 5QJ5, displayed in grey and sensory chromophore, *p*-coumaric acid linked to a cysteine displayed in green. PYP is a blue light sensing protein where the PAS domain core is fused to the C-terminus histidine kinase (HK) domain. B) Aligned structures of PYP (grey, PDB 5QJ5) and *Avena sativa* phototropin 1 LOV 2 domain (red, PDB 2V0U) indicating an identical PAS domain fold. C) Aligned structures of different LOV domain proteins indicating an identical PAS domain core. *Erythrobacter litoralis* protein 222 (EL222, light responsive transcription factor) LOV domain displayed in teal, PDB 3P7N. *Erythrobacter litoralis* protein 346 (EL346, LOV domain regulated histidine kinase) LOV domain displayed in purple, PDB 4R38. *Brucella abortus* LOV histidine kinase (HK) LOV domain displayed in black, PDB 3T50. *Bacillus subtilis* YtvA protein LOV domain displayed in blue, PDB 2PR5. *Neurospora crassa* Vivid (VVD) displayed in orange, PDB 2PD7. AsLOV2 displayed in red, PDB 2V0U. For clarity, a single FMN molecule of AsLOV2 is displayed in yellow. D) Photochemical reaction of isoalloxazine ring system with proximal cysteine residue located. E) Schematic representation of different LOV domains found in nature with I) indicating canonical short LOV domain proteins such as VVD, II) indicating prototypical LOV domain proteins with N-terminus LOV and C-terminus output domain such as *Bacillus subtilis* YtvA and III) indicating inversed topology LOV domain proteins where the output domain is positioned to the N-terminus of LOV domain such as in stramenopiles found Aureochrome1a proteins.



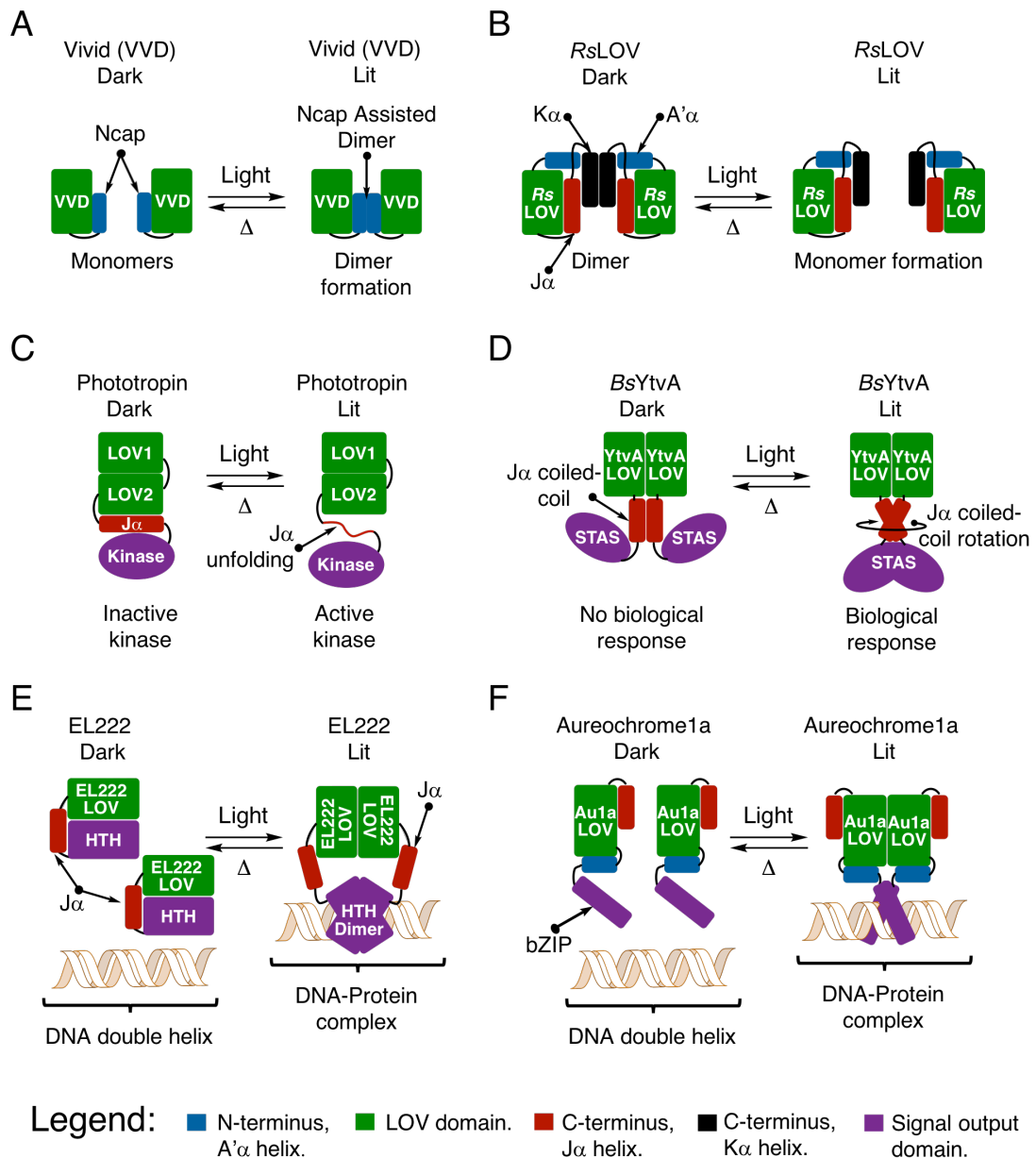
While the signal output domains fused to LOV domains are varied, a group of 'stand-alone' short LOV domains also control different biological responses in fungi, bacteria and plants. This is usually achieved through light dependent protein-protein interactions. One of the best studied examples includes *Neurospora crassa* protein Vivid (VVD)<sup>95</sup>. In the dark state, VVD is monomeric in the solution<sup>95</sup>. Excitation of the protein with blue light leads to the formation of cysteinyl-FAD photoadduct that triggers dimerisation, Figure 1.7 A, and in the presence of White Collar Complex Protein I (WC1) a formation of VVD-WC1 heterodimer<sup>95,106</sup>. Formation of this complex inhibits the formation of White Collar Complex (WCC) changing transcription of genes in response to light and leading to the entrainment of the circadian rhythms<sup>107</sup>. Light dependent oligomerisation is a common mechanism amongst different LOV domain proteins, yet opposite behaviour has been described for *Rhodobacter sphaeroides* LOV protein (*RsLOV*)<sup>108</sup>, Figure 1.7 B. For *RsLOV* light was shown to trigger dimer dissociation. The biological protein-protein interaction targets and the signalling mechanism to date are unknown. Nonetheless, experiments performed on wild type (WT) and  $\Delta$ *RsLOV* *R. sphaeroides* strains indicated major differences in gene products in response to blue light<sup>109</sup>.

There is a number of different LOV signal effector domains directly fused to the LOV core. These include protein kinases<sup>110</sup>, histidine kinases<sup>94,111,112</sup>, sulphate transporter anti  $\sigma$  factor antagonist (STAS) domains<sup>113</sup>, DNA binding domains such as basic leucine zipper (bZIP)<sup>103</sup>, helix-loop-helix (HLH) and helix-turn-helix (HTH)<sup>114</sup>. In such fused proteins, activation of LOV domain by light allows direct activation of fused protein partner. Several studies have provided structural understanding suggesting that the adduct formation results in changes of the LOV domain  $\beta$ -sheet structure and dynamics triggering changes in of A' $\alpha$  and J $\alpha$ <sup>112,115,116</sup>. One example demonstrating such mechanism includes *Avena sativa* phototropin 1 LOV2 (*AsLOV2*) where the formation of cysteinyl-FMN photoadduct results in liberation and unfolding of J $\alpha$ -helix linker connecting LOV2 with C-terminus kinase domain, Figure 1.7 C<sup>110,115</sup>. This unfolding event, observed by NMR and circular dichroism (CD) studies, appears to be conserved amongst different phototropins found in plants and green algae suggesting a common ancestor. Light activated kinase then allows autophosphorylation as well as phosphorylation of serine (Ser) and threonine (Thr) residues on other proteins<sup>110,117</sup>. This then leads to an observed biological changes including opening of leaf stroma for gaseous exchange, chloroplast localisation, phototropism, leaf blade flattening and orientation<sup>117</sup>.

Excluding an unfolding event, a different LOV domain regulatory mechanism has been described for *Bacillus subtilis* YtvA (*BsYtvA*)<sup>113,118,119</sup>. The protein is dimeric in the solution where the C-terminal J $\alpha$  helices form a coiled-coil structure controlling the activity of STAS domain. X-ray studies of isolated *BsYtvA* LOV domain indicated differences between dark grown and dark grown crystal exposed to light prior harvesting revealing J $\alpha$  coiled-coil rotation of 7°<sup>113,120,121</sup>. Further studies using electron paramagnetic resonance (EPR) probe labelling revealed distances in line with X-ray crystallography observations<sup>121</sup>. This therefore strengthened the view on a mechanism where LOV dimer rearrangement from dark to lit state allows J $\alpha$  coiled-coil rotation dictating the activity of C-terminus STAS domain, Figure 1.7 D. *BsYtvA* appeared to control stressosome response to light<sup>118</sup>.

Another mechanism how LOV domains control biological response includes ability to cage DNA binding domains allowing regulation of DNA binding in response to blue light<sup>114</sup>. One such studied example includes *Erythrobacter litoralis* protein 222 (EL222). EL222 contains N-terminus LOV domain protein linked to a C-terminus HTH DNA binding domain *via* J $\alpha$  helix<sup>114,122–124</sup>, Figure 1.7 F. Electrophoretic mobility shift assays (EMSA) have shown 10 fold enhanced DNA binding in response to blue light. As the first structurally characterised full length LOV domain protein, its structure and NMR characterisation revealed yet another mechanism of light control<sup>114</sup>. It was shown that the LOV domain directly cages HTH domain *via* hydrophobic interactions formed between LOV domain  $\beta$ -sheet surface and helix of HTH domain<sup>114</sup>. This therefore proposed a mechanism where direct steric hindrance inhibits DNA binding. It was hypothesised that light activation allows liberation of HTH facilitating EL222 dimerisation onto the DNA, hence EL222 acting as a light responsive transcription factor. In response to light, EL222 was shown to upregulate genes in response to photoprotection and photodamaged DNA repair proteins.

An inversed topology LOV domain transcription factors found in stramenopiles also act as light responsive transcription factors controlling cell cycle and photomorphogenesis<sup>103</sup>. EMSA experiments indicated light responsive DNA binding. Au1a proteins contain N-terminus unstructured domain followed by DNA binding basic leucine zipper (bZIP) domain and C-terminus LOV domain<sup>114</sup>. Studies have shown that isolated Au1a LOV domains undergo dimerisation upon photoexcitation. Structures of isolated Au1a LOV (*Au1a<sub>LOV</sub>*) domains have been reported highlighting N-terminus A' $\alpha$  and C-terminus J $\alpha$ <sup>104,116</sup>. Further Fourier Transform Infrared (FT-IR) studies have suggested A' $\alpha$  and J $\alpha$  unfolding upon cysteinyl-FMN photoadduct formation<sup>104,116</sup>.



**Figure 1.7 Structural mechanism of LOV domain activation by light.**

A) indicates short LOV domain light induced dimerisation observed for a model protein, *Neurospora crassa* Vivid (VVD). B) indicates *Rhodobacter sphaeroides* LOV (RsLOV). In the dark state protein is dimeric, while upon photoexcitation the dimer dissociates. RsLOV contains an additional C-terminal helix assigned as K $\alpha$ , where C-terminal J $\alpha$  and K $\alpha$  form a dimer interface helix bundle (PDB code 4HJ3). C) Activation of plant phototropin I, LOV domain controlled serine/threonine kinase. Blue light causes structural changes within LOV2 domain in light dependent liberation and unfolding of C-terminus J $\alpha$  helix from LOV domain  $\beta$ -sheet surface triggering activation of the kinase domain. D) structural mechanism of *Bacillus subtilis* YtvA, where light causes minor LOV domain rearrangement and rotation of C-terminus coiled-coil J $\alpha$  helices activating C-terminus STAS domain. E) activation of *Erythrobacter litoralis* transcription factor 222 (EL222). DNA binding Helix-Turn-Helix (HTH) domain is caged by LOV domain, where light dependent liberation of HTH domain allows EL222 dimerisation onto the DNA. F) An inverted topology LOV domain transcription factor Aureochrome 1a (Au1a), where proposed LOV domain dimerisation allows DNA binding.

Considering dimerisation and possible unfolding events taking place, a hypothesis has been put forward to explain how Au1a proteins function. The hypothesis suggests that upon light activation LOV domain driven dimerisation places two bZIP domains in close proximity triggering tighter DNA binding, but to date there is no conclusive evidence<sup>125</sup>. Furthermore, the lack of detailed structural information with respect to the full-length protein prevents any further understanding. So far it is still unclear how Au1a<sub>LOV</sub> domain cages DNA binding domain and controls light dependent DNA binding<sup>105,125,126</sup>.

As blue light sensory modules, LOV domains follow several different mechanisms to enforce light dependent protein behaviour although sharing an identical protein fold, Figure 1.6 A to C, and Figure 1.7 A to F. These include light dependent changes in the oligomerisation state as observed for *NcVVD* and *RsLOV*, unfolding events taking place for *AsLOV2*, rotation of coiled-coil J $\alpha$  in *BsYtvA* and direct caging of the DNA binding domains in EL222. This therefore asks a question, how does a cysteinyl-flavin covalent bond in a conserved fold dictates such versatile structural changes and what are the molecular bases.

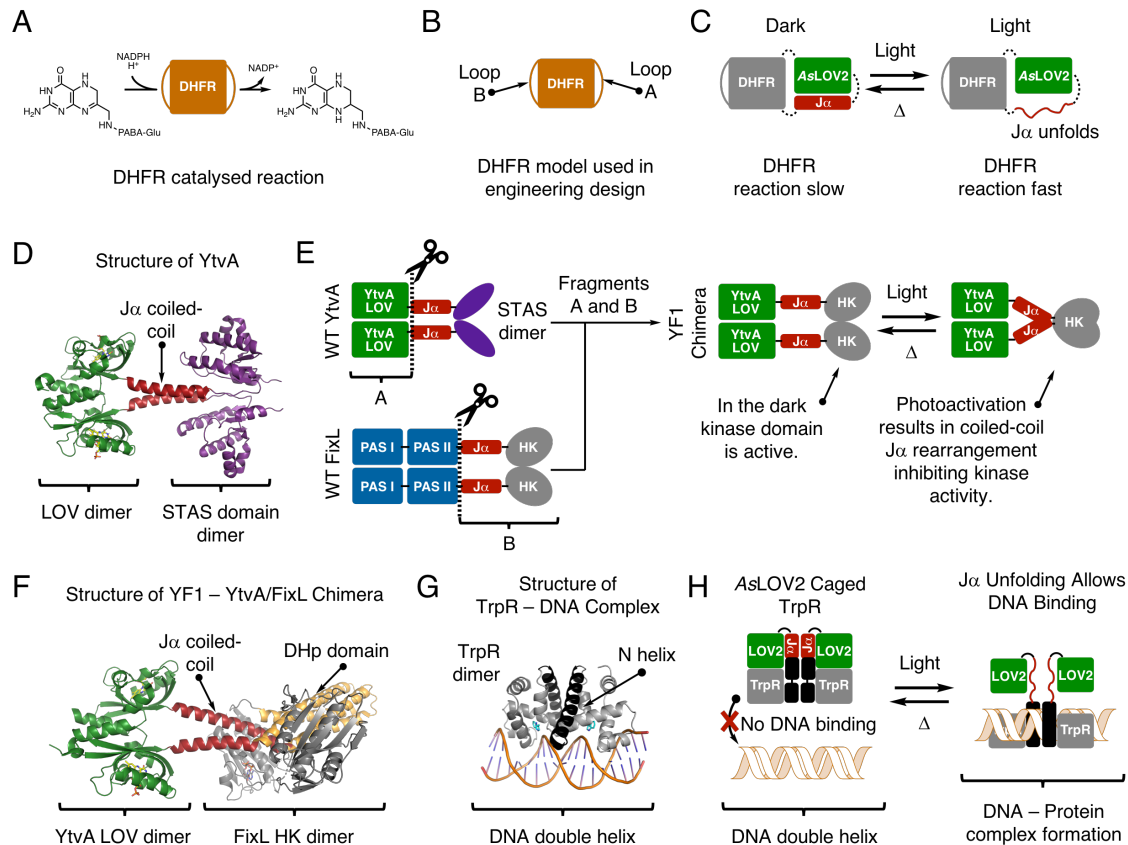
## 1.6 Engineered LOV Photoreceptors

The modular arrangement of sensory LOV domain proteins and the effector modules found in nature has inspired development of many synthetic photocaged systems<sup>127–132</sup>. Ability to control living systems using light as a stimuli provides a traceless, non-invasive approach in a spatiotemporal and dose dependent manner<sup>31,133–135</sup>. Development of all of these novel systems has relied on current understanding of LOV domain structural mechanism of activation<sup>136</sup>. One of the best examples illustrating the applicability of LOV domain proteins to control complex biological systems has been described for metabolically engineered yeast *Saccharomyces cerevisiae*<sup>137</sup>. The fast cycling light responsive transcription factor EL222 was used as a logic gate allowing rapid light dependent switch in the metabolism. Initial experiments demonstrated that production of desired high value metabolite, isobutanol, inhibited yeast growth. To overcome this challenge, a strain of yeast was engineered to contain a light switchable metabolic system where blue light activated EL222 promotes expression of pyruvate decarboxylase 1 (PDC1) allowing ethanol fermentation, hence allowing rapid growth using glucose as a carbon source. At the same time photoactivated EL222 repressed expression of acetolactate synthase (ILV2) inhibiting isobutanol production, hence preventing buildup of isobutanol in the media. When the required biomass accumulated, blue lights were switched off resulting in decreased PDC1 transcription as well as de-repressing ACS transcription in turn triggering

isobutanol production. Using this system, moderate titers of isobutanol were produced. Further improvement in the isobutanol production was achieved by subjecting cultures to cycles of light and dark. Exposure of cultures to short pulses of blue light allowed transient production PDC1 alleviating metabolic arrest leading to increasing glucose consumption and the net production of isobutanol. This work elegantly demonstrated the applicability of light controlled transcription factors to overcome challenges associated with *in vivo* production of high value compounds.

Furthermore, LOV domains have also been used to photocage metabolic enzymes, but to date no tightly light controlled system has been developed<sup>132,138</sup>. For instance, AsLOV2 has been used to photocage *E. coli* dihydro folate reductase (*EcDHFR*)<sup>132</sup>. *EcDHFR* catalyses the reduction of 7,8-dihydrofolate (DHF) to 5,6,7,8-tetrahydrofolate (THF) consuming a molecule of NADPH as a redox cosubstrate<sup>139</sup>, Figure 1.8 A. In order to develop light responsive enzyme, allosteric photocaging of *EcDHFR* was achieved by inserting AsLOV2 sequence into allosteric loop regions including loop A and B, Figure 1.8 B. By inserting AsLOV2 into loop A<sup>132</sup>, Figure 1.8 C, it was observed that DHFR reaction was two-fold faster upon photoactivation possibly due to the unfolding of the AsLOV2 J $\alpha$ , Figure 1.8 C. This example illustrated a potential mechanism to photocage desired proteins. Nonetheless, it should be noted that this engineered chimera indicated approximately 60-fold slower catalytic turnover where the hydride transfer rate was approximately 1,000 times slower.

Another engineered LOV domain chimera has been engineered by effector module swapping<sup>120,140</sup>, Figure 1.8 F. As mentioned earlier, LOV domains belong to PAS domain superfamily of proteins that commonly contain fused C-terminus effector modules<sup>141</sup>. This therefore suggests a common underlying structural mechanism of signal transduction from the PAS/LOV domain core. This underlying mechanism has been shown to hold true for PAS/LOV domain proteins that use C-terminal coiled-coil rearrangements as a signal transduction mechanism<sup>120,140,141</sup>. This was evident after the generation of a light responsive chimera between *BsYtvA* LOV domain<sup>120,140</sup> and sensory histidine kinase (HK) domain of oxygen-sensitive PAS domain protein *Bradyrhizobium japonicum* FixL (*BjFixL*), Figure 1.8 F. The engineered YF1 chimera indicated an astonishing 1,000-fold light dependent decrease in the kinase activity. These results therefore provided a novel mechanism to control protein activity with light and to date this approach has been used in the development of multiple different light responsive enzymes where  $\alpha$ -helical rearrangements trigger changes in the protein activity.



**Figure 1.8 Early engineering of light responsive proteins.**

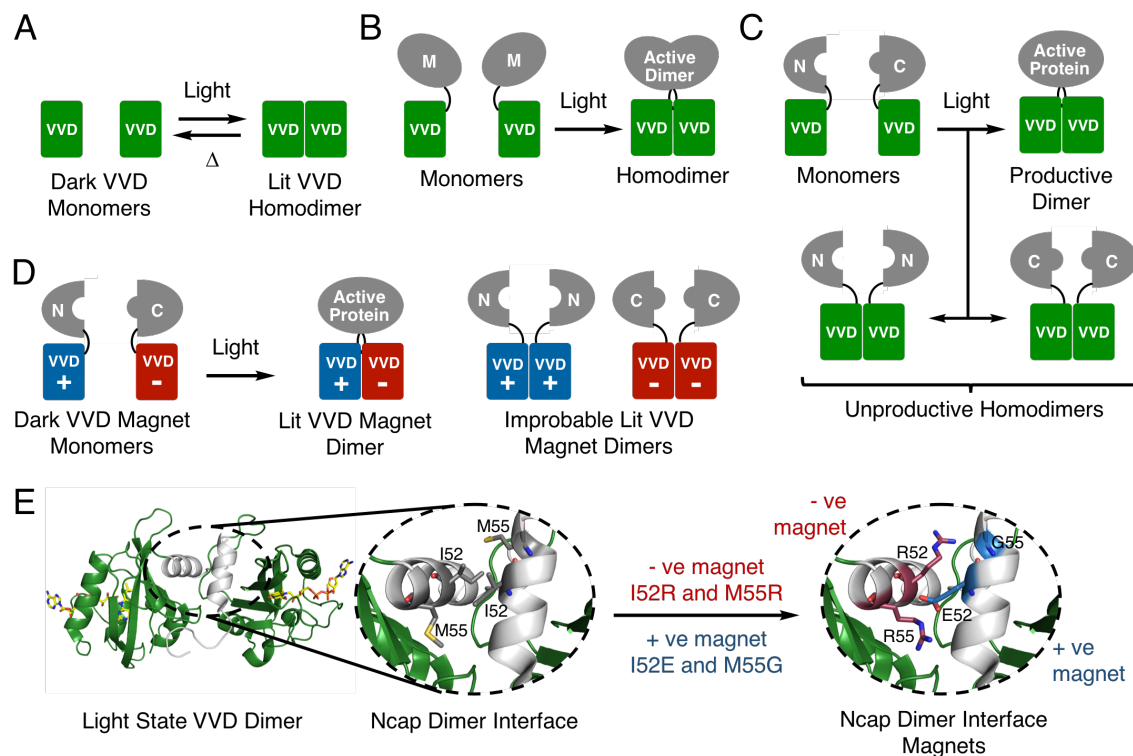
A) Reaction catalysed by *Escherichia coli* dihydrofolate reductase (*Ec*DHFR) where 7,8-dihydrofolate (DHF) is reduced to 5,6,7,8-tetrahydrofolate (THF) consuming a molecule of NADPH as a redox cosubstrate. B) Cartoon representation of *Ec*DHFR indicating two allosteric loops A and B. C) Chimera generated by inserting AsLOV2 sequence into loop A of *Ec*DHFR. In the dark state folded J $\alpha$  of AsLOV2 inhibits *Ec*DHFR reaction by preventing possible loop movements. Upon photoactivation J $\alpha$  unfolds liberating loop A facilitating faster reaction. D) Solution NMR structure of YtvA (PDB 2MWG) indicating dimeric LOV domain core (green), coiled-coil J $\alpha$  helices (red) and STAS domain (purple). E) Generation of YF1 chimera, where YtvA LOV domain core (fragment A) was fused to the FixL C-terminus histidine kinase (HK) domain including coiled-coil forming J $\alpha$  helices (fragment B). Generated YF1 chimera indicated light dependent 1,000-fold decrease in the kinase activity. F) Crystal structure of YF1 chimera (PDB 4GCZ) indicating YtvA LOV domain core (green), J $\alpha$  coiled-coil linker (red) and FixL histidine kinase dimer including dimerisation/histidine phosphotransfer (DHp) domain (orange) and catalytic/ATP-binding (CA) domain (grey). G) Crystal structure of TrpR and DNA complex with N-terminus helix depicted in black (PDB 1TRO). H) Cartoon representation indicating photocaging of TrpR with AsLOV2. In dark state, folded J $\alpha$  helix and LOV domain core sterically hinder DNA binding. Upon photoactivation, J $\alpha$  unfolds liberating TrpR and allowing DNA binding

Light responsive transcription factors are ubiquitous LOV domain proteins found in many microorganisms. Nonetheless, light responsive engineered transcription factors have been developed. The earliest example demonstrating how a LOV domain can be used to allosterically cage a dedicated transcription factor has been demonstrated for *E. coli* tryptophan repressor (*Ec*TrpR)<sup>142</sup>. *Ec*TrpR is a Trp biosynthesis repressor where high

intracellular Trp concentrations result in *EcTrpR*-Trp dimeric complex formation triggering conformational change and tight DNA binding<sup>143,144</sup>. Earlier structural studies of *EcTrpR*-Trp-DNA complex indicated N-terminal helix, Figure 1.8 G<sup>145</sup>. This observation suggested a potential mechanism where the steric hindrance of N-terminus helix would inhibit DNA binding, Figure 1.8 G and H. Considering these observations, a chimera was generated where *AsLOV2* was fused to *EcTrpR*<sup>142</sup>. Construct was generated by joining the C-terminus of the *AsLOV2* J $\alpha$  helix with the N-terminus of *EcTrpR* from phenylalanine 22 (F22) residue generating LovTAP. This construct showed light dependent enhanced DNA binding following a mechanism depicted in Figure 1.8 H. In the dark, J $\alpha$  packs against the *AsLOV2*  $\beta$ -sheet surface sterically hindering DNA binding. Upon photoactivation, unfolding of J $\alpha$  alleviates steric hindrance allowing tighter DNA binding. The initial LovTAP construct indicated poor light dependent DNA binding,  $\sim$  5.5-fold enhancement upon photoactivation. Further improvements were achieved by rationally designing *AsLOV2* constructs with a more tightly packed J $\alpha$  helix increasing the dynamic range,  $\sim$ 70 fold increased DNA binding upon photoactivation<sup>146</sup>. These studies therefore provided evidence that through understanding of LOV domain structural mechanism of activation it is possible to engineer novel LOV domain chimeras also allowing to rationally increase the dynamic range. This therefore highlights the importance of understanding of LOV domain photochemistry which in many cases is limited to only few proteins, for instance *AsLOV2*, VVD and YtvA, and larger repertoire is required.

Light dependent LOV domain dimerisation has also been exploited in the developed of optogenetic tools. Considering known *NcVVD* light dependent dimerisation, Figure 1.9 A, this mechanism has been used to develop light responsive systems where dimerisation allows activation of the target protein<sup>147</sup>, Figure 1.9 B. This approach has been used to develop light responsive transcription repressor using *E. coli* LexA as a starting point<sup>147</sup>. LexA is a transcription repressor requiring LexA homodimerisation for activity and DNA binding<sup>148</sup>. This mechanistic understanding has been used to generate a chimera where *NcVVD* was fused to the LexA C-terminus, where it was hypothesised that light dependent *NcVVD* dimerisation would trigger LexA DNA binding in a light dependent manner. Screening of multiple fusion constructs identified a chimera referred to as LEVI construct. *In vivo* studies indicated  $\sim$ 10,000-fold lower recombinant protein production when cultures were subjected to blue light indicating a very tightly controlled system. Although it was an astonishing feat, LOV domain triggered homodimerisation limits applicability, Figure 1.9 C. For instance, if split protein segments including separate N- and C-terminus are fused to *NcVVD*, light dependent homodimerisation of *NcVVD* would result

in inhomogeneous dimers including active heterodimers and inactive homodimers, Figure 1.9 C. To overcome this challenge, yet another system, referred to as magnets, has been developed allowing light dependent heterodimerisation of engineered *NcVVD* mutants<sup>133</sup>, Figure 1.9 D. Light dependent heterodimerisation of *NcVVD* was achieved by introducing mutations at the N-terminus Ncap (A'α helix) dimer interface by replacing hydrophobic residues with positively (-ve magnet) or negatively (+ve magnet) charged residues<sup>133</sup>, Figure 1.9 E.



**Figure 1.9 Engineering of VVD magnets.**

A) Light triggered dimerisation of *Neurospora crassa* Vivid (*NcVVD*). B) An approach to engineer optogenetic tools dependent on light induced dimerisation. Here, light dependent dimerisation of *NcVVD* allows generation of an active homodimer. C) Indicates limitations of light induced dimerisation approach in engineering of light responsive proteins. If split protein constructs, for instance N- and C-terminus, are fused to *NcVVD*, light induced dimerisation can result in generation of an active heterodimeric protein (top figure) and inactive homodimers. D) Generation of *NcVVD* magnets system where +ve VVD construct can only dimerise with -ve *NcVVD* construct in light dependent manner. This system overcomes challenges explained in C). E) Structure of lit state dimer of *NcVVD* (PDB 3RH8) indicating dimer interface trough Ncap (A'α). Inspection of the dimer interface indicates hydrophobic residues including isoleucine52 (I52) and methionine55 (M55). Replacement of these hydrophobic residues with corresponding charged arginine (R) and glutamate (E) allowed generation of +ve and -ve *NcVVD* magnets.



Considering introduction of polar residues at the dimer interface, no homodimerisation was observed due to same charge repulsion. This heterodimerisation driven activation of the protein activity was elegantly shown for a split firefly luciferase. In the reported work, N- and C-terminus fragments of luciferase were fused to the corresponding +ve and -ve magnets<sup>133</sup>. From bioluminescence studies it was evident that light dependent magnet heterodimerisation triggered N- and C-terminus luciferase fragment dimerisation restoring luciferase activity.

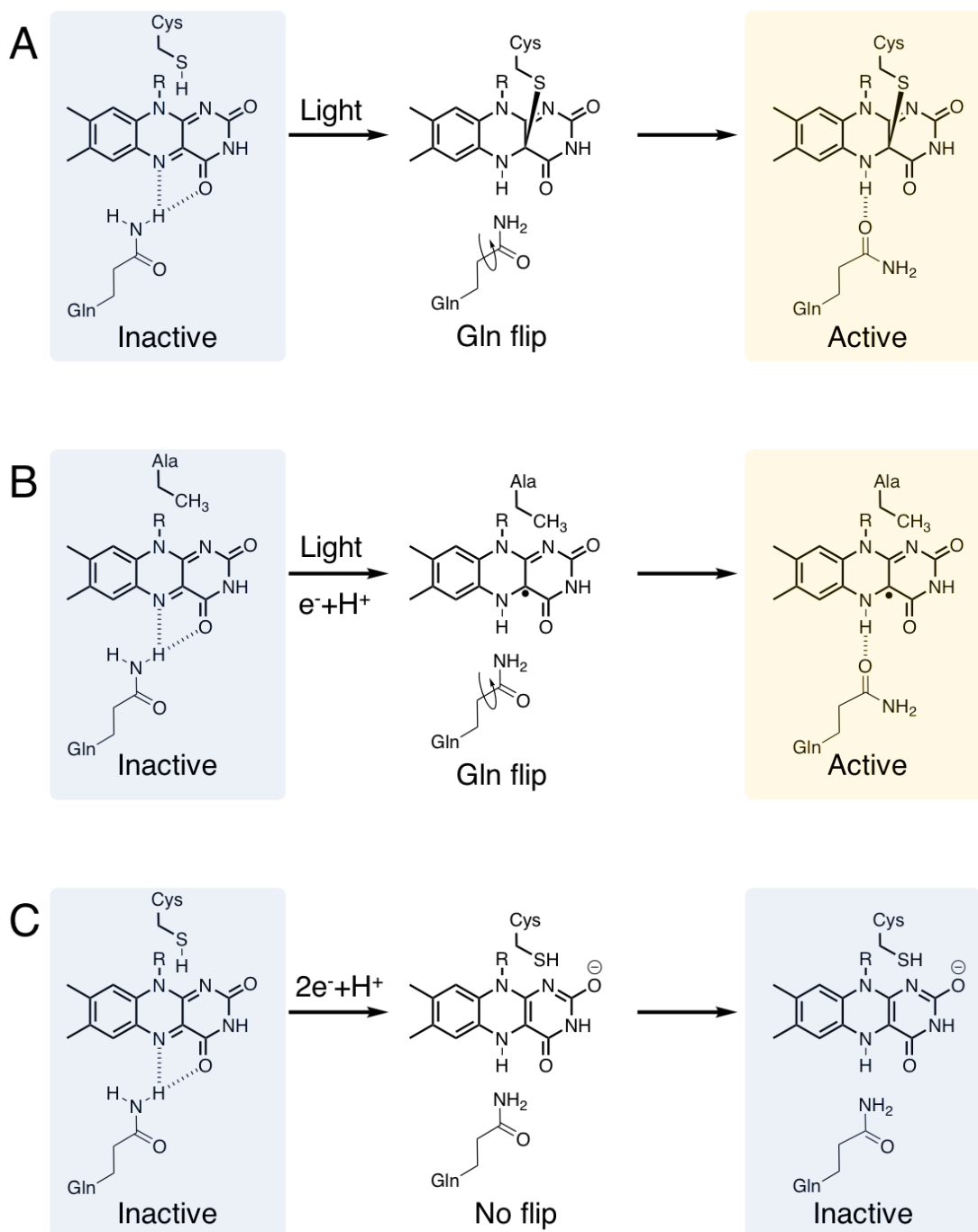
To date, the understanding of LOV domain mechanism of activation has been applied in the development of multiple light controlled systems. Unlike small molecule ligands, light provides non-invasive, dose dependent and traceless approach to manipulate complex systems in a spatiotemporal manner. Furthermore, considering the fact that LOV domains revert back to dark adapted state provides a dynamic mechanism where without a stimuli biological response ceases in a time dependent manner. In order to assist with the rational design of engineered systems, better understanding of LOV domain structural mechanisms of activation is required. Considering that photoactivated proteins undergo large scale structural changes, a method to rapidly probe structural changes in a cost effective manner is required. To date, no inexpensive method has been developed, where the current approach relies largely on time-consuming NMR experiments requiring protein labelling with <sup>15</sup>N, <sup>13</sup>C and <sup>2</sup>H.

### **1.7 LOV Domain Light Sensing, Structural and Mechanistic Insights**

As mentioned earlier, LOV domains act as photosensory modules able to allosterically activate or inactivate inter- and intramolecular protein partners in a light dependent manner. The structure of the LOV domain core is highly conserved, but the mechanism of signal transduction and subsequent activation of the effector domains is not<sup>149</sup>. Effector domains are regulated through proposed rearrangements of either N- and C-terminal auxiliary helical extensions<sup>149</sup>. These auxiliary helices including A'α (Ncap) and Jα (Ccap) vary widely and pack to the β-sheet surface of the LOV domain core. To date, the allosteric mechanism how the covalent adduct formation causes residue rearrangements within the flavin binding pocket and how these changes extend to the allosteric surface remains less well characterised. This is mostly attributed to the lack of high-resolution structures describing the lit state of LOV domain proteins. So far the hypothesis that has been put forward suggests that the photoadduct formation facilitates positional changes of the conserved glutamine residue located on the final LOV domain β-strand (βI) in the plane of

the flavin ring system<sup>150–152</sup>. Protonation at the N5 position, Figure 1.10 A, has been suggested to dictate glutamine reorientation resulting in changes of the polarity of a hydrogen bond donor/acceptor networks<sup>150–152</sup>. This change then triggers structural rearrangements within LOV domain core and changes in the dynamics of the  $\beta$ -sheet surface. Mutagenesis studies have shown that this glutamine residue is vital for correct LOV function in distantly related LOV domain protein including *AsLOV2*<sup>151</sup> and *NcVVD*<sup>95,151,152</sup>. Furthermore, the crystal structure of the lit state *NcVVD* has suggested the rotation of the glutamine carboxamide side chain<sup>106</sup>, although the resolution of the crystal structure (2.75 Å) is too low to assert the rotamer identity with certainty. Nonetheless, recent molecular dynamics (MD) studies supports this hypothesis suggesting that protonation of the N5 upon lit state formation triggers such structural rearrangements<sup>152</sup>.

Flavin protonation and sequential glutamine flip hypothesis is furthermore supported by studies of photochemically reduced LOV domain proteins where the reactive cysteine residue was replaced by a chemically inert alanine<sup>153,154</sup>. Studies of *NcVVD* and chimera YF1 protein mutants, where the active site cysteines were replaced with an alanine, indicated that photochemical generation of a neutral flavin semiquinone radical resulted in light dependent activation *in vitro* and *in vivo*, respectively<sup>153,154</sup>. For *NcVVD* C108A mutant, light dependent generation of neutral semiquinone radical resulted in generation of a dimeric species identical to the WT protein. Furthermore, electron paramagnetic resonance (EPR) studies indicated that the distance between the two radical centres was in line with crystallographic lit state dimer. Additionally, YF1 C62A mutant indicated light dependent activity comparable to the WT construct *in vivo*. From these studies, it was concluded that generation of neutral semiquinone flavin radical and protonation of the N5 suffice to activate LOV domain proteins. This conclusion is supported by later molecular dynamics (MD) studies of dark, lit and reduced FAD semiquinone states of *NcVVD* indicating that protonation of N5 is enough to trigger structural changes through the conserved glutamine flip<sup>152</sup>. This therefore strengthens the hypothesis that protonation of N5 is the key factor facilitating LOV domain structural changes. To date, similar observations have been reported for *Au1a<sub>LOV</sub>* domain proteins indicating that the generation of neutral semiquinone radical species triggers dimerisation<sup>155</sup>.



**Figure 1.10 Proposed mechanism of LOV domain activation by light.**

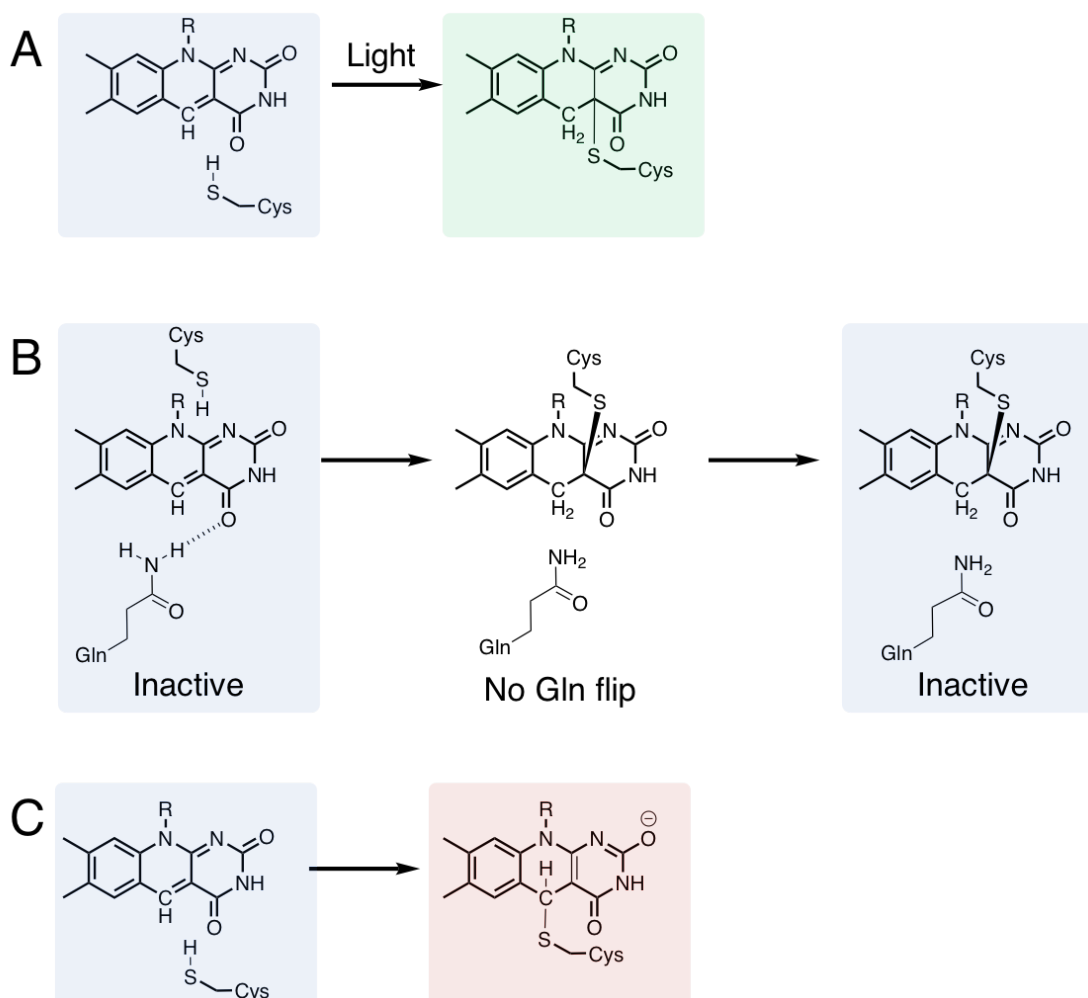
A) Photochemical generation of cysteinyl-flavin photoadduct triggers a flip of a conserved glutamine (Gln) sidechain ultimately resulting in changes in hydrogen bond networks across LOV domain core. B) Photochemically generated flavin semiquinone radical has also been shown to trigger structural changes in LOV domain due to generation of protonated flavin N5. C) Chemical reduction of flavin in LOV domain core. Although reduced flavin is protonated an N5, the expected structural changes is not enforced.

In general, the mechanism how the lit state governs structural change has been suggested to involve the mechanism depicted in the Figure 1.10. It is evident that upon the photoadduct formation FMN-N5 becomes protonated. This key change therefore triggers glutamine repositioning allowing structural changes within the LOV domain core extending to the allosteric regions including N- and C-terminal auxiliary helices. Noteworthy, ambiguity regarding this mechanism does exist. A study of *Caulobacter crescentus* LOV-histidine kinase (CdLOV-HK) indicated that chemical reduction of the FMN to FMNH<sup>-</sup> resulted in the histidine kinase (HK) domain activity similar to that of the dark state protein<sup>156</sup>. Considering that FMNH<sup>-</sup> is a protonated species at the N5, it would be expected that HK would be activated upon reduction. This intriguing observation therefore raises question if protonation of the N5 is the only determinant allowing activation of LOV domain proteins. Nonetheless, to the best of my knowledge, no further information exists in the literature describing LOV domain proteins containing reduced flavin as a cofactor and information if these proteins are active.

## 1.8 Project Aims and Purpose

In this work, it was decided to use 5-deazaflavin mononucleotide (5dFMN) to probe the structural mechanism of LOV domain activation by light<sup>157</sup>, Figure 1.11 A. This analogue is capable of forming a photochemical cysteinyl-5dFMN photoadduct, but to date it was unknown if 5dFMN forms the covalent photoadduct at an identical C4a position as in native FMN. Molecular dynamic studies suggested a covalent bond formation at C4a in line with the proposed radical pair mechanism<sup>157</sup>, but no structural information was provided to support theoretical calculations. It is also possible that light could trigger Michael addition generating cysteinyl-5dFMN-C5 covalent photoadduct, Figure 1.11 C. Furthermore, no structural information has been reported for 5dFMN reconstituted LOV domains in the literature. Considering current standing hypothesis, 5dFMN lacks hydrogen bond acceptor in the dark state and donor in the lit state, respectively. This chemical difference therefore suggests that this cofactor should be inactive where photoactivation would not trigger any observable structural rearrangements. Alternatively, if protonation of flavin-N5 is not the only requirement for light dependent structural changes, 5dFMN reconstituted LOV domain should indicate observable structural changes. From a chemical viewpoint, it is also possible that steric effects such as puckering of the conjugated flavin ring system away from planarity and interruption of conjugated aromatic ring system could contribute

to net structural changes. Understanding this fundamental aspect LOV domain structural changes could potentially assist with the rational design of novel optogenetic tools.



**Figure 1.11 Photochemical reaction of LOV domain proteins incorporated with flavin analogue.**

A) Proposed photochemical reaction between a highly conserved cysteine residue located in LOV domain core and a flavin analogue, 5-deaza-flavin (5dFMN). B) Proposed mechanism suggesting that due to a lack of hydrogen bond acceptor in the dark and donor in the lit state would not trigger activation of LOV domain protein. This is due to proposed mechanism where it has been suggested that protonation of flavin at N5 position triggers Gln residue flip furthermore facilitating structural changes in LOV domain core. C) Proposed light triggered Michael addition taking place in LOV domain reconstituted with 5dFMN.

## 1.9 References

1. Mansoorabadi, S. O., Thibodeaux, C. J. & Liu, H. The Diverse Roles of Flavin Coenzymes; Nature's Most Versatile Thespians. *J. Org. Chem.* **72**, 6329–6342 (2007).
2. Walsh, C. T. & Wencewicz, T. A. Flavoenzymes: Versatile Catalysts in Biosynthetic Pathways. *Nat. Prod. Rep.* **30**, 175-200 (2013).
3. MacHeroux, P., Kappes, B. & Ealick, S. E. Flavogenomics - A Genomic and Structural View of Flavin-Dependent Proteins. *FEBS J.* **278**, 2625-2634 (2011).
4. Bacher, A., Eberhardt, S., Fischer, M., Kis, K. & Richter, G. Biosynthesis of Vitamin B2 (Riboflavin). *Annu. Rev. Nutr.* **20**, 153-67 (2000).
5. Fischer, M. & Bacher, A. Biosynthesis of Vitamin B2: Structure and Mechanism of Riboflavin Synthase. *Arch. Biochem. Biophys.* **474**, 252-65 (2008).
6. Abbas, C. A. & Sibirny, A. A. Genetic Control of Biosynthesis and Transport of Riboflavin and Flavin Nucleotides and Construction of Robust Biotechnological Producers. *Microbiol. Mol. Biol. Rev.* **75**, 321-60 (2011).
7. Mráček, T., Drahota, Z. & Houštěk, J. The Function and the Role of the Mitochondrial Glycerol-3-Phosphate Dehydrogenase in Mammalian Tissues. *Biochim. Biophys. Acta.* **1827**, 401-410 (2013).
8. Yeh, J. I., Chinte, U. & Du, S. Structure of Glycerol-3-Phosphate Dehydrogenase, an Essential Monotopic Membrane Enzyme Involved in Respiration and Metabolism. *PNAS* **105**, 3280-3285 (2008).
9. Sun, F., Huo, X., Zhai, Y., Wang, A., Xu, J., Su, D., Bartlam, M. & Rao, Z. Crystal Structure of Mitochondrial Respiratory Membrane Protein Complex II. *Cell* **121**, 1043-1057 (2005).
10. Liu, S., Neidhardt, E. A., Grossman, T. H., Ocain, T. & Clardy, J. Structures of Human Dihydroorotate Dehydrogenase in Complex with Antiproliferative Agents. *Structure* **8**, 25-33 (2000).
11. Eswaramoorthy, S., Bonanno, J. B., Burley, S. K. & Swaminathan, S. Mechanism of Action of a Flavin-Containing Monooxygenase. *PNAS* **103**, 9832-9837 (2006).
12. Krueger, S. K. & Williams, D. E. Mammalian Flavin-Containing Monooxygenases: Structure/Function, Genetic Polymorphisms and Role in Drug Metabolism. *Pharmacol. Ther.* **106**, 357-387 (2005).
13. Beam, M. P., Bosserman, M. A., Noinaj, N., Wehenkel, M. & Rohr, J. Crystal Structure of Baeyer-Villiger Monooxygenase MtmOIV, the Key Enzyme of the Mithramycin Biosynthetic Pathway. *Biochemistry* **48**, 4476-4487 (2009).

14. Campbell, Z. T., Weichsel, A., Montfort, W. R. & Baldwin, T. O. Crystal Structure of the Bacterial Luciferase/Flavin Complex Provides Insight into the Function of the Alpha Subunit. *Biochemistry* **48**, 6085-6094 (2009).
15. Baldwin, T. O., Christopher, J. A., Raushel, F. M., Sinclair, J. F., Ziegler, M. M., Fisher, A. J. & Rayment, I. Structure of Bacterial Luciferase. *Curr. Opin. Struct. Biol.* **5**, 798-809 (1995).
16. Fisher, A. J., Thompson, T. B., Thoden, J. B., Baldwin, T. O. & Rayment, I. The 1.5-Å Resolution Crystal Structure of Bacterial Luciferase in Low Salt Conditions. *J. Biol. Chem.* **271**, 21956-21968 (1996).
17. Harper, S. M., Neil, L. C. & Gardner, K. H. Structural Basis of a Phototrophin Light Switch. *Science* **301**, 1541-1544 (2003).
18. De Colibus, L. & Mattevi, A. New Frontiers in Structural Flavoenzymology. *Curr. Opin. Struct. Biol.* **16**, 722-728 (2006).
19. Karthikeyan, S., Zhou, Q., Mseeh, F., Grishin, N. V., Osterman, A. L. & Zhang, H. Crystal Structure of Human Riboflavin Kinase Reveals a  $\beta$ -barrel Fold and a Novel Active Site Arch. *Structure* **11**, 265-273 (2003).
20. Sebastián, M., Lira-Navarrete, E., Serrano, A., Marcuello, C., Velázquez-Campoy, A., Lostao, A., Hurtado-Guerrero, R., Medina, M. & Martínez-Júlvez, M. The FAD Synthetase from the Human Pathogen *Streptococcus pneumoniae*: A Bifunctional Enzyme Exhibiting Activity-dependent Redox Requirements. *Sci. Rep.* **7**, 7609 (2017).
21. Ernst, O. P., Lodowski, D. T., Elstner, M., Hegemann, P., Brown, L. S. & Kandori, H. Microbial and Animal Rhodopsins: Structures, Functions, and Molecular Mechanisms. *Chem. Rev.* **114**, 126-163 (2014).
22. Demarsy, E. & Fankhauser, C. Higher Plants Use LOV to Perceive Blue Light. *Curr. Opin. Plant Biol.* **12**, 69-74 (2009).
23. Hasegawa, K., Masuda, S. & Ono, T. A. Light Induced Structural Changes of a Full-Length Protein and its BLUF Domain in YcgF(Blrp), a Blue-Light Sensing Protein that Uses FAD (BLUF). *Biochemistry* **45**, 3785-3793 (2006).
24. Chaves, I. *et al.* The Cryptochromes: Blue Light Photoreceptors in Plants and Animals. *Annu. Rev. Plant Biol.* **62**, 335-364 (2011).
25. Heijde, M. & Ulm, R. Reversion of the Arabidopsis UV-B photoreceptor UVR8 to the Homodimeric Ground State. *PNAS* **110**, 1113-1118 (2013).

26. Favory, J. J. *et al.* Interaction of COP1 and UVR8 Regulates UV-B-induced Photomorphogenesis and Stress acclimation in *Arabidopsis*. *EMBO J.* **28**, 591-601 (2009).
27. Möglich, A., Yang, X., Ayers, R. A. & Moffat, K. Structure and Function of Plant Photoreceptors. *Annu. Rev. Plant Biol.* **61**, 21-47 (2010).
28. Montgomery, B. L. & Lagarias, J. C. Phytochrome Ancestry: Sensors of Bilins and Light. *Trends Plant Sci.* **7**, 357-366 (2002).
29. Tilbrook, K., Arongaus, A. B., Binkert, M., Heijde, M., Yin, R. & Ulm, R. The UVR8 UV-B Photoreceptor: Perception, Signaling and Response. *Arab. B.* **11**, e0164 (2013).
30. Hayes, S., Velanis, C. N., Jenkins, G. I. & Franklin, K. A. UV-B Detected by the UVR8 Photoreceptor Antagonizes Auxin Signaling and Plant Shade Avoidance. *PNAS* **111**, 11894-11899 (2014).
31. Deisseroth, K. Optogenetics: 10 Years of Microbial Opsins in Neuroscience. *Nat. Neurosci.* 2015 **8**, 1213-1225 (2015).
32. Terakita, A. The Opsins. *Genome Biol.* **6**, 213 (2005).
33. Zhang, F. *et al.* The microbial opsin family of optogenetic tools. *Cell* **147**, 1446-1457 (2011).
34. Bridwell-Rabb, J. & Drennan, C. L. Vitamin B12 in the Spotlight Again. *Curr. Opin. Chem. Biol.* **37**, 63-70 (2017).
- Jung, J. H. *et al.* Phytochromes Function as thermosensors in *Arabidopsis*. *Science* **354**, 886-889 (2016).
36. Li, J., Li, G., Wang, H. & Wang Deng, X. Phytochrome Signaling Mechanisms. *Arab. B.* **9**, e0148 (2011).
37. Swartz, T. E. *et al.* Blue-Light-Activated Histidine Kinases: Two-Component Sensors in Bacteria. *Science* **317**, 1090-1093 (2007).
38. Cao, Z., Buttani, V., Losi, A. & Gärtner, W. A blue light inducible two-component signal transduction system in the plant pathogen *Pseudomonas syringae* pv. tomato. *Biophys. J.* **94**, 897-905 (2008).
39. Mascher, T., Helmann, J. D. & Udden, G. Stimulus Perception in Bacterial Signal-Transducing Histidine Kinases. *Microbiol. Mol. Biol. Rev.* **70**, 910-938 (2006).
40. Casino, P., Miguel-Romero, L. & Marina, A. Visualizing autophosphorylation in histidine kinases. *Nat. Commun.* **5**, 3258 (2014).



41. Delgado-Nixon, V. M., Gonzalez, G. & Gilles-Gonzalez, M. A. Dos, a heme-binding PAS protein from *Escherichia coli*, is a direct oxygen sensor. *Biochemistry* **39**, 2685-2691 (2000).
42. Janausch, I. G., Zientz, E., Tran, Q. H., Kröger, A. & Uden, G. C4-dicarboxylate carriers and sensors in bacteria. *Biochim. Biophys. Acta.* **1553**, 39-56 (2002).
43. Uden, G., Strecker, A., Kleefeld, A. & Kim, O. Bin. C4-Dicarboxylate Utilization in Aerobic and Anaerobic Growth. *EcoSal Plus* **7** (2016).
44. Singh, K., Senadheera, D. B. & Cvitkovitch, D. G. An intimate link: Two-component signal transduction systems and metal transport systems in bacteria. *Future Microbiol.* **9**, 1283-1293 (2014).
45. Ahmad, M. & Cashmore, A. R. HY4 gene of *A. thaliana* encodes a protein with characteristics of a blue-light photoreceptor. *Nature* **366**, 162-166 (1993).
46. Brettel, K. & Byrdin, M. Reaction mechanisms of DNA photolyase. *Curr. Opin. Struct. Biol.* **20**, 693-701 (2010).
47. Maul, M. J., Barends, T. R., Glas, A. F., Cryle, M. J., Domratcheva, T., Schneider, S., Schlichting, I. & Carell, T. Crystal structure and mechanism of a DNA (6-4) photolyase. *Angew. Chem. Int. Ed. Engl.* **47**, 10076–10080 (2008).
48. Yu, X., Liu, H., Klejnot, J. & Lin, C. The Cryptochrome Blue Light Receptors. *Arab. B.* **8**, e0135 (2010).
49. Cashmore, A. R., Jarillo, J. A., Wu, Y. J. & Liu, D. Cryptochromes: Blue light receptors for plants and animals. *Science* **284**, 760-765 (1999).
50. Lin, C. Plant blue-light receptors. *Trends Plant Sci.* **5**, 337-342 (2000).
51. Tamada, T., Kitadokoro, K., Higuchi, Y., Inaka, K., Yasui, A., de Ruiter, P. E., Eker, A. P. & Miki, K. Crystal structure of DNA photolyase from *Anacystis nidulans*. *Nat. Struct. Biol.* **4**, 887-891 (1997).
52. Heelis, P. F. The photophysical and photochemical properties of flavins (isoalloxazines). *Chem. Soc. Rev.* **11**, 15-39 (1982).
53. Edwards, A. M. Structure and general properties of flavins. *Methods Mol. Biol.* **1146**, 3-13 (2014).
54. Zoltowski, B. D. & Gardner, K. H. Tripping the light fantastic: Blue-light photoreceptors as examples of environmentally modulated protein-protein interactions. *Biochemistry* **50**, 4-16 (2011).
55. Kennis, J. T. & Groot, M. L. Ultrafast spectroscopy of biological photoreceptors. *Curr. Opin. Struct. Biol.* **17**, 623-630 (2007).

56. Losi, A. & Gärtner, W. Old chromophores, new photoactivation paradigms, trendy applications: Flavins in blue light-sensing photoreceptors. *Photochem Photobiol.* **87**, 491-510 (2011).
57. Nöll, G. Spectroscopic investigation of flavoproteins: Mechanistic differences between (electro)chemical and photochemical reduction and oxidation. *J. Photochem. Photobiol. B.* **152**, 71-81 (2008).
58. Kao, Y. T., Tan, C., Song, S. H., Öztürk, N., Li, J., Wang, L., Sancar, A. & Zhong, D.. Ultrafast dynamics and anionic active states of the flavin cofactor in cryptochrome and photolyase. *J. Am. Chem. Soc.* **130**, 7695-7701 (2008).
59. Fedorov, R., Schlichting, I., Hartmann, E., Domratcheva, T., Fuhrmann, M. & Hegemann, P. Crystal structures and molecular mechanism of a light-induced signaling switch: The Phot-LOV1 domain from *Chlamydomonas reinhardtii*. *Biophys. J.* **84**, 2474–2482 (2003).
60. Torra, J., Burgos-Caminal, A., Endres, S., Wingen, M., Drepper, T., Gensch, T., Ruiz-González, R. & Nonell, S. Singlet oxygen photosensitisation by the fluorescent protein Pp2FbFP L30M, a novel derivative of *Pseudomonas putida* flavin-binding Pp2FbFP. *Photochem. Photobiol. Sci.* **14**, 280-287 (2015).
61. Shu, X., Lev-Ram, V., Deerinck, T. J., Qi, Y., Ramko, E. B., Davidson, M. W., Jin, Y., Ellisman, M. H. & Tsien, R. Y. A genetically encoded tag for correlated light and electron microscopy of intact cells, tissues, and organisms. *PLoS Biol.* **9**, e1001041 (2011).
62. Torra, J., Lafaye, C., Signor, L., Aumonier, S., Flors, C., Shu, X., Nonell, S., Gotthard, G. & Royant, A. Tailing miniSOG: structural bases of the complex photophysics of a flavin-binding singlet oxygen photosensitizing protein. *Sci. Rep.* **9**, 2428 (2019).
63. Öztürk, N., Song, S. H., Selby, C. P. & Sancar, A. Animal Type 1 cryptochromes: Analysis of the redox state of the flavin cofactor by site-directed mutagenesis. *J. Biol. Chem.* **283**, 3256-3263 (2008).
64. Heijde, M. *et al.* Characterization of two members of the cryptochrome/photolyase family from *Ostreococcus tauri* provides insights into the origin and evolution of cryptochromes. *Plant Cell Environ.* **33**, 1614-1626 (2010).
65. Franz, S. *et al.* Structure of the bifunctional cryptochrome aCRY from *Chlamydomonas reinhardtii*. *Nucleic Acids Res.* **46**, 8010-8022 (2018).

66. in, C. & Shalitin, D. Cryptochrome structure and signal transduction. *Annu. Rev. Plant Biol.* **54**, 469-496 (2003).
67. Lin, C., Robertson, D. E., Ahmad, M., Raibekas, A. A., Jorns, M. S., Dutton, P. L. & Cashmore, A. R. Association of flavin adenine dinucleotide with the Arabidopsis blue light receptor CRY1. *Science* **269**, 968-970 (1995).
68. Partch, C. L. & Sancar, A. Cryptochromes and circadian photoreception in animals. *Methods in Enzymol.* **393**, 726-745 (2005).
69. Lin, C., Top, D., Manahan, C. C., Young, M. W. & Crane, B. R. Circadian clock activity of cryptochrome relies on tryptophan-mediated photoreduction. *PNAS* **115**, 3822-3827 (2018).
70. Song, S. H., Oztürk, N., Denaro, T. R., Arat, N. O., Kao, Y. T., Zhu, H., Zhong, D., Reppert, S. M. & Sancar, A. Formation and function of flavin anion radical in cryptochrome 1 blue-light photoreceptor of monarch butterfly. *J. Biol. Chem.* **282**, 17608-17612 (2007).
71. Ozturk, N., Selby, C. P., Annayev, Y., Zhong, D. & Sancar, A. Reaction mechanism of *Drosophila* cryptochrome. *PNAS* **108**, 516-521 (2011).
72. Liu, H., Liu, B., Zhao, C., Pepper, M. & Lin, C. The action mechanisms of plant cryptochromes. *Trends Plant Sci.* **16**, 684-691 (2011).
73. Anderson, S., Dragnea, V., Masuda, S., Ybe, J., Moffat, K., Bauer, C. Structure of a novel photoreceptor, the BLUF domain of AppA from *Rhodobacter sphaeroides*. *Biochemistry* **44**, 7998-8005 (2005).
74. Gomelsky, M. & Klug, G. BLUF: A novel FAD-binding domain involved in sensory transduction in microorganisms. *Trends Biochem Sci.* **27**, 497-500 (2002).
75. Jung, A., Domratcheva, T., Tarutina, M., Wu, Q., Ko, W. H., Shoeman, R. L., Gomelsky, M., Gardner, K. H. & Schlichting, I. Structure of a bacterial BLUF photoreceptor: Insights into blue light-mediated signal transduction. *PNAS* **102**, 12350-12355 (2005).
76. Masuda, S. Light detection and signal transduction in the BLUF photoreceptors. *Plant Cell Physiol.* **54**, 171-179 (2013).
77. Masuda, S. & Bauer, C. E. AppA is a blue light photoreceptor that antirepresses photosynthesis gene expression in *Rhodobacter sphaeroides*. *Cell* **110**, 613-623 (2002).
78. Bonetti, C. *et al.* The role of key amino acids in the photoactivation pathway of the Synechocystis Slr1694 BLUF domain. *Biochemistry* **48**, 11458-11469 (2009).

79. Gauden, M. *et al.* Hydrogen-bond switching through a radical pair mechanism in a flavin-binding photoreceptor. *PNAS* **103**, 10895-10900 (2006).
80. Unno, M., Sano, R., Masuda, S., Ono, T. A. & Yamauchi, S. Light-induced structural changes in the active site of the BLUF domain in AppA by Raman spectroscopy. *J. Phys. Chem. B* **109**, 12620-12626 (2005).
81. Unno, M., Masuda, S., Ono, T. A. & Yamauchi, S. Orientation of a key glutamine residue in the BLUF domain from AppA revealed by mutagenesis, spectroscopy, and quantum chemical calculations. *J. Am. Chem. Soc.* **128**, 5638-5639 (2006).
82. Domratcheva, T., Grigorenko, B. L., Schlichting, I. & Nemukhin, A. V. Molecular models predict light-induced glutamine tautomerization in BLUF photoreceptors. *Biophys. J.* **94**, 3872-3879 (2008).
83. Udvarhelyi, A. & Domratcheva, T. Glutamine rotamers in BLUF photoreceptors: A mechanistic reappraisal. *J. Phys. Chem. B* **117**, 2888-2897 (2013).
84. Collette, F., Renger, T. & Schmidt Am Busch, M. Revealing the functional states in the active site of BLUF photoreceptors from electrochromic shift calculations. *J. Phys. Chem. B* **118**, 11109-11119 (2014).
85. Khrenova, M. G., Nemukhin, A. V. & Domratcheva, T. Photoinduced electron transfer facilitates tautomerization of the conserved signaling glutamine side chain in BLUF protein light sensors. *J. Phys. Chem. B* **117**, 2369-2377 (2013).
86. Domratcheva, T., Hartmann, E., Schlichting, I. & Kottke, T. Evidence for Tautomerisation of Glutamine in BLUF Blue Light Receptors by Vibrational Spectroscopy and Computational Chemistry. *Sci. Rep.* **6**, 22669 (2016).
87. Barends, T. R. M. *et al.* Structure and mechanism of a bacterial light-regulated cyclic nucleotide phosphodiesterase. *Nature* **459**, 1015-1018 (2009).
88. Jenal, U., Reinders, A. & Lori, C. Cyclic di-GMP: Second messenger extraordinaire. *Nat Rev Microbiol.* **15**, 271-284 (2017).
89. Ohki, M. *et al.* Structural insight into photoactivation of an adenylate cyclase from a photosynthetic cyanobacterium. *PNAS* **113**, 6659-6664 (2016).
90. McDonough, K. A. & Rodriguez, A. The myriad roles of cyclic AMP in microbial pathogens: From signal to sword. *Nat Rev Microbiol.* **10**, 27-38 (2012).
91. Lukacs, A. *et al.* BLUF domain function does not require a metastable radical intermediate state. *J. Am. Chem. Soc.* **136**, 4605-4615 (2014).
92. Philip, A. F., Kumauchi, M. & Hoff, W. D. Robustness and evolvability in the functional anatomy of a PER-ARNT-SIM (PAS) domain. *PNAS* **107**, 17986-17991 (2010).

93. Banerjee, A., Herman, E., Kottke, T. & Essen, L. O. Structure of a Native-like Aureochrome 1a LOV Domain Dimer from *Phaeodactylum tricornutum*. *Structure* **24**, 171-178 (2016).
94. Banerjee, A., Herman, E., Kottke, T. & Essen, L. O. Structure of a Native-like Aureochrome 1a LOV Domain Dimer from *Phaeodactylum tricornutum*. *Structure* **24**, 171–178 (2016).
95. Rivera-Cancel, G., Ko, W., Tomchick, D. R., Correa, F. & Gardner, K. H. Full-length structure of a monomeric histidine kinase reveals basis for sensory regulation. *PNAS* **111**, 17839–17844 (2014).
96. Zoltowski, B. D., Schwerdtfeger, C., Widom, J., Loros, J. J., Bilwes, A. M., Dunlap, J. C. & Crane, B. R. Conformational Switching in the Fungal Light Sensor Vivid. *Science* **316**, 1054–1057 (2007).
97. Crosson, S. & Moffat, K. Structure of a flavin-binding plant photoreceptor domain: Insights into light-mediated signal transduction. *PNAS* **98**, 2995–3000 (2001).
98. Möglich, A., Ayers, R. A. & Moffat, K. Structure and Signaling Mechanism of Per-ARNT-Sim Domains. *Structure* **17**, 1282–1294 (2009).
99. Stadler, A. M. *et al.* Photoactivation Reduces Side-Chain Dynamics of a LOV Photoreceptor. *Biophys. J.* **110**, 1064-1074 (2016).
100. Glantz, S. T., Carpenter, E. J., Melkonian, M., Gardner, K. H., Boyden, E. S., Wong, G. K. & Chow, B. Y. Functional and topological diversity of LOV domain photoreceptors. *PNAS* **113**, E1442-E1451 (2016).
101. Röllén, K. *et al.* Signaling States of a Short Blue-Light Photoreceptor Protein PpSB1-LOV Revealed from Crystal Structures and Solution NMR Spectroscopy. *J. Mol. Biol.* **428**, 3721-3736 (2016).
102. Halavaty, A. S. & Moffat, K. N- and C-terminal flanking regions modulate light-induced signal transduction in the LOV2 domain of the blue light sensor phototropin 1 from *Avena sativa*. *Biochemistry* **46**, 14001-14009 (2007).
103. Herman, E. & Kottke, T. Allosterically regulated unfolding of the A'alpha helix exposes the dimerization site of the blue-light-sensing aureochrome-lov domain. *Biochemistry* **54**, 1484-1492 (2015).
104. Takahashi, F. *et al.* AUREOCHROME, a photoreceptor required for photomorphogenesis in stramenopiles. *PNAS* **104**, 19625-19630 (2007).
105. Herman, E., Sachse, M., Kroth, P. G. & Kottke, T. Blue-light-induced unfolding of the Jalpha helix allows for the dimerization of aureochrome-LOV from the diatom *Phaeodactylum tricornutum*. *Biochemistry* **52**, 3094-3101 (2013).

106. Heintz, U. & Schlichting, I. Blue light-induced LOV domain dimerization enhances the affinity of aureochrome 1a for its target DNA sequence. *Elife* **5**, e11860 (2016).
107. Vaidya, A. T., Chen, C. H., Dunlap, J. C., Loros, J. J. & Crane, B. R. Structure of a light-activated LOV protein dimer that regulates transcription. *Sci. Signal.* **4**, ra50 (2011).
108. Chen, C.-H., DeMay, B. S., Gladfelter, A. S., Dunlap, J. C. & Loros, J. J. Physical interaction between VIVID and white collar complex regulates photoadaptation in *Neurospora*. *PNAS*. **107**, 16715-16720 (2010).
109. Conrad, K. S., Bilwes, A. M. & Crane, B. R. Light-induced subunit dissociation by a light-oxygen-voltage domain photoreceptor from *Rhodobacter sphaeroides*. *Biochemistry* **52**, 378-391 (2013).
110. Metz, S., Jäger, A. & Klug, G. Role of a short light, oxygen, voltage (LOV) domain protein in blue light-and singlet oxygen-dependent gene regulation in *Rhodobacter sphaeroides*. *Microbiology* **158**, 368-379 (2012).
111. Harper, S. M., Christie, J. M. & Gardner, K. H. Disruption of the LOV-Jalpha helix interaction activates phototropin kinase activity. *Biochemistry* **43**, 16184-16192 (2004).
112. Bonomi, H. R. *et al.* Light regulates attachment, exopolysaccharide production, and nodulation in *Rhizobium leguminosarum* through a LOV-histidine kinase photoreceptor. *PNAS* **109**, 12135-12140 (2012).
113. Rinaldi, J. *et al.* The beta-scaffold of the LOV domain of the brucella light-activated histidine kinase is a key element for signal transduction. *J. Mol. Biol.* **420**, 112-127 (2012).
114. Möglich, A. & Moffat, K. Structural Basis for Light-dependent Signaling in the Dimeric LOV Domain of the Photosensor YtvA. *J. Mol. Biol.* **373**, 112-126 (2007).
115. Nash, A. I., McNulty, R., Shillito, M. E., Swartz, T. E., Bogomolni, R. A., Luecke, H. & Gardner, K.H. Structural basis of photosensitivity in a bacterial light-oxygen-voltage/helix-turn-helix (LOV-HTH) DNA-binding protein. *PNAS* **108**, 9449-9454 (2011).
116. Yao, X., Rosen, M. K. & Gardner, K. H. Estimation of the available free energy in a LOV2-J alpha photoswitch. *Nat. Chem. Biol.* **4**, 491–7 (2008).
117. Herman, E., Sachse, M., Kroth, P. G. & Kottke, T. Blue-light-induced unfolding of the J-alpha helix allows for the dimerization of aureochrome-LOV from the diatom *Phaeodactylum tricornutum*. *Biochemistry* **52**, 3094-3101 (2013).

118. Christie, J. M. Phototropin Blue-Light Receptors. *Annu. Rev. Plant Biol.* **58**, 21-45 (2007).
119. Jurk, M. *et al.* The Switch that Does Not Flip: The Blue-Light Receptor YtvA from *Bacillus subtilis* Adopts an Elongated Dimer Conformation Independent of the Activation State as Revealed by a Combined AUC and SAXS Study. *J. Mol. Biol.* **403**, 78–87 (2010).
120. Berntsson, O. *et al.* Time-Resolved X-Ray Solution Scattering Reveals the Structural Photoactivation of a Light-Oxygen-Voltage Photoreceptor. *Structure* **25**, 933-938 (2017).
121. Diensthuber, R. P., Bommer, M., Gleichmann, T. & Möglich, A. Full-length structure of a sensor histidine kinase pinpoints coaxial coiled coils as signal transducers and modulators. *Structure* **21**, 1127-1136 (2013).
122. Engelhard, C., Diensthuber, R. P., Möglich, A. & Bittl, R. Blue-light reception through quaternary transitions. *Sci. Rep.* **7**, (2017).
123. Rivera-Cancel, G., Motta-Mena, L. B. & Gardner, K. H. Identification of natural and artificial DNA substrates for light-activated LOV-HTH transcription factor EL222. *Biochemistry* **51**, 10024–10034 (2012).
124. Zoltowski, B. D., Motta-Mena, L. B. & Gardner, K. H. Blue light-induced dimerization of a bacterial LOV-HTH DNA-binding protein. *Biochemistry* **52**, 6653-6661 (2013).
125. Motta-Mena, L. B. *et al.* An optogenetic gene expression system with rapid activation and deactivation kinetics. *Nat. Chem. Biol.* **10**, 196-202 (2014).
126. Akiyama, Y., Nakasone, Y., Nakatani, Y., Hisatomi, O. & Terazima, M. Time-Resolved Detection of Light-Induced Dimerization of Monomeric Aureochrome-1 and Change in Affinity for DNA. *J. Phys. Chem. B* **120**, 7360-7370 (2016).
127. Banerjee, A. *et al.* Allosteric communication between DNA-binding and light-responsive domains of diatom class i aureochromes. *Nucleic Acids Res.* **44**, 5957–5970 (2016).
128. Christie, J. M., Gawthorne, J., Young, G., Fraser, N. J. & Roe, A. J. LOV to BLUF: Flavoprotein contributions to the optogenetic toolkit. *Mol. Plant* **5**, 533–544 (2012).
129. Strickland, D., Moffat, K. & Sosnick, T. R. Light-activated DNA binding in a designed allosteric protein. *PNAS.* **105**, 10709-10714 (2008).
130. Mart, R. J., Meah, D. & Allemann, R. K. Photocontrolled Exposure of Pro-apoptotic Peptide Sequences in LOV Proteins Modulates Bcl-2 Family Interactions. *ChemBioChem* **17**, 698-701 (2016).

131. Niopek, D., Wehler, P., Roensch, J., Eils, R. & Di Ventura, B. Optogenetic control of nuclear protein export. *Nat. Commun.* **7**, 1-9 (2016).
131. Renicke, C., Schuster, D., Usherenko, S., Essen, L. O. & Taxis, C. A LOV2 domain-based optogenetic tool to control protein degradation and cellular function. *Chem. Biol.* **20**, 619-626 (2013).
132. Lee, J., Natarajan, M., Nashine, V. C., Socolich, M., Vo, T., Russ, W. P., Benkovic, S. J. & Ranganathan, R. Surface Sites for Engineering. *Science* **322**, 438-442 (2008).
133. Lee, J., Natarajan, M., Nashine, V. C., Socolich, M., Vo, T., Russ, W. P., Benkovic, S. J. & Ranganathan, R. Surface Sites for Engineering. *Science* **322**, 438-442 (2008).
134. Kawano, F., Suzuki, H., Furuya, A. & Sato, M. Engineered pairs of distinct photoswitches for optogenetic control of cellular proteins. *Nat. Commun.* **6**, 1-8 (2015).
135. Pudasaini, A., El-Arab, K. K. & Zoltowski, B. D. LOV-based optogenetic devices: light-driven modules to impart photoregulated control of cellular signaling. *Front. Mol. Biosci.* **2**, 1-15 (2015).
136. Grosenick, L., Marshel, J. H. & Deisseroth, K. Closed-loop and activity-guided optogenetic control. *Neuron* **86**, 106-139 (2015).
137. Losi, A., Gardner, K. H. & Möglich, A. Blue-Light Receptors for Optogenetics. *Chem. Rev.* **118**, 10659-10709 (2018).
138. Zhao, E. M., Zhang, Y., Mehl, J., Park, H., Lalwani, M. A., Toettcher, J. E. & Avalos, J. L. Optogenetic regulation of engineered cellular metabolism for microbial chemical production. *Nature* **555**, 683-687 (2018).
139. Gehrig, S., Macpherson, J. A., Driscoll, P. C., Symon, A., Martin, S. R., MacRae, J. I., Kleinjung, J., Fraternali, F. & Anastasiou, D. An engineered photoswitchable mammalian pyruvate kinase. *FEBS Journal* **284**, 2955-2980 (2017).
140. Boehr, D. D., McElheny, D., Dyson, H. J. & Wrightt, P. E. The dynamic energy landscape of dihydrofolate reductase catalysis. *Science* **313**, 1638-1642 (2006).
141. Möglich, A., Ayers, R. A. & Moffat, K. Design and Signaling Mechanism of Light-Regulated Histidine Kinases. *J. Mol. Biol.* **385**, 1433-1444 (2009).
142. Taylor, B. L. & Zhulin, I. B. PAS domains: internal sensors of oxygen, redox potential, and light. *Microbiol. Mol. Biol. Rev.* **63**, 479-506 (1999).
143. Strickland, D., Moffat, K. & Sosnick, T. R. Light-activated DNA binding in a designed allosteric protein. *PNAS* **105**, 10709-10714 (2008).



144. Hurlburt, B. K. & Yanofsky, C. trp Repressor/trp operator interaction. Equilibrium and kinetic analysis of complex formation and stability. *J. Biol. Chem.* **267**, 16783-16789 (1992).
145. in, L., Yang, J. & Carey, J. Thermodynamics of Ligand Binding to trp Repressor. *Biochemistry* **32**, 7302-7309 (1993).
146. Otwinowski, Z. *et al.* Crystal structure of trp repressor/operator complex at atomic resolution. *Nature* **335**, 321-329 (1988).
147. Strickland, D., Yao, X., Gawlak, G., Rosen, M. K., Gardner, K. H., Sosnick, T. R. Rationally improving LOV domain-based photoswitches. *Nat. Methods* **7**, 623-626 (2010).
148. Chen, X. *et al.* An extraordinary stringent and sensitive light-switchable gene expression system for bacterial cells. *Cell Res.* **26**, 854-857 (2016).
149. Zhang, A. P. P., Pigli, Y. Z. & Rice, P. A. Structure of the LexA-DNA complex and implications for SOS box measurement. *Nature* **466**, 883-886 (2010).
150. Herrou, J. & Crosson, S. Function, structure and mechanism of bacterial photosensory LOV proteins. *Nat. Rev. Microbiol.* **9**, 713-723 (2011).
151. Jones, M. A., Feeney, K. A., Kelly, S. M. & Christie, J. M. Mutational analysis of phototropin 1 provides insights into the mechanism underlying LOV2 signal transmission. *J. Biol. Chem.* **282**, 6405-6414 (2007).
152. Nash, A. I., Ko, W. H., Harper, S. M. & Gardner, K. H. A conserved glutamine plays a central role in LOV domain signal transmission and Its duration. *Biochemistry* **47**, 13842-13849 (2008).
153. Ganguly, A., Thiel, W. & Crane, B. R. Glutamine Amide Flip Elicits Long Distance Allosteric Responses in the LOV Protein Vivid. *J. Am. Chem. Soc.* **139**, 2972-2980 (2017).
154. Yee, E. F. *et al.* Signal transduction in light-oxygen-voltage receptors lacking the adduct-forming cysteine residue. *Nat. Commun.* **6**, 10079 (2015).
155. Tsukuno, H., Ozeki, K., Kobayashi, I., Hisatomi, O. & Mino, H. Flavin-Radical Formation in the Light-Oxygen-Voltage-Sensing Domain of the Photozipper Blue-light Sensor Protein. *J. Phys. Chem. B* **122**, 8819-8823 (2018).
156. Purcell, E. B., McDonald, C. A., Palfey, B. A. & Crosson, S. An analysis of the solution structure and signaling mechanism of LovK, a sensor histidine kinase integrating light and redox signals. *Biochemistry* **49**, 6761–6770 (2010).
157. Mansurova, M., Scheercousse, P., Simon, J., Kluth, M. & Gärtner, W.

Chromophore Exchange in the Blue Light-Sensitive Photoreceptor YtvA from *Bacillus subtilis*. *ChemBioChem* **12**, 641-646 (2011).





## 2 General Materials and Methods

All chemicals and growth media components, otherwise specified, were purchased from Sigma-Aldrich, Fisher Scientific, Melford or Apollo Scientific and were used without further purification. All solutions including media components were prepared following manufacturer's guidelines.

### 2.1 Preparation of Liquid and Solid Culture Media

All liquid or solid media were sterilised by steam autoclaving, 121 °C for 15 min at 2.0 Bar pressure, in Astell autoclave following manufacturer's guidelines. Post sterilisation, media was cooled to an appropriate temperature (50 to 60 °C for solid media and room temperature for liquid media) prior addition of additives including filter (0.22 µm) sterilised antibiotics. All sterilised media and media components were stored at room temperature in closed sterile vessels. Additives in the following Table 2.1 were filter sterilised, appropriately aliquoted into sterile containers and stored at - 20 °C for future use. Additives marked with \* were not filter sterilised, contained ≥ 50 % v/v ethanol.

**Table 2.1: Commonly used additives in growth media including antibiotics.**

20 % w/v D-glucose	20 g per 100 mL
10 % w/v D-lactose	10 g per 100 mL
20 % w/v L-arabinose	20 g per 100 mL
×1,000 kanamycin sulphate (Kan)	50 mg/mL
×1,000 sodium ampicillin (Amp)	100 mg/mL
×1,000 chloramphenicol* (Cam)	25 mg/mL in 90 % v/v ethanol
×1,000 tetracycline* (Tet)	10 mg/mL in 70 % v/v ethanol
1 M * (IPTG)	120 mg/mL in 50 % v/v ethanol

#### 2.1.1 Lysogeny Broth (LB) Liquid Media Preparation

LB media is a complex media commonly used for bacterial culture growth and protein overexpression providing moderately high protein yields and culture densities. It is also commonly used to grow cultures for plasmid isolation.

**LB media was prepared by mixing per litre of diH<sub>2</sub>O:**

- 10 g tryptone 1 % w/v
- 5 g yeast extract 0.5 % w/v
- 10 g NaCl 1 % w/v
- pH adjusted to 7.0 to 7.4 with 5 M KOH

**2.1.2 Defined Minimal Media (M9) Preparation**

Defined minimal media (M9) is commonly used to label proteins and/or metabolites. It is a defined media, where all components are known and can be added with enriched isotope contents allowing desired labelling. For protein nuclear magnetic resonance (NMR) experiments, especially triple resonance experiments, proteins are commonly enriched with <sup>13</sup>C using U-<sup>13</sup>C-D-glucose, U-<sup>13</sup>C-glycerol or U-<sup>13</sup>C sodium acetate as a sole carbon source. For <sup>15</sup>N-enrichment <sup>15</sup>NH<sub>4</sub>Cl or (<sup>15</sup>NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> are commonly used as a sole nitrogen source. If required, perdeuteration of the protein or metabolites is achieved by using perdeuterated carbon or nitrogen sources while growing cultures in <sup>2</sup>H<sub>2</sub>O prepared media. To ensure stable growth, media components are usually added prior growth. Note: growth depends on natural biosynthetic pathways to supply required metabolites, hence slower growth and lower final optical densities are usually observed.

**M9 media was prepared by mixing per litre of diH<sub>2</sub>O (M9 salts):**

- 6 g Na<sub>2</sub>HPO<sub>4</sub> 42 mM
- 3 g KH<sub>2</sub>PO<sub>4</sub> 22 mM
- 0.5 g NaCl 8.5 mM
- pH adjusted to 7.4 to 7.8 with 5 M KOH

**To 1 L of M9 salts the following components were added:**

- 1 g/L NH<sub>4</sub>Cl, from 0.2 g/mL filter sterilised stock
- 10 mL ×100 trace element solution, autoclave sterilised
- 2.5 mL 1 M MgSO<sub>4</sub>, autoclave sterilised
- 0.1 mL 1 M CaCl<sub>2</sub>, autoclave sterilised
- 2 g/L D-glucose, from 20 % w/v filter sterilised stock
- 1 mL ×1,000 cofactor and cofactor precursor mix, filter sterilised
- Required antibiotic(s)
- 1:100 dilution of overnight preculture for large growths or a single colony for preculture

#### **×100 Trace Element Solution preparation:**

- 2.517 g Na<sub>2</sub>EDTA.2H<sub>2</sub>O                      6.75 mM
- 0.022 g MnSO<sub>4</sub>.7H<sub>2</sub>O                      0.10 mM
- 0.107 g FeCl<sub>2</sub>.4H<sub>2</sub>O                      0.54 mM
- 0.119 g CoCl<sub>2</sub>.6H<sub>2</sub>O                      0.50 mM
- 0.029 g ZnSO<sub>4</sub>.5H<sub>2</sub>O                      0.10 mM
- 0.005 g CuSO<sub>4</sub>.5H<sub>2</sub>O                      0.02 mM
- 0.350 g H<sub>3</sub>BO<sub>3</sub>                      5.66 mM
- 0.094 g Na<sub>2</sub>MoO<sub>4</sub>.2H<sub>2</sub>O                      0.40 mM
- 22.5 mL 0.1 M HCl
- diH<sub>2</sub>O to 1 L

0.107 g FeCl<sub>2</sub>.4H<sub>2</sub>O was dissolved in 22.5 mL 0.1 M HCl. The solution was topped with diH<sub>2</sub>O to 950 mL. To a yellow clear solution, Na<sub>2</sub>EDTA and the following listed components were added. After addition of all of the components, solution was topped to 1 L with diH<sub>2</sub>O and was autoclave sterilised. Solution was stored at room temperature.

#### **×1,000 cofactor and cofactor precursor mix preparation:**

- 1.1 mg/mL thiamine hydrochloride, thiamine diphosphate precursor (ThDP)
- 1.1 mg/mL niacin or nicotinamide, nicotinamide adenine dinucleotide precursor (NAD<sup>+</sup>/ NADP<sup>+</sup>)
- 1.1 mg/mL pyridoxal phosphate disodium salt, cofactor (PLP)
- 1.1 mg/mL D-biotin, cofactor
- 1.1 mg/mL *para*-amino benzoic acid, dihydrofolate precursor (DHF)
- 1.1 mg/mL calcium pantothenate, Coenzyme A precursor (CoA)

The solution was mixed, pH adjusted to 7.0 to 8.0 with 1 M NaOH, bright yellow solution was then filter sterilised and aliquots of 1 mL stored at -20 °C.

#### **1 M MgSO<sub>4</sub> preparation, autoclave sterilised:**

- 120.37 g MgSO<sub>4</sub>
- 1 L diH<sub>2</sub>O

#### **1 M CaCl<sub>2</sub> preparation, autoclave sterilised:**

- 27.75 g CaCl<sub>2</sub>
- 0.25 L diH<sub>2</sub>O

### 2.1.3 Preparation of Super Optimal Broth (SOB) and Super Optimal Broth with Catabolite Repression (SOC) Media

SOB media is commonly used to prepare high density precultures and overnight growths for plasmid isolation and recombinant protein production.

**SOB media was prepared by adding the following components per litre of diH<sub>2</sub>O:**

- 20 g tryptone 2 % w/v
- 5 g yeast extract 0.5 w/v
- 0.584 g NaCl 10 mM
- 0.186 g KCl 2.5 mM
- 2.4 g MgSO<sub>4</sub> 20 mM
- pH was adjusted to 7.4 with 5 M KOH

**SOC media was prepared:**

- To 1 L of sterilised SOB media 20 mL of 20 % w/v filter sterilised D-glucose was added.

### 2.1.4 Preparation of Terrific Broth (TB) Media

TB media is commonly used to express challenging proteins or achieve high protein yields per lower growth volumes. It is an exceptionally rich media with an additional carbon source, glycerol.

**TB media was prepared by dissolving the following components in 900 mL diH<sub>2</sub>O:**

- 12 g tryptone, 1.2 % w/v
- 24 g yeast extract 2.4 % w/v
- 5 g glycerol 0.5 % w/v
- 2.4 g MgSO<sub>4</sub> 20 mM
- pH adjusted to 7.4 with 5 M KOH

**and in 100 mL:**

- 2.31 g KH<sub>2</sub>PO<sub>4</sub> 0.17 M final concentration
- 12.54 g K<sub>2</sub>HPO<sub>4</sub> 0.72 M final concentration

Media and phosphate buffer were autoclave sterilised separately, and upon cooling aseptically mixed.



### 2.1.5 Preparation of Autoinduction Media – LB Based

Autoinduction media contains multiple components allowing high bacterial cell density to be achieved. It is commonly used to produce large quantities of protein per volume of culture. Unlike LB, SOB or TB expression medias that depend on induction of *lacUV5* operon located T7 RNA polymerase with isopropyl- $\beta$ -D-thiogalactoside (IPTG) at a specific growth point, autoinduction media contains multiple carbon sources. Carbon sources include D-glucose as a primary preferential carbon source, D-lactose or L-arabinose as a secondary carbon source and glycerol as a late stage aerobic carbon source<sup>154</sup>. In early growth, preferred D-glucose minimises basal expression of *lacUV5* and *araBAD* operons and their corresponding engineered plasmid systems through catabolite suppression by low intracellular levels of cyclic adenosine monophosphate (cAMP). Upon depletion of D-glucose, cAMP levels increase facilitating higher basal expression of genes involved in an alternative carbon source utilisation. Entrance of an alternative carbon source (D-lactose or L-arabinose) leads to activation of corresponding catabolite utilisation operons leading to gene of interest expression (for more information and protocols, please see sections 2.5 and 2.7). When the secondary carbon source is depleted, glycerol is used as an aerobic carbon source to maintain high density culture.

**LB based autoinduction media was prepared as described below per 1 L of media:**

- 10 g tryptone, 1 % w/v
- 5 g yeast extract 0.5 % w/v
- 2.4 g MgSO<sub>4</sub> 20 mM
- 5 g glycerol 0.5 % w/v
- Dissolved in 800 mL diH<sub>2</sub>O
- pH adjusted to 7.8 with 5 M KOH
- diH<sub>2</sub>O added to 900 mL

**Phosphate buffer was prepared:**

- 6 g Na<sub>2</sub>HPO<sub>4</sub> 42 mM
- 3 g KH<sub>2</sub>PO<sub>4</sub> 22 mM
- 0.5 g NaCl 8.5 mM
- 1 g (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> 0.1 % w/v
- 90 mL diH<sub>2</sub>O
- pH adjusted to 7.4 to 7.8 with 5 M KOH
- diH<sub>2</sub>O added to 100 mL

Media and phosphate buffers were autoclave sterilised separately and upon cooling mixed aseptically. To the media, the following carbon sources and additives were added:

- 5 mL 20 % w/v D-glucose 1 g/L
- 30 mL 10 % w/v D-lactose / L-arabinose 3 g/L

Prior inoculation with appropriate culture, antibiotics were added including 100  $\mu$ g/mL Amp or 100  $\mu$ g/mL Kan.

## 2.1.6 Preparation of Solid Luria Broth (LB) and Super Optimal Broth with Catabolite Repression (SOC) Agar Plates

**LB agar plates were prepared by adding per litre:**

- 10 g tryptone 1 % w/v
- 5 g yeast extract 0.5 % w/v
- 10 g NaCl 1 % w/v
- 800 mL diH<sub>2</sub>O
- pH adjusted to 7.0 to 7.4 with 5 M KOH
- 15 g agar granules 1.5 % w/v
- diH<sub>2</sub>O added to 1 L

Magnetic stirrer was added, suspension briefly mixed, autoclave sterilised and cooled to approximately 50 to 60 °C while gently stirring. To cooled media the required antibiotics were added, from  $\times 1,000$  stock, aseptically:

- 1 mL kanamycin
- 1 mL chloramphenicol
- 1 mL tetracycline
- 1.8 mL ampicillin (excess added due to thermal instability of the antibiotic)

Media was gently stirred to ensure uniform antibiotic distribution and media separated into sterile Pertri dishes aseptically, ~ 10 mL per plate. Plates were left to set at room temperature. Plates containing tetracycline were kept in the dark. All plates were stored at 4 °C for up to two weeks and prior use equilibrated to a room temperature for 1 h.

**SOC plates were prepared by adding per litre:**

- 20 g tryptone, 2 % w/v
- 5 g yeast extract, 0.5 w/v
- 0.584 g NaCl, 10 mM
- 0.186 g KCl, 2.5 mM
- 2.4 g MgSO<sub>4</sub>, 20 mM
- 800 mL diH<sub>2</sub>O
- pH was adjusted to 7.4 with 5 M KOH
- 15 g agar granules added
- diH<sub>2</sub>O added to 1 L

Magnetic stirrer was added, suspension briefly mixed, autoclave sterilised and cooled to 50 to 60 °C while gently stirring. 20 mL of 20 % w/v D-glucose and required antibiotics were added then plates were prepared and stored as described for LB agar plates, section 2.1.6 above.

## **2.2 Preparation of Competent Cells, Transformation and Recombinant Protein Expression**

BL21 (DE3) were purchased from Novagen. XL1-Blue cells were purchased from Agilent Technologies. TOP10 and DH5 $\alpha$  strains were purchased from Invitrogen. BL21-AI (Invitrogen) was a kind gift from Dr Louis Luk, Cardiff University, UK. All cell lines were prepared chemically competent inhouse for future use, please see sections 2.2.3 and 2.2.4 below. Chemically competent cells were stored at a stable - 80 °C. For purposes of unnatural amino acid incorporation, BL21-AI cells were modified by incorporation of an additional plasmid, see section 2.2.6.

### **2.2.1 *Escherichia coli* (*E. coli*) DNA Work Strains**

Cloning strains are summarised in the Table 2.2 below. These were routinely used for plasmid amplification, plasmid storage and cloning purposes. Cells were prepared chemically competent,  $1 \times 10^6$  to  $1 \times 10^9$  cfu/ $\mu$ g of plasmid DNA, and were prepared as described in 2.2.3 and 2.2.4. Most notably, as described by the manufacturer these strains are deficient in recombinase I (*recA1*) and endonuclease I (*endA1*) increasing stability and minimising degradation of plasmid constructs introduced into the cells.

**Table 2.2: *Escherichia coli* (*E. coli*) strains used for plasmid amplification.**

<i>E. coli</i> strain	Antibiotic Resistance
XL1-Blue	Tetracycline
DH5a	None
TOP10	None

### 2.2.2 *Escherichia coli* (*E. coli*) Protein Overexpression Strains

Strains used are listed in Table 2.3 below. Predominantly, BL21 (DE3) strain was routinely used to overproduce large quantities of protein of interest. BL21-AI cells, where AI stands for L-arabinose induction, were used to produce protein of interest containing unnatural amino acid, 3,5-difluorotyrosine. All strains used for overproduction of proteins in this work were null mutants of Lon protease (*lon*, ATP dependent serine protease) and outer membrane protease T (*ompT*, opmtin family) minimising degradation of overexpressed proteins. BL21-AI cells, derivatives of BL21, contain T7 RNAPol in *araBAD* locus, allowing induction of T7 RNAPol expression with L-arabinose rather than D-lactose or IPTG used for BL21 (DE3) cells. Considering that L-arabinose catabolism system is more stringent, basal expression of T7 RNAPol is minimal.

**Table 2.3: *Escherichia coli* (*E. coli*) strains used for overproduction of proteins.**

<i>E. coli</i> strain	Antibiotic resistance
BL21-AI	Tetracycline
BL21-AI pEVOL- <i>F<sub>n</sub></i> YR- <i>E3</i>	Tetracycline and Chloramphenicol*
BL21(DE3)	None

\*Indicates presence of an additional plasmid and antibiotic resistance carried as a selective marker.

### 2.2.3 Preparation of Competent Cells

Chemically competent cells were prepared following general CaCl<sub>2</sub> protocol. Two different CaCl<sub>2</sub> buffers were used, CCC buffer A and CCC buffer B, see below. All steps in preparation were carried out aseptically using sterile media, buffers and containers.

Competent cells were prepared by streaking out a glycerol stock or an agar stab of a cell line onto LB or SOC agar plates supplemented with an appropriate antibiotic(s) using flame sterilised metal loop. Plates were incubated for 16 h at 37 °C and a single

colony was used to inoculate 5 mL LB, SOB or SOC media supplemented with a required antibiotic(s). Preculture was grown for 16 h at 30 to 37 °C and 0.5 mL was used to inoculate 50 mL of corresponding media containing required antibiotic(s). Cells were grown to an optical density at 600 nm (OD<sub>600</sub>) of 0.3 to 0.4 (early exponential growth phase) or 0.8 to 1.1 (mid exponential growth phase) at 25 to 37 °C. Upon reaching required OD<sub>600</sub>, cultures were chilled on ice for 20 min, spun down at 4,000 RPM for 10 min at 4 °C, gently resuspended in 10 mL ice-cold CCC buffer A, incubated on ice for 10 minutes, spun down again, supernatant aseptically decanted and cell pellet gently resuspended in 2.5 mL (for OD<sub>600</sub> of 0.3 to 0.4 growth) or 5.0 mL (for OD<sub>600</sub> of 0.8 to 1.1 growth) ice-cold CCC buffer B. Cells were aliquoted (50 to 100 µL) into Eppendorf tubes chilled to 4 °C, flash frozen in liquid nitrogen and stored at - 80 °C.

**Calcium Competent Cells (CCC) buffer A:**

- 10 mM piperazine-*N,N*-bis(2ethanesulphonic acid) (PIPES)
- 100 mM CaCl<sub>2</sub>
- pH adjusted to 7.0 with 1 M KOH

**Calcium Competent Cells (CCC) buffer B:**

- 10 mM PIPES
- 100 mM CaCl<sub>2</sub>
- 15 % v/v glycerol
- pH adjusted to 7.0 with 1 M KOH

Buffers were filter sterilised and stored at 4 °C.

#### 2.2.4 Preparation of Supercompetent Cells

Supercompetent cells were prepared almost identically as competent cells, see section 2.3.3 above for more detail. Minor differences were that cells were grown to an OD<sub>600</sub> of 0.3 to 0.4 and TFB I and TFB II buffers were used instead. Composition of TFB I and TFB II buffers is indicated below. Buffer were filter sterilised and stored at 4 °C.

**TFB I:**

- 30 mM potassium acetate
- 100 mM RbCl
- 10 mM CaCl<sub>2</sub>
- 50 mM MnCl<sub>2</sub>
- pH adjusted to 5.8 with dilute acetic acid

**TFB II:**

- 10 mM piperazine-*N,N*-bis(2ethanesulphonic acid) (PIPES)
- 75 mM CaCl<sub>2</sub>
- 10 mM RbCl
- 15 % v/v glycerol
- pH adjusted to 6.8 with 1 M KOH

**2.2.5 Transformation, Preparation of Cell Culture Inoculate and Glycerol Stocks**

Chemically competent cells were transformed by thawing cells on ice, 0.1 to 1.0  $\mu$ L of vector (plasmid, 50 to 100 ng/ $\mu$ L) was added, cells gently mixed, incubated on ice for 10 to 30 minutes, heat shocked at 42 °C for 45 s and placed on ice for 2 min. To transformed cells, 1 mL of room temperature SOC media was added and cells left to recover shaking at 37 °C. 100  $\mu$ L of the culture (or centrifuge concentrated cell pellet, 6,000 RPM for 2 to 10 min) was plated onto an appropriate antibiotic LB or SOC agar plates and were incubated at 37 °C for 16 h or at room temperature for 48 h.

Overnight precultures were prepared by placing a single colony into 5 to 100 mL media (M9, LB, TB, SOB or SOC) supplemented with an appropriate antibiotic(s). Culture was left to grow for 16 h at 25 to 37 °C shaking.

Glycerol stocks were prepared by placing 0.8 mL of fresh overnight culture into a sterile Eppendorf tube, adding 0.2 mL sterile 50 % v/v glycerol, gently mixing and storing at - 80 °C.

To prepare an overnight culture from a glycerol stock, a stab using a sterile pipette tip or a whole glycerol stock was placed into a media supplemented with an appropriate antibiotic(s). The culture was left to grow overnight at 25 to 37 °C for 16 h shaking vigorously.

### 2.2.6 Preparation of BL21-AI pEVOL-*F<sub>n</sub>YR-E3* Cells

BL21-AI pEVOL-*F<sub>n</sub>YR-E3* were prepared by transforming chemically competent BL21-AI cells with pEVOL-*F<sub>n</sub>YR-E3* plasmid, see section 2.2.5 for detailed transformation protocol. Transformed cells were plated onto SOC agar plates supplemented with 25  $\mu\text{g/mL}$  Cam and 12.5  $\mu\text{g/mL}$  Tet. Plates were incubated for 16 h at 37 °C, a single colony was picked and used to inoculate 50 mL SOC media (25  $\mu\text{g/mL}$  Cam, 12.5  $\mu\text{g/mL}$  Tet) incubated overnight at 25 °C shaking 250 RPM. The following day, 0.5 mL of preculture was used to inoculate 50 mL SOC media supplemented with 25  $\mu\text{g/mL}$  Cam, cells grown to OD<sub>600</sub> of 0.3 to 0.4 and competent cells prepared as described in section 2.2.3. BL21-AI pEVOL-*F<sub>n</sub>YR-E3* cells were then used to incorporate unnatural amino acids into proteins as described in section 2.1.8. Please note that in order to minimise toxicity of pEVOL-*F<sub>n</sub>YR-E3* vector, all overnight precultures were grown in SOC media (catabolite suppression) at lower temperatures, 25 °C.

### 2.2.7 Transformation of Cloning Constructs, Preparation of Overnight Growths and Sequencing of Constructs

Vector constructs after Golden Gate assembly (New England Biolabs, USA), Gibson assembly (New England Biolabs, USA), Site Directed Mutagenesis (SDM) or blunt end ligation were transformed by placing 2.5 to 10  $\mu\text{L}$  of the reaction mixture (maximum 10 % of the reaction mixture per volume of supercompetent cells) into on ice thawed supercompetent XL1-Blue, DH5 $\alpha$  or TOP10 cells. Cells were gently mixed, incubated on ice for 30 min, heat shocked at 42 °C for 45 s, placed on ice for 2 min and 1 mL of room temperature SOC was added. Cells were incubated at 37 °C shaking vigorously, concentrated by centrifugation (6,000 RPM for 2 to 10 min), pellet resuspended in 100  $\mu\text{L}$  SOC media (residual) and cells plated onto LB or SOC agar plates containing required antibiotic(s). Plates were incubated at 37 °C for 16 h, a single colony was picked and used to inoculate 5 mL of LB, SOB or SOC media (in sterile 20 to 50 mL centrifuge tubes) containing required antibiotic(s). Cultures were grown overnight at 37 °C shaking vigorously, concentrated by centrifugation and plasmid isolated using Qiagen MiniPrep kit (Qiagen, USA). Isolated plasmid, 15  $\mu\text{L}$  of 50 to 100 ng/ $\mu\text{L}$ , was validated by sequencing, Eurofins Genomics, prior future use. Validated plasmids were transformed as described in section 2.2.5.

### 2.2.8 Expression of Recombinant Proteins in M9, LB, TB and Autoinduction Medias, Cell Culture Harvesting and Cell Paste Storage

In general, large growth cultures were prepared by placing small volume (1:100 or 1:50 dilution) of an appropriate overnight preculture into room temperature media containing required antibiotic(s) and additives aseptically. For more detailed protocols for individual proteins, please see corresponding chapter materials and methods.

LB cultures were grown at 37 °C shaking 250 RPM to an OD<sub>600</sub> of 0.6 to 0.8, induced with 0.5 to 1.0 mM IPTG and depending on protein overproduction grown for 16 h at 16 to 25 °C shaking 250 RPM.

TB cultures were grown at 37 °C shaking 250 RPM to an OD<sub>600</sub> of 1.5 to 2.0, induced with 1.0 mM IPTG and cultures grown overnight at 16 °C shaking 300 RPM.

Autoinduction cultures were grown at 37 °C for 4 to 6 h shaking 300 RPM, temperature then decreases to 16 °C and cultures grown for an additional 24 to 48 h. Lower temperatures facilitates higher  $pO_2$  in media supporting higher densities in TB and autoinduction medias.

M9 cultures were grown at 37 °C shaking 250 RMP to an OD<sub>600</sub> of 0.6 to 1.0, induced with 0.5 to 1.0 mM IPTG and cultures grown overnight at 16 to 25 °C shaking. For  $^2H_2O$  growths, cultures were usually induced at OD<sub>600</sub> of 1.0 and allowed to grow at a decreased temperature for 24 to 36 h.

Cultures were harvested by centrifugation (6,000 RPM for 10 min at 4 °C), washed once with ~ 50 mL pellet storage buffer (50 mM Tris.HCl, 150 mM NaCl, pH 8.0) and cell paste stored at - 20 °C for future use. This wash was important to ensure correct cell lysate pH in sequential protein isolation steps. This step was crucial for autoinduction cultures where carbohydrate metabolism can result in overproduction of acetic acid and lower pH.

### 2.3 Preparation, Maintenance and Regeneration of Gravity and Fast Protein Liquid Chromatography (FPLC) Columns

Columns were packed and regenerated following manufacturer's guidelines. Resins and columns were purchased from GE Healthcare. Columns were maintained and stored following recommended guidelines. Gravity drip columns were ran manually. Packed or prepacked column were run on Fast Protein Liquid Chromatography system, NGC 10 Medium Pressure Liquid Chromatography (MPLC) system, Bio-Rad Laboratories, equipped with multiple wavelength detector. Wavelengths commonly used include 260 (DNA/RNA), 280 (protein aromatic residues), 390 (cofactors including FMN) and 450 nm



(oxidised FMN maximum visible spectrum absorbance), respectively. FPLC System was operated following manufacturer's guidelines and column specifications.

### 2.3.1 Immobilised Metal Affinity Chromatography (IMAC) Gravity Drip Column and Prepacked Column Preparation and Regeneration

Immobilised metal affinity chromatography (IMAC) gravity drip column was prepared by loading 3-5 mL of Ni<sup>2+</sup>-HiLoad (GE Healthcare) resin into 30 mL empty gravity drip column. Beads were left to settle and the column packed under gravity using 20 % v/v ethanol as a mobile phase. For column preservation 20 % v/v ethanol storage solution was used. Before protein purification column was washed with 5 to 10 column volumes (CV) of diH<sub>2</sub>O and equilibrated with 5 CV of required buffer. For protein purification, crude clarified cell lysate was filtered through 0.22 or 0.45 µm filters, loaded onto the column and step gradient of imidazole in a buffer of interest (20 to 500 mM of imidazole diluted accordingly from 1 M stock in a required buffer with pH adjusted) was usually applied to elute the protein of interest containing oligo histidine tag (His<sub>6</sub>-tag). The column was regenerated by applying 2 to 5 CV 0.1 M NaOH, 5 CV diH<sub>2</sub>O, 1 CV 0.1 M sodium ethylenediaminetetraacetate (EDTA) pH 8.0, 5 CV diH<sub>2</sub>O, 1 CV 0.1 M NiCl<sub>2</sub> and 5 CV diH<sub>2</sub>O. Column was washed with and stored in 20 % v/v ethanol at 4 °C.

Prepacked HiLoad 1 mL columns were used either manually using a syringe or column was connected to FPLC system (NGC MPLC System, Bio-Rad). Protein samples were loaded onto the column by passing the sample through 0.22 µm filter. Purification and column regeneration was performed as described for gravity drip column.

### 2.3.2 Resource Q (ResQ) Column Preparation and Regeneration

Resource Q (ResQ, GE Healthcare) anion exchange column, 8 mL, was packed following manufacturer's guidelines. Briefly, 8 mL of resin stored in 20 % v/v ethanol was loaded into empty Tricorn 10/100 column. The beads were left to settle and additional resin added. Column was compacted by flowing 1 M NaCl solution at 5 mL/min rate. Column was regenerated by washes with 3 CV 1 M NaCl, 6 CV diH<sub>2</sub>O, 3 CV 1 M NaOH, 6 CV diH<sub>2</sub>O, 3 CV 1 M HCl, 6 CV diH<sub>2</sub>O and stored in 1 M NaCl solution. This regeneration protocol was commonly used to completely regenerate the column after 5 to 10 column uses or if poor performance was observed.

Proteins were purified following anion exchange chromatography principles. Column was usually equilibrated with 3 to 5 CV low salt buffer (20 to 50 mM NaCl), net negatively charged protein adsorbed to the positively charged resin, unbound proteins removed by additional 3 CV washes with low salt buffer and proteins eluted with a linear

gradient of high salt buffer from 0 to 100 % (1 M NaCl) up to 50 CVs. Protein samples were collected as 2 to 4 mL fractions. After run, column was briefly regenerated by washing with 3 CV diH<sub>2</sub>O or low salt buffer, 2 CV 1 M NaOH, 5 CV diH<sub>2</sub>O or low salt buffer and column stored in high salt buffer (1 M NaCl) for short term storage. For long term storage, column was equilibrated into 20 % v/v ethanol solution.

### 2.3.3 Gel Filtration Chromatography Purification

Prepacked 10/300 Superdex 75 or 200 analytical gel filtration columns (GE Healthcare) were used to characterise protein samples to estimate molecular size and oligomerisation states(s). Columns were run at a maximum flow rate of 0.5 mL/min following pressure and flow guidelines provided by the manufacturer. Columns were stored in 20 % v/v ethanol, where prior experiments columns were equilibrated with 2 CV of required buffer. Columns were regenerated either by 2 CV wash with 0.5 M NaOH or 6 M guanidine hydrochloride (usually buffered with 50 mM Tris.HCl with pH adjusted to 8.0), followed by 2 CV wash with diH<sub>2</sub>O and or 2 CV of buffer of interest.

Proteins of interest and corresponding protein-DNA complexes were characterised by analytical gel filtration chromatography (10/300 Superdex 75 or 200) at a concentration ranges from 10 to 2,000  $\mu$ M injecting 100  $\mu$ L of the sample. Prior loading samples onto the column, samples were briefly spun down at maximum speed (14,000 RPM for 10 min) to remove aggregates. For dark state experiments, samples were, if required, kept in the dark for up to 16 h to ensure complete recovery back to the dark-adapted state. All handling was under dim red lights and containers plus the column wrapped with 5 layers of aluminium foil. For lit state experiments, proteins samples were photoactivated for up to 5 to 60 min to ensure complete photoactivation as determined by UV-Vis characterisation. For DNA binding experiments, protein samples were photoactivated for 10 to 30 minutes on ice to allow DNA binding. To ensure protein photoactivation during the experiment, column was illuminated with blue LED lights (450  $\pm$  10 nm  $\times$  48 0.08 Watt power individual 12 V LEDs). All experiments were run on NGC MPLC System (Bio-Rad) equipped with multiwavelength detector using 260 (DNA/RNA maxima), 280 (protein aromatic residue maxima), 390 (flavin photoadduct maxima) and 450 nm (oxidised flavin maxima) wavelengths.

Prepacked HiPrep 10/600 Sephacryl S-300 column was used to purify some proteins and protein-DNA complexes. Column was stored in 20 % v/v ethanol and prior purification equilibrated with 2 CV buffer of interest. Protein samples were purified following manufacturer's guidelines (flow rate and maximum pressure). Fractions collected

were investigated by SDS-PAGE and proteins samples  $\geq 95\%$  pure were pooled and used for further experiments. Column was regenerated, if required, identically to the analytical gel filtration columns.

## **2.4 Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE)**

### **Running Buffer, Sample Loading Buffer and Polyacrylamide Gel Preparation**

#### **2.4.1 Preparation of Stock Solutions and Reducing SDS-PAGE Denaturing Loading Dye**

10 % w/v sodium dodecyl sulfate preparation:

- 10 g sodium dodecyl sulfate (SDS)
- 100 mL diH<sub>2</sub>O

Suspension was heated to 42 °C in a water bath, gently stirred to dissolve SDS and stored at room temperature. If precipitation was observed, the solution was reheated prior use.

10 % w/v ammonium persulfate (APS) preparation:

- 0.1 mg ammonium persulfate
- 1 mL diH<sub>2</sub>O

Solution was usually prepared fresh and for short term stored at 4 °C.

×10 SDS-PAGE running buffer:

- |                             |         |
|-----------------------------|---------|
| - 30.3 g Tris base,         | 0.25 M  |
| - 144 g glycine,            | 1.92 M  |
| - 10 g SDS,                 | 1 % w/v |
| - diH <sub>2</sub> O to 1 L |         |

Suspension was stirred until fully dissolved and stored at room temperature.

×1 SDS-PAGE running buffer:

- 100 mL ×10 SDS-PAGE running buffer
- 900 mL diH<sub>2</sub>O

The solution was freshly prepared before running each SDS-PAGE experiment.

#### **2.4.2 Preparation of SDS-PAGE Loading Dye**

×4 Reducing Protein Sample (denaturing) loading buffer:

- |                                     |             |
|-------------------------------------|-------------|
| - 2.5 mL 1 M Tris.HCl pH 6.8        | 0.25 M      |
| - 1 g SDS                           | 10 % w/v    |
| - 0.5 mL 0.3 % w/v bromophenol blue | 0.015 % w/v |

- 4 mL glycerol 40 % v/v
- 0.5 mL  $\beta$ -mercapto ethanol (BME) 0.2 M
- diH<sub>2</sub>O to 10 mL

×2 Reducing Protein Sample (denaturing) loading buffer:

- 2.5 mL 1 M Tris.HCl pH 6.8 0.125 M
- 1 g SDS 5 % w/v
- 0.5 mL 0.3 % w/v bromophenol blue 0.015 % w/v
- 4 mL glycerol 20 % v/v
- 0.5 mL  $\beta$ -mercapto ethanol (BME) 0.1 M
- diH<sub>2</sub>O to 20 mL

Freshly prepared loading buffers were aliquoted as 0.5 mL stocks and stored at -20 °C for future use.

#### 2.4.3 Preparation of Protein Samples for SDS-PAGE Analysis

Protein samples for SDS-PAGE were prepared by mixing the protein sample with reducing protein sample loading buffer appropriately. Samples were heated at 90 °C for 1 to 5 min to denature proteins, samples cooled to room temperature and loaded into individual SDS-PAGE plate wells.

#### 2.4.4 Preparation of Acrylamide SDS-PAGE Stacking and Resolving Gels

5 % stacking gel for 2 gel plates:

- 2.85 mL diH<sub>2</sub>O
- 1.25 mL 0.5 M Tris.HCl pH 6.8
- 0.85 mL 30 % acrylamide, 1 % bis-acrylamide
- 50  $\mu$ L 10 % w/v SDS

12 % resolving gel for 2 gel plates:

- 3.4 mL diH<sub>2</sub>O
- 2.5 mL 1.5 M Tris.HCl pH 8.8
- 4.0 mL 30 % acrylamide, 1 % bis-acrylamide
- 100  $\mu$ L 10 % w/v SDS

Solution were gently mixed and polymerisation was initiated by addition of:

- 100  $\mu$ L 10 % w/v ammonium persulphate (APS)
- 20  $\mu$ L *N,N,N',N'*- tetramethylenediamine (TEMED)

Solution was gently mixed again and gels cast between two glass plates suitable for Bio-Rad SDS-PAGE system. To prevent from drying, a thin layer of isopropanol was applied. When polymerised, isopropanol was removed, stacking gel was placed, well comb inserted and the stacking gel left to polymerise.

Gels were run in ×1 SDS-PAGE running buffer at a constant voltage of 150 V for 1 h. Gels were stained with Coomassie Blue – G250 stain, see section 2.5.6. Gels were imaged with ChemiDoc system, Bio-Rad, using manufacturer's setting for protein staining technique.

#### 2.4.5 Tricine SDS-PAGE Preparation

Tricine gels were performed to allow better resolution and band intensity for much smaller proteins (2 to 20 kDa in size) including isolated LOV domains and truncated LOV domain constructs. Several stock solutions were prepared:

##### ×10 Anode (Outside) buffer:

- 242.28 g Tris base 2 M
- 800 mL diH<sub>2</sub>O
- pH to 8.8 with concentrated HCl
- diH<sub>2</sub>O to 1 L

##### ×10 Cathode (Inner) buffer:

- 121.14 g Tris base 1 M
- 179.17 g Tricine 1 M
- 10 g SDS 1 % w/v
- diH<sub>2</sub>O to 1 L

##### Tricine gel casting (TGC) buffer was prepared:

- 90.86 g Tris base 3 M
- 200 mL diH<sub>2</sub>O
- pH to 8.45 with concentrated HCl
- 5 g SDS 2 % w/v
- diH<sub>2</sub>O to 250 mL

Tricine gels were prepared as two different components, stacking (4.5 %) and resolving (12 %) gels.

**Stacking (4.5 %) gel for 2 plates:**

- 1.0 mL TGC buffer
- 2.4 mL diH<sub>2</sub>O
- 0.6 mL 30 % acrylamide, 1 % bis-acrylamide

**Resolving (12 %) gel for 2 plates:**

- 3.3 mL TGC buffer
- 2.7 mL diH<sub>2</sub>O
- 4.0 mL 30 % acrylamide, 1 % bis-acrylamide

Solutions were gently mixed and polymerisation was initiated:

- 100  $\mu$ L 10 % w/v APS
- 20  $\mu$ L *N,N,N',N'*-tetramethylethylenediamine (TEMED)

Solutions were gently mixed and gels cast between two glass plates. Resolving gel was overlaid with a layer of isopropanol preventing drying. When resolving gel polymerised, isopropanol was removed, stacking gel was placed, 10 or 15 well comb inserted and gel left to polymerise.

Gels were run in a corresponding  $\times 1$  Cathode and Anode running buffers at a constant voltage of 120 V for 20 minutes and 160 V for 40 min. Gels were stained with Coomassie Blue – G250 stain, see section 2.5.6 for detailed protocol.

#### 2.4.6 Preparation of Coomassie Blue Stain – G250 and Staining of Protein SDS-PAGE Gels

Coomassie blue stain was prepared following general protocol. Briefly:

- 50 to 60 mg of Coomassie Brilliant Blue – G250 dissolved in minimal volume of absolute ethanol (10 to 20 mL)
- 450 mL diH<sub>2</sub>O and 3 mL concentrated HCl added
- Solution was stirred for 1 to 16 h
- diH<sub>2</sub>O added to 500 mL

This stain solution was stored in the dark and was used to stain protein gels following general protocol:

- Removed gels were boiled and washed  $\times 3$  with H<sub>2</sub>O
- Coomassie blue stain was applied (20 to 40 mL) and briefly heated

- Gel was left to stain for 2 to 20 min
- Stain was removed and background stain removed by boiling several times with H<sub>2</sub>O

Gels were imaged by with ChemiDoc system, Bio-Rad, using settings for Coomassie Blue Stain. Images were processed by the software provided by the manufacturer.

## 2.5 Chapter 3 Materials and Methods

Please note that Materials and Methods have been published in a peer reviewed journal Biochemistry under the name of 'Non-canonical Chromophore Reveals Structural Rearrangements of Light-Oxygen-Voltage Domain on Photoactivation'

Authors: Mindaugas E. Kalvaitis, Luke A. Johnson, Robert J. Mart, Pierre Rizkallah and Rudolf K. Allemann

### 2.5.1 *OdAu1a* construct generation

The pETM11 expression plasmid of full-length *Ochromonas danica* Aureochrome1a, codon-optimised for expression in *Escherichia coli* (*E. coli*), was kindly provided by Dr Harald Janovjak (IST Austria). Two further pETM11 plasmids encoding truncated forms of *OdAu1a* (*OdAu1a*<sub>LOV</sub> and *OdAu1a*<sub>bZIPLOV</sub>) were constructed by Dr Luke Johnson, Cardiff University, UK, using primers described for each construct. Briefly, the desired segments of the *OdAu1a* gene and pETM11 plasmid backbone were amplified by PCR using primers (*OdAu1a*<sub>LOV</sub> forward GTTACGCGGTCTCCCATGGATTACAGCCTGGTGAAGGCCCTG and reverse GCATGGTGGTCTCCGGCCGCTTACTTCTCGTTCTCC, *OdAu1a*<sub>bZIPLOV</sub> forward GTTACGCGGTCTCCCATGCGCACCCCTGCCACGCAAC and reverse GCAACTGGGTCTCCCATGGCGCCTGAAAATAAAGATTCTC), which appended complementary *Bsal* restriction sites to 3' and 5' ends. PCR products were gel purified using Qiagen Kit and ligated using Golden Gate assembly methods (*Bsal* and T4 Ligase from New England BioLabs, USA) and confirmed by sequencing from the T7 promoter (Eurofins Scientific, UK). Protein sequences for each construct are detailed in protein expression and purification section of this chapter.

### 2.5.2 Protein expression and purification

*OdAu1a* constructs were expressed using BL21 (DE3) *E. coli* (Merck) transformed with the appropriate plasmid and grown in LB media supplemented with glucose (1% w/v) and kanamycin (50  $\mu$ g/mL for rich media, 100  $\mu$ g/mL for minimal and autoinduction media). Auto-induction media was used for expression of natural abundance *OdAu1a*<sub>bZIPLOV</sub> and *OdAu1a*<sub>LOV</sub>. Cultures were grown in the dark for 4 h at 37 °C followed by 25 °C for 20 h. M9 minimal media containing glucose (2 g/L) and <sup>15</sup>N-ammonium chloride (1 g/L, Sigma Aldrich) was used for uniform <sup>15</sup>N labelling. M9 cultures were grown in the dark at 37 °C to an OD<sub>600</sub> of 0.8, then induced with isopentenyl thiogalactose (0.5 mM, IPTG, Melford) and grown for a further 16 h at 25 °C. Cells were harvested by centrifugation, washed with ice-cold buffer (50 mM Tris, 150 mM NaCl, pH 8.0), centrifuged again and the resulting cell pellets stored at -20 °C.

*OdAu1a*<sub>LOV</sub> was purified by suspending cell pellets in buffer A (20 mM HEPES, 20 mM NaCl, 0.3 mM TCEP, pH 7.8) supplemented with egg hen white lysozyme (1 mg/mL, Sigma Aldrich) and phenylmethylsulfonyl fluoride (1 mM, PMSF, Melford). Cells were lysed by sonication (5 minutes, pulsed 5 seconds on, 10 seconds off for a total time of 15 minutes), clarified by centrifugation at 38,000  $\times$  g for 30 min at 4 °C, and loaded onto 5 mL Ni<sup>2+</sup>-NTA column (GE Healthcare). A stepwise gradient of buffer with increasing concentrations of imidazole (0, 20, 40, 60 mM imidazole) were applied over the resin, eluting 4 column volumes (CV) of each fraction to wash (0-40 mM) *OdAu1a*<sub>LOV</sub> bound to the resin. One column volume of buffer containing flavin mononucleotide (0.1 to 0.5 mM, depending on the volume of clarified lysate applied to the column, FMN, Melford) was then applied to the column, and allowed to equilibrate for 30 min to saturate the protein with FMN. The resin was then washed with 5 CV of buffer to remove excess FMN followed by 4 CV of buffers containing imidazole (100, 200 mM) to elute *OdAu1a*<sub>LOV</sub>. Imidazole was removed by dialysis against further buffer (20 mM HEPES, 20 mM NaCl, 0.3 mM TCEP, pH 7.8) in the dark at 4 °C, tobacco etch virus protease (TEV-His<sub>6</sub> protease) with immobilised metal affinity chromatography (IMAC) was added to cleave the affinity tag away from *OdAu1a*<sub>LOV</sub> for 16 h. The sample was then loaded onto Resource Q (ResQ) anion exchange column (GE Healthcare) equilibrated with buffer A and purified further, eluting with a linear gradient of sodium chloride (1 M) over 6 column volumes. Purified *OdAu1a*<sub>LOV</sub> was passed through a 1 mL HiTrap Ni<sup>2+</sup> column (GE Healthcare) and was exchanged into either DNA binding buffer (20 mM HEPES, 100 mM NaCl, 10 mM MgCl<sub>2</sub>, 1.0 mM TCEP, pH 7.4), crystallisation buffer (50 mM MES, 100 mM NaCl, 20 mM NaOAc, 20 mM MgCl<sub>2</sub>, 5 mM EDTA, 5 mM dithiothreitol DTT pH 6.0) or NMR buffer (20 mM MES,



1 mM EDTA, 0.5 mM TCEP, 0.05 % NaN<sub>3</sub>, pH 6.0) as appropriate. Protein concentration was determined by UV absorbance using an FMN extinction coefficient of 12,600 M<sup>-1</sup>cm<sup>-1</sup> and samples were stored at a stable -80 °C until used.

*OdAu1a*<sub>bZIPLOV</sub> was purified by suspending cell pellets in a buffer containing a slightly higher salt concentration than used for *OdAu1a*<sub>LOV</sub> (20 mM HEPES, 150 mM NaCl, 0.3 mM TCEP, pH 7.8) supplemented with egg hen white lysozyme (1 mg/mL), PMSF (5 mM), benzoase DNase (1 U/mL, Sigma Aldrich), and magnesium chloride (5 mM). Ni<sup>2+</sup>-NTA purification was performed as outlined for *OdAu1a*<sub>LOV</sub>, but the resin was washed with lower concentrations of imidazole (0-25 mM) and eluted with EDTA (50 mM EDTA) due to *OdAu1a*<sub>bZIPLOV</sub> instability at elevated imidazole concentrations. Digestion with TEV-His<sub>6</sub> was performed overnight at 4 °C in buffer (20 mM HEPES, 150 mM NaCl, 0.3 mM TCEP, pH 7.8) and the protease and His<sub>6</sub> tag were removed by passing the solution through a 1 mL HiTrap Ni<sup>2+</sup> (GE Healthcare) column equilibrated with buffer. *OdAu1a*<sub>bZIPLOV</sub> was purified using a HiPrep™ 16/60 Sephacryl S-300 HR column equilibrated with DNA binding buffer (20 mM HEPES, 100 mM NaCl, 10 mM MgCl<sub>2</sub>, 1.0 mM TCEP, pH 7.4). Purified *OdAu1a*<sub>bZIPLOV</sub> was concentrated to 500 μM and stored at -80 °C. Samples for electrophoretic mobility shift assay experiments (EMSA) were supplemented with 10 % v/v glycerol and stored frozen until used.

**Wild type *OdAu1a* protein sequence, red IMAC tag and green TEV cleavage site:**

Please note, underlined region indicates bZIP-LOV region of the protein with LOV domain sequence highlighted in blue for clarity.

MKHHHHHHHPMSDYDIPTT**ENLYFOG**AMGTSKQQLPPPIFGVLGDEKQVARNGIISLVDI  
FDDFLFSGDRNQPSNTASSSSHAQESVSGKDEENDYDSNDDEGDSDDGKRRKRSRTLPR  
NMTEEQKIERRERNREHAKRSRVRKKFLLESLOHSHVRALEEEENEKLRNAIRENLQGEAEQ  
LLTRCSCGGPSVIA SDPNTATRTLDDPDYSLVKALQTAQQNFVISDPSPDNPIVYASQGFL  
TLTG YALSEVLGRNCRFLQGPETDPKAVEKVRKGLERGEDTTVLLNYRKDGSTFWNQ  
LFIAALRDGEGNVVNYLGVQCKVSEDYAKAFLKNEENEK

Number of amino acids: 315

Theoretical molecular weight from amino acid sequence post TEV cleavage: 35273.02 Da

**OdAu1a<sub>bZIPLOV</sub> construct protein sequence, red IMAC tag and green TEV cleavage site:**

Please note, LOV domain sequence is underlined and highlighted in blue for clarity.

MKHHHHHHHPMSDYDIPTT**ENLYFQG**AMRTLPRNMTEEQKIERRERNREHAKRSRVRKK  
FLLESLQHSVRALLEENEKLRNAIRENLQGEAEQLLTRCSCGGPSVIASDPNTATRTLDDP  
DYSLVKALQTAQQNFVISDPSIPDNPIVYASQGFLTTLTG YALSEVLGRNCRFLQGPETDPK  
AVEKVRKGLERGEDTTVLLNRYKDGSTFWNQLFIAALRDGEGNVVNYLGVQCKVSED  
YAKAFLKNEENEK

Number of amino acids post TEV cleavage: 227

Theoretical molecular weight from amino acid sequence post TEV cleavage: 25688.91 Da

**OdAu1a<sub>LOV</sub> construct protein sequence, red IMAC tag and green TEV cleavage site:**

MKHHHHHHHPMSDYDIPTT**ENLYFQG**AMDYSLVKALQTAQQNFVISDPSIPDNPIVYASQ  
GFLTTLTG YALSEVLGRNCRFLQGPETDPKAVEKVRKGLERGEDTTVLLNRYKDGSTFW  
NQLFIAALRDGEGNVVNYLGVQCKVSEDYAKAFLKNEENEK

Number of amino acids post TEV cleavage: 135

Theoretical molecular weight from amino acid sequence post TEV cleavage: 14,997.9 Da.

In protein sequence depicted in green bold and underlined regions indicates tobacco etch virus (TEV) protease cleavage site. TEV cleaves after residue Q resulting with a construct that contains G residue. All molecular weights are indicated below the sequences.

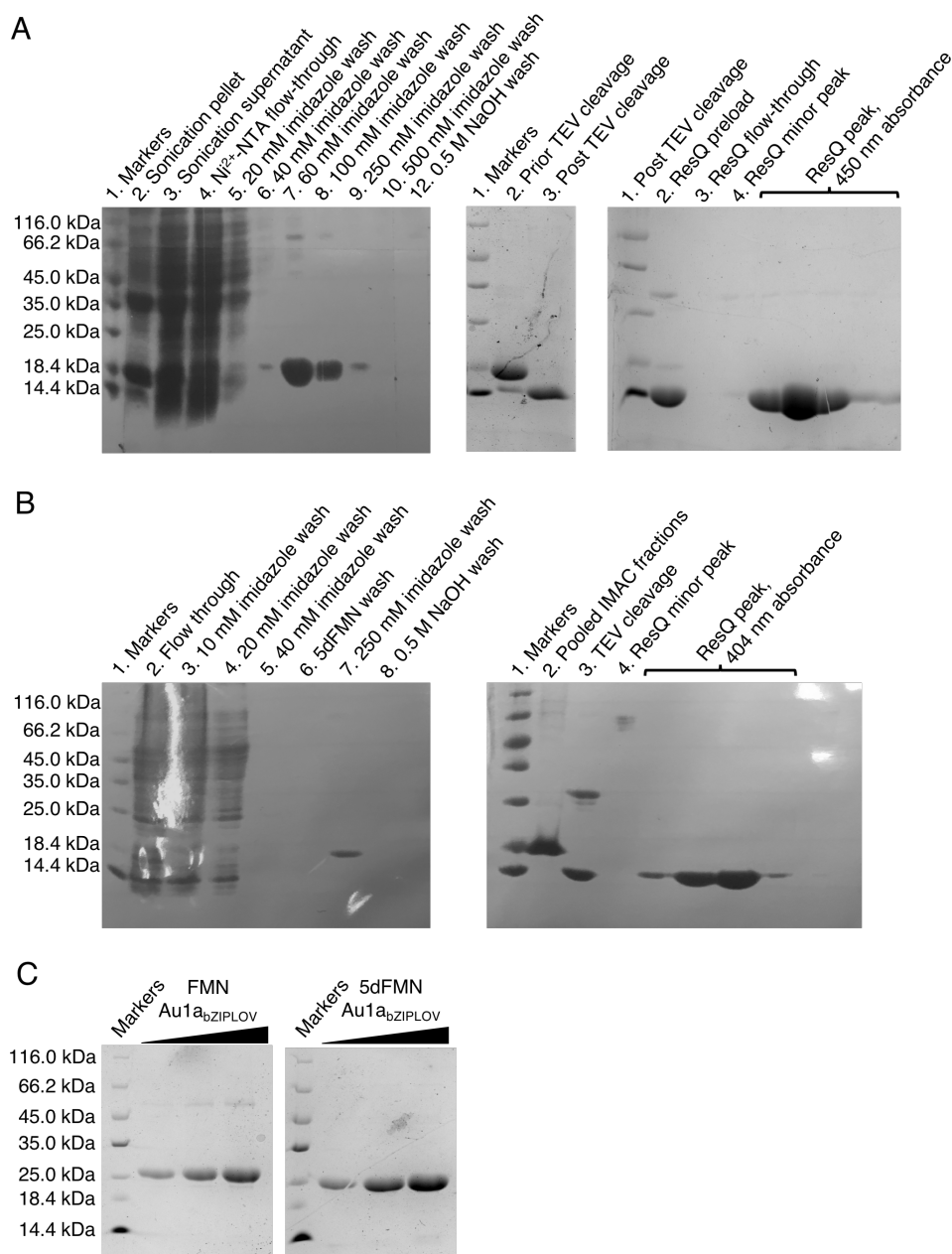
### 2.5.3 FMN to 5dFMN chromophore exchange

Chromophore exchange was achieved by on-column refolding which was challenging at the beginning. Initial attempts to use weaker chaotrope such as 8 M urea or 6 M guanidine hydrochloride were unsuccessful to remove native FMN indicating that protein was extremely stable. Further optimisation achieved conditions described in this work. Buffers were as used previously for *OdAu1a<sub>LOV</sub>* or *OdAu1a<sub>bZIPLOV</sub>* (see above). Cell lysates were loaded onto a 5 mL Ni<sup>2+</sup>-NTA column and washed with 3 CV of buffer, followed by washing with 2 CV containing imidazole (20 mM). The resin was then washed with 2 CV of buffer

containing guanidine hydrochloride (6 M, Sigma Aldrich) followed by 10-20 column volumes of buffer containing guanidinium thiocyanate (up to 3 M, Melford) until no FMN was evident by UV-visible spectroscopy in the eluent, followed by a further 2 CV. Proteins were refolded by slow addition of single CV of buffers containing decreasing concentrations of guanidine hydrochloride (6, 5, 4, 3, 2, 1 and 0.0 M). Residual chaotrope was removed by washing with 5 CV of buffer. 1 CV of buffer containing FMN (0.5 mM, refolding control) or 5dFMN (0.5 mM) was added to the resin and allowed to equilibrate for 30 min. Protein samples were then eluted and further purified as previously described for FMN containing protein including anion exchange for LOV domain construct (*OdAu1<sub>aLOV</sub>*) and size exclusion chromatography (SEC) for DNA binding construct (*OdAu1<sub>aZIPLOV</sub>*). The extent of exchange was assessed by UV-visible spectra of dark-adapted protein samples using extinction coefficients at 475 nm of  $\epsilon = 10,500 \text{ M}^{-1}\text{cm}^{-1}$  for FMN and  $\epsilon = \sim 0 \text{ M}^{-1}\text{cm}^{-1}$  for 5dFMN and 406 nm  $\epsilon = 10,500 \text{ M}^{-1}\text{cm}^{-1}$  for 5dFMN. This reconstitution method resulted in < 1% of the native FMN being retained. Cofactor exchanged protein samples were further purified and stored as FMN containing samples.

#### 2.5.4 Liquid chromatography-mass spectrometry

Concentrated protein samples were diluted to approximately 1  $\mu\text{M}$  in buffer (20 mM Tris pH 8.0). For lit state experiments, protein samples were transferred into clear vials and irradiated with 450 nm light for 1 hour to ensure full lit-state generation for FMN and 5dFMN samples, respectively. For dark state experiments, protein samples were transferred into amber vials to prevent photo-adduct formation. For analysis of 5dFMN crystals, harvested crystals in loops were sequentially dipped into three portions of crystallization buffer (50 mM MES, 100 mM NaCl, 20 mM NaOAc, 20 mM  $\text{MgCl}_2$ , 5 mM EDTA, 5 mM DTT pH 6.0) to wash off some of the precipitant solution, then the loops were individually washed three times with 2  $\mu\text{L}$  of further buffer. The loop was then shaken into 50  $\mu\text{L}$  dissolution buffer (20 mM Tris, pH 8.0 buffer). Liquid chromatography mass spectrometry (LC-MS) was performed on a Waters Synapt G2-Si quadrupole time of flight mass spectrometer coupled to a Waters Acquity H-Class UPLC system. The column, an Acquity UPLC protein BEH C4 (300 Å 1.7  $\mu\text{m}$  x 2.1 mm x 100 mm), was operated in reverse phase at a constant temperature of 60 °C. The gradient employed was 95%  $\text{H}_2\text{O}$  supplemented with 0.1%  $\text{HCO}_2\text{H}$  to 35% MeCN supplemented with 0.1%  $\text{HCO}_2\text{H}$  over 50 min. Data was collected in positive electrospray ionization mode and analysed using the Waters MassLynx software version 4.1. Deconvolution of protein charged states was obtained using the maximum entropy processing software.



**Figure 2.1 SDS-PAGE analysis of *OdAu1a<sub>LOV</sub>* purification.**

A) SDS-PAGE analysis of FMN *OdAu1a<sub>LOV</sub>* purification by immobilised metal affinity chromatography (IMAC) and anion exchange chromatography. Gel one shows that protein appeared to elute at imidazole concentrations of 60 to 250 mM. Gel two indicates investigations of His<sub>6</sub>-tagged *OdAu1a<sub>LOV</sub>* cleavage with tobacco etch virus (TEV) protease to remove IMAC tag. Lane 1 indicates markers, lane 2 protein prior TEV cleavage (0 h time when TEV was added, *OdAu1a<sub>LOV</sub>* to TEV protease ratio in mg 40:1), lane 3 indicates cleavage product after 16 h incubation with TEV protease at 4 °C also indicating a decrease in molecular mass indicating complete cleavage. Gel three indicates FMN *OdAu1a<sub>LOV</sub>* purification by anion exchange chromatography indicating elution peak which indicated 450 nm absorbance maxima indicating presence of protein bound of FMN. B) Is identical to (A), but purification of 5dFMN reconstituted *OdAu1a<sub>LOV</sub>* purification. Please note, for anion exchange chromatography 404 nm absorbance maxima peak was collected indicating protein containing 5dFMN as a cofactor. C) SDS-PAGE analysis of purified *OdAu1a<sub>bZIPLOV</sub>* construct containing either FMN or 5dFMN.

### 2.5.5 NMR spectroscopy

$U\text{-}^{15}\text{N}$ -labelled protein samples were concentrated to 0.4 to 0.6 mM in NMR buffer (20 mM MES, 1 mM EDTA, 0.5 mM TCEP, 0.05 %  $\text{NaN}_3$ , pH 6.0) and 10 %  $\text{D}_2\text{O}$  was added. For dark state experiments, protein samples were transferred into amber-coloured NMR tubes. For lit-state experiments, protein samples were transferred into clear NMR tubes and illuminated.  $^1\text{H}\text{-}^{15}\text{N}$  TROSY-HSQC spectra recorded on a DPX-600 MHz Bruker NMR spectrometer equipped with a cryoprobe and preamplifiers.

### 2.5.6 Circular dichroism spectroscopy

Native, FMN to FMN and FMN to 5dFMN exchanged protein samples at 20  $\mu\text{M}$  were buffer exchanged into potassium phosphate buffer (10 mM, pH 7.0) by dialysis overnight in the dark at 4 °C, concentrated and circular dichroism (CD) spectra recorded in triplicate at 25 °C using an Applied Photophysics (Surrey, UK) Chirascan instrument. For dark-state experiments, spectra of protein samples were recorded under dim red light. For lit-state switch experiment, protein samples were illuminated with 450 nm light until UV-visible spectroscopy results indicated full conversion (~1 min for FMN and ~ 30 minutes for 5dFMN samples) before spectra were recorded. For *OdAu1a<sub>bZIPLOV</sub>*, 40  $\mu\text{M}$  protein samples were required to reliably observe structural changes.

### 2.5.7 Analytical gel filtration chromatography

*OdAu1a<sub>LOV</sub>* stocks were diluted to 200, 150, 100, 50 and 10  $\mu\text{M}$  in DNA binding buffer and loaded as 100  $\mu\text{L}$  injections onto a Superdex 75 5/150 GL (GE Healthcare) or Superdex 200 5/150 GL (GE Healthcare) column equilibrated with DNA binding buffer attached to a NGC MPLC System (Bio-Rad). Chromatograms were recorded at 260 and 280 nm wavelengths and additional wavelengths appropriate for either FMN (390 and 450 nm) or 5dFMN (360 and 400 nm). All experiments were performed at 20 °C at a flow rate of 0.5 mL/min. Dark experiments were performed by wrapping the column and injection syringe with aluminium foil and preparing the samples under dim red light. Illuminated experiments were performed using 450 nm LEDs to illuminate the sample prior to the experiment and the column throughout chromatography. A standard curve was determined using Gel Filtration Standard (Bio-Rad) following manufacturer's instructions, for results please see Appendix Figures 3.20 and 3.21. For *OdAu1a<sub>bZIPLOV</sub>* DNA binding experiments, desalted 24 bp DNA oligonucleotides (Sigma Aldrich), were dissolved in DNA binding buffer to 500

$\mu$ M. Oligonucleotides were denatured at 95 °C for 5 min and slowly annealed by cooling from 95 to 20 °C at 1 °C/min. Annealed oligonucleotide stocks were stored at -20 °C. For DNA binding experiments, 75  $\mu$ L of 100  $\mu$ M protein (dark or illuminated) was mixed with 75  $\mu$ L DNA diluted to the appropriate concentration in DNA binding buffer, incubated in the dark or under continuous blue light for 30 min at 20 °C, centrifuged at 32,000 g for 10 min and loaded onto the Superdex 200 5/150 GL column. *OdAu1a<sub>bZIPLOV</sub>* DNA binding sequence forward 5' TGT AGC GTC TGA CGT GGT TCC CAC 3' and reverse 5' TGT AGC GTC TGA CGT GGT TCC CAC 3'.

### 2.5.8 Electrophoretic mobility shift assays

EMSA experiments were performed using 7 % native polyacrylamide gels: 3.5 mL 30% acrylamide/bis-acrylamide (29:1, Sigma Aldrich), 7.5 mL ddH<sub>2</sub>O, 5 mL 1.5 M Tris.HCl pH 8.8, 200  $\mu$ L 10 % w/v ammonium persulfate (APS, Acros Chemicals), 20  $\mu$ L N,N,N',N'-tetramethylethylenediamine (TEMED, Acros Chemicals). 5' TAMRA labelled 24 base DNA sequence (100  $\mu$ M in ddH<sub>2</sub>O, Eurofins, UK) was mixed with 1.1 equivalents of the complimentary unlabelled DNA (Sigma Aldrich), heated for 5 min at 95 °C and cooled to 20 °C at 1 °C/min. 50  $\mu$ M annealed stocks were further diluted with DNA binding buffer to 1  $\mu$ M and stored at -20 °C. DNA binding studies were performed on a 50  $\mu$ L scale by mixing 1  $\mu$ L of 1  $\mu$ M labelled DNA (20 nM final concentration) into DNA binding buffer containing glycerol (10 % v/v). Bovine serum albumin solution (1  $\mu$ L, 1 mg/mL) and 0.2  $\mu$ L of poly-dI-dC (ThermoFisher Scientific) were added to prevent non-specific interactions. *OdAu1a<sub>bZIPLOV</sub>* was added from a 5  $\mu$ M stock, to a final concentration between 100 and 600 nM. Protein and DNA samples, either in the dark or under continuous blue light illumination, were incubated on ice for 5 min and then loaded into pre-run polyacrylamide gels. Gels were run in 0.5x TAE buffer containing 10 mM MgCl<sub>2</sub> for 45 min at 150 V at 4 °C under dim red light or illuminated with 450 nm light as appropriate. Images were visualized on multichannel UV-visible imager (Bio-Rad) using different filters for flavin fluorescence and TAMRA fluorescence. Images were processed and analysed using software provided by the manufacturer. *OdAu1a<sub>bZIPLOV</sub>* DNA binding sequence forward 5' TAMRA dye TGT AGC GTC TGA CGT GGT TCC CAC 3' and reverse 5' TGT AGC GTC TGA CGT GGT TCC CAC 3'.

### 2.5.9 Crystallisation of *OdAu1a<sub>LOV</sub>*

Dark-state crystals of *OdAu1a<sub>LOV</sub>* containing FMN or 5dFMN were grown under dim red light by either sitting-drop or hanging-drop vapor diffusion at room temperature. Protein in crystallography buffer (10 mg/mL) was mixed 1:1 with 0.5 to 2.0  $\mu$ l precipitant (10 to 25 % w/v PEG 1, 2 or 3k, 100 mM sodium acetate or sodium citrate pH 4.5 to 4.9, 0.1 M ammonium chloride) with a well volume of 50 to 100  $\mu$ L. Crystallisation trials were performed using an Oryx 4 robot (Douglas Instruments). Dark-state protein samples were supplemented with excess FMN or 5dFMN from 1 mM stock in crystallisation buffer with a 4:1 protein to cofactor ratio yielding the best diffracting crystals. Crystal growth was usually observed overnight with maximum growth occurring after 7 days. Crystals were harvested under dim red light using ethanediol (EDO) as cryoprotectant and were flash frozen in liquid nitrogen. 5dFMN crystals were investigated by SDS-PAGE and protein mass spectroscopy using Synapt G2-Si (Waters, USA). For illumination experiments, dark state crystals were exposed to intense blue light, 450 nm prior to harvesting, cryoprotected with EDO and flash frozen in liquid N<sub>2</sub>.

Lit-state crystals were prepared by illuminating 15 mg/mL *OdAu1a<sub>LOV</sub>* with 450 nm light prior to and during crystallisation by sitting-drop vapor diffusion. 0.4 to 0.7  $\mu$ l of protein in crystallography buffer (15 mg/mL) was mixed with 0.5 to 1.0  $\mu$ l precipitant (1.0 to 3.0 M disodium malonate pH 7.0, 200 mM Tris.OAc buffer pH 7.5 to 8.0). 0.1 to 0.3  $\mu$ l of 100 mM Co<sup>3+</sup>(NH<sub>3</sub>)<sub>6</sub> was used as an additive. Crystal growth under continuous blue light illumination was observed overnight or up to seven days later. To ensure crystal growth, variable crystallisation was performed varying addition of protein, precipitant and polyamine additive increasing chances of hit. It should be noted that crystal growth was challenging and not all set-ups resulted in crystal growth. Crystals were harvested and flash-frozen in liquid nitrogen without cryoprotection considering the high concentration of disodium malonate was used as a precipitant. Some harvested 5dFMN-containing *OdAu1a<sub>LOV</sub>* crystals were dissolved and checked by mass spectrometry to ensure they contained 5dFMN.

Diffraction datasets were collected at Diamond Light Source (Oxfordshire, UK) on beamlines I04-1, I03 or I24 with resolutions ranging from 1.36 to 1.97 Å (Tables 1 and 2). Crystals were kept at 100 K during data collection in a cold nitrogen stream. Wavelengths used are indicated in Tables 3.1 and 3.2. Data reduction was completed with Xia2 using XDS, or DIALS or Autoproc at Diamond Light Source. Data analysis and handling was carried out using CCP4. Molecular replacement was performed using Aureochrome 1a LOV domain from *Phaeodactylum tricornum* as a search model, accession code 5A8B,



using CCP4 Phaser or MolRep tools. Model building and structural refinements was achieved by alternate cycles of Coot and REFMAC5, adding solvent molecules, then FMN or 5dFMN molecules where appropriate large positive electron density peaks were observed at later stages. 5dFMN and 5dFMN-cysteine photoadduct (5dFMN-Cys) structural models were built using JLigand, then refined by molecular modelling using FMN restraints as an initial model and replacing N5 with a carbon atom (CH, C5). In dark state structures C5 was treated as  $sp^2$  hybridised, and in light state as  $sp^3$  hybridised, respectively.

## 2.6 Chapter 4 Materials and Methods

### 2.6.1 Protein constructs

Sections in red indicate affinity tags used for immobilised metal affinity chromatography (IMAC). Protein section in black indicates protein constructs generated post tobacco etch virus (TEV) protease cleavage with indicated molecular weights. Tryptophan residues, W, are indicated by bold font and underlined. Molecular mass of the construct and number of W residues is indicated below the protein sequences.

#### ***Avena sativa* LOV2 sequence, short (AsLOV2):**

MGSSHHHHHHSSGLVPRGSHMENLYFQSEFLATTLERIEKNFIITDPRLPDNPPIIFASDS  
FLQLTEYSREEILGRNCRFLQGPETDRATVRKIRDAIDNQTEVTVQLINYTKSGKKFWNL  
FHLQPMRDQKGDVQYFIGVQLDGEHVRDAAEREGVMLIKKTAENIDEAAKEL

Number of amino acids: 146

Molecular mass post TEV cleavage: 16,875.08

Number of Tryptophan residues: 1

#### ***Avena sativa* LOV2 sequence, extended (AsLOV2<sub>FL</sub>):**

MGSSHHHHHHHTYKLILNGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWYDDAT  
KTFTVTEHMENLYFQSEFLATTLERIEKNFIITDPRLPDNPPIIFASDSFLQLTEYSREEILGR  
NCRFLQGPETDRATVRKIRDAIDNQTEVTVQLINYTKSGKKFWNLFHLQPMRDQKGDV  
QYFIGVQLDGEHVRDAAEREGVMLIKKTAENIDEAAKELPDANLRPEDLWANHG

Number of amino acids: 161

Molecular mass post TEV cleavage: 18,561.86

Number of Tryptophan residues: 2



### 2.6.2 Protein construct generation

Initial AsLOV2 clone with V416I mutation was located in pNCO plasmid with N-terminus Hisact tag (GenBank EF493211). AsLOV2 gene was amplified from pNCO plasmid using a set of primers with BsaI restriction sites to allow sub-cloning into modified pET28a vectors. Cloning of the construct was performed by Dr Luke Johnson, Cardiff University. For generation of AsLOV2<sub>FL</sub> construct with an extended C-terminus J $\alpha$  helix, an additional 15 residue extension was performed by Dr Robert Mart, Cardiff University, furthermore generating a construct with N-terminus GB1 domain in order to enhance protein expression yields. All isolated plasmids were verified by sequencing (Eurofins) and these were used for protein production.

### 2.6.3 Gene expression and protein purification

Chemically competent *Escherichia coli* (*E. coli*) BL21 (DE3) (Agilent) were transformed with modified pET28a containing *Avena sativa* phototropin LOV2 domain (AsLOV2) or modified pET28a plasmid containing an extended AsLOV2 (AsLOV2<sub>FL</sub>) construct fused to C-terminus of GB1 domain. All constructs were expressed and purified identically. To assist with the NMR and CD studies, all constructs had V416I mutation yielding a more slowly photocycling protein as previously described. Transformed cells were plated onto Super Optimum Broth with Catabolite Suppression (SOC) agar plates supplemented with 50  $\mu$ g/mL Kanamycin (Kan, Melford). A single colony was used to inoculate 50 mL of defined minimal media (M9 media) supplemented with 100  $\mu$ g/mL Kan and grown overnight at 37 °C shaking vigorously. This preculture was used to inoculate large scale M9 media, 1:100 dilution, cultures grown to an OD<sub>600</sub> of 0.8 to 1.0 at 37 °C and induced with 0.5 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG, Melford) at 16 °C for 16 h. For labelling studies with 5-fluorotryptophan, previously published procedure was followed (Curtis-Marof, R., Doko, D., Rowe, M. L., Richards, K. L., Williamson, R. A. & Howard M. J. *Org Biomol Chem*. 2014 Jun 21;12(23):3808-3812). Briefly, M9 cultures were grown to an OD of 0.6 to 0.8, 60 mg of 5-fluoroindole (Fluorochem) dissolved in 0.5 mL dimethyl sulfoxide (Sigma-Aldrich) was added and cells grown for 15-30 minutes at a decreased 16 °C temperature. Protein production was induced with 0.5 mM IPTG and cells grown at 16 °C for 16 h. Cells were harvested by centrifugation, washed once with 20 mM Tris.HCl, 150 mM NaCl, pH 8.0 buffer and pellets stored at - 20 °C.

2 L culture pellets were thawed on ice, resuspended in 40 mL purification buffer A (50 mM Tris.HCl, 300 mM NaCl, 5 mM  $\beta$ -mercaptoethanol, 0.05 % w/v NaN<sub>3</sub>, pH 8.2)

supplemented with 1 mg/mL egg hen white lysozyme and 5 mM phenylmethanesulfonyl fluoride (PMSF, Melford). Cells were lysed by sonication, clarified by centrifugation (18,500 RMP, 30 minutes, 4 °C), lysate passed through 0.22  $\mu$ m filter and loaded onto 5 mL Ni<sup>2+</sup>-NTA (GE Healthcare) gravity drip column equilibrated with buffer A. Loaded sample was washed with 5 column volumes (CV) buffer A followed by 4 CV washes with buffer A supplemented with 20 and 40 mM imidazole. 1 CV of 0.1 mM flavin mononucleotide (FMN) solution in buffer A was applied, column equilibrated for 30 min, washed extensively to remove excess flavin and protein eluted with buffer A supplemented with 200 mM imidazole. The eluted protein sample was concentrated to 2 mL using spin column (10 kDa molecular weight cut off (MWCO), GE Healthcare), tobacco etch virus (TEV) protease added (TEV protease to protein ratio 1:40 in mg) and protein sample dialyzed overnight against a dialysis buffer (20 mM Tris.HCl, 150 mM NaCl, 1 mM dithiothreitol (DTT), pH 8.0), 4 L at 4 °C. Protein sample was then passed through 1 mL Ni<sup>2+</sup>-HiTrap™ column (GE Healthcare) equilibrated with dialysis buffer, eluent concentrated in a spin column (10 kDa MWCO) and buffer exchanged by three cycles of dilution and concentration into NMR buffer (50 mM NaPi, 100 mM NaCl, 0.05 % w/v NaN<sub>3</sub>, pH 6.0). For circular dichroism (CD) studies, 1 mL of protein sample at approximately 100  $\mu$ M concentration was dialyzed against 1 L of 10 mM potassium phosphate buffer pH 7.0 for 16 h at 4 °C.

#### 2.6.4 Cofactor exchange

Native FMN was exchanged with 5-deaza flavin mononucleotide (5dFMN) or FMN cofactor (refolding control) by spin column refolding. FMN was removed by concentrating final purified protein sample (5 mL 100 to 300  $\mu$ M concentration) to 1 mL in a 6 mL centrifuge spin column with 10 kDa MWCO membrane (GE Healthcare). Denaturation buffer (50 mM Tris.HCl, 6 M guanidine hydrochloride (Gnd.HCl), pH 8.2) was added and supplemented with fresh DTT to a 10 mM final concentration. DTT was added to prevent oxidation of cysteine (Cys450) residue during the refolding steps. Protein sample was concentrated and washed with denaturation buffer plus fresh DTT with the eluent analysed by UV-Vis until no FMN absorbance was evident. At that stage, protein samples were refolded by an addition 50 mM Tris.HCl, 10 mM DTT, pH 8.2 buffer containing decreasing amounts of Gnd.HCl, 5, 4, 3 M of Gnd.HCl, respectively, concentrating protein sample to 1 mL and diluting to 6 mL with an appropriate buffer. At the final stage, when protein was in 50 mM Tris.HCl, 10 mM DTT, approximately 3 M Gnd.HCl, 1 mL sample was diluted to 6 mL with buffer A containing an excess cofactor, 0.5 to 1.0 mM 5dFMN (FMN to 5dFMN exchange)

or 0.5 to 1.0 mM FMN (refolding control) adding slowly with rapid mixing. Protein sample was then concentrated to 3 mL, sample diluted to 6 mL using buffer A and small amounts of protein aggregates removed by passing through 0.22  $\mu\text{m}$  filter. Refolded protein samples were then dialyzed overnight against NMR (50 mM sodium phosphate buffer, 100 mM NaCl, 0.05% w/v  $\text{NaN}_3$ , pH 6.0) or CD (10 mM potassium phosphate buffer, pH 7.0) buffers to remove residual amounts of Gnd.HCl. For NMR studies, approximately 6 mL of sample was dialysed against 4 L NMR buffer or 0.1 mL sample against 1 L CD buffer for 16 h at 4 °C, respectively. All refolding steps were performed under dim red lights with the containers wrapped with 5 layers of aluminium foil. For 5dFMN samples, dialysis containers were wrapped with 5 layers of aluminium foil to ensure no photoactivation of the protein sample.

#### 2.6.5 UV-Vis characterisation

Protein samples in NMR buffer were characterised by UV-Vis spectroscopy using Shimadzu UV-Vis spectrometer (Shimadzu). Protein concentration was determined using FMN  $\epsilon$  of 12,600  $\text{M}^{-1}\cdot\text{cm}^{-1}$  and 5dFMN of 10,500  $\text{M}^{-1}\cdot\text{cm}^{-1}$ . Relaxation kinetics were performed in NMR buffer pH 6.0 at concentrations of 12.5  $\mu\text{M}$  for FMN and up to 50  $\mu\text{M}$  for 5dFMN samples, respectively. To investigate percentage of a possible FMN contaminant left post protein refolding, 50  $\mu\text{M}$  5dFMN sample was investigated by photoactivating the protein sample with blue LED lights until photostationary phase was achieved, then allowing protein to relax over 12 to 24 h period to investigate if dark state FMN peaks would be observed. From such characterisation, no FMN absorbance could be detected meaning FMN contamination was < 1 %. For general relaxation experiments, protein samples were illuminated with 450 nm LED lights until a photostationary state was observed. Protein concentration was kept 12.5  $\mu\text{M}$  for FMN samples. Multiple spectra were collected in 3 min intervals from 600 to 200 nm wavelength for up to 12 hours to determine relaxation kinetics. For 5dFMN samples, individual UV-Vis spectra were collected every hour for 24 hours to investigate photochemical reversion back to the dark adapted state. All experiments were performed at a stable 20 °C temperature. For data analysis to plot exponential decay of FMN containing protein samples, 447 nm absorbance maxima was plotted vs time, 3 min timepoints, using Prism8 software (GraphPad). First order decay curves were fitted yielding  $R^2$  values >0.999. From first order decay curves half-life values were determined.

### 2.6.6 Circular dichroism experiments

Circular dichroism experiments were performed using Chirascan spectrometer (Applied Photophysics) with protein samples at a  $\sim 20 \mu\text{M}$  concentration placed in a 1 mm path length quartz cuvette (Fisher Scientific). For dark state experiments, protein samples were handled under dim red light. For lit state experiments, protein samples were illuminated with 450 nm LED lights until a photostationary phase was achieved as determined by UV-Vis reference spectra. Five individual dark state and three lit state CD spectra, 200 to up to 400 nm, were collected for individual samples and averaged using Excel. Relaxation experiments were performed at wavelengths of 200 to 350 nm by taking CD spectra individually every 3 minutes at 2 nm scan intervals for up to 12 hours to ensure full recovery to the dark state for FMN sample. Hourly spectra were collected to investigate possible recovery of 5dFMN reconstituted protein. During data analysis, 280 nm absorbance values were used to correctly adjust minor protein concentration variabilities from experiment to experiment. 280 nm absorbance reading was taken prior each experiment for protein concentration adjustment if required. All experiments were performed at 20 °C. All data handling and data plotting was performed using Excel.

### 2.6.7 Nuclear Magnetic Resonance (NMR) experiments

All protein samples investigated were exchanged into NMR buffer (50 mM sodium phosphate buffer, 100 mM NaCl, 0.05% w/v  $\text{NaN}_3$ , pH 6.0).<sup>17</sup> All NMR experiments were performed on a DPX-600 MHz Bruker NMR spectrometer equipped with a cryoprobe and preamplifiers, Cardiff University, School of Chemistry.  $^{19}\text{F}$  experiments were performed by tuning  $^1\text{H}$  channel to  $^{19}\text{F}$  nuclei. The ringing effect of the  $^{19}\text{F}$  spectra initially observed was adjusted by an introduction of a 1.5 ms delay prior FID collection.  $90^\circ$  pulse was manually calibrated. All spectra were collected at a stable 293 K temperature. Internal standards used in the experiments included buffered 5 mM 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS, Sigma-Aldrich) and 10  $\mu\text{M}$  sodium trifluoroacetate (TFA, Sigma-Aldrich). Protein concentrations ranged from 0.1 to 1.2 mM depending on an experiment. Data was analysed and processed using TopSpin software and relaxation kinetics investigated using TopSpin DynamicsCentre2.5.6 software (Bruker). All  $^{19}\text{F}$  spectra reported were referenced to TFA peak as corresponding chemical shifts in ppm. For all NMR experiments, protein samples were supplemented with 10% v/v  $^2\text{H}_2\text{O}$ .

For dark state 5dFMN and FMN experiments, NMR tube containing the protein sample was protected from light by wrapping NMR tube with a tissue and further five layers

of aluminium foil creating a protective sheath. Sample preparation was performed under dim red light conditions. Under reduced light conditions (room lights switched off), an insert harbouring an optic fiber connected to 473 nm Multimode Fiber-Coupled Laser (ThorLabs) was inserted into the NMR tube, where the light source was switched off. Prepared sample was manually placed into the magnet under dim ambient light conditions (room lights switched off) with a delay time no longer than 1 min to minimise protein exposure to the light. The dark state spectra were collected accordingly. In order to photoactivate protein inside the magnet, 473 nm Multimode Fiber-Coupled Laser light source was switched on and protein samples equilibrated appropriately. Following this methodology, light intensity was adjusted accordingly. In order to photoactivate 5dFMN, dark state samples were irradiated for up to 30 minutes to ensure protein photoactivation.

#### 2.6.8 Protein analysis by Liquid Chromatography Mass Spectroscopy (LC-MS)

Dark state protein samples were from a concentrated stock were diluted to approximately 1  $\mu$ M using 20 mM Tris.HCl buffer, pH 8.0. Throughout all protein preparation steps, protein was handled under dim red lights. For dark state experiments, protein samples were placed into amber LC-MS vials to prevent protein photoactivation. For lit state experiments, protein samples were placed into clear LC-MS vials and protein samples photoactivated using 450 nm LED lights, up to 60 minutes. Liquid chromatography mass spectrometry (LC-MS) was performed on a Waters Synapt G2-Si quadrupole time of flight mass spectrometer coupled to a Waters Acquity H-Class UPLC system, Cardiff University, School of Chemistry. The column, an Acquity UPLC protein BEH C4 (300 Å 1.7  $\mu$ m x 2.1 mm x 100 mm), operated in reverse phase at a constant temperature of 60 °C. The gradient employed was 95% H<sub>2</sub>O supplemented with 0.1% HCO<sub>2</sub>H to 35% MeCN supplemented with 0.1% HCO<sub>2</sub>H over 50 min. Data was collected in a positive electrospray ionisation mode and analysed using the Waters MassLynx software version 4.1. Deconvolution of protein charged states was obtained using the maximum entropy processing software. LC-MS results can be seen in the Appendix Figures Chapter 4.

### 2.7 Chapter 5 Materials and Methods

#### 2.7.1 Protein construct generation and plasmid isolation

Full length *Ochromonas danica* Aureochrome1a (*OdAu1a*) protein in pET-M11 vector was a kind gift from Dr Harald Janovjak, IST Austria. *OdAu1a* amino acid sequence (UniProt C5NSW6) is shown below. pET-M11 plasmid was used to generate *OdAu1a*<sub>LOV</sub> constructs

as previously described. For the site directed mutagenesis of *OdAu1a<sub>LOV</sub>* to introduce point mutations QuickChange kit (Agilent, USA) protocols were followed. For primers used and the corresponding mutations please see Chapter 5 Appendix Table 5.1. Introduction of amber stop codons in the *OdAu1a<sub>LOV</sub>* was achieved by Master of Science student Stuart Yule and Postdoctoral Research Associate Dr Luke Johnson, both at Cardiff University, United Kingdom. Amber stop codons were introduced by replacing corresponding tyrosine residue codons, tyrosine182 (Y182) and tyrosine300 (Y300), with an amber stop codon (TAG). *OdAu1a<sub>LOV</sub>* construct with truncated N-terminus A' $\alpha$  ( $\Delta$ A' $\alpha$ *OdAu1a<sub>LOV</sub>*) was generated using a set of primers depicted in Table 5.1. After polymerase chain reaction (PCR), the amplified linear construct was ligated with T4 ligase (New England Biolabs, USA) following manufacturers guidelines. *OdAu1a<sub>LOV</sub>* construct with truncated C-terminus J $\alpha$  was achieved by replacing serine297 (S297) codon with a stop codon generating  $\Delta$ J $\alpha$ *OdAu1a<sub>LOV</sub>*. After each construct preparation by molecular techniques, the parental plasmid was destroyed by addition of DpnI restriction enzyme (New England Biolabs, USA) which recognises and cleaves after methylated adenine (<sup>m</sup>A) in G<sup>m</sup>ATC sequence. Treated products were then used to transform chemically competent DH5 $\alpha$  (New England Biolabs, USA) or XL1-Blue (Agilent, USA) cells which were plated onto super optimum broth with catabolite suppression (SOC) agar plates supplemented with 50 mg/L kanamycin sulphate (Kan, Melford, UK) and for XL1-Blue cells additional antibiotic, 10 mg/L tetracycline hydrochloride (Tet, Melford, UK). Colonies were allowed to develop for 16 h at 37 °C.

The following day, a single colony was used to inoculate 5 mL of SOC media supplemented with appropriate antibiotic(s) and cultures left to grow for 16 h at 37 °C shaking vigorously. Cultures were harvested by centrifugation and plasmids isolated using Qiagen MiniPrep kit (Qiagen, USA) following manufacturer's guidelines. Generation of corresponding mutations or protein constructs was confirmed by sequencing (Eurofins, UK).

In all protein sequences, bolded sequence indicates immobilised metal affinity chromatography (IMAC) oligo histidine (His<sub>6</sub>) tag removed by a treatment with tobacco etch virus (TEV) protease. Bolded, underlined and highlighted residues, Y and W, indicate targets used for <sup>19</sup>F NMR labelling.

**Table 2.4  $\Delta A'\alpha$ OdAu1a<sub>LOV</sub> and  $\Delta J\alpha$ OdAu1a<sub>LOV</sub> construct generation primers.**

Construct	Forward Primer	Reverse Primer
$\Delta A'\alpha$ OdAu1a <sub>LOV</sub>	5'AGCGCCCTGAAAATAAAGATTC TCAGTAGTGGGG3'	5'CAGCAGAACTTCGTTATCTCTGAT CCGAGCATTCCGG 3'
$\Delta J\alpha$ OdAu1a <sub>LOV</sub>	5'GGGCGTTCAGTGCAAAGTA <u>TG</u> <u>AG</u> GAGGACTATGCTAAAGC 3'	5'GCTTTAGCATAGTCCTC <u>TC</u> CATACTT TGCACTGAACGCCC 3'

Full length *Ochromonas danica* Aureochrome1a (*OdAu1a*)

**MKHHHHHHHPMSDYDIPTTENLYFQG**AMGTSKQQLPPPIFGVLGDEKQVARNGIISLV  
DIFDDFLFSGDRNQPSNTASSSSHAQESVSGKDEENDYDSNDDEGDSDDGKRRKRS  
RTLPRNMTEEQKIERRERNREHAKRSRVRKKFLLESLQHSHVRALEEENEKLRNAIRENL  
QGAEQLLTRCSCGGPSVIASDPNTATRTLDDPDYSLVKALQTAQQNFVISDPSIPDNPI  
VYASQGFLTLTGIALSEVLGRNCRFLQGPETDPKAVEKVRKGLERGEDTTVLLNYRK  
DGSTFWNQLFIAALRDGEGNVVNYLGVQCKVSEDYAKAFLKNEENEK

**Number of amino acids without His<sub>6</sub> tag: 315**

**Molecular weight: 35273.02**

*Ochromonas danica* Aureochrome1a LOV domain (*OdAu1a<sub>LOV</sub>*)

**MKHHHHHHHPMSDYDIPTTENLYFQG**AMD**Y**SLVKALQTAQQNFVISDPSIPDNPIVYASQ  
GFLTLTGIALSEVLGRNCRFLQGPETDPKAVEKVRKGLERGEDTTVLLNYRKDGSTF  
**W**NQLFIAALRDGEGNVVNYLGVQCKVSE**D**YAKAFLKNEENEK

**Number of amino acids without His<sub>6</sub> tag: 135**

**Molecular weight: 14997.90**

*Ochromonas danica* Aureochrome1a LOV domain  $\Delta J\alpha$  ( $\Delta J\alpha$ OdAu1a<sub>LOV</sub>)

**MKHHHHHHHPMSDYDIPTTENLYFQG**AMDYSLVKALQTAQQNFVISDPSIPDNPIVYASQ  
GFLTLTGIALSEVLGRNCRFLQGPETDPKAVEKVRKGLERGEDTTVLLNYRKDGSTF  
WNQLFIAALRDGEGNVVNYLGVQCKV

**Number of amino acids without His<sub>6</sub> tag: 119**

**Molecular weight: 13100.88**

*Ochromonas danica* Aureochrome1a LOV domain tyrosine300 stop (TAG)  
(*OdAu1a<sub>LOV300TAG</sub>*)

**MKHHHHHHHPMSDYDIPTTENLYFQG**AMDYSLVKALQTAQQNFVISDPSIPDNPIVYASQ  
GFLTLTGIALSEVLGRNCRFLQGPETDPKAVEKVRKGLERGEDTTVLLNYRKDGSTF  
WNQLFIAALRDGEGNVVNYLGVQCKVSE**D**

**Number of amino acids without IMAC tag: 122**

**Molecular weight: 13432.16**



### 2.7.2 Gene expression and storage of culture pellets

For generations of natural abundance unlabelled protein constructs including *OdAu1a*, *OdAu1a<sub>LOV</sub>*,  $\Delta A'\alpha OdAu1a_{LOV}$  and  $\Delta J\alpha OdAu1a_{LOV}$  chemically competent *Escherichia coli* (*E. coli*) BL21 (DE3) cells were transform with corresponding pET-M11 plasmids, transformed cells were plated onto Luria-Bertani (LB, Melford, UK) agar plates supplemented with 50 mg/L Kan and colonies were allowed to develop for 16 h at 37 °C. The following day, a single colony was used to inoculate 100 mL of LB media supplemented with 50 mg/L Kan or 100 mL terrific broth (TB, Melford) supplemented with 100 mg/L Kan with these precultures grown at 37 °C shaking vigorously. This preculture, 1:50 dilution, was then used to inoculate larger 1 L cultures of LB or TB supplemented with required amount of Kan. Cultures were grown to an optical density at 600 nm (OD600) of 0.6 in LB or 1.5 in TB at 37 °C shaking vigorously and were induced with 0.5 mM isopropyl- $\beta$ -D-thiogalactoside (IPTG, Melford, UK) at 25 °C. Cultures were grown for 16 h at a decreased temperature, were harvested by centrifugation (4k RPM, 30 min, 4 °C), pellets washed once with 50 mL ice-cold culture wash buffer (50 mM Tris.HCl pH 8.0, 150 mM NaCl) and once again concentrated by centrifugation. Remaining pellets were stored -20 °C for the future use.

For labelling of *OdAu1a*, *OdAu1a<sub>LOV</sub>* and  $\Delta J\alpha OdAu1a_{LOV}$  with 5-fluoro-L-tryptophan (5fW) previously published protocols were followed. Briefly, a single colony of BL21 (DE3) cells was used to inoculate 100 mL of minimal defined media (M9) supplemented with 100 mg/L Kan and cultures grown for 16 h at 37 °C shaking vigorously. The following day this preculture was used to inoculate 1 L of M9 media supplemented with 100 mg/L Kan in 1:50 dilution. Cultures were grown at 37 °C shaking vigorously to an OD600 of 0.6 to 0.8 at which point temperature was decreased to 25 °C, 60 mg/L of 5-fluoroindole (FluoroChem, UK) dissolved in 1 mL dimethyl sulfoxide (MSO) was added, cultures grown for an additional 15 to 30 min at a decreased temperature and protein production was induced by addition of IPTG (0.5 mM final concentration). Cultures were grown for 16 h, harvested by centrifugation, washed once with ice-cold cell culture wash buffer and remaining pellets stored at -20 °C.

For generation of U-<sup>15</sup>N labelled *OdAu1a*, *OdAu1a<sub>LOV</sub>* and  $\Delta J\alpha OdAu1a_{LOV}$  proteins, a single colony of BL21 (DE3) cells was used to inoculate 100 mL of M9 media supplemented with 100 mg/L Kan and cultures grown for 16 h at 37 °C shaking vigorously. The following day this preculture, 1:100 dilution, was used to inoculate 1 L of M9 media containing 1 g/L <sup>15</sup>NH<sub>4</sub>Cl (Sigma-Aldrich) as a sole nitrogen source. Cultures were supplemented with 100 mg/L Kan and were shaking vigorously at 37 °C to an OD600 of



0.8 at which point 0.5 mM IPTG was added and cultures grown at 25 °C for 16 h. Cultures were harvested by centrifugation, washed once with ice-cold cold cell culture buffer and pellets stored at -20 °C for future use.

For labelling of *OdAu1a<sub>LOV</sub>* with 3,5-difluoro-L-tryptophan (2fY), BL21 AI cells (Invitrogen) were transformed with pEVOL-*F<sub>n</sub>YR-E3* plasmid (a kind gift from Dr JoAnne Stubbe, Massachusetts Institute of Technology), transformed cells were plated onto SOC agar plates supplemented with 35 mg/L chloramphenicol (Cam) and colonies were allowed to develop for 16 h at 37 °C. A single colony of BL21 AI, now harbouring pEVOL-*F<sub>n</sub>YR-E3* plasmid, was used to inoculate 5 mL of SOC media supplemented with 35 mg/L Cam and culture was left to grow at 25 °C for 16 h shaking vigorously. The following day, 0.5 mL of this preculture was used to inoculate 50 mL SOC media plus 35 mg/L Cam, cultures grown at 37 °C to an OD600 of 0.3 and chemically competent cells prepared. Briefly, culture was chilled on ice for 20 min, chilled culture was then harvested by centrifugation 4k RPM at 4 °C, resuspended in CCC buffer A, incubated on ice for 10 min and concentrated again at 4 °C by centrifugation. The remaining pellet was gently resuspended into 2.5 mL CCC buffer B, suspension aliquoted as 50 to 100 µL stocks into chilled Eppendorf tubes and flash frozen in liquid nitrogen. All steps were performed aseptically, and cell aliquots were stored at a stable -80 °C temperature. In order to incorporate 2fY, chemically competent BL21-AI cells harbouring pEVOL-*F<sub>n</sub>YR-E3* were transformed with pET-M11 plasmids containing Y182TAG and Y300TAG stop codon *OdAu1a<sub>LOV</sub>* mutants. Transformed cells were then plated onto SOC agar plates supplemented with 50 mg/L Kan and 35 mg/L Cam and incubated for 16 h at 37 °C. The following day, a single colony was used to inoculate 50 mL SOC media supplemented with 50 mg/L Kan and 35 mg/L Cam. Cultures were left to grow at 25 °C for 16 h and this preculture was used to inoculate 0.5 L modified LB based autoinduction media, see section 2.3.6. Autoinduction media was grown to an OD600 of 1.0 at which point 2 mM 2fY was added and temperature was decreased to 25 °C. Cultures were grown at a decreased temperature for 24 h, was harvested by centrifugation, was washed once with ice-cold cell culture pellet buffer and remaining pellets stored at -20 °C. For preparation of *OdAu1a<sub>LOV300TAG</sub>* control construct, protein production was achieved identically as described above, but addition of 2fY was omitted. Lack of unnatural amino in the cultures resulted in generation of a truncated protein construct where Y300TAG codon simply acted as a stop codon.

2fY amino acid was synthesised and purified as previously described, where the system and improved protocol was kindly provided by Dr JoAnne Stubbe, Massachusetts Institute of Technology, USA. After purification, freeze dried amino acid was stored at room

temperature. To be used for incorporation of unnatural amino acid, stock was weighed and resuspended using 0.1 M NaOH solution to a concentration of 1 M.

#### **Calcium Competent Cells (CCC) buffer A:**

- 10 mM piperazine-*N,N*-bis(2ethanesulphonic acid) (PIPES)
- 100 mM CaCl<sub>2</sub>
- pH adjusted to 7.0 with 1 M KOH

#### **Calcium Competent Cells (CCC) buffer B:**

- 10 mM PIPES
- 100 mM CaCl<sub>2</sub>
- 15 % v/v glycerol
- pH adjusted to 7.0 with 1 M KOH

### **2.7.3 Protein purification**

All protein constructs were purified under dim red lights and identically apart from full length *OdAu1a* constructs, where the lysis buffer was supplemented with benzoase DNase (1 U/mL, Sigma Aldrich). Cell pellets containing proteins of interest were resuspended in 50 to 100 mL buffer A (50 mM Tris.HCl, 500 mM NaCl, 0.5 mM tris(2-carboxyethyl)phosphine (TCEP), pH 8.0) supplemented with 1 mg/mL egg hen white lysozyme (Sigma-Aldrich) and 1 mM phenylmethylsulfonyl fluoride (PMSF, Melford, UK). Cells were lysed by sonication on ice-water slurry for a total sonication time of 5 minutes with 5 s sonication pulses and 10 s resting time. Lysates were clarified by centrifugation at 40,000 g for 30 min at 4 °C, clarified lysate passed through 0.22 µm filter and loaded onto 5 mL Ni<sup>2+</sup>-NTA column (GE Healthcare) equilibrated with buffer A. Bound protein sample was extensively washed with buffer A (up to 20 column volumes, CV) and step gradient of buffer A supplemented with imidazole was applied, 20 and 40 mM imidazole, 10 CV individual washes. 1 CV of 0.1 mM flavin mononucleotide (FMN, Melford, UK) in buffer A was added, column equilibrated for 30 minutes and excess FMN removed by extensive washes with buffer A, approximately 10-20 CV. Corresponding proteins were eluted with buffer A supplemented with 250 mM imidazole, yellow protein sample was concentrated by spin column (10 kDa molecular weight cut off membranes, MWCO, GE Healthcare) to 2 mL, tobacco etch virus (TEV) protease was added (1:50 ration in mg TEV

to target protein, respectively) and the sample were dialysed at 4 °C for 16 h against 4 L dialysis buffer (20 mM Tris.HCl, 150 mM NaCl, 0.5 mM TCEP, pH 8.0) using 12-14 kDa MWCO dialysis membrane (Thermo Fisher Scientific). The following day protein samples were passed through 1 mL Ni<sup>2+</sup>-NTA column and flow through was collected. Yellow protein fractions were then buffer exchanged by concentrating protein using spin columns and diluting sample 12 volumes with NMR buffer (20 mM 2-(*N*-morpholino) ethanesulfonic acid (MES), 0.5 mM TCEP, 1 mM disodium(ethylenedinitrilo)tetraacetate (Na<sub>2</sub>EDTA), 5 mM NaN<sub>3</sub>, pH 6.0). This step was repeated three times to ensure that protein was completely buffer exchanged into desired buffer. Protein samples were then concentrated to 500 to 2,000  $\mu$ M concentration and were stored at -80 °C for future use.

#### 2.7.4 UV-Visible light experiments

All UV-Vis experiments were performed in a Shimadzu UV-2600 UV-Vis spectrometer in Hellma Suprasil Quartz Semi-Micro 1 cm pathlength 1 mL quartz cuvettes. Experiments were performed at 20 °C taking spectral absorbance from 200 to 600 nm using buffer as a blank. For time course experiments, spectra were recorded at 10 min intervals for up to 12 h to investigate kinetics of reversion of the lit state back to the dark adapted state. To estimate protein concentration, FMN bound to LOV domain molar extinction coefficient was used,  $\epsilon = 12,600 \text{ M}^{-1}.\text{cm}^{-1}$ .

To activate proteins with light, protein samples in a clear container or in a quartz cuvettes were illuminated with 450  $\pm$  10 nm light ( $\times 24$  0.08 Watt power individual 12 V LED lights at 20 to 60 cm distance away from the sample) for 1 to 5 min. All samples, where appropriate, were handled under safe dim red light.

Data of UV-Vis experiments was analysed using Excel, processed and kinetics plotted using Prism software (GraphPad, USA) to determine corresponding decay half-lives from fitted first order decay curves. Data handling has kept at a minimum to minimise errors.

#### 2.7.5 Protein production and purification for X-ray crystallography

To ensure reproducible crystallisation,  $\Delta J\alpha.OdAu1a_{LOV}$  construct was expressed, purified and crystallised as previously described. Yellow crystal were evident overnight that diffracted 2.0 Å. Crystals were harvested and cryoprotected under dim red lights with precipitant mixture supplemented with 30% v/v ethanediol (EDO). Crystals were flash frozen in liquid nitrogen and datasets collected at Diamond Light Source (Oxfordshire, UK)

on beamline IO3. Data reduction was completed within Diamond Light Source with Xia2 using XDS, or DIALS or Autoproc. For further structural analysis, datasets which satisfied criteria including completeness  $\geq 98\%$ ,  $I/\sigma(I)$ , multiplicity,  $R$ -merge and  $R$ -measured, total measurements, and unique reflections were selected and used for structural analysis. Data analysis and handling was carried out using CCP4i2 package. Structures were solved by molecular replacement using *OdAu1*<sub>aLOV</sub> domain (PDB 6I20) as a search model either using Phaser or MolRep tools only selectin protein chain as a search model. Structure model building and structural refinements was achieved by alternate cycles of Coot and REFMAC5, adding solvent molecules, then FMM molecules where appropriate large positive electron density peaks were observed at later stages. For dataset collection, statistics and refinement, please see Chapter 5 Appendix Table 5.2.

Attempts were made to photoactivate protein crystals prior harvesting to investigate potential structural changes. Unfortunately, crystals obtained were not large enough to withstand photoactivation and shattered upon exposure to blue light. For future experiments, larger crystals would have to be grown.

#### 2.7.6 NMR experiments

$U\text{-}^{15}\text{N}$ -labelled protein samples were concentrated to 0.4 to 0.6 mM in NMR buffer (20 mM MES, 1 mM EDTA, 0.5 mM TCEP, 5 mM  $\text{NaN}_3$ , pH 6.0) and 10 %  $\text{D}_2\text{O}$  was added. All NMR experiments were performed on a DPX-600 MHz Bruker NMR spectrometer equipped with a cryoprobe and preamplifiers, Cardiff University, School of Chemistry.  $^{19}\text{F}$  experiments were performed by tuning  $^1\text{H}$  channel to  $^{19}\text{F}$  nuclei. The ringing effect of the  $^{19}\text{F}$  spectra initially observed was adjusted by an introduction of a 1.5 ms delay prior FID collection.  $90^\circ$  pulse was manually calibrated. Spectra were collected at a stable  $15^\circ\text{C}$ ,  $25^\circ\text{C}$  or  $35^\circ\text{C}$  temperature. Internal standards used in the experiments included buffered 5 mM 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS, Sigma-Aldrich) and  $10\ \mu\text{M}$  sodium trifluoroacetate (TFA, Sigma-Aldrich). Protein concentrations ranged from 0.1 to 0.6 mM dependent on an experiment. For  $^{19}\text{F}$  1D experiments, protein concentration was at least  $100\ \mu\text{M}$ . Data was analysed and processed using TopSpin software. NMR experiments, protein samples were supplemented with at least 10% v/v  $\text{D}_2\text{O}$ . For solvent induced isotope shift (SIIS) experiments, protein samples were buffer exchanged into NMR buffer containing either 50 or 100%  $\text{D}_2\text{O}$ . 100%  $\text{D}_2\text{O}$  buffer stock was prepared by freeze drying 100 mL NMR buffer, resuspending the powder into 25 mL  $\text{D}_2\text{O}$  (99%, Sigma-Aldrich) and repeating this process two more times to ensure that most of the  $^1\text{H}$  nuclei were replaced with  $^2\text{H}$  nuclei. The final freeze dried buffer was resuspended into 100 mL  $\text{D}_2\text{O}$  and pD

adjusted using perdeuterated hydrochloric acid (DCI) and if required perdeuterated sodium hydroxide (NaOD) from 1 M stock (Sigma-Aldrich). pD was adjusted using a pH metre briefly equilibrated into 5 mL D<sub>2</sub>O bottle and approximation formula applied where  $pD = pH + 0.4$  from which observed pH was used to correctly adjust pD.

For dark state experiments, NMR tube containing the protein sample was protected from light by wrapping NMR tube with a tissue and further five layers of aluminium foil creating a protective sheath. Sample preparation was performed under dim red light conditions. Under reduced light conditions, an insert harbouring an optic fiber connected to 473 nm Multimode Fiber-Coupled Laser (ThorLabs) was inserted into the NMR tube, where the light source was switched off. Prepared sample was manually placed into the magnet under dim ambient light conditions (room lights switched off) with a delay time no longer than 1 min to minimise protein exposure to the light. The dark state spectra were collected accordingly. In order to photoactivate protein inside the magnet, 473 nm Multimode Fiber-Coupled Laser light source was switched on and protein samples equilibrated appropriately. Following this methodology, light intensity was adjusted accordingly.

#### 2.7.7 Liquid Chromatography Mass Spectroscopy (LC-MS) experiment

Protein samples from a concentrated stock were diluted to approximately 1  $\mu$ M using 20 mM Tris.HCl buffer, pH 8.0. Throughout all protein preparation steps, protein was handled under dim red lights. Liquid chromatography mass spectrometry (LC-MS) was performed on a Waters Synapt G2-Si quadrupole time of flight mass spectrometer coupled to a Waters Acquity H-Class UPLC system, Cardiff University, School of Chemistry. The column, an Acquity UPLC protein BEH C4 (300 Å 1.7  $\mu$ m x 2.1 mm x 100 mm), operated in reverse phase at a constant temperature of 60 °C. The gradient employed was 95% H<sub>2</sub>O supplemented with 0.1% HCO<sub>2</sub>H to 35% MeCN supplemented with 0.1% HCO<sub>2</sub>H over 50 min. Data was collected in a positive electrospray ionisation mode and analysed using the Waters MassLynx software version 4.1. Deconvolution of protein charged states was obtained using the maximum entropy processing software.

#### 2.7.8 Analytical gel filtration chromatography experiments

In all experiments, protein samples were in NMR buffer and loaded as 100  $\mu$ L injections onto a Superdex 75 5/150 GL (GE Healthcare) column equilibrated with NMR buffer attached to a NGC MPLC System (Bio-Rad). Protein concentrations used in each

experiment are indicated in the corresponding figures. Chromatograms were recorded at 390 nm (lit state maxima) and 450 nm (dark state maxima) wavelengths. All experiments were performed at 20 °C at a flow rate of 0.5 mL/min. Dark experiments were performed by wrapping the column and the injection syringe with aluminium foil and preparing the samples under dim red light. Illuminated experiments were performed using 450 nm LEDs to illuminate the sample prior to the experiment and the column throughout chromatography. A standard curve was determined using Gel Filtration Standard (Bio-Rad) following manufacturer's instructions, but unfortunately not all standards yielded resolved peaks. This was probably due to NMR buffer not being suitable for the proteins used. Furthermore, NMR buffer does not contain any salt (NaCl) increasing changes of protein aggregation and interaction with the matrix.

#### 2.7.9 Circular dichroism experiments

Protein sample at concentration 400 to 2,000  $\mu$ M in NMR buffer were diluted to 20  $\mu$ M using potassium phosphate buffer (10 mM, pH 7.0) and circular dichroism (CD) spectra recorded in triplicate at 25 °C using an Applied Photophysics (Surrey, UK) Chirascan instrument. For dark-state experiments, spectra of protein samples were recorded under dim red light, where protein samples were allowed to recover for up to 24 h at 20 °C. For dark to lit state switch experiments, protein samples were illuminated with 450 nm light until UV-visible spectroscopy results indicated full conversion (~1 min) before spectra were recorded. All data analysis and handling was performed using Excel.



### 3 Structural Studies of *Ochromonas danica* Aureochrome1a Reconstituted with 5dFMN

This work has been published in a peer reviewed journal Biochemistry under the name of 'Non-canonical Chromophore Reveals Structural Rearrangements of Light-Oxygen-Voltage Domain on Photoactivation'

Authors: Mindaugas E. Kalvaitis, Luke A. Johnson, Robert J. Mart, Pierre Rizkallah and Rudolf K. Allemann

Protein Data Bank entries reported in this chapter are available, free access, under accession codes of **6I20**, **6I21**, **6I22**, **6I23**, **6I24** and **6I25**. Raw data and refinement files (CCP4i2) are available upon request, please contact Prof Rudolf K Allemann. All data has been deposited to Allemann's Group External drive (Cardiff University, Wales, UK). A copy of the published manuscript can be found at the end of this thesis with a permission from ACS journal Biochemistry. The manuscript is also available online, open access, Kalvaitis M E *et al. Biochemistry* 2019, 58, 22, 2608-2616.

#### 3.1 Introduction

As members of the Per-ARNT-Sim (PAS) superfamily, Light-Oxygen-Voltage (LOV) photoreceptors act as blue-light sensing modules controlling a wide range of biological processes including phototropism, phototaxis, circadian rhythms, acclimatisation to the light, metabolic changes and stress responses<sup>1-7</sup>. The modular arrangement of sensory LOV domain proteins with intra- or intermolecular effectors found in nature<sup>4,5,8-10</sup> have inspired many synthetic designs<sup>11-16</sup>. These engineered constructs usually exhibit varying levels of photo-responsiveness, a result usually associated with the lack of understanding regarding the mechanisms of allosteric control and how the chromophore dictates structural changes<sup>11-19</sup>. In order to improve the rational design of optogenetic tools and fully exploit the potential of LOV domains, an understanding of the structural changes upon the photoadduct formation and how it controls allosteric regions within the protein is required.

In LOV domains the flavin chromophore binding site is enclosed by a five antiparallel  $\beta$ -sheets flanked by  $\alpha$ -helices and loops<sup>7,20</sup>. Blue-light absorption triggers the formation of a covalent photoadduct between the flavin isoalloxazine ring at the C4a position and the thiol of a conserved cysteine residue<sup>20</sup>, Figure 3.1 A. The flanking N- and



C-terminal helices in different LOV domains are denoted as A' $\alpha$  and J $\alpha$  respectively, and have been shown to control the activity of effector modules *via* light dependent structural changes<sup>21-29</sup>. Although structurally conserved, LOV domains have been shown to control structural rearrangements following different mechanisms. The most studied examples illustrating this versatility include C-terminal J $\alpha$  helix unfolding observed in *Avena sativa* Phototropin 1 LOV2 (AsLOV2)<sup>22,23</sup>, J $\alpha$  rotation and effector domain rearrangement in *Bacillus subtilis* YtvA (BsYtvA)<sup>25-27</sup>, and homodimerisation in *Neurospora crassa* vivid (NcVVD)<sup>6,28,29</sup>. Notably, isolated NcVVD undergoes dimerisation, while *in vivo* it was suggested to interfere with white collar complex (WC) formation through light induced heterodimerisation with the white collar protein 1 (WC1)<sup>6</sup>. The molecular basis of how such diverse results are obtained from blue light driven formation of cysteinyl-FMN adduct remains unclear<sup>7,20,28,30</sup>. One hypothesis suggests protonation of N5 of the flavin cofactor upon the photoadduct formation, changing N5 from a hydrogen bond acceptor in the dark state to a donor in the lit state, triggers a 'flip' of the highly conserved glutamine (Gln) sidechain resulting in rearrangements of hydrogen bond donor/acceptor networks across the PAS domain core<sup>31-33</sup>. The current resolution of the crystal structures of lit state LOV domains has been too low (>2.7 Å) to determine the position and the rotamer identity of the conserved Gln residues with certainty<sup>29,34</sup>, but molecular dynamics (MD) simulations on NcVVD have provided support for the N5 protonation/Gln flip hypothesis<sup>33</sup>. This hypothesis is furthermore supported by an observation where the photochemical generation of a neutral semiquinone radical resulted in activating proteins including NcVVD<sup>35</sup>. This observation was furthermore used to study Aureochrome1a LOV domain dimerisation in the solution with the results indicating that two flavin neutral semiquinone radical species were in line with crystallographic dimer arrangement<sup>36</sup>. Additionally, site directed mutagenesis of the key Gln residue confirmed the importance for the function of distantly related LOV domains including NcVVD and AsLOV2<sup>28,32</sup> suggesting a common underlying mechanism.

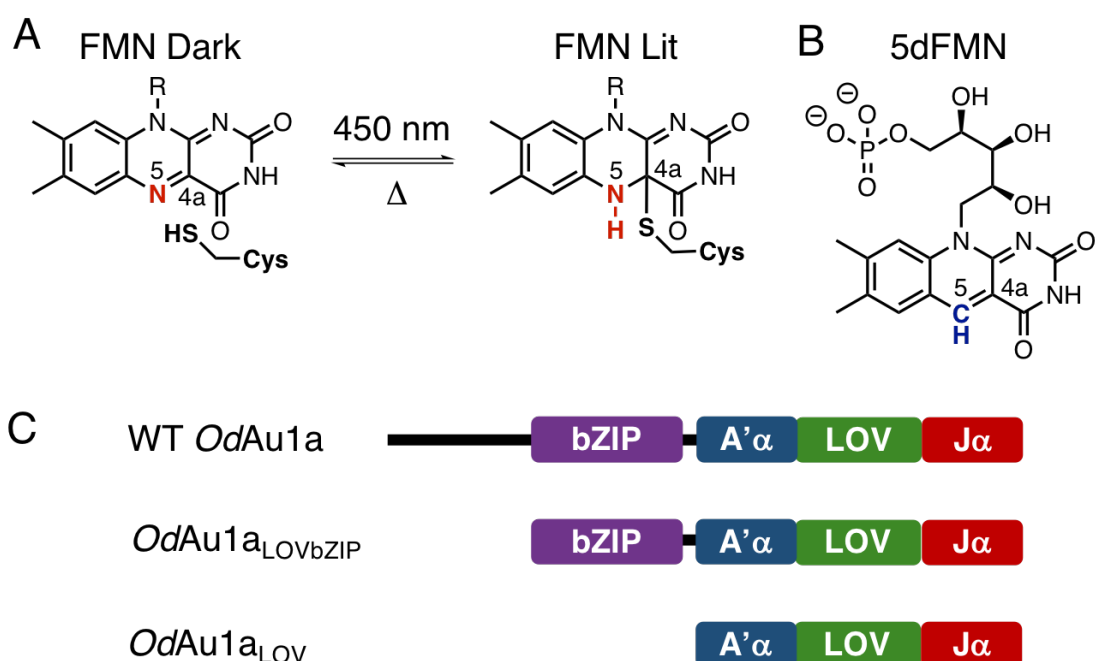
Whilst the importance of the conserved Gln is well established, alternative hypotheses on how it dictates light induced structural changes have been suggested. MD simulations of phototropin LOV domains generated a different conformation for the conserved Gln sidechain, altering the hydrogen bonding network to flanking helices<sup>37-39</sup>. Other recent reports propose that further Gln sidechain orientations are involved in LOV domain activation through hydrogen bonds extending from the FMN-O4<sup>39-43</sup>. A recent study highlighted that LOV domain of *Arabidopsis thaliana* ZEITLUPE, a LOV domain photoreceptor found in higher plants, indicated multiple states of the conserved Gln

residue<sup>40</sup>. In the dark state, the conserved Gln was a short distance from FMN-O4, 2.7 Å, but in the lit state the hydrogen bond weakened yielding multiple states around the FMN-O4. Notably, none of the structures indicated a potential Gln to FMN-N5 hydrogen bond or Gln flip. Conversely, the crystal structure of *Pseudomonas putida* BS1 LOV (*PpBS1<sub>LOV</sub>*) domain in the dark state could not pinpoint the conserved Gln residue, however it was evident in the lit state structure where Gln was in a hydrogen bond distance to the FMN-O4<sup>43</sup>. These observations therefore suggest that multiple different factors including changes in hydrogen bonding extending from FMN-O4, apart from protonation of N5 in the lit state, potentially contribute to the structural changes.

Reflecting on the importance of the potential hydrogen bonding associated with N5 of the flavin and the challenges associated with studying the lit state of thermally reverting LOV domains, in this work it was decided to use 5-deazaflavin mononucleotide (5dFMN, Figure 3.1 B) analogue that has been previously shown to form a stable photochemical cysteinyl-5dFMN adduct in *BsYtvA*<sup>44</sup> and used to study catalytic mechanisms of different flavoproteins<sup>45-47</sup>. Prior to this work, it was not known whether the lit state of 5dFMN containing LOV photoreceptors function similarly to FMN containing proteins and if the cysteinyl-5dFMN photoadduct forms at the C4a position. It was therefore decided to examine the effect of 5dFMN incorporation on the photochemistry and the function of Aureochrome1a (Au1a) transcription factor identified in *Ochromonas danica*. Uniquely, the use of 5dFMN allowed structural changes of the protein to be investigated while avoiding the use of site directed mutagenesis. Considering the current standing hypothesis explaining LOV domain photoactivation, if protonation at the N5 is essential to facilitate structural changes, 5dFMN reconstituted protein should be inactive. Conversely, if other factors including puckering of the C4a position or interruption of the conjugated isoalloxazine ring system contribute, structural changes would be observed.

Au1a photoreceptors comprise a family of LOV domain containing transcription factors found in photosynthetic stramenopiles regulating the cell cycle, growth and photomorphogenesis<sup>48,49</sup>. Au1a transcription factors contain an N-terminal unstructured region followed by a DNA binding basic leucine zipper (bZIP) domain and C-terminal LOV domain<sup>48,49</sup>, Figure 3.1 C. This domain architecture is inverted in comparison to most other LOV photoreceptors. For Au1a the N-terminal A'α helix rather than the C-terminal Jα connects the effector domain to the photosensory LOV core<sup>48,49</sup>, Figure 3.1 C. Fourier transform infrared spectroscopy (FTIR) studies on the isolated Au1a LOV domain from *Phaeodactylum tricornutum* (*PtAu1a<sub>LOV</sub>*) suggested a stepwise unfolding of A'α and Jα helices upon photochemical cysteinyl-FMN covalent bond formation resulting in the

LOV domain dimerisation<sup>21,22</sup>. The X-ray structure at 2.7 Å resolution of light-grown LOV domain suggested A'α helix rearrangement across the β-sheet surface facilitating dimerisation<sup>34</sup>. To date, the full length Au1a has not been crystallised obscuring further understanding. Low resolution structural studies employing small angle X-ray scattering (SAXS) experiments on a *PtAu1a* construct with a truncated unstructured region suggested significant shape and volume changes implying intramolecular DNA binding bZIP/LOV interactions<sup>34</sup>. Steric caging of the DNA binding domain, for instance observed in *Erythrobacter litoralis* protein 222 (EL222)<sup>50,51</sup>, could therefore complement LOV domain-driven dimerisation resulting in light dependent DNA binding. Additionally, a recent time resolved DNA binding study has suggested that light driven Au1a<sub>LOV</sub> domain dimerisation could be the major driving force controlling DNA binding<sup>52</sup>. From the kinetic measurements performed on dilute protein samples, the measurements were in line with light dependent Au1a dimerisation followed by DNA binding.



**Figure 3.1 Photochemical reaction of *OdAu1a*.**

A) Photochemical reaction of FMN with proximal cysteine residue resulting in a cysteinyl-FMN photoadduct formation at C4a position of isoalloxazine ring system. B) Structure of 5dFMN. C) Schematic representation of *Ochromonas danica* Aureochrome1a (*OdAu1a*) protein structure where solid lines indicate unstructured region, residues 1 to 94, basic leucine zipper (bZIP) DNA binding domain, residues 95 to 180, and C-terminal LOV domain, residues 181 to 312. Figure was adapted from publication with minor changes<sup>53</sup>.

In this work the functional analysis and the first high-resolution crystal structures of a *OdAu1a<sub>LOV</sub>* domain with either FMN or 5dFMN identified three distinct conformations for the conserved Gln293 linking the chromophore binding site to the allosteric A' $\alpha$  helix. The structures provide glimpses into how the covalent adduct formation triggers structural rearrangements also providing further understanding in which factors contribute in achieving the signaling active lit state conformation. This key Gln residue is highly conserved in LOV domains and these observations may provide further understanding beyond the Au1a family of LOV domain photoreceptors.

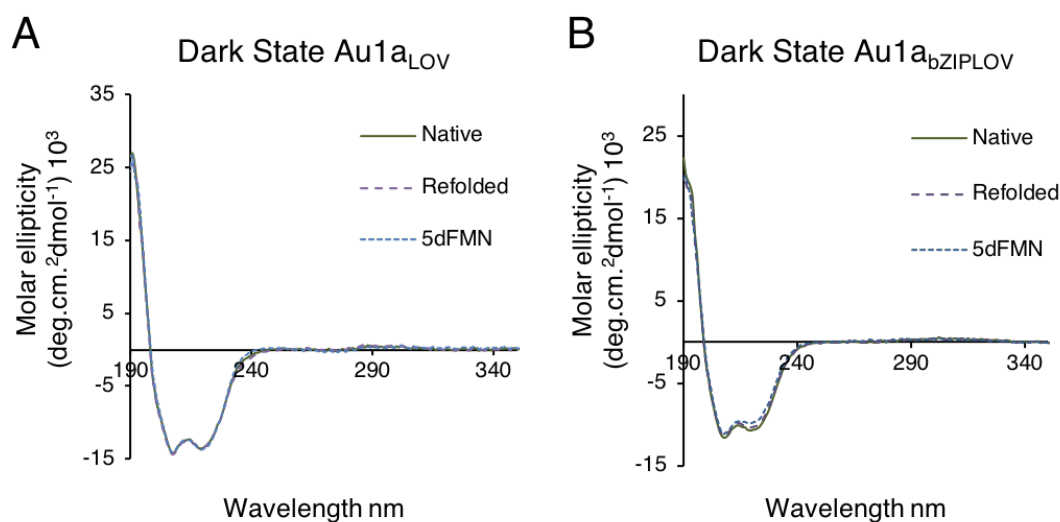
### 3.2 Reconstitution of *Ochromonas danica* Aureochrome1a with 5-deaza Flavin Mononucleotide (5dFMN)

To investigate the effects of 5dFMN incorporation into a LOV domain protein, two constructs of *Ochromonas danica* Au1a comprising the isolated LOV domain (*OdAu1a<sub>LOV</sub>*), and the LOV domain with intact DNA binding bZIP domain (*OdAu1a<sub>bZIPLOV</sub>*) were generated, Figure 3.1 C. To reconstitute *OdAu1a* constructs with 5dFMN, on-column unfolding with 6M guanidinium chloride followed by 3M guanidine thiocyanate was required to fully elute the native FMN chromophore. Initial attempts using urea or guanidinium chloride failed to fully remove FMN as determined by UV-Vis characterisation. For the 5dFMN reconstituted protein, all purification steps were performed under dim red lights to prevent photoactivation. The 5dFMN analogue was incorporated on the column once FMN and denaturant were completely removed. Upon FMN removal, decreasing concentrations of guanidinium chloride were applied and bound protein washed with buffer to remove residual chaotrope. At this point 5dFMN or FMN (refolding control) was applied, equilibrated for 30 min, and excess cofactor removed by washes with purification buffer. To mitigate potential oxidative damage to cysteine residues, buffers contained reducing agent tris(2-carboxyethyl)phosphine (TCEP) at 0.5 mM concentration.

To investigate the fidelity of the refolding methodology, protein samples were characterised by circular dichroism (CD) and nuclear magnetic resonance (NMR) spectroscopy. CD spectra with FMN and 5dFMN overlapped without any major differences with an FMN sample that had not undergone refolding. This suggests that no large-scale secondary structural differences occur upon refolding, Figure 3.2 A and B. *OdAu1a<sub>LOV</sub>* <sup>1</sup>H-<sup>15</sup>N transverse relaxation-optimised spectroscopy (TROSY) heteronuclear single quantum coherence (HSQC) spectra of the dark state FMN and 5dFMN displayed similar peak dispersion indicative of similarly folded proteins, Figure 3.3 A, B and C. The majority of

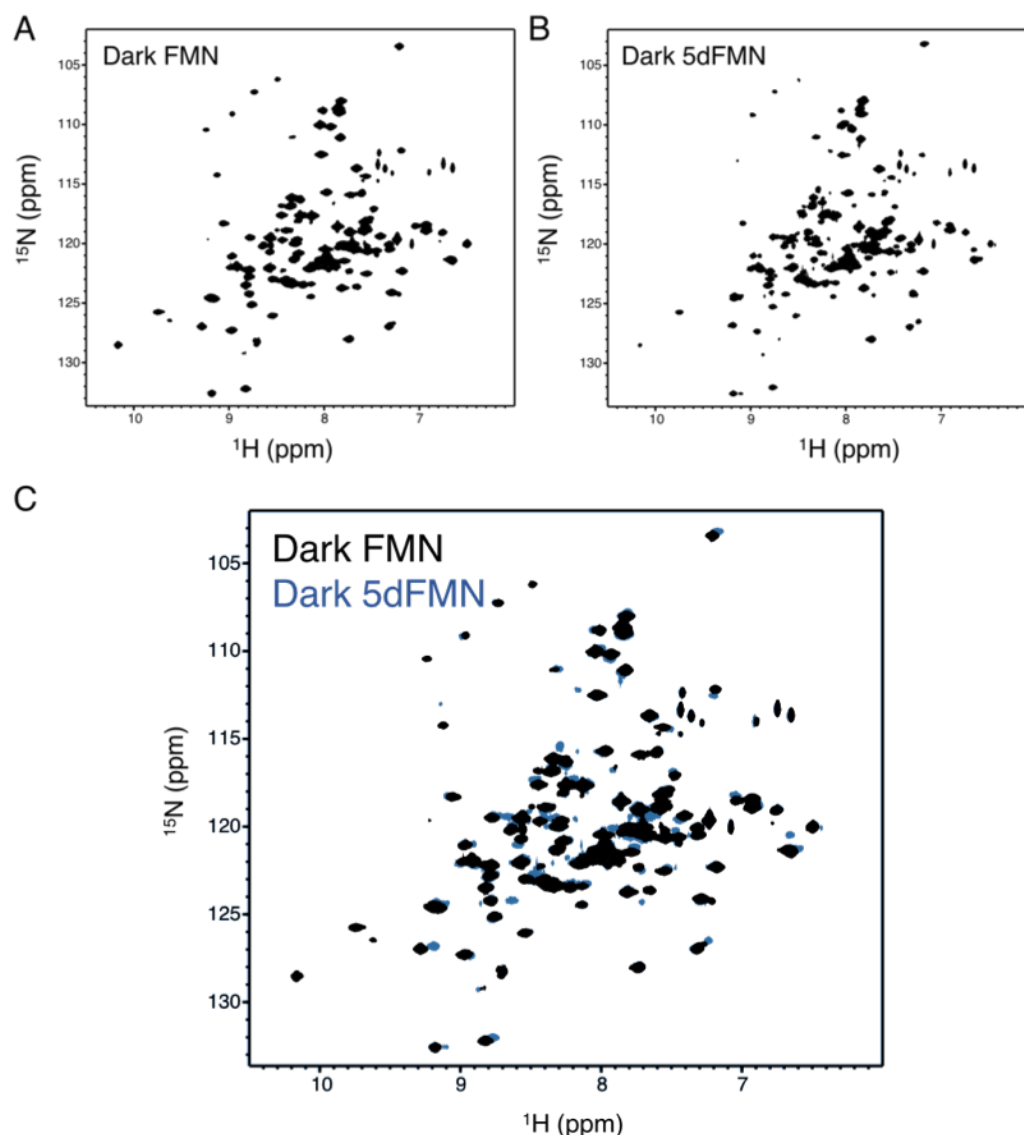
resonances were identical for FMN or 5dFMN *OdAu1a<sub>LOV</sub>* demonstrating that the protein adopts the same overall structure, Figure 3.3 C. Minor chemical shift differences were observed for a small number of resonances, most likely due to the electrostatic differences between the two cofactors, Figure 3.3 C.

To investigate the fidelity of the refolding methodology, protein samples were characterised by circular dichroism (CD) and nuclear magnetic resonance (NMR) spectroscopy. CD spectra with FMN and 5dFMN overlapped without any major differences with an FMN sample that had not undergone refolding. This suggests that no large-scale secondary structural differences occur upon refolding, Figure 3.2 A and B. *OdAu1a<sub>LOV</sub>* <sup>1</sup>H-<sup>15</sup>N transverse relaxation-optimised spectroscopy (TROSY) heteronuclear single quantum coherence (HSQC) spectra of the dark state FMN and 5dFMN displayed similar peak dispersion indicative of similarly folded proteins, Figure 3.3 A, B and C. The majority of resonances were identical for FMN or 5dFMN *OdAu1a<sub>LOV</sub>* demonstrating that the protein adopts the same overall structure, Figure 3.3 C. Minor chemical shift differences were observed for a small number of resonances, most likely due to the electrostatic differences between the two cofactors, Figure 3.3 C.



**Figure 3.2 Circular dichroism (CD) spectra of FMN and 5dFMN *OdAu1a<sub>LOV</sub>* and *OdAu1a<sub>ZIPLOV</sub>*.**

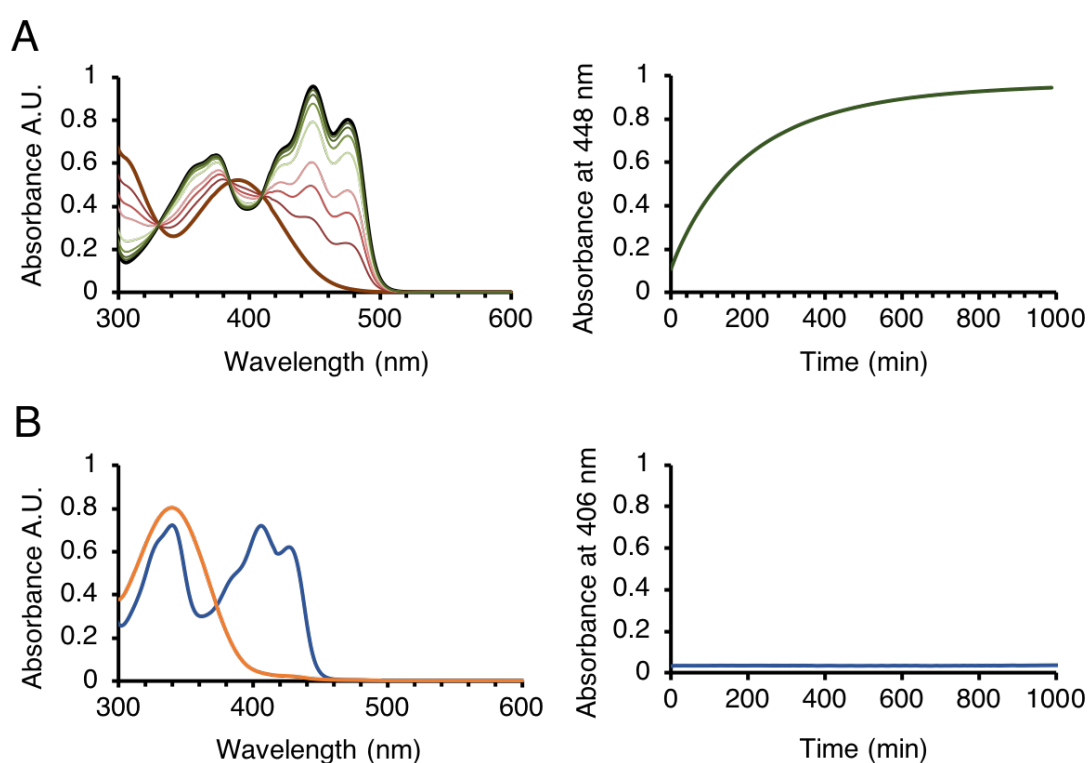
A) Overlapped CD spectra of dark state of native (not refolded), refolded FMN containing (refolding control) and refolded 5dFMN containing *OdAu1a<sub>LOV</sub>* domains and B) *OdAu1a<sub>ZIPLOV</sub>* construct. The figure was adapted from publication with minor changes<sup>53</sup>.



**Figure 3.3**  $^{15}\text{N}$ - $^1\text{H}$  TROSY-HSQC NMR characterisation of dark FMN and 5dFMN *OdAu1a<sub>LOV</sub>*. A) Dark state FMN *OdAu1a<sub>LOV</sub>*. B) Dark state 5dFMN *OdAu1a<sub>LOV</sub>*. C) Dark state FMN (black) and 5dFMN (blue) containing *OdAu1a<sub>LOV</sub>* overlapped. Please note that majority of the resonances overlapped well indicating an identical structural fold and correctly refolded protein. Few minor differences in chemical shift perturbation between FMN and 5dFMN *OdAu1a<sub>LOV</sub>* samples were evident most likely due to the electrostatic differences between the two cofactors. The figure was adapted from publication with minor changes<sup>157</sup>.

Refolded and 5dFMN reconstituted *OdAu1a<sub>LOV</sub>* and *OdAu1a<sub>bZIPLOV</sub>* produced the characteristic blue shifted vibrational triplet by UV-Vis spectroscopy of oxidized 5dFMN with maxima at 385, 406 and 423 nm, Figure 3.4 and Appendix Figure 3.18. UV-Vis spectra of protein samples using approximate extinction coefficients at 475 nm of  $\epsilon = 10,500 \text{ M}^{-1}\text{cm}^{-1}$  for FMN and  $\epsilon = \sim 0 \text{ M}^{-1}\text{cm}^{-1}$  for 5dFMN and 406 nm  $\epsilon = 10,500 \text{ M}^{-1}\text{cm}^{-1}$  for 5dFMN showed that the refolding and reconstitution method resulted in less than 1% of the native FMN retained. Photoactivation of FMN and 5dFMN samples with 450 nm light

resulted in conversion to the lit-state, with a single maximum at 385 nm for FMN and 345 nm for 5dFMN samples, Figure 3.4 A and B. 5dFMN *OdAu1a<sub>LOV</sub>* and *OdAu1a<sub>bZIPLOV</sub>* lit-states were stable for over 7 days in agreement with the proposed radical mechanism that involves formation of cyteiny-C4a photoadduct<sup>44</sup>. The high  $pK_a$  of protons at C4a position means that the rate of deprotonation is significantly reduced limiting the rate of thermal decay back to the dark-adapted state. FMN *OdAu1a<sub>LOV</sub>* thermally relaxed to the dark state with a half-life of 112 min at pH 7.4. Reversion of 5dFMN to the original dark-state was achieved with 330 nm light allowing cycling between lit and dark states for at least 5 times, without significant photodamage, Appendix Figure 3.18.



**Figure 3.4 UV-Vis spectra of FMN and 5dFMN containing *OdAu1a<sub>LOV</sub>*.**

A) Indicates UV-Vis relaxation of FMN *OdAu1a<sub>LOV</sub>* indicating a vibrational FMN triplet with three maxima at 424, 448 and 475 nm, dark green solid line. Upon photoactivation bleaching of the vibrational triplet was observed yielding lit state spectra, dark red solid line, which restored back to the dark-adapted state. Relaxation kinetics were plotted by recording 448 nm (FMN maxima) absorbance every 10 min and plotting vs time. Relaxation half-life was 112 min. B) Investigations of 5dFMN *OdAu1a<sub>LOV</sub>* lit state stability. Dark adapted state is displayed as a blue solid line with vibrational triplet maxima at 385, 406 and 423 nm. Lit state is displayed as an orange line. Potential relaxation was investigated by investigating changes at 406 nm absorbance of the lit state sample over period of time indicating no observable fluctuations. This observation indicated a lit state locked protein. The figure was adapted from publication with minor changes<sup>53</sup>.

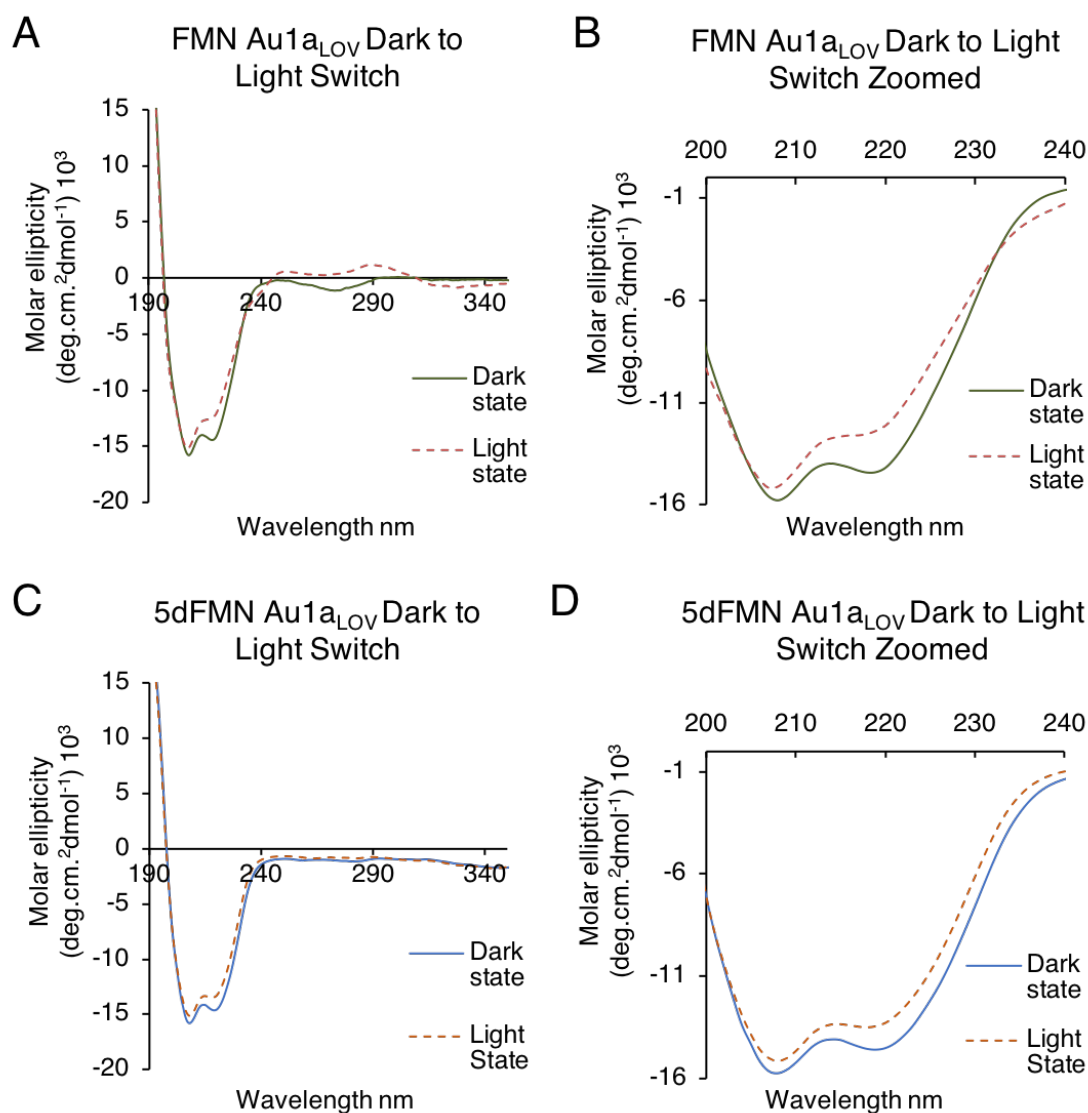
As identified by UV-Vis absorbance, illumination of 5dFMN *OdAu1a<sub>LOV</sub>* and *OdAu1a<sub>bZIPLOV</sub>* provided a stable complex that did not decay back to the dark adapted state. To establish that the photochemical covalent adduct was formed, mass spectrometry of FMN and 5dFMN *OdAu1a<sub>LOV</sub>* prior to and after illumination was performed. Dark and lit state FMN *OdAu1a<sub>LOV</sub>* samples indicated a single species with molecular mass (Mr) of 14,997.0 Da, Appendix Figure 3.19, corresponding to an apoprotein. On the other hand, lit 5dFMN *OdAu1a<sub>LOV</sub>* identified a species with a molecular mass of 15,542.5 Da not observed for the dark state sample, Appendix Figure 3.20, suggesting that cysteinyl-5dFMN adduct was stable under MS conditions. In the case of *OdAu1a<sub>bZIPLOV</sub>* constructs, lit state FMN and 5dFMN covalent adducts were not observed suggesting that the isolated LOV domain was more stable under MS conditions in comparison to the *OdAu1a<sub>bZIPLOV</sub>* construct, Appendix Figures 3.19 and 3.20.

### 3.3 CD and Nuclear Magnetic Resonance Investigations of FMN and 5dFMN Aureochrome1a LOV Domain Structural Changes

To investigate if 5dFMN could trigger light dependent structural changes CD spectra of *OdAu1a<sub>LOV</sub>* and *OdAu1a<sub>bZIPLOV</sub>* with FMN and 5dFMN cofactors were recorded under dark and lit states. When illuminated with blue light FMN-containing samples revealed a decrease in  $\alpha$ -helical content. FMN *OdAu1a<sub>LOV</sub>* had a  $14.2 \pm 0.8$  % change from dark to lit state at 220 nm minima, and  $5.4 \pm 0.8$  % at 210 nm, Figure 3.5 A and B. 5dFMN containing *OdAu1a<sub>LOV</sub>* displayed a smaller difference upon photoactivation of  $8.5 \pm 1.0$  % at 220 nm and  $3.7 \pm 1.0$  % at 210 nm, Figure 3.5 C and D. Although it is difficult to directly compare 5dFMN and FMN samples, this suggests that 5dFMN is unable to completely induce the secondary structural changes observed for the FMN *OdAu1a<sub>LOV</sub>* sample.

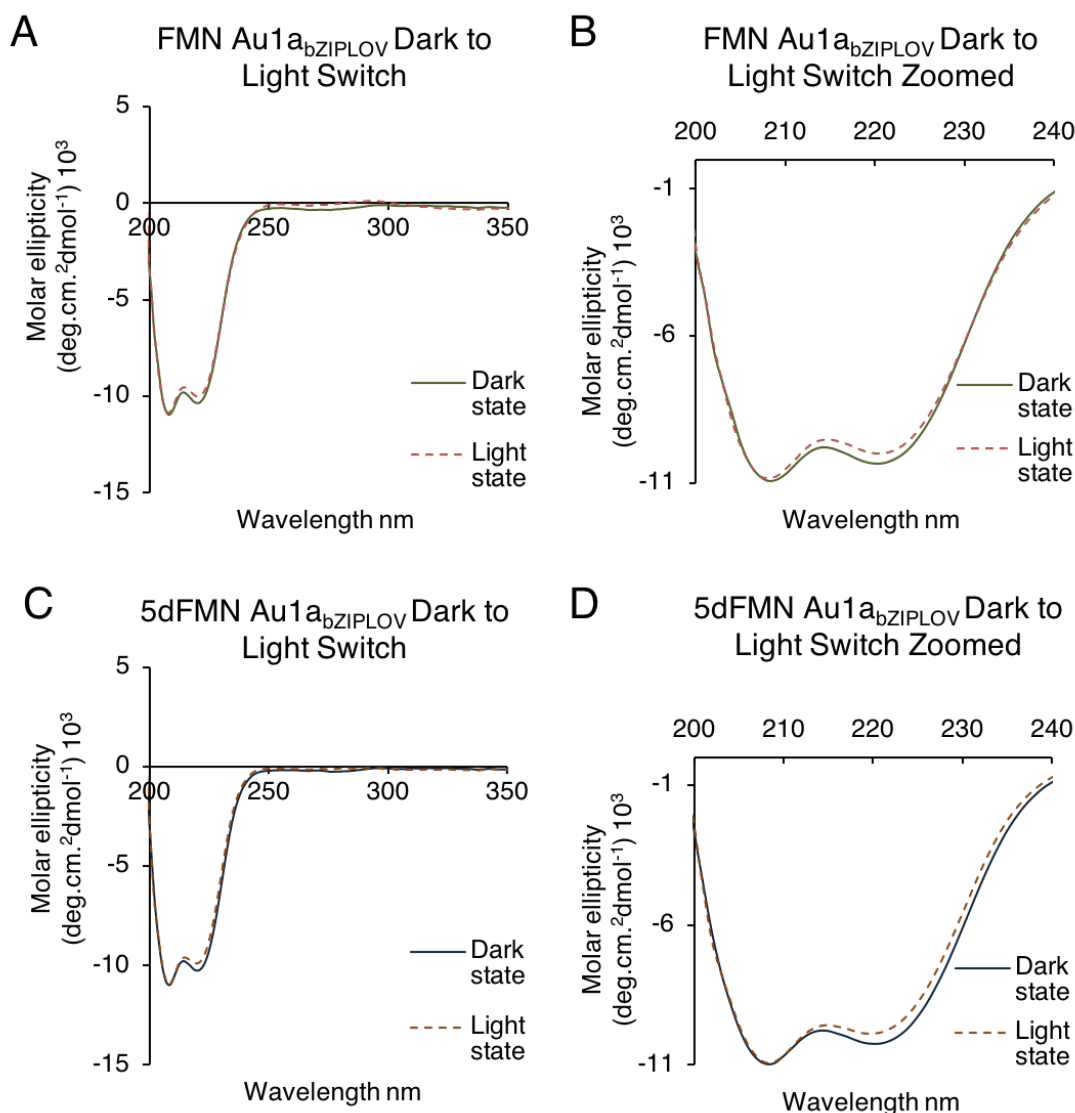
Similar, but less prominent, results were observed for *OdAu1a<sub>bZIPLOV</sub>* indicating that the secondary structural changes predominately occur within the LOV domain core. For FMN *OdAu1a<sub>bZIPLOV</sub>* a light-dark difference of  $3.4 \pm 1.4$  % and  $1.7 \pm 1.4$  % at 220 and 210 nm, were observed, Figure 3.6 A and B. 5dFMN containing protein did not result in significant structural changes, Figure 3.6 C and D, beyond the error of the experiment. This result was probably due to fact that *OdAu1a<sub>bZIPLOV</sub>* construct contains an  $\alpha$ -helical bZIP domain which masks the global changes in the protein. This could also be caused by oxidation of cysteine residues located in the bZIP region of the protein, and preventing light induced structural changes.





**Figure 3.5 Circular dichroism studies of *OdAu1a<sub>LOV</sub>*.**

A) CD spectra FMN *OdAu1a<sub>LOV</sub>* dark (green solid line) and lit (red dashed line) states, respectively. B) Zoomed CD spectra region in A). C) CD spectra 5dFMN *OdAu1a<sub>LOV</sub>* dark (blue solid line) and lit (orange dashed line) states, respectively. D) Zoomed CD spectra region in C). The figure was adapted from publication with minor changes<sup>53</sup>.

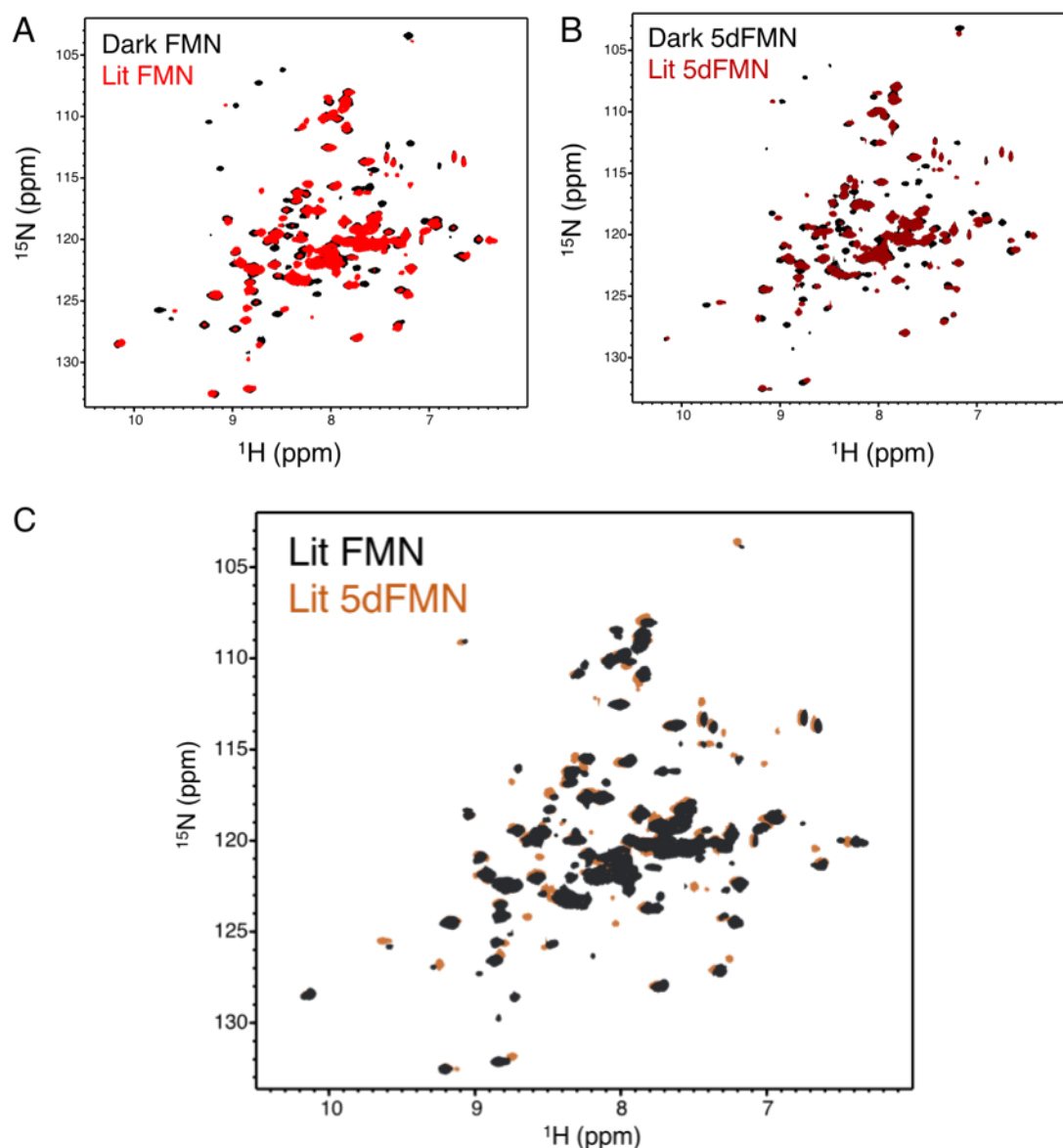


**Figure 3.6 Circular dichroism studies of *OdAu1a<sub>bZIPLOV</sub>*.**

A) CD spectra FMN *OdAu1a<sub>bZIPLOV</sub>* dark (green solid line) and lit (red dashed line) states, respectively. B) Zoomed CD spectra region in A). C) CD spectra 5dFMN *OdAu1a<sub>bZIPLOV</sub>* dark (blue solid line) and lit (orange dashed line) states, respectively. D) Zoomed CD spectra region in C). The figure was adapted from publication with minor changes<sup>53</sup>.

Native FMN containing and 5dFMN reconstituted *OdAu1a<sub>LOV</sub>* protein samples were also investigated by NMR. Illumination of FMN *OdAu1a<sub>LOV</sub>* resulted in a number of chemical shift perturbations across the <sup>1</sup>H-<sup>15</sup>N TROSY-HSQC spectra, Figure 3.7 A. Due to the relaxation of the lit state FMN *OdAu1a<sub>LOV</sub>* during data acquisition, minor peaks corresponding to the dark state were evident making complete interpretation difficult and any assignment of the resonances impossible. Illumination of 5dFMN *OdAu1a<sub>LOV</sub>* resulted in similar chemical shift perturbations to those observed for the FMN-containing sample, Figure 3.7 B and C. Nonetheless, minor differences in the chemical shifts between 5dFMN

and FMN lit states was evident indicative of minor structural differences, Figure 3.6 C. This could be a product of cofactor electrostatic variance, yet due to the lack of stability of the FMN lit state, these differences could not be explored further. Nonetheless, NMR data suggested that 5dFMN was able to dictate some structural changes but not to an extent observed for FMN containing *OdAu1a<sub>LOV</sub>*.

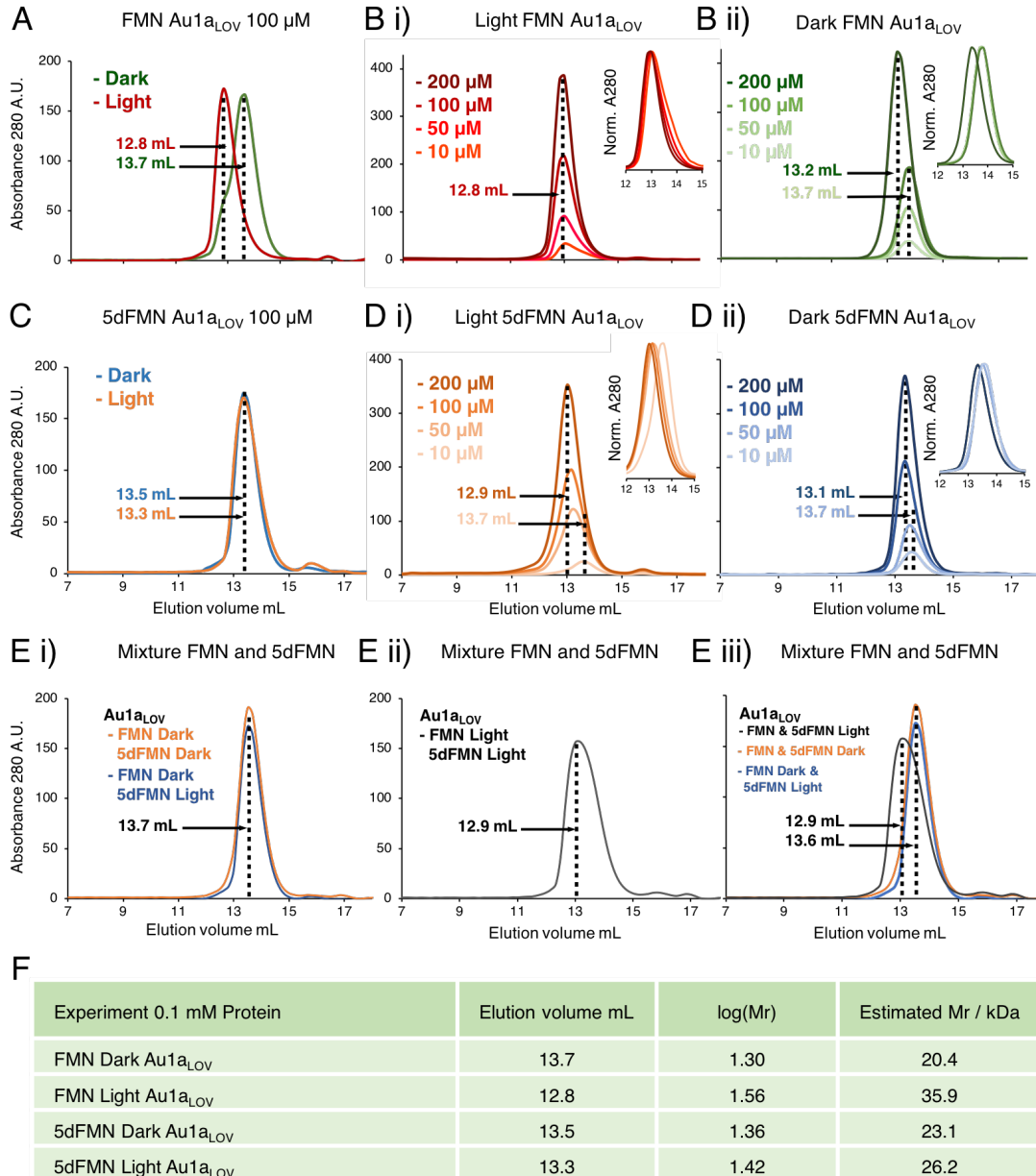


**Figure 3.7**  $^{15}\text{N}$ - $^1\text{H}$  TROSY-HSQC NMR characterisation of lit FMN and 5dFMN *OdAu1a<sub>LOV</sub>*. A) Overlapped spectra of dark (back) and lit (red) state FMN *OdAu1a<sub>LOV</sub>*. B) Overlapped spectra of dark (back) and lit (dark red) state 5dFMN *OdAu1a<sub>LOV</sub>*. C) Lit state FMN (black) and 5dFMN (orange) containing *OdAu1a<sub>LOV</sub>* overlapped. The figure was adapted from publication with minor changes<sup>53</sup>.

### 3.4 Solution Characterisation of Aureochrome1a Dimerisation and DNA Binding

As identified by UV-Vis absorbance, NMR and CD, illumination of 5dFMN *OdAu1a* provides a stable species that only reverts after illumination with ~330 nm light. In order to probe structural differences further, it was decided to investigate if 5dFMN *OdAu1a* would be functionally active. Isolated *Au1a<sub>LOV</sub>* domains undergo a light dependent dimerisation and full-length *Au1a* or *Au1a<sub>bZIPLOV</sub>* constructs show light dependent DNA binding<sup>21,22,34,49,52,54,55</sup>. Therefore analytical gel filtration chromatography experiments were undertaken to investigate light dependent behaviour of *OdAu1a<sub>LOV</sub>* and *OdAu1a<sub>bZIPLOV</sub>* with either FMN or 5dFMN cofactors. In accordance with the literature for *Au1a* homologues<sup>21,22,34</sup>, *OdAu1a<sub>LOV</sub>* appeared to undergo light dependent LOV domain dimerisation, Figure 3.8 A. Dark state FMN *OdAu1a<sub>LOV</sub>* sample eluted at 13.7 mL with an approximate molecular mass of 20.4 kDa corresponding to a monomer, Figure 3.8 A and F. Light activated FMN *OdAu1a<sub>LOV</sub>* elution maxima shifted to 12.8 mL suggesting formation of a dimer, Figure 3.8 A and F. However, the dark and lit states of 5dFMN reconstituted *OdAu1a<sub>LOV</sub>* eluted in between the FMN dark and lit states, at 13.3 and 13.5 mL, Figure 3.8 C and F. This therefore suggested a proportion of monomer and dimer states in exchange on the timescale of the size exclusion experiments.

In order to further understand *OdAu1a<sub>LOV</sub>* domain, concentration dependent behaviour was characterised. Light state FMN *OdAu1a<sub>LOV</sub>* from 200 to 10  $\mu$ M eluted at ~ 12.8 mL suggesting a formation of a relatively strong dimer, Figure 3.8 B i. Dark state FMN samples from 100 to 10  $\mu$ M yielded an elution maxima of ~ 13.7 mL, while 200  $\mu$ M sample eluted at 13.2 mL suggesting a weak oligomerisation at elevated concentrations, Figure 3.8 B ii. Characterisation of 5dFMN *OdAu1a<sub>LOV</sub>* in the lit state indicated a concentration-dependent behaviour. As the concentration decreased, the samples eluted later, ranging from 12.9 to 13.7 mL, Figure 3.8 D i. The dark state 5dFMN sample showed a similar behaviour to the lit state samples, with the elution changing from 13.1 to 13.4 mL as concentration decreased, Figure 3.8 D ii. At lower concentrations, elution maxima shifted to 13.7 mL consistent with dark state monomeric FMN *OdAu1a<sub>LOV</sub>*, Figure 3.8 D ii. For light activated FMN and 5dFMN *OdAu1a<sub>LOV</sub>* samples at 50 and 10  $\mu$ M concentrations, FMN showed asymmetry towards later elution whereas 5dFMN towards earlier elution. This observation suggested a possible dimer to monomer exchange on the timescale of the experiments.



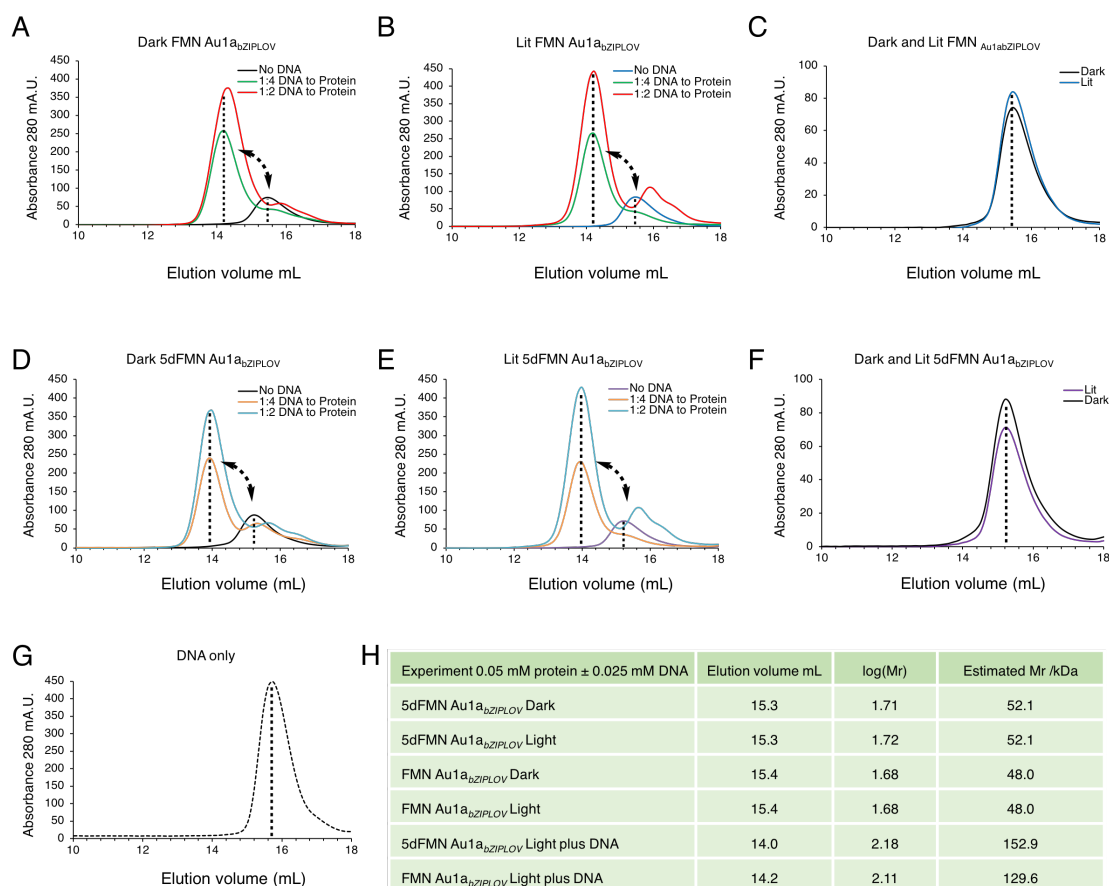
**Figure 3.8 Analytical size exclusion chromatography characterisation of *OdAu1a<sub>LOV</sub>*.**

Individual protein concentrations are indicated in the figure. A) FMN containing *OdAu1a<sub>LOV</sub>*, 100 µM, in dark (green line) and lit (red line) states. The dark state elution maxima of 13.7 mL (estimated 20.4 kDa from calibration curve, theoretical mass 15 kDa), shifted to 12.8 mL (estimated mass of 35.9 kDa, theoretical size of a dimer 30 kDa) when protein sample was photoactivated prior and during the experiment. B i) Lit-state FMN-containing *OdAu1a<sub>LOV</sub>* (200 to 10 µM) indicating an elution maximum of 12.8 ± 0.1 mL suggesting relatively strong dimer formation. B ii) Dark state FMN-containing *OdAu1a<sub>LOV</sub>* samples at 100 to 10 µM yielded an elution maximum of 13.7 ± 0.1 mL, while the 200 µM sample had an elution maximum of 13.2 mL. C) 5dFMN-containing *OdAu1a<sub>LOV</sub>* (100 µM) eluted at 13.5 mL in the dark state (blue line) and 13.3 mL in the lit-state (orange line), in-between the FMN-containing *OdAu1a<sub>LOV</sub>* dark and lit states. D i) Lit state 5dFMN *OdAu1a<sub>LOV</sub>* exhibited strongly concentration dependent elution maxima, where 200 µM sample eluted at 12.9 mL, 100 µM at 13.3 mL, 50 µM at 13.5 mL and 10 µM at 13.7 mL, respectively, with asymmetry of elution peaks towards earlier elution. D ii) A dark state 5dFMN *OdAu1a<sub>LOV</sub>* sample at 200 µM eluted at 13.1 mL, 100 µM sample at 13.4 mL, and for 50 and 10 µM samples the elution maxima shifted to 13.7 mL consistent with elution of a monomer.

E i) Size exclusion chromatography trace of a 50  $\mu$ M individual concentration mixture of dark state FMN and 5dFMN *OdAu1a<sub>LOV</sub>* (orange) and dark state FMN and lit state 5dFMN *OsAu1a<sub>LOV</sub>* (blue). E ii) Size exclusion chromatography trace of a 50  $\mu$ M individual concentration mixture of lit state FMN and 5dFMN *OdAu1a<sub>LOV</sub>*. E iii) indicates overlapped traces of E i) and E ii). F) A table representing elution maxima volumes and estimated molecular mass from size exclusion chromatography column calibration curve in Appendix Figure 22. The figure was adapted from publication with minor changes for clarity<sup>53</sup>.

From these results it was decided to investigate mixtures of FMN and 5dFMN *OdAu1a<sub>LOV</sub>* at 50  $\mu$ M individual concentrations. At 50  $\mu$ M concentration, the FMN sample showed a clear difference in the elution between lit and dark states. It was hypothesised that if 5dFMN provides lit state like behaviour and fast exchange dimer formation, it could potentially form a more stable dimer with lit state FMN sample. As expected, the mixture of dark state FMN and 5dFMN had an elution maxima of a monomer, Figure 3.8 E i. Lit state 5dFMN and dark FMN *OdAu1a<sub>LOV</sub>* indicated an elution maxima indicative of a monomer, Figure 3.8 E i. Lit FMN and 5dFMN *OdAu1a<sub>LOV</sub>* mixture showed a broad peak, Figure 3.8 E ii, with a maxima at 12.9 mL and tailing towards later elution. This data was difficult to interpret as the convolution of the traces for 50  $\mu$ M FMN, dimer, and 5dFMN, monomer, *OdAu1a<sub>LOV</sub>* in lit state may result in a similar outcome. This therefore indicated that the lit state FMN and 5dFMN *OdAu1a<sub>LOV</sub>* could not form a stable dimer. The dark 5dFMN and lit FMN *OdAu1a<sub>LOV</sub>* states could not be reliably investigated due to the set-up of the experiment requiring blue light irradiation of the size exclusion chromatography column during the functional assay.

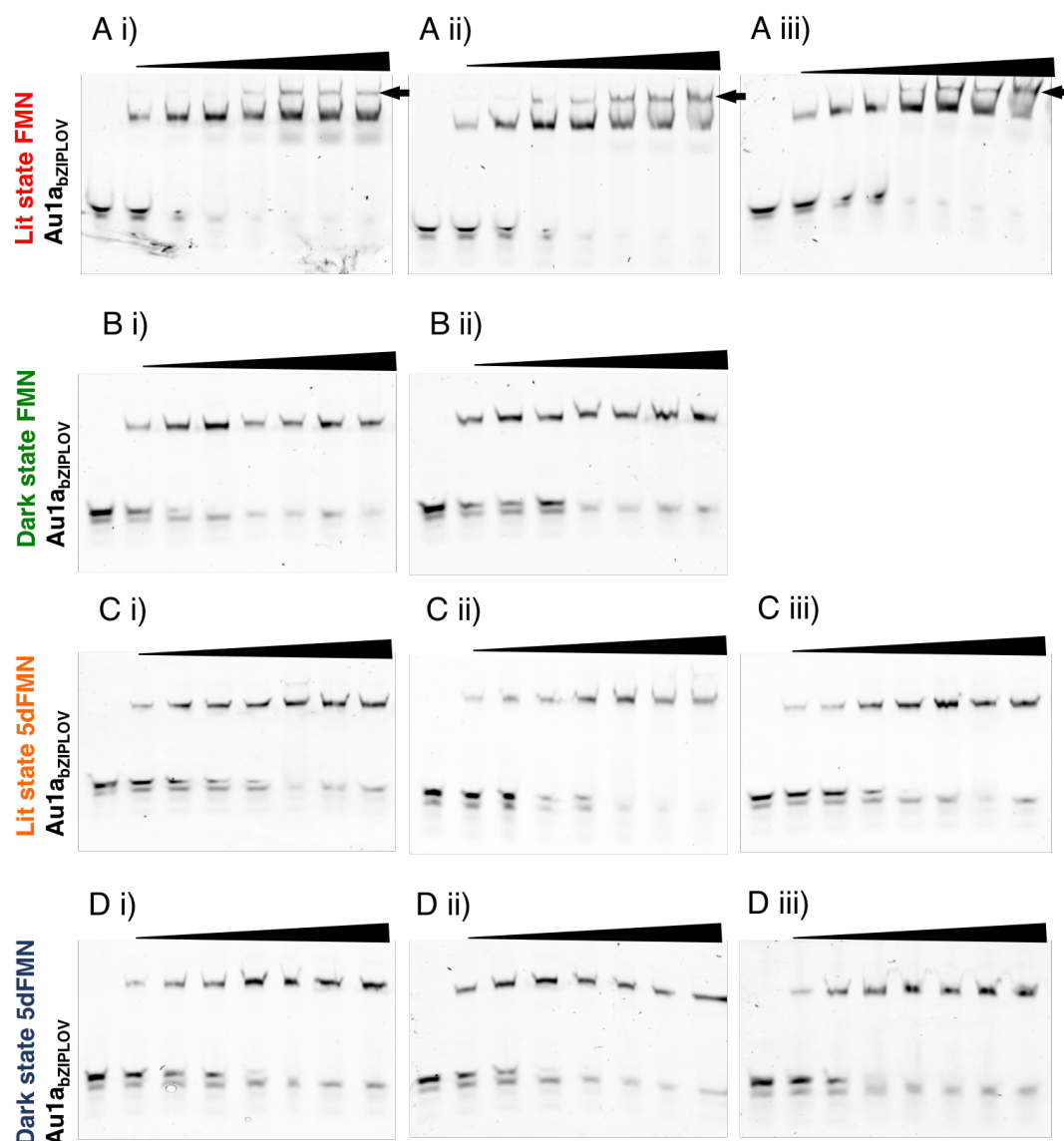
For *OdAu1a<sub>bZIPLOV</sub>* containing either FMN or 5dFMN, no changes in oligomerisation state were observed upon photoactivation. A constant elution volume of 14.5 mL corresponding to an estimated size of 72.4 kDa was observed, suggesting that *OdAu1a<sub>bZIPLOV</sub>* exists as a dimer irrespective of illumination state, Figure 3.9 C and D. Strong DNA binding was observed regardless of the cofactor and resulted in the elution volume shifting from 14.5 to approximately 12.5 mL, Figure 3.9. This shift resulted in the estimated molecular mass of approximately 150 kDa where the estimated mass of the complex was 66.3 kDa (DNA only 14.9 kDa and dimer 51.4 kDa). Although this could suggest a potential higher state oligomerisation onto the DNA, it must be taken into the account that shape of the complex also determines estimated molecular size, where a rod-shaped complex can result in a higher estimated molecular size due to molecular tumbling resulting in larger net radius. These results therefore suggest that at elevated concentrations *OdAu1a<sub>bZIPLOV</sub>* binds target DNA tightly regardless of the illumination state.



**Figure 3.9** Size exclusion chromatography characterisation of DNA binding of FMN or 5dFMN-containing *OdAu1a<sub>bZIPLOV</sub>* constructs.

50  $\mu$ M protein was used throughout experiments, while target DNA at 12.5 (1 DNA to 4 protein ratio, 1:4) and 25  $\mu$ M (1 DNA to 2 protein ratio, 1:2). A) Dark and B) Lit state FMN-containing *OdAu1a<sub>bZIPLOV</sub>* with increasing amounts of DNA titrated. C) Overlaid dark and lit state FMN *OdAu1a<sub>bZIPLOV</sub>* in the absence of DNA. D) Dark and E) Lit state 5dFMN *OdAu1a<sub>bZIPLOV</sub>* with increasing amounts of DNA added. F) Overlaid dark and lit state 5dFMN *OdAu1a<sub>bZIPLOV</sub>* in the absence of DNA. G) Trace indicating DNA alone (control). H) A table representing elution maxima volumes and estimated molecular mass from size exclusion chromatography column calibration curve, see Appendix Figure 23. Please note the size of *OdAu1a<sub>bZIPLOV</sub>* monomer from amino acid sequence 25.7 kDa, dimer 51.4 kDa and DNA only 14.1 kDa. The figure was adapted from publication with minor changes<sup>53</sup>.

In order to investigate light responsiveness of *OdAu1a<sub>bZIPLOV</sub>* containing either FMN or 5dFMN and the corresponding light dependent DNA binding, electrophoretic mobility shift assays (EMSA) were undertaken, Figure 3.10. Unlike in size exclusion experiments, EMSA experiments require significantly less protein allowing investigations of DNA binding at more physiologically relevant concentrations. Initial trials suggested that like for other *Au1a* homologues addition of  $Mg^{2+}$  ions was required in DNA binding buffers and resulted in a tighter DNA binding and higher light responsiveness<sup>49</sup>.



**Figure 3.10 DNA binding studies of FMN and 5dFMN containing *OdAu1a<sub>bZIPLOV</sub>* construct.** The first lane contains TAMRA labelled DNA only (500 nM) and the subsequent lanes have increasing *OdAu1a<sub>bZIPLOV</sub>* concentrations (0.4 to 12  $\mu$ M). In the experiments where lit state was investigated, electrophoresis tank was continuously illuminated with blue light. In experiments where dark state was investigated, protein sample was handled under dim red light and containers were wrapped in three layers of aluminium foil. All experiments were performed at 4 °C. A) Lit state FMN *OdAu1a<sub>bZIPLOV</sub>* experiment in triplicate. B) Dark-state FMN *OdAu1a<sub>bZIPLOV</sub>* in duplicate. C) Lit 5dFMN and D) dark state 5dFMN *OdAu1a<sub>bZIPLOV</sub>* experiments in triplicates. Please note that due to challenges associated with protein loading and handling under dim red lights, for the dark state FMN *OdAu1a<sub>bZIPLOV</sub>* experiment one gel was omitted for further data analysis as errors in loading were observed.

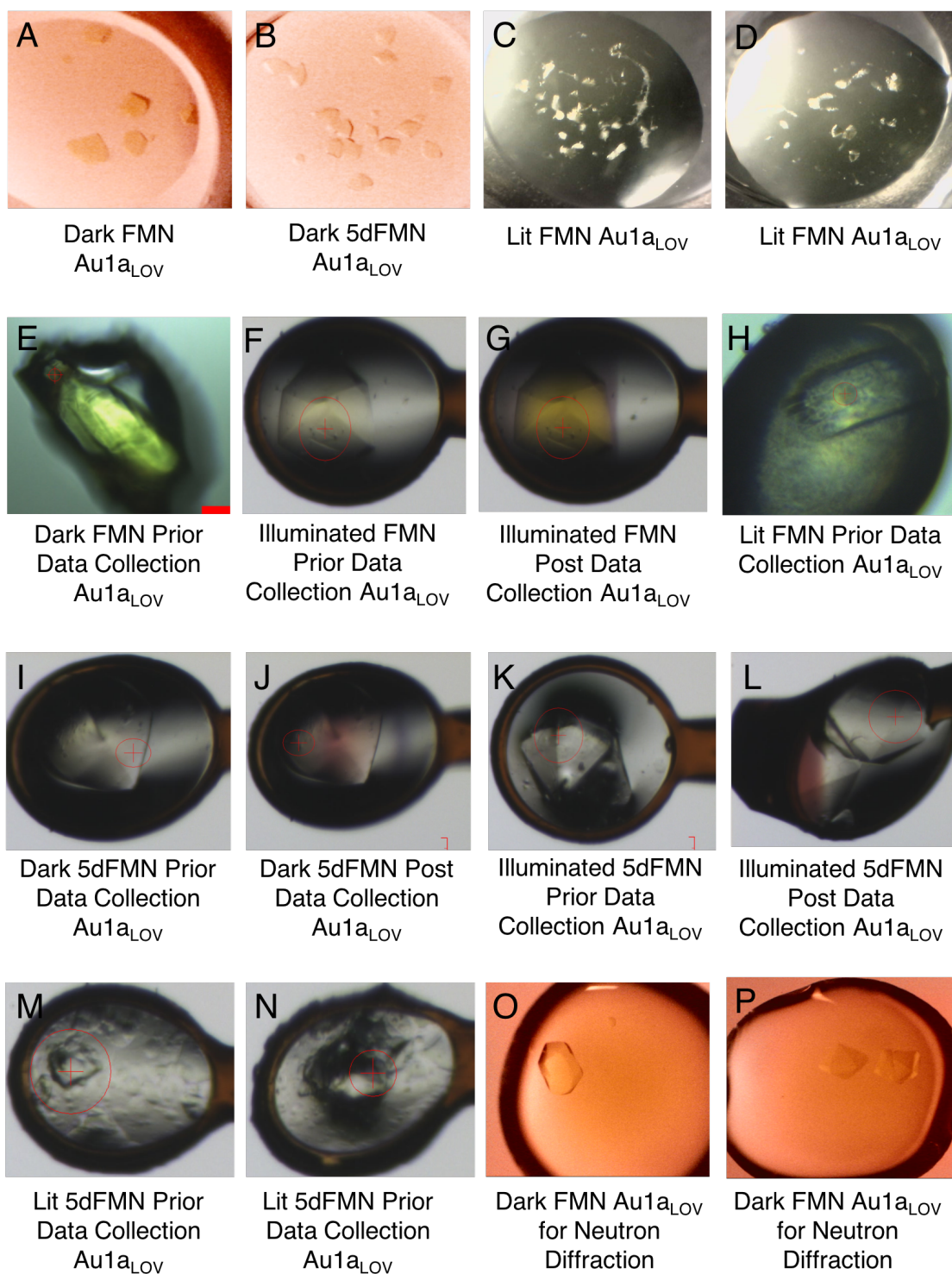
In line with gel filtration experiments, tight DNA binding was observed even in the dark, Figure 3.10 A. When comparing light and dark experiments, lit state FMN sample showed a complete saturation of DNA binding, Figure 3.10 A, while this was not observed



for the dark state sample, Figure 3.10 B. Furthermore, the photoactivated FMN sample also indicated previously observed 'super shift' behaviour<sup>49</sup>, Figure 3.10 A black arrows, not observed for the dark state sample. For 5dFMN reconstituted *OdAu1a<sub>bZIPLOV</sub>*, the dark state sample required photoactivation for 60 minutes prior to the addition of DNA to minimise damage caused to the dye. Both dark and light state 5dFMN results showed similar behaviour, Figure 3.10 C and D, essentially identical DNA binding to that of dark state FMN containing *OdAu1a<sub>bZIPLOV</sub>*, Figure 3.10 B, C and D. This therefore suggested that 5dFMN was unable to provide light responsiveness, and contrary to the lit FMN sample, did not indicate higher oligomer state 'super shift' bands. Considering the in solution characterisation, 5dFMN provided some light dependent changes but not to the extent observed for FMN-containing protein samples.

### 3.5 Structural Studies of Au1aLOV Containing FMN and 5dFMN

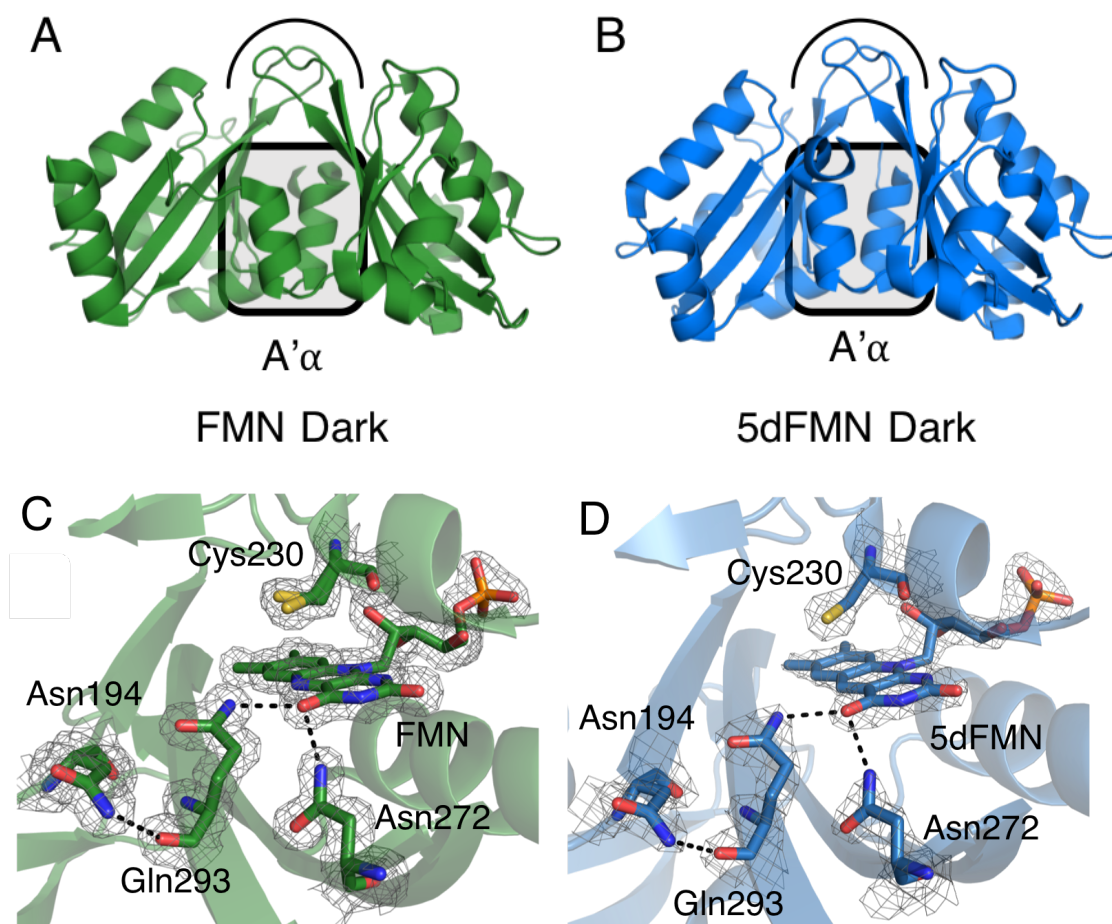
Although UV-Vis spectroscopy indicated that 5dFMN is able to form a cysteinyl covalent adduct, to the best of our knowledge, no direct experimental evidence for the formation of a cysteinyl-5dFMN covalent adduct at the C4a position has been achieved. To address this protein crystallisation was undertaken. To determine the structural changes of *OdAu1a<sub>LOV</sub>* at an atomic resolution, crystallisation was attempted with either FMN or 5dFMN under different illumination conditions. Crystals were grown and harvested under dark conditions (dark state), grown under dark and then illuminated with blue light prior harvesting (illuminated state) and grown under blue light conditions (lit state). Initial attempts to crystallise *OdAu1a<sub>LOV</sub>* using commercial screens including Molecular Dimension Morpheus, Structure 1+2, JCSG, PACT, and Hampton Research Natrix or Index failed to yield either dark or lit state crystals. A protein databank (PDB) search of multiple LOV domain proteins including different Au1a<sub>LOV</sub> homologues suggested much lower pH values for the dark state crystals. Considering this, an inhouse screen varying 10 to 20 % w/v PEG 1, 2 or 3K, 0.1 M NH<sub>4</sub>Cl, 0.1 M sodium acetate or 0.1 M disodium citrate pH 4.5 to 4.9 was prepared. Using this screen, crystals of *OdAu1a<sub>LOV</sub>* in the dark were achieved overnight, Figure 3.11 A and B, for FMN and 5dFMN samples. Attempts to crystallise the lit state under identical conditions did not yield any robust crystals. In some cases there was growth of needle clusters, Figure 3.11 C and D. Further screening was undertaken in an attempt to crystallise the lit state protein.



**Figure 3.11 Images of  $OdAu1a_{LOV}$  crystals.**

Crystals of dark state A) FMN and B) 5dFMN containing  $OdAu1a_{LOV}$ . C) and D) indicate needle clusters observed when  $OdAu1a_{LOV}$  was attempted to be grown under intense blue light irradiation throughout the experiment. E) to N) indicate images of  $OdAu1a_{LOV}$  crystals taken in loops at Diamond Light Source and appropriate information can be found in the figure. O) and P) indicate large crystals grown to investigate if dark state FMN  $OdAu1a_{LOV}$  crystals large enough for the neutron diffraction could be grown. Crystal size lengthwise in O) was 1.2 mm and sideways 0.8 mm. P) diamond shaped crystals observed where the diameter was 0.9 mm. In O) and P) the droplet volume was 6  $\mu$ L consisting of 3  $\mu$ L of protein and 3  $\mu$ L of precipitant mix.

The highest resolution, 1.36 Å, diffraction data for the dark state FMN *OdAu1a<sub>LOV</sub>* was obtained from a single crystal in space group  $P 2_1 2_1 2_1$ , Figure 3.11 E. Crystals in space group  $P 3_2 2 1$  were also observed, but yielded much poorer diffraction with resolution of around 2.0 to 2.5 Å. The highest resolution for 5dFMN *OdAu1a<sub>LOV</sub>* dark state was obtained under identical conditions in space group  $P 3_2 2 1$ , Figure 3.11 I.



**Figure 3.12 X-ray crystal structures of dark state FMN and 5dFMN *OdAu1a<sub>LOV</sub>*.**

A) 1.36 Å structure of the dark state FMN *OdAu1a*. The asymmetric unit contained four monomers as parallel dimers (green) with A'α positioned across the β-sheet surface (PDB 6l20). B) 1.97 Å structure of dark-state 5dFMN-containing *OdAu1a* with dimer (PDB 6l23) similar to dark-state FMN showing similar A'α positioning across β-sheet. RMSD value for Chains A for dark state FMN and 5dFMN structures was obtained as 0.271 (782 to 782 atoms). Electron density maps displayed as a grey mesh for FMN or 5dFMN containing *OdAu1a<sub>LOV</sub>* at  $\sigma = 1$ . Black dashed lines indicate predicted hydrogen bonding networks in distances up to 3.2 Å. C) Dark state FMN and D) Dark state 5dFMN.

The structures were solved by molecular replacement using *Phaeodactylum tricornutum* Au1a<sub>LOV</sub> domain dark state structure, PDB entry 5A8B, as a search model. In order to minimise potential bias, structures were solved using a search model excluding FMN chromophore, additives or solvent molecules. Rounds of structure building and refinement were achieved by building a correct sequence of amino acid residues, then in later cycles inserting water molecules followed by an insertion of a chromophore in the final cycles where a large positive electron density was observed. Comparing both structures, the LOV domain core appeared to be identical, Figure 3.12 A and B. The electron density for N terminus A'α and C terminus Jα had less defined maps in the 5dFMN *OdAu1a<sub>LOV</sub>* structure. No structural differences could be identified when interpreting electron density maps within the chromophore binding pocket, Figure 3.12 C and D, confirming an identical mode of cofactor binding and residue arrangement.

When dark grown crystals were illuminated with light prior to harvesting, crystals diffracted in a space group P 6<sub>4</sub> 2 2, Figure 3.10 F, G, K and L. Previous reports have identified that high energy X-rays rapidly decay the cysteinyl-FMN photoadduct<sup>29,34</sup>. To overcome this problem, datasets from a single crystal in multiple sections are usually collected to mitigate damage and then are compiled to a full dataset<sup>29,34</sup>. Although this approach is useful, it usually results in poorer overall resolution. In order to investigate minor structural changes, a balance between the resolution and the degree of the covalent adduct was considered. Eventually, it was decided to collect datasets from the best diffracting crystals until a full dataset was collected. Upon solving the structure, a single monomer per asymmetric unit was observed. Considering symmetry partners, a dimer could be identified, Figure 3.13 A and B, which was almost identical to the dark-adapted state dimer. A well-defined N-terminal A'α helix and Jα were identified from the electron density maps. While refining the structures, closer inspection of the cofactor binding site showed density for a partial covalent bond between Cys230 and FMN or 5dFMN. Positioning of the Cys230 residue also indicated the covalent adduct formed between C4a position for both FMN and 5dFMN, providing the first direct evidence that 5dFMN photochemically reacts similarly to FMN. Multiple trial and error attempts resulted in a 30% occupancy for the covalent cysteinyl-FMN or 5dFMN photoadduct, where no red electron density was observed between Cys230 to FMN or 5dFMN C4a. Reflecting on the dark and illuminated structures, differences in Cys230 thiol distance from C4a of FMN or 5dFMN were observed. For the dark state FMN structures, Cys230 thiol group was 3.9 Å away from the C4a, while for dark state 5dFMN 3.6 Å. To further minimise bias, FMN and 5dFMN

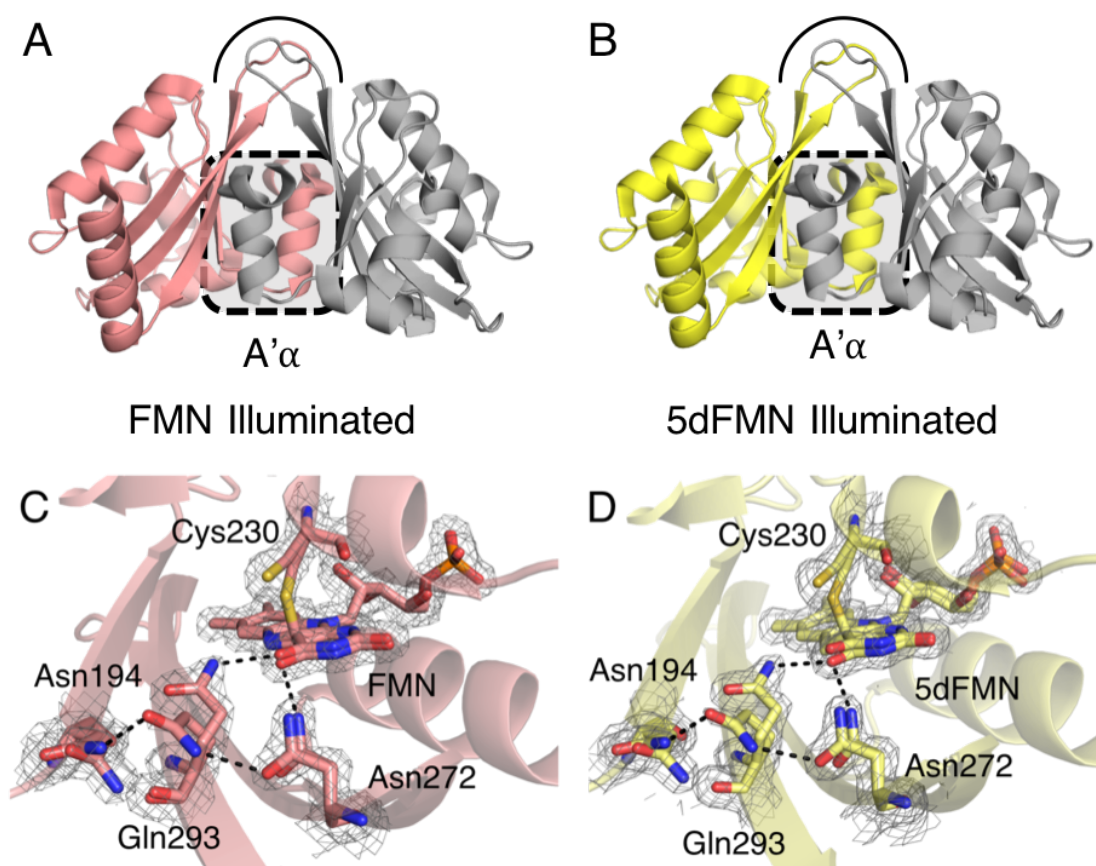
were modelled as 70% for the dark state and 30% for the lit state orientation of the cofactor, respectively.

In both cases the model's net result was 30% conformer with cysteinyl-FMN or 5dFMN adduct with bond length of 1.8 Å, and 70% dark-like state with Cys230 thiol 3.0 Å away from C4a. This distance of 3.0 Å was slightly shorter than expected for the dark state conformation, but this outcome was probably a product of X-ray dependent scission throughout dataset collection.

In addition to multiple cofactor poses, in both illuminated FMN and 5dFMN containing structures, the conserved Gln293 and Asn272 residues showed multiple states not observed in the dark state structures. Additionally, the electron density map for Asn194, a residue positioned in a loop region between A' $\alpha$  and  $\beta$ A, was diffused suggestive of multiple conformations. Furthermore, Gln293 appeared to swing out of FMN binding pocket into a new position with 20% occupancy determined by rounds of refinements, forming a possible hydrogen bond with the Asn194 sidechain, Figure 3.12 B and E. Additionally, Asn272, a residue that hydrogen bonds to O4 of the FMN ring also indicated multiple states,  $\sim$  3.0 Å and 3.4 Å away, Figure 3.13 C and D. Distinct conformational poses of Asn194, Gln293 and Asn272 suggested the formation of a hydrogen bond network between these three residues. Although the electron density map for swung out Gln293 is low, elimination of this conformer from refinement leads to a large positive electron density trace as observed in the early stages of refinement, Figure 3.15 B and E.

To ensure that 5dFMN reconstituted *OdAu1a<sub>LOV</sub>* did not enrich FMN-containing protein in the protein crystal lattice, illuminated 5dFMN *OdAu1a<sub>LOV</sub>* crystals were harvested and characterised by LC-MS. From UV-Vis characterisation, it was evident that less than 1% of contaminant protein was observed, but during crystallisation it is possible that FMN containing protein could get enriched. LC-MS results suggested that the dominant protein species was the covalent 5dFMN-*OdAu1a<sub>LOV</sub>* photoadduct suggesting that the crystals contained cysteinyl-5dFMN *OdAu1a<sub>LOV</sub>* protein, Appendix Figure 3.21 A and B. Due to presence of polyethylene glycol (PEG) in the crystallisation condition, FMN crystals were not characterised in the same way in the sequential experiment. It was observed that after the initial run residual PEG caused poor signal to noise (column contamination).



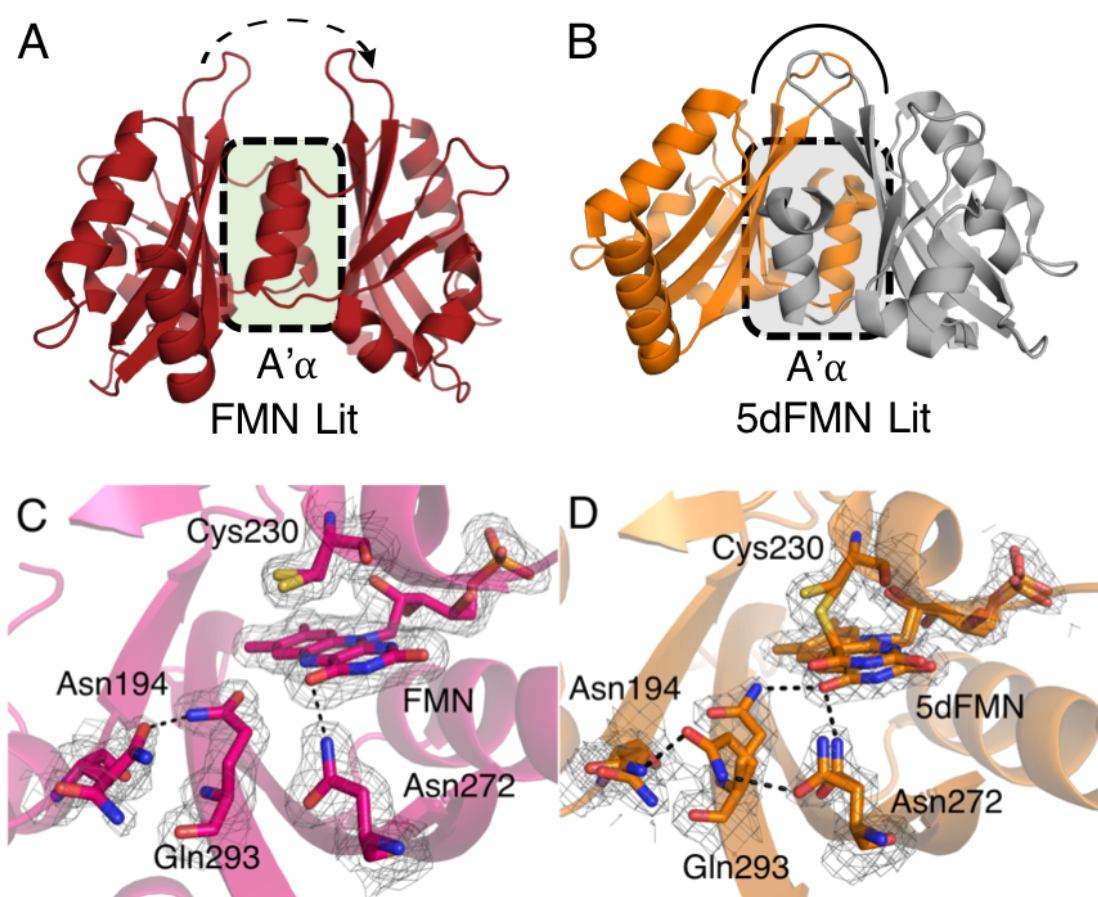


**Figure 3.13 X-ray crystal structures of illuminated state FMN and 5dFMN *OdAu1a<sub>LOV</sub>*.**

A) 1.50 Å structure of illuminated crystals of dark grown FMN *OdAu1a*. The asymmetric unit contained a single monomer (pink), forming a parallel dimer similar to the dark-state when considering a symmetry equivalent (grey). B) 1.43 Å structure of illuminated 5dFMN containing *OdAu1a* with a symmetry partner equivalent dimer displayed in grey. RMSD value for Chains A for illuminated state FMN and 5dFMN structures was obtained as 0.056 (932 to 932 atoms). Electron density maps displayed as a grey mesh for FMN or 5dFMN containing *OdAu1a<sub>LOV</sub>* at  $\sigma = 1$ . Black dashed lines indicate predicted hydrogen bonding networks in distances up to 3.2 Å. C) Illuminated FMN and D) Illuminated 5dFMN.

Lastly, when crystals were grown under blue light irradiation, no growth was observed under identical conditions excluding growth of some needle clusters. The 5dFMN sample showed some crystal growth, but these crystals dissolved when cryoprotection was attempted. Light state crystal growth was achieved under different conditions, 1.0 to 3.0 M disodium malonate pH 7.0 and 200 mM Tris-acetate buffer pH 7.5 to 8.0 supplemented  $\text{Co}^{3+}$  hexammine chloride as an additive, yielding monoclinic crystals in space group C 1 2 1 for FMN and crystals in space group P 6<sub>4</sub> 2 2 for 5dFMN samples. As harvested crystals were in 2.0 to 3.0 M disodium malonate, no cryoprotection was required. For 5dFMN, crystal growth was evident overnight, while for FMN sample crystal growth took up to seven days. When solving the structure of lit-state FMN *OdAu1a<sub>LOV</sub>*, no

covalent bond character could be observed suggesting a much faster covalent bond scission during data collection. Likewise, 5dFMN also indicated weaker covalent bond character. It is possible that this faster scission is due to the elevated pH of the crystallisation conditions.



**Figure 3.14 X-ray crystal structures of lit state FMN and 5dFMN *OdAu1a<sub>LOV</sub>*.**

A) 1.67 Å structure of lit state grown FMN *OdAu1a* featuring the unique dimer arrangement with A'α repositioning across the β-sheet surface and loop region rearranging. The asymmetric unit contained four monomers as parallel dimers. One dimer was identical to the dark state structure. The second dimer highlighted different A'α arrangement indicated here. B) 2.00 Å structure of light grown 5dFMN-containing *OdAu1A* showing A'α helix arrangement similar to the illuminated proteins with symmetry equivalent indicated in grey. RMSD value for Chain C of FMN lit and Chain A of 5dFMN lit was obtained as 0.561 (807 to 807 atoms). Electron density maps displayed as a grey mesh for FMN or 5dFMN containing *OdAu1a<sub>LOV</sub>* at σ = 1. Black dashed lines indicate predicted hydrogen bonding networks in distances up to 3.2 Å. C) Light-grown FMN and D) Light-grown 5dFMN. Please note that for 5dFMN structure the hydrogen bond networks are identical to the illuminated state.

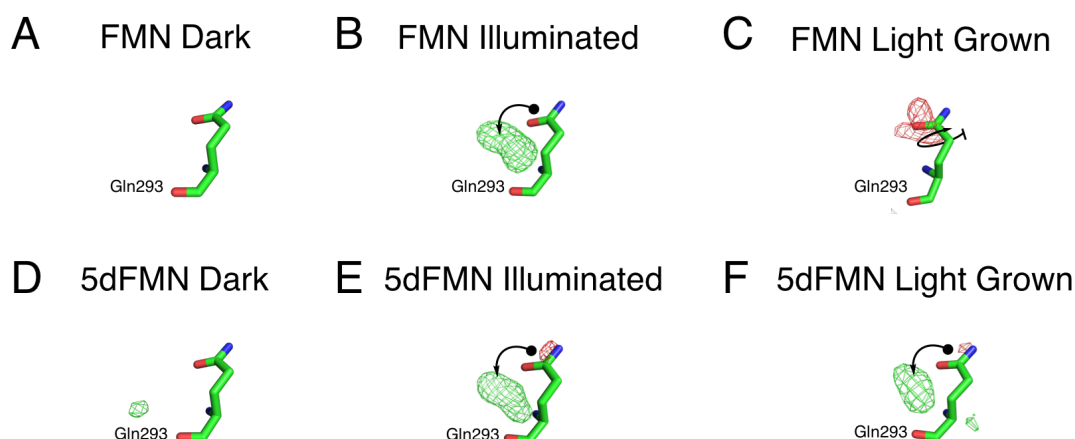
The lit state FMN *OdAu1a<sub>LOV</sub>* structure showed four monomers per asymmetric unit arranged as parallel dimers. Two of the monomers appeared to be identical to the dark state structure (chains A and B), while the second pair contained a different A' $\alpha$  arrangement, Figure 3.14 A (chains C and D). This indicates that the dimer rearrangement occurred through changes in the A' $\alpha$ . Closer inspection of the FMN pocket within the unique dimer showed a different arrangement for residues directly neighbouring the chromophore. Initially, Gln293 was modelled as in the dark state structure. Rounds of refinements showed diffuse electron density map for Asn194, placing the residue in two distinct 'V' shaped poses at 50% individual occupancy. Unlike in the illuminated structure, where Gln293 swings out of the FMN binding pocket, in the lit-state FMN structure Asn194 swings to close proximity to Gln293, Figure 3.14 C. In order to satisfy this arrangement of Asn194 and Gln293 the side chain of Gln293 needs to 'flip' in orientation. Refinement with only the dark-like Gln293 orientation showed negative electron density supporting the additional conformation where Gln293 is 'flipped', Figure 3.15 C. This observation is of a third unique arrangement, where Asn194 swings into the pocket forming a hydrogen bond with 'flipped' Gln293 under lit state conditions.

Light grown 5dFMN *OdAu1a<sub>LOV</sub>* crystals did not show any dimer rearrangement through changes in the A' $\alpha$ . Although the structure could only be refined to 2.0 Å, the lit state structure was almost identical to illuminated 5dFMN structure showing identical Asn194, Gln293 and Asn272 arrangements, Figure 3.14 B and D. As such it appears that 5dFMN can form the 'swing' conformation as observed for FMN illuminated structures but not 'flip' conformation observed in the light grown crystals with FMN. Therefore, it appears that 5dFMN is capable of dictating some small structural changes near the chromophore but is unable to achieve the large-scale dimer rearrangement as observed for the lit state FMN *OdAu1a<sub>LOV</sub>*. Importantly FMN and 5dFMN *OdAu1a<sub>LOV</sub>* were crystallised under identical conditions, which means that structural differences observed were not due to differences in crystallisation conditions such as use of precipitant, pH, additives and salt.

### 3.6 Discussion and Conclusions

Reconstitution of the Au1a light-dependent LOV domain transcription factor with 5dFMN led to generation of a stable cysteinyl-5dFMN photoadduct. The first high-resolution crystal structures of 5dFMN reconstituted *OdAu1a<sub>LOV</sub>* domain confirmed that cysteinyl-5dFMN photoadduct forms at the C4a position as observed for the native FMN. Despite adduct formation, dimerisation was not evident for the 5dFMN-containing *OdAu1a<sub>LOV</sub>* protein upon





**Figure 3.15 Modelling of Gln293 swing and flip states from early refinement.**

All figures show normalized ( $\sigma = 0$ , stdev = 1) positive (green-mesh) and negative (red-mesh) electron density peaks at  $\sigma = 3$  at early stages of refinement prior to introduction of Gln293 swing or flip states, respectively. A) Dark state, B) Illuminated and C) Lit grown states of FMN *OdAu1a<sub>LOV</sub>*, respectively. D) Dark state, E) Illuminated and F) Lit grown states of 5dFMN *OdAu1a<sub>LOV</sub>*, respectively. Figure was adapted from publication<sup>53</sup>.

illumination. EMSA and analytical size exclusion experiments showed that the 5dFMN analogue failed to direct large scale conformational changes needed for DNA binding. This provides support that protonation at the N5 position of isoalloxazine ring system plays an essential role in driving the lit state conformational changes needed for activation or to lock protein in a lit state conformation. On the other hand, CD and to a lesser extent NMR experiments suggested that some structural changes do occur when 5dFMN *OdAu1a<sub>LOV</sub>* is photoactivated. Crystal structures show some changes in residues neighbouring the chromophore in both 5dFMN and FMN proteins upon illumination. Unfortunately, large-scale changes were not evident. Crystal structures obtained here provide further details regarding the mechanism of LOV domain activation. MD of phototropin LOV domains suggested that the conserved Gln residue ‘swings’ out of the FMN binding pocket to facilitate structural rearrangements<sup>37-39</sup>, yet no direct structural observation has been reported. The crystal structures with FMN and 5dFMN directly observe this ‘swing’ conformation. The MD studies also identified hydrogen bond changes within the FMN binding pocket upon photoactivation, especially around the FMN-O4 ring and suggested that these directly regulate the LOV domain  $\beta$ -sheet surface and dynamics.

The most notable observations in this study were changes in the dimer arrangement between dark, illuminated and lit state structures of FMN containing *OdAu1a<sub>LOV</sub>*. Superimposition of chain A of dark state structure (PDB 6I20) with illuminated structure chain A (PDB 6I21) yields a value of 0.471 Å (814 of 814 atoms) suggesting

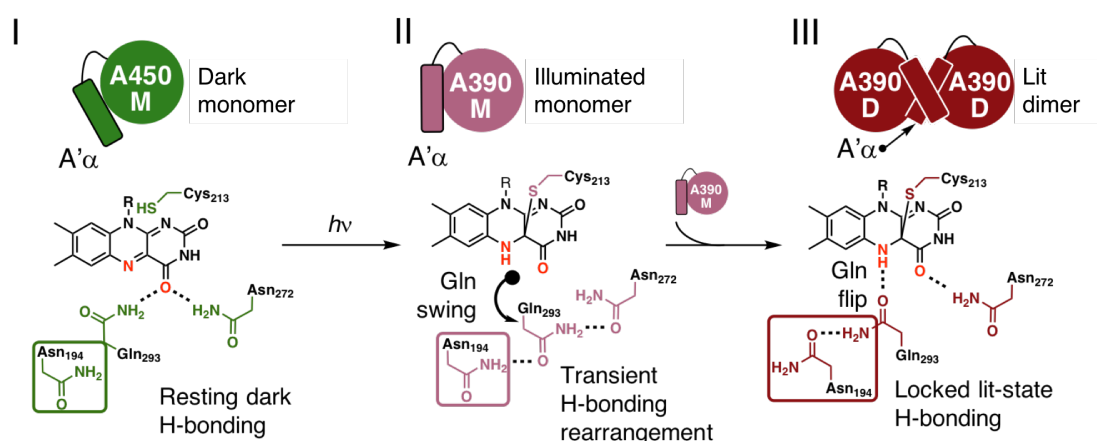
some small structural differences. Comparison of dark state FMN chain A with lit state chain C (unique dimer, PDB 6I22) gives RMSD a value of 0.367 Å (769 to 769 atoms), a value smaller than observed when comparing dark and illuminated structures. Nonetheless, significant changes in the dimer arrangement through N-terminal A'α is evident. Comparison of FMN illuminated and FMN lit chain C structures provides RMSD value of 0.531 Å (819 to 819 atoms) suggesting the largest structural difference. Considering lower than 1.2 Å RMSD values, it was decided to investigate thermodynamic parameters of potential dimers using PDBePISA software (PDBe) provided by Protein Data Bank ([www.ebi.ac.uk/PDBe](http://www.ebi.ac.uk/PDBe)). Dark state FMN structure indicated four monomers per asymmetric unit, chains A and C plus chains B and D forming parallel dimers, respectively. Engaged interfaces for chains A and C resulted in calculated  $\Delta^iG = -13.2$  kcal/mol and for chains B and D  $\Delta^iG = -15.1$  kcal/mol suggesting potential dimers in the solution. For illuminated FMN structure  $\Delta^iG = -13.2$  kcal/mol was also evident when symmetry mate was considered, with a dimer arrangement similar to that observed in the dark state. Lit state FMN structure indicated four monomers per asymmetric unit where dark-like chains A and B resulted in a dimeric arrangement with  $\Delta^iG = -12.6$  kcal/mol, while unique dimer formed between chains C and D with rearranged N-terminal A'α resulted in  $\Delta^iG = -13.2$  kcal/mol suggesting a potential dimer. Interestingly, a possible dimer arrangement (docking) between dark-like chains A or B with unique dimer chain C or D results in much weaker  $\Delta^iG$ . For instance a formation of docked interactions between chain A and chain C results in  $\Delta^iG = -7.9$  kcal/mol, a significantly smaller value, suggesting that individual chains are not exchangeable and are locked in a conformation allowing dimerisation.

Comparison of dark (PDB 6I23) and illuminated (PDB 6I24) structures of 5dFMN *OdAu1a<sub>LOV</sub>* yields RMSD 0.376 Å (839 to 839 atoms), while dark and lit (PDB 6I25) state 5dFMN structures yields RMSD of 0.364 Å (804 to 804 atoms). As expected, alignment of illuminated and lit state structures of 5dFMN *OdAu1a<sub>LOV</sub>* gives RMSD value of 0.129 Å (859 to 859 atoms) suggesting almost identical structures. Considering PISA analysis, dark state 5dFMN structure indicated  $\Delta^iG = -12.3$  kcal/mol, where illuminated and lit state structures  $\Delta^iG \sim -13$  kcal/mol suggesting potential dimers in the solution in dark or lit states.

When FMN and 5dFMN *OdAu1a<sub>LOV</sub>* structures were compared, dark state structures using chains A yielded RMSD values of 0.271 Å (782 to 782 atoms) indicating very closely related structures. When illuminated structures were investigated, RMSD values of 0.056 Å (932 to 932 atoms) was observed indicating almost an identical residue arrangements. Lastly, comparison of lit state structures (chain C for FMN and chain A for

5dFMN) resulted in RMSD of 0.561 (807 to 807 atoms) indicating a degree of difference between two structures. This therefore highlights a fact that 5dFMN protein failed to achieve the lit state FMN conformation. Nonetheless, these results have to be interpreted with caution, as during crystallisation experiment high concentration of protein and crystal packing can result in artefacts such as forced dimers that are not present in the solution.

Taking these studies and considering a previous Au1a<sub>LOV</sub> homologue lit state structure<sup>34</sup>, it is possible to propose a stepwise mechanism, Figure 3.16, explaining the intermediate photoresponsiveness of 5dFMN. In the resting dark state O4 of FMN and 5dFMN hydrogen bond to the sidechain NH<sub>2</sub> of Gln293 and Asn272. Upon illumination, interruption of the conjugated isoalloxazine ring systems detaches Asn272 and Gln293. This weakening in the hydrogen bonds allows a possible transient Gln293 swing out of the FMN binding pocket forming a hydrogen bond with an allosteric Asn194 creating an intermediate hydrogen bond network between Asn194, Gln293 and Asn272. Considering that this orientation was observed for 5dFMN containing protein therefore explains partial light responsiveness. After recruitment of Asn194, the formation of the dimer across possible A'α helices for *OdAu1a<sub>LOV</sub>* is stabilised by yet another conformation where Asn194 forms a hydrogen bond with a 'flipped' and repositioned Gln293 locking the lit-state conformation. 5dFMN is unable to act as hydrogen bond donor at the C5, therefore is unable to lock the lit state dimer.



**Figure 3.16 Proposed structural mechanism of Au1a<sub>LOV</sub> activation by light.**

In the resting state, Gln293 and Asn272 form hydrogen bonds with the FMN-O4 locking protein in a monomeric state (State I). Upon the photoadduct formation, a planar aromatic isoalloxazine ring system is interrupted unanchoring Gln293 and Asn272 weakening of hydrogen bonds to the FMN-O4. This therefore allows Gln293 residue to swing out of the flavin binding pocket forming transient hydrogen bond network with Asn194 and Asn272 (State II). Upon dimerisation, the dimer is stabilised by Asn194 and Gln293 hydrogen bond and possible Gln293 side chain oxygen to protonated FMN-N5 (State III). In the figure, allosteric Asn194 residue, located in a latch region

between A' $\alpha$  and  $\beta$ A strand is indicated in a box. Possible hydrogen bond networks are indicated in the figure as dashed lines.

Additionally, Gln flip stabilised by Asn residue was previously observed in a *PtAu1a<sub>LOV</sub>* homologue crystalized under blue light irradiation (PDB code 5DKL)<sup>34</sup>. It is therefore possible to hypothesise that this unique residue arrangement could be important for Au1a signalling. Furthermore, this Asn residue position, where the residue is located between A' $\alpha$  and A $\beta$ , is also found in *AsLOV2*, where irradiation with light has been shown to result in the unfolding of A' $\alpha$  followed by the unfolding of J $\alpha$  helix, respectively<sup>56</sup>.

Previous studies suggested that a conserved Gln residue senses protonation state of FMN-N5 and dictates structural changes<sup>33</sup>. Site directed mutagenesis indicated that this key Gln was vital for functional light responsiveness, but the results were difficult to interpret<sup>28,32</sup>. Studies on *AsLOV2* where Gln was replaced by Asn resulted in a partially active protein as determined by CD and NMR experiments. Furthermore, replacement of Gln with an aliphatic Leu residue unable to hydrogen bonds with either FMN-O4 or protonated FMN-N5 also resulted in a partially light responsive protein suggesting that Gln is not an absolute requirement<sup>32</sup>. Considering our observations with 5dFMN reconstituted *OdAu1a<sub>LOV</sub>*, it is possible that the lack of hydrogen bond acceptor or donor at the fifth position creates pseudo light responsive protein similar to the *AsLOV2* mutants. This therefore suggests that protonation of FMN-N5 is not the only requirement to trigger observable structural changes and other factors such as hydrogen bonding to FMN-O4 and possibly puckering or interruption of isoalloxazine ring system contributing to an extent. Structural studies of *ZEITLUPE* and *PpBS1<sub>LOV</sub>* indicating changes extending from FMN-O4 rather than FMN-N5 suggesting that in some LOV proteins to FMN-O4 extending hydrogen bonding and light dependent changes could contribute more<sup>40,43</sup>. This therefore asks a question if 5dFMN would be more functionally responsive in these proteins in contrast to *OdAu1a<sub>LOV</sub>*.

In summary, our results are in agreement with previous MD simulations and Fourier transform infrared spectroscopy (FT-IR) experiments that identified hydrogen bond changes to the FMN-O4 being important for the regulation of LOV domain activation especially in the early stages post photoadduct formation<sup>37-39,41,42</sup>. Intermediate structural changes observed in illuminated FMN and 5dFMN *OdAu1a<sub>LOV</sub>* therefore suggest that changes in hydrogen bonds extending from FMN-O4 contribute towards achieving the lit state conformation. Reflecting on the applicability of LOV domains in the development of new optogenetic systems, understanding of these fundamental aspects of dark to lit state switching proves scaffold to assist with the rational design of LOV domain-controlled proteins.

### 3.7 References

1. Crosson, S., Rajagopal, S. & Moffat, K. The LOV Domain Family: Photoresponsive Signaling Modules Coupled to Diverse Output Domains. *Biochemistry* **42**, 2-10 (2003).
2. Ito, S., Song, Y. H. & Imaizumi, T. LOV domain-Containing F-box Proteins: Light-Dependent Protein Degradation Modules in Arabidopsis. *Molecular Plant* **5**, 573-582 (2012).
3. Krauss U, Minh, B. Q., Losi, A., Gärtner, W., Eggert, T., von Haeseler, A. & Jaeger, K. E. Distribution and Phylogeny of Light-Oxygen-Voltage-Blue-Light-Signaling Proteins in the Three Kingdoms of Life. *J. Bacteriol.* **191**, 7234-7242 (2009).
4. Glantz, S. T., Carpenter, E. J., Melkonian, M., Gardner, K. H., Boyden, E. S., Wong, G. K. & Chow, B. Y. Functional and Topological Diversity of LOV Domain Photoreceptors. *PNAS* **113**, E1442-E1451 (2016).
5. Bonomi, H. R., Posadas, D. M., Paris, G., Carrica Mdel, C., Frederickson, M., Pietrasanta, L. I., Bogomolni, R. A., Zorreguieta, A. & Goldbaum, F. A. Light Regulates Attachment, Exopolysaccharide Production, and Nodulation in *Rhizobium leguminosarum* Through a LOV-Histidine Kinase Photoreceptor. *PNAS* **109**, 12135-12140 (2012).
6. Chen, C. H., DeMay, B. S., Gladfelter, A. S., Dunlap, J. C. & Loros, J. J. Physical Interaction Between VIVID and White Collar Complex Regulates Photoadaptation in *Neurospora*. *PNAS* **107**, 16715-16720 (2010).
7. Möglich, A., Ayers, R. A. & Moffat, K. Structure and Signaling Mechanism of Per-ARNT-Sim Domains. *Structure* **17**, 1282-1294 (2009).
8. Cao, Z., Buttani, V., Losi, A. & Gärtner, W. A Blue Light Inducible Two-Component Signal Transduction System in the Plant Pathogen *Pseudomonas syringae* pv. tomato. *Biophys. J.* **94**, 897-905 (2008).
9. Purcell, E. B., Siegal-Gaskins, D., Rawling, D. C., Fiebig, A. & Crosson, S. A Photosensory Two-Component System Regulates Bacterial Cell Attachment. *PNAS* **104**, 18241-18246 (2007).
10. Herrou, J. & Crosson, S. Function, Structure and Mechanism of Bacterial Photosensory LOV Proteins. *Nat. Rev. Microbiol.* **9**, 713-723 (2011).
11. Christie, J. M., Gawthorne, J., Young, G., Fraser, N. J. & Roe, A. J. LOV to BLUF: Flavoprotein Contributions to the Optogenetic Toolkit. *Mol. Plant* **5**, 533-544

- (2012).
12. Strickland, D., Moffat, K. & Sosnick, T. R. Light-Activated DNA Binding in a Designed Allosteric Protein. *PNAS* **105**, 10709-10714 (2008).
  13. Mart, R. J., Meah, D. & Allemann, R. K. Photocontrolled Exposure of Pro-apoptotic Peptide Sequences in LOV Proteins Modulates Bcl-2 Family Interactions. *ChemBioChem* **17**, 698-701 (2016).
  14. Niopek, D., Wehler, P., Roensch, J., Eils, R. & Di Ventura, B. Optogenetic Control of Nuclear Protein Export. *Nat. Commun.* **7**, 1-9 (2016).
  15. Renicke, C., Schuster, D., Usherenko, S., Essen, L. O. & Taxis, C. A LOV2 Domain-Based Optogenetic Tool to Control Protein Degradation and Cellular Function. *Chem. Biol.* **20**, 619-626 (2013).
  16. Lee, J., Natarajan, M., Nashine, V. C., Socolich, M., Vo, T., Russ, W. P., Benkovic, S. J. & Ranganathan, R. Surface Sites for Engineering Allosteric Control in Proteins. *Science* **322**, 438-442 (2008).
  17. Strickland, D., Yao, X., Gawlak, G., Rosen, M. K., Gardner, K. H. & Sosnick, T. R. Rationally Improving LOV Domain Based Photoswitches. *Nat. Methods* **7**, 623-626 (2010).
  18. Gehrig, S., Macpherson, J. A., Driscoll, P. C., Symon, A., Martin, S. R., MacRae, J. I., Kleinjung, J., Fraternali, F. & Anastasiou, D. An Engineered Photoswitchable Mammalian Pyruvate Kinase. *FEBS J.* **284**, 2955-2980 (2017).
  19. Kawano, F., Suzuki, H., Furuya, A. & Sato, M. Engineered Pairs of Distinct Photoswitches for Optogenetic Control of Cellular Proteins. *Nat. Commun.* **6**, 1-8 (2015).
  20. Crosson, S. & Moffat, K. Structure of a Flavin-Binding Plant Photoreceptor Domain: Insights into Light-Mediated Signal Transduction. *PNAS* **98**, 2995-3000 (2001).
  21. Herman, E., Sachse, M., Kroth, P. G. & Kottke, T. Blue-Light-Induced Unfolding of the Jalpha Helix Allows for the Dimerization of Aureochrome-LOV from the diatom *Phaeodactylum tricornutum*. *Biochemistry* **52**, 3094-3101 (2013).
  22. Herman, E. & Kottke, T. Allosterically Regulated Unfolding of the A'alpha Helix Exposes the Dimerization Site of the Blue-Light-Sensing Aureochrome-lov domain. *Biochemistry* **54**, 1484-1492 (2015).
  23. Harper, S. M., Christie, J. M. & Gardner, K. H. Disruption of the LOV-Jalpha Helix Interaction Activates Phototropin Kinase Activity. *Biochemistry* **43**, 16184-16192 (2004).

24. Harper, S. M., Neil, L. C. & Gardner, K. H. Structural Basis of a Phototropin Light Switch. *Science* **301**, 1541-1544 (2003).
25. Möglich, A. & Moffat, K. Structural Basis for Light-Dependent Signaling in the Dimeric LOV Domain of the Photosensor YtvA. *J. Mol. Biol.* **373**, 112-126 (2007).
26. Berntsson, O. *et al.* Time-Resolved X-Ray Solution Scattering Reveals the Structural Photoactivation of a Light-Oxygen-Voltage Photoreceptor. *Structure* **25**, 933–938 (2017).
27. Engelhard, C., Diensthuber, R. P., Möglich, A. & Bittl, R. Blue-Light Reception Through Quaternary Transitions. *Sci. Rep.* **7**, 1385 (2017).
28. Zoltowski, B. D., Schwerdtfeger, C., Widom, J., Loros, J. J., Bilwes, A. M., Dunlap, J. C. & Crane B. R. Conformational Switching in the Fungal Light Sensor Vivid. *Science* **316**, 1054-1057 (2007).
29. Vaidya, A. T., Chen, C. H., Dunlap, J. C., Loros, J. J. & Crane, B. R. Structure of a Light-Activated LOV Protein Dimer that Regulates Transcription. *Sci. Signal.* **4**, ra50 (2011).
30. Diensthuber, R. P., Bommer, M., Gleichmann, T. & Möglich, A. Full-Length Structure of a Sensor Histidine Kinase Pinpoints Coaxial Coiled Coils as Signal Transducers and Modulators. *Structure* **21**, 1127-1136 (2013).
31. Jones, M. A., Feeney, K. A., Kelly, S. M. & Christie, J. M. Mutational Analysis of Phototropin 1 Provides Insights Into the Mechanism Underlying LOV2 Signal Transmission. *J. Biol. Chem.* **282**, 6405-6414 (2007).
32. Nash, A. I., Ko, W. H., Harper, S. M. & Gardner, K. H. A Conserved Glutamine Plays a Central Role in LOV Domain Signal Transmission and Its Duration. *Biochemistry* **47**, 13842-13849 (2008).
33. Ganguly, A., Thiel, W. & Crane, B. R. Glutamine Amide Flip Elicits Long Distance Allosteric Responses in the LOV Protein Vivid. *J. Am. Chem. Soc.* **139**, 2972-2980 (2017).
34. Heintz, U. & Schlichting, I. Blue Light-Induced LOV Domain Dimerization Enhances the Affinity of Aureochrome 1a for Its Target DNA Sequence. *Elife* **5**, e11860 (2016).
35. Yee, E. F., Diensthuber, R. P., Vaidya, A. T., Borbat, P. P., Engelhard, C., Freed, J. H., Bittl, R., Möglich, A. & Crane, B. R. Signal transduction in light-oxygen-voltage receptors lacking the adduct-forming cysteine residue. *Nat. Commun.* **6**, 10079 (2015).
36. Tsukuno, H., Ozeki, K., Kobayashi, I., Hisatomi, O. & Mino, H. Flavin-Radical

- Formation in the Light-Oxygen-Voltage-Sensing Domain of the Photozipper Blue-light Sensor Protein. *J. Phys. Chem. B* **122**, 8819-8823 (2018).
37. Freddolino, P. L., Dittrich, M. & Schulten, K. Dynamic Switching Mechanisms in LOV1 and LOV2 Domains of Plant Phototropins. *Biophys. J.* **91**, 3630-3639 (2006).
  38. Peter, E., Dick, B. & Baeurle, S. A. Mechanism of Signal Transduction of the LOV2-J $\alpha$  Photosensor from *Avena sativa*. *Nat. Commun.* **1**, 122–127 (2010).
  39. Freddolino, P. L., Gardner, K. H. & Schulten, K. Signaling Mechanisms of LOV Domains: New Insights from Molecular Dynamics Studies. *Photochem. Photobiol. Sci.* **12**, 1158-1170 (2013).
  40. Pudasaini, A., Shim, J. S., Song, Y. H., Shi, H., Kiba, T., Somers, D. E., Imaizumi, T. & Zoltowski, B. D. Kinetics of the LOV Domain of ZEITLUPE Determine Its Circadian Function in *Arabidopsis*. *Elife* **6**, 1-27 (2017).
  41. Gil, A. A. *et al.* Femtosecond to Millisecond Dynamics of Light Induced Allostery in the *Avena sativa* LOV Domain. *J Phys Chem B.* **121**, 1010-1019 (2017).
  42. Zayner, J. P., Antoniou, C. & Sosnick, T. R. The Amino-Terminal Helix Modulates Light-Activated Conformational Changes in AsLOV2. *J. Mol. Biol.* **419**, 61-74 (2012).
  43. Röllén, K. *et al.* Signaling States of a Short Blue-Light Photoreceptor Protein PpSB1-LOV Revealed from Crystal Structures and Solution NMR Spectroscopy. *J. Mol. Biol.* **428**, 3721-3736 (2016).
  44. Mansurova, M., Scheercousse, P., Simon, J., Kluth, M. & Gärtner, W. Chromophore Exchange in the Blue Light-Sensitive Photoreceptor YtvA from *Bacillus subtilis*. *ChemBioChem* **12**, 641-646 (2011).
  45. Fisher, J., Spencer, R. & Walsh, C. Enzyme-Catalyzed Redox Reactions with the Flavin Analogues 5-Deazariboflavin, 5-Deazariboflavin 5'-Phosphate, and 5-Deazariboflavin 5'-Diphosphate, 5'  $\rightarrow$  5'-Adenosine Ester. *Biochemistry* **15**, 1054-1064 (1976).
  46. Hemmerich, P., Massey, V. & Fenner, H. Flavin and 5-deazaflavin: A Chemical Evaluation of 'Modified' Flavoproteins with Respect to the Mechanisms of Redox Biocatalysis. *FEBS L.* **84**, 5-21 (1977).
  47. Hedison, T. M., Leferink, N. G. H., Hay, S. & Scrutton, N. S. Correlating Calmodulin Landscapes with Chemical Catalysis in Neuronal Nitric Oxide Synthase Using Time-Resolved FRET and a 5-deazaflavin Thermodynamic Trap. *ACS Catal.* **6**, 5170-5180 (2016).



48. Takahashi, F., Yamagata, D., Ishikawa, M., Fukamatsu, Y., Ogura, Y., Kasahara, M., Kiyosue, T., Kikuyama, M., Wada, M. & Kataoka, H. AUREOCHROME, a Photoreceptor Required for Photomorphogenesis in Stramenopiles. *PNAS* **104**, 19625-19630 (2007).
49. Banerjee, A., Herman, E., Serif, M., Maestre-Reyna, M., Hepp, S., Pokorny, R., Kroth, P. G., Essen, L. O. & Kottke, T. Allosteric Communication Between DNA-Binding and Light-Responsive Domains of Diatom Class i Aureochromes. *Nucleic Acids Res.* **44**, 5957-5970 (2016).
50. Nash, A. I., McNulty, R., Shillito, M. E., Swartz, T. E., Bogomolni, R. A., Luecke, H. & Gardner, K. H. Structural Basis of Photosensitivity in a Bacterial Light-Oxygen-Voltage/Helix-Turn-Helix (LOV-HTH) DNA-Binding Protein. *PNAS* **108**, 9449-54 (2011).
51. Zoltowski, B. D., Motta-Mena, L. B. & Gardner, K. H. Blue Light-Induced Dimerization of a Bacterial LOV-HTH DNA-Binding Protein. *Biochemistry* **52**, 6653-6661 (2013).
52. Akiyama, Y., Nakasone, Y., Nakatani, Y., Hisatomi, O. & Terazima, M. Time-Resolved Detection of Light-Induced Dimerization of Monomeric Aureochrome-1 and Change in Affinity for DNA. *J. Phys. Chem. B* **120**, 7360–7370 (2016).
53. Kalvaitis, M. E., Johnson, L. A., Mart, R. J., Rizkallah, P. & Allemann, R. K. A Noncanonical Chromophore Reveals Structural Rearrangements of the Light-Oxygen-Voltage Domain upon Photoactivation. *Biochemistry* **58**, 2608-2616 (2019).
54. Banerjee, A., Herman, E., Kottke, T. & Essen, L. O. Structure of a Native-like Aureochrome 1a LOV Domain Dimer from *Phaeodactylum tricornutum*. *Structure* **24**, 171-178 (2016).
55. Mitra, D., Yang, X. & Moffat, K. Crystal structures of aureochrome1 LOV suggest new design strategies for optogenetics. *Structure* **20**, 698-706 (2012).
56. Halavaty, A. S. & Moffat, K. N- and C-terminal Flanking Regions Modulate Light-Induced Signal Transduction in the LOV2 Domain of the Blue Light Sensor Phototropin 1 from *Avena sativa*. *Biochemistry* **46**, 14001-14009 (2007)

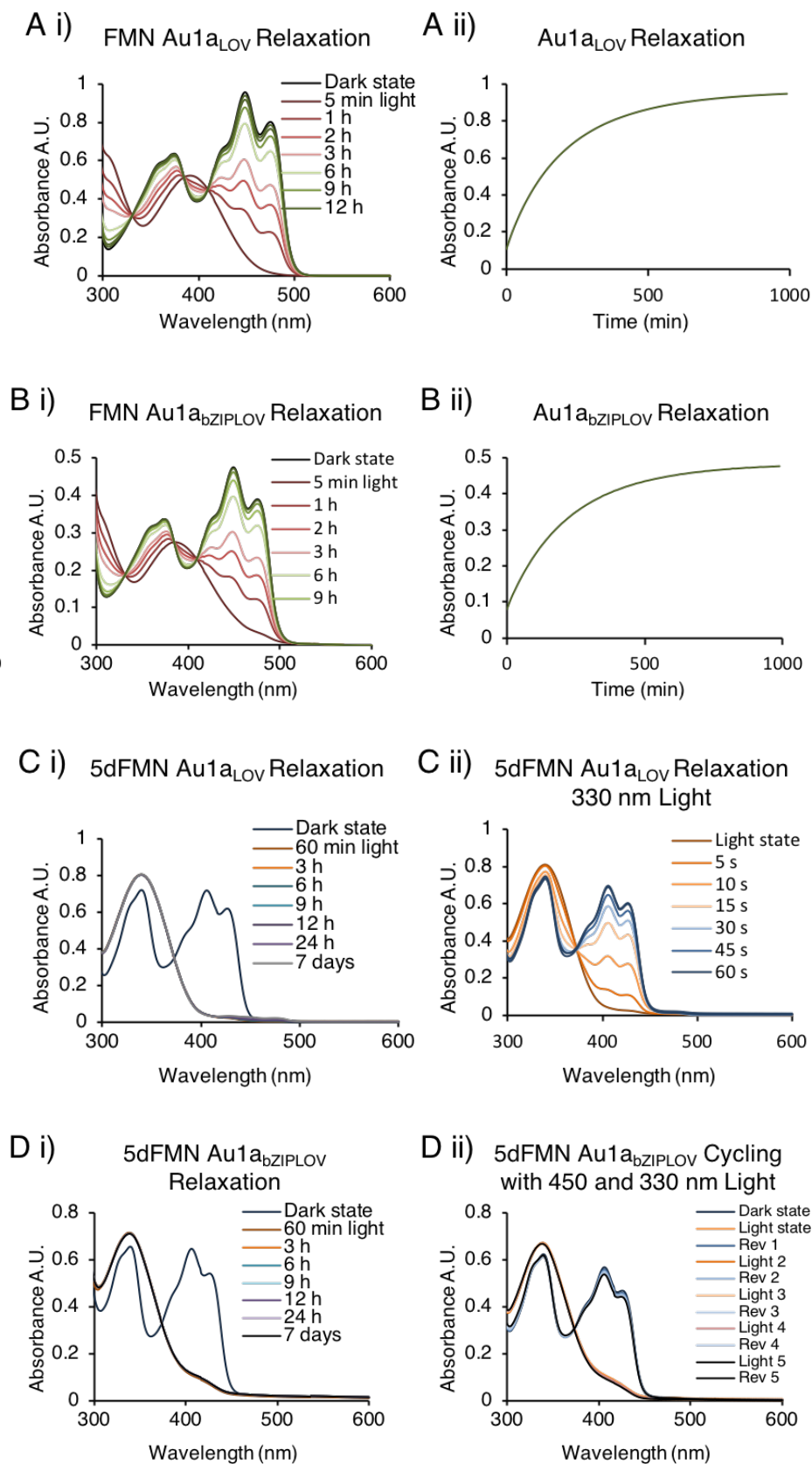
### 3.8 Appendix

**Table 3.1 Data sets collected for FMN-containing *OdAu1a<sub>LOV</sub>*.**

PDB Entry	FMN Dark	FMN Illuminated	FMN Light
<b>Data Collection</b>			
DLS Visit	mx14843-7	mx14843-5	mx14843-15
Date	10-08-2017	12-05-2017	29-03-2018
Diamond Beamline	I24	I04-1	I03
Wavelength Å	0.96858	0.97950	0.97625
Data Set No	156_4	186_1	162_2
Accession Code	6I20	6I21	6I22
pH	4.5 to 4.9	4.5 to 4.9	7.5 to 8.0
Crystallisation Condition	0.1 M sodium acetate, 0.1 M ammonium chloride 15-20% PEG 2000	0.1 M sodium acetate, 0.1 M ammonium chloride I 15-20% PEG 2000	1.5-3 M sodium malonate, 0.1 M Tris.acetate
<b>Crystal Data</b>			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	65.68, 104.04, 105.94	106.39, 106.39, 67.49	150.18, 91.51, 99.84
<i>h</i> , <i>k</i> , <i>l</i>	90.0, 90.0, 90.0	90.0, 90.0, 120.0	90.0, 131.34, 90.0
Space group	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P 6 <sub>4</sub> 2 2	C 1 2 1
Resolution (Å)	1.37 – 55.54	1.50 – 92.14	1.66 – 71.05
Outer shell	1.37 – 1.41	1.50 – 1.54	1.66 – 1.70
<i>R</i> -merge (%)	10.9 (127.8)	7.2 (266.0)	8.3 (102.9)
<i>R</i> -meas (%)	11.2 (137.7)	7.4 (272.5)	8.9 (120.1)
CC1/2	0.997 (0.546)	1.000 (0.644)	0.997 (0.543)
<i>I</i> / $\sigma(I)$	15.2 (1.3)	18.2 (1.3)	12.0 (0.9)
Completeness (%)	99.9 (98.7)	100.0 (100)	98.0 (99.3)
Multiplicity	18.7 (6.8)	20.7 (21.3)	6.1 (3.7)
Total Measurements	2,858,037 (35,281)	754,935 (37,382)	714,461 (32,079)
Unique Reflections	152,478 (7,382)	36,515 (2,637)	117,045 (8,720)
Wilson B-factor(Å <sup>2</sup> )	16.3	22.2	26.7
<b>Refinement Statistics</b>			
Non-H Atoms	5,107	1,343	4,747
R-work reflections	144,858	34,209	111,292
R-free reflections	7,522	1,886	5,750
R-work/R-free	0.151 / 0.175	0.173 / 0.205	0.203 / 0.225
<b>rms deviations</b>			
Bond lengths (Å)	0.013	0.016	0.014
Bond Angles (°)	1.659	2.006	1.717
<sup>1</sup> Coordinate error	0.033	0.056	0.073
Mean B value (Å <sup>2</sup> )	23.9	26.4	37.5
<b>Ramachandran Statistics</b>			
Favoured/Allowed/Outliers	477 / 5 / 0	118 / 1 / 0	499 / 11 / 0
%	99.0 / 1.0 / 0.0	99.2 / 0.8 / 0.0	97.8 / 2.2 / 0.0

**Table 3.2 Data sets for 5dFMN-containing *OdAu1a<sub>Low</sub>*.**

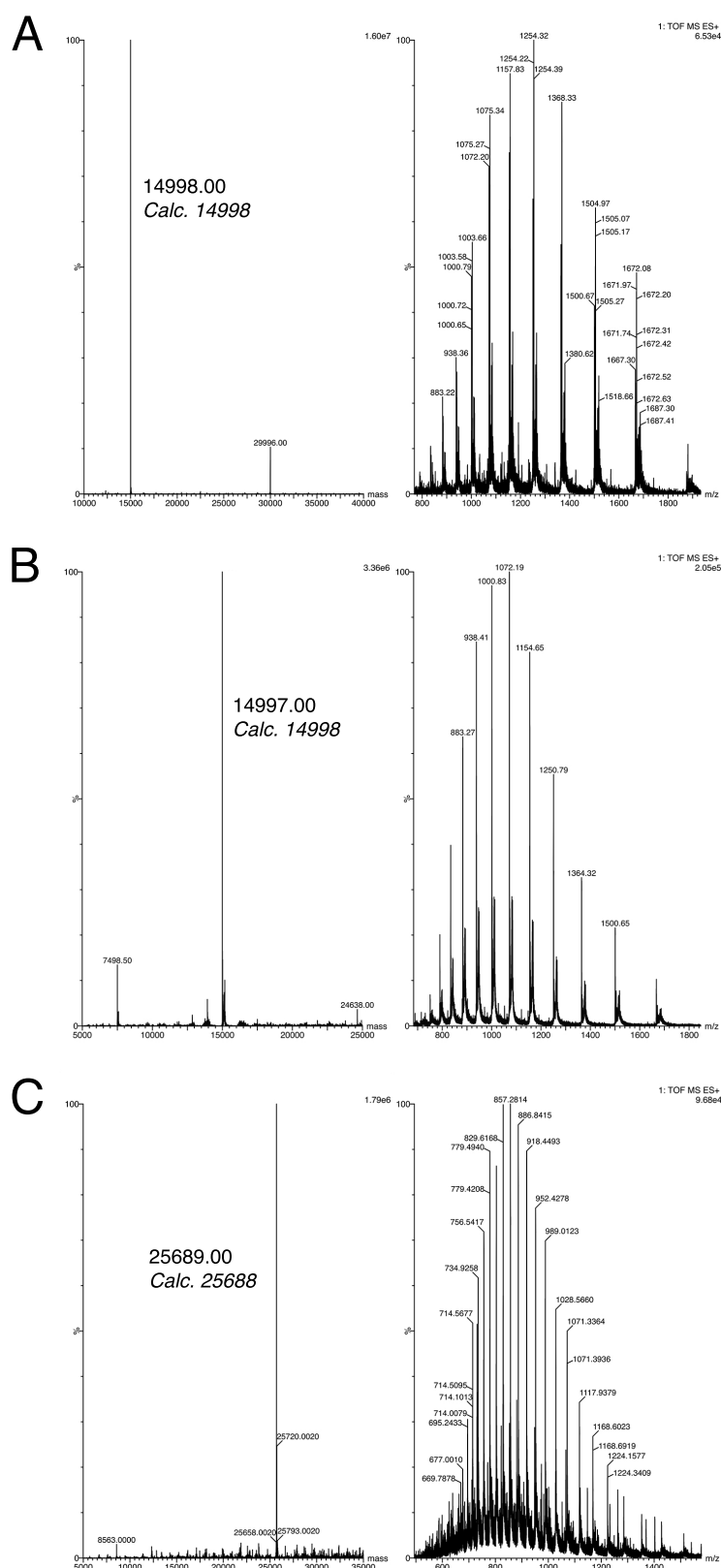
PDB Entry	5dFMN Dark	5dFMN Illuminated	5dFMN Light
<b>Data Collection</b>			
DLS Visit	mx14843-5	mx14843-5	mx14843-5
Date	12-05-2017	12-05-2017	12-05-2017
Diamond Beamline	I04-1	I04-1	I04-1
Wavelength	0.97950	0.97950	0.97950
Data Set No	156_1	183_1	154_1
Accession Code	6I23	6I24	6I25
pH	4.5 to 4.9	4.5 to 4.9	7.5 to 8.0
Crystallisation Condition	0.1 M sodium acetate, 0.1 M ammonium chloride 15-20% PEG 2000	0.1 M sodium acetate, 0.1 M ammonium chloride 15-20% PEG 2000	1.5-3 M sodium malonate, 0.1 M Tris.acetate
<b>Crystal Data</b>			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	104.56, 104.56, 68.01	106.33, 106.33, 67.49	106.29, 106.29, 67.86
<i>h</i> , <i>k</i> , <i>l</i>	90.0, 90.0, 120.0	90.0, 90.0, 120.0	90.0, 90.0, 120.0
Space group	P 3 <sub>1</sub> 2 1	P 6 <sub>4</sub> 2 2	P 6 <sub>4</sub> 2 2
Resolution (Å)	1.97 – 52.28	1.43 – 67.49	1.97 – 54.62
Outer shell	1.97 – 2.02	1.43 – 1.47	1.97 – 2.02
<i>R</i> -merge (%)	7.0 (145.1)	8.7 (217.1)	39.0 (308.5)
<i>R</i> -meas (%)	7.3 (152.1)	8.9 (222.9)	40.0 (31.61)
CC1/2	0.999 (0.839)	0.999 (0.741)	0.995 (0.592)
<i>I</i> / $\sigma(I)$	18.2 (1.8)	16.1 (1.1)	7.9 (1.4)
Completeness (%)	99.7 (99.7)	100.0 (100)	100.0 (100.0)
Multiplicity	11.1 (11.2)	20.6 (19.7)	20.9 (20.9)
Total Measurements	337,877 (25,005)	867,173 (41,996)	345,367 (24,640)
Unique Reflections	29,340 (2.235)	41,996 (3,051)	16,503 (1,181)
Wilson B-factor(Å <sup>2</sup> )	36.3	17.0	17.5
<b>Refinement Statistics</b>			
Non-H Atoms	2,297	1,401	1,311
R-work reflections	27,817	39,844	15,792
R-free reflections	1,501	2,113	674
R-work/R-free	0.187 / 0.220	0.151 / 0.178	0.168 / 0.200
<b>rms deviations</b>			
Bond lengths (Å)	0.014	0.015	0.014
Bond Angles (°)	1.717	1.931	1.697
<sup>1</sup> Coordinate error	0.139	0.035	0.119
Mean B value (Å <sup>2</sup> )	48.8	25.2	23.9
<b>Ramachandran Statistics</b>			
Favoured/allowed/Outliers	247 / 5 / 0	110 / 1 / 0	123 / 1 / 0
%	98.0 / 2.0 / 0	99.1 / 0.9 / 0	99.2 / 0.8 / 0



**Figure 3.17 UV-visible spectra of FMN and 5dFMN *OdaU1a* constructs.**

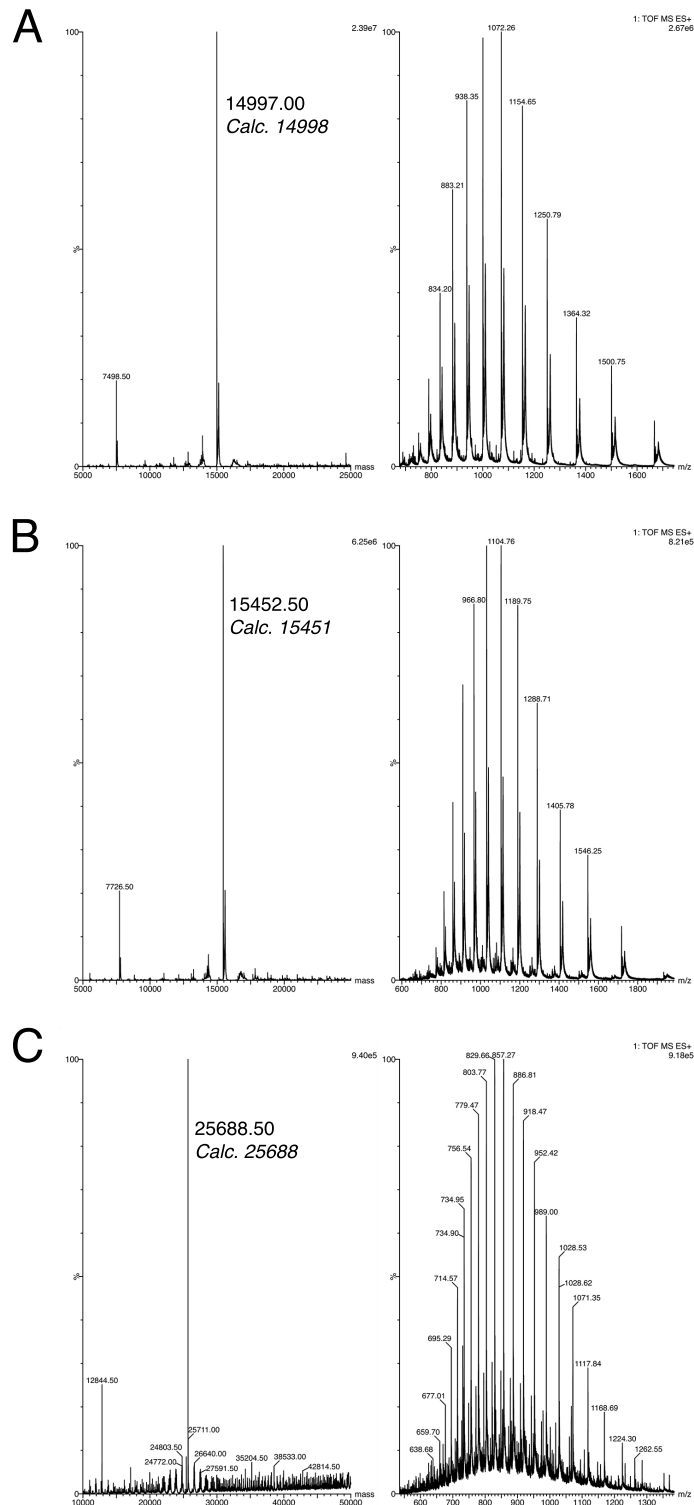
All experiments were performed in DNA binding buffer, 20 mM HEPES, 100 mM NaCl, 10 mM MgCl<sub>2</sub>, 1.0 mM TCEP, pH 7.4, at 20 °C. A) FMN *OdaU1a<sub>LOV</sub>* and B) FMN *OdaU1a<sub>bZIPLOV</sub>* reversion experiments. Individual scan times are depicted as coloured lines and are accompanied by a trace

indicating the absorbance at 447 nm over time. The half-life of *OdAu1a<sub>LOV</sub>* is 112 min. *OdAu1a<sub>bZIPLOV</sub>* also as an apparent half-life of 112 min, but does not fit well to the first order kinetics, possibly due to dimerisation effects. Spectra of C) 5dFMN *OdAu1a<sub>LOV</sub>* and D) 5dFMN *OdAu1a<sub>bZIPLOV</sub>* were recorded every hour for 24 hours. After 24 h, protein samples and blank buffer were stored in the dark at 4 °C for 7 days and an additional spectrum was collected. E) Photoreversion of 5dFMN *OdAu1a<sub>LOV</sub>* using 330 nm light, where coloured lines indicate exposure time. 60 seconds was sufficient to completely revert the photoadduct. F) Cycling of 5dFMN *OdAu1a<sub>bZIPLOV</sub>* from lit to dark adapted states using 450 nm light to photoactivate the protein and 330 nm light to revert protein to dark adapted state. After five cycles only 5.6 % of the original absorbance at 406 nm was lost. The figure was adapted from publication<sup>53</sup>.



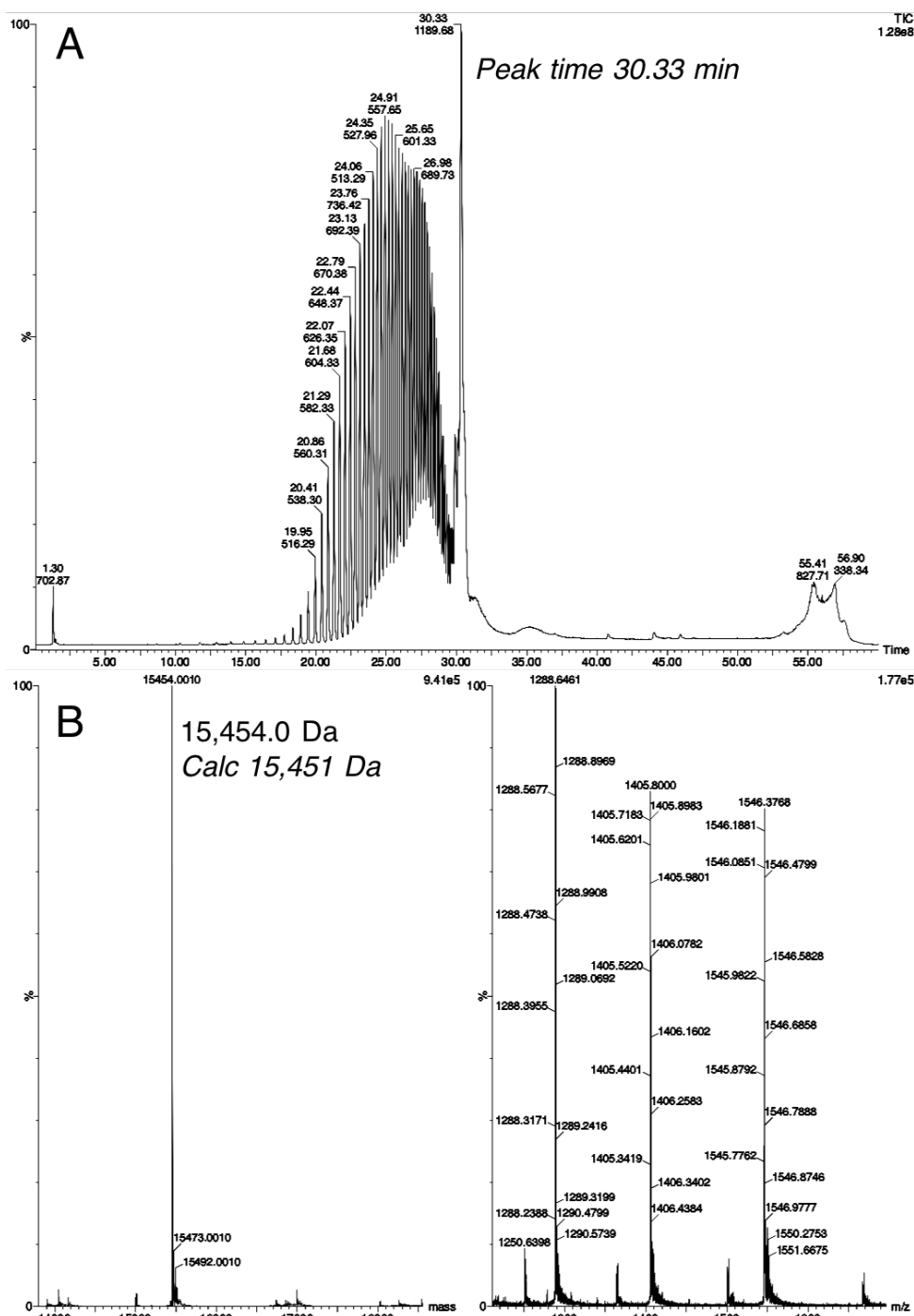
**Figure 3.18 Mass spectra of FMN *OdAu1a<sub>LOV</sub>* and *OdAu1a<sub>bZIPLOV</sub>*.**

Deconvoluted mass spectra of A) Dark state and B) Illuminated *OdAu1a<sub>LOV</sub>* indicating species with a molecular mass of 14,997.0 AMU, consistent with apoprotein. C) Deconvoluted mass spectra of dark state FMN *OdAu1a<sub>bZIPLOV</sub>* indicating a single species with molecular mass of 25,689.0 AMU, consistent with apoprotein. Please note error of MS was 1 atomic mass units (AMU).



**Figure 3.19 Mass spectra of 5dFMN incorporated *OdAu1a<sub>LOV</sub>* and *OdAu1a<sub>ZIPLOV</sub>*.**

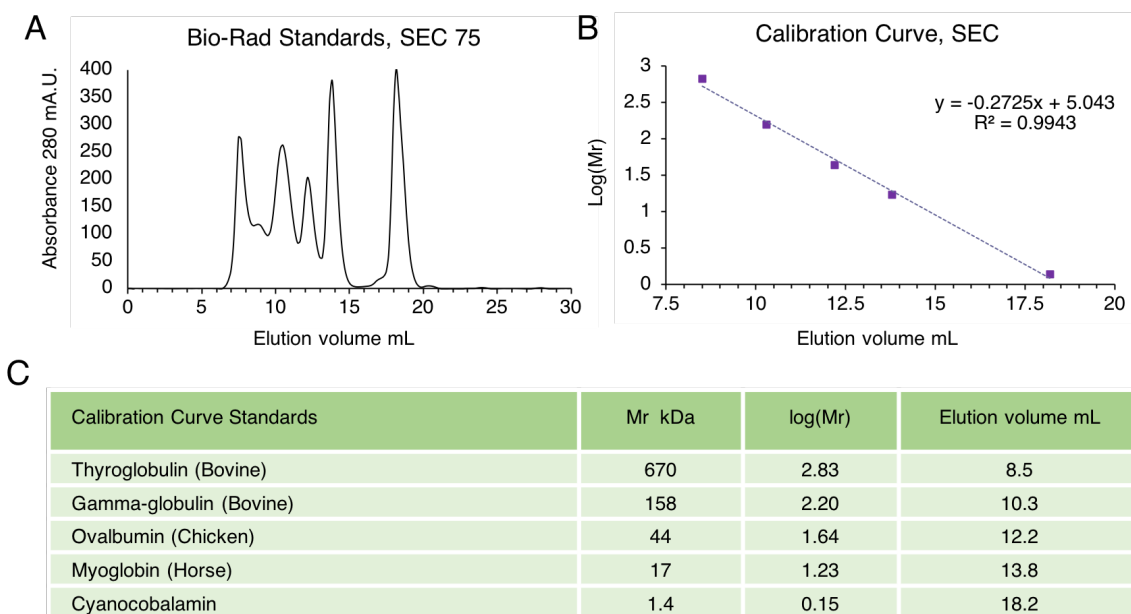
Deconvoluted mass spectra of A) Dark state *OdAu1a<sub>LOV</sub>* indicating a molecular mass of 14,997.0 AMU, consistent with apoprotein. B) Illuminated *OdAu1a<sub>LOV</sub>* indicating a species with molecular mass of 15,452.5 AMU corresponding to the predicted mass of a 5dFMN covalent photoadduct. C) Dark-state 5dFMN*OdAu1a<sub>ZIPLOV</sub>* indicating a single species with molecular mass of 25,688.5 AMU corresponding to mass of apoprotein. The error of experiment was 1 AMU.



**Figure 3.20 Mass spectrum of harvested lit state 5dFMN containing *OdAu1a<sub>LOV</sub>* crystals.**

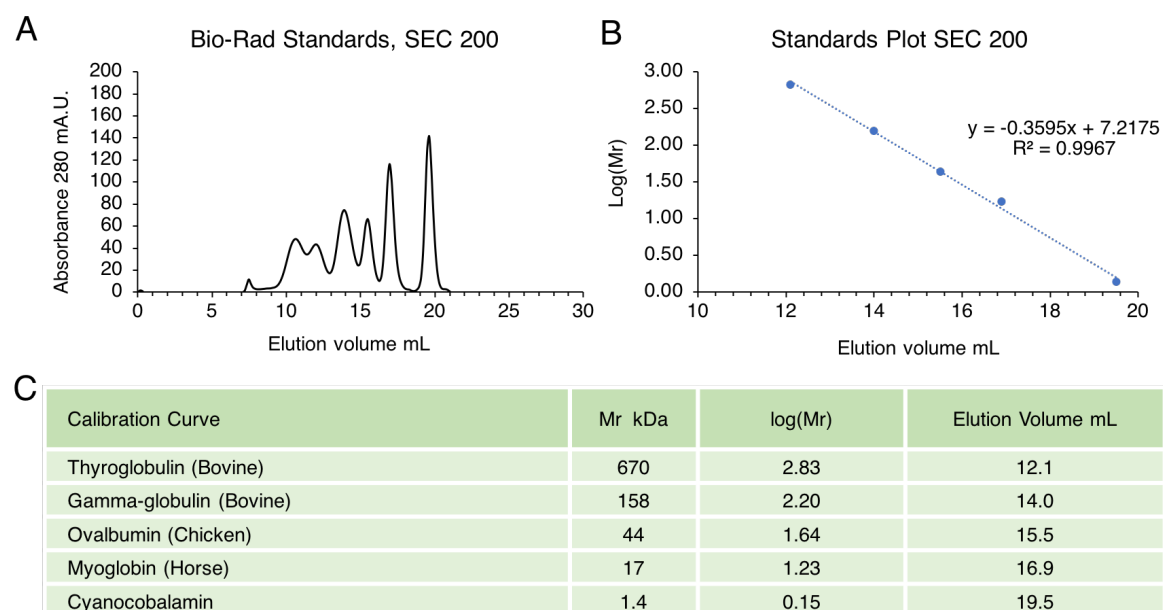
5-deaza flavin mononucleotide containing crystal were grown in dark state and were illuminated for 30 min with 450 nm LED lights. To ensure full photoactivation, protein crystals were harvested, gently washed 3 times with crystallization buffer and were dissolved in 20 mM Tris.HCl buffer pH 8.0 for MS analysis. A) UV chromatograph of lit state 5dFMN *OdAu1a<sub>LOV</sub>* crystal. Sharp chromatograph peaks indicate polyethylene glycol (PEG) species and peak at 30.33 min indicates a protein peak. B) As observed in the figure, deconvoluted mass spectrum indicates the mass of a protein closely corresponding to a *OdAu1a<sub>LOV</sub>* cysteinyl-5dFMN photoadduct. Unfortunately, due to PEG loading sequential analysis of FMN crystals could not be performed. The error of experiment was 1 AMU.





**Figure 3.21 Analytical size exclusion chromatography calibration curve of Superdex 75 column.**

A) Chromatogram of gel filtration standards of known molecular mass. Protein aggregates at 7.5 mL were considered as the void volume. B) Calibration curve of standard proteins. C) Table of protein standards, corresponding elution maxima and molecular sizes.



**Figure 3.22 Analytical size exclusion chromatography calibration curve of Superdex 200 column.**

A) Chromatogram of gel filtration standards of known molecular mass. Protein aggregates at 12.4 mL were considered as the void volume. B) Calibration curve of standard proteins. C) Table of protein standards, corresponding elution maxima and molecular sizes.



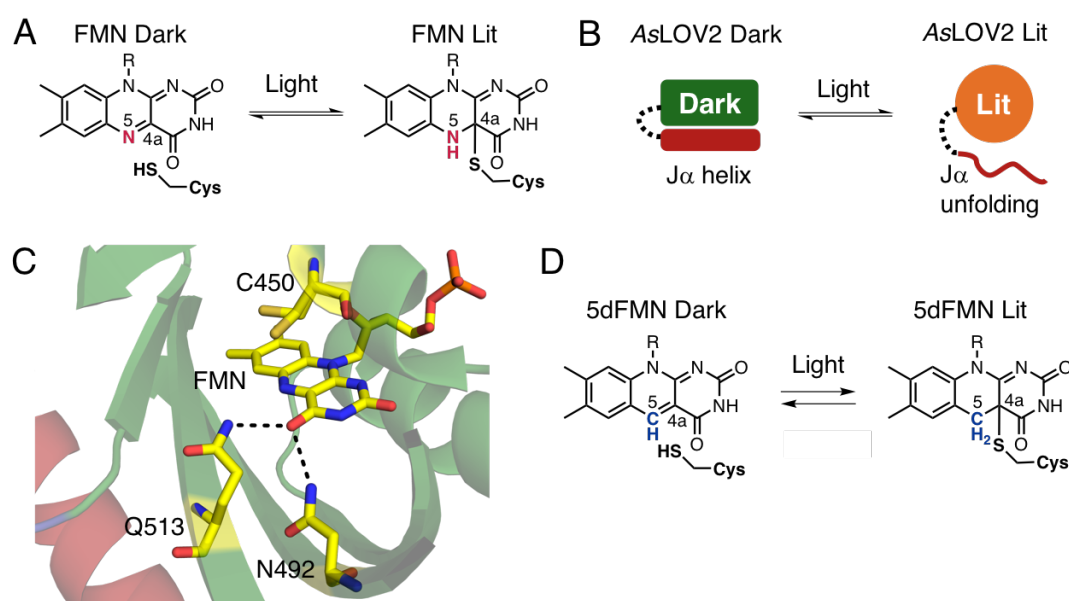
## 4 Understanding of LOV Domain Signalling – Testing Proposed Hypothesis in *Avena sativa* Phototropin LOV2 Domain (AsLOV2)

LOV domains have been used in the design and development of light controlled proteins to provide non-invasive, dose dependent and traceless approaches to study biological systems. The widespread use of engineered photoreceptors is hindered by challenges associated with assessing functionality and light-responsiveness. In this work 5-fluorotryptophan labelling in conjugation with  $^{19}\text{F}$  NMR spectroscopy was used to pinpoint the structural mechanism of *Avena sativa* phototropin LOV2 domain (AsLOV2) photoactivation using the natural FMN and synthetic 5dFMN cofactors. Through monitoring  $^{19}\text{F}$  NMR chemical shift differences and circular dichroism (CD) spectra, the results provide evidence that interruption of the conjugated isoalloxazine ring system participates in the activation of LOV domain proteins. Using  $^{19}\text{F}$  NMR chemical shifts to follow the photoactivation of LOV domain proteins opens a new avenue to study LOV domain conformational changes. This approach could be applied to the study engineered LOV domain proteins above the size limit of conventional backbone  $^{15}\text{N}$  and  $^{13}\text{C}$  NMR methods.

### 4.1 Introduction

Development of engineered LOV domain proteins relies on our understanding of LOV domain photochemistry and the structural mechanism of activation<sup>1-7</sup>. An easy method to assist more detailed functional and structural characterisations of LOV domain proteins is required for the validation of engineered constructs. One of the most commonly used LOV domain proteins in the development of photocaged proteins is *Avena sativa* phototropin LOV2 domain (AsLOV2)<sup>2-7</sup>. As a canonical LOV domain protein, AsLOV2 binds flavin cofactor within the hydrophobic protein core<sup>8,9</sup>. Upon exposure to blue light, excitation of FMN occurs and results in the formation of a covalent adduct between C4a of FMN and a cysteine residue<sup>8,9</sup>, Figure 4.1 A. Formation of the photoadduct then leads to the unfolding of A' $\alpha$  helix followed by the release and the unfolding of C-terminal J $\alpha$  helix B<sup>8,9</sup>, Figure 4.1. This principle of light activation has been used in the development of multiple photo-controlled systems, but what properties of the photoadduct drive the unfolding of the C-terminus remain unclear.

In the resting dark state, glutamine513 (Q513) and asparagine492 (N492) hydrogen bond to the FMN-O4<sup>9</sup>, Figure 4.1 C. It has been suggested that upon the photoadduct formation, protonation of the FMN-N5 results in Q513 side chain to flip orientation. This in turn repositions the LOV  $\beta$ -sheet surface structure and alters its dynamics, which drives unfolding of the J $\alpha$  helix through a conserved hydrogen bond network<sup>8-10</sup>. The importance of Q513 is supported by mutagenesis studies, where replacement of Q513 with leucine (Q513L) or asparagine (Q513N) resulted in attenuating light-responsiveness<sup>10</sup>. Most notably, the Q513L mutant displayed pseudo lit state characteristics by CD spectroscopy, yet it was incapable of forming hydrogen bonds to FMN-O4 or protonated cysteinyl-FMN-N5. This therefore suggests that other factors excluding FMN-N5 protonation and Q513 flip contribute to the unfolding of the J $\alpha$  helix. The protonation hypothesis is supported by the studies of other model LOV domain proteins including *Neurospora crassa* Vivid (NcVVD)<sup>11</sup> where protonation of the N5 was also suggested to dictate the structural changes of activation<sup>12,13</sup>.



**Figure 4.1 *Avena sativa* phototropin LOV2 (AsLOV2) activation by light.**

A) LOV domain photochemistry between FMN and a cysteine residue forming cysteinyl-FMN photoadduct. B) Structural mechanism of AsLOV2 activation by light results in C-terminus J $\alpha$  helix unfolding. C) Crystal structure of dark state (PDB: 2V0U) AsLOV2 domain displaying hydrogen bonds between FMN-O4 and the protein core plus a conserved cysteine 450 (C450). D) Photochemical reaction between 5dFMN and a cysteine residue generating cysteinyl-5dFMN photoadduct.

Opposed to the changes in the protonation state of FMN-N5 driving AsLOV2 structural activation, Fourier Transform Infrared (FT-IR) experiments, time resolved FT-IR experiments and molecular dynamics (MD) studies on AsLOV2 identified additional changes in hydrogen bonds to the FMN-O4 from Q513 and N492 and which occur prior unfolding of the J $\alpha$  helix<sup>14-17</sup>. Furthermore, these studies could not identify potential hydrogen bond of Q513 side chain carbonyl to the protonated cysteinyl-FMN-N5. To probe if structural rearrangements are a result of the FMN-N5 protonation or simply interruption of a conjugated isoalloxazine ring system, it was decided to use 5-deaza flavin mononucleotide, 5dFMN<sup>18</sup>, Figure 4.1 D, as a chromophore analogue incorporated into AsLOV2. 5dFMN lacks the capacity to hydrogen bond at the fifth position, C5, Figure 4.1 D. If protonation acts as the sole driver for AsLOV2 domain activation, 5dFMN should provide no structural change upon the photoadduct formation and no light control over unfolding of the J $\alpha$  helix. Conversely, if interruption of the conjugated isoalloxazine ring system leads to the changes in hydrogen bonding extending from the FMN-O4 position and if these changes contribute to AsLOV2 domain activation, 5dFMN should provide detectable light-responsiveness. Earlier structural studies on 5dFMN reconstituted *Ochromonas danica* Aureochrome1a LOV domain (*OdAu1a<sub>LOV</sub>*) have shown that the covalent adduct forms at the C4a position as for the natural FMN cofactor<sup>19</sup>. Nonetheless, no large-scale structural changes for the lit state 5dFMN *OdAu1a<sub>LOV</sub>* were observed, but some residue rearrangements were evident which suggests a partial 5dFMN light-responsiveness. Considering the applicability and the importance of AsLOV2 in the development of novel optogenetic tools<sup>4-7</sup>, further understanding of how the photoadduct governs structural changes is necessary to assist in the rational design of constructs.

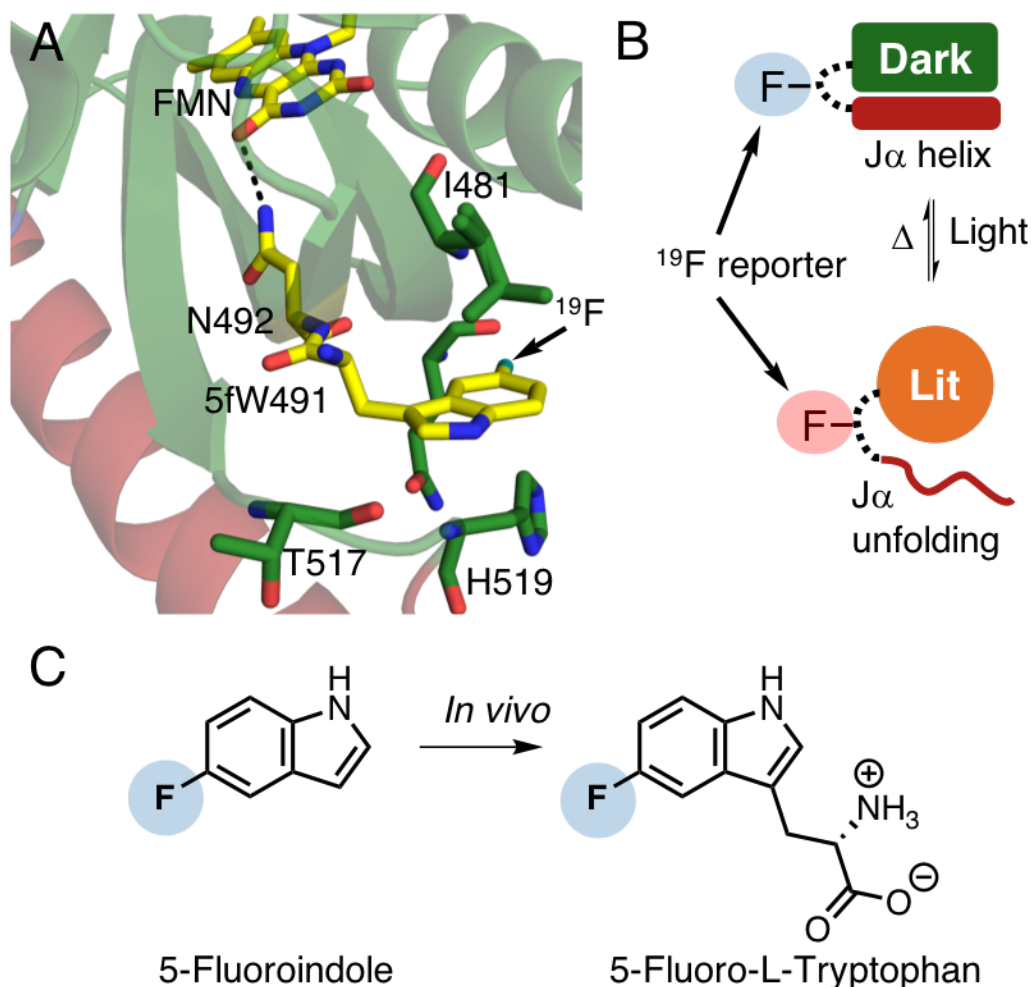
To probe structural changes, it was decided to use <sup>19</sup>F labelling to investigate AsLOV2 photoactivation. Protein <sup>19</sup>F labelling has been used to investigate protein folding<sup>20</sup> and structural changes to assess protein-protein interactions<sup>21,22</sup>. Considering current literature and to the best of our knowledge, <sup>19</sup>F labelling of LOV has not been previously undertaken to investigate light dependent structural changes. Considering the applicability of AsLOV2 in protein engineering, development of a <sup>19</sup>F labelling strategy in combination with <sup>19</sup>F NMR chemical shifts reporting J $\alpha$  unfolding would allow rapid assessment of the functionality of engineered AsLOV2 proteins that is not limited by protein size like conventional NMR methods requiring U-<sup>15</sup>N/<sup>13</sup>C labelling.

Fluorine probes are incorporated into proteins either by supplementation of growth nutrients with required fluorinated amino acids<sup>23-25</sup>, incorporating fluorinated amino acids via genetic code expansion methodology<sup>26</sup>, or by labelling of purified protein with the

desired fluorine probes<sup>27,28</sup>. The  $^{19}\text{F}$  atom van der Waals radius is only 18% larger than for a  $^1\text{H}$  atom<sup>29</sup>, hence incorporation of  $^{19}\text{F}$  probe in the corresponding amino acid residue usually results in only minor net structural perturbations<sup>23-25</sup>. Unlike  $^1\text{H}$  nuclei,  $^{19}\text{F}$  nuclei are rarely found in biological systems making background contamination unlikely<sup>30</sup>. Furthermore, the  $^{19}\text{F}$  nucleus is a spin  $\frac{1}{2}$  species, has natural abundance of almost 100% and a gyromagnetic ratio 83% that of  $^1\text{H}$  nucleus allowing high sensitivity in the  $^{19}\text{F}$  NMR spectroscopy<sup>30</sup>.  $^{19}\text{F}$  incorporation into fluorinated aliphatic and aromatic amino acid residues displays a large range of chemical shifts dependent on the local chemical environment including hydrophobicity, electrostatics and dielectric constants<sup>25,31</sup>. This sensitivity manifests as approximately 100 times wider chemical shifts observed for the  $^{19}\text{F}$  nuclei in comparison to the  $^1\text{H}$  making  $^{19}\text{F}$  NMR especially useful in the study of protein conformational changes. The intrinsic sensitivity to the chemical environment arises from  $^{19}\text{F}$  nuclei having six outer electrons that enforce different shielding effects as a product of electron motions around the nucleus. For instance, the chemical shift for  $^{19}\text{F}$  nuclei in a hydrophobic environment such as protein core is substantially different to the chemical shift of  $^{19}\text{F}$  exposed to water with a much higher dielectric constant<sup>21</sup>. This therefore makes 1D  $^{19}\text{F}$  NMR data easier to interpret providing a wealth of information regarding the local chemical environment. This methodology was therefore applied in the study of AsLOV2 incorporated with 5dFMN.

## **4.2 UV-Visible, Circular Dichroism and Nuclear Magnetic Resonance Studies of Native and $^{19}\text{F}$ AsLOV2**

Initially, AsLOV2 crystal structure (PDB code: 2V0U) was investigated to pinpoint any key residues that could be used for the labelling studies. Inspection indicated a tryptophan491 (W491) side-chain located in a latch region between the LOV domain core and  $\text{J}\alpha$  helix, Figure 4.2 A and B. W491 is partially shielded from the solvent by neighboring residues, and  $\text{J}\alpha$  undocking and unfolding most likely causes the residue to get exposed to the solvent. Additionally, W491 is upstream from N492, a residue hydrogen bonding to FMN-O4, in an ideal position to investigate structural changes as a result of hydrogen bond changes upon the photoadduct formation.



**Figure 4.2 5-fluorotryptophan labelling of AsLOV2.**

A) Structure of 5fW491 modelled into AsLOV2 crystal structure (PDB 2V0U) with the hydrogen bond between FMN-O4 and N492 highlighted. B) Schematic representation of AsLOV2 highlighting the  $^{19}\text{F}$  probe located in a loop region between LOV domain core and  $\text{J}\alpha$  helix. C) Structures of 5-fluoroindole and 5-fluoro-L-tryptophan (5fW). 5fW labelling is achieved by introducing 5-fluoroindole into the media where in *Escherichia coli* tryptophan synthase generates 5fW which gets incorporated into the proteins.

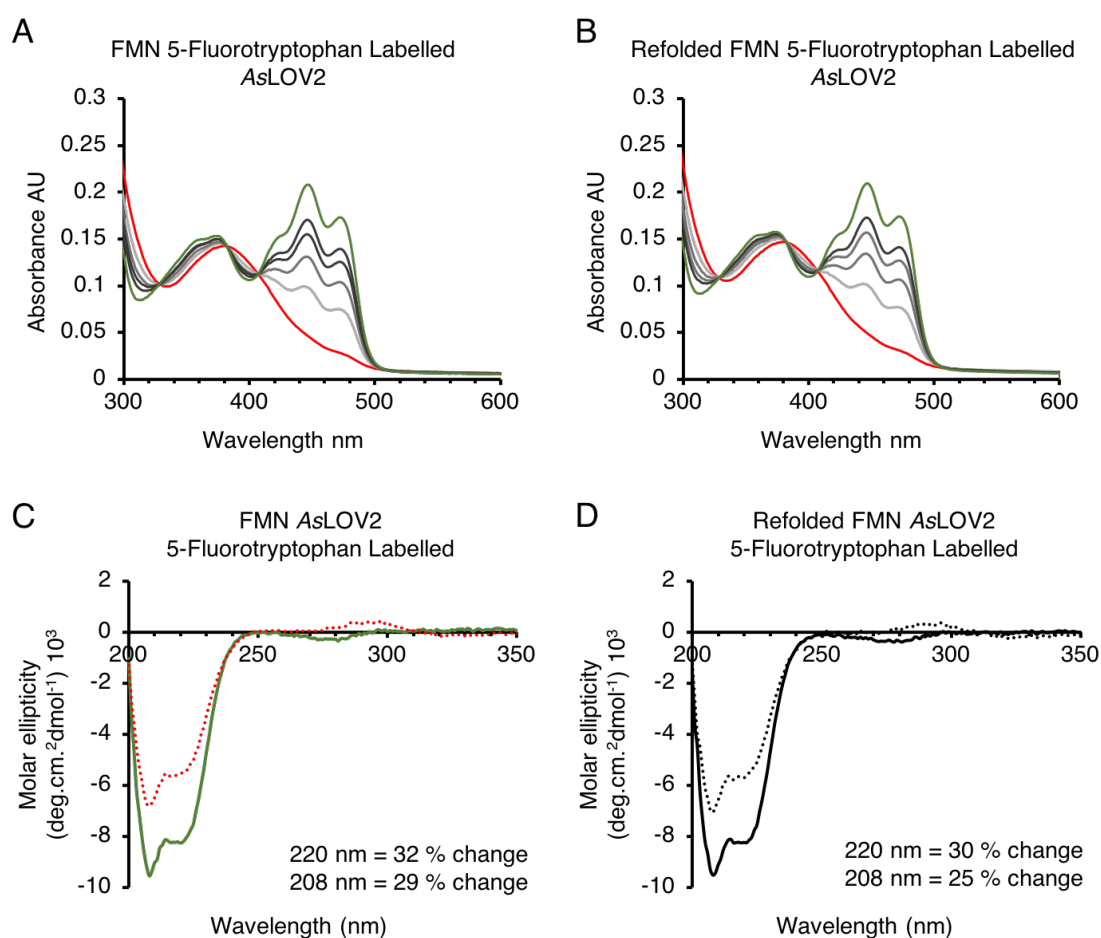
In order to investigate structural changes using W491 as a reporter, 5-fluoroindole was used to enrich the protein with 5-fluorotryptophan (5fW) following previously described procedures<sup>20</sup>, Figure 4.2 C. 5-Fluoroindole is a relatively cheap reagent, where per 1 L of minimal defined media (M9) culture only 60 mg is used, yielding a net cost for protein labelling of £0.26 (5 g Fluorochem £22.00). Reflecting on  $^{19}\text{F}$  nuclei sensitivity to the local environment, 5fW was hypothesised to provide understanding at a single residue resolution. On the other hand, conventional  $^{15}\text{N}$ - $^1\text{H}$ -heteronuclear single-quantum coherence (HSQC) spectra for FMN and 5dFMN AsLOV2 could result in significantly

different spectra making data difficult to interpret without a complete assignment of the resonances. This challenge with data interpretation has been observed in a previously mentioned mutagenesis study of AsLOV2, Q513N and Q513L mutations, respectively<sup>10</sup>. Although  $^{15}\text{N}$ - $^1\text{H}$  HSQC NMR spectra were collected, interpretation was hindered by significant differences in the chemical shifts from the wild-type protein. Time consuming residue assignment would therefore have to be repeated for each mutant in order to establish any conclusive interpretations. Considering 5fW labelling, data interpretation was expected to be less challenging, and therefore it was hypothesised that if 5dFMN behaves identically to FMN, identical chemical shifts would be observed in the  $^{19}\text{F}$  NMR spectra. Conversely, if no or partial light responsiveness would be evident for 5dFMN AsLOV2, this would be reflected in the  $^{19}\text{F}$  NMR spectra.

The cofactor exchange and refolding of AsLOV2 was achieved by spin column refolding. Briefly, the purified protein sample was concentrated in a spin column and washed with denaturation buffer (50 mM Tris.HCl, 6 M guanidine hydrochloride (Gnd.HCl)) supplemented with fresh dithiothreitol (DTT) to protect cysteine residue from oxidation. The protein sample was washed with chaotrope until spin column flow through did not show any FMN absorbance as determined by UV-Vis spectroscopy. Refolding was achieved by applying decreasing amounts of chaotrope, 5, 4, 3 M of Gnd.HCl, respectively. When protein was in a buffer containing around 3 M Gnd.HCl, the sample was diluted with a buffer containing an excess cofactor, 5dFMN or FMN (refolding control) by adding slowly with rapid mixing.

Firstly, the protein refolding methodology was investigated by FMN to FMN exchange. UV-Vis and CD experiments demonstrated that refolding yields a functioning protein without altered photo-responsiveness, Figure 4.3 and Appendix Figures 4.10 and 4.21. Labelling with 5fW was confirmed by Liquid-Chromatography Mass-Spectroscopy (LC-MS), Appendix Figures 4.15 to 4.18. Furthermore, UV-Vis and CD studies indicated that 5fW label did not interfere with protein folding and light responsiveness, Appendix Figures 4.20 and 4.21, resulting in lit to dark state recovery half-lives of 5.7 min for AsLOV2, 5.7 min for 5fW AsLOV2 and 5.6 min for refolded 5fW AsLOV2. When 5fW labelled FMN AsLOV2 was photoactivated and CD spectrum collected, a loss of the  $\alpha$ -helical content was observed at 220 and 208 nm minima, respectively, Figures 4.3 C and D, in line with  $\text{J}\alpha$  helix unfolding post photoactivation<sup>8,9</sup>.



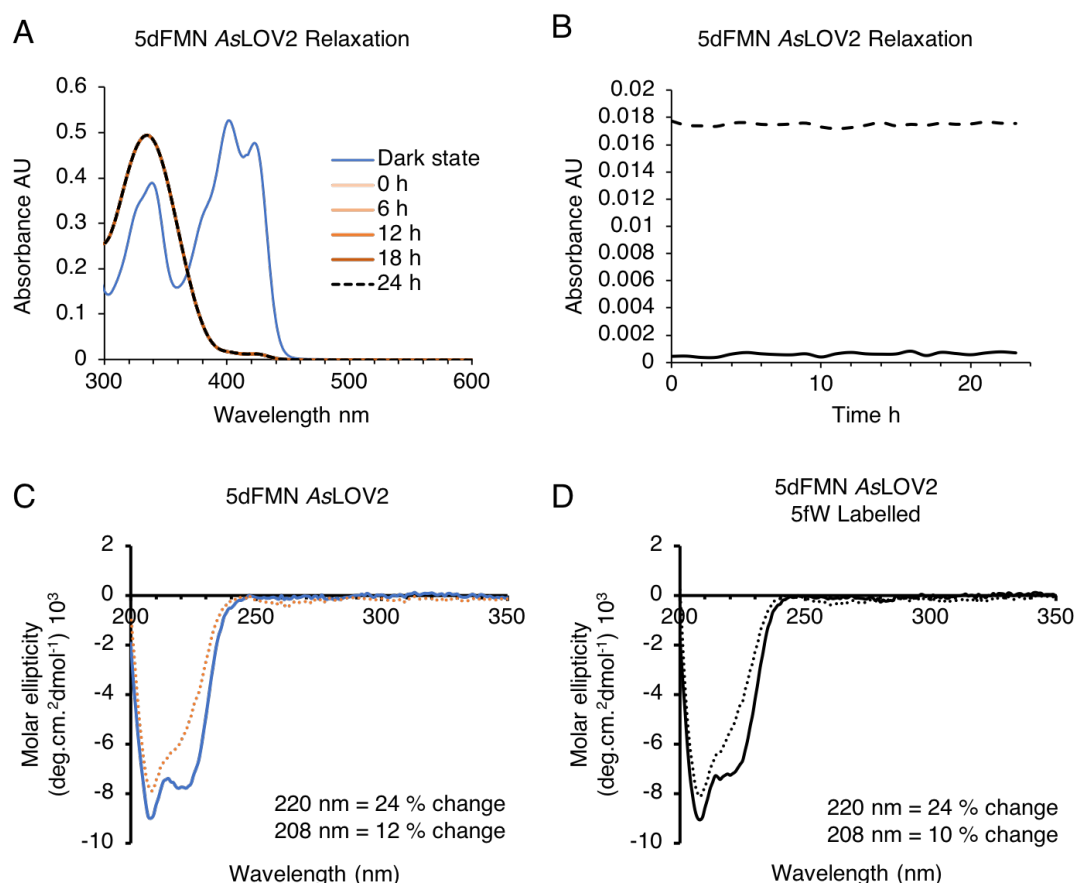


**Figure 4.3 UV-Vis and CD characterisation of FMN AsLOV2.**

A) 5fW labelled FMN AsLOV2 and B) refolded 5fW labelled FMN AsLOV2 relaxing from lit, red line, to dark state, green line. Spectra were collected for 87 min taking a reading every 3 minutes. For clarity not all data is shown. C) CD spectra of 5fW labelled FMN AsLOV2 switch from dark (green solid line) to lit (red dots) states. D) CD spectra of refolded 5fW labelled FMN AsLOV2 switch from dark (black solid line) to lit (black dots) states. Observed secondary structural changes are indicated in the corresponding figures.

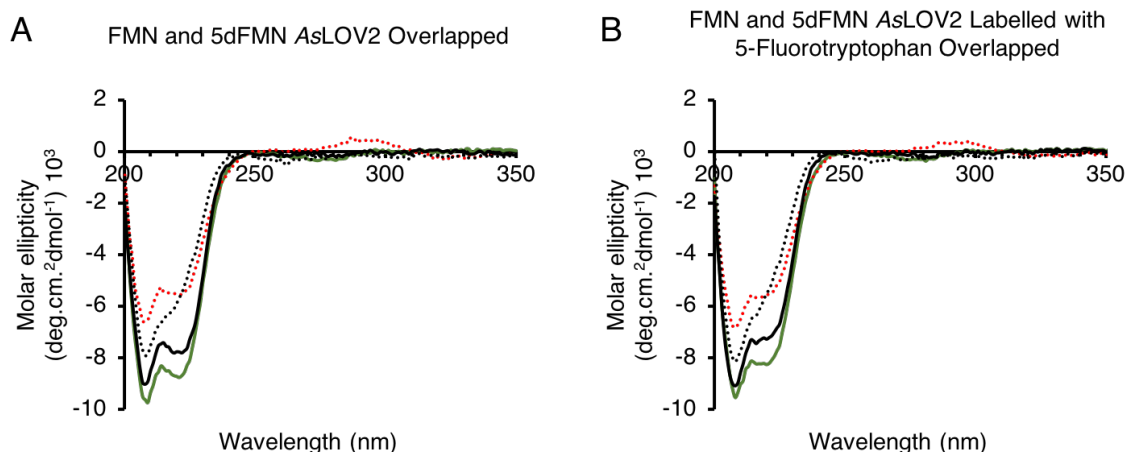
Considering reproducible refolding methodology, the effects of 5dFMN introduction were investigated. The photobleaching of the vibration triplet of 5dFMN was observed by UV-Vis spectroscopy, Figure 4.4 A. The cysteinyl-5dFMN photoadduct was stable for at least 24 hours, Figure 4.4 A and B. Considering UV-Vis data, 5dFMN AsLOV2 sample could not indicate any observable FMN contamination, Figure 4.4 B. LC-MS analysis indicated a stable lit state cysteinyl-5dFMN photoadduct mass, Appendix Figures 4.16 and 4.18. These results indicated that lit state 5dFMN, unlike lit state FMN, AsLOV2 was stable under the LC-MS conditions, Appendix Figures 4.14 and 4.16. Photoactivation of 5dFMN AsLOV2 indicated changes in the CD spectrum, Figure 4.4 C and D, 24% signal loss at 220 nm and 12% at 208 nm, respectively, indicative of lower than FMN, but observable

structural changes, Figure 4.5. Additionally, time course CD experiments indicated that for FMN AsLOV2, unfolding of the J $\alpha$  was reversible, Figure 4.6 A, while photoactivated 5dFMN AsLOV2 CD yielded lit state locked conformation judged from time course experiment, Figure 4.6 B.



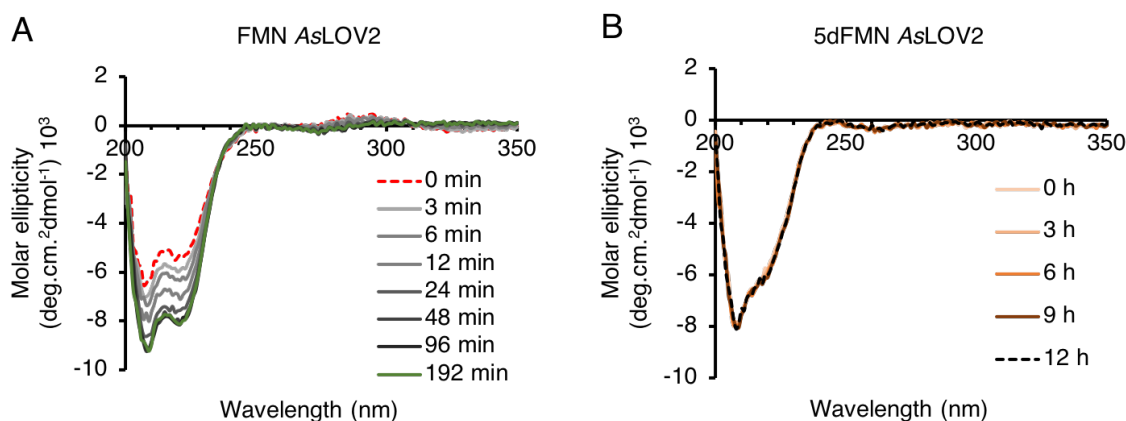
**Figure 4.4 UV-Vis and CD characterisation of 5dFMN AsLOV2.**

A) Time course investigation of lit state 5dFMN AsLOV2 stability, where blue line indicates dark state while orange solid lines plus 24 h time point are displayed as dashed black line, respectively. B) Investigation of potential FMN contamination, dashed line indicates absorbance at 401 nm (5dFMN maxima) and solid line absorbance at 447 nm (FMN maxima), indicating no detectable changes at 447 nm absorbance. C) CD spectra of 5fW labelled 5dFMN AsLOV2 switch from dark (blue solid line) to lit (orange dots) states. D) CD spectra of refolded 5fW labelled 5dFMN AsLOV2 switch from dark (black solid line) to lit (black dots) states. Observed secondary structural changes are indicated in the corresponding figures.



**Figure 4.5 CD of AsLOV2 containing FMN and 5dFMN as a cofactor.**

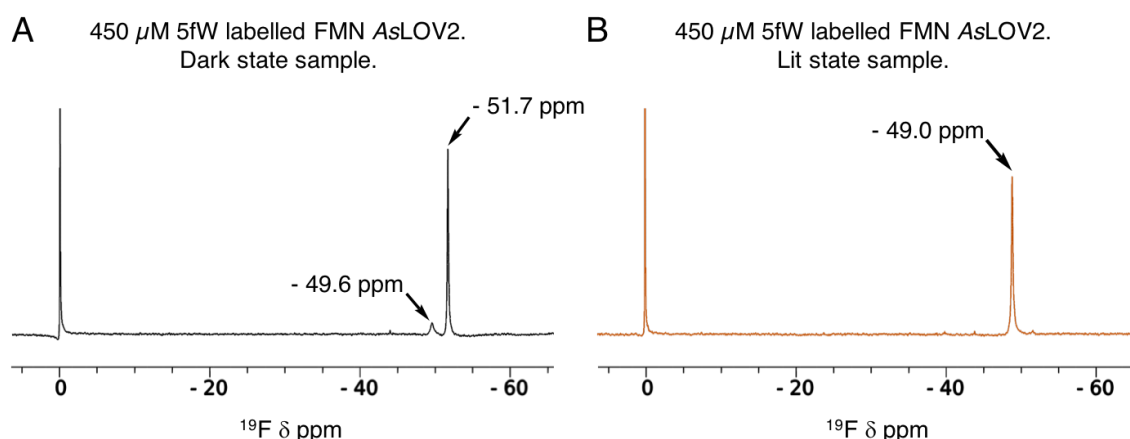
A) FMN AsLOV2 in dark (solid green line) and lit (red dots) states, respectively. 5dFMN AsLOV2 dark (solid black line) and lit (black dots) states, respectively. Percentage difference between dark FMN and dark 5dFMN at 220 nm was 11%, and at 208 nm 6%. Lit FMN and lit 5dFMN difference at 220 nm was 6% and at 208 nm 16%. B) 5-fluorotryptophan (5fW) labelled protein samples containing FMN AsLOV2 in dark (solid green line) and lit (red dots) states, respectively. 5fW labelled 5dFMN AsLOV2 dark (solid black line) and lit (black dots) states, respectively. Percentage difference between dark FMN and dark 5dFMN at 220 nm was 12%, and at 208 nm 5%. Lit FMN and lit 5dFMN difference at 220 nm was 1% and at 208 nm 16%.



**Figure 4.6 CD following FMN and 5dFMN containing AsLOV2 relaxation.**

A) FMN AsLOV2 relaxation from lit to dark state, respectively. Individual spectra were taken every 3 min for 192 min. For clarity, not all spectra are displayed. B) 5dFMN reconstituted AsLOV2, where individual spectra were taken every hour for a total time of 12 h. For clarity, not all spectra are displayed. For 5dFMN sample, no relaxation was evident indicating a lit state locked secondary structure.

To judge the functionality of 5dFMN by  $^{19}\text{F}$  NMR, 5fW labelled AsLOV2 was incorporated with both FMN and 5dFMN chromophores.  $^{19}\text{F}$  1D NMR spectra were recorded to investigate the chemical shift changes for the native FMN, refolded FMN (refolding control) and 5dFMN protein samples. For native and refolded FMN AsLOV2, large  $^{19}\text{F}$  chemical shift changes were observed upon irradiation, Figure 4.7 A and B, and Appendix Figure 4.22. In the dark state, a major peak with a  $\delta$  of -51.7 ppm and a minor peak at  $\delta$  of -49.6 ppm were observed, Figure 4.7 A. Integration of the minor peak gave an area 15% that of the dark state major peak. Photoactivation resulted in the disappearance of the two peaks yielding a new peak with  $\delta$  of -49.0 ppm, Figure 4.7 B. Identical results were observed for the refolded protein sample, Appendix Figure 4.22, furthermore supporting the fidelity of the refolding methodology. Such chemical shift changes were probably a result from 5fW sidechain becoming exposed to the solvent. Similar  $^{19}\text{F}$  chemical shift differences were observed in the literature when investigating protein folding<sup>21,32</sup> where the formation of the hydrophobic protein core resulted in upfield  $^{19}\text{F}$  chemical shifts.

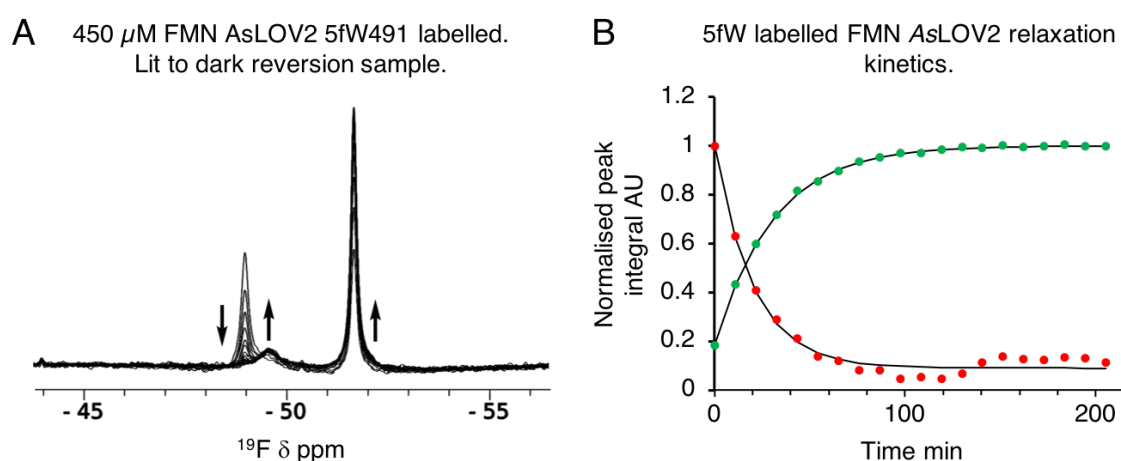


**Figure 4.7  $^{19}\text{F}$  1D NMR characterisation of 5fW labelled FMN AsLOV2.**

$^{19}\text{F}$  NMR spectra of FMN AsLOV2 under A) dark and B) lit states, respectively. All chemical shifts reported in the text and represented in the figure were referenced to the trifluoro acetate (TFA) peak adjusted to  $\delta$  0.0 ppm.

1D  $^{19}\text{F}$  NMR were then recorded every 10.8 min after irradiation to determine the half-life of the decay back to dark state, Figure 4.8 A and B. Relaxation kinetics showed a longer half-life than that observed by UV-Vis experiments (5.7 min), Figure 4.8 A and B. Surprisingly, the decay of the lit peak,  $\delta$  -49.0 ppm, and growth of the dark peak,  $\delta$  -51.7 ppm, showed different half-life kinetics. The decay of the lit peak at  $\delta$  -49.0 ppm had a half-life of 13.9 min, while the growth of the dark peak at  $\delta$  -51.7 ppm had a half-life of 21.2

min. This was probably a result of long data acquisition during NMR experiments, as data collection for 512 scans was 10.8 min, much longer than the half-life of reversion back to the dark state of 5.7 min. Considering that the NMR spectrum is represented as a sum of multiple FIDs, large errors could be introduced. For the lit peak decay, the experiment began with 100% peak intensity at time = 0 min, then reduced over time, but what was observed was an integral between timepoints time = 0 min to time = 10 min. By contrast, for the growth of dark peak, experiment at time = 0 min yielded no signal, while at time = 10 min resulted in multiple FIDs with high integrals resulting in an overestimated peak intensity. Alternatively, it is also likely that fewer datapoints, every 10.8 min, for the data fit resulted in much larger errors. Unfortunately due to time restraints these observations were not explored further.

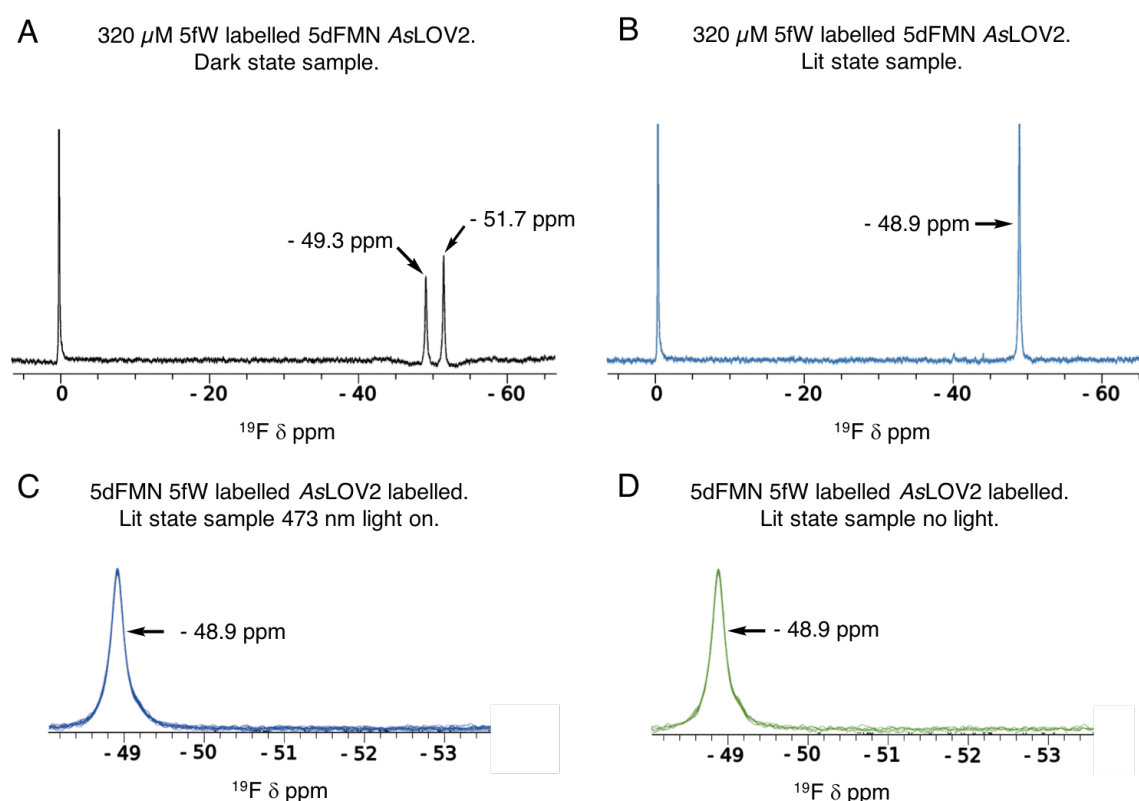


**Figure 4.8  $^{19}\text{F}$  1D NMR relaxation kinetics of 5fW labelled AsLOV2.**

A)  $^{19}\text{F}$  NMR relaxation of AsLOV2 from lit to dark state, respectively, with arrows indicating corresponding peak decay (down) or growth (up). B) Plotted relaxation kinetics of 5fW labelled AsLOV2 with dots in green indicating growth of dark peak at  $\delta$  -51.7 ppm, half-life 21.2 min, and red dots decay of lit peak at  $\delta$  -49.0 ppm, half-life of 13.9 min. Solid black line indicates plotted first order decay curves.

Strikingly, the dark state 5dFMN reconstituted 5fW AsLOV2 1D  $^{19}\text{F}$  NMR spectrum indicated two peaks with similar peak heights, Figure 4.9 A. One of the dark peaks had a  $\delta$  of -51.7 ppm, Figure 4.9 A, identical  $\delta$  to the dark state FMN, Figure 4.7 A. The second peak had  $\delta$  of -49.3 ppm, Figure 4.9 A, similar to the dark state FMN minor peak,  $\Delta\delta$  of -0.3 ppm. Furthermore, the integral of the peak at  $\delta$  -49.3 ppm was 82% that of peak at  $\delta$  -51.7 ppm. When 5dFMN AsLOV2 sample was photoactivated, a change in  $^{19}\text{F}$  NMR spectrum was observed, Figure 4.9 B, resulting in a single peak with  $\delta$  of -48.9 ppm,  $\Delta\delta$  of -0.1 ppm in comparison to the lit FMN sample. The 5dFMN lit state peak  $^{19}\text{F}$  chemical shifts were independent of irradiation with 473 nm via the optic fiber providing evidence

that chemical shifts were indicative of structural changes and not due to blue light irradiation, Figure 4.9 C and D. The 15 individual spectra collected in the absence of 473 nm light indicated lit state locked protein without observable decay, Figure 4.9 D. Nonetheless, the spectrum of the dark state sample implied that 5dFMN, unlike FMN, containing AsLOV2 sample had a significant portion of an unknown secondary conformer of 5fW491, while in the lit state the conformer was almost identical to that of FMN AsLOV2. These observations are supported by the CD spectra of 5dFMN sample having lower global  $\alpha$ -helical content suggesting a population of unfolded J $\alpha$  helix hence yielding a prominent secondary 5fW491 peak. In summary, CD and NMR results suggested that 5dFMN was able to enforce structural changes prompting further investigations.



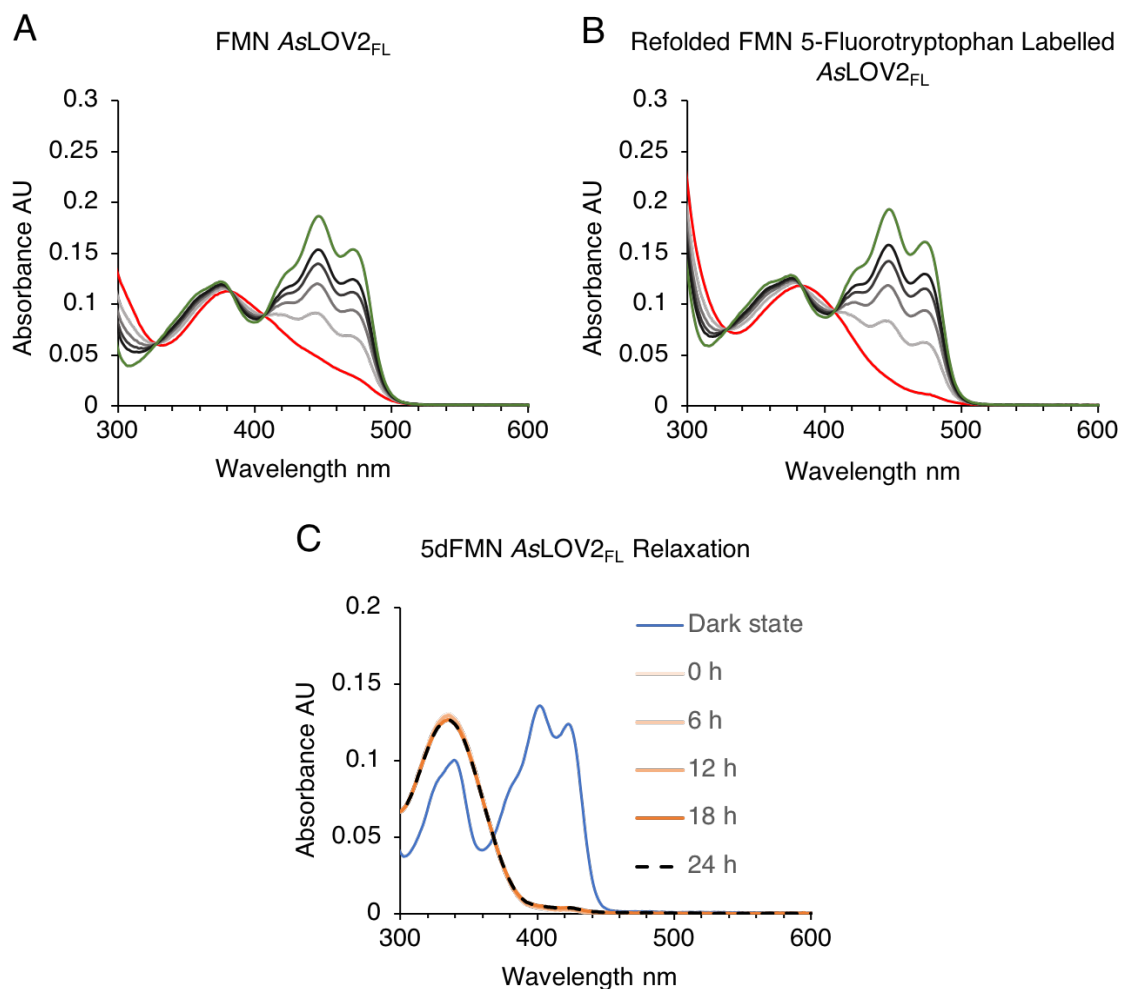
**Figure 4.9  $^{19}\text{F}$  1D NMR characterisation of 5fW labelled 5dFMN AsLOV2.**

$^{19}\text{F}$  NMR spectra of 5dFMN AsLOV2 under A) dark and B) lit states, respectively. C) Five overlapped individual  $^{19}\text{F}$  NMR spectra of lit state 5dFMN AsLOV2 irradiated with 473 nm light during the experiment. D) Fifteen overlapped individual  $^{19}\text{F}$  NMR spectra of lit state 5dFMN AsLOV2 in the absence of light indicating lit state locked protein. Please note that time to collect a single spectrum was 10.8 min. All chemical shifts reported were referenced to the TFA peak adjusted to  $\delta$  0.0 ppm.

### 4.3 Investigations of AsLOV2 Structural Changes with Elongated J $\alpha$ Helix

Considering the fact that 5dFMN fails to provide a hydrogen bond acceptor in the dark and donor in the lit states, the previous NMR and CD results suggest that FMN-N5 protonation in the lit state is not the only requirement for the unfolding of J $\alpha$  to occur. The partial functional responsiveness of 5dFMN supports the hypothesis that the covalent adduct formation changes the hydrogen bonding around (5d)FMN-O4 and that these changes participate in the structural change liberating the J $\alpha$  helix. Nonetheless, the lower  $\alpha$ -helical content in conjugation with a prominent  $^{19}\text{F}$  spectrum secondary peak suggested an unknown conformer representative of a population of unfolded J $\alpha$ . The initial AsLOV2 construct was 15 residues shorter in comparison to the construct with previously characterised AsLOV2 J $\alpha$  construct<sup>8</sup>. To directly compare current findings to the structural changes observed previously, a construct incorporating the additional 15 residues of the WT protein was generated. Notably, this full length AsLOV2 (AsLOV2<sub>FL</sub>) contained an additional C-terminal tryptophan 557 (W557).

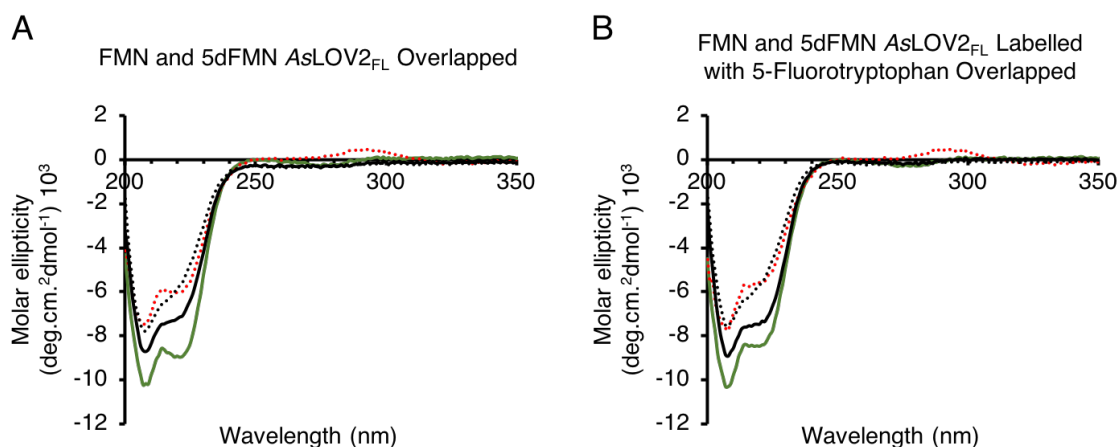
The AsLOV2<sub>FL</sub> construct was investigated by UV-Vis and CD, Figures 4.8 and 4.9. FMN AsLOV2<sub>FL</sub> had a relaxation half-life of 5.8 min, Figure 4.10 A, whilst FMN AsLOV2<sub>FL</sub> labelled with 5fW had slightly shorter half-lives, 5.1 min for non-refolded protein and 5.2 min for the refolding control, Figure 4.10 A and B, and Appendix Figure 4.23. CD studies indicated an average light dependent  $\alpha$ -helical loss of 34% at 220 nm and 26% at 208 nm, Figure 4.11 and Appendix Figure 24. Dark 5dFMN AsLOV2<sub>FL</sub> CD spectra were identical to the shorter AsLOV2 in showing a lower signal at 220 nm and 208 nm, Figure 4.11. 5dFMN AsLOV2<sub>FL</sub> showed an average light dependent change in the CD spectrum, 21% at 220 nm and 13% at 208 nm, respectively, Figure 4.11. Correct 5fW labelling was confirmed by LC-MS, indicating that the majority of the protein was labelled with 5fW, see Appendix Figure 4.19. Cysteinyl-5dFMN lit state adduct was also observed by LC-MS, yielding the expected AsLOV2<sub>FL</sub>-5dFMN photoadduct mass, Appendix Figure 4.18 C. Nonetheless, species corresponding to apo AsLOV2<sub>FL</sub> was also evident.



**Figure 4.10 UV-Vis characterisation of AsLOV2<sub>FL</sub>.**

A) FMN AsLOV2<sub>FL</sub> relaxing from lit, red solid line, to dark, green solid line. B) Refolded and 5fW labelled FMN AsLOV2<sub>FL</sub> relaxing from lit, red dashed line, to dark, green dashed line. In A) and B) spectra were collected for 87 min taking a reading every 3 minutes. For clarity not all data is shown. C) 5dFMN AsLOV2<sub>FL</sub> relaxation from lit, orange solid lines plus 24 h time point dashed black line, to dark, blue solid line. No relaxation was evident by UV-Vis experiments.

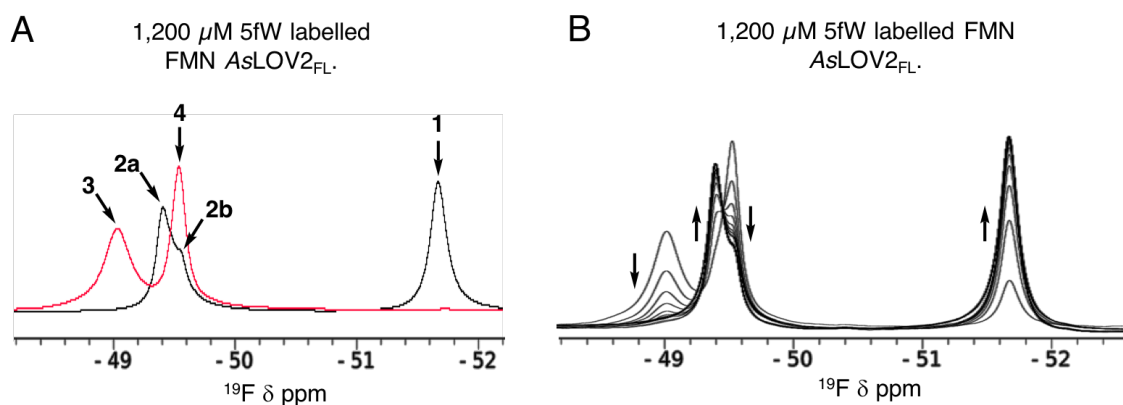




**Figure 4.11 CD characterisation of AsLOV2<sub>FL</sub> containing FMN and 5dFMN as a cofactor.**

A) FMN AsLOV2<sub>FL</sub> in dark (solid green line) and lit (red dots) states. 5dFMN AsLOV2<sub>FL</sub> dark (solid black line) and lit (black dots) states. Percentage difference between dark FMN and dark 5dFMN at 220 nm was 19%, and at 208 nm 14%. Lit FMN and lit 5dFMN difference at 220 nm was 2% and at 208 nm 4%. B) 5fW labelled FMN AsLOV2<sub>FL</sub> in dark (solid green line) and lit (red dots) states. 5fW labelled 5dFMN AsLOV2<sub>FL</sub> dark (solid black line) and lit (black dots) states. Percentage difference between dark FMN and dark 5dFMN at 220 nm was 12%, and at 208 nm 14%. Lit FMN and lit 5dFMN difference at 220 nm was 3% and at 208 nm 2%.

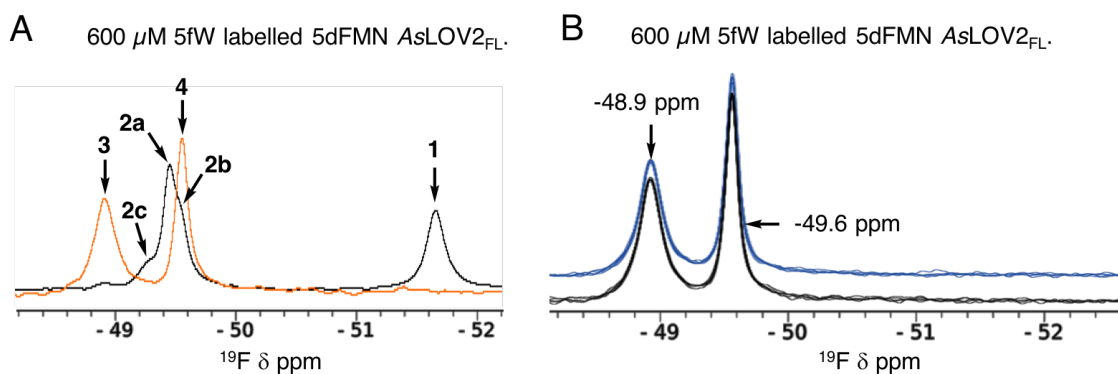
Considering that the CD results suggested that 5dFMN was able to dictate structural changes in AsLOV2<sub>FL</sub>, 1D <sup>19</sup>F NMR experiments were undertaken noting the previously assigned 5fW491 peak allowing C-terminal 5fW557 to be readily pinpointed. <sup>19</sup>F NMR spectrum of the dark state FMN AsLOV2<sub>FL</sub> displayed two distinct peaks, Figure 4.12 A. Peak 1 had a  $\delta$  of -51.7 ppm, identically to the peak observed in the dark state AsLOV2 spectra, Figure 4.7 A. The second peak, peak 2, appeared to be two overlapped peaks with peak 2a maxima at  $\delta$  -49.4 ppm and 2b shoulder at  $\delta$  -49.6 ppm, Figure 4.12 A. The peak 2a most probably corresponds to the C-terminal and solvent exposed 5fW557. The shoulder peak 2b at  $\delta$  -49.6 ppm most probably represents the minor peak observed for 5fW491 in the AsLOV2 construct, Figure 4.7 A. Irradiation of the sample yielded a lit state spectrum with two peaks, peaks 3 and 4, Figure 4.12 A. Peak 3 had a  $\delta$  of -49.0 ppm, identical to the previously observed lit state 5fW491, while peak 4 had a  $\delta$  of -49.6 ppm most likely corresponding to the lit state 5fW557. As for the initial construct, reversion from lit to dark states could be observed by <sup>19</sup>F NMR, Figure 4.12 B. The growth of the dark peak 1 at  $\delta$  -51.7 ppm yielded half-lives of 20.1 min. Lit to dark state switch was identical to the refolded AsLOV2<sub>FL</sub>, Appendix Figure 4.24, furthermore confirming reproducible refolding methodology.



**Figure 4.12  $^{19}\text{F}$  1D NMR characterisation of 5fW labelled FMN AsLOV2<sub>FL</sub>.**

A)  $^{19}\text{F}$  NMR spectra of FMN AsLOV2<sub>FL</sub>, dark spectrum in black and lit in red, respectively. Observed peaks are labelled with arrows with chemical shifts described in the text. B)  $^{19}\text{F}$  NMR relaxation of 5fW labelled AsLOV2<sub>FL</sub> from lit to dark state, respectively, with arrows indicating corresponding peak growth (up arrows) and decay (down arrows). Taking integral of dark peak growth at  $\delta$  -51.7 ppm, half-life of 20.1 min was determined. All chemical shifts reported were referenced to the TFA peak adjusted to 0.0 ppm.

When 5dFMN AsLOV2<sub>FL</sub> sample was investigated,  $^{19}\text{F}$  NMR spectrum displayed an identical to dark state FMN peak 1 with  $\delta$  -51.7 ppm, Figure 4.13 A. Peak 2a had a slightly altered  $\delta$  of -49.5 ppm, a  $\Delta\delta$  of -0.1 ppm in comparison to FMN maxima. As observed for the FMN AsLOV2<sub>FL</sub>, peak 2 indicated a shoulder peak with  $\delta$  -49.6 ppm, Figure 4.13 A peak 2b, which could represent a lit state like 5fW557 as observed upon photoactivation. This observation furthermore supported a population of unfolded J $\alpha$ . Unlike the FMN containing AsLOV2<sub>FL</sub> sample, the 5dFMN sample displayed another shoulder peak, 2c, with  $\delta$  within -49.3 ppm region, Figure 4.13 A. Shoulder peak 2c had  $\delta$  similar to that observed in a shorter 5dFMN AsLOV2 construct, Figure 4.9 A. The presence of shoulder 2c therefore suggested that AsLOV2<sub>FL</sub> contained a secondary conformer for 4fW491 providing further evidence for a population of unfolded J $\alpha$ . This was in line with CD data indicating lower global  $\alpha$ -helical content for 5dFMN in the dark. Photoactivation of 5dFMN AsLOV2<sub>FL</sub> resulted in only two  $^{19}\text{F}$  peaks observed with peak 3 corresponding to the lit state 5fW491 observed previously with  $\delta$  -48.9 ppm, Figures 4.13 A and 4.9 B, respectively. Peak 4, which most likely corresponds to lit state 5fW557, had  $\delta$  -49.6 ppm, identical to the FMN sample. Considering CD and NMR experiments, these provided further evidence that for AsLOV2<sub>FL</sub> construct, 5dFMN was able to control a degree of light dependent structural changes.



**Figure 4.13  $^{19}\text{F}$  1D NMR characterisation of 5fW labelled 5dFMN AsLOV2<sub>FL</sub>.**

A)  $^{19}\text{F}$  NMR spectra of 5dFMN containing 5fW labelled AsLOV2<sub>FL</sub> with dark spectra in black and lit in orange, respectively. Observed peaks are labelled with arrows with chemical shifts described in the text. B)  $^{19}\text{F}$  NMR spectra of lit state 5dFMN AsLOV2<sub>FL</sub> indicating lit state locked protein. Blue colour indicates ten overlapped individual spectra while lit state 5dFMN AsLOV2<sub>FL</sub> sample was irradiated with 473 nm light and black colour indicates ten overlapped individual spectra in the absence of light. Please note that time to collect a single spectrum was 10.8 min. All chemical shifts reported were referenced to the TFA peak adjusted to 0.0 ppm.

#### 4.4 Conclusion and Discussion

The lower global  $\alpha$ -helical content observed for 5dFMN AsLOV2 by the CD experiments suggested a population of unfolded J $\alpha$ . CD results are furthermore supported by the presence of a prominent secondary peak for 5fW491 in 1D  $^{19}\text{F}$  NMR spectra in the dark state for both constructs investigated. The observed secondary conformer of 5fW491 had a similar, but not identical, chemical shift to the lit state sample suggestive that the photoactivation results in further structural change possibly due to changes in the hydrogen bonding extending from the 5dFMN-O4. These results also suggested that 5dFMN was unable to fully lock J $\alpha$  helix and act identically to the FMN AsLOV2 under dark state conditions.

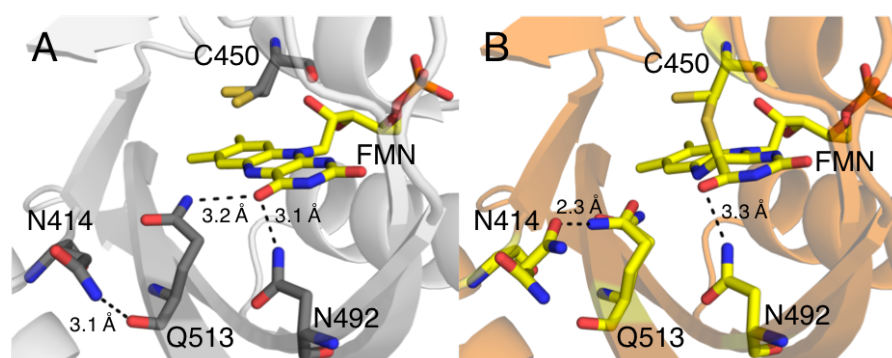
Nonetheless, CD and NMR results provided evidence that 5dFMN was capable of dictating a degree of light dependent structural changes for AsLOV2 even though 5dFMN was unable to act as a hydrogen donor or acceptor at the fifth position. These observations therefore suggest that hydrogen-bonding extending from the (5d)FMN-O4 and possibly interruption of the conjugated isoalloxazine ring system participate in driving the structural changes observed upon the photoadduct formation. Considering that hydrogen bonding networks extending from FMN-O4 have been proposed for BLUF domain protein photoactivation<sup>33,34</sup>, it is possible that LOV domains use a similar mechanism. Notably, structural studies of LOV domain proteins including ZEITLUPE and *Pp*SB1 have suggested that the hydrogen bonding extending from FMN-O4 position were important in

directing the structural rearrangements leading to the lit state conformation<sup>35,36</sup>. A recent ultra-high resolution structure of LOV domain protein from thermophilic organism *Chloroflexus aggregans* also identified two distinct conformations of the conserved Q148 residue hydrogen bonding to FMN-O4, although the importance of this hydrogen bonding arrangement was not investigated with regards to the signaling state<sup>37</sup>.

While 5dFMN incorporated into AsLOV2 provided unique observations indicating partial light responsiveness, further questions can be raised. For *OdAu1a*<sub>LOV</sub> domain construct, 5dFMN dark state behaved more similarly to the dark state FMN, while the lit 5dFMN indicated an intermediate behaviour in comparison to the lit FMN protein<sup>19</sup>. Additionally, for *OdAu1a*<sub>LOV</sub>, no differences in hydrogen bonding to FMN or 5dFMN-O4 could be identified in the dark state<sup>19</sup>, but it is not known if these observations can be applied to AsLOV2. For AsLOV2, 5dFMN appeared to provide more of a lit state like behaviour considering the lower global  $\alpha$ -helical content observed by CD and prominent secondary 5fW491 peak observed in the <sup>19</sup>F NMR spectrum. For 5dFMN AsLOV2 activation with light resulted in the lit state closely resembling that of a lit state FMN sample indicating light dependent unfolding of the J $\alpha$ . This therefore raises a question why these two LOV domain proteins behave differently. This question can be attempted to be answered examining the structural mechanism of activation. Au1a LOV domains have been shown to undergo light dependent dimerisation<sup>19,38-40</sup> suggesting that lit state dimeric conformation needs to be locked. This therefore suggests a possible mechanism where FMN-N5 protonation upon the photoadduct formation allows lit state conformation to be locked facilitating dimerisation. In contrast, AsLOV2 undergoes light dependent unfolding of the J $\alpha$  helix<sup>8,10,41</sup>. In resting dark state, AsLOV2<sub>FL</sub> has been shown to populate a small percentage of unfolded J $\alpha$  state, 1.6%<sup>41</sup>. This therefore suggests that the unfolding event needs to be locked in the dark state and alleviated in the lit state, respectively. Upon cysteinyl-5dFMN formation, putative changes in hydrogen bonding around 5dFMN-O4 via Q513 and N492 might suffice to result in entropically favourable J $\alpha$  unfolding. On the other hand, the observed population of unfolded J $\alpha$  for 5dFMN AsLOV2 in the dark state could be a product of two different factors. Firstly, the lack of a hydrogen bond acceptor at the 5-position could destabilise dark state Q513 position resulting in a population of unanchored J $\alpha$ . Secondly, chemical differences between FMN and 5dFMN can result in hydrogen bonds around 5dFMN-O4 position being weaker than for FMN-O4, hence resulting in the dark state not fully locked and a large percentage of unfolded J $\alpha$ .

The degree of light responsiveness observed for 5dFMN AsLOV2 is in line with previous FTIR, time resolved FTIR experiments and MD simulations suggesting that when

the cysteinyl-FMN photoadduct forms, a weakening of Q513 and N492 hydrogen bonding to FMN-O4 is observed<sup>14-16,42</sup>. Elongation of the FMN-O4 to N492 hydrogen bond, from 3.1 to 3.3 Å, has been observed in the room temperature structure of AsLOV2<sup>9</sup>, Figure 4.14 A and B. Furthermore, Q513 was modelled in two different conformations where neither orientation was in the hydrogen bond distance to the protonated FMN-N5 or even FMN-O4, Figure 4.14 A and B. Considering these reports, it is possible to suggest that protonation at N5 may not be the only determinant to switch AsLOV2 into the light activated state and that hydrogen bonding extending from the FMN-O4 also contribute significantly.



**Figure 4.14 Room temperature X-ray structures of AsLOV2 in dark and lit states indicating changes in hydrogen bonds extending from the FMN-O4.**

A) Room temperature structure of dark state AsLOV2 (PDB 2V1A, resolution 1.65 Å) indicating hydrogen bonds extending from FMN-O4 highlighting Q513 and N492 hydrogen bonds. N414 is also shown which hydrogen bonds to backbone carbonyl of Q513. B) Room temperature structure of lit state AsLOV2 (PDB 2V1B, resolution 1.55 Å) indicating hydrogen bonds extending from FMN-O4 highlighting elongated N492 hydrogen bond. The two conformers of Q513 do not form hydrogen bonds with FMN-O4 or now protonated N5. Instead, one of the Q513 conformers appears to reposition in close proximity to the secondary conformer of N414 with distance of only 2.3 Å, too short for a hydrogen bond most likely indicating partial residue dual occupancy.

Overall, by using <sup>19</sup>F 1D NMR, light dependent chemical shifts were observed for 5fW491 and 5fW557 suggesting that <sup>19</sup>F NMR could be applied to rapidly investigate novel construct light responsiveness to assist in structural investigations of engineered optogenetic tools. Experiments requiring the use of U-<sup>15</sup>N/ <sup>13</sup>C labelling require prior peak resonance assignments to pinpoint structural changes and are restricted by protein size where large constructs cannot be easily investigated. Although CD could be applied in a similar way, potential issues could arise. If no CD spectra changes occur upon the photoactivation it cannot be concluded if Jα is folded or unfolded. Additionally, if the fusion protein is large in size, for instance 2-4 times the size of AsLOV2, net changes in the mean residue ellipticity could be too small to be observed with certainty. Considering the low cost, lack of size restriction, facile implementation of <sup>19</sup>F labelling and <sup>19</sup>F NMR experiments, this opens yet another way to study LOV domain proteins.

## 4.5 References

1. Losi, A., Gardner, K. H. & Möglich, A. Blue-Light Receptors for Optogenetics. *Chem. Rev.* **118**, 10659-10709 (2018).
2. Lungu, O. I., Hallett, R. A., Choi, E. J., Aiken, M. J., Hahn, K. M. & Kuhlman, B. Designing Photoswitchable Peptides Using the AsLOV2 Domain. *Chem. Biol.* **19**, 507-517 (2012).
3. Strickland, D., Yao, X., Gawlak, G., Rosen, M. K., Gardner, K. H. & Sosnick, T. R. Rationally Improving LOV Domain-Based Photoswitches. *Nat. Methods* **7**, 623-626 (2010).
4. Wu, Y. I., Frey, D., Lungu, O. I., Jaehrig, A., Schlichting, I., Kuhlman, B. & Hahn K. M. A genetically Encoded Photoactivatable Rac Controls the Motility of Living Cells. *Nature* **461**, 104-108 (2009).
5. Lee, J., Natarajan, M., Nashine, V. C., Socolicj, M., Russ, W. P., Benkovic, S. J. & Ranganathan, R. Surface Sites for Engineering Allosteric Control in Proteins. *Science* **322**, 438-442 (2008).
6. Lee, D., Hyun, J. H., Jung, K., Hannan, P. & Kwon, H. B. A Calcium- and Light-Gated Switch to Induce Gene Expression in Activated Neurons. *Nat. Biotechnol.* **35**, 858-863 (2017).
7. Wang, W., Wildes, C. P., Pattarabanjird, T., Sanchez, M. I., Glober, G. F., Matthews, G. A., Tye, K. M. & Ting, A. Y. A light- and calcium-gated transcription factor for imaging and manipulating activated neurons. *Nat. Biotechnol.* **35**, 864-871 (2017).
8. Harper, S. M., Neil, L. C. & Gardner, K. H. Structural Basis of a Phototropin Light Switch. *Science* **301**, 1541-15444 (2003).
9. Halavaty, A. S. & Moffat, K. N- and C-terminal Flanking Regions Modulate Light-Induced Signal Transduction in the LOV2 Domain of the Blue Light Sensor Phototropin 1 from *Avena sativa*. *Biochemistry* **46**, 14001-14009 (2007).
10. Nash, A. I., Ko, W. H., Harper, S. M. & Gardner, K. H. A Conserved Glutamine Plays a Central Role in LOV Domain Signal Transmission and Its Duration. *Biochemistry* **47**, 13842-13849 (2008).
11. Vaidya, A. T., Chen, C. H., Dunlap, J. C., Loros, J. J. & Crane, B. R. Structure of a Light-Activated LOV Protein Dimer that Regulates Transcription. *Sci. Signal.* **4**, ra50 (2011).
12. Yee, E. F., Dienstrhuber, R. P., Vaidya, A. T., Borbat, P. P., Engelhard, C., Freed, J. H., Bittl, R., Möglich, A. & Crane, B. R. Signal Transduction in Light-

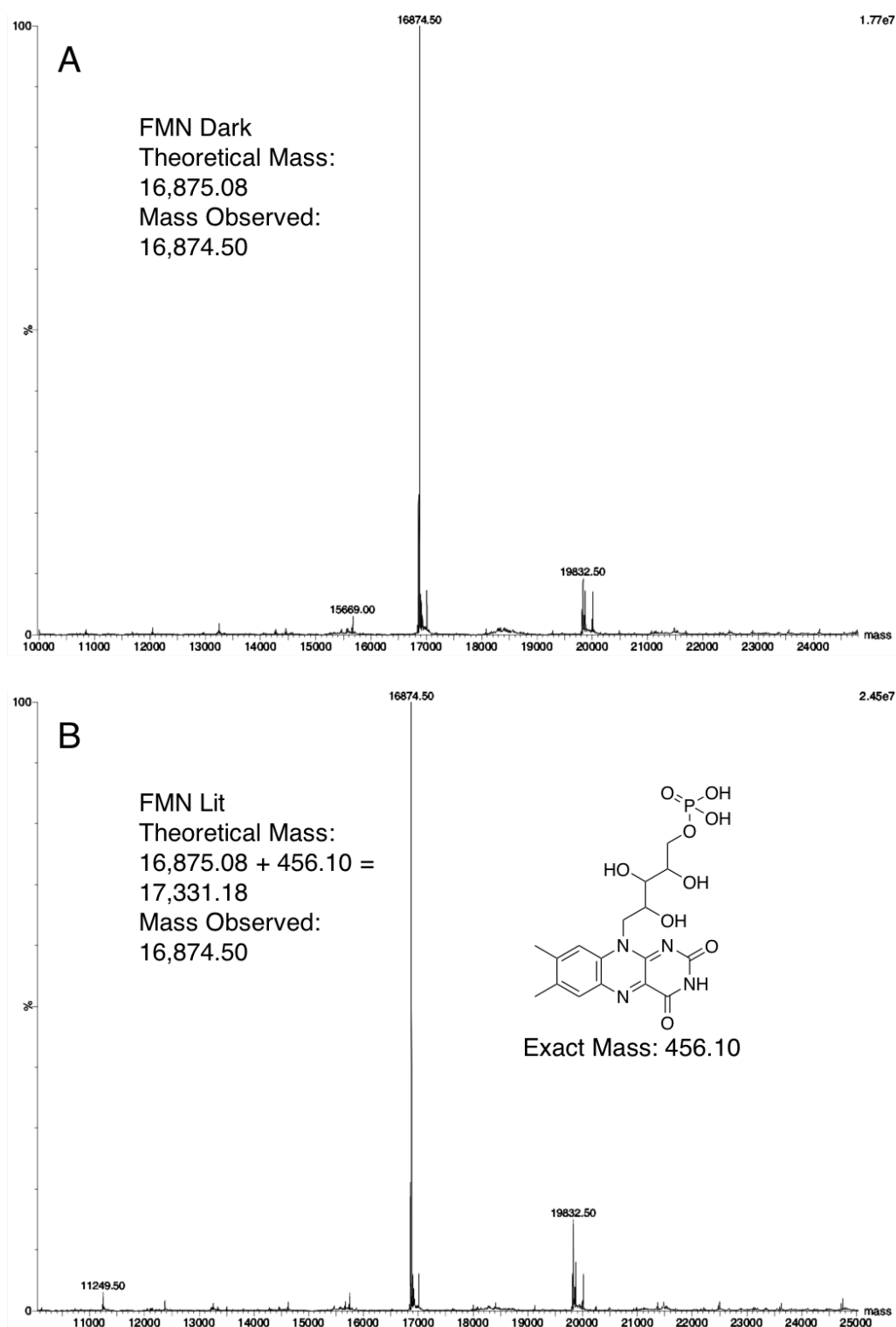
- Oxygen-Voltage Receptors Lacking the Adduct-Forming Cysteine Residue. *Nat. Commun.* **6**, 10079 (2015).
13. Ganguly, A., Thiel, W. & Crane, B. R. Glutamine Amide Flip Elicits Long Distance Allosteric Responses in the LOV Protein Vivid. *J. Am. Chem. Soc.* **139**, 2972-2980 (2017).
  14. Alexandre, M. T. A., Van Grondelle, R., Hellingwerf, K. J. & Kennis, J. T. M. Conformational Heterogeneity and Propagation of Structural Changes in the LOV2/J $\alpha$  Domain from *Avena sativa* Phototropin 1 as Recorded by Temperature-Dependent FTIR Spectroscopy. *Biophys. J.* **97**, 238-247 (2009).
  15. Zayner, J. P., Antoniou, C. & Sosnick, T. R. The Amino-Terminal Helix Modulates Light-Activated Conformational Changes in As LOV2. *J. Mol. Biol.* **419**, 61-74 (2012).
  16. Freddolino, P. L., Gardner, K. H. & Schulten, K. Signaling mechanisms of LOV domains: New Insights from Molecular Dynamics Studies. *Photochem. Photobiol. Sci.* **12**, 1158-1170 (2013).
  17. Gil, A. A. *et al.* Femtosecond to Millisecond Dynamics of Light Induced Allostery in the *Avena sativa* LOV Domain. *J. Phys. Chem. B* **121**, 1010-1019 (2017).
  18. Mansurova, M., Scheercousse, P., Simon, J., Kluth, M. & Gärtner, W. Chromophore Exchange in the Blue Light-Sensitive Photoreceptor YtvA from *Bacillus subtilis*. *ChemBioChem* **12**, 641–646 (2011).
  19. Kalvaitis, M. E., Johnson, L. A., Mart, R. J., Rizkallah, P. & Allemann, R. K. A Noncanonical Chromophore Reveals Structural Rearrangements of the Light-Oxygen-Voltage Domain upon Photoactivation. *Biochemistry* **58**, 2608-2616 (2019).
  20. Curtis-Marof, R., Doko, D., Rowe, M. L., Richards, K. L., Williamson, R. A. & Howards, M. J. <sup>19</sup>F NMR Spectroscopy Monitors Ligand Binding to Recombinantly Fluorine-Labelled b'x from Human Protein Disulphide Isomerase (hPDI). *Org. Biomol. Chem.* **12**, 3808-3812 (2014).
  21. Bann, J. G., Pinkner, J., Hultgren, S. J. & Frieden, C. Real-Time and Equilibrium (19)F-NMR Studies Reveal the Role of Domain-Domain Interactions in the folding of the Chaperone PapD. *PNAS* **99**, 709-714 (2002).
  22. Aramini, J. M., Hamilton, K., Ma, L. C., Swapna, G. V. T., Leonard, P. G., Ladbury, J. E., Krug, R. M. & Montelione, G. T. (19)F NMR Reveals Multiple Conformations at the Dimer Interface of the Nonstructural Protein 1 Effector Domain from Influenza A Virus. *Structure* **22**, 515-525 (2014).

23. Lian, C., Le, H., Montez, B., Patterson, J., Harrell, S., Laws, D., Matsumura, I., Pearson, J. & Oldfield, E. Fluorine-19 Nuclear Magnetic Resonance Spectroscopic Study of Fluorophenylalanine- and Fluorotryptophan-Labeled Avian Egg White Lysozymes. *Biochemistry* **33**, 5238-5245 (1994).
24. Hoeltzli, S. D. & Frieden, C. <sup>19</sup>F NMR Spectroscopy of [6-<sup>19</sup>F]Tryptophan-Labeled *Escherichia coli* Dihydrofolate Reductase: Equilibrium Folding and Ligand Binding Studies. *Biochemistry* **33**, 5502-5509 (1994).
25. Pomerantz, W. C., Wang, N., Lipinski, A. K., Wang, R., Cierpicki, T. & Mapp, A. K. Profiling the Dynamic Interfaces of Fluorinated Transcription Complexes for Ligand Discovery and Characterization. *ACS Chem. Biol.* **7**, 1345-1350 (2012).
26. Oyala, P. H., Ravichandran, K. R., Funk, M. A., Stucky, P. A., Stich, T. A., Drennan, C. L., Britt, R. D. & Stubble, J. Biophysical Characterization of Fluorotyrosine Probes Site-Specifically Incorporated into Enzymes: *E. coli* Ribonucleotide Reductase As an Example. *J. Am. Chem. Soc.* **138**, 7951-7964 (2016).
27. Manglik, A. *et al.* Structural Insights Into the Dynamic Process of  $\beta$ 2-Adrenergic Receptor Signaling. *Cell* **161**, 1101-1111 (2015).
28. Liu, J. J., Horst, R., Katritch, V., Stevens, R. C. & Wüthrich, K. Biased Signaling Pathways in  $\beta$ 2-Adrenergic Receptor Characterized by <sup>19</sup>F-NMR. *Science* **335**, 1106-1110 (2012).
29. O'Hagan, D. & Rzepa, H. S. Some Influences of Fluorine in Bioorganic Chemistry. *Chem. Comm.* **7**, 645-652 (1997).
30. Buer, B. C. & Marsh, E. N. G. Fluorine: A New Element in Protein Design. *Protein Sci.* **21**, 453-462 (2012).
31. Marsh, E. N. G. & Suzuki, Y. Using <sup>19</sup>F NMR to Probe Biological Interactions of Proteins and Peptides. *ACS Chem. Biol.* **9**, 1242-1250 (2014).
32. Evanics, F., Bezsonova, I., Marsh, J., Kitevski, J. L., Forman-Kay, J. D. & Prosser, R. S. Tryptophan Solvent Exposure in Folded and Unfolded States of an SH3 Domain by <sup>19</sup>F and <sup>1</sup>H NMR. *Biochemistry* **45**, 14120-14128 (2006).
33. Fujisawa, T., Takeuchi, S., Masuda, S. & Tahara, T. Signaling-State Formation Mechanism of a BLUF Protein PapB from the Purple Bacterium *Rhodospseudomonas palustris* Studied by Femtosecond Time-Resolved Absorption Spectroscopy. *J. Phys. Chem. B* **118**, 14761-14773 (2014).
34. Lukacs, A. *et al.* BLUF Domain Function Does Not Require a Metastable Radical Intermediate State. *J. Am. Chem. Soc.* **36**, 4605-4615 (2014).



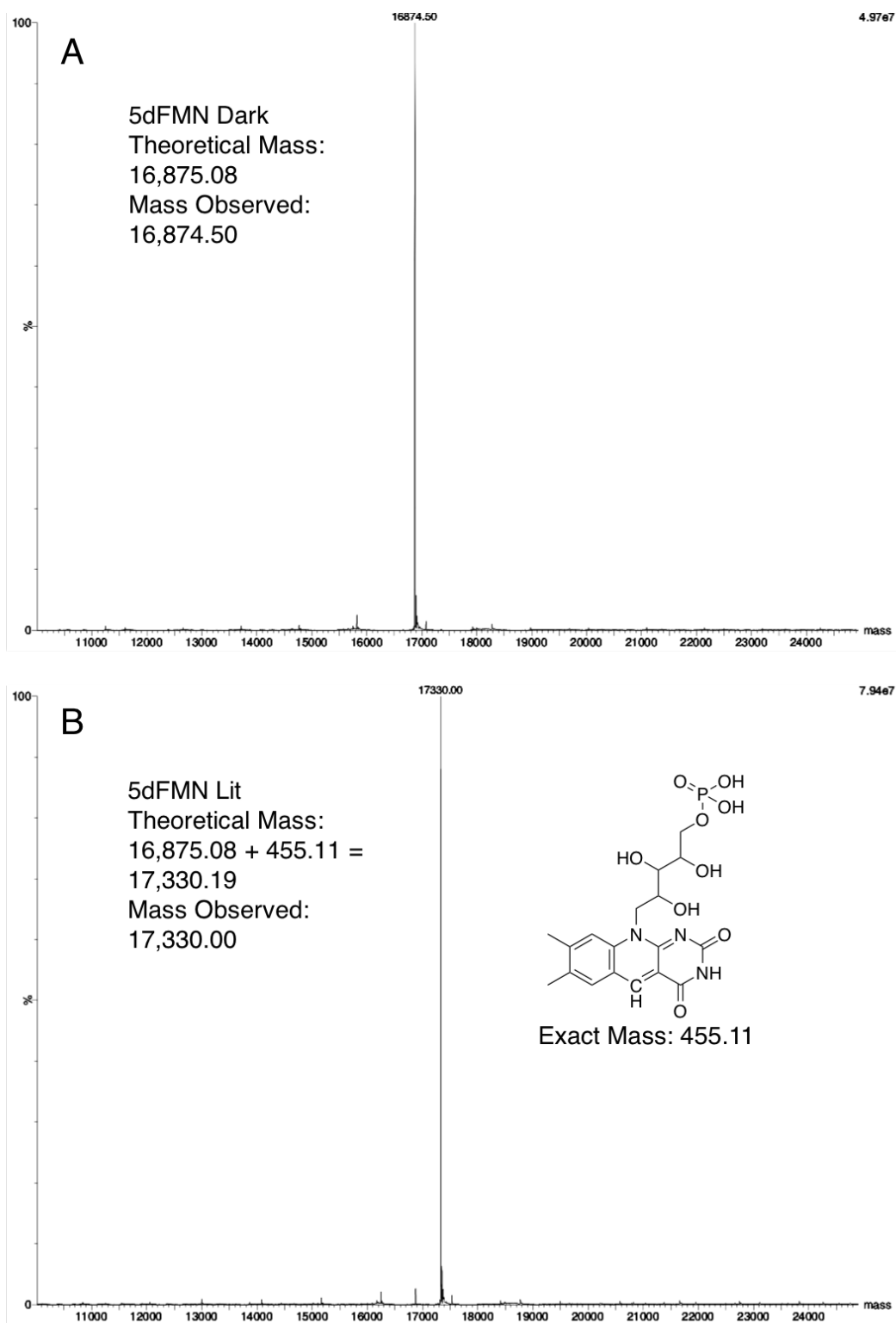
35. Pudasaini, A., Shim, J. S., Song, Y. H., Shi, H., Kiba, T., Somers, D. E., Imaizumi, T. & Zoltowski, B. D. Kinetics of the LOV Domain of ZEITLUPE Determine Its Circadian Function in Arabidopsis. *Elife* **6**, e21646 (2017).
36. Röllén, K. *et al.* Signaling States of a Short Blue-Light Photoreceptor Protein PpSB1-LOV Revealed from Crystal Structures and Solution NMR Spectroscopy. *J. Mol. Biol.* **428**, 3721-3736 (2016).
37. Nazarenko, V. *et al.* Thermostable Flavin-Based Fluorescent Protein from *Chloroflexus aggregans*: a Framework for Ultra-High Resolution Structural Studies. *Photochem. Photobiol. Sci.* **18**, 1793-1805 (2019).
38. Heintz, U. & Schlichting, I. Blue Light-Induced LOV Domain Dimerization Enhances the Affinity of Aureochrome 1a for Its Target DNA Sequence. *Elife* **5**, e11860 (2016).
39. Akiyama, Y., Nakasone, Y., Nakatani, Y., Hisatomi, O. & Terazima, M. Time-Resolved Detection of Light-Induced Dimerization of Monomeric Aureochrome-1 and Change in Affinity for DNA. *J. Phys. Chem. B* **120**, 7360–7370 (2016).
40. Banerjee, A., Herman, E., Kottke, T. & Essen, L. O. Structure of a Native-like Aureochrome 1a LOV Domain Dimer from *Phaeodactylum tricornutum*. *Structure* **24**, 171–178 (2016).
41. Yao, X., Rosen, M. K. & Gardner, K. H. Estimation of the Available Free Energy in a LOV2-J Alpha Photoswitch. *Nat. Chem. Bio.* **4**, 491-497 (2008).
42. Koyama, T., Iwata, T., Yamamoto, A., Sato, Y., Matsuoka, D., Tokutomi, S. & Kandori, H. Different Role of the Jalpha Helix in the Light-Induced Activation of the LOV2 Domains in Various Phototropins. *Biochemistry* **48**, 7621–7628 (2009).
43. Mart, R. J., Meah, D. & Allemann, R. K. Photocontrolled Exposure of Pro-apoptotic Peptide Sequences in LOV Proteins Modulates Bcl-2 Family Interactions. *ChemBioChem* **17**, 698-701 (2016).

## 4.6 Appendix



**Figure 4.15 Liquid chromatography mass spectroscopy (LC-MS) analysis of FMN AsLOV2.**

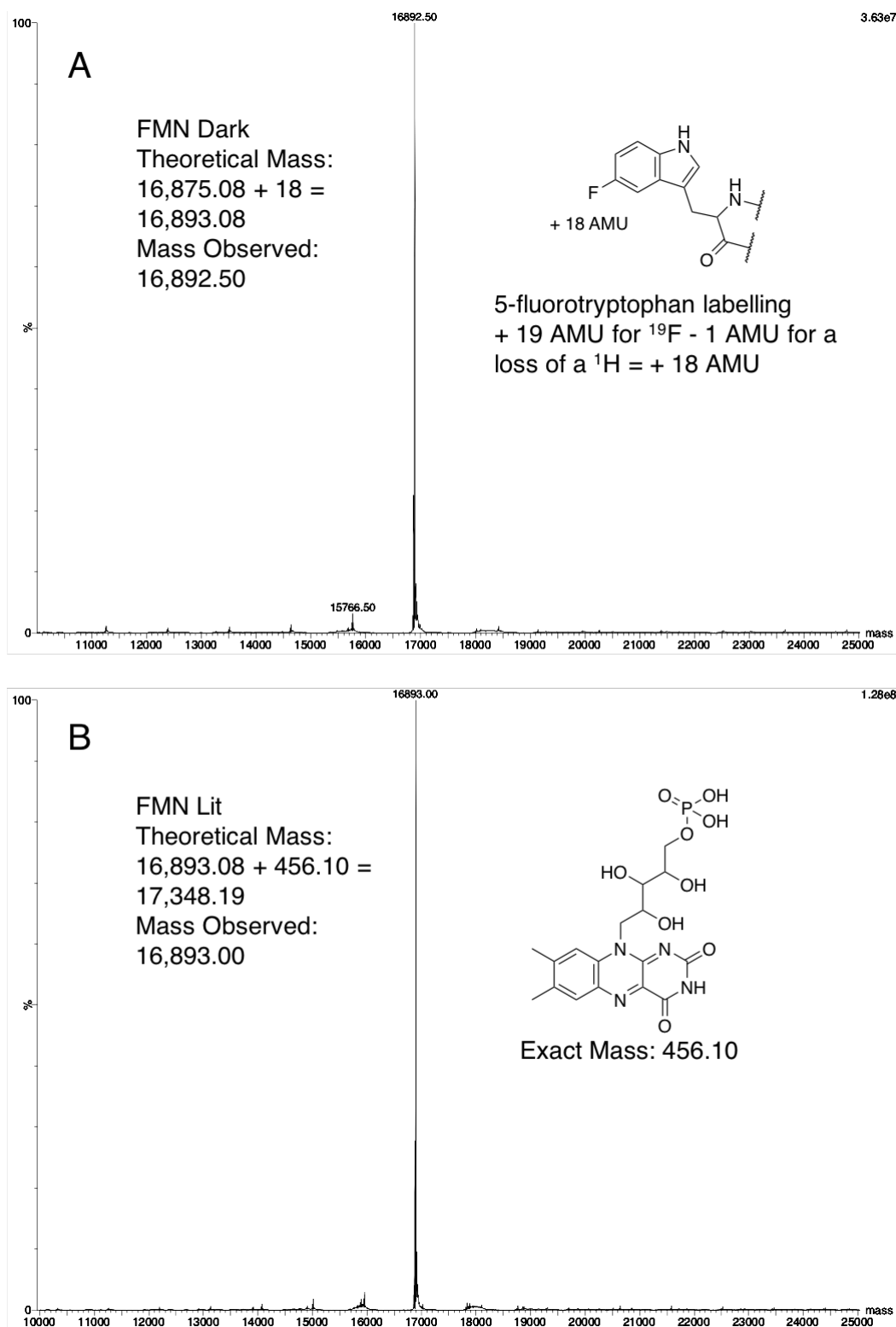
A) Indicates results of dark state protein with observed mass 16,874.5 atomic mass units (AMU), 0.5 AMU away from the theoretical mass of 16,875.08 determined from protein amino acid sequence. B) Indicates results of lit state protein sample yielding a mass of 16,874.50 AMU suggestive of no cysteinyl-FMN covalent adduct with a theoretical mass of  $16,875 + 456$  (exact mass of FMN 456.10 AMU) = 17,331 AMU. Please note error of LC-MS is 1.0 AMU.



**Figure 4.16 Liquid chromatography mass spectroscopy (LC-MS) analysis of 5dFMN AsLOV2.**

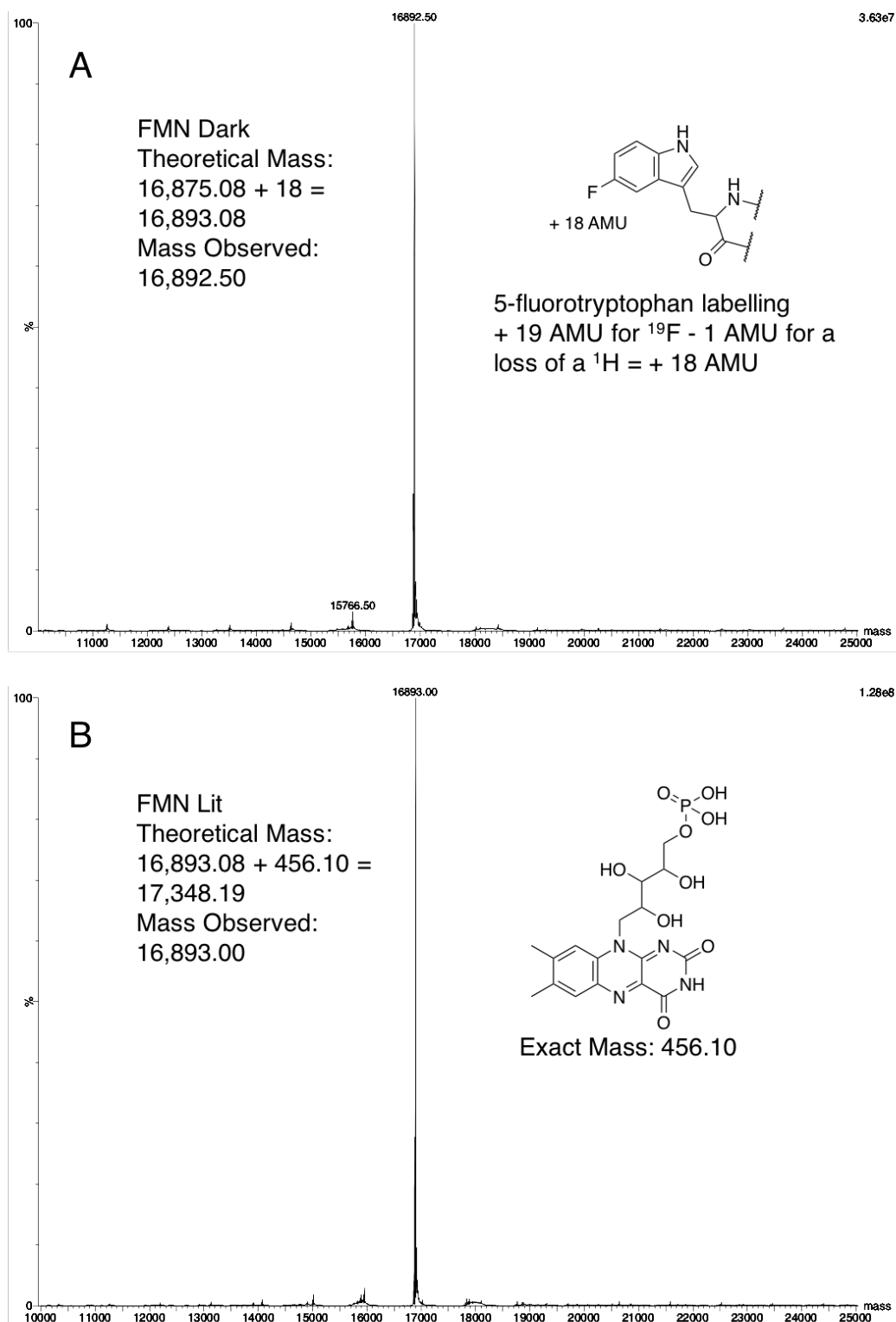
A) Indicates results of dark state protein sample yielding a protein with a mass of 16,874.5 AMU, 0.5 AMU away from the theoretical mass of 16,875 determined from protein amino acid sequence.

B) Indicates results of lit state protein sample yielding a protein with a mass of 17,330 AMU suggestive of cysteinyl-5dFMN covalent adduct with a theoretical mass of 16,875 + 455 (exact mass of 5dFMN 455.11 AMU) = 17,330 AMU. Please note that the error of this LC-MS analysis is 1.0 AMU.



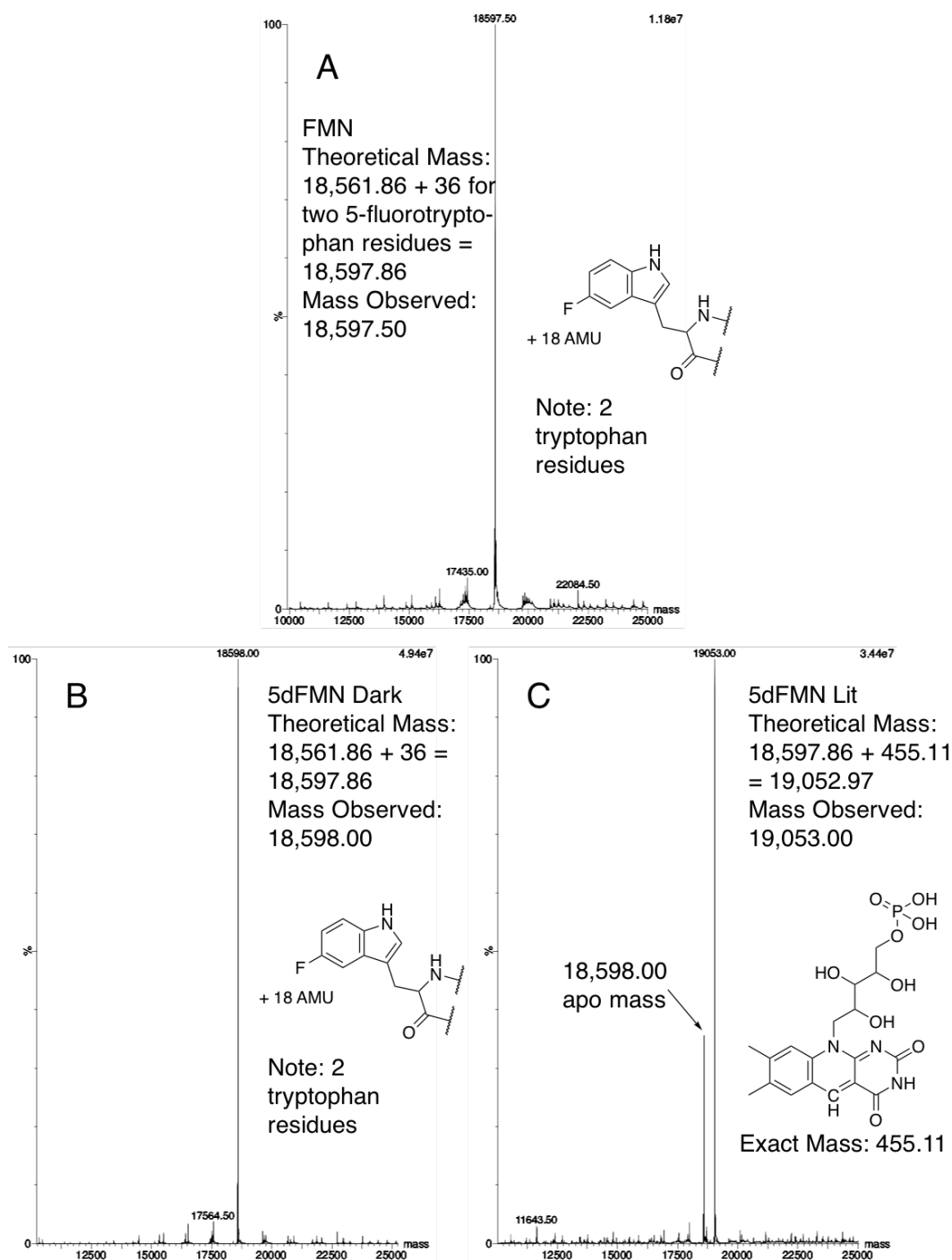
**Figure 4.17 Liquid chromatography mass spectroscopy (LC-MS) analysis of FMN containing AsLOV2 protein labelled with 5fW.**

A) Indicates results of dark state protein sample yielding a protein with a mass of 16,892.50 AMU, 0.5 AMU away from the theoretical mass of 16,893.08 determined from protein amino acid sequence. B) Indicates results of lit state sample yielding a protein with a mass of 16,893.00 AMU suggestive of no cysteinyl-FMN covalent adduct with a theoretical mass of  $16,893 + 456$  (exact mass of FMN 456.10 AMU) = 17,349 AMU. Please note error of LC-MS is 1.0 AMU.



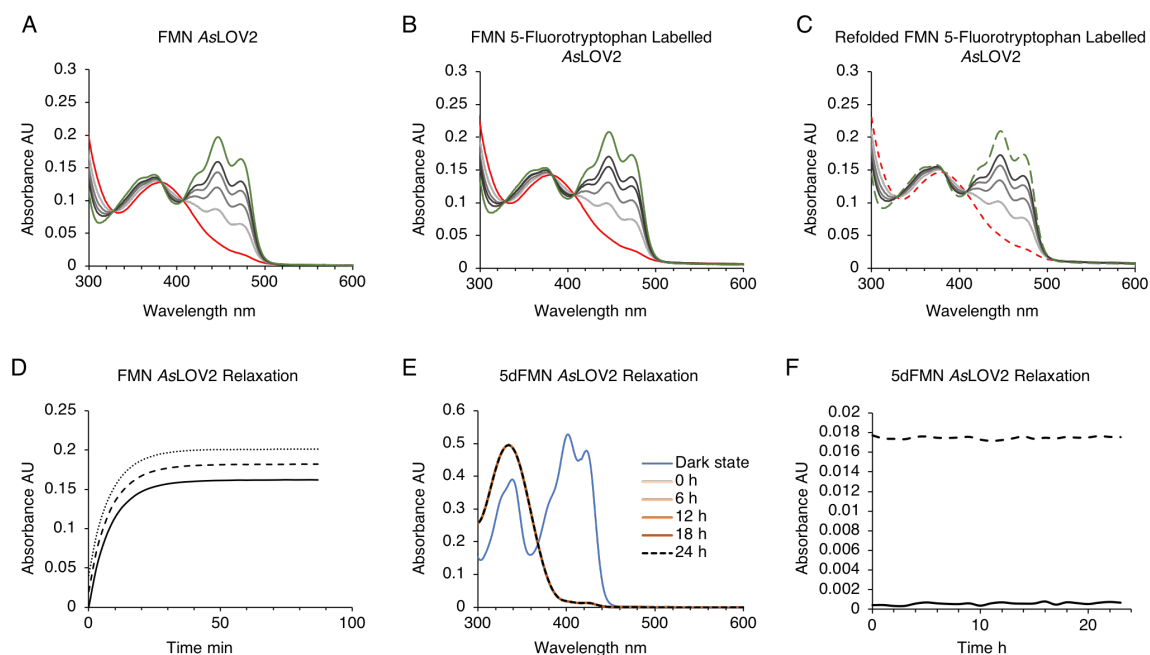
**Figure 4.18 Liquid chromatography mass spectroscopy (LC-MS) analysis of 5dFMN containing AsLOV2 labelled with 5fW.**

A) Dark state protein indicating a protein with a mass of 16,892.50 AMU, 0.5 AMU away from the theoretical mass of 16,893.08 determined from the protein amino acid sequence. B) Indicates results of lit state sample yielding a protein with a mass of 17,348.00 AMU suggestive of cysteinyl-5dFMN covalent adduct with a theoretical mass of  $16,893 + 455$  (exact mass of 5dFMN 455.11 AMU) = 17,348 AMU. Please note error of LC-MS is 1.0 AMU.



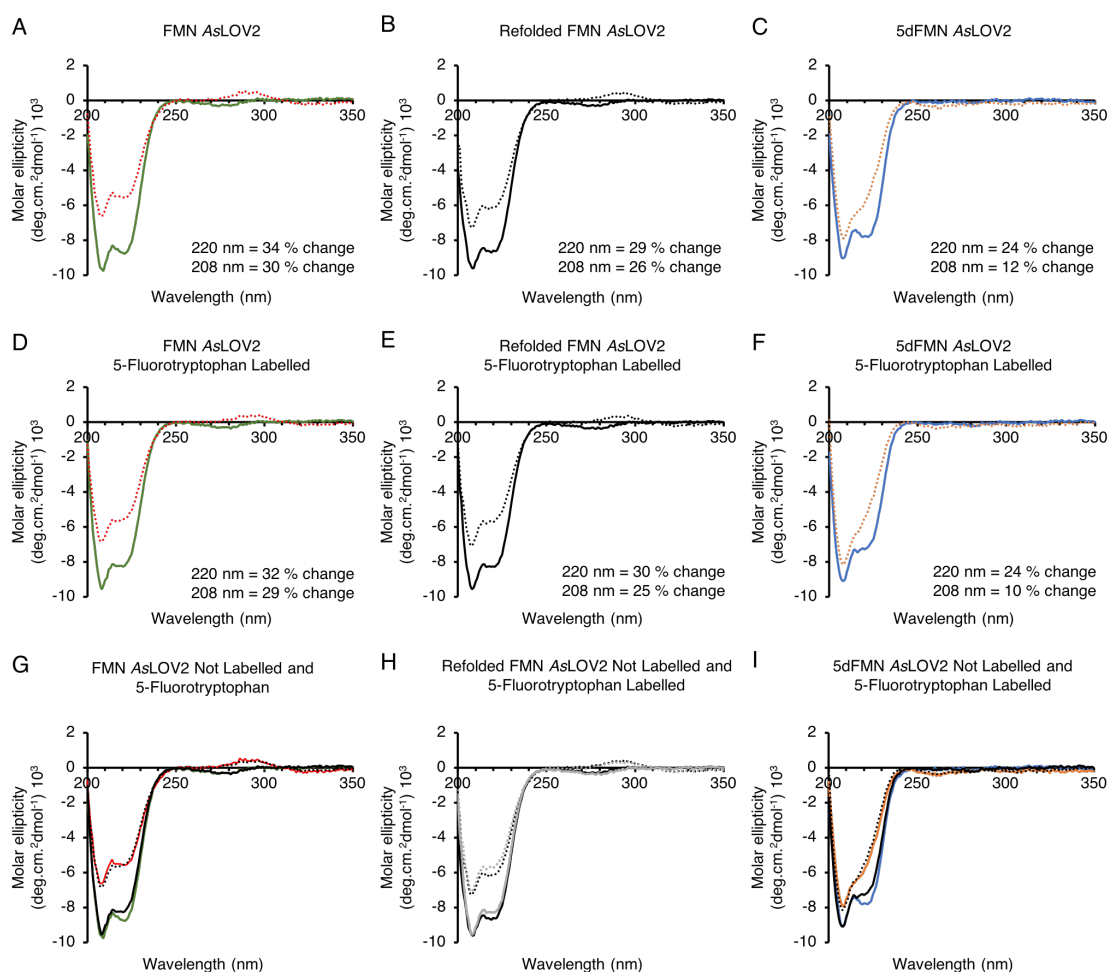
**Figure 4.19 Liquid chromatography mass spectroscopy (LC-MS) analysis of FMN and 5dFMN containing AsLOV2<sub>FL</sub> proteins labelled with 5fW.**

Please note that protein contained two 5fW residues. A) Indicates results of dark state FMN protein sample yielding a protein with a mass of 18,597.50 AMU. B) Indicates results of dark state 5dFMN sample yielding a protein with a mass of 18,598.00 AMU suggestive of no cysteinyl-5dFMN covalent adduct. C) Lit state results with a theoretical mass of  $18,598.00 + 455$  (exact mass of 5dFMN 455.11 AMU) = 19,053 AMU. Please note that higher degree, ~ 40%, of apo protein was detected suggestive of not fully photoactivated protein. Error of LC-MS is 1.0 AMU.



**Figure 4.20 UV-Vis characterisation of FMN and 5dFMN containing AsLOV2.**

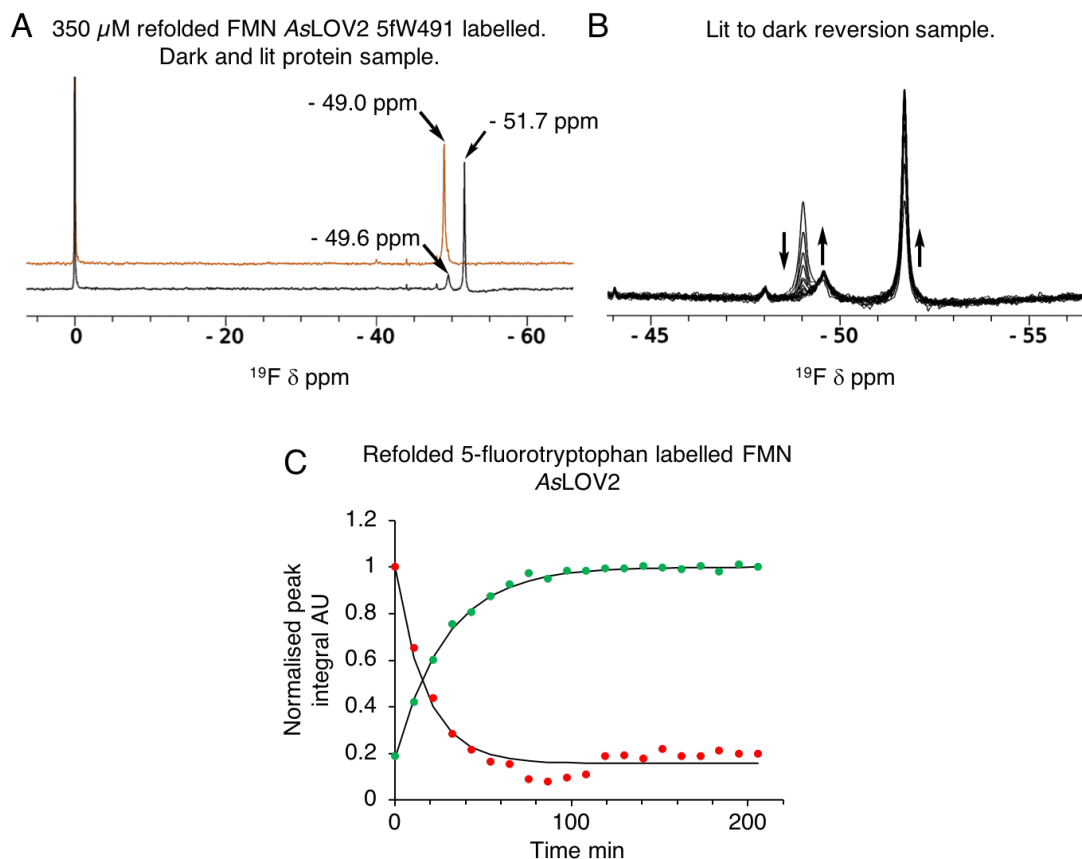
For FMN samples protein concentration was  $12.5 \mu\text{M}$  and individual scans taken every 3 min for 87 min. For clarity, only 0, 3, 6, 9, 12 and 87 min scans are shown in the corresponding figures. For 5dFMN AsLOV2,  $50 \mu\text{M}$  sample was used to investigate refolding methodology to detect FMN. All protein samples were illuminated with 450 nm LED lights until a photostationary phase was observed. A) FMN AsLOV2 relaxing from lit, red solid line, to dark, green solid line. Relaxation half-life time was 5.7 min. B) FMN AsLOV2 5-fluorotryptophan labelled relaxing from lit, red solid line, to dark, green solid line. Relaxation half-life time was 5.7 min. C) Refolded FMN AsLOV2 relaxing from lit, red dashed line, to dark, green dashed line. Relaxation time was 5.6 min. D) Overlapped first order decay of spectra of A) B) and C) at 447 nm maxima absorbance. For clarity decay curves were offset by 0.025 AU. E) Investigation of lit state 5dFMN AsLOV2 stability, with blue solid line indicating dark state spectra and orange solid lines plus 24 h time point dashed black line indicate lit state. No relaxation was evident by UV-Vis experiments. F) Plot of 5dFMN AsLOV2 stability over time course to investigate potential FMN contamination, dashed line indicates absorbance at 401 nm (5dFMN maxima) and solid line absorbance at 447 nm (FMN maxima), indicating no major changes at 447 nm absorbance.



**Figure 4.21 CD studies of AsLOV2 containing FMN or 5dFMN as a cofactor.**

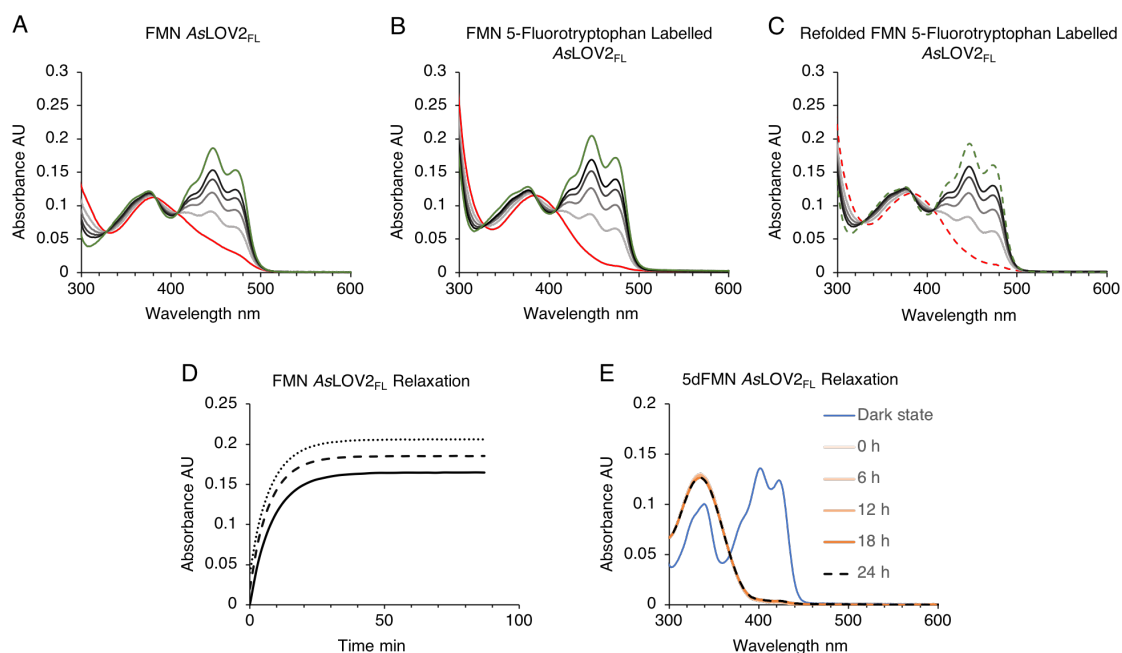
Please note dead time to reach 220 nm wavelength was 3 min and 30 s. Changes at 220 and 208 nm between dark and lit states are indicated in the figure. A) FMN AsLOV2 switch from dark (green solid line) to lit (red dots) state. B) Refolding FMN AsLOV2 control switch from dark (black solid line) to lit state (black dots). C) 5dFMN AsLOV2 dark (blue solid line) and lit (orange dots) states. D) FMN AsLOV2 labelled with 5fW switch from dark (green solid line) to lit (red dots) states. E) Refolded FMN AsLOV2 labelled with 5fW control switch from dark (black solid line) to lit (black dots) states. F) 5dFMN AsLOV2 labelled with 5fW switch from dark (blue solid line) to lit (orange dots) states. G) Overlapped FMN AsLOV2 (dark state green solid line and lit state red solid line) and FMN AsLOV2 labelled with 5fW (dark state solid black line and lit state black dots). H) Overlapped refolding control of FMN AsLOV2 (dark state solid black line and lit state black dots) and FMN AsLOV2 labelled with 5fW (dark state solid grey line and lit state grey dots). I) Overlapped 5dFMN AsLOV2 (dark state blue solid line and lit state solid orange line) and 5dFMN AsLOV2 labelled with 5fW (dark state solid black line and lit state black dots). Please note that in G) H) and I) the spectra indicate minor differences in CD spectra at 220 nm that could be a product of 5fW labelling or due to the experimental error.





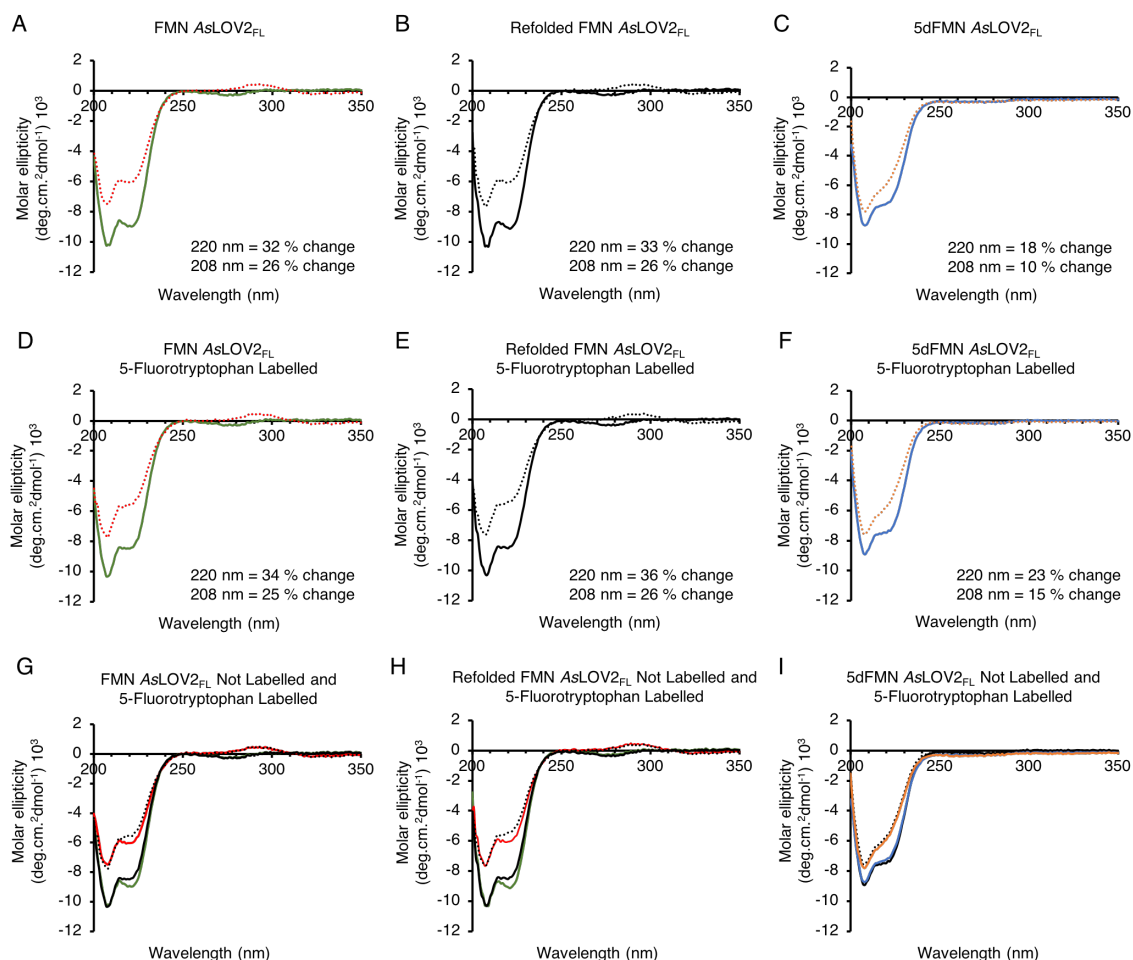
**Figure 4.22**  $^{19}\text{F}$  1D NMR characterisation of refolded 5fW labelled FMN AsLOV2.

A)  $^{19}\text{F}$  NMR spectra of refolded FMN AsLOV2, with dark spectra in black and lit in orange, respectively. Observed peaks are labelled with arrows. B)  $^{19}\text{F}$  NMR relaxation of refolded AsLOV2 from lit to dark state, respectively, with arrows indicating corresponding peak decay (down) or growth (up). C)  $^{19}\text{F}$  NMR relaxation kinetics of refolded AsLOV2 with dots in green indicating growth of dark peak at  $\delta$  -51.7 ppm, half-life 21.1 min, and red dots decay of lit peak at  $\delta$  -49.0 ppm, half-life of 14.1 min.



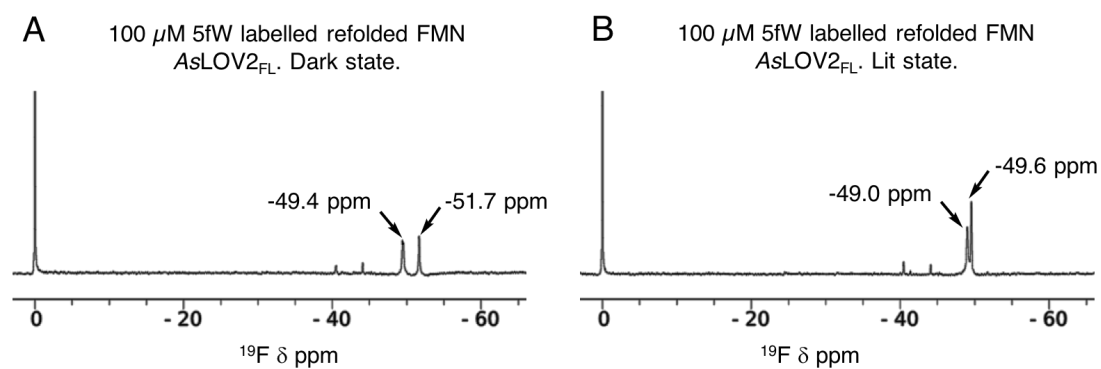
**Figure 4.23 UV-Vis characterisation of AsLOV2<sub>FL</sub>.**

For FMN samples protein concentration was 12.5  $\mu\text{M}$  and individual scans taken every 3 min for 87 min. For clarity, only 0, 3, 6, 9, 12 and 87 min scans are shown in the corresponding figures. All protein samples were illuminated with 450 nm LED lights until a photostationary phase was achieved. A) FMN AsLOV2<sub>FL</sub> relaxing from lit, red solid line, to dark, green solid line. Relaxation half-life time was 5.8 min. B) FMN AsLOV2<sub>FL</sub> 5-fluorotryptophan labelled relaxing from lit, red solid line, to dark, green solid line. Relaxation half-life time was 5.1 min. C) Refolded and 5fW labelled FMN AsLOV2<sub>FL</sub> relaxing from lit, red dashed line, to dark, green dashed line. Relaxation time was 5.2 min. D) Overlapped first order decay of spectra of A), B) and C) at 447 nm maxima absorbance. For clarity decay curves were offset by 0.025 AU. E) Investigation of lit state 5dFMN AsLOV2<sub>FL</sub> stability, with blue solid line indicating dark state spectra and orange solid lines plus 24 h time point dashed black line indicate lit state. No relaxation was evident by UV-Vis experiments.



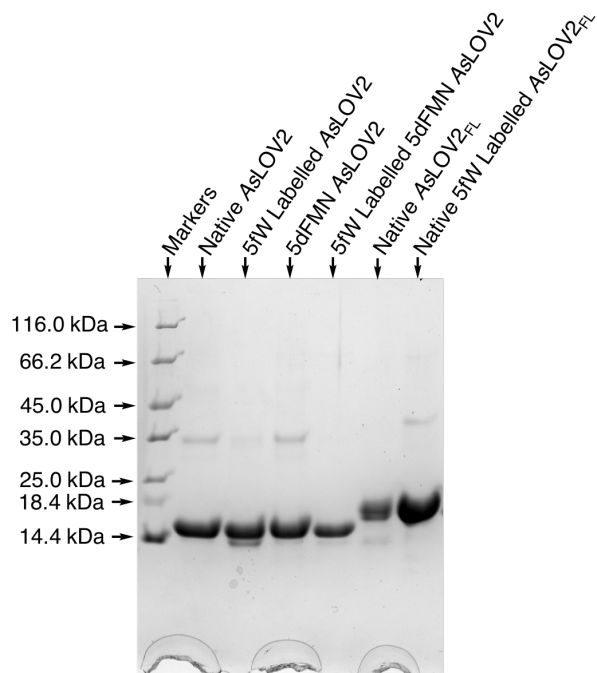
**Figure 4.24 CD studies of AsLOV2<sub>FL</sub> containing FMN or 5dFMN as a cofactor.**

Please note dead time to reach 220 nm wavelength was 3 min and 30 s. Changes at 220 and 208 nm between dark and lit states are indicated in the figure. A) FMN AsLOV2<sub>FL</sub> switch from dark (green solid line) to lit (red dots) state. B) Refolding FMN AsLOV2<sub>FL</sub> control switch from dark (black solid line) to lit state (black dots). C) 5dFMN AsLOV2<sub>FL</sub> dark (blue solid line) and lit (orange dots) states. D) FMN AsLOV2<sub>FL</sub> labelled with 5fW switch from dark (green solid line) to lit (red dots) states. E) Refolded FMN AsLOV2<sub>FL</sub> labelled with 5fW control switch from dark (black solid line) to lit (black dots) states. F) 5dFMN AsLOV2<sub>FL</sub> labelled with 5fW switch from dark (blue solid line) to lit (orange dots) states. G) Overlapped FMN AsLOV2<sub>FL</sub> (dark state green solid line and lit state red solid line) and FMN AsLOV2<sub>FL</sub> labelled with 5fW (dark state solid black line and lit state black dots). H) Overlapped refolding control of FMN AsLOV2<sub>FL</sub> (dark state solid green line and lit state solid red line) and FMN AsLOV2<sub>FL</sub> labelled with 5fW (dark state solid black line and lit state black dots). I) Overlapped 5dFMN AsLOV2<sub>FL</sub> (dark state blue solid line and lit state solid orange line) and 5dFMN AsLOV2<sub>FL</sub> labelled with 5fW (dark state solid black line and lit state black dots). Please note that in G) H) and I) the spectra indicate minor differences in CD spectra at 220 nm that could be a product of 5fW labelling or due to the experimental error.



**Figure 4.25  $^{19}\text{F}$  1D NMR characterisation of refolded 5fW labelled FMN AsLOV2<sub>FL</sub>.**

$^{19}\text{F}$  NMR spectra of FMN AsLOV2<sub>FL</sub> under A) dark and B) lit states, respectively. All chemical shifts reported in the text and represented in the figure were referenced to the TFA peak adjusted to  $\delta$  0.0 ppm.



**Figure 4.26 SDS-PAGE analysis of purified AsLOV2 and AsLOV2<sub>FL</sub> protein samples.**

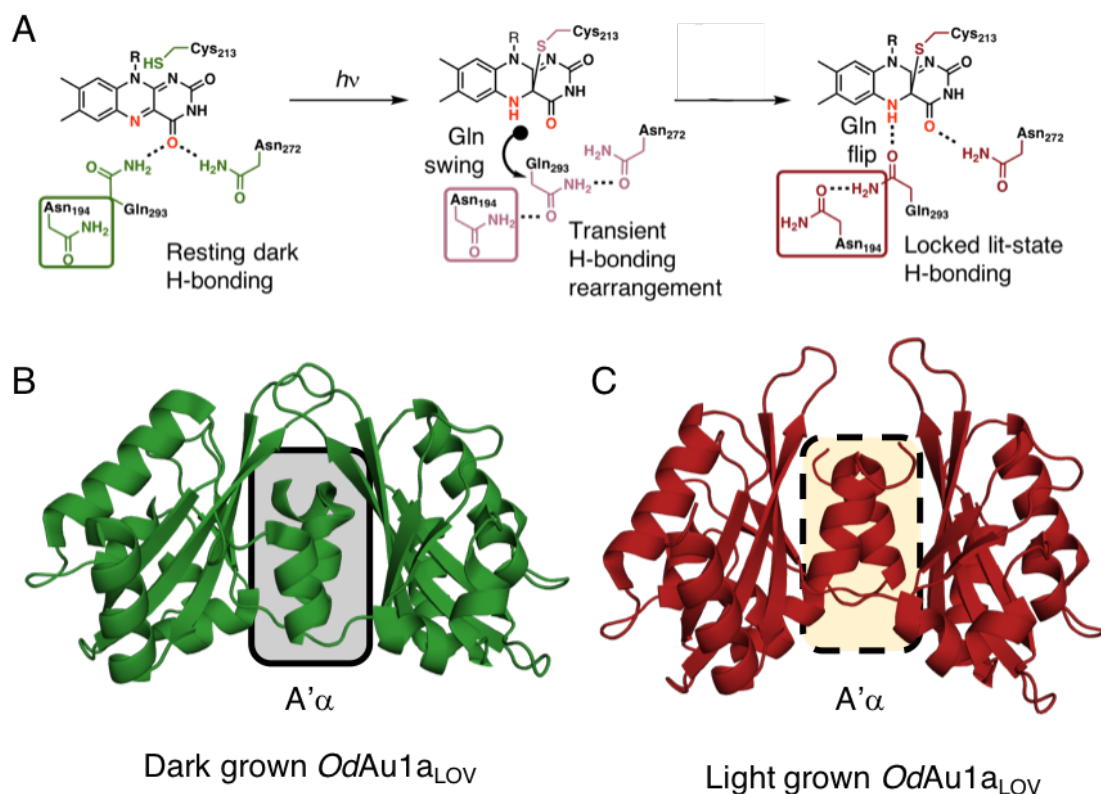
Lanes indicating particular protein samples loaded. Please note that the upper band at around 35 kDa marker for AsLOV2 and below 45 kDa marker for AsLOV2<sub>FL</sub> was probably a dimer generated due to a formation of intermolecular disulphide between C450 residues. Protein markers and corresponding molecular masses are indicated in the figure.



## 5 In Solution Structural Studies of *Ochromonas danica* Full Length Aureochrome1a and Isolated LOV Domain

### 5.1 Introduction

Aureochromes act as blue-light sensing transcription factors binding DNA *via* the N-terminal bZIP domains to regulate acclimatisation to light, photomorphogenesis and the cell cycle<sup>1-5</sup>. Although a number of X-ray crystal structures of Au1a<sub>LOV</sub> have been obtained the large-scale structural changes that drive DNA binding remain undetermined<sup>6-9</sup>. Our previous studies on the isolated *Ochromonas danica* Au1a LOV domain (*OdAu1a*<sub>LOV</sub>) provided structural insights into the residue rearrangements within the FMN binding pocket upon illumination<sup>9</sup>. Structures obtained under the dark, dark then illuminated and lit conditions showed multiple conformations for residues Q293, N272 and N194 defined as dark, 'swing' and 'flip' states, respectively. These structures suggest that formation of the cysteinyl-FMN photoadduct triggers rearrangement of hydrogen bond networks extending from the FMN to the allosteric A' $\alpha$ , Figure 5.1 A. Furthermore, only minor structural differences were evident for the A' $\alpha$  helix, which was observed to rearrange across the LOV domain  $\beta$ -sheet surface between dark and lit states<sup>9</sup>, Figure 5.1 B and C. This therefore raised a question if these structural rearrangements were representative of *OdAu1a*<sub>LOV</sub> dimers in the solution and were not an artefact of different crystallisation condition. To investigate *OdAu1a*<sub>LOV</sub> signalling mechanism and to pinpoint the essential protein regions involved in dictating dimerisation, in this work the solution behaviour of *OdAu1a*<sub>LOV</sub> was investigated by NMR studies. Due to applied <sup>19</sup>F NMR methodology, evidence was provided to suggest that <sup>19</sup>F labelling can be applied in the study of large full length *OdAu1a* construct. Studies on *PtAu1a*<sub>LOV</sub> have demonstrated light dependent dimerisation as observed for *OdAu1a* homologue<sup>7,9-11</sup>. A construct with truncated J $\alpha$  helix ( $\Delta$ J $\alpha$ *PtAu1a*<sub>LOV</sub>) appeared monomeric in both dark and lit states, indicating that the C-terminal helix is essential for dimerisation of *PtAu1a*<sub>LOV</sub><sup>11</sup>. With the A' $\alpha$  helix removed  $\Delta$ A' $\alpha$ *PtAu1a*<sub>LOV</sub> showed light dependent dimerisation at low concentrations, but at elevated concentrations formed only a dimer<sup>10</sup>. Fourier transform infrared spectroscopy (FT-IR) suggested that A' $\alpha$  unfolds upon illumination suggesting that the A' $\alpha$  positioning in the dark state blocks the LOV domain dimerisation surface<sup>10</sup>. This key mechanistic proposal is supported by structural studies of *PtAu1a*<sub>LOV</sub> in dark and lit states indicating A' $\alpha$  rearrangement across the  $\beta$ -sheet surface<sup>7</sup>.



**Figure 5.1 Photoactivation and structure of *Ochromonas danica* Aureochrome1a LOV domain.**

A) Photochemical mechanism of *Ochromonas danica* Au1a<sub>LOV</sub> (*OdAu1a<sub>LOV</sub>*) activation by light. In resting state, Gln293 and Asn272 form hydrogen bonds with the FMN-O4 locking protein in a monomeric state. Upon the photoadduct formation, a planar aromatic isoalloxazine ring system is interrupted unanchoring Gln293 and Asn272 due to weakening of hydrogen bonds to the FMN-O4. This therefore allows Gln293 residue to swing out ('swing state') of the flavin binding pocket forming transient hydrogen bond network with Asn194 and Asn272. Upon dimerisation, the dimer is stabilised by Asn194 and Gln293 hydrogen bond and possible Gln293 side chain oxygen to protonated FMN-N5 ('flip state'). In the figure, allosteric Asn194 residue, located in a latch region between A'α and βA strand is indicated in a box. Possible hydrogen bond networks are indicated in the figure as dashed lines. Structure of B) dark state dimer formed between chains A and C (PDB code: 6I20) and C) lit state unique dimer formed between chains C and D (PDB code: 6I22) *OdAu1a<sub>LOV</sub>* indicating A'α rearrangement across β-sheet surface (labelled with highlighted boxes). Comparisons of chain A of 6I20 (dark state) and chain C of 6I22 (lit state) indicates closely related structures with RMSD value of 0.367 Å (769 to 769 atoms), yet arrangement of A'α helix across β-sheet surface is observed. Notably, Proteins Interfaces Structures and Assemblies (PISA, online software provided by PDBe at [www.ebi.ac.uk/pdbe](http://www.ebi.ac.uk/pdbe)) analysis estimates  $\Delta^iG = -13.2$  kcal/mol for dark grown dimer and  $\Delta^iG = -12.6$  kcal/mol for unique light grown dimer, respectively, suggesting both dimeric states stable in the solution.

Additionally, crystal structures also indicated that repositioned A'α makes residue contacts with Jα helix explaining why  $\Delta J\alpha$ .*PtAu1a<sub>LOV</sub>* was monomeric<sup>7,10</sup>. In agreement with X-ray structure, deuterium exchange mass spectroscopy also indicated a much slower hydrogen to deuterium exchange for the helical A'α section under illuminated in comparison to dark conditions. This therefore suggested that upon lit state dimer formation

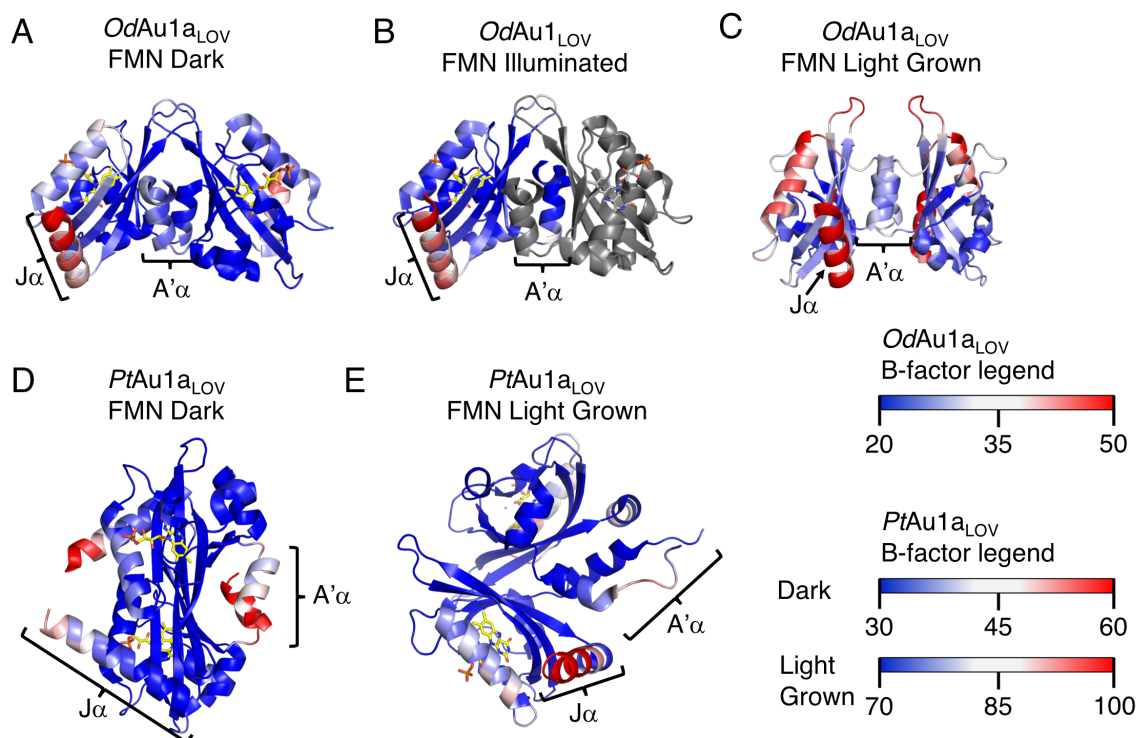


A'α becomes shielded from the solvent<sup>7</sup>. The hydrogen to deuterium exchange was rapid for the Jα helix irrespective of the photoactivation state suggesting that the C-terminus Jα is always accessible to the solvent<sup>7</sup>.

In this work it was decided to investigate the mechanism of *OdAu1a<sub>LOV</sub>* activation by light *via* site directed mutagenesis and to study the importance of auxiliary A'α and Jα helices in the signalling mechanism. NMR backbone assignment of *OdAu1a<sub>LOV</sub>* under dark and lit state conditions (undertaken by Dr Luke Johnson, Cardiff University) have highlighted two unique observations. First of all, the A'α corresponding resonances are not observed in <sup>15</sup>N-<sup>1</sup>H heteronuclear single quantum coherence (HSQC) spectra under lit and dark states. Secondly, experiments under the dark and lit states show negligible chemical shift changes for the Jα helix suggesting that the Jα is uninvolved in dimerisation. If the Jα helix is not essential, deletion of this section would yield a small functioning LOV domain protein (13.1 kDa) that could be applied in the development of LOV domain controlled proteins. In order to understand role of A'α and Jα helices, it was decided to proceed with targeted fluorine (<sup>19</sup>F) labelling for <sup>19</sup>F NMR analysis as conventional NMR methods were impossible. Unlike X-ray crystallography, <sup>19</sup>F NMR allowed us to probe structural changes in the solution within the full-length protein as well.

## 5.2 Assessment of *Ochromonas danica* Aureochrome1a Crystal Structures to Guide Construct Design

Previous X-ray crystallography investigations of *OdAu1a<sub>LOV</sub>* and *PtAu1a<sub>LOV</sub>* suggested that the A'α helix rearranges across β-sheet surface upon illumination<sup>7,9</sup>, Figure 5.2. In order to understand possible changes and to identify more flexible regions, the crystal structures previously obtained were investigated reflecting on B-factors, Figure 5.2. B-factors can be applied to the X-ray scattering term for each atom or a group of atoms including ligands describing a degree to which electron density is spread out<sup>12</sup>. Higher B-factors can be indicative of a static or dynamic mobility within a structure where higher relative values are suggestive of flexible regions<sup>13</sup>. Nonetheless, B-factors must be interpreted with caution due to the dependence on resolution which can be influenced by the percentage solvent content, noises arising from crystal lattice, crystal packing differences, errors in model building and the type of structural refinement resulting in up to 15% errors<sup>14,15</sup>. Tight packing can also artificially increase B-factors and prevent large scale structural changes from taking place as observed for *Halobacterium salinarum* halorhodopsin when switching from dark (ground) to lit (excited) states<sup>16</sup>.



**Figure 5.2 Structural B-factor analysis of *Ochromonas danica* and *Phaeodactylum tricornutum* Aureochrome1a LOV (Au1a<sub>LOV</sub>) domains.**

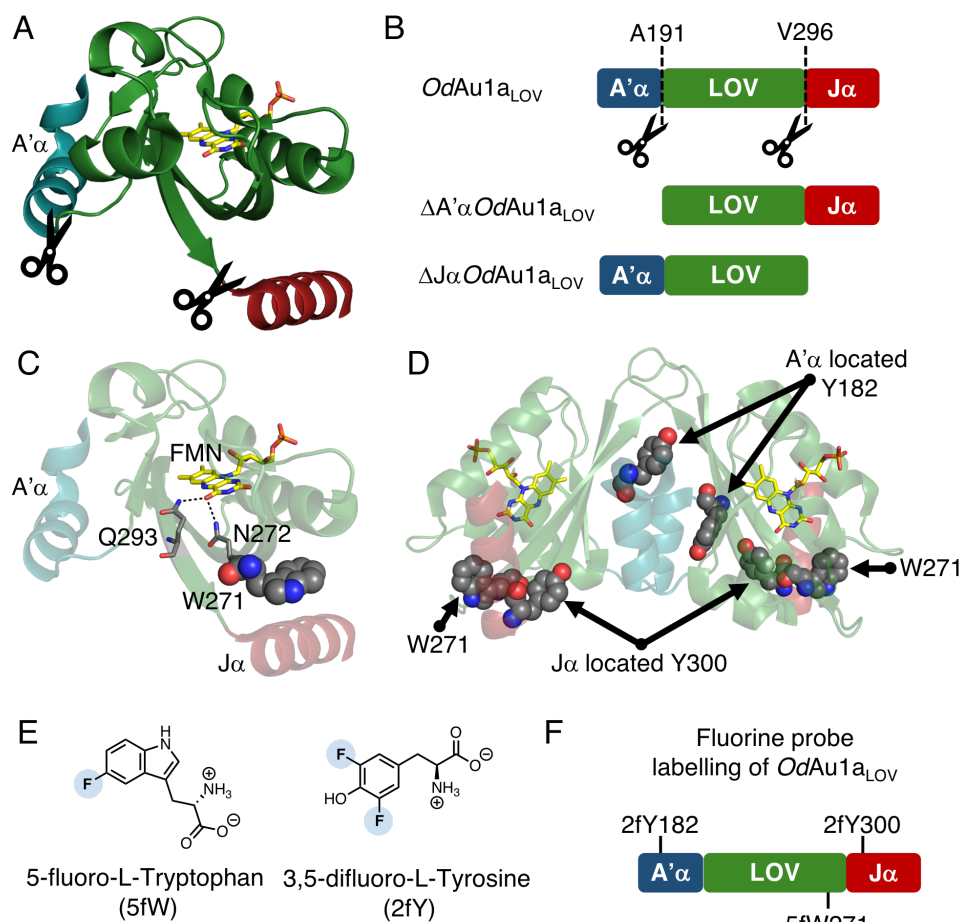
In all figures, B-factors are displayed as heat maps across the appropriate range defined. N-terminal A'α and C-terminal Jα helices are labelled for clarity. A) PDB 6I20, B) PDB 6I21 and C) PDB 6I22 describe Au1a<sub>LOV</sub> crystal structures under different illumination conditions. D) PDB 5DKK and E) PDB 5DKL crystal structures of PtAu1a<sub>LOV</sub>. The FMN cofactor is displayed as yellow sticks.

When OdAu1a<sub>LOV</sub> was investigated, the LOV domain core and A'α showed lower global B-factor than the Jα helix, Figure 5.2 A, B and C. All structures displayed high Jα helix B-factors (35 to 50) indicating that this region has significant heterogeneity irrespective of a photoactivation state, Figure 5.2 A, B and C. PtAu1a<sub>LOV</sub> structures show overall higher B-factors for both the dark state (30-60) and lit state (70-100) structures<sup>7</sup> in comparison to OdAu1a<sub>LOV</sub><sup>9</sup>, Figure 5.2 D and E. Higher B-factors are probably a result of much poorer resolution. In the dark state PtAu1a<sub>LOV</sub> A'α shows higher relative B-factors (45 to 60), than the Jα helix, Figure 5.2 D. Lit state structure shows the opposite behaviour with lower B-factors for the A'α now participating in dimer interface and higher B-factors for the Jα helix. These results suggested that for OdAu1a<sub>LOV</sub> the Jα helix is flexible irrespective of the photoactivation state while for PtAu1a<sub>LOV</sub> Jα becomes more flexible upon lit state dimer formation.

Considering the B-factors analysis, it would appear that the flexibility of A' $\alpha$  and J $\alpha$  helices is different between *PtAu1a<sub>LOV</sub>* and *OdAu1a<sub>LOV</sub>* homologues. Within the literature there is debate over the role of ancillary helices and whether the LOV domain dimerisation drives Aureochrome activation, or whether the bZIP domain defines dimerization and LOV domain alters the dimer structure through the A' $\alpha$  and J $\alpha$  helices<sup>5-11</sup>. As *OdAu1a* appears to behave differently to its homologues under crystallography condition it may yield new results that could help to clarify this debate. To investigate the role of ancillary helices, *OdAu1a<sub>LOV</sub>* constructs with truncated A' $\alpha$  and J $\alpha$  were designed, Figure 5.3 A and B. Deletion of *OdAu1a<sub>LOV</sub>* A' $\alpha$  was designed at a position similar to *PtAu1a<sub>LOV</sub>*<sup>10</sup> yielding  $\Delta$ A' $\alpha$ *OdAu1a<sub>LOV</sub>*, Figure 5.3 B. The J $\alpha$  deletion was designed by reflecting on the *OdAu1a<sub>LOV</sub>* crystal structures and identifying valine296 (V296) as the terminal  $\beta$ I-sheet residue. Replacement of serine297 (S297) with a stop codon yielded  $\Delta$ J $\alpha$ *OdAu1a<sub>LOV</sub>*.

In addition to truncation constructs alternative avenues to study the importance of A' $\alpha$  and J $\alpha$  helices in solution were considered, especially as large-scale structural changes are prevented within the crystal lattice. Our previous studies of *AsLOV2* have highlighted the applicability of <sup>19</sup>F labelling in the study of LOV domain conformational changes. Identically to *AsLOV2*, *OdAu1a<sub>LOV</sub>* also contains a tryptophan (W271) residue sidechain located in a latch region between the LOV domain core and the J $\alpha$  helix, Figure 5.3 C. *AsLOV2* results showed that unfolding of the J $\alpha$  resulted in a large-scale chemical shift perturbations for 5-fluoro-L-tryptophan (5fW labelled) protein. This therefore suggested that *OdAu1a<sub>LOV</sub>* could be studied similarly, with chemical shift changes expected if the J $\alpha$  is involved in dimerisation or undergoes unfolding event.

As the *OdAu1a* primary sequence contains only a single tryptophan residue, the A' $\alpha$  could not be investigated *via* 5fW labelling. Upon inspection of the crystal structure, it was evident that A' $\alpha$  and J $\alpha$  contain tyrosine residues, Figure 5.3 D, in ideal positions to study structural changes. Unspecific labelling of fluorinated tyrosine residues, similarly to 5fW, has been reported in the literature<sup>17,18</sup>. This approach was avoided due to presence of six tyrosine residues in the *OdAu1a<sub>LOV</sub>* sequence making assignment of <sup>19</sup>F signals challenging and requiring extensive site directed mutagenesis to pinpoint specific resonances. To overcome this, expansion of the genetic code methodology was considered. Previous studies of ribonucleotide reductase (RNR) catalytic mechanism involving proton coupled electron transfer (PCET) across tyrosine residues has exploited site specific incorporation of fluorotyrosine residues<sup>19,20</sup>.



**Figure 5.3 *Ochromonas danica* Aureochrome1a LOV domain ( $OdAu1a_{LOV}$ ) construct design.**

A) Structure of dark state  $OdAu1a_{LOV}$  (PDB code 6I20) depicting designed truncation constructs indicated with scissors, where  $A'\alpha$  is coloured in teal and  $J\alpha$  in red, respectively. B) Schematic representation of designed  $OdAu1a_{LOV}$  constructs indicating residues after which truncation was achieved.  $\Delta A'\alpha$  was designed after alanine191 (A191) including protein regions between glutamine192 (Q192) to lysine310 (K310) generating  $\Delta A'\alpha OdAu1a_{LOV}$ .  $\Delta J\alpha$  was designed after valine296 (V296) by replacing serine297 (S297) codon with a stop codon generating  $\Delta J\alpha OdAu1a_{LOV}$ . C) Structure of dark state  $OdAu1a_{LOV}$  indicating FMN and residues including glutamine293 (Q293) plus asparagine272 (N272) that hydrogen bond to FMN-O4 depicted as sticks. Tryptophan271 (W271) residue sidechain located in a latch region between LOV core and  $J\alpha$ , similarly to *Avena sativa* LOV2 (AsLOV2), is depicted as spheres. D) Structure of dark state  $OdAu1a_{LOV}$  (PDB code 6I20) indicating residues used  $^{19}F$  labelling studies including tyrosine182 (Y182), W271 and Y300. E) Chemical structures of 5-fluoro-L-tryptophan (5fW) and 3,5-difluoro-L-tyrosine (2fY) used for  $^{19}F$  labelling studies. F) Schematic representation of  $OdAu1a_{LOV}$  labelling with  $^{19}F$  probes for structural studies indicating 2fY182 located in the  $A'\alpha$  helix, 5fW271 located in LOV domain core and 2fY300 located in  $J\alpha$  helix, respectively.

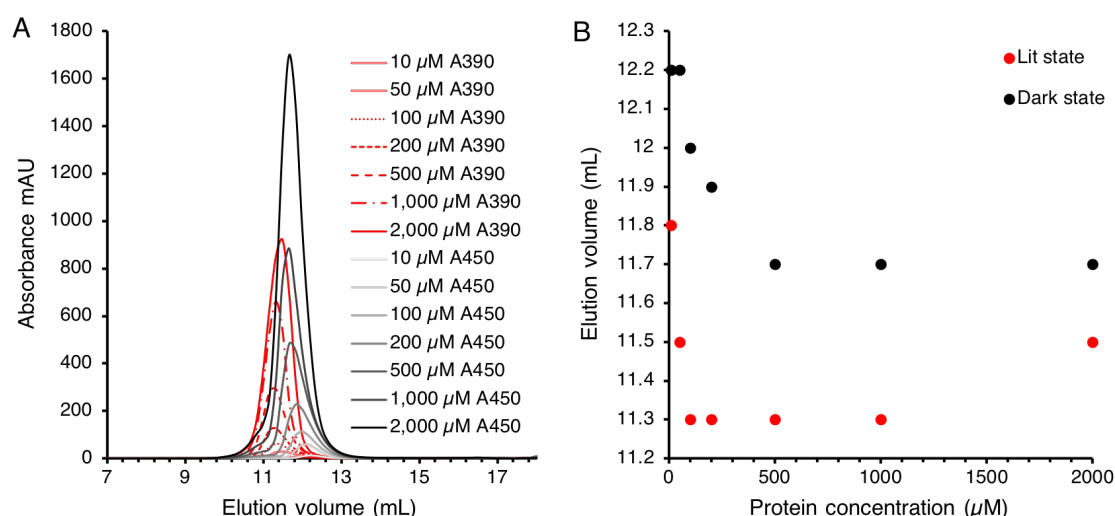
Using this methodology it was decided to label tyrosine182 (Y182) and tyrosine300 (Y300), Figure 5.3 D, with 2,5-difluoro-L-tyrosine (2fY), Figure 5.3 E. This approach was hypothesised to allow understanding of the  $A'\alpha$  helix behaviour in the

solution which could not be detected by the conventional  $^{15}\text{N}$ - $^1\text{H}$  TROSY-HSQC approach. Furthermore, considering that  $\text{J}\alpha$  was hypothesised to undergo minor structural changes, Y300 labelling with 2fY300 was also considered to investigate C-terminus structural change at a single residue resolution.

### 5.3 Structural Studies of 5-fluoro-L-Tryptophan Labelled *Ochromonas danica* Aureochrome1a and Aureochrome1a LOV Domain Constructs

#### 5.3.1 Labelling of *OdAu1a* and *OdAu1a<sub>LOV</sub>* Constructs with 5-fluoro-L-tryptophan

As NMR experiments require high protein concentrations the oligomerisation of *OdAu1a<sub>LOV</sub>* was investigated to determine its effect. Concentration dependent behaviour was investigated by size exclusion chromatography (SEC) up to 2,000  $\mu\text{M}$  in lit and dark states. Results showed light dependent shift of approximately 0.5 mL in the elution volume maxima at concentrations of 1,000 to 10  $\mu\text{M}$ , Figure 5.4 A and B. At 2,000  $\mu\text{M}$ , change in the elution volume was significantly smaller, 0.2 mL, with the lit elution peak shifting to a later elution volume. These observations indicate that under NMR like conditions *OdAu1a<sub>LOV</sub>* shows a clear light dependent shift in SEC trace.

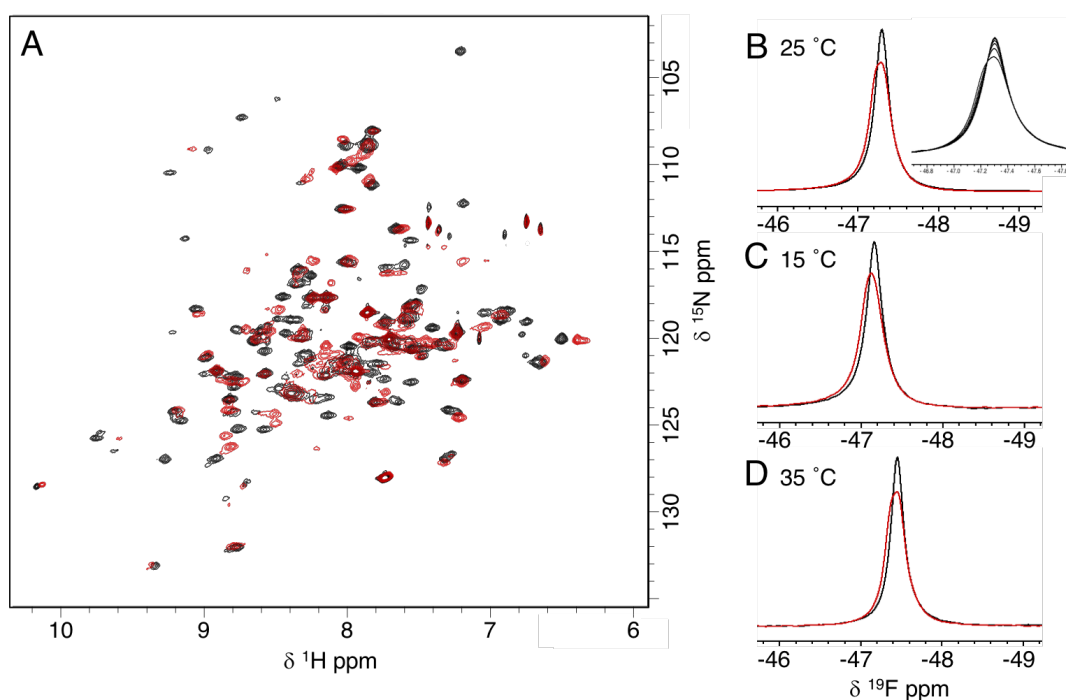


**Figure 5.4 Size exclusion chromatography (SEC) investigations of *OdAu1a<sub>LOV</sub>* oligomerisation change upon photoactivation under NMR like conditions.**

Two different absorbance maxima were followed including A450 (dark state FMN absorbance) and A390 (lit state FMN absorbance). Following A450 and A390 allowed high protein concentrations during the experiments considering that 280 nm traces yield absorbance above the detection levels. A) SEC75 traces of *OdAu1a<sub>LOV</sub>* under dark (grey and black solid lines) and lit (red lines) state conditions, respectively. B) Plot of *OdAu1a<sub>LOV</sub>* elution maxima in mL vs protein concentration. Black dots indicate dark state and red dots lit state samples, respectively. All experiments were performed in the NMR buffer pH 6.0 at room temperature.

Labelling with 5fW was undertaken following previously published methodologies<sup>21,22</sup> to investigate structural changes of the J $\alpha$  helix. If structural changes and the unfolding event are similar to AsLOV2, photoactivation would result in large  $^{19}\text{F}$  chemical shift perturbations. *OdAu1a<sub>LOV</sub>*  $^{15}\text{N}$ - $^1\text{H}$  TROSY HSQC experiments indicated that 5fW labelling did not interfere with protein folding and light responsiveness, Figure 5.5 A, and 5fW labelled and not labelled spectra overlapped almost identically, Appendix Figures 5.19 and 5.20. Labelling with 5fW was confirmed by liquid chromatography mass spectroscopy (LC-MS), Appendix Figure 5.21. When  $^{19}\text{F}$  NMR was collected for the dark state, a single peak with a chemical shift of  $\delta$  -47.30 ppm was observed with respect to the trifluoroacetate (TFA) peak, Figure 5.5 B. Illumination did not result in any large scale chemical shift perturbations indicating a slightly broader peak with  $\delta$  of -47.29 ppm,  $\Delta\delta$  0.01 ppm, Figure 5.5 B. Considering the observed chemical shifts for 5fW271, the result suggested minor J $\alpha$  structural change at the latch region. Relaxation of the protein from the lit to the dark state could be followed by  $^{19}\text{F}$  NMR, but half-lives could not be determined due to overlapping resonances, Figure 5.5 B.

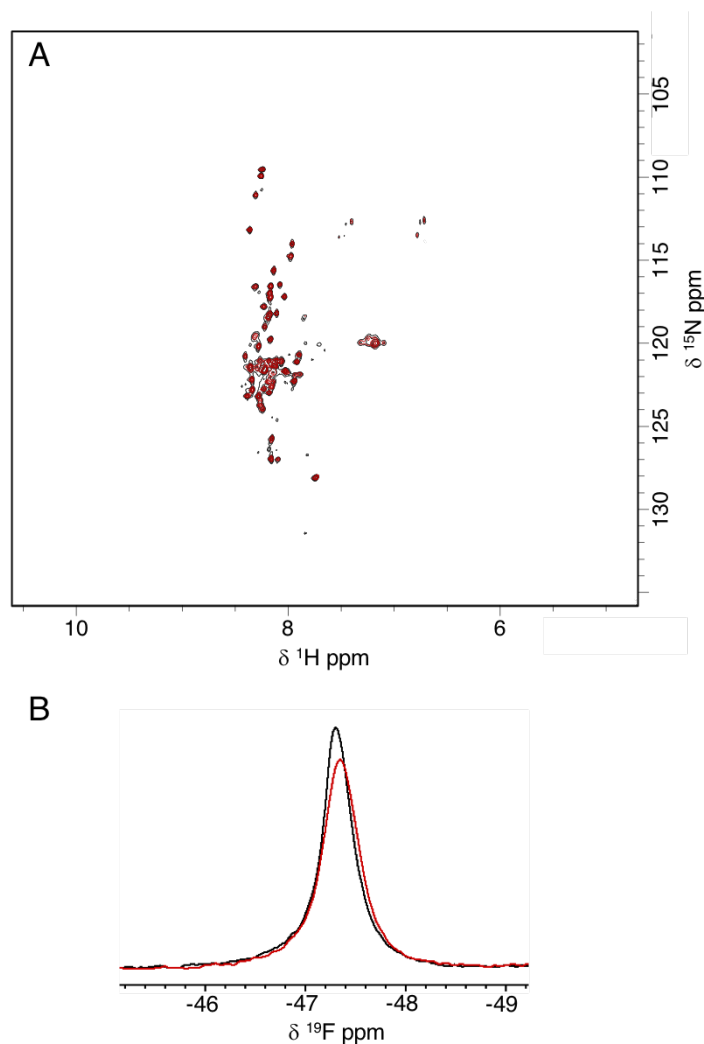
In order to investigate chemical shift differences further, temperature dependent behaviour upon photoactivation was investigated. Initial data was collected at 25 °C, therefore it was decided to investigate if upon cooling, 15 °C, or warming up the sample, 35 °C, peak intensities and shapes would change. The broadening of the lit state  $^{19}\text{F}$  NMR peak observed for 5fW271 *OdAu1a<sub>LOV</sub>* at 25 °C was most likely a product of dimer formation where larger complex influenced the relaxation rate due to slower molecular tumbling. Alternatively, it was also possible that the observed broader peak was a product of overlapping two peaks with an exchange rate on the timescale of the NMR experiment and this could be potentially resolved by altering experimental temperature.  $^{19}\text{F}$  NMR spectra at 15 °C was similar to that at 25 °C indicating  $\delta$  of -47.16 ppm in the dark and  $\delta$  of -47.13 ppm in the lit states, respectively, with a  $\Delta\delta$  of -0.03 ppm, Figure 5.5 C. Upon warming protein sample to 35 °C, no further changes were evident yielding chemical shift for dark and lit states of  $\delta$  -47.45 ppm, Figure 5.5 D.



**Figure 5.5 NMR investigations of 5fW271 labelled *OdAu1a*<sub>LOV</sub>.**

A)  $^{15}\text{N}$ - $^1\text{H}$  TROSY-HSQC spectra of dark state, black, and lit state, red, *OdAu1a*<sub>LOV</sub> labelled with 5fW271. Upon photoactivation, large-scale chemical shift perturbations were evident indicative of chemical changes. B), C) and D) indicated  $^{19}\text{F}$  NMR spectra of 5fW271 labelled *OdAu1a*<sub>LOV</sub> where dark state coloured in black and lit state is in red at B) 25 °C  $\delta$  -47.30 ppm dark and  $\delta$  -47.29 ppm, C) 15 °C  $\delta$  -47.16 ppm dark and  $\delta$  -47.13 ppm lit and D) 35 °C  $\delta$  -47.45 ppm dark and lit states, respectively. Top right corner of B) indicates multiple NMR spectra (512 scans) indicating relaxation of lit 5fW271 labelled *OdAu1a*<sub>LOV</sub> back to dark adapted state suggesting time dependent changes. Results indicate no significant changes in the chemical shifts or appearance of different peaks. Chemical shifts reported were referenced to trifluoroacetate (TFA) peak adjusted to  $\delta$  0.00 ppm.

To investigate if *OdAu1a*<sub>LOV</sub> results were representative for full length *OdAu1a* protein, 5fW271 labelling was undertaken. *OdAu1a*  $^{15}\text{N}$ - $^1\text{H}$  TROSY-HSQC did not show any resonances corresponding to the LOV domain core. The only resonances detected corresponded to the unstructured protein regions with a population of resonances at approximately 8.3 ppm in the proton dimension, Figure 5.6 A. No major differences if the chemical shifts could be detected upon photoactivation. Considering that NMR experiments required 400  $\mu\text{M}$  protein concentrations it is likely that the resulting dimer (approximately 70 kDa in size) was too large to yield a spectrum with resolved resonances. When the protein sample was investigated by  $^{19}\text{F}$  NMR experiments, the dark state indicated  $\delta$  of -47.31 ppm, almost identical to the *OdAu1a*<sub>LOV</sub>. Comparison of 5fW271 labelled *OdAu1a*<sub>LOV</sub> and *OdAu1a* showed no significant chemical variance in the  $^{19}\text{F}$  nuclei suggesting almost identical chemical environment.



**Figure 5.6 NMR studies of full length *OdAu1a*.**

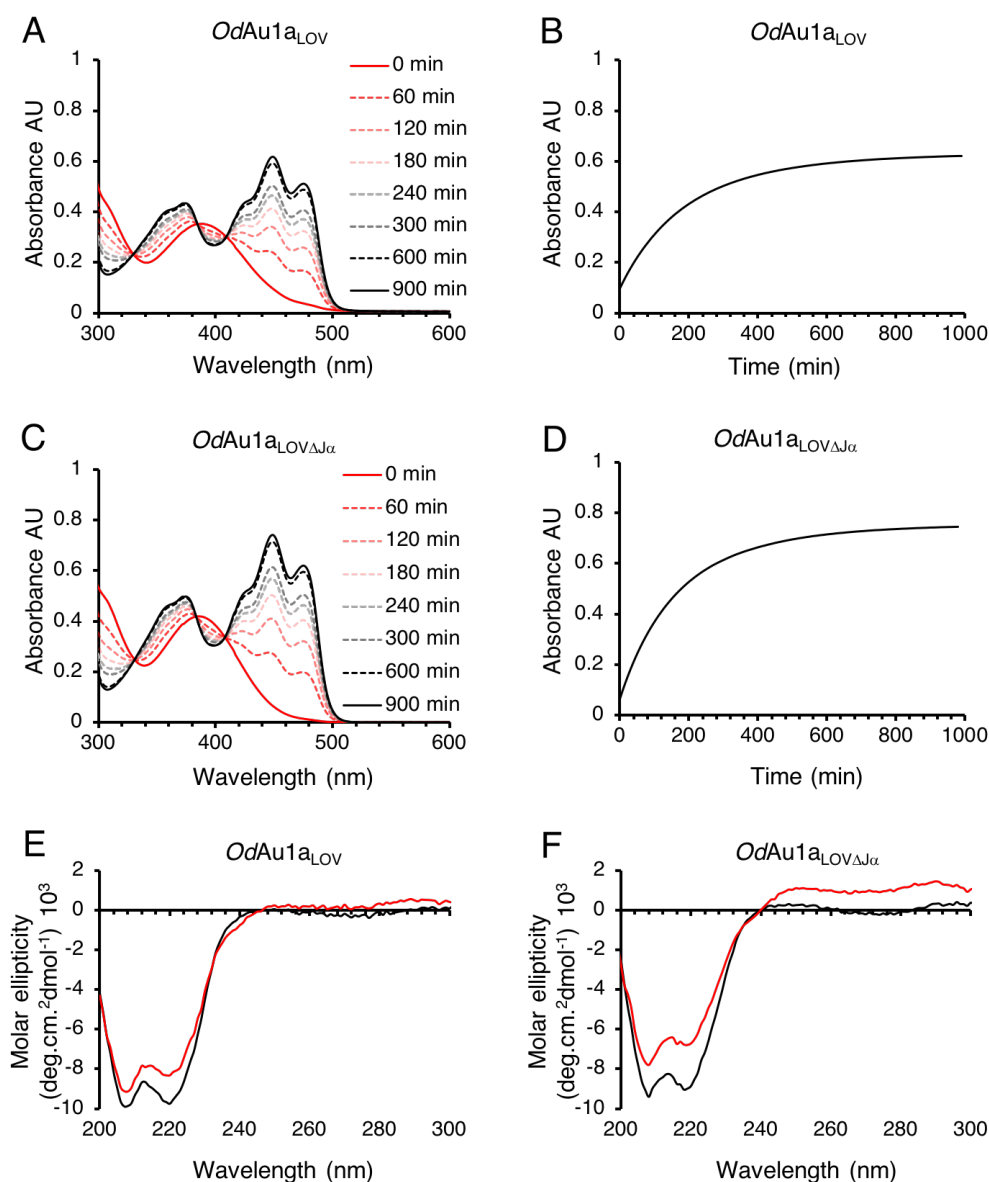
A)  $^{15}\text{N}$ - $^1\text{H}$  TROSY-HSQC spectra of dark state, black, and lit state, red, *OdAu1a*. Upon photoactivation, no large scale chemical shift perturbations were evident. Presence of chemical shift perturbation in the middle of the spectra, around 8.3 ppm in the  $^1\text{H}$  dimension, is indicative of unfolded N-terminal protein region. No resonances corresponding to LOV domain core were evident. B)  $^{19}\text{F}$  NMR spectra of 5fW271 labelled *OdAu1a* at 25 °C with dark spectra in black,  $\delta$  -47.31 ppm, and lit spectra in red,  $\delta$  -47.35 ppm, respectively. Results indicate no significant changes in the chemical shifts. Chemical shifts reported were referenced to trifluoroacetate (TFA) peak adjusted to  $\delta$  0.00 ppm.

### 5.3.2 Structural Investigations of $\Delta\text{J}\alpha\text{OdAu1a}_{\text{LOV}}$ Construct

Truncation of  $\text{J}\alpha$  influenced half-life relaxation kinetics resulting in half-life of 125 min, 15 min shorter than *OdAu1a<sub>LOV</sub>*, Figure 5.7 A to D. Characterisation by circular dichroism (CD) spectroscopy indicated larger  $\alpha$ -helical content loss than that observed for *OdAu1a<sub>LOV</sub>* construct, Figure 5.7 E and F, suggestive that the dominant structural change upon photoactivation was possibly the unfolding of the  $\text{A}'\alpha$  helix. Investigations of the protein construct by  $^{15}\text{N}$ - $^1\text{H}$  TROSY-HSQC indicated a spectrum with resonances overlapping with



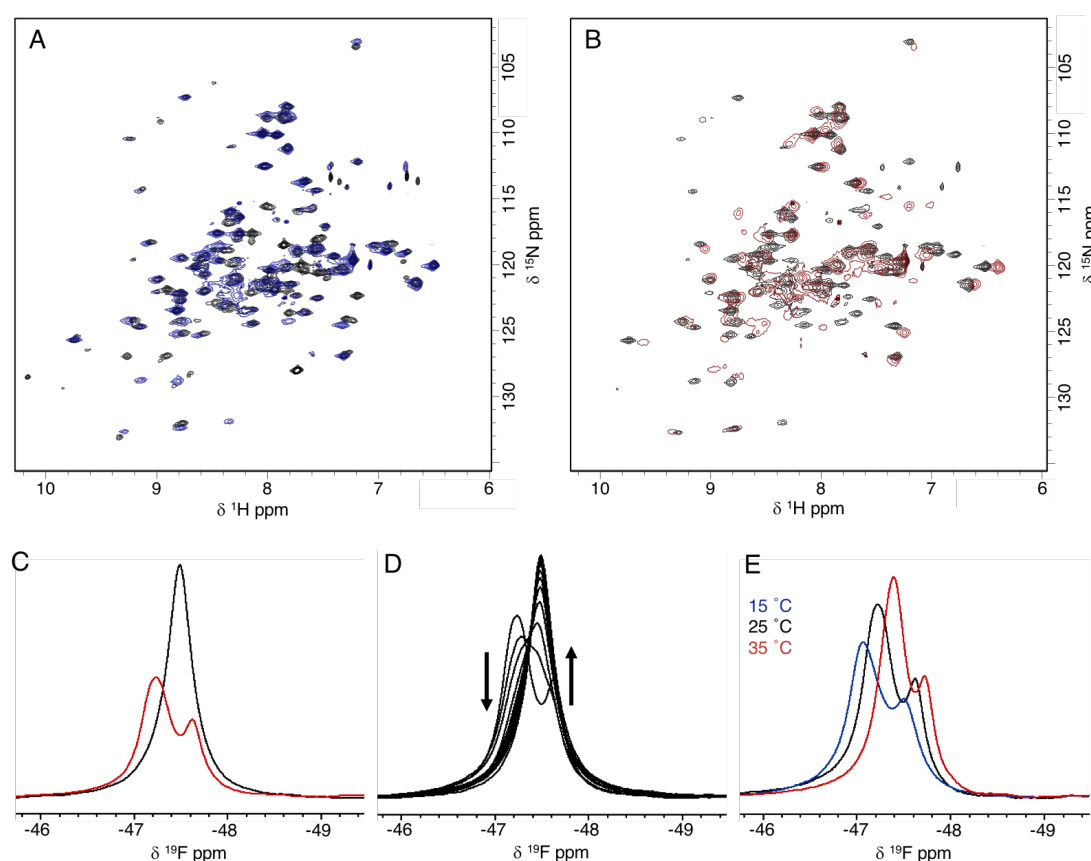
*OdAu1a<sub>LOV</sub>* <sup>15</sup>N-<sup>1</sup>H TROSY-HSQC spectrum excluding missing resonances of the J $\alpha$  helix, Figure 5.8 A. Overlapping NMR resonances, Figure 5.8 A, therefore indicate that J $\alpha$  truncation did alter or change the LOV domain structural integrity or fold. This observation was opposite to *AsLOV2* NMR studies where truncation of the J $\alpha$  helix resulted in altered resonances corresponding to the dark state LOV domain core<sup>23</sup>.



**Figure 5.7 Characterisation of *OdAu1a<sub>LOV</sub>* and  $\Delta$ J $\alpha$ *OdAu1a<sub>LOV</sub>* by UV-Vis and CD.**

A) *OdAu1a<sub>LOV</sub>* relaxing from lit, red solid line, to dark, black solid line. Relaxation half-life was 140 min. B) Plotted relaxation of *OdAu1a<sub>LOV</sub>* from lit to dark states. Raw data fitted to the first order decay. C)  $\Delta$ J $\alpha$ *OdAu1a<sub>LOV</sub>* relaxing from lit, red line, to dark, black line. Relaxation half-life was 125 min. D) Plotted relaxation of  $\Delta$ J $\alpha$ *OdAu1a<sub>LOV</sub>* from lit to dark states. Raw data fitted to the first order. E) CD spectra of *OdAu1a<sub>LOV</sub>* switching from dark, black line, to lit, red line, states. Change at 220 nm minima was 15% and 208 nm minima 8%. F) CD spectra of  $\Delta$ J $\alpha$ *OdAu1a<sub>LOV</sub>* switching from dark, black line, to lit, red line, states. Change at 220 nm minima was 22% and 208 nm minima 17%.

Labelling of the truncated variant with 5fW was undertaken. It was expected that the lack of the J $\alpha$  helix would result in the residue sidechain exposed to the solvent and no significant changes in  $^{19}\text{F}$  chemical shift upon photoactivation. 5fW271 labelling did not interfere with protein folding or light responsiveness resulting in a functioning protein as determined by  $^{15}\text{N}$ - $^1\text{H}$  TROSY-HSQC, Figure 5.8 B and Appendix Figures 5.22 and 5.23, with labelling confirmed by LC-MS, Appendix Figure 24. As expected, dark  $^{19}\text{F}$  NMR spectra showed a single peak with  $\delta$  -47.50 ppm,  $\sim$ 0.20 ppm upfield from *OdAu1a* and *OdAu1a<sub>LOV</sub>* peaks.

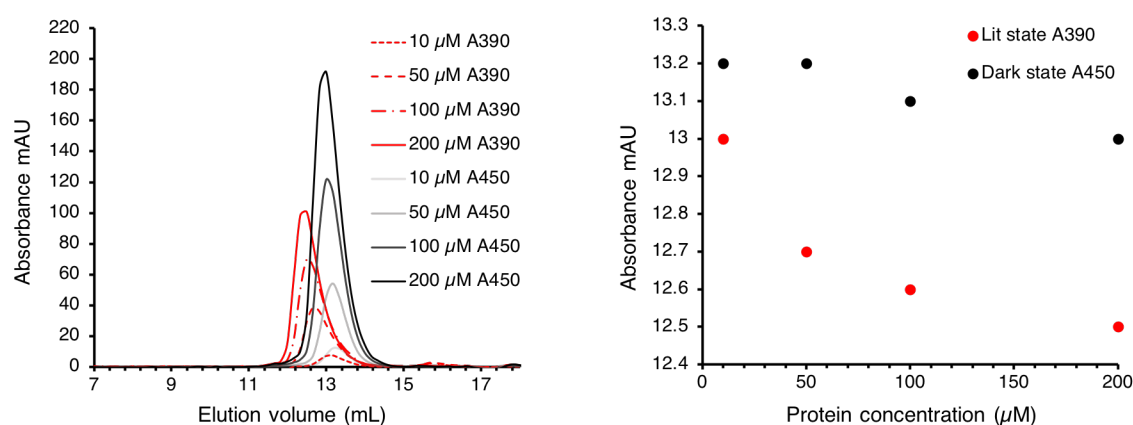


**Figure 5.8 NMR investigations of  $\Delta\text{J}\alpha\text{OdAu1a}_{\text{LOV}}$ .**

A)  $^{15}\text{N}$ - $^1\text{H}$  TROSY-HSQC spectra of dark state *OdAu1a<sub>LOV</sub>* (black) and  $\Delta\text{J}\alpha\text{OdAu1a}_{\text{LOV}}$  (blue). B)  $^{15}\text{N}$ - $^1\text{H}$  TROSY-HSQC spectra of 5fW271 labelled  $\Delta\text{J}\alpha\text{OdAu1a}_{\text{LOV}}$  under dark (black) and lit (red) state conditions. Upon photoactivation, large scale chemical shift perturbations were evident indicative of structural changes. C)  $^{19}\text{F}$  NMR spectra of 5fW271 labelled  $\Delta\text{J}\alpha\text{OdAu1a}_{\text{LOV}}$  with dark state is in black and lit state is in red. D)  $^{19}\text{F}$  NMR spectra indicating reversion of 5fW271 labelled  $\Delta\text{J}\alpha\text{OdAu1a}_{\text{LOV}}$  from lit to dark states, respectively. Arrow down indicates decaying lit state and arrow up growth of a peak for dark state 5fW271. Due to overlapping resonances, half-life could not be determined. E) Indicates lit state 5fW271  $\Delta\text{J}\alpha\text{OdAu1a}_{\text{LOV}}$  peaks at 15 °C (blue), 25 °C (black) and 35 °C (red). No major differences in peak shape or intensities were evident. Chemical shifts reported here were referenced to TFA peak adjusted to  $\delta$  0.00 ppm.

Unexpectedly, photoactivation resulted in a significant change in the  $^{19}\text{F}$  NMR spectra yielding two overlapping peaks with maxima at  $\delta$  -47.24 and -47.63 ppm, respectively, a  $\Delta\delta$  of -0.39 ppm Figure 5.8 C. Time course  $^{19}\text{F}$  NMR experiments showed reversion of  $^{19}\text{F}$  chemical shifts, Figure 5.8 D, but due to overlapping resonances the half-life could not be determined. Temperature dependent experiments, at 15 and 35 °C, did not show any significant differences in peak shapes or intensities, Figure 5.8 E. The  $\Delta\delta$  differences between the two lit state  $^{19}\text{F}$  peaks did differ slightly with temperature. At 15 °C the  $\Delta\delta$  was -0.43 ppm, at 25 °C -0.39 ppm and at 35 °C -0.33 ppm indicative of two distinct conformations that upon warming up the sample approach fast exchange regime. These peaks are likely to arise from two possible conformations of the indole ring, a product of removal of sterically hindering J $\alpha$  helix.

Changes in the  $\Delta\text{J}\alpha\text{OdAu1a}_{\text{LOV}}$  oligomerisation state upon photoactivation were confirmed by analytical gel filtration chromatography. Titration experiments, from 200 to 50  $\mu\text{M}$  of the protein under NMR like conditions indicated light dependent differences in the elution volume of 0.5 mL, Figure 5.9, indicative of a functional protein. Notably, at 10  $\mu\text{M}$   $\Delta\text{J}\alpha\text{OdAu1a}_{\text{LOV}}$  indicated a much smaller shift of 0.2 mL suggesting that at lower concentrations light dependent changes in oligomerisation state can be hindered. As expected, overall elution volumes for dark and lit state  $\Delta\text{J}\alpha\text{OdAu1a}_{\text{LOV}}$  samples were later than for  $\text{OdAu1a}_{\text{LOV}}$  construct.



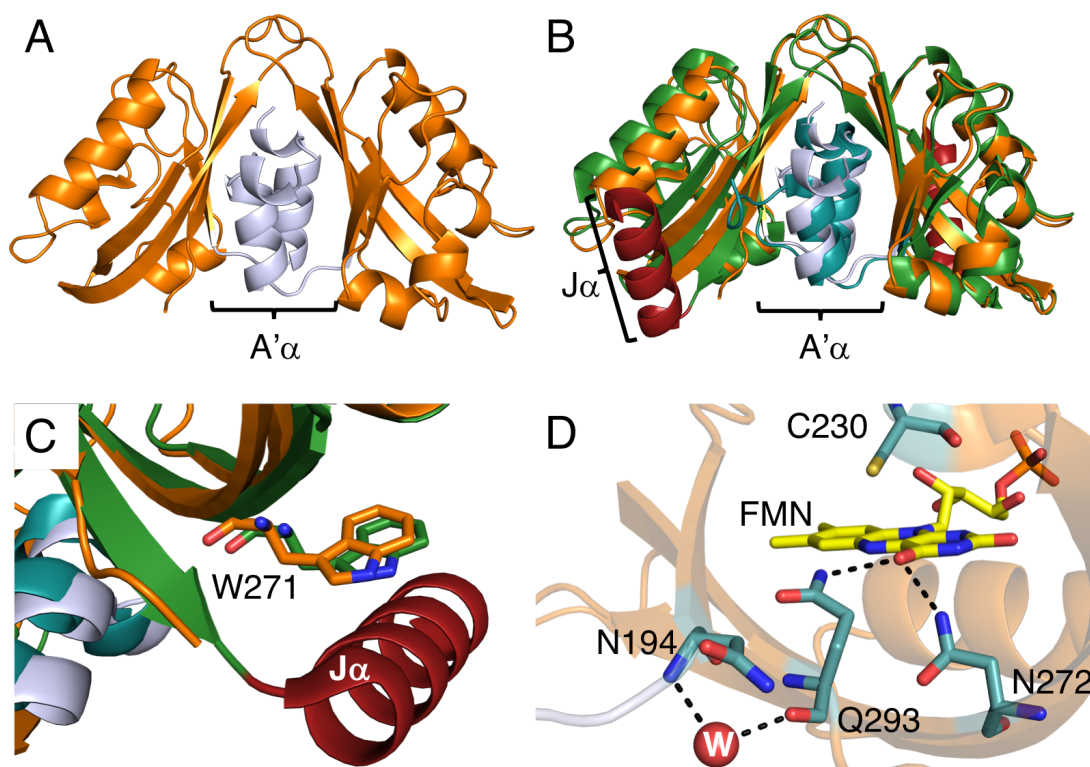
**Figure 5.9 Size exclusion chromatography (SEC) investigations of  $\Delta\text{J}\alpha\text{OdAu1a}_{\text{LOV}}$  oligomerisation change upon photoactivation under NMR like conditions.**

A) SEC75 traces of  $\Delta\text{J}\alpha\text{OdAu1a}_{\text{LOV}}$  under dark (grey and black solid lines) and lit (red lines) state conditions, respectively. B) Plot of  $\Delta\text{J}\alpha\text{OdAu1a}_{\text{LOV}}$  elution maxima in mL vs protein concentration. Black dots indicate dark state and red dots lit state samples, respectively. All experiments were performed in the NMR buffer pH 6.0 at room temperature.

To further understand  $\Delta J\alpha OdAu1a_{LOV}$  constructs, protein was crystallised under identical conditions as for  $OdAu1a_{LOV}$ <sup>9</sup> resulting in yellow crystals in a space group  $P 3_1 2 1$  diffracting to a 2.0 Å resolution. Molecular replacement using  $OdAu1a_{LOV}$  (PDB 6I20) indicated two monomers per asymmetric unit, Figure 5.10 A. The A'α helix was positioned across β-sheet surface as previously observed, Figure 5.10 B. W271 sidechain appeared to be in a single conformation similarly to  $OdAu1a_{LOV}$  construct explaining a similar <sup>19</sup>F chemical shift environment, Figure 5.10 C. The most notable differences between  $OdAu1a_{LOV}$  and  $\Delta J\alpha OdAu1a_{LOV}$  crystal structures was the positioning and the hydrogen bond network of N194 residue. For  $OdAu1a_{LOV}$  N194 sidechain hydrogen bonds to the backbone carbonyl of Q293, but for the  $\Delta J\alpha OdAu1a_{LOV}$  construct N194 was 3.3 – 3.4 Å away in a different orientation. This N194 to Q293 hydrogen bond was replaced by an ordered water molecule in close proximity to the backbone amine of N194 and backbone carbonyl of Q293. Nonetheless, it is interesting to note that although lacking previously hypothesised hydrogen bond network between N194 sidechain and backbone carbonyl Q293,  $\Delta J\alpha OdAu1a_{LOV}$  was functional at concentrations above 50 μM.

Analysis of the dimeric structure of chain A of  $\Delta J\alpha OdAu1a_{LOV}$  and chain A of dark state FMN  $OdAu1a_{LOV}$  indicated RMSD value of 0.332 Å (743 to 743 atoms) suggesting a related structure as seen in the figure 5.10 A and B. PISA analysis of the structure furthermore indicated dimer interface energy  $\Delta^iG = -14.3$  kcal/mol, a value similar to dark state FMN  $OdAu1a_{LOV}$  structure of  $\Delta^iG = -13.2$  kcal/mol. This structural analysis therefore suggests that such dimeric arrangement is possibly stable in the solution even under dark state conditions. Attempts were made to try to crystallise protein under continues blue irradiation, but no crystallisation was evident. Furthermore, it was also pursue to illuminate crystals with blue light prior harvesting, but rapid crystal degradation was evident shortly after illumination with visible damaged observed.

Attempts to purify  $\Delta A'\alpha OdAu1a_{LOV}$  construct were unsuccessful as the protein was found in the inclusion bodies. Multiple attempts to purify protein under denaturing conditions yielded protein, but upon the removal of the chaotrope precipitation was evident. This therefore suggested that unlike  $\Delta J\alpha OdAu1a_{LOV}$ ,  $\Delta A'\alpha OdAu1a_{LOV}$  was prone to aggregation.

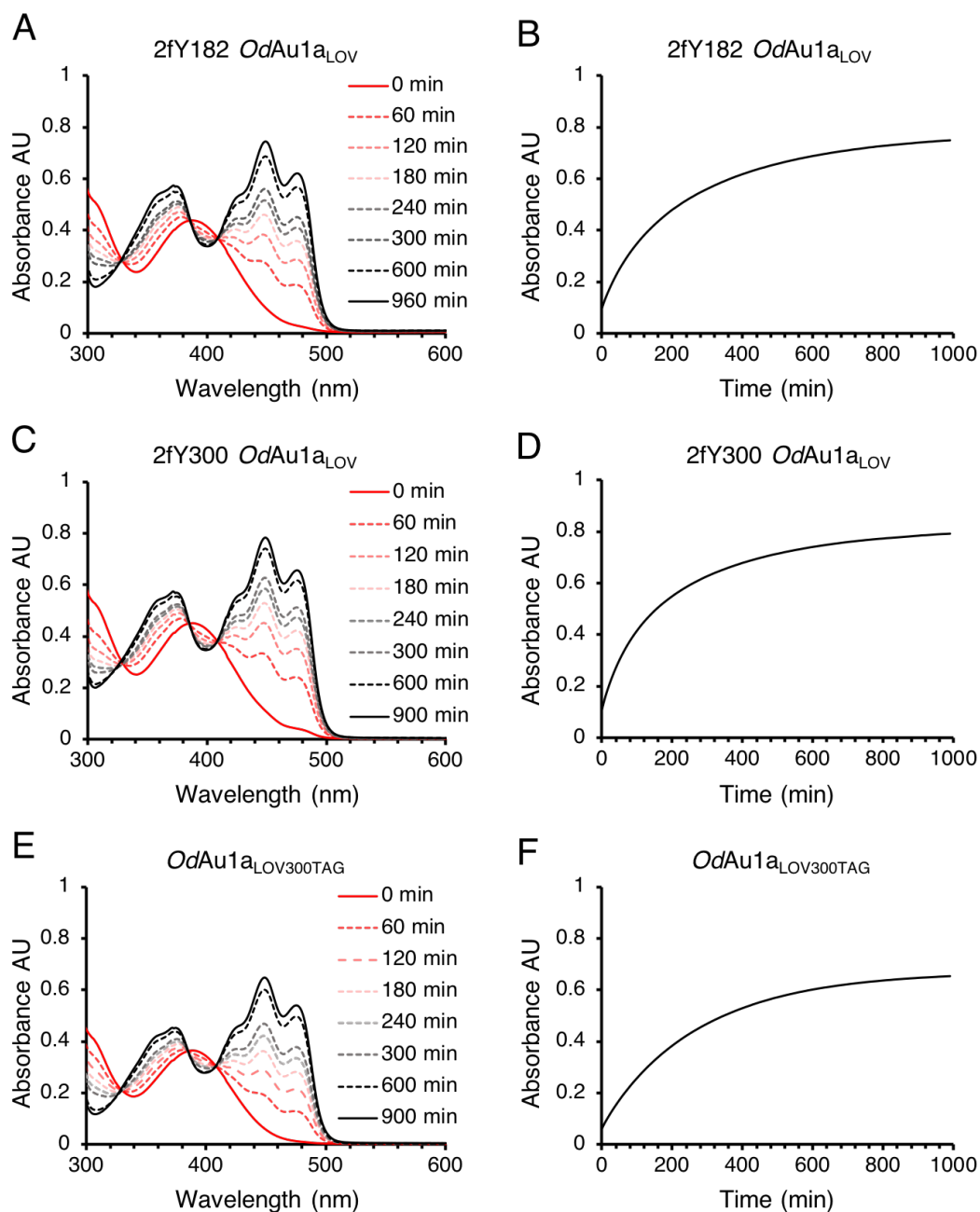


**Figure 5.10 X-ray crystallography studies of  $\Delta J\alpha OdAu1a_{LOV}$  construct.**

A) 2.0 Å resolution crystal structure of  $\Delta J\alpha OdAu1a_{LOV}$ . LOV domain core is displayed in orange and N-terminal  $A'\alpha$  in grey-blue. PISA analysis of the dimer interchange indicated dimer interface energy  $\Delta'G = -14.3$  kcal/mol suggesting possible dimeric structure in the solution. B) Overlaid structures of  $OdAu1a_{LOV}$  (PDB code 6I20, LOV domain core displayed in green,  $A'\alpha$  in teal and  $J\alpha$  in red, respectively) and  $\Delta J\alpha OdAu1a_{LOV}$  with colouring as in A). RMSD value of 0.332 Å (743 to 743 atoms) was determined when structures were overlaid (chains A). C) Structure and positioning of W271 residue in  $OdAu1a_{LOV}$  and  $\Delta J\alpha OdAu1a_{LOV}$  constructs, respectively, with colouring identical to figure B). D)  $\Delta J\alpha OdAu1a_{LOV}$  structure depicting hydrogen bond networks within the FMN binding pocket and allosteric regions indicating residues N194, C230, N272 and Q293 displayed as sticks. Hydrogen bonds in distance of 2.9 to 3.2 Å are depicted as black dashed lines. No hydrogen bond could be observed between N194 sidechain and backbone carbonyl of Q293, but a bridging water molecule was observed (red sphere labelled W).

#### 5.4 Structural Studies of Aureochrome1a LOV Domain Labelled with 3,5-difluoro-L-Tyrosine

In order to provide structural insights for the  $A'\alpha$  and  $J\alpha$  helices, incorporation of 3,5-difluoro-L-tyrosine (2fY) was undertaken. 2fY was incorporated as described in the literature with minor modifications<sup>19,20</sup>. 2fY was incorporated by replacing Y182 and Y300 codons with amber stop codons (TAG) allowing modified orthogonal *Methanococcus jannaschii* tyrosyl-tRNA synthetase/ tRNA to incorporate the unnatural amino acid into the protein<sup>19,20</sup>. This approach allowed targeting of appropriate tyrosine residue for the <sup>19</sup>F labelling studies.



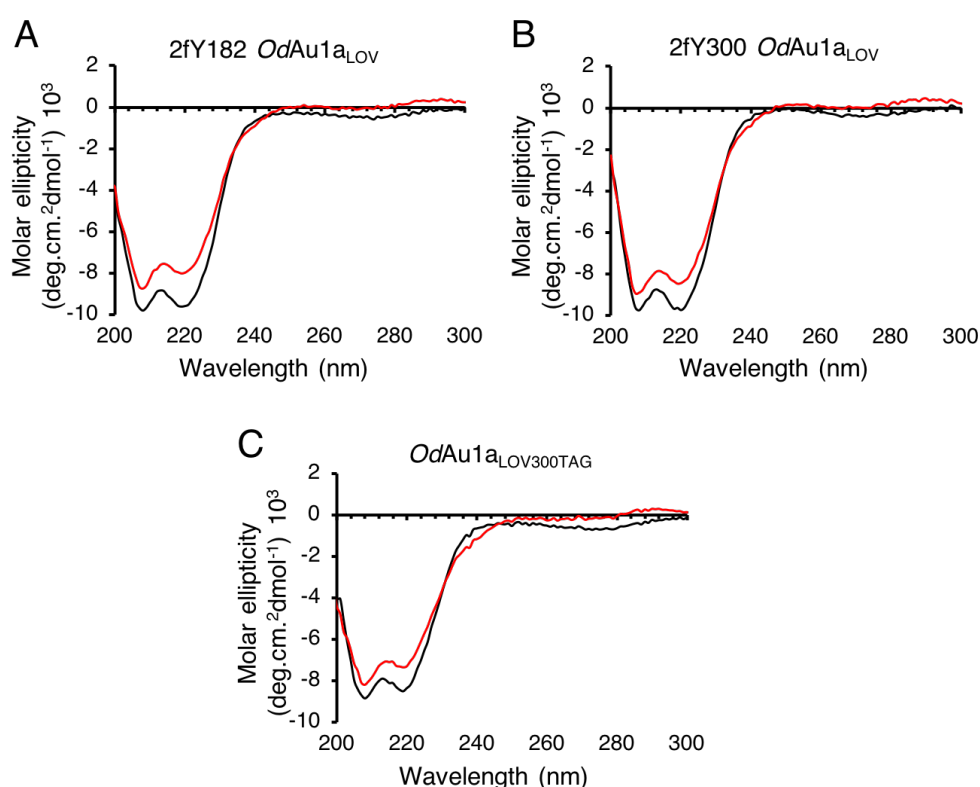
**Figure 5.11 Characterisation of 2fY182 and 2fY300 labelled *OdAu1a<sub>LOV</sub>* by UV-Vis.**

2fY labelled *OdAu1a<sub>LOV</sub>* relaxing from lit, red solid line, to dark, black solid line where A) 2fY182 and C) 2fY300 *OdAu1a<sub>LOV</sub>*, and E) *OdAu1a<sub>LOV300TAG</sub>* control. Plotted relaxation of 2fY labelled *OdAu1a<sub>LOV</sub>* at 448 nm from lit to dark states where B) 2fY182 half-life 171 min, D) 2fY300 half-life 143 min and F) *OdAu1a<sub>LOV300TAG</sub>* half-life 188 min. Raw data fitted to the first order decay.

Investigation of the protein constructs including 2fY182 and 2fY300 *OdAu1a<sub>LOV</sub>*, and *OdAu1a<sub>LOV300TAG</sub>* (control truncation sample where 2fY addition was omitted in the protein induction step) yielded photocycling protein constructs with altered half-life kinetics in comparison to the wild type protein. The half-life of 2fY182 containing protein was 171

min, 2fY300 143 min and *OdAu1a*<sub>LOV300TAG</sub> 188 min, Figure 5.11 A to E. Considering that half-life of wild type *OdAu1a*<sub>LOV</sub> was 140 min, it was evident that truncation as well as introduction of 2fY influenced protein recovery.

Introduction of 2fY was confirmed by LC-MS experiments yielding protein mass plus two corresponding <sup>19</sup>F atoms, Appendix Figure 5.25 A and B. 2fY300 *OdAu1a*<sub>LOV</sub> indicated 15% contaminant of *OdAu1a*<sub>LOV300TAG</sub>, appendix Figure 5.25 B and C. This was a result of not complete 2fY incorporation, where TAG codon can be used either for unnatural amino acid incorporation or act as a termination (stop) codon<sup>19,20</sup>. Considering that *Escherichia coli* (*E. coli*) cells were not amber recognising prokaryotic translation termination factor (pTTF) null phenotype, the 2fY charged tRNA that recognises UAG codon had to compete with endogenous pTTF. To overcome this challenge with inhomogeneous protein generation, C-terminus His<sub>6</sub>-tag can be used allowing purification of only the full length and unnatural amino acid charged protein. This was not considered due to C-terminal His<sub>6</sub>-tag requiring additional C-terminal residues with potential undesired consequences.



**Figure 5.12 Characterisation 2fY labelled *OdAu1a*<sub>LOV</sub> and *OdAu1a*<sub>LOV300TAG</sub> by CD.**

CD spectra of A) 2fY182 and B) 2fY300 *OdAu1a*<sub>LOV</sub> switching from dark, black line, to lit, red line, states. Changes in CD spectra minima A) 220 nm 16% and 208 nm 10%, and B) 220 nm 14% and 208 nm 8%. C) Changes in CD spectra for *OdAu1a*<sub>LOV300TAG</sub> switching from dark, black line, to lit, red line, states. Change at 220 nm minima was 13% and 208 nm minima 8%.



CD characterisation indicated light dependent secondary structural changes comparable to the wild type protein, Figure 5.12. This observation therefore suggests that 2fY probe or truncation of the *OdAu1a<sub>LOV</sub>* at Y300 position did not interfere with protein folding or light dependent structural changes.

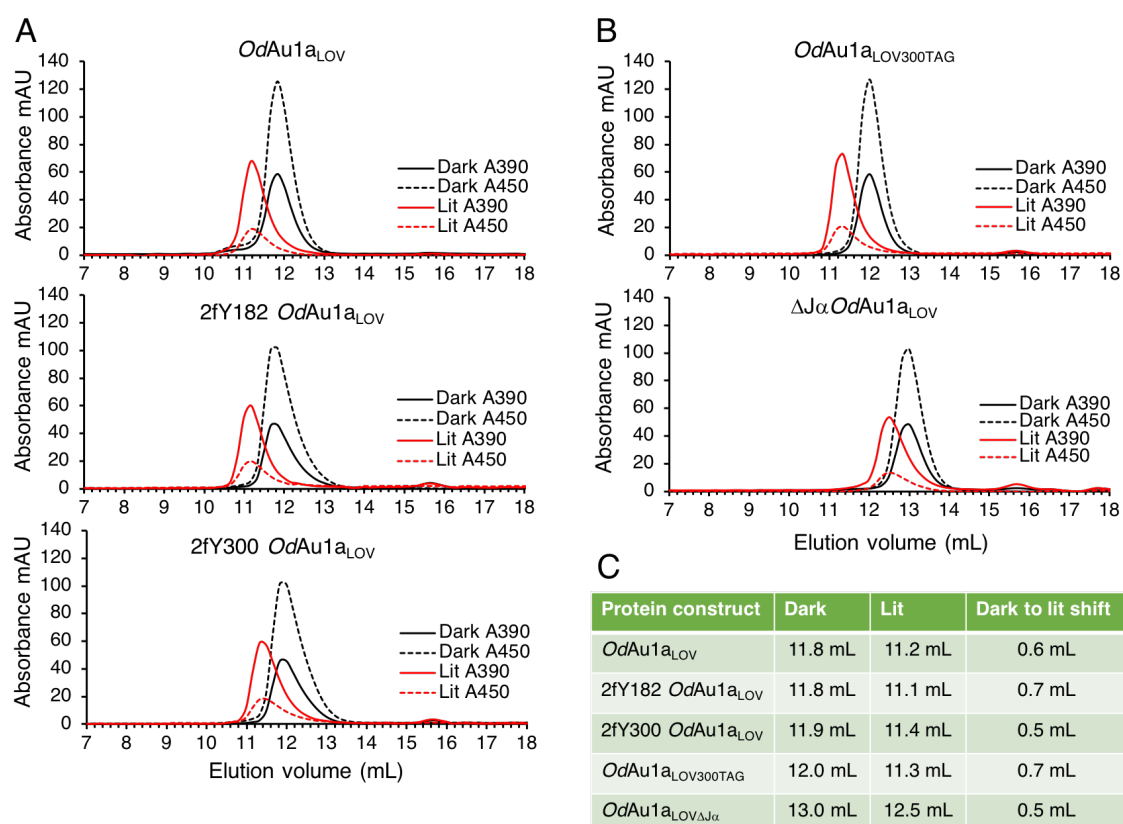
Considering influenced recovery kinetics, prior proceeding with the NMR experiments it was decided to confirm that 2fY labelling did not interfere with previously observed SEC dark to lit state elution volume shifts. SEC experiments using 100  $\mu$ M protein samples were in line with light dependent changes in the elution volume, Figure 5.13. In comparison to the wild type protein, minorly altered elution volumes were evident for 2fY182 *OdAu1a<sub>LOV</sub>*, 2fY300 *OdAu1a<sub>LOV</sub>* and *OdAu1a<sub>LOV300TAG</sub>* suggestive of different protein volumes. To confirm that this was not a product of experiment to experiment variation or errors, each SEC experiment was performed in triplicate indicating that observed differences were reproducible, Appendix Figures 5.26 to 5.30. 2fY182 *OdAu1a<sub>LOV</sub>* results, Figure 5.13 A and C, indicated 0.1 mL earlier lit state elution with dark to lit state shift of 0.7 mL. Conversely, 2fY300 *OdAu1a<sub>LOV</sub>* indicated 0.1 mL later dark state and 0.2 mL lit state elution, Figure 5.13 A and C, with dark to lit shift of 0.5 mL.

Although it was unclear why different elution volume maxima were observed, it was evident that introduction of 2fY influenced potential protein dark state monomer and lit state dimer shapes without influencing light dependent oligomerisation change. The most notable and unexpected result was obtained for *OdAu1a<sub>LOV300TAG</sub>* construct. The molecular size of the construct was comparable to the  $\Delta$ J $\alpha$  deletion construct with only three extra residues. The elution volumes observed for *OdAu1a<sub>LOV300TAG</sub>* were more representative of *OdAu1a<sub>LOV</sub>* construct, Figure 5.13 A, B and C, in comparison to  $\Delta$ J $\alpha$ *OdAu1a<sub>LOV</sub>*.

Considering  $^1\text{H}$  and  $^{19}\text{F}$  nuclei size differences, it was unlikely that the observed volume differences were due to size alone unless  $^{19}\text{F}$  probe prevented tight residue packing. Considering that introduction of two  $^{19}\text{F}$  atoms results in tyrosine hydroxy group pKa value change from 10.0 (tyrosine) to 7.2 (2fY), it was likely that changes in the pKa influenced hydrogen bonding<sup>20</sup>. Changes in hydrogen bonding have been observed in the  $\beta$ -subunit of *E. coli* ribonucleotide reductase (RNR) near the diiron-oxo centre where 2fY indicated different orientation<sup>20</sup>. This therefore suggests that introduction of 2fY possibly altered hydrogen bonding extending from corresponding tyrosine residues in turn influencing molecular volumes. Although minor SEC differences were observed, considering light dependent changes in SEC elution volumes and observable changes in the CD signal, it was decided to proceed with  $^{19}\text{F}$  studies for 2fY182 *OdAu1a<sub>LOV</sub>*, but not for 2fY300 *OdAu1a<sub>LOV</sub>*. 2fY300 *OdAu1a<sub>LOV</sub>* was not characterised due to inhomogeneous



sample and time limitations. It was hypothesised that presence of *OdAu1a*<sub>LOV300TAG</sub> contaminant would result in multiple dimers that could also exchange, see Figure 5.14, making NMR resonance results challenging to interpret.

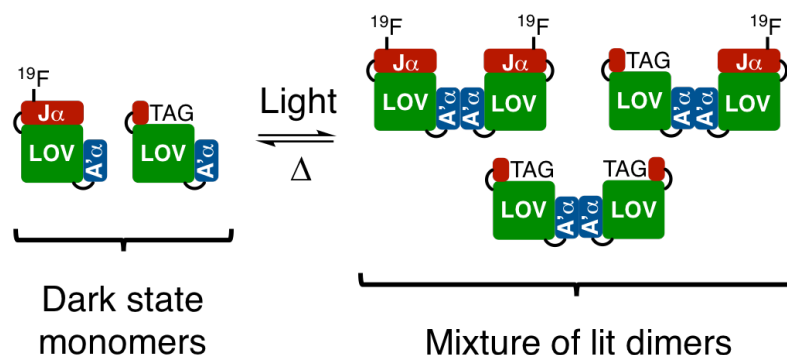


**Figure 5.13 SEC investigations of 2fY labelled *OdAu1a*<sub>LOV</sub> and *OdAu1a*<sub>LOV300TAG</sub>.**

A) SEC75 traces of 100  $\mu$ M top *OdAu1a*<sub>LOV</sub>, middle 2fY182 *OdAu1a*<sub>LOV</sub> and bottom 2fY300 *OdAu1a*<sub>LOV</sub> under dark (black lines) and lit (red lines) state conditions, respectively. B) SEC75 traces of 100  $\mu$ M top *OdAu1a*<sub>LOV300TAG</sub> and bottom  $\Delta J\alpha$  *OdAu1a*<sub>LOV</sub>. C) Values of all elution volume maxima. All experiments were performed in the NMR buffer pH 6.0 at room temperature. Changes upon photoactivation were investigated following absorbance values at 450 and 390 nm confirming complete protein photoactivation where absorbance at 390 nm was higher than at 450 nm.

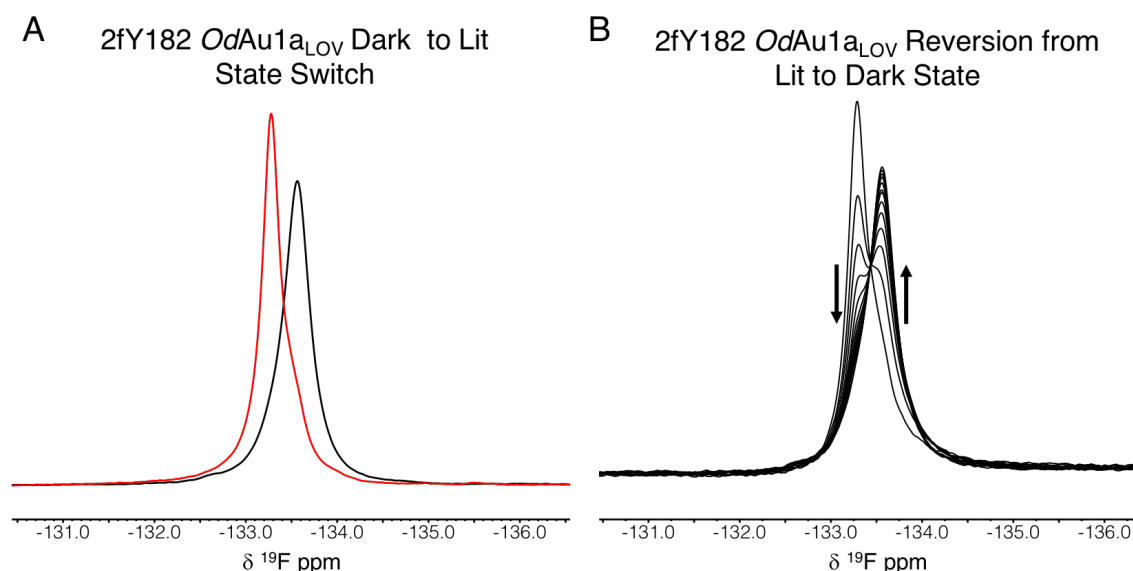
When 2fY182 *OdAu1a*<sub>LOV</sub> protein was investigated by  $^{19}\text{F}$  NMR spectroscopy, a small downfield chemical shift difference of  $\Delta\delta$  -0.29 ppm was evident suggestive of a structural change, Figure 5.15 A. Reversion back to the dark-adapted state could be followed by  $^{19}\text{F}$  NMR, but due to overlapping peaks, half-life could not be accurately determined. Nonetheless,  $\delta$  changes were not large, hence it could not be concluded if chemical shift were in line with residue sidechain environment change from a hydrophobic to solvent exposed conformations. In order to pinpoint and confirm  $^{19}\text{F}$  NMR observations, it was decided to proceed with solvent induced isotope shift (SIIS) analysis by replacing  $^1\text{H}_2\text{O}$  ( $\text{H}_2\text{O}$ ) to  $^2\text{H}_2\text{O}$  ( $\text{D}_2\text{O}$ ) buffer as a solvent<sup>24,25</sup>.  $\text{H}_2\text{O}$  and  $\text{D}_2\text{O}$  have different dielectric constants therefore influencing chemical shift of the  $^{19}\text{F}$  nuclei<sup>24,25</sup>. For instance, it has

been reported that exchange of H<sub>2</sub>O to D<sub>2</sub>O results in an upfield <sup>19</sup>F NMR signal for fluorinated aromatic residues<sup>24,25</sup>. This therefore means that buried <sup>19</sup>F nuclei show small or none, whilst solvent exposed large SIIS providing an easy methodology to study protein conformational changes. To minimise the bias, spectra were not referenced to TFA peak adjusted to 0.00 ppm and TFA  $\delta$  were reported as observed.



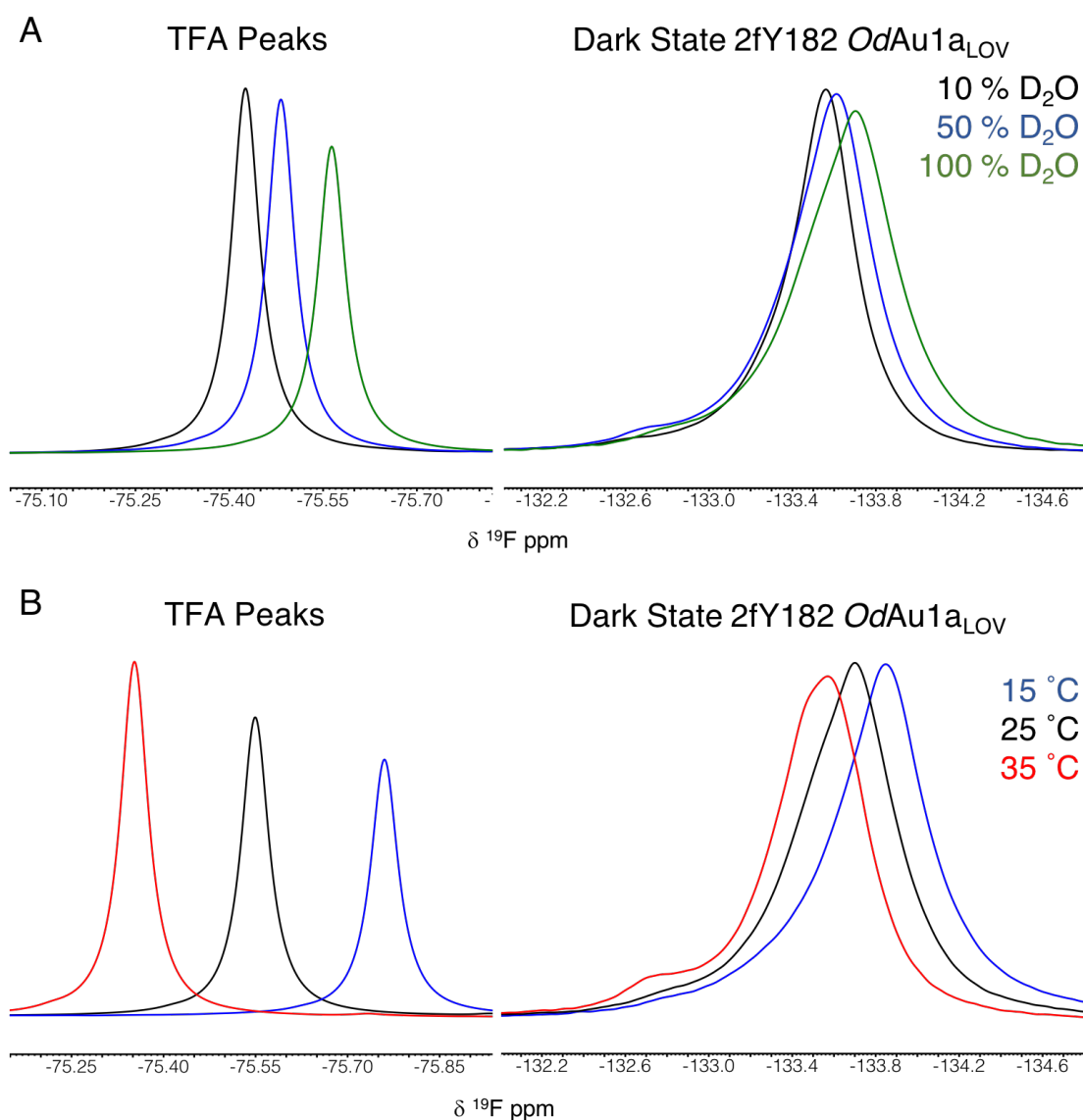
**Figure 5.14 Hypothesised light dependent multiple dimer formation for 2fY300 *OdAu1a<sub>LOV</sub>* sample.**

2fY300 *OdAu1a<sub>LOV</sub>* sample contains approximately 15% *OdAu1a<sub>LOV300TAG</sub>* contaminant. In the dark state, 2fY300 *OdAu1a<sub>LOV</sub>* and *OdAu1a<sub>LOV300TAG</sub>* exist as monomers. Photoactivation triggers dimerisation resulting in the formation of multiple dimeric states. Furthermore, it is not known if lit state dimers can rapidly exchange resulting in further complications.



**Figure 5.15 <sup>19</sup>F NMR investigations of 2fY182 *OdAu1a<sub>LOV</sub>*.**

A) <sup>19</sup>F NMR spectra of 2fY182 *OdAu1a<sub>LOV</sub>* with dark state is in black and lit state is in red. The dark state maxima was at -133.58 ppm and lit state maxima -133.29 ppm. B) <sup>19</sup>F NMR spectra indicating reversion of 2fY182 *OdAu1a<sub>LOV</sub>* from lit to dark state, respectively. Arrow down indicates decaying lit state and arrow up growth of a peak for dark state 2fY182 *OdAu1a<sub>LOV</sub>*. All spectra were collected at 25 °C in a buffer supplemented with 10% D<sub>2</sub>O.



**Figure 5.16 SIIS <sup>19</sup>F NMR investigations of dark state 2fY182 *OdAu1a<sub>LOV</sub>*.**

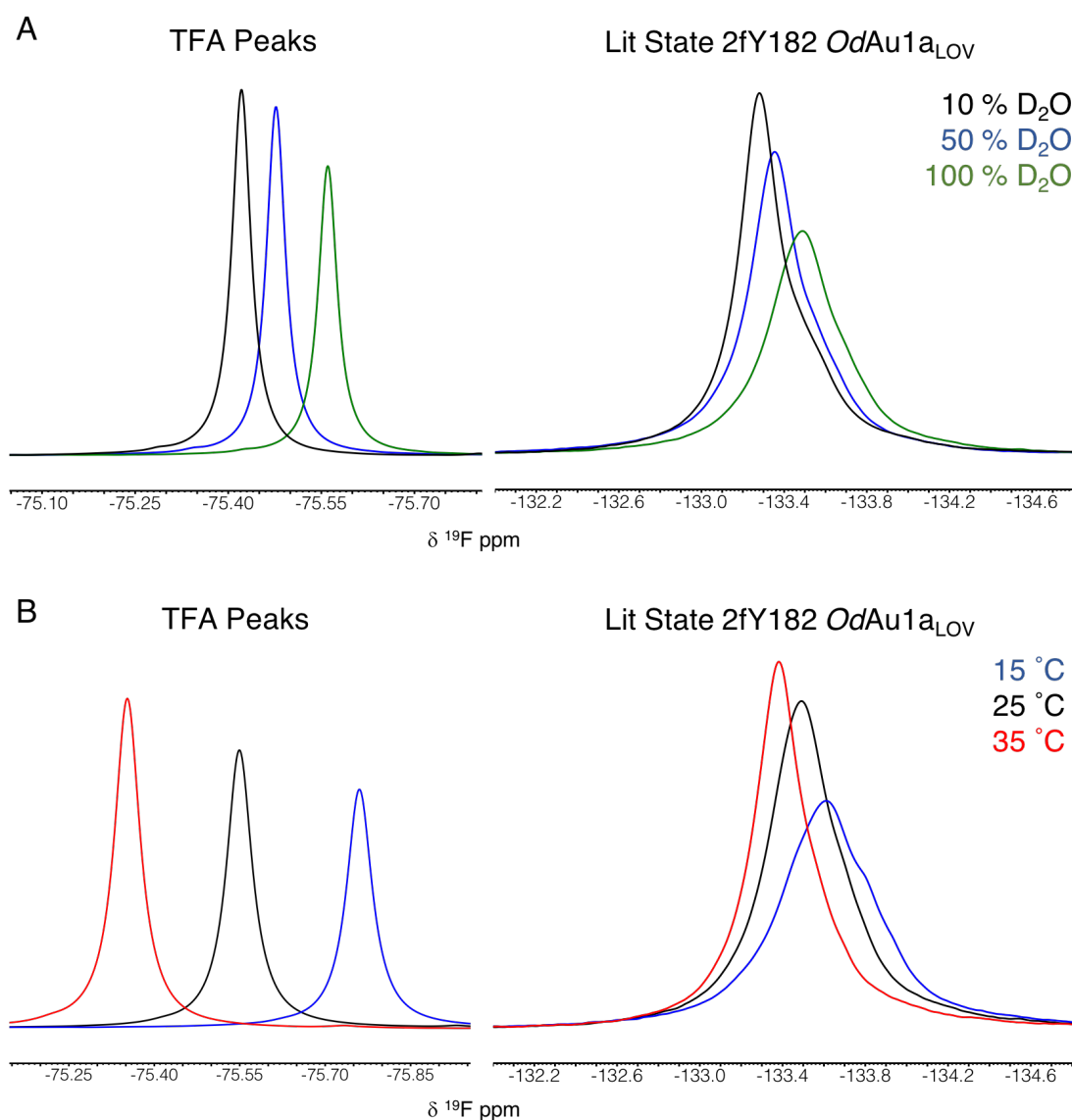
A) <sup>19</sup>F NMR spectra of dark state 2fY182 *OdAu1a<sub>LOV</sub>* with variable percentage of D<sub>2</sub>O containing buffer, where black 10%, blue 50% and green 100% D<sub>2</sub>O, respectively. The left section of the spectrum indicates TFA peaks with maxima -75.43 ppm for 10%, -75.49 ppm 50% and -75.57 ppm 100% of D<sub>2</sub>O, respectively. The right section of the spectrum indicates 2fY182 peak with maxima -133.58 ppm 10%, -133.63 ppm 50% and -133.72 ppm 100% of D<sub>2</sub>O, respectively. B) <sup>19</sup>F NMR spectra of 2fY182 *OdAu1a<sub>LOV</sub>* at 15 °C, blue, 25 °C, black, and 35 °C, red, respectively. The left section of the spectrum indicates TFA peaks with maxima -75.78 ppm at 15 °C, -75.57 ppm at 25 °C and -75.38 ppm at 35 °C, respectively. The right section of the spectrum indicates 2fY182 peak with maxima -133.86 ppm at 15 °C, -133.72 ppm at 25 °C and -133.72 ppm at 35 °C. Please note at 35 °C a prominent overlapping peak is observed with  $\delta$  around -132.60 ppm suggestive of a small population of a secondary conformer for 2fY182 under dark state conditions. All temperature dependent experiments were performed on a protein sample containing 100% D<sub>2</sub>O buffer.

Prior proceeding with SIIS experiments, 2fY SIIS was investigated to determine maximum value observed for a free amino acid in the solution under NMR conditions, Appendix Figure 5.31. Replacing 10% D<sub>2</sub>O buffer to 100% D<sub>2</sub>O buffer resulted in SIIS values of  $\Delta \delta$  of -0.13 ppm for TFA and  $\Delta \delta$  of -0.15 ppm for 2fY amino acid. Considering these results, dark state 2fY182 *OdAu1a<sub>LOV</sub>* was investigated. H<sub>2</sub>O to D<sub>2</sub>O exchange resulted in altered chemical shifts for TFA and 2fY182,  $\Delta \delta$  -0.14 ppm for TFA and  $\Delta \delta$  -0.14 ppm for 2fY182, Figure 5.16 A. Considering these observations, it was evident that 2fY182 under dark state conditions indicated SIIS similar to that of a free amino acid suggestive of accessibility to the solvent.

Excluding SIIS observed for the dark state 2fY182 peak, a small shoulder peak was also evident upon increase in D<sub>2</sub>O content to 100%, Figure 5.16 A, in  $\delta$  region -132.70 ppm suggestive of an alternative conformation downfield from the dark and the lit state 2fY182 signal. Considering these observations, it was decided to investigate if altering temperature would result in changes in peak shapes or intensities to provide evidence if alternative conformations exist for the dark state A' $\alpha$ . Cooling from 25 to 15 °C resulted in an upfield shift for TFA and 2fY182 resonances without altering peak shapes or intensities, Figure 5.16 B. As expected, warming up to 35 °C indicated downfield shift for TFA and 2fY182 also indicating an additional more prominent minor peak at  $\delta$  -132.60 to -133.80 ppm suggestive of an alternative conformation. Furthermore, peak sharpening was not observed at 35 °C suggesting multiple 2fY182 conformations under dark state condition.

Furthermore, lit state 2fY182 chemical shifts were also investigated. Replacement of H<sub>2</sub>O with D<sub>2</sub>O resulted in a larger change in  $\Delta \delta$  for 2fY182 resonances yielding downfield SIIS of -0.21 ppm, Figure 5.17 A, a value larger than for a free 2fY amino acid in the solution. Considering dark state results, it was evident that upon exposure to light, the conformational changes result in 2fY182 getting exposed to the solvent and present in a different chemical environment influencing <sup>19</sup>F resonance signal. Lastly, when temperature dependent behaviour was investigated, no additional peaks were evident. As expected, cooling the sample to 15 °C resulted in broadening and warming to 35 °C in sharpening of 2fY signal, Figure 5.17 B.

These results therefore suggested significant structural changes taking place in the A' $\alpha$  region of the protein, which could be in line with dimerisation event occurring. Nonetheless, no significant broadening of peaks or weakening of the signals was evident. This therefore raises a question if the dimerisation does take place or this could be a product of dimer rearrangement as seen in the crystal structures.



**Figure 5.17 SIIS  $^{19}\text{F}$  NMR investigations of lit state 2fY182 *OdAu1a<sub>LOV</sub>*.**

A)  $^{19}\text{F}$  NMR spectra of lit state 2fY182 *OdAu1a<sub>LOV</sub>* with variable percentage of  $\text{D}_2\text{O}$  containing buffer, where black 10%, blue 50% and green 100%  $\text{D}_2\text{O}$ , respectively. The left section of the spectrum indicates TFA peaks with maxima -75.43 ppm for 10%, -75.49 ppm 50% and -75.57 ppm 100% of  $\text{D}_2\text{O}$ , respectively. The right section of the spectrum indicates 2fY182 peak with maxima -133.29 ppm 10%, -133.37 ppm 50% and -133.50 ppm 100% of  $\text{D}_2\text{O}$ , respectively. B)  $^{19}\text{F}$  NMR spectra of 2fY182 *OdAu1a<sub>LOV</sub>* at 15 °C, blue, 25 °C, black, and 35 °C, red, respectively. The left section of the spectrum indicates TFA peaks with maxima -75.78 ppm at 15 °C, -75.57 ppm at 25 °C and -75.38 ppm at 35 °C, respectively. The right section of the spectrum indicates 2fY182 peak with maxima -133.63 ppm at 15 °C, -133.50 ppm at 25 °C and -133.39 ppm at 35 °C. All temperature dependent experiments were performed on a protein sample containing 100%  $\text{D}_2\text{O}$  buffer.

## 5.5 Discussion and Conclusions

In this work,  $^{19}\text{F}$  labelling and consequent NMR chemical shift investigations were used to pinpoint structural changes in the solution for *OdAu1a<sub>LOV</sub>*. As discussed for *AsLOV2* in Chapter 4,  $^{19}\text{F}$  NMR provides an alternative methodology to study of light dependent LOV domain conformational changes.

First of all, strong evidence has been provided to suggest that for *OdAu1a<sub>LOV</sub>* C-terminal  $\text{J}\alpha$  is not required for in solution light dependent conformational changes.  $^{15}\text{N}$ - $^1\text{H}$  TROSY-HSQC experiments confirmed that truncation of the  $\text{J}\alpha$  does not result in perturbed LOV domain core resonances indicative of an identical fold. Furthermore, photoactivation triggered global chemical shift perturbations suggesting major structural changes taking place similarly to the WT LOV domain.

Attempts to generate SEC calibration curve were unsuccessful under NMR like conditions. Therefore, it was not possible to conclude if changes in SEC traces were in line with dimerisation. Previous studies, Chapter 3, suggested light dependent dimerisation of *OdAu1a<sub>LOV</sub>*, but structural studies indicated parallel dimers under dark and light grown conditions, respectively. Furthermore, RMSD analysis resulted in low score values furthermore suggesting closely related structures, yet between dark and light grown states dimer arrangement across  $\beta$ -sheet surface was evident (see Chapter 3, section 3.6 for more details). In line with the dimer arrangement changes in the crystal structures, PISA (PDBe) analysis suggested dimeric arrangement with  $\Delta G = -12$  to  $-14$  kcal/mol. Considering that SEC experiments failed to determine molecular size of the species, it is possible to suggest that elution volume changes under NMR conditions could be a product of dimerisation, as previously suggested, or could also be a result of dimer rearrangement and the change in molecular volume of the dimer, Figure 5.18 A, hence resulting in changes in SEC traces. Poor calibration curve traces were probably due to NMR buffer being more acidic (pH 6.0) and containing no salt increasing changes of protein aggregation and interaction with the matrix (20 mM MES, 1 mM EDTA, 1 mM TCEP, 5 mM  $\text{NaN}_3$  pH 6.0), see Appendix Figure 4.32 for more details. Lastly, experiments performed on *OdAu1a<sub>LOV300TAG</sub>* control construct also indicated light dependent changes in the SEC elution volumes providing further evidence that  $\text{J}\alpha$  was not required for light dependent structural changes.

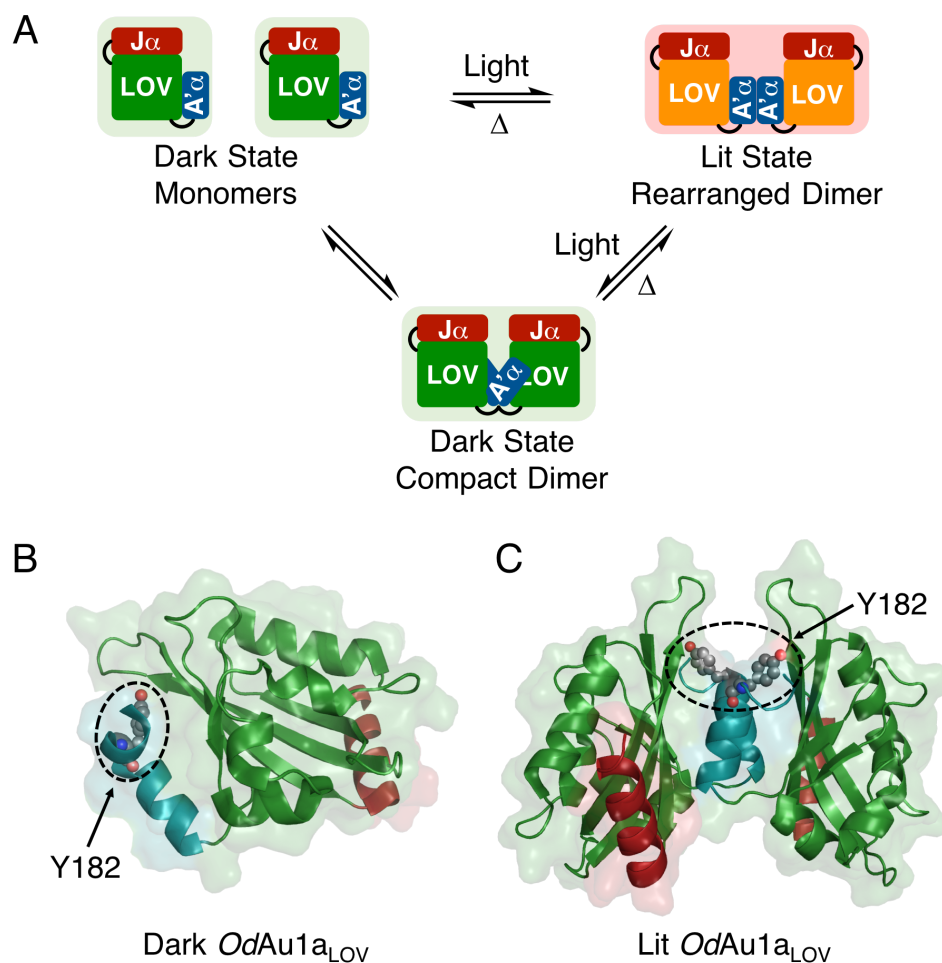
Considering  $^{19}\text{F}$  experiments, 5fW271 labelled *OdAu1a<sub>LOV</sub>* and *OdAu1a* indicated only minor chemical shift perturbation upon photoactivation. The W271 indole group is located in the latch region between the LOV domain core and the  $\text{J}\alpha$ , therefore these

observations suggested only minor structural changes occurring upon photoactivation. On the other hand,  $^{19}\text{F}$  NMR of 5fW271 labelled  $\Delta\text{J}\alpha\text{OdAu1a}_{\text{LOV}}$  indicated different set of chemical shift perturbations in the lit state. Although it was a puzzling observation, it was likely a product of higher degrees of freedom of the indole group. Considering that the removal of the  $\text{J}\alpha$  alleviates steric hindrance, it was likely that this allowed multiple orientations of the sidechain in the lit state, hence different  $^{19}\text{F}$  nuclei chemical environments.

X-ray crystallography studies of the  $\Delta\text{J}\alpha\text{OdAu1a}_{\text{LOV}}$  indicated a unique observation. Previously, it was suggested that a hydrogen bond between N194 sidechain to the backbone carbonyl of Q293 was required to lock protein in the dark state conformation. The crystal structure of  $\Delta\text{J}\alpha\text{OdAu1a}_{\text{LOV}}$  indicated a different scenario. First of all, N194 to Q293 hydrogen bond was not evident, whilst protein crystallised as previously observed for the dark state  $\text{OdAu1a}_{\text{LOV}}$ . Secondly, this hydrogen bond was replaced by a bridging water molecule suggesting a possible rescue mechanism to lock  $\text{A}'\alpha$  in the dark state orientation. This observation therefore provides evidence that hydrogen bonding extending from N194 might not be an absolute requirement to lock protein the dark state.

Finally, by applying 2fY labelling solution  $\text{A}'\alpha$  structural changes upon photoactivation could be studied. 2fY182  $\text{OdAu1a}_{\text{LOV}}$  indicated moderate  $^{19}\text{F}$  NMR chemical shifts and large SIIS changes upon photoactivation. Notably, higher intensity peak was evident for the lit state 2fY182  $\text{OdAu1a}_{\text{LOV}}$  in comparison to the dark state sample. Opposite results were evident for 5fW271  $\text{OdAu1a}_{\text{LOV}}$ . Broadening and weakening of 5fW271 signals were probably a product of the changes in oligomerisation state or changes in the dimer arrangement where the larger complex formation influenced molecular tumbling hindering  $^{19}\text{F}$  nuclei relaxation. Considering these observations, it is evident that  $\text{A}'\alpha$  at Y182 residue becomes partially unfolded allowing higher degree of freedom, hence higher intensity of the  $^{19}\text{F}$  NMR signal. Unfortunately, due to time restraints it was not possible to investigate this observation further.

To provide further structural understanding,  $^{19}\text{F}$  NMR results were mapped onto crystal structures of  $\text{OdAu1a}_{\text{LOV}}$  including dark (PDB 6I20) and lit (PDB 6I22) states. In the dark state structure, Y182 sidechain is packed against  $\beta$ -sheet surface and is partially shielded from the solvent in three of the monomers (chains A, B and D)<sup>9</sup>. In one of the monomers (chain C) Y182 appears to pack against the  $\beta$ -sheet surface of a symmetry partner suggestive of a crystallographic artefact<sup>9</sup>. This structural observation is in line with observed -0.14 ppm SIIS results, Figure 5.18 B, suggesting sidechain accessibility to the solvent. Furthermore, dark state 2fY182  $^{19}\text{F}$  NMR results also indicated a minor downfield



**Figure 5.18 Mapping of 2fY182  $^{19}\text{F}$  NMR results onto crystal structure of *OdAu1a*<sub>LOV</sub>.**

In the figures N-terminal A'α helix is depicted in teal and C-terminal Jα in red. Protein accessible surface areas (ASAs) are depicted at a transparency level of 80%. A) Schematic representation of possible *OdAu1a*<sub>LOV</sub> changes in oligomerisation or changes of the dimer arrangement. In the dark state, protein can exist as a monomer where exposure to the light triggers dimerisation (lit state dimer formation). Alternatively, protein under NMR conditions can form or exists as a dark state compact dimer where exposure to the light triggers rearrangement of A'α helix across dimer interface allowing formation of the lit state larger dimer. Please note that the lit state dimer is larger in size than the dark state dimer. Considering that SEC experiments could not determine molecular size of the dark and the lit state protein, small elution volume shift could be a product of either dimerisation or possibly changes of the dimeric arrangement. Crystal structures of *OdAu1a*<sub>LOV</sub> under B) dark (PDB 6I20) and C) lit (PDB 6I22) state conditions, respectively. In the figure tyrosine182 (Y182) residues are indicated with arrows and are depicted as spheres with atoms appropriately coloured including carbon grey, oxygen red and nitrogen blue, respectively. In C) A'α helix is 'sandwiched' between two LOV domains with Y182 residue exposed to the solvent.

peak suggesting yet uncharacterised low population conformation that probably cannot be identified by X-ray structural studies. Once again, due to limited time this observation could not be investigated further by experiments including  $^{19}\text{F}$  relaxation dispersion to quantitatively investigate conformational dynamics. Lit state 2fY182 results were also



mapped onto the lit state *OdAu1a*<sub>LOV</sub> structure, Figure 5.18 B. Unique lit state dimer indicates A'α helices sandwiched between the two LOV domain cores where Y182 sidechain is located in a cavity providing an environment for possible residue rearrangements, Figure 5.18 B. This observation is in line with a large -0.21 ppm SIIS observed for the lit state 2fY182 suggestive of a high degree of exposure to the solvent. Considering that this cavity provides an environment for residue sidechain rotation explains higher lit state <sup>19</sup>F NMR signal. In summary, 2fY182 labelling and solution <sup>19</sup>F NMR results provide evidence to support *OdAu1a*<sub>LOV</sub> structural changes observed by X-ray crystallography.

A clearer view of *OdAu1a* signalling mechanism is emerging. Structural studies of *OdAu1a*<sub>LOV</sub> provided insights suggesting different to *PtAu1a*<sub>LOV</sub> light dependent dimerisation mechanism. First of all, our results indicate that for *OdAu1a*<sub>LOV</sub> C-terminal Jα is not required for protein dimerisation and shows only minor structural changes upon photoactivation as observed by NMR studies. Furthermore, considering ΔJα*OdAu1a*<sub>LOV</sub> results it is evident that this small, 13.1 kDa in size, protein construct can be applied in the development of novel optogenetic tools. Secondly, by applying variable <sup>19</sup>F labelling methodology in conjugation with <sup>19</sup>F NMR spectroscopy we provide an alternative avenue in study LOV domain conformational changes. The utility of <sup>19</sup>F NMR spectroscopy has been demonstrated for in solution structural studies of N-terminal A'α helix supporting crystallographic observations. Considering that <sup>19</sup>F NMR is not limited by size as conventional <sup>15</sup>N/<sup>13</sup>C NMR, <sup>19</sup>F methodology provides a technique in the study of full length *OdAu1a* protein.

## 5.6 References

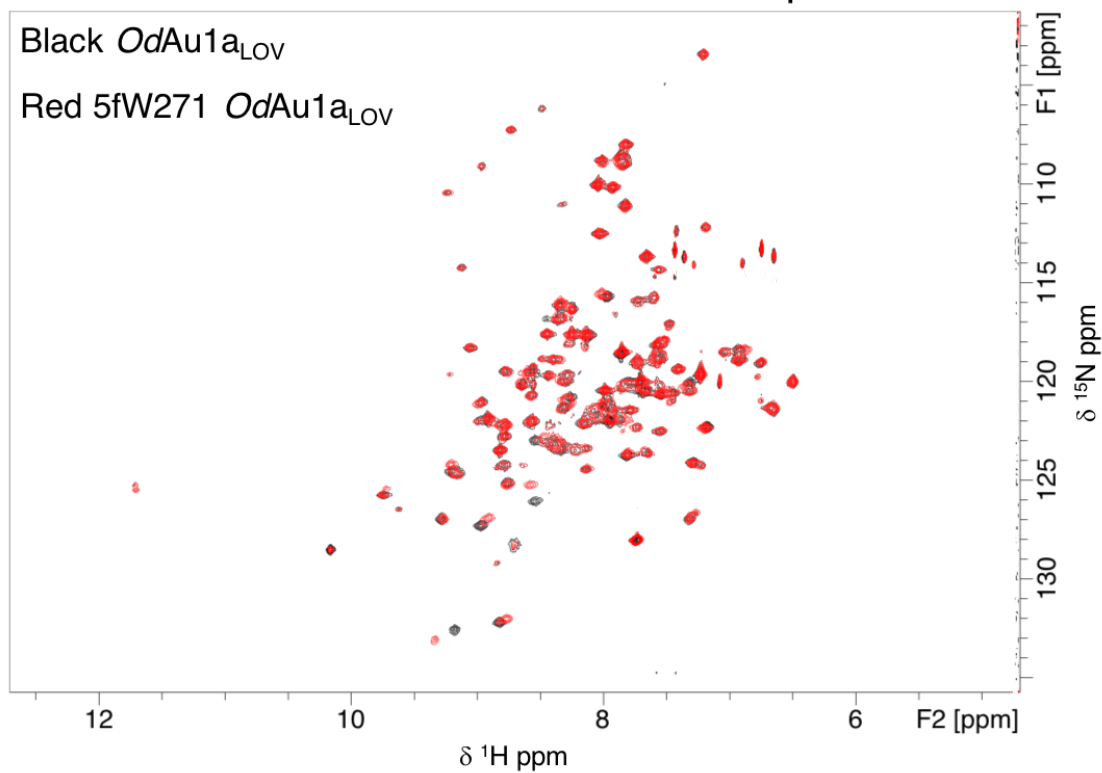
1. Takahashi, F. *et al.* AUREOCHROME, a Photoreceptor Required for Photomorphogenesis in Stramenopiles. *PNAS* **104**, 19625-19630 (2007).
2. Mitra, D., Yang, X. & Moffat, K. Crystal Structures of Aureochrome1 LOV Suggest New Design Strategies for Optogenetics. *Structure* **20**, 698-706 (2012).
3. Krauss, U., Krauss U, Minh, B. Q., Losi, A., Gärtner, W., Eggert, T., von Haeseler, A. & Jaeger, K. E. Distribution and Phylogeny of Light-Oxygen-Voltage-Blue-Light-Signaling Proteins in the Three Kingdoms of Life. *J. Bacteriol.* **191**, 7234-7242 (2009).
4. Takahashi, F. Blue-Light-Regulated Transcription Factor, Aureochrome, in Photosynthetic Stramenopiles. *J. Plant Res.* **129**, 189-197 (2016).
5. Banerjee, A., Herman, E., Serif, M., Maestre-Reyna, M., Hepp, S., Pokorny, R., Kroth, P. G., Essen, L. O. & Kottke T. Allosteric Communication Between DNA-binding and Light-Responsive Domains of Diatom Class i Aureochromes. *Nucleic Acids Res.* **44**, 5957-5970 (2016).
6. Banerjee, A., Herman, E., Kottke, T. & Essen, L. O. Structure of a Native-like Aureochrome 1a LOV Domain Dimer from *Phaeodactylum tricornutum*. *Structure* **24**, 171-178 (2016).
7. Heintz, U. & Schlichting, I. Blue Light-Induced LOV Domain Dimerization Enhances the Affinity of Aureochrome 1a for its Target DNA Sequence. *Elife* **5**, e11860 (2016).
8. Akiyama, Y., Nakasone, Y., Nakatani, Y., Hisatomi, O. & Terazima, M. Time-Resolved Detection of Light-Induced Dimerization of Monomeric Aureochrome-1 and Change in Affinity for DNA. *J. Phys. Chem. B* **120**, 7360-7370 (2016).
9. Kalvaitis, M. E., Johnson, L. A., Mart, R. J., Rizkallah, P. & Allemann, R. K. A Noncanonical Chromophore Reveals Structural Rearrangements of the Light-Oxygen-Voltage Domain upon Photoactivation. *Biochemistry* **58**, 2608-2616 (2019).
10. Herman, E. & Kottke, T. Allosterically Regulated Unfolding of the A'Alpha Helix Exposes the Dimerization Site of the Blue-Light-Sensing Aureochrome-LOV Domain. *Biochemistry* **54**, 1484-1492 (2015).
11. Herman, E., Sachse, M., Kroth, P. G. & Kottke, T. Blue-Light-Induced Unfolding of the J-Alpha Helix Allows for the Dimerization of Aureochrome-LOV from the diatom *Phaeodactylum tricornutum*. *Biochemistry* **52**, 3094–3101 (2013).

12. Carugo, O. How Large B-factors Can Be in Protein Crystal Structures. *BMC Bioinformatics* **19**, 61 (2018).
13. Radivojac, P., Obradovic, Z., Smith, D.K., Zhu, G., Vucetic, S., Brown, C.J., Lawson, J.D. & Dunker, A. K. Protein Flexibility and Intrinsic Disorder. *Protein Sci.* **13**, 71-80 (2004).
14. Sun, Z., Liu, Q., Qu, G., Feng, Y. & Reetz, M. T. Utility of B-Factors in Protein Science: Interpreting Rigidity, Flexibility, and Internal Motion and Engineering Thermostability. *Chem. Rev.* **19**, 1626-1665 (2019).
15. Kuriyan, J., Karplus, M. & Petsko, G. A. Estimation of Uncertainties in X-ray Refinement Results by Use of Perturbed Structures. *Proteins Struct. Funct. Genet.* **2**, 1-12 (1987).
16. Schreiner, M., Schlesinger, R., Heberle, J. & Niemann, H. H. Structure of Halorhodopsin from *Halobacterium salinarum* in a New Crystal Form that Imposes Little Restraint on the E-F loop. *J. Struct. Biol.* **190**, 373-378 (2015).
17. Khan, F., Kuprov, I., Craggs, T. D., Hore, P. J. & Jackson, S. E. 19F NMR Studies of the Native and Denatured States of Green Fluorescent Protein. *J. Am. Chem. Soc.* **128**, 10729-10737 (2006).
18. Sun, Z. Y., Truong, H. T., Pratt, E. A., Sutherland, D. C., Kulig, C. E., Homer, R.J., Groetsch, S. M., Hsue, P. Y. & Ho, C. A 19F-NMR Study of the Membrane-Binding Region of D-Lactate Dehydrogenase of *Escherichia coli*. *Protein Sci.* **2**, 1938-1947 (1993).
19. Minnihan, E. C., Young, D. D., Schultz, P. G. & Stubbe, J. Incorporation of Fluorotyrosines into Ribonucleotide Reductase Using an Evolved, Polyspecific Aminoacyl-tRNA Synthetase. *J. Am. Chem. Soc.* **133**, 15942-15945 (2011).
20. Oyala, P. H., Ravichandran, K. R., Funk, M. A., Stucky, P. A., Stich, T. A., Drennan, C.L., Britt, R.D. & Stubbe, J. Biophysical Characterization of Fluorotyrosine Probes Site-Specifically Incorporated into Enzymes: *E. coli* Ribonucleotide Reductase As an Example. *J. Am. Chem. Soc.* **138**, 7951-7964 (2016).
21. Curtis-Marof, R., Doko, D., Rowe, M. L., Richards, K. L., Williamson, R. A. & Howard, M. J. 19F NMR Spectroscopy Monitors Ligand Binding to Recombinantly Fluorine-Labelled b'x from Human Protein Disulphide Isomerase (hPDI). *Org. Biomol. Chem.* **12**, 3808-3812 (2014).
22. Richards, K. L., Rowe, M. L., Hudson, P. B., Williamson, R. A. & Howard, M. J. Combined Ligand-Observe <sup>19</sup>F and Protein-Observe <sup>15</sup>N,<sup>1</sup>H-HSQC NMR Suggests

- Phenylalanine as the Key  $\delta$ -Somatostatin Residue Recognized by Human Protein Disulfide Isomerase. *Sci. Rep.* **6**, 19518 (2016).
23. Harper, S. M., Neil, L.C. & Gardner, K. H. Structural Basis of a Phototropin Light Switch. *Science* **301**, 1541-1544 (2003).
24. Hou, Y., Hu, W., Li, X., Skinner, J. J., Liu, D. & Wüthrich K. Solvent-Accessibility of Discrete Residue Positions in the Polypeptide Hormone Glucagon by  $^{19}\text{F}$ -NMR Observation of 4 Fluorophenylalanine. *J Biomol NMR* **68**, 1-6 (2017).
25. Danielson, M.A. & Falke, J.J. Use of  $^{19}\text{F}$  NMR to Probe Protein Structure and Conformational Changes. *Annu Rev Biophys Biomol Struct.* **25**, 163-195 (1996).

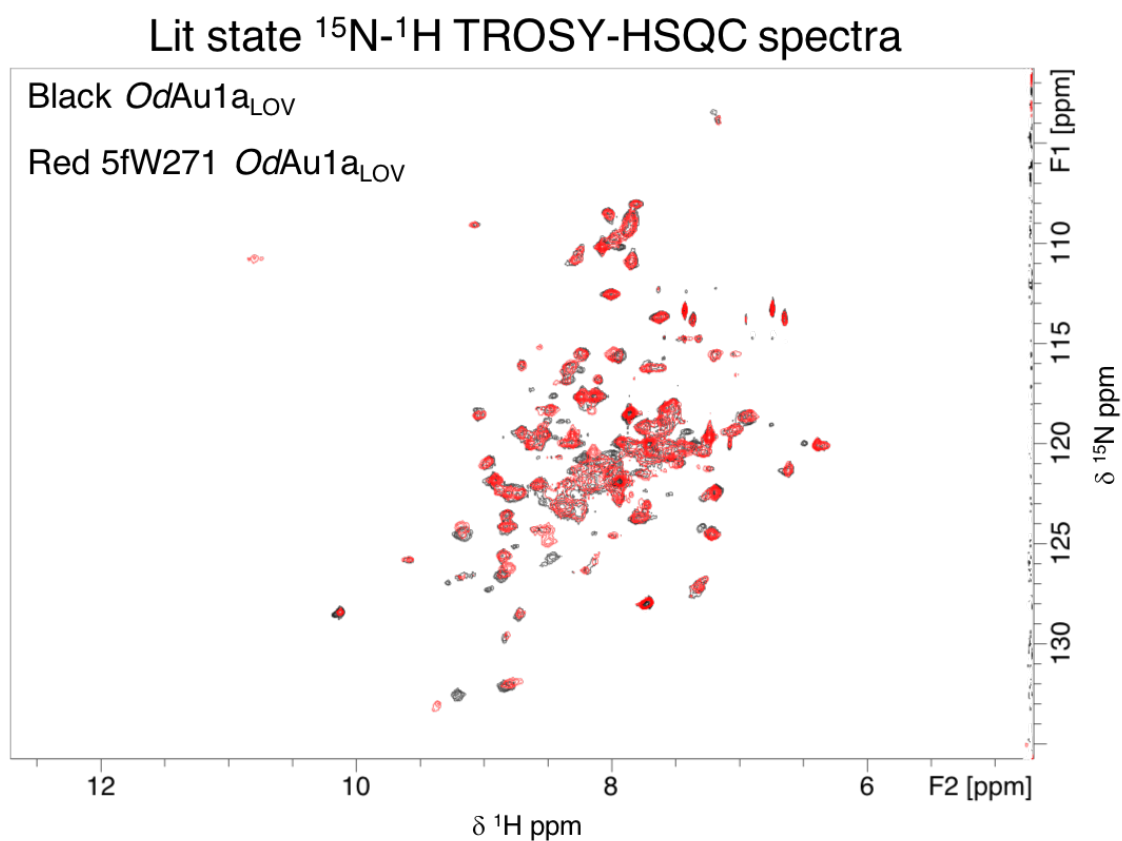
## 5.7 Appendix

### Dark state $^{15}\text{N}$ - $^1\text{H}$ TROSY-HSQC spectra



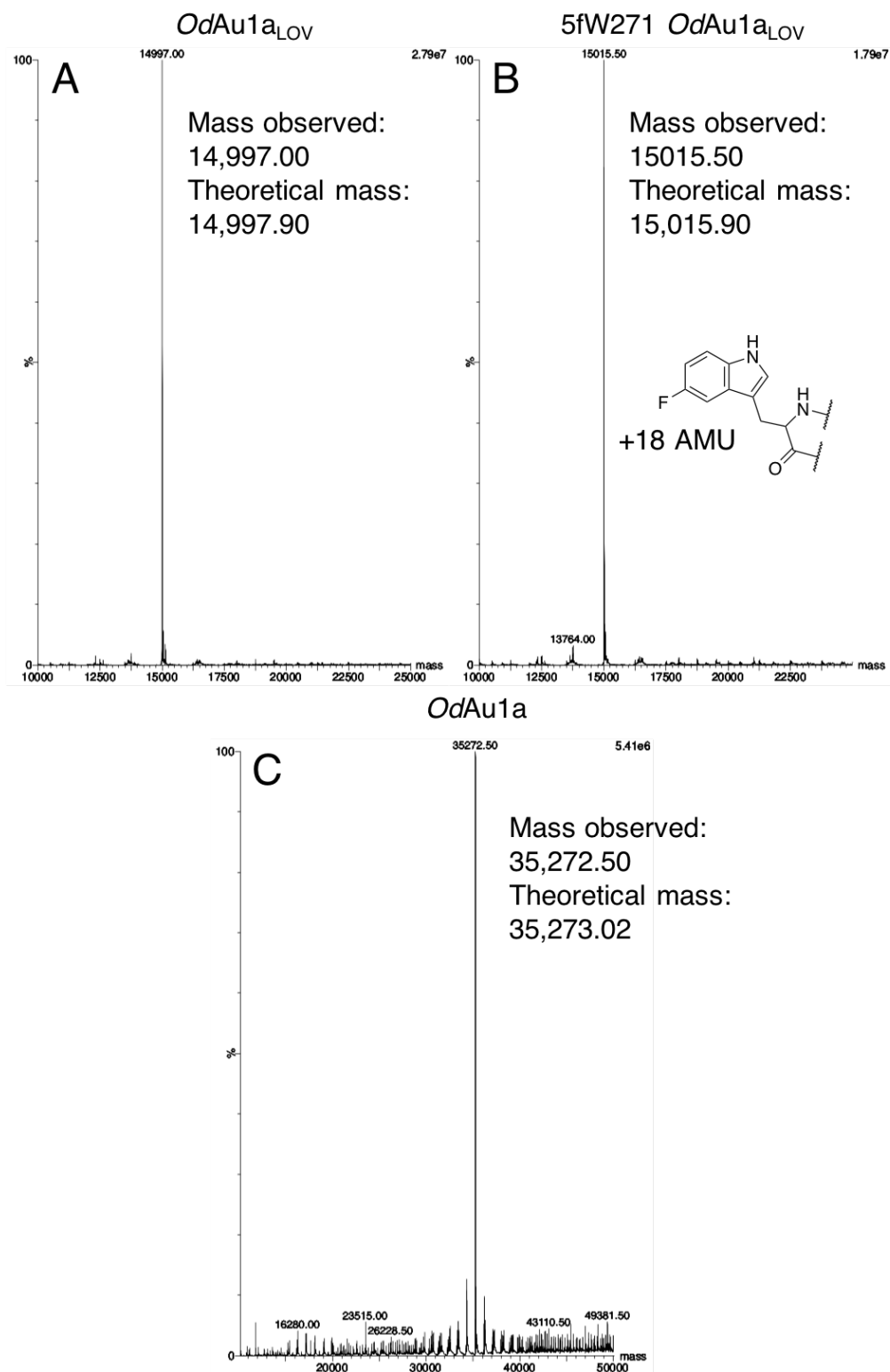
**Figure 5.19** Dark state  $^{15}\text{N}$ - $^1\text{H}$  TROSY-HSQC spectra of *OdAu1a*<sub>LOV</sub> and 5fW271 labelled *OdAu1a*<sub>LOV</sub>.

Dark state  $^{15}\text{N}$ - $^1\text{H}$  TROSY-HSQC spectra of 500  $\mu\text{M}$  *OdAu1a*<sub>LOV</sub>, black, and 400  $\mu\text{M}$  5fW271 labelled *OdAu1a*<sub>LOV</sub>, red. The figure indicates that the majority of resonances overlapped indicative of an identical structural fold.



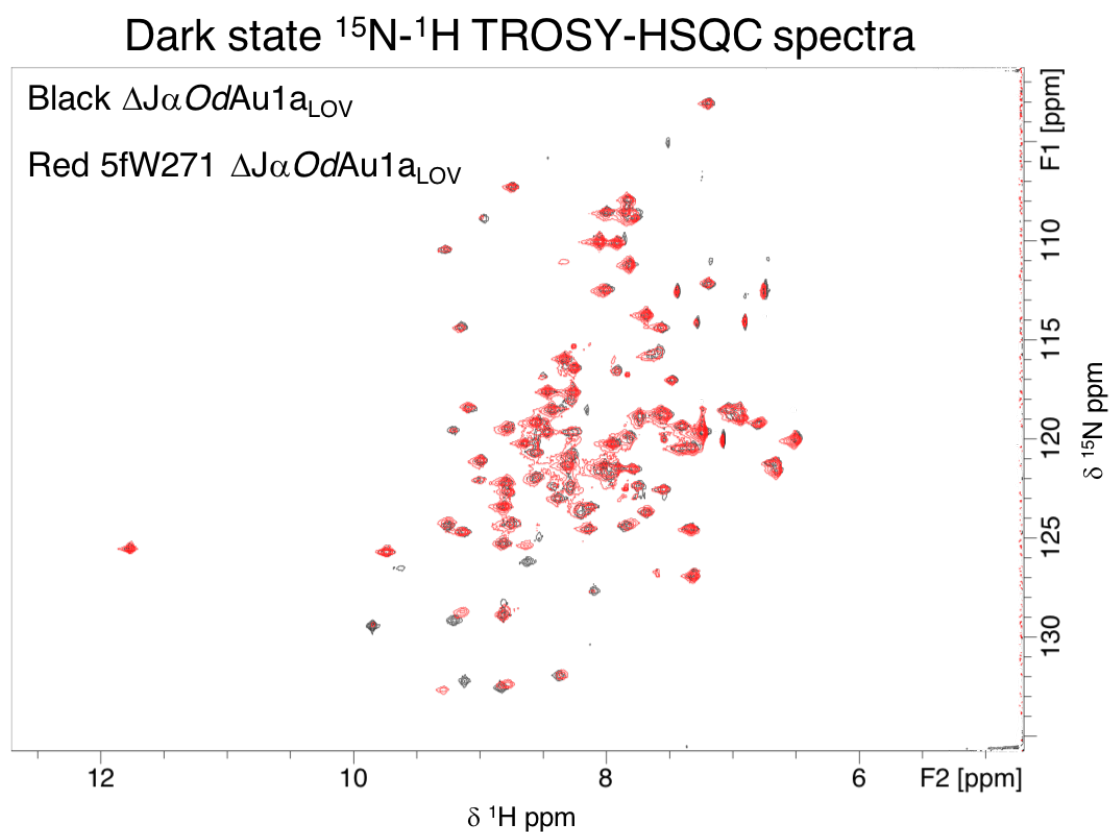
**Figure 5.20** Lit state  $^{15}\text{N}$ - $^1\text{H}$  TROSY-HSQC spectra of *OdAu1a<sub>LOV</sub>* and 5fW271 labelled *OdAu1a<sub>LOV</sub>*.

Lit state  $^{15}\text{N}$ - $^1\text{H}$  TROSY-HSQC spectra of 500  $\mu\text{M}$  *OdAu1a<sub>LOV</sub>*, black, and 400  $\mu\text{M}$  5fW271 labelled *OdAu1a<sub>LOV</sub>*, red. The figure indicates that the majority of resonances overlapped indicative of an identical structural fold.



**Figure 5.21** LC-MS analysis of *OdAu1a*<sub>LOV</sub> and *OdAu1a*.

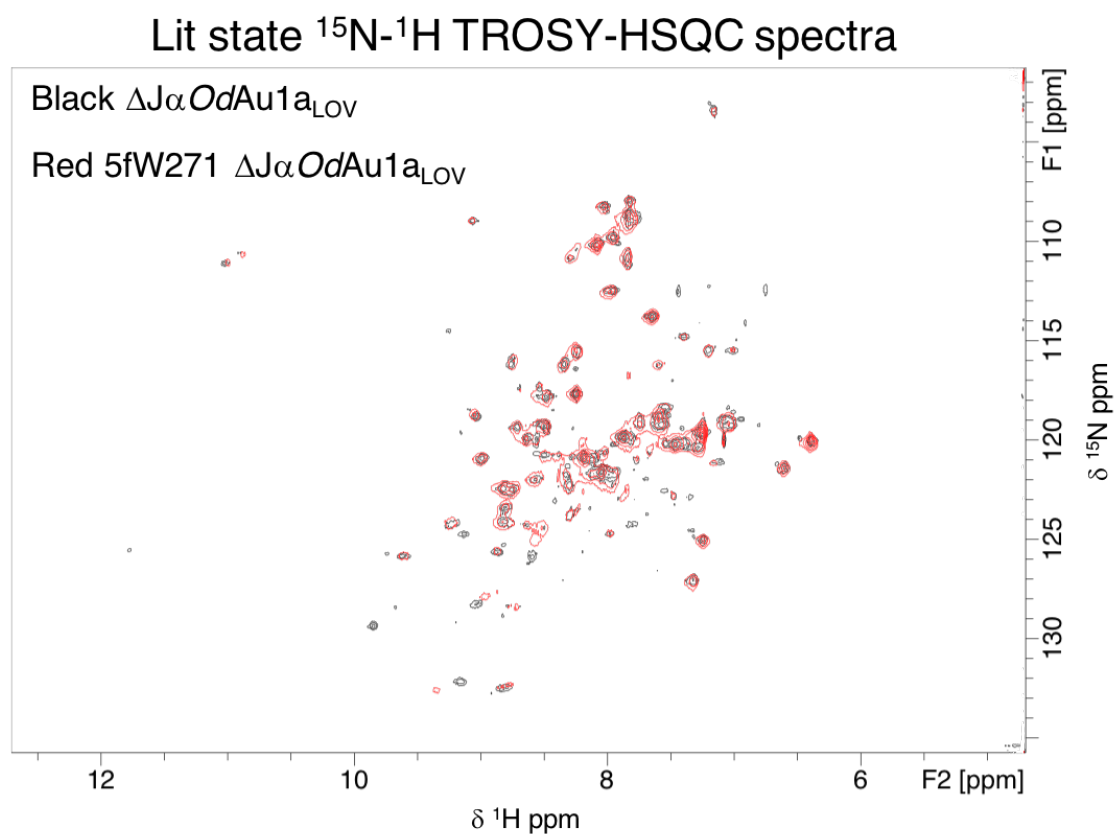
Molecular masses observed for A) *OdAu1a*<sub>LOV</sub>, B) *OdAu1a*<sub>LOV</sub> labelled with 5-fluoro-L-tryptophan (5fW) with the structure indicated in the figure and C) *OdAu1a*. Please note the error of the experiment was 1.0 atomic mass units (AMU).



**Figure 5.22** Dark state  $^{15}\text{N}$ - $^1\text{H}$  TROSY-HSQC spectra of  $\Delta\text{J}\alpha\text{OdAu1a}_{\text{LOV}}$  and 5fW271 labelled  $\Delta\text{J}\alpha\text{OdAu1a}_{\text{LOV}}$ .

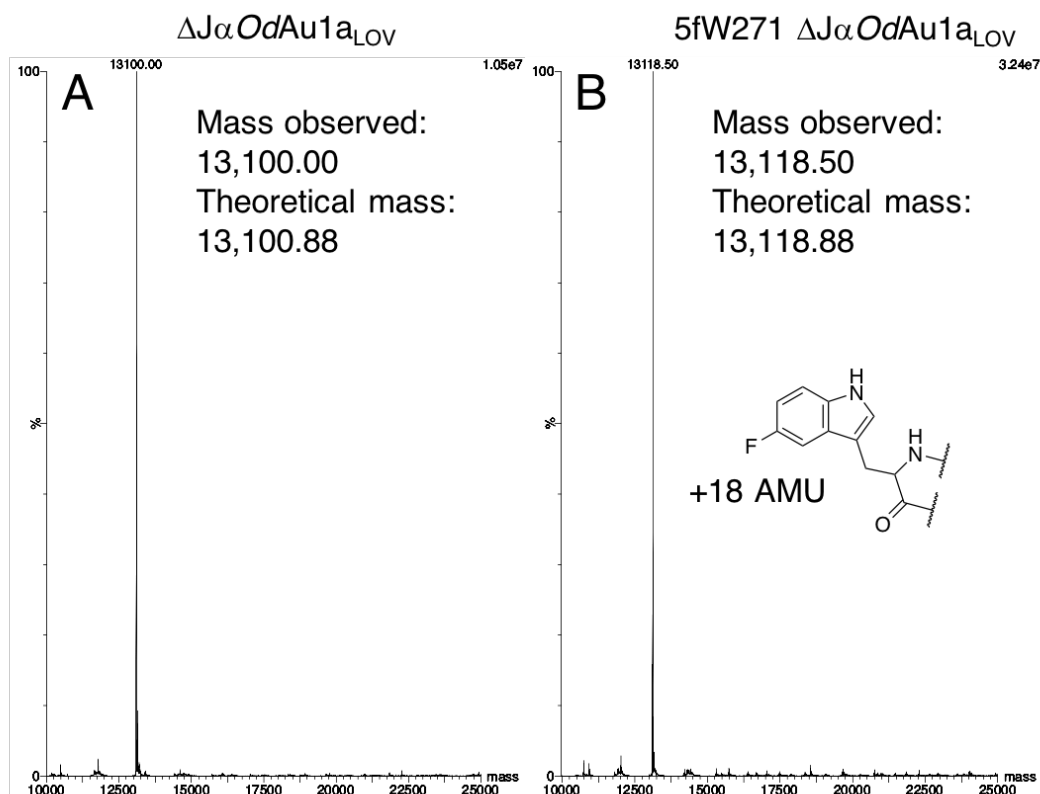
Dark state  $^{15}\text{N}$ - $^1\text{H}$  TROSY-HSQC spectra of 400  $\mu\text{M}$   $\Delta\text{J}\alpha\text{OdAu1a}_{\text{LOV}}$ , black, and 400  $\mu\text{M}$  5fW271 labelled  $\Delta\text{J}\alpha\text{OdAu1a}_{\text{LOV}}$ , red. The figure indicates that the majority of resonances overlapped indicative of an identical structural fold.





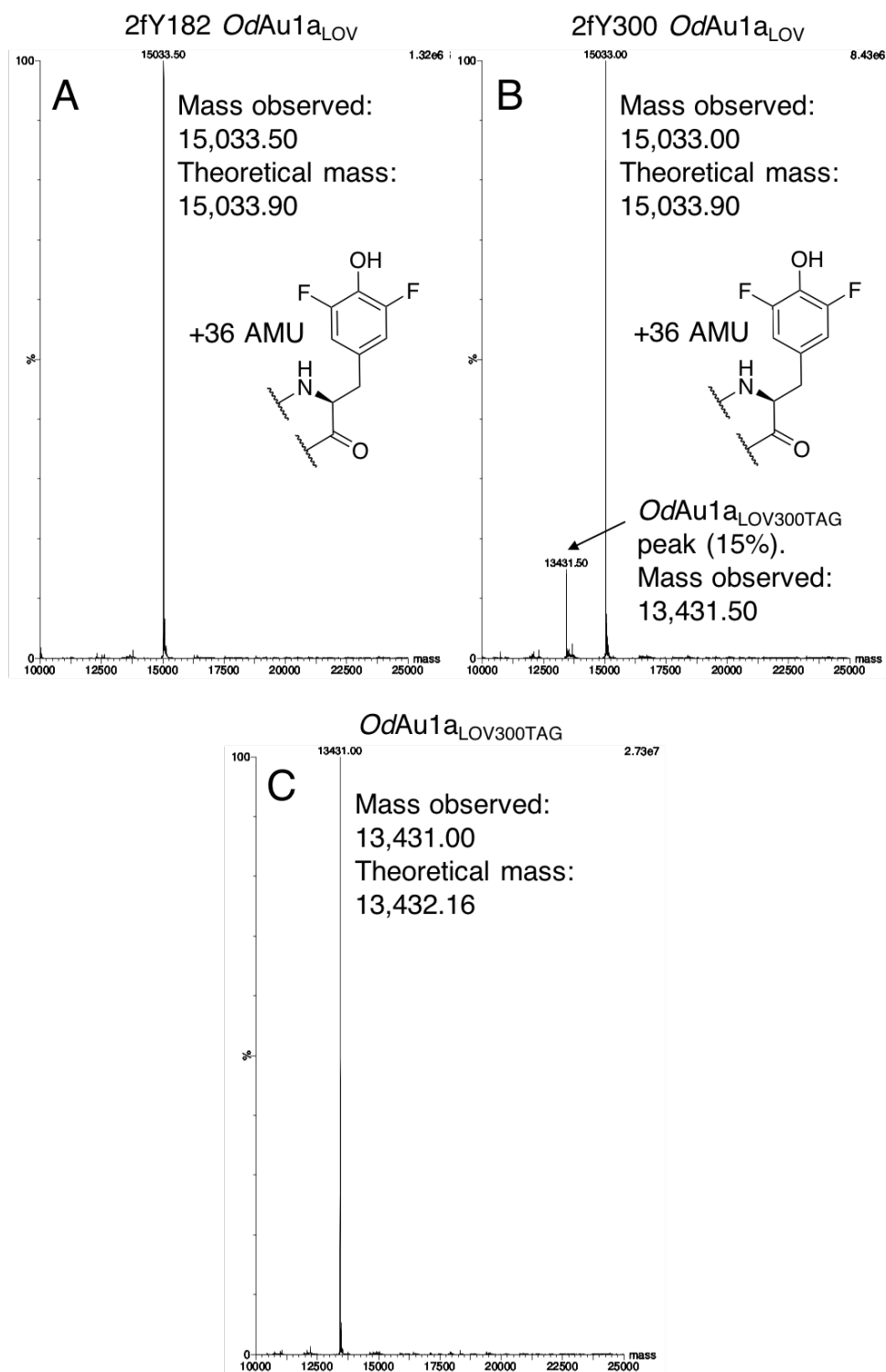
**Figure 5.23** Lit state  $^{15}\text{N}$ - $^1\text{H}$  TROSY-HSQC spectra of  $\Delta\text{J}\alpha\text{OdAu1a}_{\text{LOV}}$  and 5fW271 labelled  $\Delta\text{J}\alpha\text{OdAu1a}_{\text{LOV}}$ .

Lit state  $^{15}\text{N}$ - $^1\text{H}$  TROSY-HSQC spectra of 400  $\mu\text{M}$   $\Delta\text{J}\alpha\text{OdAu1a}_{\text{LOV}}$ , black, and 400  $\mu\text{M}$  5fW271 labelled  $\Delta\text{J}\alpha\text{OdAu1a}_{\text{LOV}}$ , red. The figure indicates that the majority of resonances overlapped indicative of an identical structural fold.



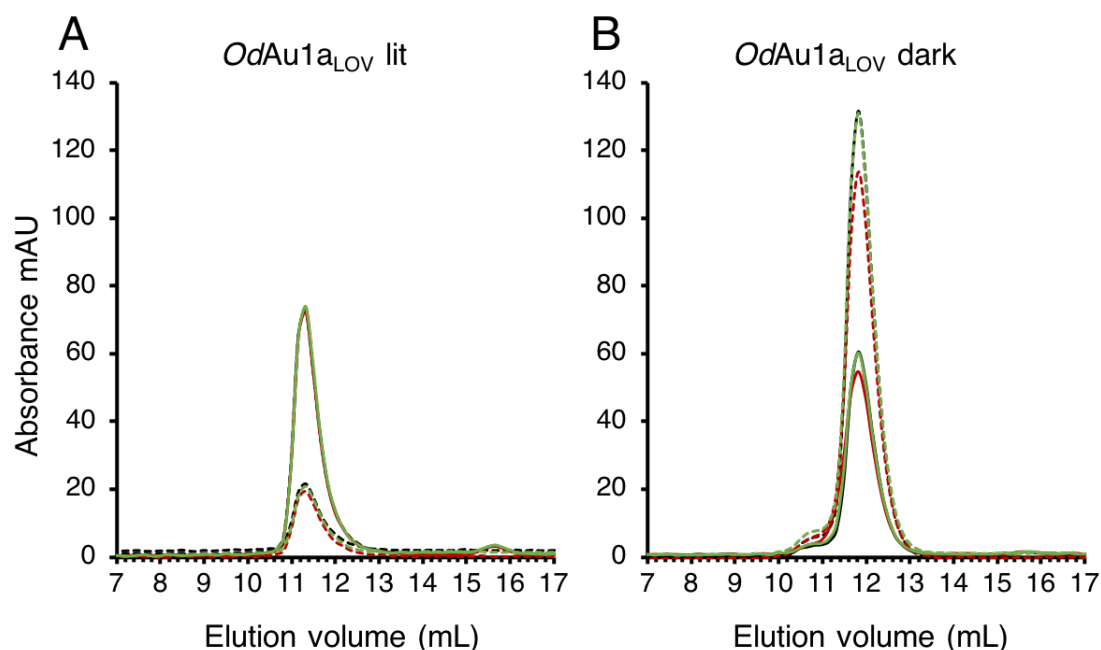
**Figure 5.24 LC-MS analysis of  $\Delta J\alpha OdAu1a_{LOV}$ .**

Molecular masses observed for A)  $\Delta J\alpha OdAu1a_{LOV}$  and B)  $\Delta J\alpha OdAu1a_{LOV}$  labelled with 5-fluoro-L-tryptophan (5fW) with structure indicated in the figure. Please note the error of the experiment was 1.0 AMU.



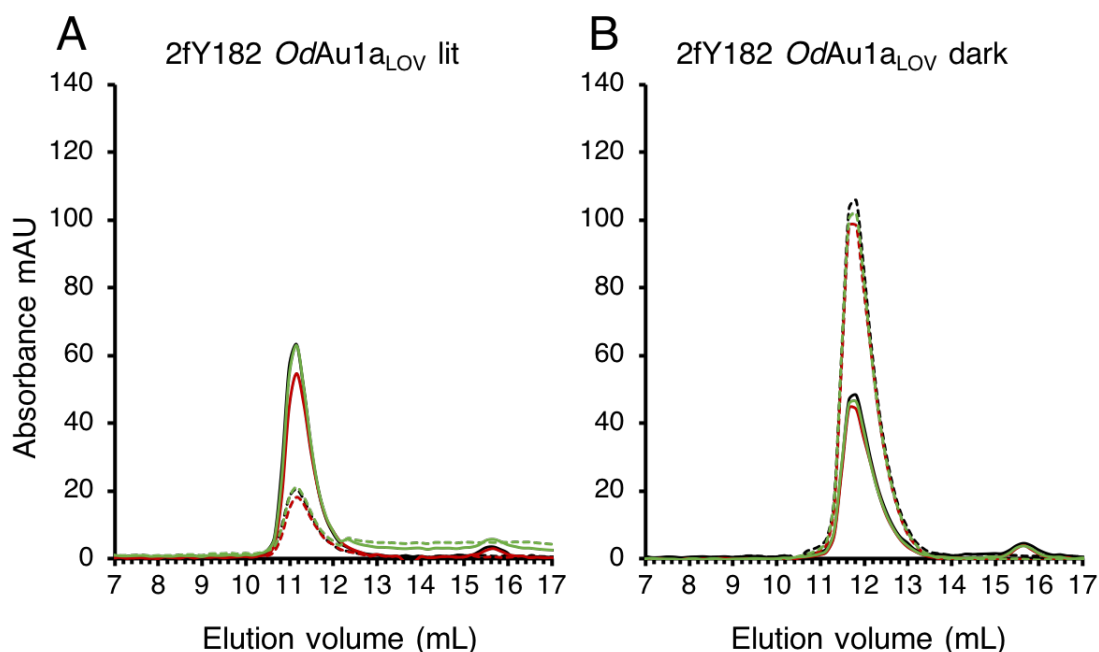
**Figure 5.25** LC-MS analysis of 2fY182 and 2fY300 labelled *OdAu1a*<sub>LOV</sub>, and *OdAu1a*<sub>LOV300TAG</sub>.

Molecular masses observed for *OdAu1a*<sub>LOV</sub> labelled with 3,5-difluoro-L-tryptophan (2fY, indicated in the figure). A) 2fY182 labelled *OdAu1a*<sub>LOV</sub> indicating a single species. B) 2fY300 labelled *OdAu1a*<sub>LOV</sub> indicating a correct mass species and a truncation product with a mass of 13,431.50 AMU, *OdAu1a*<sub>LOV300TAG</sub> 15% intensity. C) *OdAu1a*<sub>LOV300TAG</sub> mass. The mass of *OdAu1a*<sub>LOV300TAG</sub> confirms that the signal observed in B) is 15% truncation product. Please note the error of the experiment was 1.0 AMU.



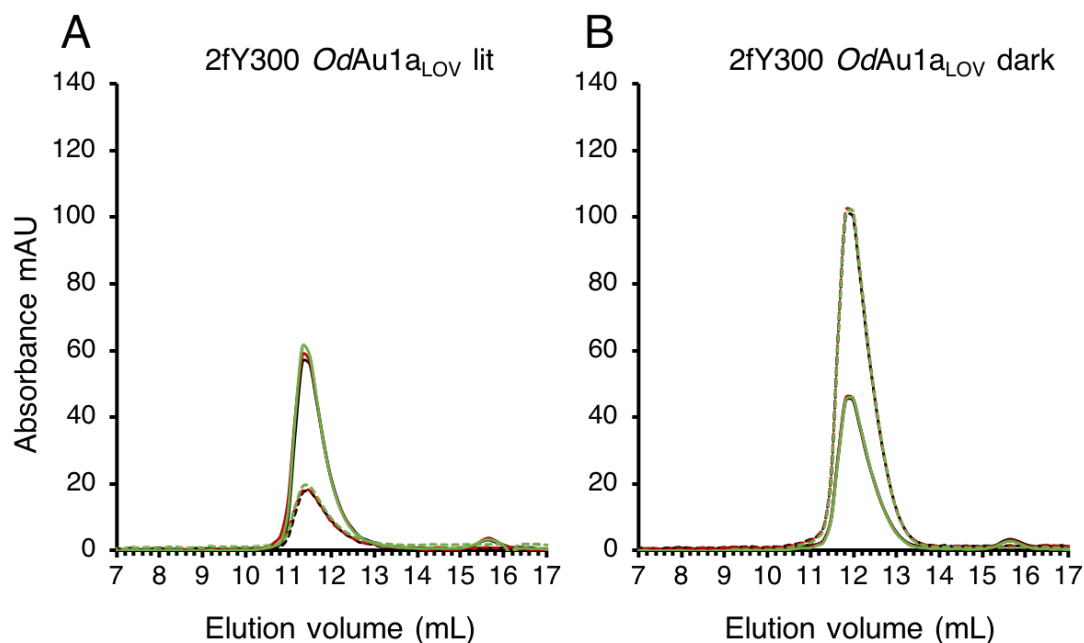
**Figure 5.26 SEC75 characterisation of *OdAu1a<sub>LOV</sub>*.**

SEC75 characterisation of 100  $\mu$ M *OdAu1a<sub>LOV</sub>*. All experiments were performed in triplicate where run 1 is coloured black, run 2 coloured red and run 3 coloured green, respectively. Throughout the experiments 450 nm (FMN dark state absorbance maxima) and 390 nm (lit state FMN absorbance maxima) absorbances were detected to confirm the photoactivation state of the protein. Absorbance at 450 nm is depicted as solid and absorbance at 390 nm as dashed lines, respectively, for each repeat. A) Three repeats of SEC75 trace of lit state *OdAu1a<sub>LOV</sub>*. Please note absorbance at 390 nm (solid lines) was higher than at 450 nm (dashed lines) confirming photoactivated protein. Elution maxima was 11.2 mL. B) Three repeats of SEC75 trace of dark state *OdAu1a<sub>LOV</sub>*. Please note absorbance at 390 nm (solid lines) was lower than at 450 nm (dashed lines) confirming dark state protein. Elution maxima was 11.8 mL.



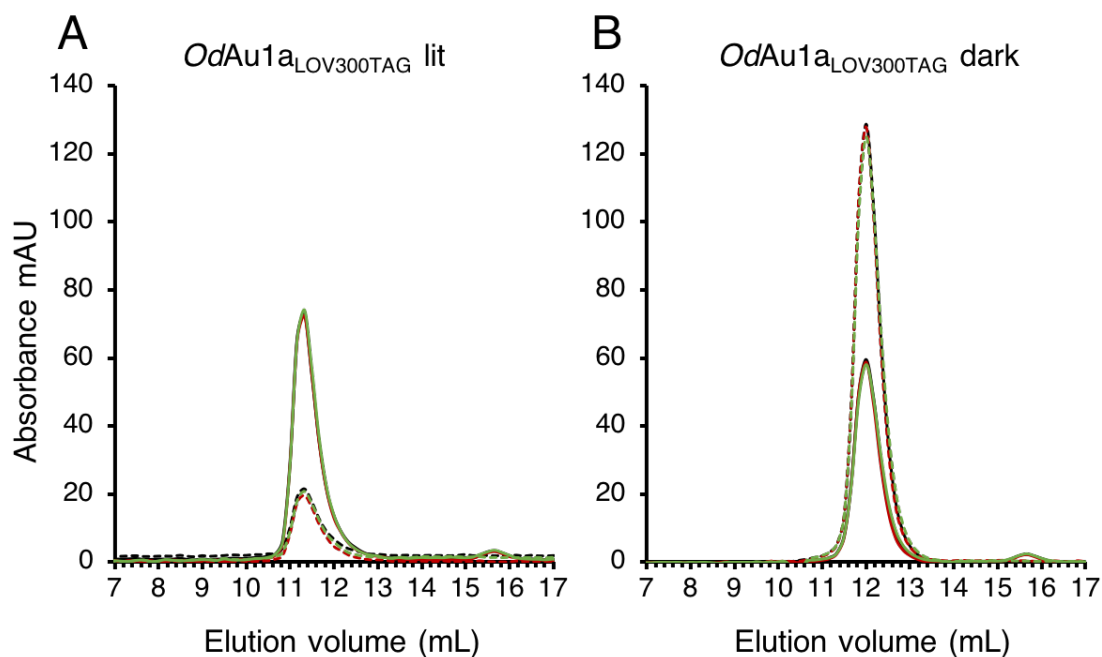
**Figure 5.27 SEC75 characterisation of 2fY182 labelled *OdAu1a*<sub>LOV</sub>.**

SEC75 characterisation of 100  $\mu$ M 2fY182 labelled *OdAu1a*<sub>LOV</sub>. All experiments were performed in triplicate where run 1 is coloured black, run 2 coloured red and run 3 coloured green, respectively. Throughout the experiments 450 nm (FMN dark state absorbance maxima) and 390 nm (lit state FMN absorbance maxima) absorbances were detected to confirm the photoactivation state of the protein. Absorbance at 450 nm is depicted as solid and absorbance at 390 nm as dashed lines, respectively, for each repeat. A) Three repeats of SEC75 trace of lit state 2fY182 *OdAu1a*<sub>LOV</sub>. Please note absorbance at 390 nm (solid lines) was higher than at 450 nm (dashed lines) confirming photoactivated protein. Elution maxima was 11.1 mL. B) Three repeats of SEC75 trace of dark state 2fY182 *OdAu1a*<sub>LOV</sub>. Please note absorbance at 390 nm (solid lines) was lower than at 450 nm (dashed lines) confirming dark state protein. Elution maxima was 11.8 mL.



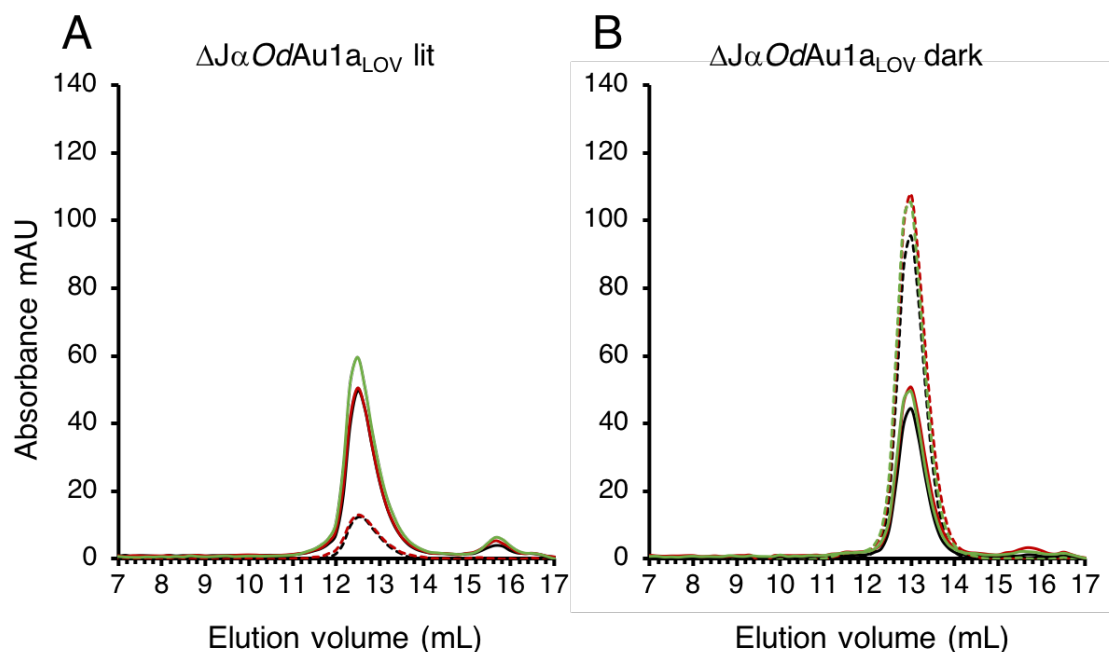
**Figure 5.28 SEC75 characterisation of 2fY300 *OdAu1a*<sub>LOV</sub>.**

SEC75 characterisation of 100  $\mu$ M 2fY300 labelled *OdAu1a*<sub>LOV</sub>. All experiments were performed in triplicate where run 1 is coloured black, run 2 coloured red and run 3 coloured green, respectively. Throughout the experiments 450 nm (FMN dark state absorbance maxima) and 390 nm (lit state FMN absorbance maxima) absorbances were detected to confirm the photoactivation state of the protein. Absorbance at 450 nm is depicted as solid and absorbance at 390 nm as dashed lines, respectively, for each repeat. A) Three repeats of SEC75 trace of lit state 2fY300 *OdAu1a*<sub>LOV</sub>. Please note absorbance at 390 nm (solid lines) was higher than at 450 nm (dashed lines) confirming photoactivated protein. Elution maxima was 11.4 mL. B) Three repeats of SEC75 trace of dark state 2fY300 *OdAu1a*<sub>LOV</sub>. Please note absorbance at 390 nm (solid lines) was lower than at 450 nm (dashed lines) confirming dark state protein. Elution maxima was 11.9 mL.



**Figure 5.29 SEC75 characterisation of *OdAu1a<sub>LOV300TAG</sub>*.**

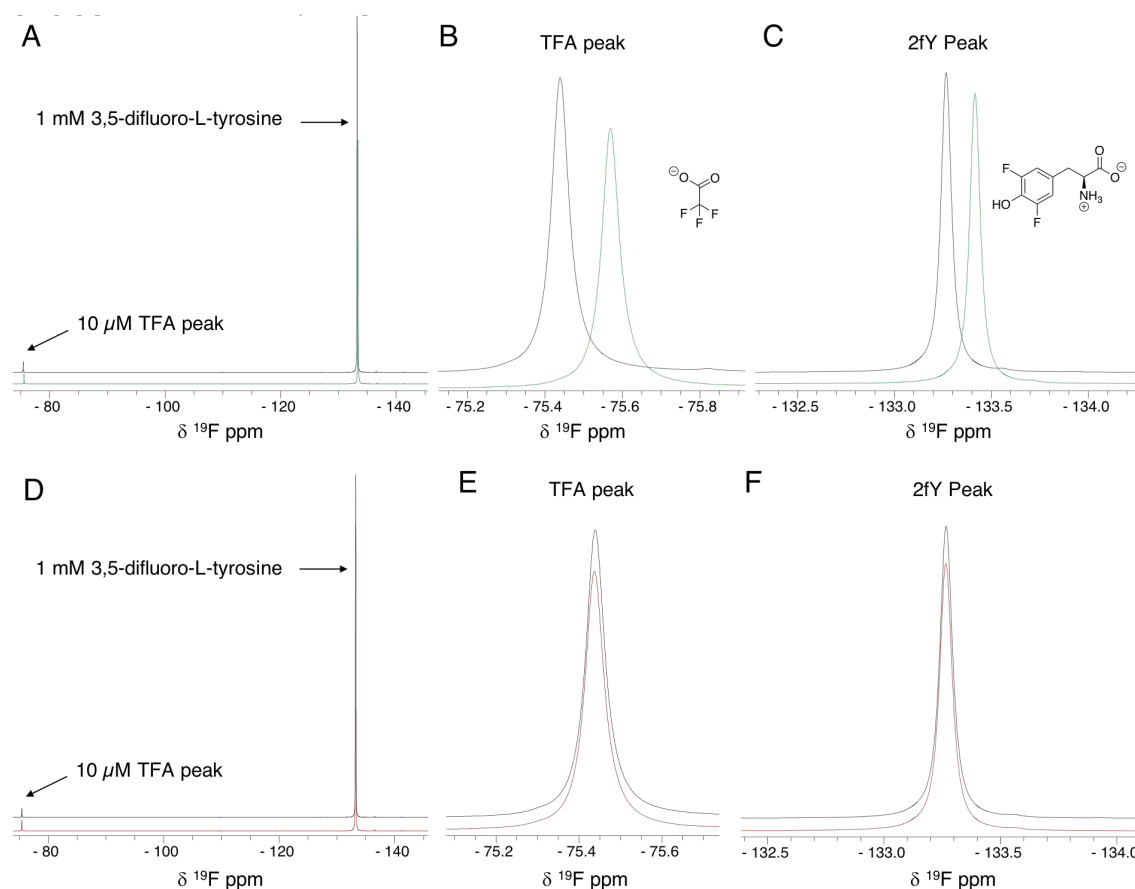
SEC75 characterisation of 100  $\mu$ M *OdAu1a<sub>LOV300TAG</sub>*. All experiments were performed in triplicate where run 1 is coloured black, run 2 coloured red and run 3 coloured green, respectively. Throughout the experiments 450 nm (FMN dark state absorbance maxima) and 390 nm (lit state FMN absorbance maxima) absorbances were detected to confirm the photoactivation state of the protein. Absorbance at 450 nm is depicted as solid and absorbance at 390 nm as dashed lines, respectively, for each repeat. A) Three repeats of SEC75 trace of lit state *OdAu1a<sub>LOV300TAG</sub>*. Please note absorbance at 390 nm (solid lines) was higher than at 450 nm (dashed lines) confirming photoactivated protein. Elution maxima was 11.3 mL. B) Three repeats of SEC75 trace of dark state *OdAu1a<sub>LOV300TAG</sub>*. Please note absorbance at 390 nm (solid lines) was lower than at 450 nm (dashed lines) confirming dark state protein. Elution maxima was 12.0 mL.



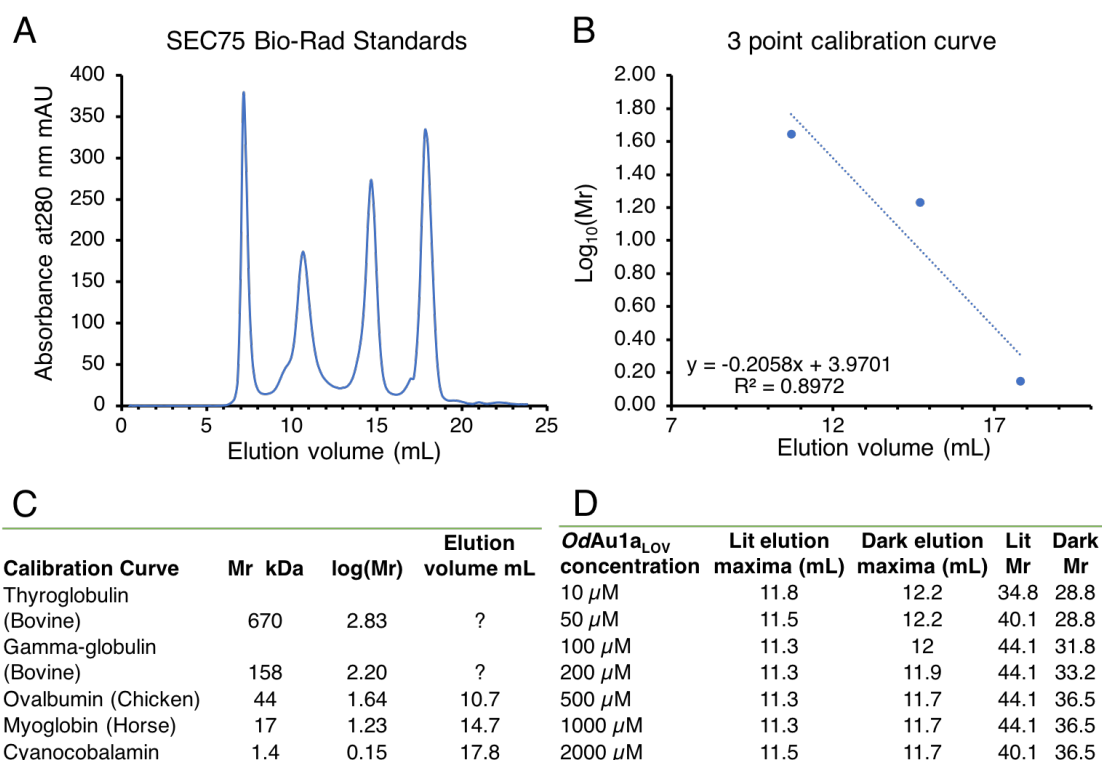
**Figure 5.30 SEC75 characterisation of  $\Delta J\alpha OdAu1a_{LOV}$ .**

SEC75 characterisation of 100  $\mu$ M  $\Delta J\alpha OdAu1a_{LOV}$ . All experiments were performed in triplicate where run 1 is coloured black, run 2 coloured red and run 3 coloured green, respectively. Throughout the experiments 450 nm (FMN dark state absorbance maxima) and 390 nm (lit state FMN absorbance maxima) absorbances were detected to confirm the photoactivation state of the protein. Absorbance at 450 nm is depicted as solid and absorbance at 390 nm as dashed lines, respectively, for each repeat. A) Three repeats of SEC75 trace of lit state  $\Delta J\alpha OdAu1a_{LOV}$ . Please note absorbance at 390 nm (solid lines) was higher than at 450 nm (dashed lines) confirming photoactivated protein. Elution maxima was 12.5 mL. B) Three repeats of SEC75 trace of dark state  $\Delta J\alpha OdAu1a_{LOV}$ . Please note absorbance at 390 nm (solid lines) was lower than at 450 nm (dashed lines) confirming dark state protein. Elution maxima was 13.0 mL.





**Figure 5.31  $^{19}\text{F}$  NMR spectra of 3,5-difluoro-L-tyrosine (2fY) in 10 and 100%  $\text{D}_2\text{O}$  buffer.** A)  $^{19}\text{F}$  NMR spectra 3,5-difluoro-L-tyrosine (2fY) in 10%  $\text{D}_2\text{O}$ , black, and 100%  $\text{D}_2\text{O}$  buffer, green, respectively. Peaks corresponding to trifluoroacetate (TFA) and 2fY are indicated with arrows. B) At TFA peak zoomed  $^{19}\text{F}$  NMR spectra indicating TFA  $\delta$  -75.44 ppm 10%  $\text{D}_2\text{O}$  and  $\delta$  -75.57 ppm 100%  $\text{D}_2\text{O}$ . SIIS upfield  $\Delta$   $\delta$  -0.13 ppm. C) At 2fY peak zoomed  $^{19}\text{F}$  NMR spectra indicating 2fY  $\delta$  -133.27 ppm 10%  $\text{D}_2\text{O}$  and  $\delta$  -133.42 ppm 100%  $\text{D}_2\text{O}$ . SIIS upfield  $\Delta$   $\delta$  -0.15 ppm. D)  $^{19}\text{F}$  NMR spectra of a control experiment investigating if illumination with 473 nm light does not perturb chemical shifts of TFA and 2fY. Black spectra indicates experiment run in the absence of light and red spectra when sample was continuously illuminated with 473 nm light, respectively. No chemical shifts were evident indicating that light did not influence  $^{19}\text{F}$  NMR signal. E) and F) are zoomed regions of spectra depicted in D) where E) indicates TFA and F) 2fY peak.



**Figure 5.32 Calibration of SEC75 column and investigation of *OdAu1a<sub>LOV</sub>* light dependent change in elution maxima.**

A) Chromatogram of gel filtration standards of known molecular mass indicating four distinct peaks only. Peak at 7.2 mL most likely indicates protein aggregates. Experiment was run in NMR like buffer (20 mM MES, 1 mM TCEP, 1 mM EDTA, 5 mM NaN<sub>3</sub>, pH 6.0), which might not be optimal for the standards used (Bio-Rad). B) Calibration curve of standard proteins only using elution maxima of three peaks observed in the chromatograph in A). Please note R<sup>2</sup> value was very poor. C) Table of protein standards, corresponding elution maxima and molecular sizes. Peaks corresponding to gamma-globulin and thyroglobulin could not be identified in the SEC75 trace. D) Estimated molecular sizes of *OdAu1a<sub>LOV</sub>* under dark and lit states, respectively. Considering that only three point calibration curve was used, accurate molecular sizes could not be determined.

**Table 5.1 Dataset for  $\Delta J\alpha OdAu1a_{LOV}$ .**

PDB Entry	$\Delta J\alpha OdAu1a_{LOV}$
<b>Data Collection</b>	
<b>Crystal Data</b>	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	105.00, 105.00, 68.12
<i>h</i> , <i>k</i> , <i>l</i>	90.0, 90.0, 120.0
Space group	P 3 <sub>1</sub> 2 1
Resolution (Å)	1.99 – 54.52
Outer shell	1.99 – 2.02
<i>R</i> -merge (%)	0.074 (1.310)
<i>R</i> -meas (%)	0.077 (1.432)
CC1/2	0.999 (0.799)
<i>I</i> / $\sigma(I)$	20.11 (2.00)
Completeness (%)	100.00 (100.00)
Multiplicity	16.73 (11.22)
Total Measurements	503,109 (16,886)
Unique Reflections	30,073 (1,505)
Wilson B-factor(Å <sup>2</sup> )	38.272
<b>Refinement Statistics</b>	
Non-H Atoms	2098
R-work reflections	30,052
R-free reflections	1,533
R-work/R-free	0.190 / 0.201
<b>rms deviations</b>	
Bond lengths (Å)	0.0074
Bond Angles (°)	1.295
Coordinate error	0.110
Mean B value (Å <sup>2</sup> )	43.551
<b>Ramachandran Statistics</b>	
Favoured/allowed/Outliers	226 / 2 / 0
%	99.12/ 0.88 / 0





## 6 Future Work

### 6.1 Chapter 3 Future Work

Considering the fact that this work provided evidence for a previously structurally uncharacterised state, one of the approaches would be to attempt to study protein with a structural technique which would not result in the decay of the covalent adduct. One such approach would be a room temperature neutron diffraction which in conjugation with a blue light optical fibre or a laser would allow photoactivation of the protein throughout data collection. Although the resolution reported here are good enough to assign side chain positions and rotamers, the observed mixed states makes it difficult to assign residue hydrogen bonds unambiguously. Unlike X-rays which interact and scatter off electrons, neutrons interact and scatter off nuclei therefore making it possible to observe hydrogen atoms allowing unambiguous assignment of hydrogen bonds. Considering the size of the crystals achieved throughout the study, Figure 3.3 O and P, and unparalleled diffraction data for any known Au1a<sub>LOV</sub> homologue, this system provides a plausible path to further understand LOV domain structural changes. Throughout the project, collaboration with Dr Mark Wilson, Redox Biology Centre, University of Nebraska, USA, was established by Prof Rudolf K Allemann in further study of *Od*Au1a<sub>LOV</sub>.

The proposed mechanism for the propagation of the lit state information through structural changes could be further explored using site-directed mutagenesis. Key asparagine residues Asn194 and Asn272 highlighted in this work could be probed using this methodology. To the best of our knowledge, no reports have investigated if these residues are crucial for Au1a<sub>LOV</sub> dimerisation.

Another way to investigate *Od*Au1a<sub>LOV</sub> would be to use NMR spectroscopy to pinpoint structural changes in the solution at individual amino acid residue resolution. Throughout this project excluding U-<sup>15</sup>N labelled *Od*Au1a<sub>LOV</sub>, U-<sup>2</sup>H-<sup>13</sup>C-<sup>15</sup>N protein sample was also prepared and was used for triple resonance experiments for protein backbone assignment by Dr Luke Johnson, Cardiff University, UK. Furthermore, optic fiber system was introduced to allow the study of photoactivated LOV domain proteins. This opens yet another possibility to understand structural changes and investigate structural changes in PAS domain core and the auxiliary A'α and Jα helices.

## 6.2 Chapter 4 Future Work

To provide further evidence towards the functional responsiveness and the aspects of structural changes for 5dFMN containing AsLOV2<sub>FL</sub>, J $\alpha$  anchoring mutants previously described in the literature could be investigated. If the hypothesis that 5dFMN in the lit state behaves more similarly to FMN through destabilisation of hydrogen bonding extending from FMN-O4, mutants with tighter J $\alpha$  caging should still provide similar light dependent <sup>19</sup>F chemical shifts and observable changes in the CD spectra. If constructs with tighter J $\alpha$  would fail to provide observable light responsiveness, this would then suggest that protonation at N5 in conjugation with the hydrogen bonds around FMN-O4 changes dictate structural changes. Conversely, mutants with weaker J $\alpha$  helix anchoring should provide more of the secondary conformer for 5fW signals as well as lower CD signal. This in turn would suggest that to lock protein in the dark state with docked C-termini, hydrogen bonding to O4 or even N5 of FMN are essential and 5dFMN is unable to fulfill these requirements. Alternatively, if dark state 5dFMN in these mutants would behave like FMN, this would then suggest that hydrogen bonding to 5dFMN suffice to lock protein in the dark-like state.

Although evidence regarding 5dFMN functionality in AsLOV2 and AsLOV2<sub>FL</sub> have been provided using <sup>19</sup>F NMR and CD, it would also be possible to attempt to pinpoint structural differences by using U-<sup>15</sup>N or U-<sup>15</sup>N/<sup>13</sup>C labelled samples to provide further structural information. By assigning 5dFMN AsLOV2<sub>FL</sub> <sup>15</sup>N-<sup>1</sup>H HSQC spectra, it would be possible to identify structural differences at individual residue resolution. Additionally, all experiments were performed at a constant 293 K temperature allowing direct comparison of structural changes to the published results. The next step would be to investigate 5fW chemical shifts and CD spectra results in respect to the temperature, from low to high. At higher temperatures, where the system contains more thermal energy, a higher population of unfolded J $\alpha$  should be observed for the dark states. If this unfolding in dark state is representative of dark minor peak for FMN (at  $\delta$  -49.6 ppm) and dark peak 1 for 5dFMN (at  $\delta$  -49.6 ppm), peak integrals in respect to major peak should increase upon heating of the sample. Plotting of peak integrals vs temperature and the corresponding slope values would allow characterisation of differences in J $\alpha$  docking for dark FMN and 5dFMN AsLOV2. If this would be indeed observed, further evidence for a weaker J $\alpha$  docking for 5dFMN AsLOV2 could be provided. Whilst this would be interesting to investigate, one notable question should be answered. If indeed the secondary dark peak observed represents a population of unfolded J $\alpha$ , why its chemical shift is different to that of lit state.

Although no direct evidence has been provided in this work, but a hypothesis can be raised explaining this observation. It is possible that upon the photoadduct formation elongated N492-FMN-O4 hydrogen bond is furthermore observed by a further chemical shift in 5fW491 resonance, which is a product of J $\alpha$  unfolding and structural changes within the PAS domain core.

Noteworthy for *OdAu1a*<sub>LOV</sub>, no differences in hydrogen bonding to FMN- or 5dFMN-O4 could be identified in the dark state, it is not known if these would be the same for *AsLOV2*. To investigate this, X-ray structural studies could be undertaken. Additionally, FT-IR experiments could be performed to investigate hydrogen bonds in dark state 5dFMN containing *AsLOV2* furthermore investigating changes upon photoactivation. This would provide key data to determine if 5dFMN is an ideal structural analogue also providing evidence for the changes in hydrogen bond networks.

### 6.3 Chapter 5 Future Work

In chapters 4 and 5, the power of <sup>19</sup>F NMR has been demonstrated in the study of LOV domain proteins. Most notably in chapter 5, conventional NMR approaches failed to provide any structural information for full length *OdAu1a* construct, most likely due to large dimeric complex being formed (~70 kDa in size). Collected two-dimensional NMR spectra could only indicated the unstructured N-terminus and no evidence for LOV core could be identified. Strikingly, <sup>19</sup>F NMR in conjugation with 5-fluoro-L-tryptophan labelling provided evidence for structural changes not taking place at the C-terminus. This therefore suggest that by applying targeted <sup>19</sup>F labelling the full-length protein could be studied providing an avenue to determine low resolution structural information for *OdAu1a* with and without target DNA. So far Aureochrome1a protein construct could not be crystallised obscuring understanding of light depended structural changes. Through targeted fluorotyrosine labelling at different sections of the protein, including basic leucine zipper (bZIP), the light depend structural changes could be investigated in an attempt to pinpoint DNA binding mechanism. If fully understood, this information could then be directly applied in the development of engineered light responsive bZIP transcription factors.



## 7 Published Work – ACS Biochemistry Manuscript

Data in Chapter 3 has been published in peer reviewed journal Biochemistry, ACS. All structural data has been deposited as Protein Data Bank entries of **6I20**, **6I21**, **6I22**, **6I23**, **6I24** and **6I25**. Raw data and refinement files (CCP4i2) are available upon request, please contact Prof Rudolf K Allemann. All data has been deposited to Allemann's Group External drive (Cardiff University, Wales, UK). A copy of the published manuscript can be found at the end of this thesis with the permission from ACS journal Biochemistry. The manuscript is also available online, open access, Kalvaitis M E *et al. Biochemistry* 2019, 58, 22, 2608-2616. Alternatively, please follow the link shown:

<https://pubs.acs.org/doi/full/10.1021/acs.biochem.9b00255>

### 7.1 Permission Request

A request to attached a copy of the manuscript to this thesis has been granted by American Chemical Society (ACS), please see a copy of the e-mail below.



Dear Mindaugas E Kalvaitis,

Your permission requested is granted and there is no fee for this reuse. In your planned reuse, you must cite the ACS article as the source, add this direct link <https://pubs.acs.org/doi/full/10.1021/acs.biochem.9b00255>, and include a notice to readers that further permissions related to the material excerpted should be directed to the ACS.

If you need further assistance, please let me know.

Sincerely,

Raquel Picar-Simpson  
ACS Publications Support  
Customer Services & Information  
Website: <https://help.acs.org/>

=====

Title: A Noncanonical Chromophore Reveals Structural Rearrangements of the Light-Oxygen-Voltage Domain upon Photoactivation  
Authors: Mindaugas E. Kalvaitis Luke A. Johnson Robert J. Mart Pierre Rizkallah Rudolf K. Allemann\*  
Biochemistry 2019, 58, 22, 2608-2616  
Publication Date: May 13, 2019  
<https://doi.org/10.1021/acs.biochem.9b00255>  
Copyright © 2019 American Chemical Society



# A Noncanonical Chromophore Reveals Structural Rearrangements of the Light-Oxygen-Voltage Domain upon Photoactivation

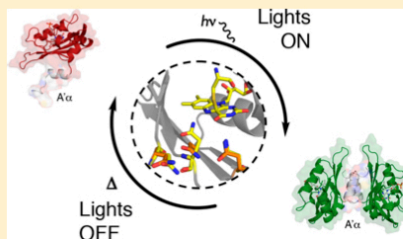
Mindaugas E. Kalvaitis,<sup>†</sup> Luke A. Johnson,<sup>†</sup> Robert J. Mart,<sup>†</sup> Pierre Rizkallah,<sup>‡</sup> and Rudolf K. Allemann<sup>\*,†</sup>

<sup>†</sup>School of Chemistry, Cardiff University, Park Place, Cardiff CF10 3AT, United Kingdom

<sup>‡</sup>School of Medicine, University Hospital Wales, Main Building, Heath Park, Cardiff CF14 4XN, United Kingdom

## Supporting Information

**ABSTRACT:** Light-oxygen-voltage (LOV) domains are increasingly used to engineer photoresponsive biological systems. While the photochemical cycle is well documented, the allosteric mechanism by which formation of a cysteinyl-flavin adduct leads to activation is unclear. Via replacement of flavin mononucleotide (FMN) with 5-deazaflavin mononucleotide (SdFMN) in the Aureochrome1a (Au1a) transcription factor from *Ochromonas danica*, a thermally stable cysteinyl-SdFMN adduct was generated. High-resolution crystal structures (<2 Å) under different illumination conditions with either FMN or SdFMN chromophores reveal three conformations of the highly conserved glutamine 293. An allosteric hydrogen bond network linking the chromophore via Gln293 to the auxiliary A'α helix is observed. With FMN, a “flip” of the Gln293 side chain occurs between dark and lit states. SdFMN cannot hydrogen bond through the C5 position and proved to be unable to support Au1a domain dimerization. Under blue light, the Gln293 side chain instead “swings” away in a conformation distal to the chromophore and not previously observed in existing LOV domain structures. Together, the multiple side chain conformations of Gln293 and functional analysis of SdFMN provide new insight into the structural requirements for LOV domain activation.



Light-oxygen-voltage (LOV) photoreceptors are members of the Per-ARNT-Sim (PAS) superfamily of proteins that act as blue-light-sensing modules, mediating a wide range of processes, including phototropism, circadian rhythms, and stress responses.<sup>1–7</sup> The modular arrangement of sensory LOV domain proteins and effectors found in nature<sup>4,5,8–10</sup> has inspired many synthetic designs.<sup>11–16</sup> Such engineered proteins exhibit varying levels of photoresponsiveness, which can be partly attributed to the incomplete understanding of the mechanisms of allosteric control employed by natural LOV domains over effector modules.<sup>12,16–19</sup> To fully exploit the photochemical potential of LOV domains for engineered systems, a comprehensive picture of the structural determinants of allostery is needed.

The structure of the LOV domain core is highly conserved, comprising a flavin chromophore binding site composed of a five-stranded, antiparallel β-sheet with ancillary helices.<sup>7,20</sup> Blue-light absorption results in the formation of reversible covalent adducts between the flavin isoalloxazine ring (C4a) and the sulfhydryl side chain of a conserved cysteine residue (Figure 1A). Flanking A'α (N-terminal) and Jα (C-terminal) helices act to relay photochemically induced changes in the LOV domain to associated effector modules.<sup>21,22</sup> Although the core LOV domain is structurally conserved, several different mechanisms of signal transduction are known. Mechanisms include Jα helix unfolding to release effector domains in *Avena*

*sativa* phototropin 1 LOV2 (AsLOV2),<sup>23,24</sup> Jα rotation and effector domain rearrangement in *Bacillus subtilis* YtvA (BsYtvA),<sup>25–27</sup> and dimerization in *Neurospora crassa* vivid (NcVVD).<sup>6,28,29</sup> The molecular basis of how such diverse results are obtained from the shared phenomenon of blue-light-driven formation of a covalent adduct between FMN and the cysteine side chain remains unclear.<sup>7,20,28,30</sup> One hypothesis suggests that protonation of N5 of the flavin cofactor, changing N5 from a hydrogen bond acceptor to a donor, causes a “flip” of the side chain of a conserved glutamine, with this change in polarity postulated to be communicated through a hydrogen bond donor/acceptor network.<sup>31–33</sup> The resolution of current crystal structures of lit-state proteins has been too low (>2.7 Å) to assert the rotamer identity with certainty.<sup>29,34</sup> Molecular dynamics (MD) simulations offer some support for the N5 protonation/glutamine flip hypothesis,<sup>35</sup> and site-directed mutagenesis of the glutamine residue confirmed its vital importance for the function of distantly related LOV domains,<sup>28,32</sup> suggesting a common underlying mechanism.

While the importance of the conserved glutamine is established, alternative hypotheses of how it governs light

Received: March 23, 2019

Revised: May 7, 2019

Published: May 13, 2019

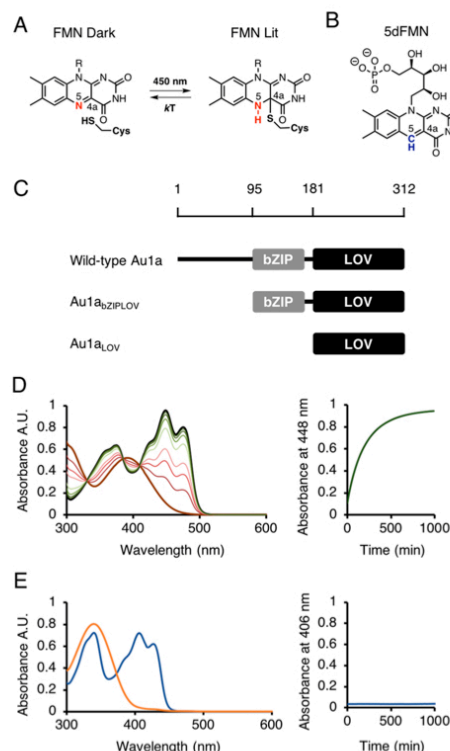


ACS Publications

© 2019 American Chemical Society

2608

DOI: 10.1021/acs.biochem.9b00255  
Biochemistry 2019, 58, 2608–2616



**Figure 1.** (A) Formation of a cysteinyl-FMN covalent adduct occurs upon absorption of blue light by flavin mononucleotide (FMN). Spontaneous thermal reversion re-forms the dark-adapted state. (B) Structure of 5-deazaflavin mononucleotide (5dFMN) with a carbon atom (blue) at position 5. (C) Domain topology of *O. danica* Aureochrome1a. Au1a<sub>bZIPLOV</sub> comprises bZIP and LOV domains, and Au1a<sub>LOV</sub> comprises only the LOV domain. UV-vis spectra of thermal reversion from the lit to dark state of (D) FMN-containing (red-green) and (E) SdFMN-containing (orange-blue) *OdAu1a<sub>LOV</sub>*. Spectra were recorded every hour for the first 3 h and then every 2 h. Reversion kinetics were monitored at 448 nm for FMN-containing *OdAu1a<sub>LOV</sub>* and 406 nm for SdFMN-containing *OdAu1a<sub>LOV</sub>*. Lit-state FMN *OdAu1a<sub>LOV</sub>* reverts to its dark state with a half-life of 112 min. No reversion to the dark state is observed for lit-state SdFMN-containing *OdAu1a<sub>LOV</sub>*.

switching have been proposed. MD simulations of phototropin LOV domains generated a different conformation for the conserved glutamine side chain, altering the hydrogen bonding network to flanking helices.<sup>35–37</sup> Other recent reports propose that further glutamine side chain orientations are involved in LOV domain activation through hydrogen bonds with O4 of the flavin ring.<sup>37–40</sup> Given the importance of the potential hydrogen bonding associated with N5 of the flavin and the challenges associated with studying the lit state of thermally reverting LOV domains, we used 5-deazaflavin mononucleotide [5dFMN (Figure 1B)], an analogue that had previously been suggested to form a stable photochemical cysteinyl-flavin adduct in BsYtvA<sup>41</sup> and successfully employed to alter the redox potentials of other flavoproteins.<sup>42–44</sup> At present, there

are no experimental data to indicate whether the lit states of SdFMN-containing LOV photoreceptors function like FMN-containing examples. We therefore decided to examine the effect of SdFMN incorporation on the photochemistry and function of Aureochrome1a (Au1a) of *Ochromonas danica*.

Aureochromes comprise a family of LOV domain-containing transcription factors found in photosynthetic stramenophiles that regulate the cell cycle and photomorphogenesis.<sup>3,45,46</sup> Au1a consists of an N-terminal unstructured region, followed by a basic leucine zipper (bZIP) domain and a C-terminal LOV domain. This domain topology is inverted compared to those of most other photoreceptors and means that the A'α helix, instead of the C-terminal Jα helix, connects effector and LOV domains. Spectroscopic and biochemical measurements of the isolated LOV domain from *Phaeodactylum tricornutum* and *Vaucheria frigida* Au1a suggest that stepwise unfolding of A'α and Jα helices upon illumination results in LOV domain dimerization.<sup>22,47</sup> Single-crystal X-ray structures of light-grown LOV domain crystals at 2.7 Å suggested the availability of the core β-sheet for use as a dimerization interface.<sup>34</sup> In dark-state structures, this dimerization site is obscured by the A'α helix. Full-length Au1a has resisted crystallization, but small-angle X-ray scattering (SAXS) of constructs whose unstructured region has been truncated shows significant volume changes that suggest intramolecular bZIP-LOV interactions.<sup>34,48</sup> Steric caging of the bZIP domain may therefore complement LOV domain-driven dimerization, which is proposed to be the driving force behind Aureochrome DNA binding.<sup>48</sup> Here, we present functional analysis and the first high-resolution crystal structures of a LOV domain with SdFMN, identifying three conformations for Gln293 of Au1a and the allosteric network linking the chromophore to the A'α helix. This glutamine is widely conserved among LOV domains, and as there are several examples in which truncations of the A'α helix directly influence effectors connected through the Jα helix, these results may have wider implications beyond the Au1a family.

## MATERIALS AND METHODS

**Protein Expression and Purification.** Standard molecular biology techniques were employed to generate *OdAu1a<sub>LOV</sub>* and *OdAu1a<sub>bZIPLOV</sub>* constructs from the wild-type *O. danica* Au1a gene (UniProt, C5NSW6\_OCHDN) using oligonucleotides detailed in Table S1. *OdAu1a*-derived proteins were obtained by heterologous expression in BL21 (DE3) *Escherichia coli* in either minimal and autoinduction medium supplemented with glucose [1% (w/v)] and kanamycin (50–100 μg/mL). Cultures were grown at 37 °C until an OD<sub>600</sub> of 0.8 was reached, induced with isopentenyl thiogalactose (0.5 mM, IPTG, Melford), and grown at 25 °C for a further 16 h. *OdAu1a* proteins were purified by Ni<sup>2+</sup>-NTA (5 mL, GE Healthcare) affinity chromatography followed by Resource Q anion exchange (GE Healthcare) chromatography using purification buffer: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES, 20 mM, pH 7.8), sodium chloride [20 mM (*OdAu1a<sub>LOV</sub>*) or 150 mM (*OdAu1a<sub>bZIPLOV</sub>*)], tris(carboxyethyl) phosphine (TCEP, 0.3 mM), and gradients of imidazole (from 20 to 500 mM) and sodium chloride (from 0 to 1 M). Chromophore exchange was performed by applying the clarified cell lysate to Ni<sup>2+</sup>-NTA resin (5 mL, GE Healthcare) and washing with 5 column volumes of purification buffer. Proteins were partially unfolded by passing this buffer supplemented with guanidine hydrochloride (6 M) over the resin. To complete FMN elution, a guanidinium thiocyanate



solution (3 M) was applied until no flavin was observed in the eluent by ultraviolet–visible (UV–vis) spectroscopy. Proteins were refolded by sequentially applying lower concentrations of guanidine hydrochloride (one column volumes of concentrations of 6, 5, 4, 3, 2, and 0 M). The resin was then washed with 5 column volumes of purification buffer, followed by incubation with 1 column volume of purification buffer containing SdFMN (0.1–0.5 mM) for 30 min. Protein samples were eluted and then further purified as previously described.

**Solution Characterization.** For circular dichroism (CD) experiments, purified protein samples were dialyzed overnight at 4 °C against potassium phosphate buffer (10 mM, pH 7.0). Spectra were collected with an Applied Photophysics Chirascan spectrophotometer. For analytical gel filtration experiments, protein samples were exchanged into gel filtration buffer [HEPES (20 mM, pH 7.4), sodium chloride (100 mM), magnesium chloride (10 mM), and TCEP (0.3 mM)] in centrifugal filter columns. All protein samples were handled in dim red light. For photoactivation, protein samples were illuminated with 450 nm light-emitting diodes (LEDs) until a steady state was reached as determined by UV–vis spectroscopy. For gel filtration experiments, analytical gel filtration columns were either wrapped in aluminum foil for dark experiments or illuminated with 450 nm LEDs for lit-state experiments. For nuclear magnetic resonance (NMR) studies, purified and uniformly  $^{15}\text{N}$ -labeled protein samples (400–600  $\mu\text{M}$ ) were exchanged into NMR buffer [2-(*N*-morpholino)-ethanesulfonic acid (MES, 20 mM, pH 6.0), ethylenediaminetetraacetate (EDTA, 1 mM), TCEP (1 mM), and sodium azide (0.05%)] and supplemented with 10% deuterium oxide. For dark-state experiments, protein samples were transferred into amber-colored NMR tubes. For lit-state experiments, protein samples were transferred into clear NMR tubes and illuminated with 450 nm LEDs. NMR spectra were recorded on a DPX-600 MHz Bruker NMR spectrometer equipped with a cryoprobe and preamplifiers.

**DNA Binding.** Light-dependent DNA binding was characterized by electrophoretic mobility gel shift assays (EMSAs) using TAMRA-labeled double-stranded DNA containing an *OdAula* recognition site (5'-TGTAGCGTC-TGACGTGGTTCAC-3'). EMSA experiments were performed at 4 °C. Dark-state experiments were performed in a room illuminated by dim red light, while for lit-state experiments, protein samples were illuminated for 5 min prior to commencing the experiment and throughout the electrophoresis experiment with 450 nm LEDs. Gels were imaged using a Bio-Rad ChemiDoc MP system (Bio-Rad Laboratories) and software provided by the manufacturer.

**Crystallography.** Purified FMN- or SdFMN-containing *OdAula*<sub>LOV</sub> was exchanged into crystallization buffer [2-(*N*-morpholino)ethanesulfonate sodium salt (MES, 50 mM, pH 6.0), sodium chloride (100 mM), magnesium chloride (20 mM), sodium acetate (20 mM), dithiothreitol (DTT, 5 mM), and EDTA (5 mM)] and concentrated to 10–15 mg/mL. Dark-state crystals were grown in plates wrapped with aluminum foil by the hanging drop method. Drops consisted of protein (2  $\mu\text{L}$ , 10 mg/mL) mixed with a reservoir solution [2  $\mu\text{L}$ , 10–20% (w/v) polyethylene glycol (PEG) with an average molecular weight of 2000 or 3000, ammonium chloride (0.1 M), and sodium acetate (0.1 M, pH 4.5–4.9) or disodium citrate, (0.1 M, pH 4.5–4.9)] suspended over further reservoir buffer (100  $\mu\text{L}$ ) in 96-well plates (Screw Top Hanging Drop Plate, Molecular Dimensions). Crystal growth

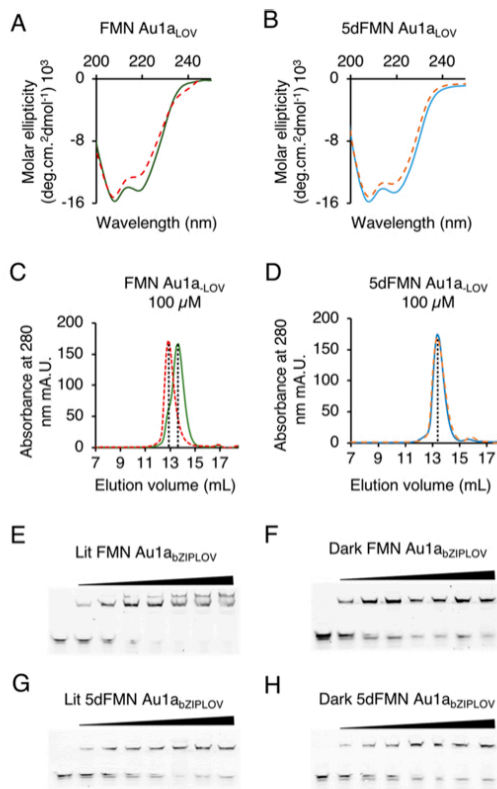
was usually evident after 16 h with maximum growth observed after 7 days. For dark-state structures, crystals were cryoprotected with ethylene glycol, harvested, and flash-frozen in liquid nitrogen under dim red light. For illumination experiments, crystals were illuminated with 450 nm LEDs for 30 min, cryoprotected with ethylene glycol, harvested, and flash-frozen in liquid nitrogen. Light-grown crystals were obtained by mixing light-state FMN- or SdFMN-containing *OdAula*<sub>LOV</sub> (1.7  $\mu\text{L}$  of a 15 mg/mL solution) with a reservoir solution [2  $\mu\text{L}$ , disodium malonate (1.5–3 M, pH 7.0) and TRIS acetate (0.1 M, pH 7.5–8.0)]. Drops were supplemented with hexamine cobalt(III) (0.3  $\mu\text{L}$ , 0.1 M). Crystals were grown under blue light and appeared after 1–7 days. Crystals were harvested without cryoprotection and flash-frozen in liquid nitrogen. Data sets were collected from a single crystal each at the Diamond Light Source synchrotron at beamlines I02, I03, and I24. Initial structures were obtained from Phaser<sup>55</sup> or MolRep<sup>55,56</sup> using the dark-state *Phaeodactylum tricornutum* *Aula*<sub>LOV</sub> domain (Protein Data Bank entry 5A8B) as a search model. Structures were determined by subjecting initial models to cycles of model building with COOT<sup>57</sup> and refinement using REFMAC5.<sup>55</sup> For final  $R_{\text{work}}$  and  $R_{\text{free}}$  values, see Tables S4 and S5.

## RESULTS

***OdAula* with SdFMN Incorporated Forms a Thermally Stable Cysteiny-Flavin Adduct.** To investigate the mechanism of LOV domain activation and the effects of introducing SdFMN (Figure 1B), two truncated versions of *O. danica* *Aula* containing the isolated LOV domain (*OdAula*<sub>LOV</sub>) and the LOV domain with the DNA binding bZIP domain (*OdAula*<sub>bZIPLOV</sub>) were constructed. The FMN cofactor of the expressed proteins was replaced with SdFMN by binding the protein to Ni-NTA resin and washing with guanidine thiocyanate followed by removal of the denaturant and incubation with SdFMN. Refolding of *OdAula*<sub>LOV</sub> and *OdAula*<sub>bZIPLOV</sub> in the presence of SdFMN produced the characteristic blue-shifted vibrational triplet of oxidized SdFMN with absorbance maxima at 385, 406, and 423 nm (Figure 1E and Figure S1C–F).<sup>41</sup> Comparing the absorbance at 406 and 475 nm indicated that >99% of the cofactor had been exchanged. *OdAula*<sub>LOV</sub> containing FMN reverted from its lit state to its dark state, with a half-life of 112 min, but no reversion was observed for *OdAula*<sub>LOV</sub> containing SdFMN even after 7 days (Figure S1C,D). Cycling between lit and dark states using 450 and 330 nm light was possible with no significant photobleaching for at least five cycles (Figure S1E,F). The stability of the SdFMN adduct was further demonstrated in liquid chromatography–mass spectrometry experiments in which species corresponding to the stable covalent cysteiny-SdFMN conjugate for *OdAula*<sub>LOV</sub> were observed but no cysteiny-FMN conjugates were evident (Figures S2 and S3).

**FMN to SdFMN Exchange Prevents Light-Induced Dimerization of *OdAula*<sub>LOV</sub>.**  $^1\text{H}$ – $^{15}\text{N}$  heteronuclear single-quantum coherence NMR spectra of  $^{15}\text{N}$ -labeled proteins confirmed that refolding with SdFMN did not lead to any large-scale structural perturbation. Illumination of *OdAula*<sub>LOV</sub> bound to SdFMN resulted in chemical shift perturbations similar to those observed for FMN (Figure S4). CD spectra also indicated that the secondary structure after refolding of SdFMN-containing *OdAula*<sub>LOV</sub> was the same as that of native *OdAula*<sub>LOV</sub> (Figure S5). Both FMN- and SdFMN-containing

*OdAul<sub>1aLOV</sub>* samples exhibited changes in their CD spectra when photoactivated. FMN-containing *OdAul<sub>1aLOV</sub>* displayed a  $14.2 \pm 0.8\%$  decrease in mean residue ellipticity at its 220 nm minimum (Figure 2A and Figure S5C), whereas SdFMN-



**Figure 2.** Circular dichroism spectra of (A) FMN-containing and (B) SdFMN-containing *OdAul<sub>1aLOV</sub>* (20 μM) in potassium phosphate buffer (10 mM, pH 7.0) under dark and light (450 nm) conditions. Green and red traces correspond to dark and lit states of FMN, respectively, while blue and orange traces correspond to dark and lit states of SdFMN, respectively. Size-exclusion chromatography of *OdAul<sub>1aLOV</sub>*. (C) FMN-containing *OdAul<sub>1aLOV</sub>* under dark (green) and illuminated (red dashed) conditions. (D) SdFMN-containing *OdAul<sub>1aLOV</sub>* under dark (blue) and illuminated (orange dashed) conditions. Electrophoresis mobility shift assays of *OdAul<sub>1aLOV</sub>* with a DNA target (40 nM) for (E) illuminated FMN-containing *OdAul<sub>1aLOV</sub>*, (F) dark-state FMN-containing *OdAul<sub>1aLOV</sub>*, (G) illuminated SdFMN-containing *OdAul<sub>1aLOV</sub>*, and (H) dark-state SdFMN-containing *OdAul<sub>1aLOV</sub>*. The first lane contains TAMRA-labeled DNA only, and subsequent lanes have increasing protein concentrations (from 0.4 to 12 μM from left to right, respectively).

containing *OdAul<sub>1aLOV</sub>* exhibited a smaller change of  $8.5 \pm 1.0\%$  at 220 nm (Figure 2B and Figure S5D). While UV-vis spectroscopy and mass spectrometry suggested complete and stable adduct formation for SdFMN-containing *OdAul<sub>1aLOV</sub>*, the secondary structural changes inferred an intermediate state. Size-exclusion chromatography was used to establish whether

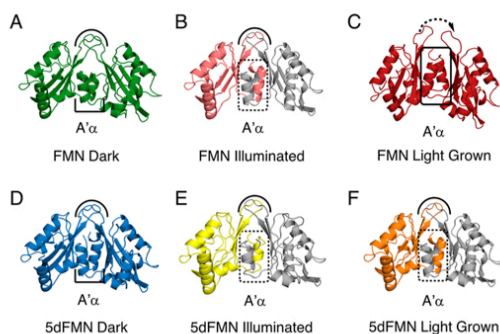
SdFMN retained the capacity to induce *OdAul<sub>1aLOV</sub>* dimerization in response to light that has been shown for other isolated Aureochrome1a LOV domains.<sup>21,22,49,50</sup> FMN-containing *OdAul<sub>1aLOV</sub>* (100 μM) showed clear light-dependent dimerization as determined by size-exclusion chromatography. Dark-state FMN-containing *OdAul<sub>1aLOV</sub>* eluted at 13.7 mL with an estimated mass of 20.4 kDa with a slight shoulder toward a larger volume, whereas the lit-state equivalent elution maxima shifted to 12.8 mL in agreement with dimerization (Figure 2C). However, 100 μM SdFMN-containing *OdAul<sub>1aLOV</sub>* showed very little shift of the elution volume with peaks at 13.5 and 13.3 mL for the dark and lit states, respectively (Figure 2D). Such an intermediate elution volume most likely represented a monomer–dimer exchange on the time scale of the size-exclusion experiments.

To probe this observation further, we conducted concentration-dependent experiments. The position of the lit-state SdFMN-containing *OdAul<sub>1aLOV</sub>* elution peak was strongly concentration-dependent (Figure S6), showing earlier elution at higher concentrations. In an identical concentration range, equivalent FMN-containing samples appeared to elute uniformly at volumes consistent with a dimer [200 to 10 μM (Figure S6)]. Although SdFMN was unable to effect efficient dimerization in *OdAul<sub>1aLOV</sub>*, the longer *OdAul<sub>1aZIPLOV</sub>* construct consistently eluted from the size-exclusion column at volumes corresponding to a dimer with both cofactors in the dark and lit states (Figures S7–S9). Strong DNA binding was observed in both states regardless of the cofactor used for 50 μM protein samples. To further probe the light responsiveness of SdFMN, lower concentrations were utilized in DNA binding assays.

Electrophoretic mobility shift assays were used to examine DNA binding by *OdAul<sub>1aZIPLOV</sub>* containing FMN or SdFMN in the dark and under illuminated conditions (Figure 2E–H). Lit-state FMN-containing *OdAul<sub>1aZIPLOV</sub>* uniquely showed a slowly migrating “supershifted” band (Figure 2E),<sup>46</sup> while experiments with dark-state FMN (Figure 2F) and both dark and illuminated SdFMN (Figure 2G,H) showed evidence of only a single slower-migrating shifted band. Having demonstrated by size-exclusion chromatography that FMN-containing *OdAul<sub>1aLOV</sub>* supports dimerization at 10 μM only when illuminated, the lower band may represent a 1:1 protein–DNA complex while the “supershifted” band most likely corresponds to the functional 2:1 complex of *OdAul<sub>1aZIPLOV</sub>*. The putative 2:1 complex was not observed when SdFMN replaced FMN. Overall, the structural and functional experiments in solution suggest that SdFMN incorporation creates a protein that can form a stable cysteinyl–SdFMN adduct that shows some structural features of a lit-state FMN-containing protein but with incomplete control over the longer-range interactions that direct DNA binding and dimer stability.

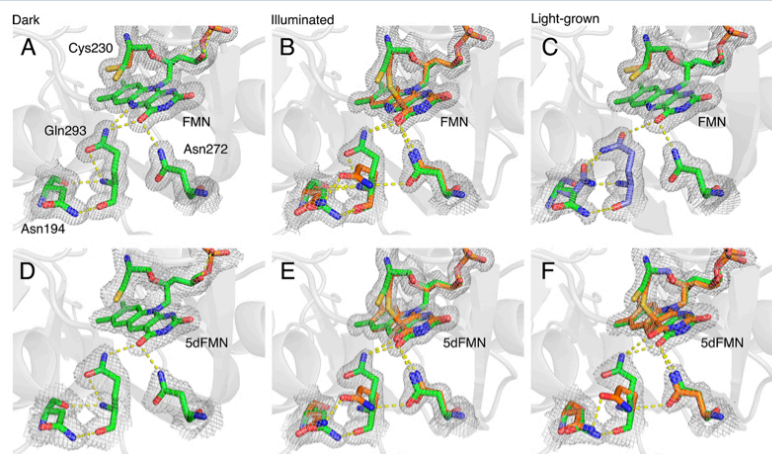
**5dFMN Forms Cysteinyl-Flavin Adducts at the C4α Position but Induces No Rearrangement of the A′α Helix.** To understand how SdFMN can mimic FMN photochemistry but is incapable of complete *OdAul<sub>1aLOV</sub>* activation, high-resolution single-crystal X-ray structures for FMN- and SdFMN-containing *OdAul<sub>1aLOV</sub>* were determined for crystals grown in the dark, in the dark and then illuminated with blue light (“illuminated”), and under steady strong blue-light exposure (“light-grown”). The highest-resolution structure of dark-state FMN-containing *OdAul<sub>1aLOV</sub>* was obtained at 1.37 Å from a single crystal in space group *P*2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> with four monomers per asymmetric unit as parallel dimers (Figure 3A).



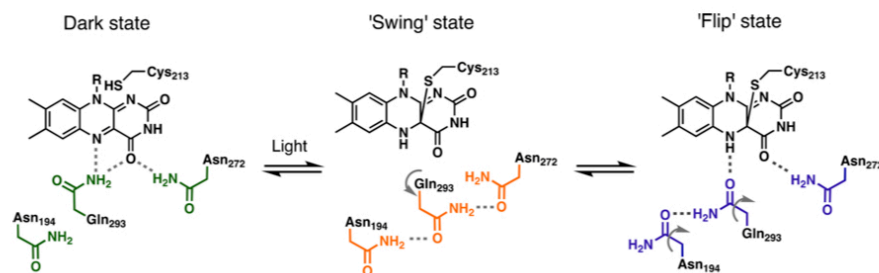


**Figure 3.** Dimer arrangements for X-ray crystal structures of FMN- and SdFMN-containing *OdAul1<sub>LOV</sub>* under dark (left), illuminated (middle), and light-grown (right) conditions. (A) The 1.37 Å structure of dark-state FMN-containing *OdAul1*. The asymmetric unit contained four monomers as parallel dimers (green) with A'α positioned across the β-sheet surface (half black box). Loops of each monomer lie close to each other (black curved line). (B) The 1.50 Å structure of illuminated crystals of dark-grown FMN-containing *OdAul1*. The asymmetric unit contained a single monomer (pink), forming a parallel dimer similar to that in the dark state when considering a symmetry equivalent (gray). (C) The 1.66 Å structure of light-grown FMN-containing *OdAul1* featuring a unique dimer arrangement with A'α being repositioned across a β-sheet surface (black box) and loop region rearrangement (dashed arrow). (D) The 1.97 Å structure of dark-state SdFMN-containing *OdAul1* with a dimer similar to dark-state FMN. (E) The 1.43 Å structure of illuminated SdFMN-containing *OdAul1* with a symmetry partner equivalent to a dimer colored gray. (F) The 2.00 Å structure of light-grown SdFMN-containing *OdAul1* showing a similar loop (black curved line) and A'α helix arrangement (black box) as for dark-state and illuminated proteins with a symmetry equivalent colored gray.

Crystals in space group *P*3<sub>1</sub>21 were also observed, but these diffracted poorly. A 1.97 Å structure of dark-state SdFMN-containing *OdAul1<sub>LOV</sub>* was obtained from crystals grown under identical conditions in space group *P*3<sub>1</sub>21 indicating a parallel dimer per asymmetric unit (Figure 3D). The identity of the cofactor had little effect on the overall LOV domain structure or the chromophore binding pocket (Figure 4A,D), confirming an identical mode of chromophore binding and no rearrangement of the surrounding environment. When dark-grown crystals were illuminated, the space group changed to *P*6<sub>2</sub>22 with a single monomer per asymmetric unit (Figure 3B,E). However, once symmetry partners were considered, symmetrical dimers almost identical to the dark-adapted state could be identified with symmetry equivalents. Inspection of the cofactor binding site of the illuminated crystals showed electron density for approximately 30% occupancy of a covalent bond between Cys230 and the cofactor for both FMN and SdFMN structures (Figure 4B,E). This occupancy that is significantly lower than indicated by UV-vis spectroscopy and MS (Figure 1D,E, and S10) is likely to be the result of a photochemical scission of the covalent adduct during data collection as reported for other LOV domain proteins.<sup>29</sup> Although the usual approach under such circumstances is to record multiple data sets from a single crystal, this usually yields much poorer resolution and was therefore not attempted. We hypothesized that higher-resolution data sets could provide unique insights into structural change. To ensure minimal bias in the cycles of structural refinement, we modeled covalent adduct structure at 30% occupancy (cysteinyl-flavin photoadduct) and the dark state at 70% occupancy, yielding two flavin and cysteine orientations. Electron density for a partial occupancy of a cysteinyl-flavin adduct at the C4a position of the isoalloxazine ring for SdFMN-containing *OdAul1<sub>LOV</sub>* was clearly observed, confirming that SdFMN forms a photochemical adduct structurally equivalent to the native chromophore.



**Figure 4.** Electron density maps (gray mesh) for FMN or SdFMN, Cys230, and residues forming hydrogen bonding networks among O4 of FMN, Asn272, Gln293, and Asn194 are displayed at the  $\sigma = 1$  level. Partial occupancies are colored by characteristic conformations observed for dark-state (green), illuminated (orange), or light-grown (purple) structures. Yellow dashed lines indicate predicted hydrogen bonding. (A) Dark-state FMN (1.36 Å). (B) Illuminated FMN (1.50 Å). (C) Light-grown FMN (1.67 Å). (D) Dark-state SdFMN (1.97 Å). (E) Illuminated SdFMN (1.43 Å). (F) Light-grown SdFMN (2.00 Å).



**Figure 5.** In the dark-state conformation (green), conserved Gln293 hydrogen bonds to O4 and N5 of FMN. Illumination with blue light results in the Gln293 “swing” state (orange) where its side chain rotates away from the FMN chromophore. Progression to the Gln293 “flip” state (purple) may occur from the “swing” state or from the dark state but cannot proceed when the protein is trapped in the crystal lattice. Rotation of the side chain of Asn194 upon formation of the “flip” state is likely to lead to conformational changes in the A’ $\alpha$  helix.

Light-grown crystals could not be obtained under the conditions used for the dark state, but alternative conditions produced monoclinic crystals in space group C121 for the FMN sample and hexagonal crystals in space group P6<sub>3</sub>22 for SdFMN. For light-grown FMN-containing *OdAul<sub>1</sub>LOV*, no cysteinyl-FMN adduct was evident in the electron density map. Electron density corresponding to a cysteinyl-SdFMN adduct was observed but was less prominent than in maps from illuminated crystals. Notably, light-grown FMN-containing *OdAul<sub>1</sub>LOV* contained four monomers per asymmetric unit as parallel dimers. Two of the monomers appeared to be identical to the dark-state structure, while the second pair showed a different A’ $\alpha$  arrangement across the  $\beta$ -sheet surface. Compared with the dark-state structures, a change in the relative positions of strand I $\beta$  (287–293) of the  $\beta$ -sheet and A’ $\alpha$  (183–189) of 12° is observed. This rearrangement does not occur in illuminated structures, probably due to crystal lattice constraints. Light-grown SdFMN-containing *OdAul<sub>1</sub>LOV* maintained a single monomer per asymmetric unit, resembling the arrangement of illuminated structures. This supported solution data that although SdFMN undergoes photochemistry similar to that of FMN, it is unable to fully activate *OdAul<sub>1</sub>LOV*. Taken together, this suggests that the A’ $\alpha$  rearrangement, observed for only light-grown FMN-containing *OdAul<sub>1</sub>LOV*, could correlate with dimerization in solution (Figure 3A,C).

**Adduct Formation Populates Different Conformations of Gln293.** Dark-state structures gave single well-defined populations of Gln293, Asn272, and Asn194 for both FMN and SdFMN, but close examination of electron density maps from illuminated and light-grown conditions yielded multiple conformations for these residues (Figure 4). In dark-state structures, Gln293 lies close to the chromophore and forms a probable hydrogen bond to the O4 position (Figure 4A,D). For illuminated structures, a 20% occupancy of a new conformation of Gln293, with its side chain away from the FMN binding pocket, was evident (Figure S22). Additional conformations of Asn194 and Asn272 are also observed. Formation of a new hydrogen bond network among these three residues creates a route for the conformation of Gln293 to be communicated to the A’ $\alpha$  helix through Asn194 (Figure 4B,E), which is located in the loop connecting A’ $\alpha$  with the LOV domain core. Examination of the FMN binding pocket of the light-grown crystal structure revealed a third arrangement for Gln293 and Asn194. In the parallel dimer with a unique A’ $\alpha$  arrangement, Asn194 moves in toward Gln293. This

coincides with a probable change in the orientation of the Gln293 side chain and the polarity of the hydrogen bond network due to flavin protonation (Figure 4C). In contrast, SdFMN-containing *OdAul<sub>1</sub>LOV* did not form this “flip” conformation but closely resembled the structure of illuminated SdFMN-containing *OdAul<sub>1</sub>LOV* with identical Asn194 and Gln293 conformations (Figure 4F). This suggests that the structural changes that we observed were not an artifact of different crystallization conditions, as identical space groups were achieved for SdFMN under both conditions. It appears that only growing crystals of FMN-containing protein under constant illumination allow the structural reorientation of the allosteric A’ $\alpha$  helix to support dimer rearrangement (Figure 3).

## DISCUSSION

Reconstitution of truncated versions of the light-dependent transcription factor *OdAul<sub>1</sub>LOV* with SdFMN led to proteins that undergo photoadduct formation to produce a thermally stable cysteinyl-SdFMN adduct. Despite clear evidence of adduct formation captured by UV spectroscopy and mass spectrometry and subsequent light-induced changes determined by CD and NMR spectroscopy, SdFMN-containing *OdAul<sub>1</sub>LOV* does not dimerize under the conditions where dimerization of FMN *OdAul<sub>1</sub>LOV* occurs. High-resolution crystal structures show identical FMN and SdFMN binding modes in dark-grown crystals, suggestive of identical chromophore binding characteristics. Crystal structures of illuminated and light-grown *OdAul<sub>1</sub>LOV* containing SdFMN provide conclusive evidence of cysteinyl photoadduct formation at the C4a position in apparent support of the radical-based mechanism proposed for the native chromophore.<sup>41</sup>

Crystal structures obtained under different illumination conditions define three distinct conformations for conserved Gln293 and its hydrogen bonding partners. For the FMN-containing light-grown crystal structure, inspection of calculated difference maps with both rotamers strongly implies a glutamine “flip” (Figure S22). Protonation of FMN and a corresponding “flip” of the glutamine side chain is a leading hypothesis for LOV activation<sup>29,32,33</sup> and has even been used to explain the activation of a LOV domain containing a neutral semiquinone flavin.<sup>51</sup> However, in our and other published light-grown structures,<sup>29,34</sup> the distance between Gln293 and N5 of FMN is longer than might be expected for a hydrogen bond (~3.3 Å). SdFMN forms the equivalent covalent adduct, but in contrast to FMN, Gln293 does not appear to “flip” in



light-grown crystals (Figure 4C,F). Taken with SdFMN's inability to induce light-dependent dimerization, this furthermore suggests that N5 protonation is a prerequisite for locking the lit-state conformation of *OdAula*<sub>LOV</sub>. Comparison of illuminated and light-grown FMN-containing *OdAula*<sub>LOV</sub> structures (Figure 4B,C) suggests that a key aspect of the Gln293 "flip" is to engage an alternative conformation of the side chain of Asn194, located between the A $\beta$  strand and A' $\alpha$  helix, in hydrogen bonding. This change in the Asn194 conformation may be the key to propagating the effects of adduct formation beyond the LOV domain, by rearranging the domain to favor dimerization and/or by releasing the A' $\alpha$  helix. Notably, this key Asn residue is also found in other *Aula* homologues<sup>34,49,50</sup> and in *AsLOV2*.<sup>52</sup>

To the best of our knowledge, our structures of illuminated crystals of SdFMN-containing *OdAula*<sub>LOV</sub> provide the first experimental evidence for a further arrangement of Gln293 and Asn194. For both FMN and SdFMN, Gln293 "swings" away from the chromophore when illuminated (Figure 5). The persistence of a "swing" conformation in SdFMN light-grown crystals suggests that it is not a crystallographic artifact generated by illuminating LOV domains trapped in the crystal lattice or a product of a different space group due to changes in crystallographic conditions but that SdFMN is unable to support progress to the "flip" state. CD measurements indicate that illumination of SdFMN-containing *OdAula*<sub>LOV</sub> results in some secondary structural changes, although not to the extent seen with FMN. Likewise, size-exclusion chromatography revealed intermediate changes upon illumination for SdFMN. These results infer the "swing" conformation of Gln293, formed in the absence of larger-scale secondary structural changes, provides a degree of activation. It therefore seems probable that the "swing" state represents an intermediate stage in activation rather than an unproductive conformation (Figure 5).

The key determinants of success for LOV domain-based optogenetic systems are the dynamic ranges of affinities and activities in the dark and lit states. Most previous work has relied on using molecular modeling to guide alterations to helix docking propensities to improve the dynamic range of optogenetic tools. Here we provide experimental insight into the molecular basis of LOV domain photoactivation. The nature of the "swing" state may be crucial for improving LOV domains by rational design; if the "swing" state is an on-path intermediate, then encouraging its formation is of key importance, placing a greater emphasis on the O4–Gln293–Asn194 axis for initial activation. This is in agreement with MD simulations and Fourier transform infrared spectroscopy experiments that identified hydrogen bond changes to O4 of the FMN ring being important for the regulation of LOV domain activation especially in the early stages after photo-adduct formation.<sup>35–37,39,40</sup> Indeed, results that inferred a role for N5 protonation by generation of a neutral flavin semiquinone radical<sup>33,51</sup> and N5-protonated reduced flavins<sup>53</sup> also predict significant polarity changes at O4.<sup>54</sup> Alternatively, if the "swing" state is an off-path intermediate, its destabilization may lead to improved switches with higher dynamic ranges. Understanding these fundamental aspects of domain activation has a potentially enormous impact for the design of new tools based on LOV domains and may allow researchers to improve the performance of designed LOV domains with multiple optogenetic applications.

## ■ ASSOCIATED CONTENT

### ■ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.9b00255.

Experimental procedures, including purification, characterization, and crystallization methods; UV–vis spectroscopy, mass spectrometry, NMR, CD, and SEC data of *OdAula* (Figures S1–S10, S22) and 5-deazaflavin mononucleotide chemoenzymatic synthesis (Figures S11–S21); and tables of DNA oligonucleotides, protein sequences, and structural refinement statistics (Tables S4 and S5) (PDF)

### Accession Codes

Protein Data Bank entries 6I20, 6I21, 6I22, 6I23, 6I24, and 6I25.

## ■ AUTHOR INFORMATION

### Corresponding Author

\*E-mail: [allemannrk@cardiff.ac.uk](mailto:allemannrk@cardiff.ac.uk).

### ORCID

Mindaugas E. Kalvaitis: 0000-0002-8175-7552

Luke A. Johnson: 0000-0002-6697-6589

Robert J. Mart: 0000-0003-2196-5840

Pierre Rizkallah: 0000-0002-9290-0369

Rudolf K. Allemann: 0000-0002-1323-8830

### Funding

This work was supported by BBSRC Grants BB/M006158/1 and BB/P009980/1 and EPSRC Grant EP/L027240/1.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

The authors thank Dr. Harald Janovjak (IST Austria) for providing the DNA encoding full-length *O. danica* Aureochromela, Dr. Yi Jin for help with structure determination and helpful discussions, and Tom Williams for assistance with mass spectrometry. The authors thank the Diamond Light Source for beamtime (Proposal mx14843) and the staff of beamlines I02, I03, and I24 for assistance with data collection.

## ■ ABBREVIATIONS

*OdAula*, *O. danica* Aureochromela; LOV, light-oxygen-voltage; bZIP, basic leucine zipper; FMN, flavin mononucleotide; PAS, Per-ARNT-Sim; SdFMN, 5-deazaflavin mononucleotide; CD, circular dichroism; SAXS, small-angle X-ray scattering; MD, molecular dynamics; NMR, nuclear magnetic resonance.

## ■ REFERENCES

- (1) Crosson, S., Rajagopal, S., and Moffat, K. (2003) The LOV Domain Family: Photoresponsive Signaling Modules Coupled to Diverse Output Domains. *Biochemistry* 42, 2–10.
- (2) Ito, S., Song, Y. H., and Imaizumi, T. (2012) LOV Domain-Containing F-Box Proteins: Light-Dependent Protein Degradation Modules in Arabidopsis. *Mol. Plant* 5, 573–582.
- (3) Krauss, U., Minh, B. Q., Losi, A., Gärtner, W., Eggert, T., Von Haeseler, A., and Jaeger, K. E. (2009) Distribution and Phylogeny of Light-Oxygen-Voltage-Blue-Light-Signaling Proteins in the Three Kingdoms of Life. *J. Bacteriol.* 191, 7234–7242.



- (4) Glantz, S. T., Carpenter, E. T., Melkonian, M., Gardner, K. H., Boyden, E. S., Wong, G. K., and Chow, B. Y. (2016) Functional and Topological Diversity of LOV Domain Photoreceptors. *Proc. Natl. Acad. Sci. U. S. A.* 113, E1442–E1451.
- (5) Bonomi, H. R., Posadas, D. M., Paris, G., Carrica, M. D. C., Frederickson, M., Pietrasanta, L. I., Bogomolni, R. A., Zorreguieta, A., and Goldbaum, F. A. (2012) Light Regulates Attachment, Exopolysaccharide Production, and Nodulation in *Rhizobium Leguminosarum* through a LOV-Histidine Kinase Photoreceptor. *Proc. Natl. Acad. Sci. U. S. A.* 109, 12135–12140.
- (6) Chen, C. H., DeMay, B. S., Gladfelter, A. S., Dunlap, J. C., and Loros, J. J. (2010) Physical Interaction between VIVID and White Collar Complex Regulates Photoadaptation in *Neurospora*. *Proc. Natl. Acad. Sci. U. S. A.* 107, 16715–16720.
- (7) Möglich, A., Ayers, R. A., and Moffat, K. (2009) Structure and Signaling Mechanism of Per-ARNT-Sim Domains. *Structure* 17, 1282–1294.
- (8) Cao, Z., Buttani, V., Losi, A., and Gärtner, W. (2008) A Blue Light Inducible Two-Component Signal Transduction System in the Plant Pathogen *Pseudomonas Syringae* Pv. Tomato. *Biophys. J.* 94, 897–905.
- (9) Purcell, E. B., Siegal-Gaskins, D., Rawling, D. C., Fiebig, A., and Crosson, S. (2007) A Photosensory Two-Component System Regulates Bacterial Cell Attachment. *Proc. Natl. Acad. Sci. U. S. A.* 104, 18241–18246.
- (10) Herron, J., and Crosson, S. (2011) Function, Structure and Mechanism of Bacterial Photosensory LOV Proteins. *Nat. Rev. Microbiol.* 9, 713–723.
- (11) Christie, J. M., Gawthorne, J., Young, G., Fraser, N. J., and Roe, A. J. (2012) LOV to BLUF: Flavoprotein Contributions to the Optogenetic Toolkit. *Mol. Plant* 5, 533–544.
- (12) Strickland, D., Moffat, K., and Sosnick, T. R. (2008) Light-Activated DNA Binding in a Designed Allosteric Protein. *Proc. Natl. Acad. Sci. U. S. A.* 105, 10709–10714.
- (13) Mart, R. J., Meah, D., and Allemann, R. K. (2016) Photocontrolled Exposure of Pro-Apoptotic Peptide Sequences in LOV Proteins Modulates Bcl-2 Family Interactions. *ChemBioChem* 17, 698–701.
- (14) Niopek, D., Wehler, P., Roensch, J., Eils, R., and Di Ventura, B. (2016) Optogenetic Control of Nuclear Protein Export. *Nat. Commun.* 7, 10624.
- (15) Renicke, C., Schuster, D., Usherenko, S., Essen, L. O., and Taxis, C. (2013) A LOV2 Domain-Based Optogenetic Tool to Control Protein Degradation and Cellular Function. *Chem. Biol.* 20, 619–626.
- (16) Lee, J., Natarajan, M., Nashine, V. C., Socolich, M., Vo, T., Russ, W. P., Benkovic, S. J., and Ranganathan, R. (2008) Surface Sites for Engineering Allosteric Control in Proteins. *Science* 322, 438–442.
- (17) Strickland, D., Yao, X., Gawlak, G., Rosen, M. K., Gardner, K. H., and Sosnick, T. R. (2010) Rationally Improving LOV Domain Based Photoswitches. *Nat. Methods* 7, 623–626.
- (18) Gehrig, S., Macpherson, J. A., Driscoll, P. C., Symon, A., Martin, S. R., MacRae, J. I., Kleijung, J., Fraternali, F., and Anastasiou, D. (2017) An Engineered Photoswitchable Mammalian Pyruvate Kinase. *FEBS J.* 284, 2955–2980.
- (19) Kawano, F., Suzuki, H., Furuya, A., and Sato, M. (2015) Engineered Pairs of Distinct Photoswitches for Optogenetic Control of Cellular Proteins. *Nat. Commun.* 6, 6256.
- (20) Crosson, S., and Moffat, K. (2001) Structure of a Flavin-Binding Plant Photoreceptor Domain: Insights into Light-Mediated Signal Transduction. *Proc. Natl. Acad. Sci. U. S. A.* 98, 2995–3000.
- (21) Herman, E., Sachse, M., Kroth, P. G., and Kottke, T. (2013) Blue-Light-Induced Unfolding of the  $\alpha$  Helix Allows for the Dimerization of Aureochrome-LOV from the Diatom *Phaeodactylum Tricornutum*. *Biochemistry* 52, 3094–3101.
- (22) Herman, E., and Kottke, T. (2015) Allosterically Regulated Unfolding of the  $\alpha$  Helix Exposes the Dimerization Site of the Blue-Light-Sensing Aureochrome-LOV Domain. *Biochemistry* 54, 1484–1492.
- (23) Harper, S. M., Christie, J. M., and Gardner, K. H. (2004) Disruption of the LOV- $\alpha$  Helix Interaction Activates Phototropin Kinase Activity. *Biochemistry* 43, 16184–16192.
- (24) Harper, S. M., Neil, L. C., and Gärtner, K. H. (2003) Structural Basis of a Phototropin Light Switch. *Science* 301, 1541–1544.
- (25) Möglich, A., and Moffat, K. (2007) Structural Basis for Light-Dependent Signaling in the Dimeric LOV Domain of the Photosensor YtvA. *J. Mol. Biol.* 373, 112–126.
- (26) Berntsson, O., Diensthuber, R. P., Panman, M. R., Björling, A., Hughes, A. J., Henry, L., Niebling, S., Newby, G., Liebi, M., Menzel, A., Henning, R., Kosheleva, I., Möglich, A., and Westenhoff, S. (2017) Time-Resolved X-Ray Solution Scattering Reveals the Structural Photoactivation of a Light-Oxygen-Voltage Photoreceptor. *Structure* 25, 933–938.
- (27) Engelhard, C., Diensthuber, R. P., Möglich, A., and Bittl, R. (2017) Blue-Light Reception through Quaternary Transitions. *Sci. Rep.* 7, 1385.
- (28) Zoltowski, B. D., Schwerdtfeger, C., Widom, J., Loros, J. J., Bilwes, A. M., Dunlap, J. C., and Crane, B. R. (2007) Conformational Switching in the Fungal Light Sensor Vivid. *Science* 316, 1054–1057.
- (29) Vaidya, A. T., Chen, C. H., Dunlap, J. C., Loros, J. J., and Crane, B. R. (2011) Structure of a Light-Activated LOV Protein Dimer That Regulates Transcription. *Sci. Signaling* 4, ra50.
- (30) Diensthuber, R. P., Bommer, M., Gleichmann, T., and Möglich, A. (2013) Full-Length Structure of a Sensor Histidine Kinase Pinpoints Coaxial Coiled Coils as Signal Transducers and Modulators. *Structure* 21, 1127–1136.
- (31) Jones, M. A., Feeney, K. A., Kelly, S. M., and Christie, J. M. (2007) Mutational Analysis of Phototropin 1 Provides Insights into the Mechanism Underlying LOV2 Signal Transmission. *J. Biol. Chem.* 282, 6405–6414.
- (32) Nash, A. I., Ko, W. H., Harper, S. M., and Gardner, K. H. (2008) A Conserved Glutamine Plays a Central Role in LOV Domain Signal Transmission and Its Duration. *Biochemistry* 47, 13842–13849.
- (33) Ganguly, A., Thiel, W., and Crane, B. R. (2017) Glutamine Amide Flip Elicits Long Distance Allosteric Responses in the LOV Protein Vivid. *J. Am. Chem. Soc.* 139, 2972–2980.
- (34) Heintz, U., and Schlichting, I. (2016) Blue Light-Induced LOV Domain Dimerization Enhances the Affinity of Aureochrome 1a for Its Target DNA Sequence. *eLife* 5, No. e11860.
- (35) Freddolino, P. L., Dittich, M., and Schulten, K. (2006) Dynamic Switching Mechanisms in LOV1 and LOV2 Domains of Plant Phototropins. *Biophys. J.* 91, 3630–3639.
- (36) Peter, E., Dick, B., and Baeurle, S. A. (2010) Mechanism of Signal Transduction of the LOV2-*Ja* Photosensor from *Avena Sativa*. *Nat. Commun.* 1, 122.
- (37) Freddolino, P. L., Gardner, K. H., and Schulten, K. (2013) Signaling Mechanisms of LOV Domains: New Insights from Molecular Dynamics Studies. *Photochem. Photobiol. Sci.* 12, 1158–1170.
- (38) Pudasaini, A., Shim, J. S., Song, Y. H., Shi, H., Kiba, T., Somers, D. E., Imaizumi, T., and Zoltowski, B. D. (2017) Kinetics of the LOV Domain of ZEITLUPE Determine Its Circadian Function in *Arabidopsis*. *eLife* 6, e21646.
- (39) Gil, A. A., Laptinok, S. P., French, J. B., Iuliano, J. N., Lukacs, A., Hall, C. R., Sazanovich, I. V., Greetham, G. M., Bacher, A., Illarionov, B., Fischer, M., Tonge, P. J., and Meech, S. R. (2017) Femtosecond to Millisecond Dynamics of Light Induced Allosteric in the *Avena Sativa* LOV Domain. *J. Phys. Chem. B* 121, 1010–1019.
- (40) Zayner, J. P., Antoniou, C., and Sosnick, T. R. (2012) The Amino-Terminal Helix Modulates Light-Activated Conformational Changes in *As* LOV2. *J. Mol. Biol.* 419, 61–74.
- (41) Mansurova, M., Scheercoose, P., Simon, J., Kluth, M., and Gärtner, W. (2011) Chromophore Exchange in the Blue Light-Sensitive Photoreceptor YtvA from *Bacillus Subtilis*. *ChemBioChem* 12, 641–646.
- (42) Fisher, J., Spencer, R., and Walsh, C. (1976) Enzyme-Catalyzed Redox Reactions with the Flavin Analogues 5-Deazariboflavin, 5-

Deazariboflavin 5'-Phosphate, and 5-Deazariboflavin 5'-Diphosphate, 5' → 5'-Adenosine Ester. *Biochemistry* 15, 1054–1064.

(43) Hemmerich, P., Massey, V., and Fenner, H. (1977) Flavin and 5-Deazaflavin: A Chemical Evaluation of "modified" Flavoproteins with Respect to the Mechanisms of Redox Biocatalysis. *FEBS Lett.* 84, 5–21.

(44) Hedison, T. M., Leferink, N. G. H., Hay, S., and Scrutton, N. S. (2016) Correlating Calmodulin Landscapes with Chemical Catalysis in Neuronal Nitric Oxide Synthase Using Time-Resolved FRET and a 5-Deazaflavin Thermodynamic Trap. *ACS Catal.* 6, 5170–5180.

(45) Takahashi, F., Yamagata, D., Ishikawa, M., Fukamatsu, Y., Ogura, Y., Kasahara, M., Kiyosue, T., Kikuyama, M., Wada, M., and Kataoka, H. (2007) AUREOCHROME, a Photoreceptor Required for Photomorphogenesis in Stramenopiles. *Proc. Natl. Acad. Sci. U. S. A.* 104, 19625–19630.

(46) Banerjee, A., Herman, E., Serif, M., Maestre-Reyna, M., Hepp, S., Pokorny, R., Kroth, P. G., Essen, L. O., and Kottke, T. (2016) Allosteric Communication between DNA-Binding and Light-Responsive Domains of Diatom Class 1 Aureochromes. *Nucleic Acids Res.* 44, 5957–5970.

(47) Herman, E., Sachse, M., Kroth, P. G., and Kottke, T. (2013) Blue-Light-Induced Unfolding of the J-Alpha Helix Allows for the Dimerization of Aureochrome-LOV from the Diatom *Phaeodactylum Tricornutum*. *Biochemistry* 52, 3094–3101.

(48) Akiyama, Y., Nakasone, Y., Nakatani, Y., Hisatomi, O., and Terazima, M. (2016) Time-Resolved Detection of Light-Induced Dimerization of Monomeric Aureochrome-1 and Change in Affinity for DNA. *J. Phys. Chem. B* 120, 7360–7370.

(49) Mitra, D., Yang, X., and Moffat, K. (2012) Crystal Structures of Aureochrome1 LOV Suggest New Design Strategies for Optogenetics. *Structure* 20, 698–706.

(50) Banerjee, A., Herman, E., Kottke, T., and Essen, L. O. (2016) Structure of a Native-like Aureochrome 1a LOV Domain Dimer from *Phaeodactylum Tricornutum*. *Structure* 24, 171–178.

(51) Yee, E. F., Diensthuber, R. P., Vaidya, A. T., Borbat, P. P., Engelhard, C., Freed, J. H., Bittl, R., Möglich, A., and Crane, B. R. (2015) Signal Transduction in Light-Oxygen-Voltage Receptors Lacking the Adduct-Forming Cysteine Residue. *Nat. Commun.* 6, 10079.

(52) Halavaty, A. S., and Moffat, K. (2007) N- and C-Terminal Flanking Regions Modulate Light-Induced Signal Transduction in the LOV2 Domain of the Blue Light Sensor Phototropin 1 from *Avena Sativa*. *Biochemistry* 46, 14001–14009.

(53) Purcell, E. B., McDonald, C. A., Palfe, B. A., and Crosson, S. (2010) An Analysis of the Solution Structure and Signaling Mechanism of LovK, a Sensor Histidine Kinase Integrating Light and Redox Signals. *Biochemistry* 49, 6761–6770.

(54) Lans, I., Frago, S., and Medina, M. (2012) Understanding the FMN Cofactor Chemistry within the Anabaena Flavodoxin Environment. *Biochim. Biophys. Acta, Bioenerg.* 1817, 2118–2127.

(55) Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 Suite and Current Developments. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 67, 235–242.

(56) Vagin, A. A., Steiner, R. A., Lebedev, A. A., Potterton, L., McNicholas, S., Long, F., and Murshudov, G. N. (2004) REFMAC 5 Dictionary: Organization of Prior Chemical Knowledge and Guidelines for Its Use. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 60, 2184–2195.

(57) Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 66, 486–501.