Time-resolved monochromatic synchrotron crystallography of a plant photoreceptor domain

Sylvain Aumonier

Communauté
UNIVERSITÉ Grenoble Alpes

THÈSE

Pour obtenir le grade de

DOCTEUR DE LA COMMUNAUTE UNIVERSITE GRENOBLE ALPES

Spécialité : **Biologie Structurale et Nanobiologie**

Arrêté ministériel : 25 mai 2016

Présentée par

Sylvain AUMONIER

Thèse dirigée par **Antoine ROYANT**
codirigée par **David von STETTEN** et **Gordon LEONARD**

préparée au sein du **Structural Biology Group,**
European Synchrotron Radiation Facility
dans **l'École Doctorale Chimie et Sciences du Vivant**

Cristallographie monochromatique résolue en temps au synchrotron d'un domaine de photorécepteur de plante

–

Time-resolved monochromatic synchrotron crystallography of a plant photoreceptor domain

Thèse soutenue publiquement le **11 Juillet 2019**,
devant le jury composé de :

Président du jury : Pr. Eva PEBAY-PEYROULA
Professeure, Université Grenoble-Alpes, France
Rapporteur : Pr. Lars-Oliver ESSEN
Professeur, Philipps University of Marburg, Germany
Rapporteur : Dr. Pascal ARNOUX
Chercheur CEA HDR, BIAM CEA Cadarache, France
Examinatrice : Pr. Elspeth GARMAN
Professeure, University of Oxford, United Kingdom
Examinatrice : Pr. Arwen PEARSON
Professeure, University of Hambourg, Germany
Directeur de thèse : Dr. Antoine ROYANT
Directeur de Recherche, CNRS, France

Acknowledgements

I would like to thank my thesis supervisors for welcoming me at the ESRF and helping me to cope with the multidisciplinary character of my PhD subject. First, Dr. Antoine Royant for his daily monitoring and for the collaborative working relationship we have developed about all the aspects of the thesis work and beyond. I also thank Dr. Gordon Leonard for his experience and advice about the experiments and the writing, and Dr. David von Stetten for his involvement in the experiment design and the technical and data processing work on the beamlines.

I warmly thank the members of my thesis defence jury for their insightful comments and questions: Prof. Eva Pebay Peyroula as president of the jury, Dr. Pascal Arnoux and Prof. Lars-Oliver Essen as referees and Prof. Arwen Pearson and Prof. Elspeth Garman as examiners.

I am grateful to Dr. Eric Girard and Dr. Laurent Terradot for being part of the thesis monitoring committee and for their useful advice about the scientific work and the overall organization of the project.

I would like to particularly thank Dr. Guillaume Gotthard for his help and coaching about the various challenges we had to face during these several years of intensive work, and Dr. Gianluca Santoni for his help with data processing and computing issues.

Many thanks to the beamline technicians of the Structural Biology group Fabien, Hugo and Thierry for their technical support on the beamlines.

I warmly thank Dr. Montserrat Soler-Lopez for welcoming me in the wet lab of the ESRF Structural Biology group, Samira, Ulrike and Jennyfer for their help in the lab and all PhD students, post-docs and scientific staff for sharing technical and theoretical knowledge.

I would like to thanks the ESRF administrative staff, and more specifically Claudine Romero which has handled all the logistic details for the PhD defence, leaving me nothing else to do than the scientific work.

I also thank Jean Susini and Christoph Mueller-Dieckmann for their significant support to my PhD project.

I am thankful to Prof. Santiago Nonell from the Chemical Institute of Sarriá (IQS) in Barcelona, Prof. Theodorus Gadella from the University of Amsterdam and to Dr. Nathan Shaner from the Scintillon Institute in San Diego, for the fruitful scientific collaborations.

Enfin je remercie les amis qui m'ont permis de décompresser durant ces 3 (presque 4 !) années de travail, avec qui j'ai partagé quelques bières et sorties de ski/montagne rendant la science plus facile à faire : Etienne, Oriane, Romain, Julien, Rémi, Benjamin, Emilie, Pauline, Justin, Laurent, Guigui, Quentin, Joyce, Maxime, Nico, Fab, Matt, Arthur, Gauthier, Stan, Lilia, Florent.

Je remercie également mes parents et ma famille qui m'ont soutenu durant toutes ces longues années d'études.

Enfin je tiens à remercier Alexandra, qui m'a soutenu jusqu'à la fin, me supportant et m'apportant une aide précieuse dans les moments plus ou moins faciles de la thèse.

Cristallographie monochromatique résolue en temps au synchrotron d'un domaine de photorécepteur de plante

Résumé de la thèse en français

1 – Présentation générale de la thèse

La cristallographie résolue dans le temps (TRX) permet l'identification à l'échelle quasi atomique des changements structuraux progressifs au sein d'une protéine lorsqu'elle réalise sa fonction, permettant la constitution de véritables films moléculaires. La diffraction Laue (polychromatique) dans les synchrotrons de 3ème génération a d'abord permis d'obtenir de tels films à une échelle de temps de 100 ps, et plus récemment les sources de rayons X de 4ème génération, les lasers à électrons libres (XFELs, pour X-ray Free Electron Lasers), ont permis d'atteindre une résolution temporelle de 100 fs en utilisant un faisceau monochromatique.

L'objectif de cette thèse était d'explorer la possibilité d'effectuer des expériences de TRX au synchrotron sur les lignes de lumière monochromatiques dédiées à la cristallographie macromoléculaire en tirant parti des tous derniers développements technologiques. Nous avons envisagé de travailler sur trois protéines photosensibles différentes, pour lesquelles j'ai dans un premier temps développé différentes stratégies de cristallogenèse afin de contrôler la taille et la forme des cristaux. J'ai également testé différentes façons d'exposer les cristaux au faisceau de rayons X à température ambiante, soit des cristaux uniques montés sur une boucle et maintenu dans un flux d'humidité contrôlée, soit des microcristaux incorporés dans un flux de graisse passant à travers le faisceau de rayons X. Nous avons ensuite évalué comment les dommages d'irradiation spécifiques pouvaient affecter les structures d'état intermédiaires de protéines à température ambiante, espèces particulièrement sensibles à température cryogénique en général. Nous avons conclu de cette étude que les dégâts d'irradiation spécifiques constituent un bien moindre problème à température ambiante qu'à température cryogénique.

Après une caractérisation initiale des différentes protéines et des moyens de présentation des cristaux au faisceau, nous avons concentré notre approche de cristallographie résolue dans le temps sur des cristaux d'un domaine de photorécepteur de plante, *At*Phot2LOV2 (le domaine LOV2 du récepteur à la lumière bleue phototropine 2 de la plante *Arabidopsis thaliana*). L'état initial de *At*Phot2LOV2 se convertit en quelques microsecondes

en un état de signalisation, ou état photo-activé, qui se relaxe en quelques centaines de secondes. En utilisant un détecteur rapide de rayons X, nous avons caractérisé structuralement le phénomène de relaxation, et montré qu'elle se développe en quelques centaines de secondes *via* une conversion de groupe d'espace. Nous avons ensuite diminué la vitesse de formation de l'état photo-activé dans les cristaux en limitant la quantité de photons nécessaires à la photoconversion, et nous avons pu visualiser l'augmentation progressive de la population en l'état photo-activé dans le cristal avec une résolution temporelle de 63 ms, en utilisant une approche basée sur l'enregistrement de données de diffraction sur moins de 100 cristaux et en tirant parti de la combinaison de jeux de données partiels, approche que nous avons nommée TR-SOX (Time-Resolved Serial Oscillation Crystallography). En résumé, notre travail ouvre la voie aux expériences de cristallographie résolue dans le temps sur les lignes de lumières monochromatiques de cristallographie macromoléculaire au synchrotron, sur des cristaux de taille usuelle pour des protéines subissant des réarrangements structuraux à l'échelle de la milliseconde.

2 – Cristallisation de protéines colorées et contrôle de la taille des cristaux pour les expériences résolues dans le temps.

La première partie de mon travail de thèse a consisté en l'expression, la purification ainsi que la cristallisation de différentes protéines qui sont potentiellement intéressantes à étudier par cristallographie résolue dans le temps. J'ai d'abord exploré la cristallisation de protéines fluorescentes rouges et de chromoprotéines (colorées, mais non fluorescentes) de type GFP. Cette étude a permis notamment la résolution de la structure de mScarlet, la protéine rouge la plus brillante mise au point à ce jour, en collaboration avec le groupe du Professeur Gadella de l'Université d'Amsterdam. Nous avons mené une étude extensive de 6 mutants de cette protéine, qui a conduit à une publication dans *Nature Methods* (Bindels *et al.*, 2017), mais qui m'a également permis de développer des stratégies de cristallogenèse que j'ai pu ensuite appliquer aux trois protéines de mon projet principal (*Dr*CBD, Twist-Cerulean et *At*Phot2LOV2) avec de légères variations dans chaque cas.

Mon travail a d'abord consisté en l'optimisation de l'expression et de la purification de ces trois protéines, afin d'avoir à disposition des échantillons d'un haut degré de pureté, permettant d'obtenir des cristaux diffractant à haute résolution. J'ai ensuite élaboré différents protocoles permettant de contrôler la taille, la forme et le nombre des cristaux, allant d'un

cristal unique de taille « macroscopique » (50 – 500 µm) pour notre approche résolue dans le temps utilisant des méthodes de collectes de données avec oscillation, à des milliers de microcristaux (< 10 µm) dans le cadre de la cristallographie sérielle pour les méthodes ayant recours à l'injection de microcristaux.

3 –Comparaison de la vitesse d'apparition des dommages d'irradiation spécifiques et globaux à température ambiante – Implications pour les expériences résolues dans le temps.

Afin d'envisager des expériences de cristallographie résolue dans le temps, nous nous sommes d'abord familiarisés avec les collectes de données synchrotron à température ambiante, en utilisant le contrôleur d'humidité HC1 pour lequel il faut estimer au mieux l'humidité relative requise pour une condition de cristallisation donnée. Parce qu'il a été montré que les états intermédiaires de réaction étaient souvent particulièrement sensibles aux dommages d'irradiation spécifiques à température cryogénique, nous nous sommes demandé comment contrôler l'étendue des dommages spécifiques à température ambiante. Comme les dégâts spécifiques sont bien plus rapides que les dommages globaux à température cryogénique, et comme les dégâts globaux à température ambiante sont bien plus rapides qu'à température cryogénique, nous craignions qu'il faille limiter drastiquement les doses absorbées par le cristal de protéine pour avoir une chance de s'affranchir de l'artefact des dégâts spécifiques. Nous avons tout d'abord caractérisé la vitesse des dommages radiatifs spécifiques à température ambiante sur des protéines pour lesquelles ces dommages spécifiques avaient déjà été bien caractérisés à température cryogénique : la protéine fluorescente cyan Cerulean ainsi que le lysozyme de blanc d'œuf de poule. La difficulté à visualiser des dommages spécifiques pour ces deux protéines à température ambiante nous a mené à définir un ensemble de paramètres permettant de comparer les vitesses d'apparition des dommages spécifiques et globaux, et ce à températures cryogénique et ambiante.

De cette étude, nous avons conclu que les vitesses d'apparition des deux types de dommages, qui sont très différentes à température cryogénique (découplage des deux types de dommages), sont proches l'une de l'autre à température ambiante (couplage, ou recouplage des deux types de dommages). Ceci indique que les dommages d'irradiation spécifiques peuvent-être évités en première approximation à température ambiante, à condition de se mettre dans les conditions d'enregistrer un jeu de données complet de diffraction par oscillation à partir un

cristal unique en atténuant suffisamment le faisceau de rayons X. Nous avons mis en application ce résultat en montrant que nous pouvions déterminer la structure à température ambiante d'un intermédiaire réactionnel du photocycle d'un domaine LOV de la phototropine 2 de la plante *Arabidopsis thaliana*, connu pour être particulièrement sensible aux dommages d'irradiation à température cryogénique. Cette partie de la thèse explorant les effets des dommages radiatifs spécifiques et globaux sur des cristaux de protéines a fait l'objet d'une publication dans *IUCrJ* (Gotthard*, Aumonier* *et al.*, 2019).

4 – Cristallographie sérielle basée sur des méthodes d'injection de microcristaux et développements préliminaires de la méthode de cristallographie résolue en temps basée sur des collectes de jeux de données par oscillation sur macro-cristaux de la protéine *At*Phot2LOV2

Afin de réaliser nos expériences de cristallographie résolue en temps, nous avons utilisé différentes méthodes de présentation des cristaux au faisceau de rayons X, et différentes façons de collecter les données. J'ai pu ainsi utiliser un injecteur de microcristaux en milieu visqueux sur la ligne de lumière ID30A-3, avec lequel j'ai pu tester des microcristaux de nos trois protéines cibles, donnant des résultats constituant une preuve de principe de la faisabilité de telles expériences, y compris sous illumination actinique. Nous avons également utilisé des détecteurs de rayons X de plus en plus rapides (du Pilatus 6M à l'Eiger 4M) pour collecter des jeux de données complets sur un cristal unique. Nous avons ainsi diminué la résolution temporelle potentielle avec laquelle un phénomène pourrait être étudié par cristallographie à température ambiante, de 30 secondes à environ 1 seconde. Tous ces développements et expériences préliminaires ont permis de réduire cette résolution temporelle à 63 millisecondes, grâce à des développements méthodologiques décrits ci-après.

5 – Cristallographie résolue en temps à temperature ambiante d'un domaine de photorécepteur de plante avec une résolution temporelle de 3 secondes – Caractérisation de la relaxation de l'état photo-activé de *At*Phot2LOV2

Les travaux de TRX sur la protéine *At*Phot2LOV2 ont été divisés en deux projets distincts. Le premier, techniquement plus facile grâce à la disponibilité du détecteur de rayons X Eiger 4M, a consisté à caractériser la relaxation de l'état photo-activé sur une échelle de temps longue (de la seconde à plusieurs minutes), et le second, techniquement plus complexe

à mettre en œuvre, a permis de visualiser l'augmentation de la population de l'état photo-activé au sein d'un cristal sur l'échelle de la milliseconde (Chapitre 6).

Dans le chapitre 3, nous avons montré qu'il était possible de déterminer à température ambiante, et en temps réel (c'est-à-dire sans avoir à maintenir l'irradiation lumineuse pendant la collecte), la structure de l'état photo-excité du domaine du photo-récepteur de plante *At*Phot2LOV2, sans détruire de façon significative la liaison covalente entre le chromophore FMN et le résidu cystéine voisin. Nous avions confirmé que cette liaison était pourtant particulièrement sensible aux dommages d'irradiation à température cryogénique. Ceci nous a donné l'idée de suivre le phénomène de relaxation de l'état photo-excité de *At*Phot2LOV2 par cristallographie résolue en temps sur une échelle de temps relativement lente, de l'ordre de quelque centaine de secondes. Notre expérience préliminaire a été effectuée en utilisant un détecteur de rayon X du type Pilatus 6M, nous permettant d'atteindre une résolution temporelle entre deux collectes de jeux de données successifs d'environ 30 secondes. La comparaison de deux jeux de données enregistrés successivement sur un même cristal à [0-30 s] et [30-60 s] a montré que l'état photo-excité est toujours présent de façon significative dans le second jeu de données, évidemment dans une proportion moindre.

L'installation d'un nouveau détecteur de rayons X plus rapide sur la ligne de lumière ID30A-3, un Eiger 4M, capable de fonctionner à une fréquence de 750 Hz, nous a donné l'opportunité d'améliorer drastiquement la durée d'une collecte, jusqu'à 2 ou 3 secondes pour un jeu complet. En enregistrant 19 points temporels entre 0 et 20 minutes, nous avons pu proposer une description du phénomène de relaxation de l'état photo-excité, dont le résultat le plus remarquable est le changement de groupe d'espace induit par le retour à l'état initial, qui n'est pas possible pour une molécule sur deux dans le cristal.

6 – Cristallographie résolue en temps et à température ambiante d'un domaine de photorecepteur de plante avec une résolution temporelle de 63 millisecondes – Développement de la méthode TR-SOX (Time-Resolved Serial Oscillation Crystallography)

Le suivi de la population en état photo-activé d'un cristal de *At*Phot2LOV2 à l'échelle de la milliseconde a constitué l'un des défis majeurs de notre étude. Lors de la caractérisation du phénomène de relaxation de la protéine (chapitre 5), nous avions observé une perte de fluorescence progressive des cristaux de *At*Phot2LOV2 lorsque nous initions l'état photo-activé en irradiant les cristaux avec une lumière bleue. Comme ce phénomène survenait sur

l'échelle de la seconde, alors que l'état photo-excité a un temps de montée de l'ordre de la microseconde en solution, nous avons émis l'hypothèse que ce que nous observions au travers de cette perte de fluorescence du cristal était en fait sa photo-transformation progressive sous une illumination bleue d'intensité modérée. En bref, la LED que nous avons utilisée ne fournit pas en quelques microsecondes la quantité de photons nécessaires pour transformer toutes les molécules du cristal. Cependant, la photo-activation progressive du cristal constitue un phénomène beaucoup plus rapide que sa relaxation (milliseconde-seconde contre seconde-minute), et nous avons cherché à échantillonner la visualisation du phénomène avec la résolution temporelle la plus courte possible, avec pour objectif avoué la dizaine de millisecondes.

Nous avons mis au point une méthode de cristallographie sérielle sur cristaux de taille usuelle (50 – 100 microns) en collectant des jeux de données complets avec oscillation sous illumination de lumière bleue, sur moins d'une centaine de cristaux différents. Ces jeux de données ont été découpés en jeux de données partiels correspondant à des points de mesure temporels successifs séparés de 63 millisecondes puis ont été recombinés entre eux afin de former des jeux de données complets pour chaque point de mesure couvrant ainsi le phénomène sur une durée totale de 4,158 secondes avec une résolution temporelle de 63 ms. Le développement de cette méthode, dite TR-SOX pour cristallographie d'oscillation sérielle résolue en temps, sa mise en application et les résultats concernant la protéine *At*Phot2LOV2 sont décrits dans le manuscrit constituant ce chapitre.

Références :

Bindels, D. S., Haarbosch, L., van Weeren, L., Postma, M., Wiese, K. E., Mastop, M., Aumonier, S., Gotthard, G., Royant, A., Hink, M. A. & Gadella Jr, T. W. J. (2017). *Nat. Methods*. 1–12.

Gotthard, G., Aumonier, S., Sanctis, D. De, Leonard, G., Stetten, D. Von & Royant, A. (2019). *IUCrJ*. **6**, https://doi.org/10.1107/S205225251900616X.

List of abbreviations

APS	Advanced Photon Source
*At*Phot2LOV2	LOV2 domain of phototropin 2 from *Arabidopsis thaliana*
BFP	Blue Fluorescent Protein
BR	Bacteriorhodopsin
CCD	Charged coupled device
CFP	Cyan Fluorescent Protein
CLEM	Correlation light and electron microscopy
CRIMS	CRystallization Information Management System
CT	Cryogenic temperature
Cryo-EM	Cryogenic temperature electron microscopy
Cryo-MX	Cryogenic temperature macromolecular crystallography
DAB	3,3'-Diamonibenzidine
DFFN	Double-flow focusing nozzle
DNA	Deoxyribonucleic acid
*Dr*BphP	*Deinococcus radiodurans* bacteriophytochrome
*Dr*CBD	*Deinococcus radiodurans* Chromophore Binding Domain
DTT	Dithiothreitol
EMBL	European Molecular Biology Laboratory
ESRF	European Synchrotron Radiation Facility
FMN	Flavin mononucleotide
FRET	Förster (or fluorescence) resonance energy transfer
FWHM	Full width at half maximum
GFP	Green Fluorescent Protein
GDVN	Gas dynamic virtual nozzle
HARE	Hit-And-REturns
HC1	Humidity controller 1
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HEWL	Hen egg-white lysozyme
HPLC	High pressure liquid chromatography
HTX	High throughput crystallization
HVE	High viscosity extrusion
IBS	Institut de Biologie Structurale
*ic*OS	*in crystallo* optical spectroscopy

IFP	Infrared Fluorescent Protein
LCP	Lipidic cubic phase
LED	Light-emitting diode
LOV	Light Oxygen and Voltage
MBM	Microbatch mixing
MD2	Microdiffractometer 2
MES buffer	2-(*N*-morpholino)-ethanesulfonic acid
miniSOG	minimal-sized Singlet Oxygen Generator
MPD	2-Methyl-2,4-pentanediol
MX	Macromolecular crystallography
NMR	Nuclear magnetic resonance
PAS	Per, ARNT, Sim
PEG	Polyethylene glycol
PDB	Protein Data Bank
PS-I	Photosystem-I
PSS	Personal safety system
PYP	Photoactive Yellow Protein
RFP	Red Fluorescent Protein
RH	Relative humidity
ROS	Reactive oxygen species
RT	Room temperature
rsEGFP2	reversibly photoswitchable Enhanced Green Fluorescent Protein 2
SASE	Self-amplified spontaneous emission
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SFX	Serial femtosecond crystallography
SLS	Swiss Light Source
SMX	Serial millisecond crystallography
SSX	Serial synchrotron crystallography
TR	Time-resolved
TR-MX	Time-resolved macromolecular crystallography
TR-SFX	Time-resolved serial femtosecond crystallography
TR-SOX	Time-resolved serial oscillation crystallography
TR-SSX	Time-resolved serial synchrotron crystallography
TRX	Time-resolved crystallography
TTL	Transistor-transistor logic
XFEL	X-ray free electron laser

Table of contents

Chapter 1 – Introduction.....1

1.1 Time-resolved protein crystallography at synchrotrons and XFELs.....3

1.1.1 Slow time-scale time-resolved crystallography by classical protein crystallography.....3

1.1.2 Time-resolved Laue diffraction: extraordinary results, yet on few systems.....5

1.1.3 Time-resolved crystallography at X-ray Free Electron Lasers (XFELs).....7

1.2 Time-resolved protein crystallography on monochromatic synchrotron beamline?.....12

1.2.1 Injection devices.....13

1.2.2 Fixed-target sample holders.....14

1.2.3 First examples of dose-resolved, or time-resolved experiments at synchrotrons.....14

1.3. TR-MX – on which biological systems?.....17

1.4. The LOV2 domain of phototropin 2 from *Arabidopsis thaliana*.....18

1.4.1 The blue light plant photoreceptors phototropin.....18

1.4.2 The LOV domains from phototropin.....19

1.4.3 Crystallographic studies on the photocycle of LOV domains.....20

1.4.4 Biotechnological applications of LOV domains.....21

1.5 Fluorescent proteins and chromoproteins.....23

1.6 *Dr*CBD: the chromophore-binding domain of the phytochrome from *Deinococcus radiodurans*.....24

1.7 Preamble to thesis manuscript.....25

1.8 References.....28

Chapter 2 – Crystallogenesis: methods and results - Coloured protein crystallization and crystal size control tuning for time-resolved study purposes.....35

2.1 Crystallisation strategies.....37

2.1.1 High throughput crystallization conditions screening.....37

2.1.2 Crystal optimization.....38

2.1.3 Limited proteolysis.....39

2.1.4 Micro-seeding protocol.....41

2.2 Application to coloured proteins.....43

2.2.1 The chromoproteins Pumpkin and Aubergine.....43

2.2.2 The Red Fluorescent Protein mScarlet.....45

2.2.3 The Cyan Fluorescent Protein Twist-Cerulean.....48

2.2.4 *At*Phot2LOV2: the LOV2 domain of phototropin 2 from *Arabidopsis thaliana*.....52
2.2.5 *Dr*CBD: the chromophore-binding domain of bacteriophytochrome from *Deinoccocus radiodurans*.....54
2.3 Crystal preservation in diffraction or spectroscopic data collection.....58
2.3.1 Cryogenic temperature data collection.....58
2.3.2 Room temperature data collection.....58
2.4 References.....59

Chapter 3 – Results – Article 1 – Comparison of the rate of specific and global radiation damage at room temperature –Implications for time-resolved crystallography.....61
Presentation of Article 1.....63
3.1 Introduction.....65
3.2. Methods.....68
3.2.1 Protein expression and purification.....68
3.2.2 Protein crystallization.....68
3.2.3 X-ray data collection.....69
3.2.4 Data reduction and structure refinement.....70
3.2.5 Online Raman spectroscopy.....71
3.2.6 *In crystallo* UV-visible absorption spectroscopy of LOV2.....71
3.3 Results.....72
3.3.1 Cerulean.....72
3.3.2 Hen Egg-White Lysozyme.....77
3.3.3 Photoadduct of the LOV2 domain from *Arabidopsis thaliana* phototropin 2.....81
3.4 Discussion.....86
3.5 References.....97

Chapter 4 – Methods & Unpublished Results – Time-Resolved Crystallography.....103
4.1 Crystal injection-based Serial Synchrotron Crystallography (SSX).....105
4.1.1 Sample preparation protocol.....105
4.1.2 Sample loading and operation of the HVE injector at the beamline.....106
4.1.3 Data collection, data integration and structure refinement.....109
4.1.4 Results obtained from Twist-Cerulean microcrystals.....109
4.1.5 Results obtained from *Dr*CBD crystals.....112
4.1.6 Results obtained from *At*Phot2LOV2 crystals.....112
4.2 Early developments of the Time-Resolved Serial Oscillation Crystallography (TR-SOX) method on *At*Phot2LOV2.....116

4.2.1 Spectroscopic characterization of the dark and light states of *At*Phot2LOV2....116
4.2.2 Proof of principle experiment: 30 s time resolution....117
4.2.3 Sub-second time-resolved crystallography....119
4.3 References....124

Chapter 5 – Results – Article 2 – Room temperature time-resolved crystallography on a plant photoreceptor domain with a 3 s time resolution – Monitoring of the slow relaxation of the light state of *At*Phot2LOV2....125
Presentation of Article 2....127
5.1 Introduction....130
5.2 Methods....131
5.2.1 Protein expression, purification and crystallization....131
5.2.2 X-ray data collection....131
5.2.3 Data reduction and structure refinement....132
5.2.4 *In crystallo* UV-visible absorption spectroscopy....132
5.3 Results....133
5.4 Discussion....142
5.5 References....144

Chapter 6 – Results – Article 3 – Room temperature time-resolved crystallography on a plant photoreceptor domain with a 63 ms time resolution – Development of the Time-Resolved Serial Oscillation Crystallography (TR-SOX) method....145
Presentation of Article 3....147
6.1 Introduction....149
6.2 Development of the TR-SOX method....151
6.3 Structural analysis and discussion....153
6.4 Conclusion....157
6.5 Methods....157
6.5.1 Protein expression and purification....157
6.5.2 Protein crystallization....158
6.5.3 *In crystallo* UV-visible absorption spectroscopy....158
6.5.4 Diffraction data collection and sorting....158
6.5.5 Diffraction data set processing....159
6.5.6 Structure determination and refinement of the dark and light states....160
6.5.7 Occupancy refinement for all time points....160

6.6 References.....161

6.7 Supplementary information.....164

Chapter 7 – Conclusion.....181

7.1 Leads for improvement of the TR-SOX method.....183

7.2 Fixed-target SSX – eliminating the relaxation waiting time using in plate diffraction.....184

7.3 Use of microfluidics in time-resolved experiments.....185

Annex 1 – Results – Article 4 – Beyond the photosensory function of *At*Phot2LOV2: turning it into a ROS generator.....187

Article 4: Tailing miniSOG: structural bases of the complex photophysics of a flavin-binding singlet oxygen photosensitizing protein.....189

Supplementary information.....199

Chapter 1

INTRODUCTION

Time-resolved crystallography

-

Coloured proteins studied during the PhD work

-

Preamble to thesis manuscript

1.1 Time-resolved protein crystallography at synchrotrons and XFELs

X-ray crystallography is currently the most straightforward way of obtaining a near-atomic resolution structure of a biological macromolecule. The PDB currently contains more than 150,000 structures essentially obtained by X-ray crystallography (89%), and to a lesser extent, by NMR (8%) and electron microscopy (2%). X-ray structures give a static picture of a macromolecule (a protein for instance) at a given time by averaging the electron density of over billions of individual molecules. Understanding the mechanism of a protein requires to determine a number of distinct structures corresponding to various states adopted by the molecule in the course of fulfilling its function. The difficulty lays in how to populate those billions of molecules in one of these states. The various techniques and methods used to achieve this goal are called kinetic crystallography (Bourgeois & Royant, 2005), which may rely on structure determination at cryogenic temperature, but also at room temperature. In the former case, one generally talks of intermediate trapping experiments, usually performed at synchrotrons. In the latter case, time-resolved experiments can be envisaged. In particular, they have been the reason for the development of Laue diffraction on proteins at synchrotrons, and for the recent focus of the Structural Biology community on X-ray Free Electron Lasers (XFELs). One shall note that the booming field of cryo-electron microscopy (cryo-EM) ("Method of the year 2015" - *Nat. Meth.* 2016; Cheng, 2018) is, by definition, restricted to the trapping of intermediates.

1.1.1 Slow time-scale time-resolved crystallography by classical protein crystallography

On monochromatic macromolecular crystallography (MX) beamlines, time-resolved crystallography has long been hampered by the combination of the two facts that diffraction data collections were taking 10-100 min, and that crystals had to be kept at cryogenic temperature, at which protein dynamics are essentially frozen out, in order to withstand radiation damage. As a consequence, the time resolution applies to the sample preparation step, not to the data recording one, and is only possible on reactions occurring on the time scale of tens of seconds. Such an experiment consists in initiated the reaction by adding a substrate or cofactor to a crystal-containing drop, from which crystals are serially fished and flash-cooled at various time points. A prototypical example is that of the human DNA polymerase η (Nakamura *et al.*, 2012) whose reaction could be monitored by classical X-ray crystallography

at time points between 40 and 300 s after mild alkalinisation of the crystal conditions and addition of Mg^{2+} (**Fig. 1.1**).

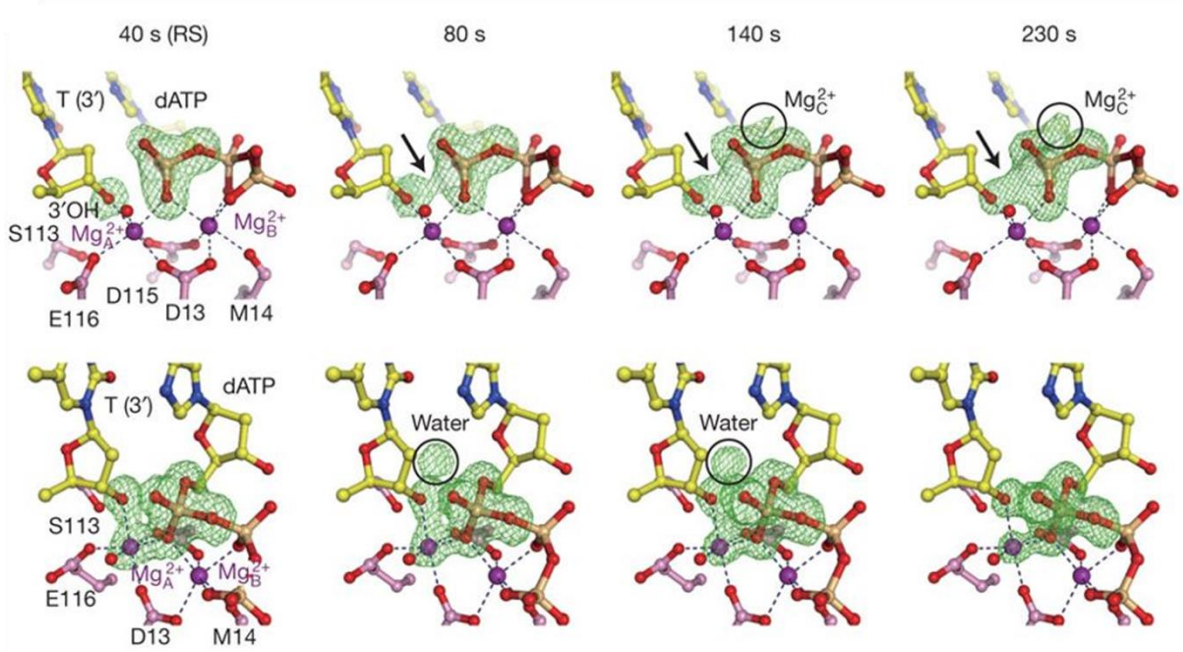


Figure 1.1: Cryotrapping experiments on crystals of DNA polymerase η. Evolution of the $2F_{obs} - F_{calc}$ map (with the terminal phosphate of dATP omitted) on the 40 to 230 s time scale showing phosphodiester bond formation (arrow). A water molecule transiently relocates, and a third Mg^{2+} ion is recruited to stabilize the intermediate (Adapted from Nakamura *et al.*, 2012).

The availability of the first large pixel X-ray detector in 2009 (Pilatus 6M, Dectris, Baden-Dättwil, Switzerland) has opened up the possibility of recording faster data sets than with CCD X-ray detectors, and data sets could then be recorded in less than a minute (frame rate of 8-25 Hz on beamline ID29 of the ESRF). Further developments have led to the Eiger series in 2015, which allow for data collections in less than a second (750 Hz frame rate). The Jungfrau detector, soon to be available, will be even faster (1.14-2.4 kHz). These detectors are so fast that it now puts the limitation of oscillation data collection on the speed of the goniometer.

Besides, the development of the wet humidity controller HC1 (sometimes called de-humidifier) (Sanchez-Weatherby *et al.*, 2009) initially destined to perform controlled dehydration crystal experiments (with the goal of increasing crystal diffraction resolution), has greatly facilitated the recording of diffraction data collection at room temperature on single

crystals (Bowler *et al.*, 2015) when compared to the capillary technique, provided the X-ray beam has been sufficiently attenuated.

The combination of fast pixel detectors and sample environment equipment ensuring preservation of crystal diffraction quality has paved the way for the development of time-resolved crystallography experiments on monochromatic synchrotron MX beamlines.

1.1.2 Time-resolved Laue diffraction: extraordinary results, yet on few systems

Laue diffraction consists of shooting a protein crystal bathing in its mother liquor in a sealed capillary at room temperature with a short, intense polychromatic X-ray pulse of 100 ps width (at the ESRF). A partial data set is obtained from each crystal, of which each image contains diffraction spots for different wavelengths, and thus less images are required from a few crystals to result in a complete data set than with a monochromatic beam. This puts however constraints on crystal quality, which need to be of low mosaicity and well-diffracting.

Time-resolved Laue diffraction experiments can be summarized as pump-probe experiments, in which the pump is a pulsed visible light laser with a short pulse width (of femtoseconds to nanoseconds), and the probe is the X-ray pulse (of typically sub-nanosecond width) (Wulff *et al.*, 1997). Delays between the pump and probe provide the time points of the time-resolved diffraction experiment, and the time resolution depends on the widths of the light and X-ray pulses, but is usually dominated by that of the X-ray pulse. The need to repeat the experiment several times on the same crystal put constraints on the reaction that can be studied. It needs to be repeatable with a fast turn-over, and obviously, the studied phenomenon needs to be a photoreaction, not an enzymatic one.

Unfortunately, these constraints have limited the number of protein, for which time-resolved Laue diffraction has led to devising molecular movies. The studies with the highest time resolution are the following: myoglobin from 100 ps to 3.14 µs with a 150 ps time resolution (Schotte *et al.*, 2003) (**Fig. 1.2**), photoactive yellow protein from from 3 ns to 3 s with a 150 ps time resolution (Bourgeois *et al.*, 2003) and from 100 ps to 1 s with a 150 ps time resolution (Schotte *et al.*, 2012), a bacterial photosynthetic reaction centre with one snapshot 3 ms after photolysis (Wöhri *et al.*, 2010) and a carbon monoxide-bound form of the oxygen sensor FixL from 1 µs to 10 ms with a 500 ns time resolution (Key *et al.*, 2007).

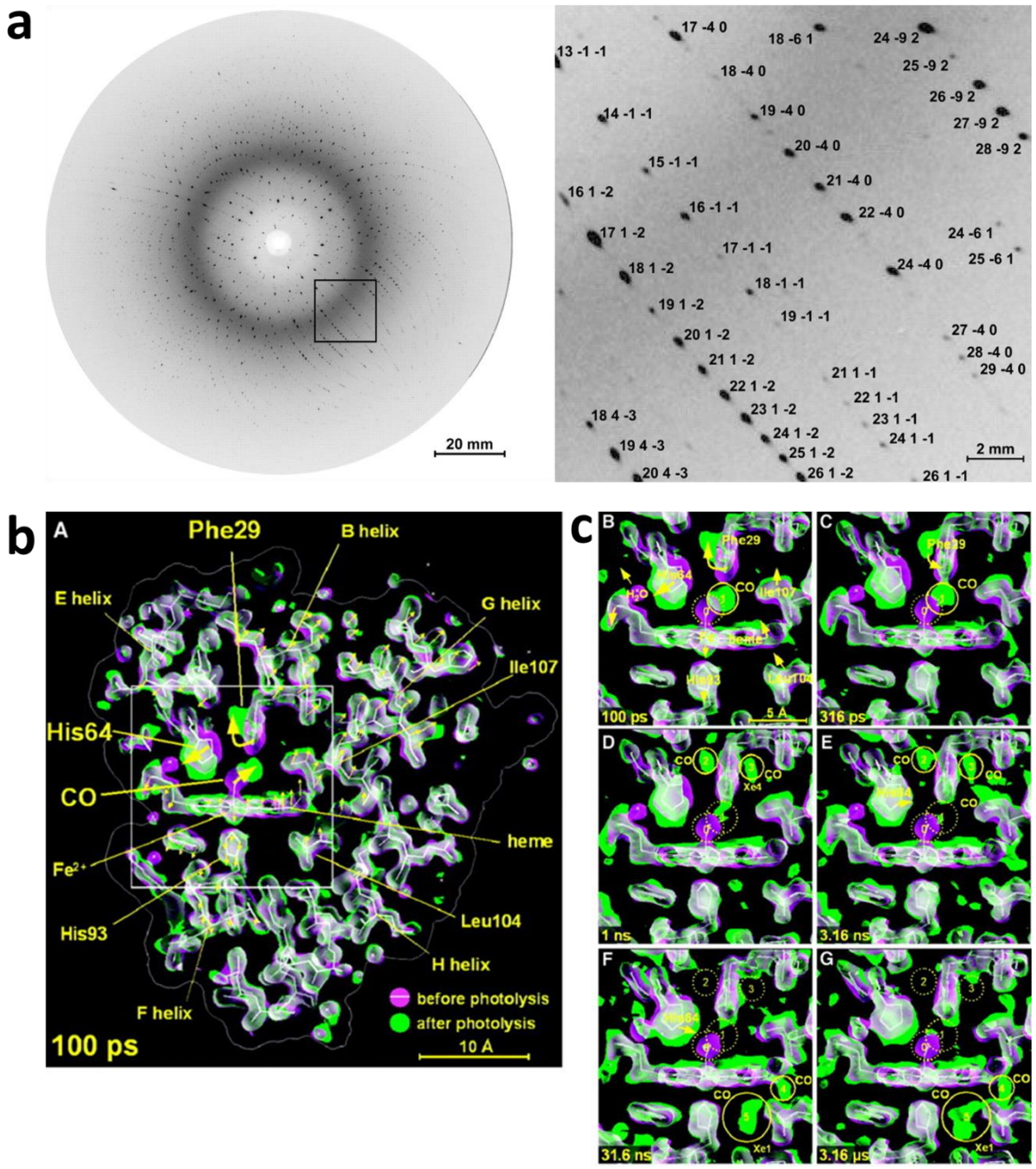


Figure 1.2: Example of a time-resolved Laue diffraction experiment on carboxy-myoglobin. (**a**) Laue diffraction pattern. (**b**) Electron density maps before (magenta, white) and 100 ps after light excitation (green, white) on the whole myoglobin molecule. (**c**) Time series of electron density maps from 100 ps to 3.16 µs after light excitation (adapted from Schotte *et al.*, 2003).

While this type of time-resolved Laue diffraction experiment has proved less successful with the absence of recent published results, it has fostered new developments in serial crystallography at synchrotrons, where a low-background micro-patterned silicon chip was used to put microcrystals in the pink beam of the BioCARS beamline at APS (Argonne,

Illinois) (Meents *et al.*, 2017). Each microcrystal saw one 100 ps X-ray pulse resulting in 100 to 1000 hits per type of sample. Eventually, only 40 to 200 crystals were necessary to result in complete data sets for each of the 5 samples

Laue diffraction has been the pioneering technique in time-resolved crystallography and helped to obtain impressive results. In order to be able to work on systems with less constraints (crystal quality, repeatability of reaction), monochromatic approaches could now be used, initially at slower time scales (µs-ms *vs.* ps). However, the advent of X-ray Free Electron Lasers suitable for Structural Biology experiments in 2010 (Chapman *et al.*, 2011; Seibert *et al.*, 2011) has opened the door for time-resolved crystallography experiments thanks to their monochromatic very short (ten to hundreds of fs) X-ray pulses.

1.1.3 Time-resolved crystallography at X-ray Free Electron Lasers (XFELs)

1.1.3a Principle of an XFEL

In an XFEL, electrons are linearly accelerated in radiofrequency cavities. Once relativistic, the passing through successive undulators results in a process known as self-amplified spontaneous emission (SASE) which generates laser-like pulses of light in the X-ray wavelength range. XFELs thus produce extremely bright (9 orders of magnitude higher brilliance than third-generation synchrotrons) ultra-short (ten to hundreds of fs) coherent X-rays pulses (Schlichting, 2015).

1.1.3b Microcrystals as preferred samples

While a few experiments have been performed on large protein crystals, the preferred samples for XFEL protein crystallography are millions of microcrystals. This arises from the fact that when a crystal is hit by the very intense X-ray pulse, it is almost instantaneously destroyed (though after producing an X-ray pattern, hence the technique is called “diffraction-before-destruction” (Neutze *et al.*, 2000)) and so the sample needs to be quickly replaced for the next X-ray pulse. As a consequence, a large variety of crystal injectors have been devised (see thereafter) and microcrystals are the most amenable to steady, stable injection through the beam. One can also add that for time-resolved pump-probe experiments, the sample absorbs at the wavelength of the pumping laser and the crystal should not be too thick in order to ensure

the most homogenous photo-conversion of molecules throughout the crystal. One advantage of microcrystals is that they often possess less impurities or geometrical disorder than bigger ones, due to the reduced growing time. However, obtaining good microcrystals may be as difficult as obtaining large crystals, and crystallogenesis is also a bottleneck for this type of diffraction experiments. As a consequence, an efficient control of crystal growth parameters is required (for instance through crystallization condition screening and micro-seeding, see **Chapter 2**).

1.1.3c Injection of microcrystals through the X-ray beam

Crystal delivery to the X-ray beam is a key step in Serial Femtosecond Crystallography (SFX) and has mostly relied on the use of injection devices. The goal is to have the crystals being carried in a medium (originally liquid, then viscous) which then has to be injected as a continuous, steady flow that crosses the X-ray beam. Ideally, each X-ray pulse should meet a crystal and result in a diffraction pattern. However, most images do not contain diffraction patterns, and the number of images with a diffraction patter over the total number of images is called the 'hit rate'. Various types of injectors have been developed in order to optimize the hit rate, which usually has a direct influence on the quality of the resulting data set.

The specificities of injectors mainly consist in the nature of the injected medium, the jet speed, and the way of stabilizing the flow. The gas dynamic virtual nozzle (GDVN) was used in the early years of SFX (DePonte *et al.*, 2008). The GDVN ideally generates a periodic straight-line stream of monodisperse droplets, whose flow is focused thanks to the dynamic forces of a coaxially co-flowing gas. The GDVN can be used with a microfluidics mixer for millisecond reactant and crystal combination in a mixing-based Time-Resolved-Serial Femtosecond Crystallography (TR-SFX) (Calvey *et al.*, 2016). Due to the high flow rate, the significant disadvantage of the GDVN is the high sample consumption.

The development of nanoflow electrospinning SFX has led to a decrease of 60 to 100-fold in sample consumption (Sierra *et al.*, 2012). In this method, the flow rate is decreased thanks to a spinning electric field which serves to focus the jet. This has been used in the first SFX study on photosystem II (Kern *et al.*, 2012). One limitation of this method is that it requires addition of either glycerol or PEG to extend the length of the jet, which increases the background on diffraction images and may affect crystal quality.

More recently, the GDVN was improved into a DFFN (double-flow focusing nozzle), which consists in a double flow-focused liquid injector, in which crystals are injected directly in their crystallization buffer into a sheath jet formed by ethanol and accelerated by a helium gas stream (Oberthuer *et al.*, 2017). This new approach is well suited for TR-SFX *via* mixing, and the sample consumption is decreased up to eight-fold compared to the GDVN.

To reduce drastically the consumption of protein sample by liquid injectors, viscous medium injectors have been progressively developed. The decrease in sample consumption is essentially due to the much lower speed of the generated jet. However, they also come with several issues that need to be considered, such as a higher background scattering by the medium, and their effect on crystal integrity (*e.g.* dissolving or disordering). The first such used media were either a mineral oil-based or a synthetic grease (SuperLube, Synco Chemical Co) (Sugahara *et al.*, 2015). The oil-sensitivity problem for some proteins, inherent to grease matrices, was overcome using a grease-free water-based hyaluronic acid matrix (Sugahara *et al.*, 2016), using hydrogels (Kovácsová *et al.*, 2017) or using agarose (Conrad *et al.*, 2015). When crystals have grown in a lipid cubic mesophase environment (*in meso* crystallization (Caffrey, 2015)), it appears advantageous to use the same lipid as carrying medium to minimize mechanical stress on the crystals due to mixing with other media (Nogly *et al.*, 2016).

Finally, an innovative injection technique has been developed on the basis of acoustic wave generation, allowing for the possibility of approaching a 100% hit rate. Roessler *et al.* have developed injectors that focus sound waves on crystallisation drops sitting on microplates, leading to the ejection of picolitre to nanolitre crystal-containing droplets in the direction of the X-ray beam (Roessler *et al.*, 2016). The key to a high hit rate resides in the synchronization of the ejection with the X-ray pulse scheme. As such, this injection technique should easily find application in TR-SFX.

1.1.3d Fixed-target data collections

While most of MX experiments at XFELs have relied on the use of injectors, which consume much sample, the possibility of presenting crystals at various positions on an X, Y translatable sample holder has been investigated to greatly minimize sample consumption. Because the sample is immobile when hit by the X-ray beam, this mode of operation is called 'fixed-target'. Fixed-target sample holders include ultra-thin silicon nitride membranes (Hunter *et al.*, 2014) and micro-patterned silicon chips (Roedig *et al.*, 2015). The latter has been used

in combination with the high-speed goniometer Roadrunner, leading to the SFX structure determination of two different viruses (Roedig *et al.*, 2017).

1.1.3e Diffraction data analysis

Ideally, a microcrystal will result in a single diffraction pattern on a single image. However, this never holds true due to low hit rates, and useful diffraction images have to be separated from empty ones. This has to be done by automated software such as *NanoPeakCell* (Coquelle *et al.*, 2015). Data are then integrated with *CrystFEL* (White *et al.*, 2012) and a data set is constructed using a Monte Carlo algorithm (Kirian *et al.*, 2010). Data reduction is performed with *Cheetah* (Barty *et al.*, 2014) for instance. In order to obtain a good quality data set, tens of thousands of images need to be merged, although this number has decreased to several thousand with the recent improvement in data analysis software.

1.1.3f Examples

The time structure of the XFEL X-ray beam immediately raised the possibility of performing pump-probe TR-SFX on very fast time scales. After initial determination of static structures (Schlichting, 2015), a number of intermediate structures, or several of them, composing genuine molecular movies have been obtained by TR-SFX from Photosystem II (PSII) (Kupitz *et al.*, 2014), Photoactive Yellow Protein (PYP) (Tenboer, 2015; Pande *et al.*, 2016), carboxymyoglobin (Barends *et al.*, 2015), bacteriorhodopsin (BR) (**Fig. 1.3**) (Nango *et al.*, 2016; Nogly *et al.*, 2018) and the reversibly photoswitchable fluorescent protein rsEGFP2 (Coquelle *et al.*, 2018). Other recent publications of static structures suggest that TR-SFX has already been attempted on various other systems, like photosystem I (PSI) (Chapman *et al.*, 2011) and the chromophore-binding domain of a bacterial phytochrome (from *Deinococcus radiodurans* – *Dr*CBD) (Edlund *et al.*, 2016) and that TR studies should follow.

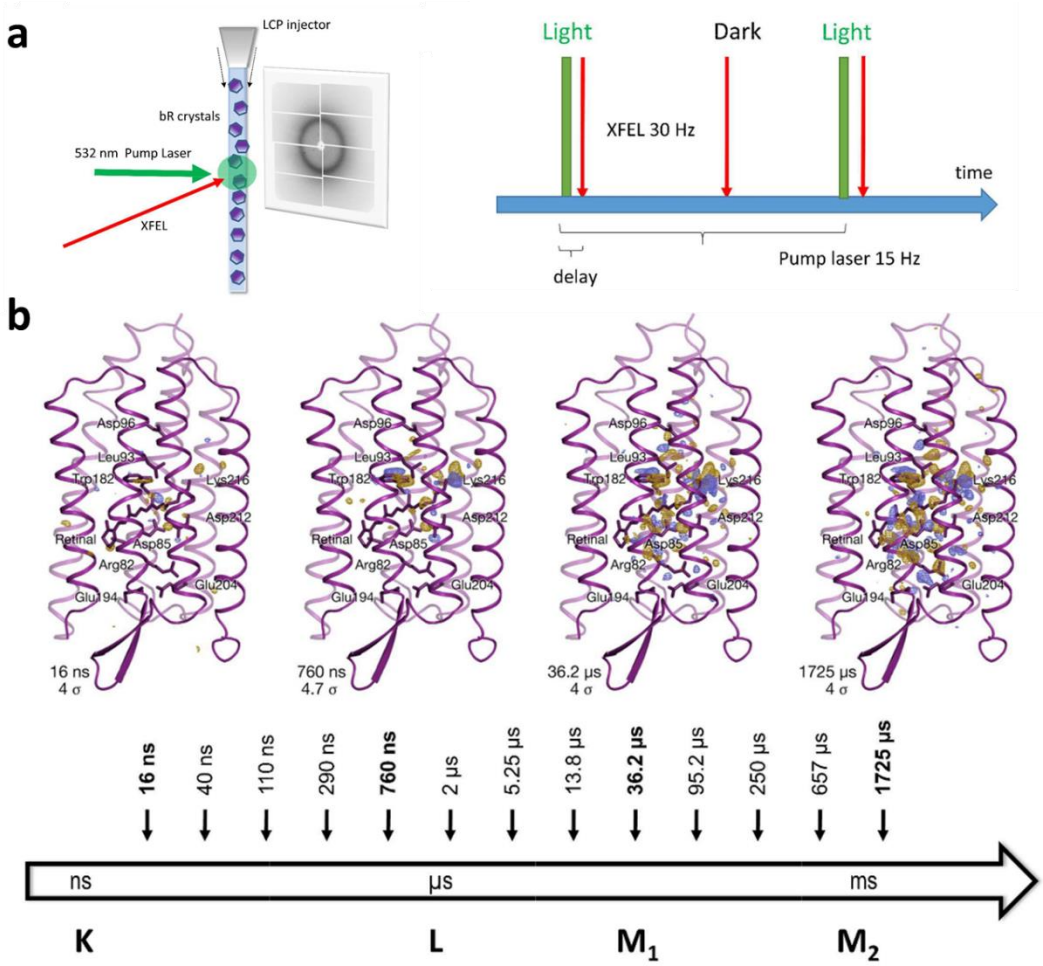


Figure 1.3: Molecular movie of bacteriorhodopsin recorded by SFX. (**a**) Principle of pump-probe experiment on a flow of bacteriorhodopsin crystals (**b**) Fourier difference maps between the dark state and four time points spanning over 5 orders of magnitude (Adapted from Nango *et al.*, 2016).

The development at XFELs of sample injection devices and data analysis software has fostered their implementation at synchrotrons, paving the way for serial crystallography on microcrystals, potentially in a time-resolved modality, yet at lower time resolution, obviously. Due to the high demand on beamtime at XFELs, the relatively low numbers of beamlines per XFEL and the still relatively high level of sample consumption, synchrotron beamlines should prove cheaper and more available alternatives for a fraction of time-resolved crystallography experiments.

1.2 Time-resolved protein crystallography on monochromatic synchrotron beamlines?

The time resolution and intensity of the X-ray beam at an XFEL cannot be matched at synchrotrons. However, there are many biological systems that can be studied at slower time scales (**Fig. 1.4**), including protein domain motions. For instance, it has been estimated that the average rate constant of an average enzyme (out of thousands of various enzymes) is k_{cat} = 10 s^{-1}, giving 100 ms as the minimal duration of a catalysis cycle (Bar-Even *et al.*, 2011).

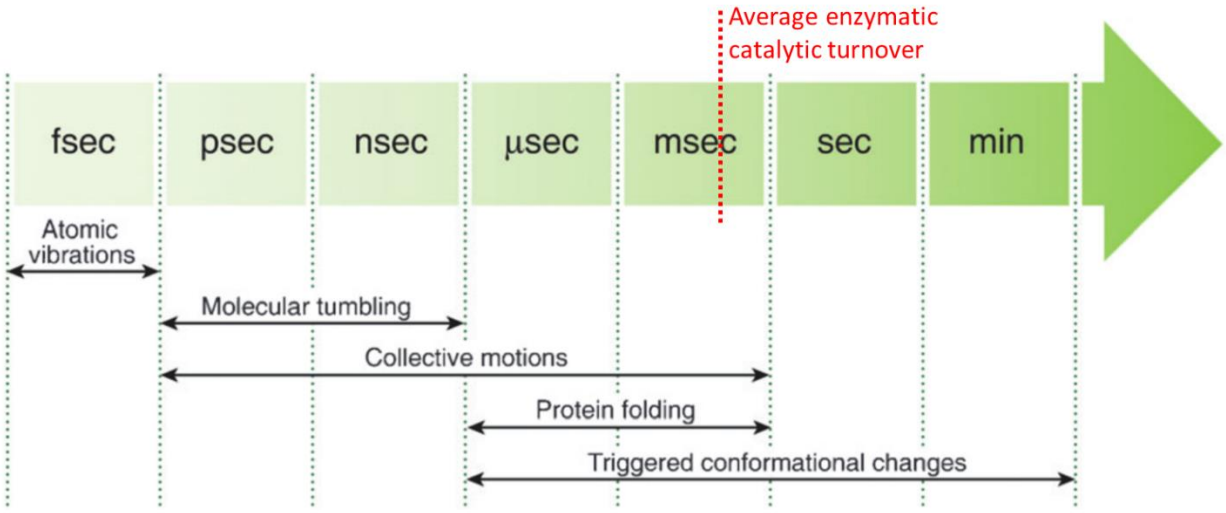


Figure 1.4: Time domains of various processes composing protein dynamics (Adapted from Ben-Nissan & Sharon, 2011).

The development of protein crystallography experiments at XFELs has spurred synchrotrons to develop and test new modes of presenting samples to the X-ray beam. While cryogenic data collections were almost exclusively performed on crystals flash-cooled on nylon/mylar loops, people have adapted crystal injection systems (usually using a viscous medium) or translatable solid supports (usually chips). They have tweaked their crystallogenesis recipes to be able to produce microcrystals rather than 50 to 500 µm large crystals. Finally, they also borrowed data analysis software. All this has eventually led to the development of time-resolved crystallography experiments, which has been coined SSX (Serial Synchrotron Crystallography) or SMX (Serial Millisecond Crystallography).

1.2.1 Injection devices

Initial experiments have been performed at PETRA III (Hamburg, Germany) on lysozyme microcrystals injected in a 100 µm diameter capillary (Stellato *et al.*, 2014). Scattering from the capillary makes it a non-optimal solution, and several other injectors have been developed, using a viscous medium in order to ensure a low flow rate, thus maximizing the number of crystals being intercepted by the X-ray beam. The carrying medium can be a lipidic cubic phase (Nogly *et al.*, 2015) (**Fig. 1.5**), high-molecular-weight poly(ethylene oxide) (Martin-Garcia *et al.*, 2017) or any high viscosity medium (Botha *et al.*, 2015). A version of the High Viscosity Extrusion sample injector described in Botha *et al.* (2015) has been installed on ID30A-3 of the ESRF.

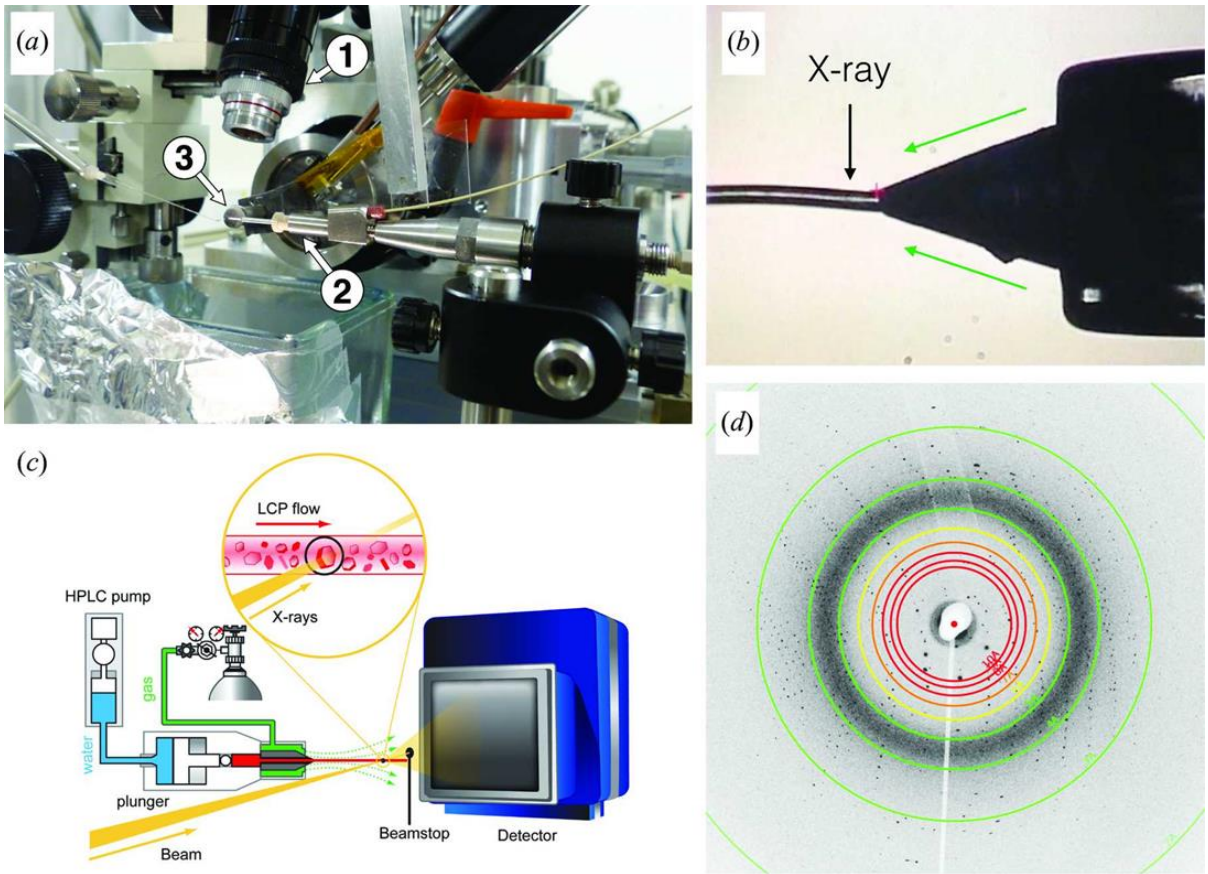


Figure 1.5: LCP injector installed on beamline ID13 of the ESRF. (**a**) Overall view of the experimental setup. (**b**) Close-up of the nozzle. (**c**) Schematics of the injection stage. (**d**) Diffraction pattern showing the contribution of the LCP to the scattered background (dark ring). (Reproduced from Nogly *et al.*, 2015).

1.2.2 Fixed-target sample holders

The idea of using chips for room temperature serial crystallography experiments at synchrotrons has been introduced shortly after the first XFEL experiments and was tested at the SLS synchrotron (Villigen, Switzerland) (Zarrine-Afsar *et al.*, 2012). Clear advantages of chips compared to injectors are that they require lesser quantities of protein sample and that the hit rate could potentially reach 100%. Many chips have been developed since then, in various materials and shapes. The material needs to scatter X-rays as little as possible in order to minimize background, hence the parts holding crystals are usually ultra-thin. The most common material is silicon nitride, but can be graphene as well (Sui *et al.*, 2016). There are two main types of chips, the grid-like one, formed of a mesh or of micro-patterned holes (Roedig *et al.*, 2015, 2016), and the microfluidics-based one. The latter can be used for *in-situ* crystallization and for mixing with substrate shortly before exposure to the X-rays (Lyubimov *et al.*, 2015; Heymann *et al.*, 2014; Khvostichenko *et al.*, 2014; Sui *et al.*, 2016).

A particularly interesting chip for time-resolved crystallography is that described in (Oghbaey *et al.*, 2016), as a prior spectroscopic mapping allows for the identification of useable grid points, which can eventually lead to a 100% hit rate. In addition, the optical setup could be used as an actinic source in a pump-probe experiment.

1.2.3 First examples of dose-resolved, or time-resolved experiments at synchrotrons

One of the first time-resolved serial crystallography experiments performed on beamline 14-ID-B at the APS synchrotron was done with a polychromatic beam on a microfluidics chip in which crystals of photoactive yellow protein were grown (Pawate *et al.*, 2015). The crystals are immobile within the chip, which is directly mounted on the beamline goniometer. Data were collected for the dark state, 10 µs and 1 ms after actinic excitation of a 485 nm laser bringing 4.8 $mJ.mm^{-2}$ energy density in a 5 ns pulse. Because of the Laue diffraction technique used, only 10 crystals were necessary to record 21 or 22 images for each time point, which were sufficient to constitute 74% complete data sets at 1.8 Å resolution with a multiplicity of 2.4. The resulting difference maps unambiguously show the progressive rearrangement of the chromophore *p*-coumaric acid.

A proof-of-principle fixed-target experiment has been performed on 46 thaumatin crystals sandwiched in-between two Kapton foils on beamline P14 at PETRA III (Schubert *et*

al., 2016). 20 data sets were reconstituted from merging together the n^{th} frames of each of the 46 20-frame oscillation data sets (frame duration: 40 ms, total data set duration: 800 ms) for an attenuated X-ray beam (low-dose series, total absorbed dose 230 kGy), and for a 10-times stronger X-ray beam (high-dose series, total absorbed dose 2.32 MGy). The low-dose series hardly shows damage to the 8 disulfide bonds of the protein at dose point #20, the high-dose series shows very obvious damage after 200 ms. While this does not exactly qualify as a time-resolved experiment but rather as a dose-resolved experiment, this experiment has showed the feasibility of looking at a structural transformation by X-ray crystallography on the 40 ms time resolution.

An injection method based on the deposition of a stream of crystals on a translating polyimide tape has allowed the demonstration of mix-and-diffuse serial crystallography on beamline P11 at PETRA III (Beyerlein *et al.*, 2017). Ahead of the deposition tip of the microfluidics injection system, a T-shaped connection enables mixing of the crystal suspension with a solution of the protein inhibitor chitotriose. Two ways of mixing resulted in two time points being investigated: the short-mixing time point of 2 s, and the long-mixing time point of 50 s. The differences in the inhibitor-bound structure for these two time points resides in the higher number of water molecules connecting the inhibitor to the protein at 2 s (11 water molecules) than at 50 s (4 molecules).

A fixed-target approach where the delay between the laser pump pulse and the X-ray probe pulse is mechanically controlled by the translation time from well to well, which are separated by fixed distances, has led to the visualization of various intermediates states of fluoroacetate halogenase binding and reaction with its fluoroacetate substrate, which has been rapidly generated by femtosecond laser-induced fragmentation of a photocaged fluoroacetate, at time points between 30 ms and 2,052 ms (Schulz *et al.*, 2018) (**Fig. 1.6**). The particular method has been called HARE, for ‘Hit-And-REturn’.

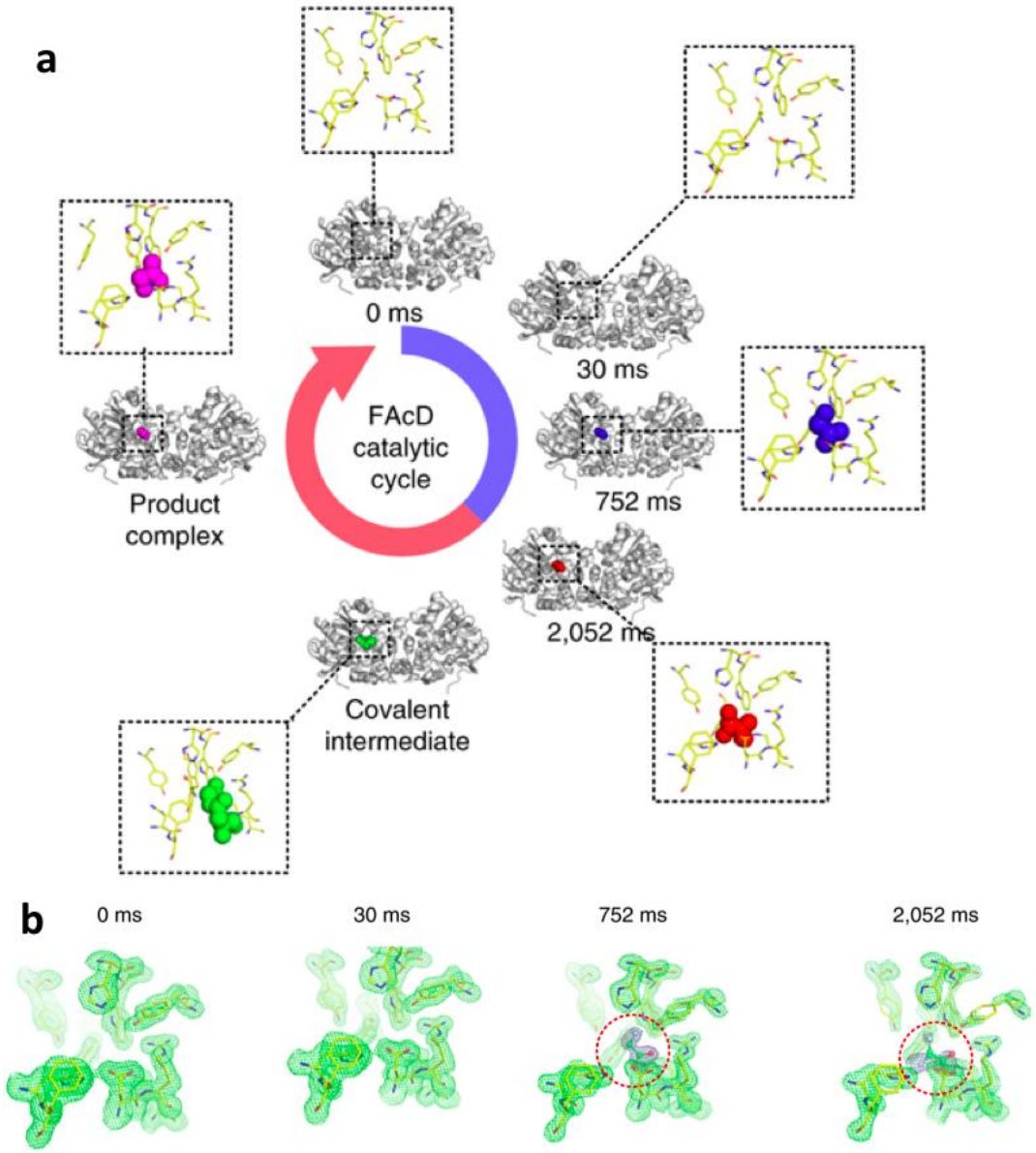


Figure 1.6: The fixed-target-based HARE approach has led to the visualisation of various binding and reaction intermediate states of fluoroacetate dehalogenase. (**a**) Description of the catalytic cycle. **(b)** Electron density in the active site of molecule A of the crystallographic dimer at various time points. (Reproduced from Schulz *et al.*, 2018).

A recent manuscript deposited in bioRxiv describes a time-resolved crystallography study on the photocycle of bacteriorhodopsin from 0 to 200 ms with a 5 ms time resolution (Weinert *et al.*, 2019), using a high-viscosity injector (Weierstall *et al.*, 2014) and the 5 ms pulse of a diode laser bringing 37.43 W/cm^2 at the sample position, performed at the Swiss Light Source synchrotron (Villigen, Switzerland). The electron density difference map calculated between the dark data set and the [10-15 ms] data set gave evidence of the occurrence of the helix movements expected in the second-half of the photocycle.

1.3 TR-MX – on which biological systems?

Investigating the detailed mechanism of a protein requires synchronization of the beginning of the reaction, which can be initiated either by light activation for light-sensitive proteins or by letting substrate molecules diffuse within crystals for other systems, most importantly enzymes, which are almost never activated by light. Only three enzymes are light activated: DNA photolyase (Sancar, 2016), protochlorophyllide oxidoreductase (Gabruk & Mysliwa-Kurdziel, 2015) and the newly-discovered fatty acid photodecarboxylase (Sorigué et al., 2017). Light is the simplest way of synchronizing a reaction, since enzymatic reactions are limited by the substrate diffusion rate *in crystallo*. As a consequence, light-sensitive proteins are the easiest systems on which to develop time-resolved crystallography methods. The blue light and red light/far-red light photoreceptors phototropin (**Paragraph 1.4**) and phytochrome (**Paragraph 1.6**) offer opportunities to study how light-absorption by a chromophore encased in a protein can lead to a cascade of structural changes eventually fulfilling the photosensory function. Fluorescent proteins however are expected not to exhibit much structural change upon light irradiation since the fluorescence properties of their chromophores are favoured by a rigid environment preventing non-radiative decay of their excited states. This being said, fluorescent proteins are imperfect, and their chromophore may undergo photo-isomerisation, which is sometimes detrimental to their function (photobleaching, *i.e.* reversible or irreversible loss of fluorescence signal) but can also be taken as an advantage for use in super-resolution microscopy (photochromism, photoswitching, photoactivation).

1.4 The LOV2 domain of phototropin 2 from *Arabidopsis thaliana*

1.4.1 The blue light plant photoreceptor phototropin

Phototropin is a plant photoreceptor sensing blue light *via* FMN chromophores and mediating various processes, including phototropism, that lead to optimizing of photosynthesis upon variations in light intensity (Christie, 2007). Phototropin acts as a light-activated protein kinase, whose autophosphorylation triggers the signalling cascade (**Fig. 1.7**). There are two types of phototropins in higher plants, Phot1 and Phot2. Phot1 is the primary photoreceptor implicated in phototropism in low-intensity blue light conditions, and Phot1 and Phot2 function redundantly in medium to high blue light conditions (Liscum *et al.*, 2014). Both Phot1 and Phot2 contains two distinct LOV domains in the N-terminal half of the protein, LOV1 and LOV2, with the serine/threonine kinase domain at the end of the C-terminal half. Both halves are separated by the J_{α} helix which disorders upon light activation, liberating the kinase domain.

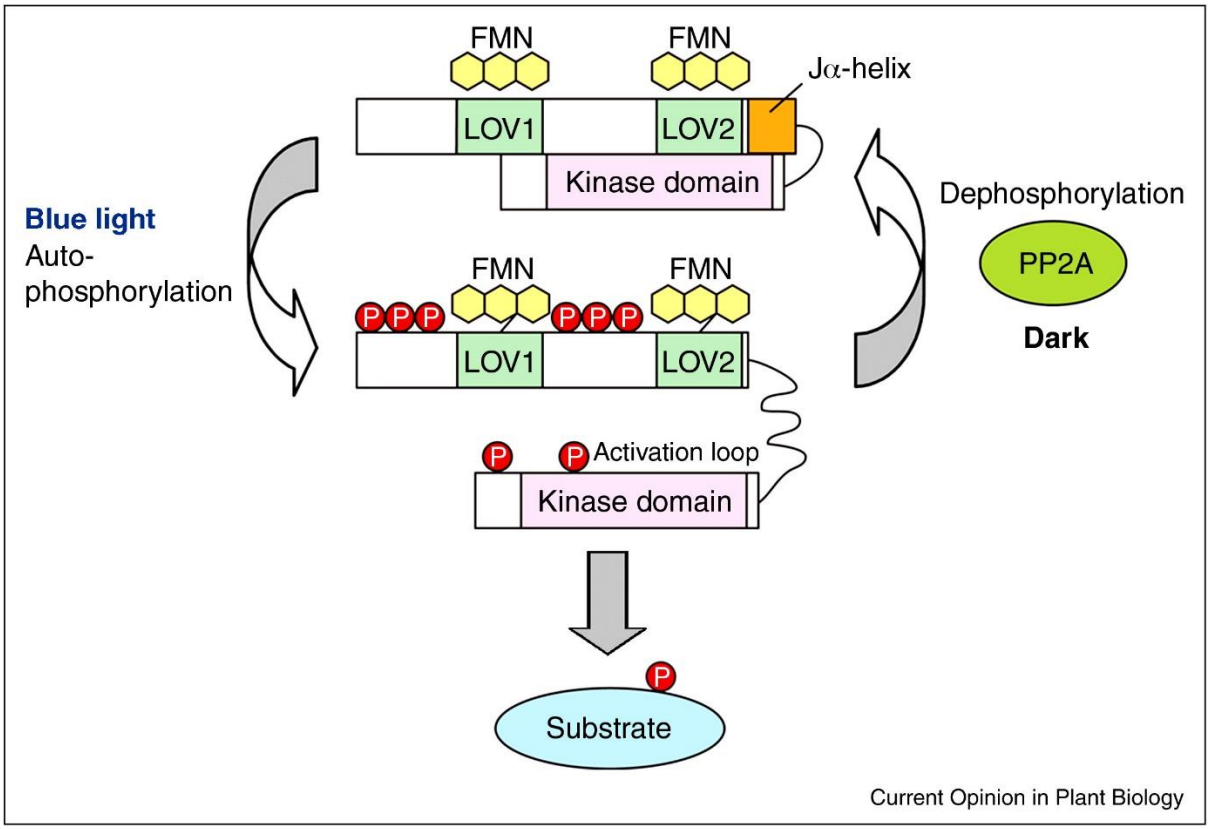


Figure 1.7: Schematic of light-induced activation of phototropin activity (reproduced from Inoue *et al.*, 2010).

1.4.2 The LOV domains from phototropin

The LOV domains belong to the superfamily of PAS (Per, ARNT, Sim) domains that are associated with bonding of cofactors and that mediate protein interactions. More specifically, they are regulated by **L**ight, **O**xygen or **V**oltage, which gives their name. They bind non-covalently to the FMN chromophore, which absorbs light maximally around 447 nm in darkness (hence, the ground state, or 'dark state'), *i.e*. in the blue part of the visible spectrum. Upon absorption of a blue light photon, there is formation of a covalent bond between the C_{4a} atom of the FMN and the S_{γ} atom of a conserved cysteine residue. The species (the signalling state, or 'light state') thus formed in microseconds absorbs light maximally at 390 nm and is long-lived as it returns to the dark state in tens to hundreds of seconds (**Fig. 1.8**).

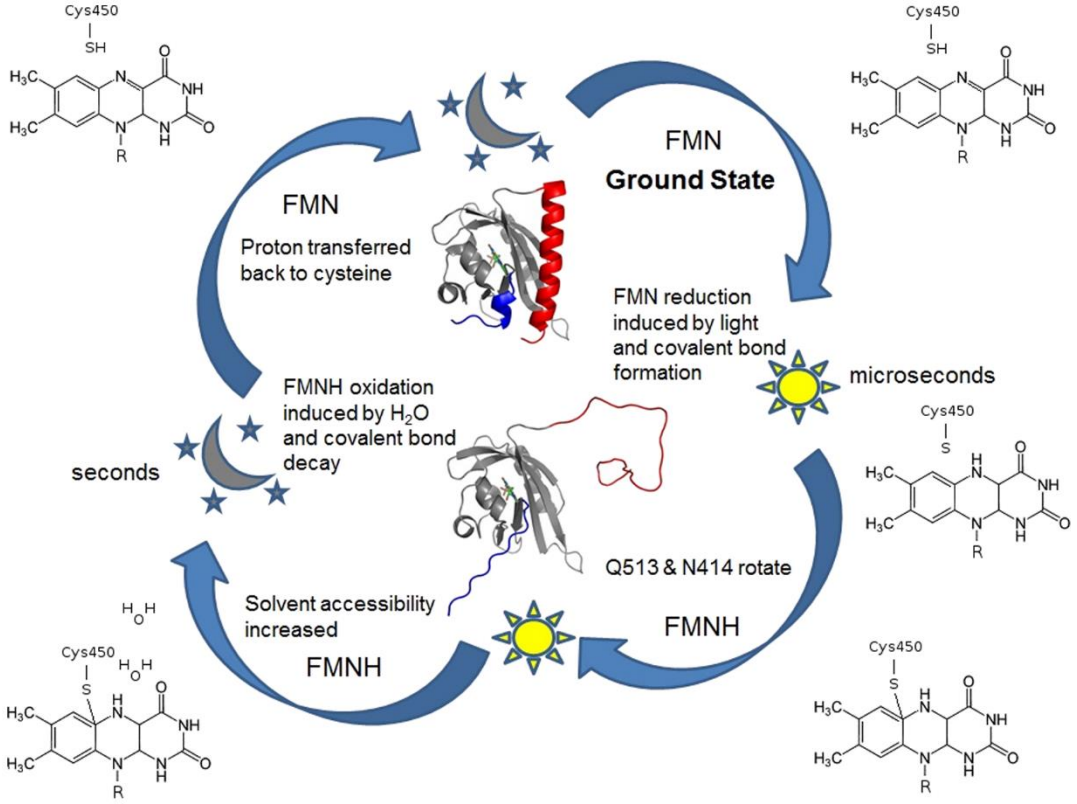


Figure 1.8: Schematic of the photocycle of a LOV2 domain of phototropin from *Avena sativa* (reproduced from Zayner & Sosnick, 2014).

The LOV1 and LOV2 domains, albeit structurally very similar, function differently (Christie, 2007). They exhibit distinct photoreaction quantum efficiencies and kinetic parameters. LOV2 has, in general, a higher efficiency than LOV1 from the same phototropin, with the ratio varying between 2 and 10. Moreover, the decay of the light state is faster for

LOV2 than LOV1 in Phot2, while slower in Phot1 (Kasahara *et al.*, 2002). LOV2 appears to play the main role in phototropin activity regulation, while the function of LOV1 is not clear, and may be of helping receptor dimerization or extend the duration of the light state.

1.4.3 Crystallographic studies on the photocycle of LOV domains

Due to the peculiarity of the light-catalysed reaction catalysed by LOV domains, *i.e.* the formation of a covalent bond, they have attracted interest in kinetic crystallography experiments destined to understand the molecular mechanism. Two pioneering projects have been performed on LOV1 and LOV2, by the groups of Ilme Schliching and Keith Moffat, respectively. The first structure of a LOV domain, that of LOV2 from the chimeric fern photoreceptor phy3, was determined at 2.7 Å resolution at room temperature (Crosson & Moffat, 2001). In the follow-up work, the structure of the photoexcited LOV2 was determined at 2.6 Å resolution under photostationary conditions at room temperature, which unambiguously showed the formation of the adduct (**Fig. 1.9a**) (Crosson & Moffat, 2002).

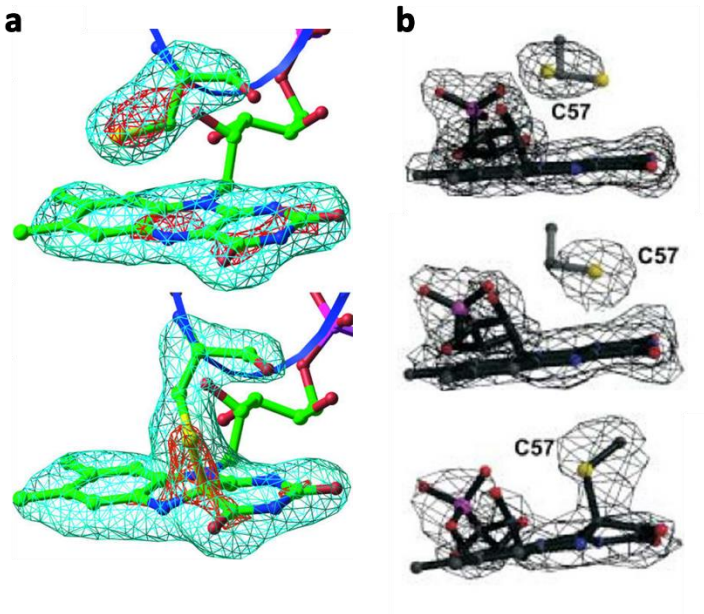


Figure 1.9: First structure determination of LOV domains from photoreceptors. (**a**) Chromophore region for the dark and light states of fern phy3LOV2. (**b**) Top to bottom: chromophore region for the dark, light (X-ray damaged) and light (composite data set) states of the LOV2 domain of Phot1LOV1 from *Chlamydomonas reinhardtii* (adapted from Crosson & Moffat, 2002 and Fedorov *et al.*, 2003).

In contrast, the various structures of the LOV1 domain of Phot1 from *Chlamydomonas reinhardtii* were determined at cryogenic temperature (Fedorov *et al.*, 2003). This study

revealed that the covalent bond between the FMN and the cysteine residue is very sensitive to X-rays, and required a composite data set approach to determine its undamaged structure (**Fig. 1.9b**).

1.4.4 Biotechnological applications of LOV domains

LOV domains encase a FMN molecule in a rather rigid environment, favouring green fluorescence upon blue light excitation (Christie, 2007). Mutation to an alanine of the reactive cysteine in the LOV2 domain of Phot2 from *A. thaliana* abolishes the photoreaction, thus improving the fluorescence quantum yield. Further engineering led to the design of iLOV, a small, genetically-encoded green fluorescent protein (Chapman *et al.*, 2008), whose structure was later determined at 1.8 Å resolution (Christie *et al.*, 2012).

Following up on the same idea, Shu and collaborators engineered the same LOV2 domain into a genetically encoded tag with two modalities (**Fig. 1.10**): (1) green fluorescence, and (2) capacity of generating singlet oxygen upon light excitation (Shu *et al.*, 2011). One notable difference is the mutation of the cysteine into a glycine. The two functions of the resulting protein, miniSOG for **mini**mal-sized **S**inglet **O**xygen **G**enerator, makes it an attractive tool for Correlated Light and Electron Microscopy experiments, allowing high resolution protein localization.

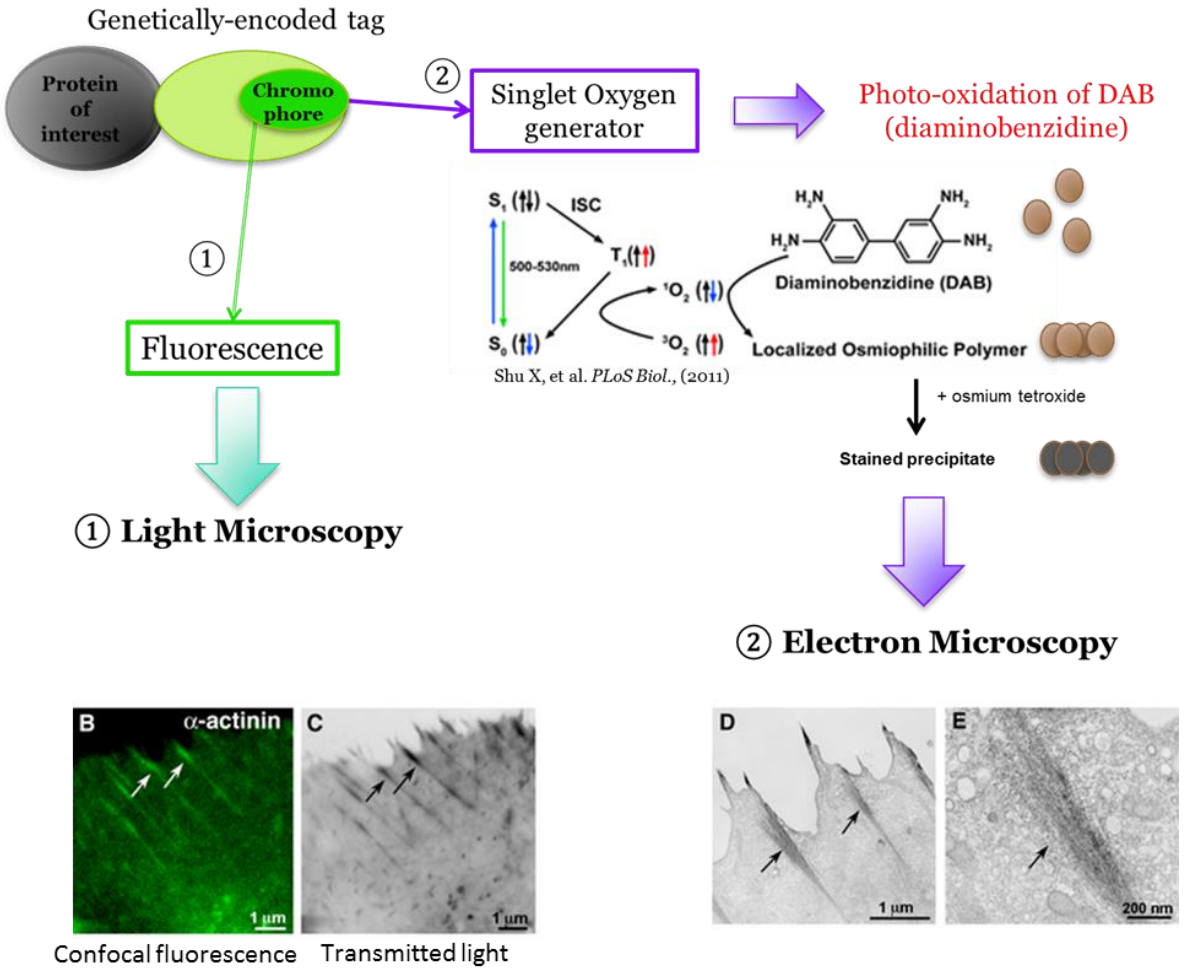


Figure 1.10: Principle of a CLEM experiment using miniSOG. Its fluorescence modality allows the identification of low-resolution features in two types of light microscopy (**1**). Irradiation of miniSOG leads to the generation of singlet oxygen which will induce the localized polymerization of DAB molecules (added to the cell medium beforehand), whose precipitates are then labelled with osmium tetroxide. Electron microscopy (**2**) is then used for high resolution localization of the osmium atoms (adapted from Shu *et al.*, 2011).

1.5 Fluorescent proteins and chromoproteins

Green Fluorescent Protein (GFP) has been first identified and isolated from the jellyfish *Aequorea victoria* in 1962 (Shimomura *et al.*, 1962). Its cloning in 1992 (Prasher *et al.*, 1992) has revolutionized the technique of fluorescence imaging by providing a genetically-encoded fluorophore to cells. Other homologues were later found in marine organisms (Matz *et al.*, 1999), some with different colours of emission, and some which are non-fluorescent, and are then called 'chromoproteins'.

The fluorescent chromophore of GFP is formed from the autocatalytic cyclisation of three consecutive amino-acid residues of its sequence Ser65-Tyr66-Gly67, which is located on an α-helix traversing an eleven-strand β-barrel. Mutation of Tyr66 into a histidine or a tryptophan, and additional rounds of mutagenesis, leads to proteins emitting blue (BFPs) or cyan (CFPs) fluorescence. One such CFP, Cerulean, has become a popular energy donor in Förster resonance energy transfer (FRET) experiments used to report protein interactions or biochemical processes within living cells (Rizzo *et al.*, 2004). A point-mutant of Cerulean, Twist Cerulean, has been shown to isomerize upon a pH jump on the sub-second time scale, in stark contrast with Cerulean, for which the pH-induced isomerisation process occurs on the hour time scale (Hélène Pasquier & Antoine Royant, data not shown). The structure of Cerulean was solved in the group of Antoine Royant and became a model protein for crystallogenesis and structural analysis (Lelimousin *et al.*, 2009; von Stetten *et al.*, 2012; Gotthard *et al.*, 2017).

Many Red Fluorescent Proteins (RFPs) have been developed after the discovery of the naturally-occurring tetrameric RFP *Ds*Red from a species of *Discosoma* (a family of sea anemones). However, these RFPs suffer from a lack of brightness, and a thorough evolution process starting from a consensus synthetic template (constituted after alignment of RFP sequences) led to the development of mScarlet, the brightest RFP in its class (Bindels *et al.*, 2017).

1.6 *Dr*CBD: The chromophore-binding domain of the phytochrome from *Deinococcus radiodurans*

Phytochromes are red/far-red light photoreceptors using bilins as chromophores, which are widely found in fungi and plants, but also in bacteria, in which they mediate phototaxis and pigmentation. The phytochrome of *Deinococcus radiodurans* (*Dr*BphP) absorbs light through a biliverdin chromophore, and has attracted much attention from structural biologists, which led to the first structure determination of a domain of a phytochrome, the chromophore-binding domain *Dr*CBD (Wagner *et al.*, 2005). As such, and because a single-point mutant of *Dr*CBD showed weak red fluorescence (Wagner *et al.*, 2008), it then attracted the attention of fluorescent protein-developing scientists, who evolved the first near-infrared fluorescent protein IFP1.4 (Shu *et al.*, 2009), which was crystallized and had its structure solved in the group of Antoine Royant (Feliks *et al.*, 2016). The photocycle of *Dr*BphP (**Fig. 1.11**) shows that the Pfr (absorbing **f**ar-**r**ed light) state builds up on the millisecond time scale from the Pr (absorbing **r**ed light) state, which is most usually the stable form. *Dr*CBD, which lacks the photosensory domain, must exhibit a photocycle with altered kinetics, still with the equivalent of a Pr and a Pfr state. *Dr*CBD crystals could thus be used to study, in a steady-state crystallography approach, the structural effects of red-light induced chromophore isomerisation.

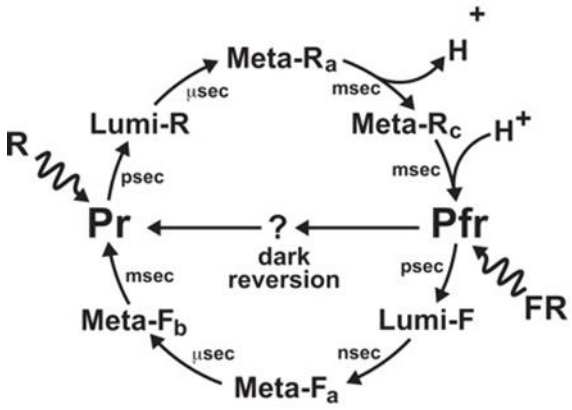


Figure 1.11: Photocycle of *Dr*BphP showing the time scales of various intermediate build-ups (adapted from Wagner *et al.*, 2008).

1.7 Preamble to thesis manuscript

At the beginning of the project, we identified three different proteins (**Fig. 1.12**) on which we could attempt to perform time-resolved crystallography experiments: (**a**) Twist-Cerulean, a single point mutant of the cyan fluorescent protein Cerulean, which had been shown to isomerize on the millisecond to second time scale upon a pH jump (unpublished results), (**b**) *Dr*CBD, the chromophore-binding domain of the phytochrome from *Deinococcus radiodurans*, for which isomerisation of the chromophore leads to an equilibrium between a red-absorbing and a far-red-absorbing form on the millisecond time scale, and (**c**) *At*Phot2LOV2, the LOV2 domain of phototropin 2 from *Arabidopsis thaliana*, which, upon blue light illumination, undergoes formation of a covalent bond between its FMN chromophore and a neighbouring cysteine residue on the microsecond time scale, which relaxes back on the 100-second level.

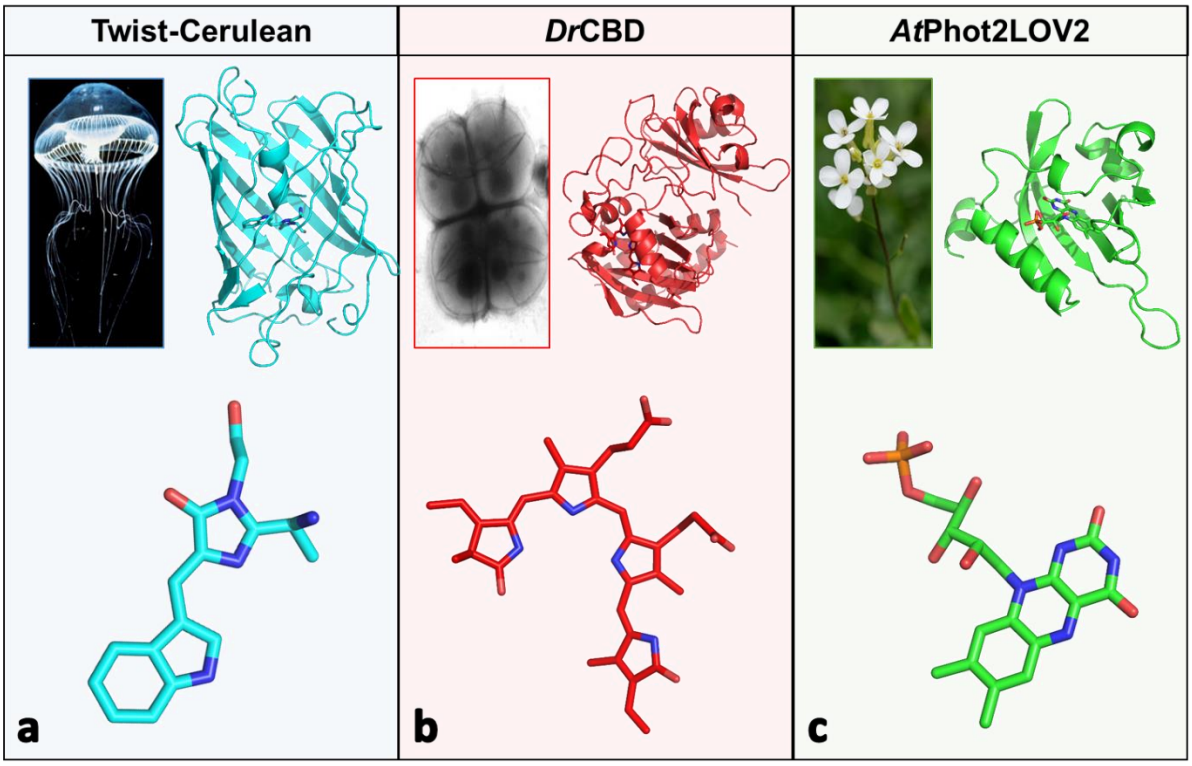


Figure 1.12: Photograph or micrograph of the origin organism, 3-dimensional structure and chromophore structure of the three proteins envisaged in this project for developing time-resolved methods: (**a**) the fluorescent protein Twist-Cerulean, (**b**) *Dr*CBD, the chromophore-binding domain of the *Deinococcus radiodurans* phytochrome and (**c**) *At*Phot2LOV2, the LOV2 domain of phototropin 2 from *Arabidopsis thaliana*.

The first part of my PhD work has consisted of expressing, purifying and learning how to crystallize these three proteins in various sizes and shapes (**Chapter 2**). Protein expression and purification first needed to be optimized, particularly for *At*Phot2LOV2, in order to provide protein sample of sufficient purity to yield well-diffracting crystals. I have then developed strategies to achieve a thorough control of nucleation, so to provide us with reasonably large crystals for our single-crystal based TR-MX experiments, and with microcrystals for injector-based SSX experiments. The crystallogenesis experiments on GFP-like chromoproteins and the mScarlet series that I performed during my Master2 degree and in the early stage of my PhD work are also described there as they constituted milestones in my training.

Envisaging time-resolved experiments led us to investigate how severely specific radiation damage could potentially hamper a structural analysis of putative intermediate states. We thus characterized and compared the build-up rate of both specific and global radiation damage at both cryogenic and room temperature for lysozyme, a common target in radiation damage studies, but also on the fluorescent protein Cerulean and the light state of *At*Phot2LOV2 shown to be particularly radiosensitive at cryogenic temperature. (**Chapter 3**).

Chapter 4 presents our attempts to use and develop instruments and methods for time-resolved crystallography experiments. On one hand, the Structural Biology group was implementing a microcrystal injector on beamline ID30A-3, on which I tried microcrystals of all three of my proteins, which led to several proof-of-principle results (**Chapter 4**, **Paragraph 4.1**). On the other hand, we exploited the faster and faster pixel X-ray detectors to see how fast we could record data sets on single crystals and thus decrease the time resolution of room temperature data collection, from ~30 s to ~1 s (**Chapter 4**, **Paragraph 4.2**).

At one point of the PhD project, I needed to focus on one of the proteins. *Dr*CBD was the least advanced and was thus left aside. The work on Twist-Cerulean was promising, but involved investigation of the use of caged protons in order to generate light-induced pH jump, a potentially time-consuming study. We thus focussed on *At*Phot2LOV2, following the initial results presented in **Chapter 4**, **Paragraph 4.2**.

The TR-MX work on *At*Phot2LOV2 was divided into two projects: the ‘easier’ one, that is the monitoring of the relaxation of its light state on the 100-second level (**Chapter 5**), and the ‘challenging one’, that is the monitoring of the build-up of its light state population in crystals on the millisecond time scale (**Chapter 6**). The results of these experiments are briefly discussed in **Chapter 7**.

Finally, my work on *At*Phot2LOV2 gave me the opportunity to join a project already advanced within the group on the structural elucidation of the mechanism of the photosensitizing protein miniSOG, a 6-point mutant of *At*Phot2LOV2. This recently published study is provided as **Annex 1**.

1.8 References

Bar-Even, A., Noor, E., Savir, Y., Liebermeister, W., Davidi, D., Tawfik, D. S. & Milo, R. (2011). *Biochemistry*. **50**, 4402–4410.

Barends, T. R. M., Foucar, L., Ardevol, A., Nass, K., Aquila, A., Botha, S., Doak, R. B., Falahati, K., Hartmann, E., Hilpert, M., Heinz, M., Hoffmann, M. C., Köfinger, J., Koglin, J. E., Kovacsova, G., Liang, M., Milathianaki, D., Lemke, H. T., Reinstein, J., Roome, C. M., Shoeman, R. L., Williams, G. J., Burghardt, I., Hummer, G., Boutet, S. & Schlichting, I. (2015). *Science*. **350**, 445–450.

Barty, A., Kirian, R. A., Maia, F. R. N. C., Hantke, M., Yoon, C. H., White, T. A. & Chapman, H. (2014). *J. Appl. Cryst.* **47**, 1118–1131.

Ben-Nissan, G. & Sharon, M. (2011). *Chem. Soc. Rev.* **40**, 3627–3637.

Beyerlein, K. R., Dierksmeyer, D., Mariani, V., Kuhn, M., Sarrou, I., Ottaviano, A., Awel, S., Knoska, J., Fuglerud, S., Jönsson, O., Stern, S., Wiedorn, M. O., Yefanov, O., Adriano, L., Bean, R., Burkhardt, A., Fischer, P., Heymann, M., Horke, D. A., Jungnickel, K. E. J., Kovaleva, E., Lorbeer, O., Metz, M., Meyer, J., Morgan, A., Pande, K., Panneerselvam, S., Seuring, C., Tolstikova, A., Lieske, J., Aplin, S., Roessle, M., White, T. A., Chapman, H. N., Meents, A. & Oberthuer, D. (2017). *IUCrJ*. **4**, 769–777.

Bindels, D. S., Haarbosch, L., van Weeren, L., Postma, M., Wiese, K. E., Mastop, M., Aumonier, S., Gotthard, G., Royant, A., Hink, M. A. & Gadella Jr, T. W. J. (2017). *Nat. Methods*. 1–12.

Botha, S., Nass, K., Barends, T. R. M., Kabsch, W., Dworkowski, F., Panepucci, E., Wang, M., Robert, L., Schlichting, I. & Doak, R. B. (2015). *Acta Cryst. D*. **71**, 387–397.

Bourgeois, D. & Royant, A. (2005). *Curr. Opin. Struct. Biol.* **15**, 538–547.

Bourgeois, D., Vallone, B., Schotte, F., Arcovito, A., Miele, A. E., Sciara, G., Wulff, M., Anfinrud, P. & Brunori, M. (2003). *Proc. Natl Acad. Sci. USA*. **100**, 8704–8709.

Bowler, M. W., Mueller, U., Weiss, M. S., Sanchez-Weatherby, J., Sorensen, T. L.-M., Thunnissen, M. M. G. M., Ursby, T., Gobbo, A., Russi, S., Bowler, M. G., Brockhauser, S., Svensson, O. & Cipriani, F. (2015). *Cryst. Growth Des.* **15**, 1043–1054.

Caffrey, M. (2015). *Acta Cryst. F* . **71**, 3–18.

Calvey, G. D., Katz, A. M., Schaffer, C. B. & Pollack, L. (2016). *Struct. Dyn.* **3**,.

Chapman, H. N., Fromme, P., Barty, A., White, T. A., Kirian, R. A., Aquila, A., Hunter, M. S., Schulz, J., Deponte, D. P., Weierstall, U., Doak, R. B., Maia, F. R. N. C., Martin, A. V, Schlichting, I., Lomb, L., Coppola, N., Shoeman, R. L., Epp, S. W., Hartmann, R., Rolles, D., Rudenko, A., Foucar, L., Kimmel, N., Weidenspointner, G., Holl, P., Liang, M., Barthelmess, M., Caleman, C., Boutet, S., Bogan, M. J., Krzywinski, J., Bostedt, C., Bajt, S., Gumprecht, L., Rudek, B., Erk, B., Schmidt, C., Hömke, A., Reich, C., Pietschner, D., Ströder, L., Hauser, G., Gorke, H., Ullrich, J., Herrmann, S., Schaller, G., Schopper, F., Soltau, H., Kühnel, K. U., Messerschmidt, M., Bozek, J. D., Hau-Riege, S. P., Frank, M., Hampton, C. Y., Sierra, R. G., Starodub, D., Williams, G. J., Hajdu, J., Timneanu, N., Seibert, M. M., Andreasson, J., Rocker, A., Jönsson, O., Svenda, M., Stern, S., Nass, K., Andritschke, R., Schröter, C. D., Krasniqi, F., Bott, M., Schmidt, K. E., Wang, X., Grotjohann, I., Holton, J. M., Barends, T. R. M., Neutze, R., Marchesini, S., Fromme, R.,

Schorb, S., Rupp, D., Adolph, M., Gorkhover, T., Andersson, I., Hirsemann, H., Potdevin, G., Graafsma, H., Nilsson, B. & Spence, J. C. H. (2011). *Nature (London)*. **470**, 73–78.

Chapman, S., Faulkner, C., Kaiserli, E., Garcia-Mata, C., Savenkov, E. I., Roberts, A. G., Oparka, K. J. & Christie, J. M. (2008). *Proc. Natl Acad. Sci. USA*. **105**, 20038–20043.

Cheng, Y. (2018). *Science*. **361**, 876–880.

Christie, J. M. (2007). *Annu. Rev. Plant Biol.* **58**, 21–45.

Christie, J. M., Hitomi, K., Arvai, A. S., Hartfield, K. A., Mettlen, M., Pratt, A. J., Tainer, J. A. & Getzoff, E. D. (2012). *J. Biol. Chem.* **287**, 22295–22304.

Conrad, C. E., Basu, S., James, D., Wang, D., Schaffer, A., Roy-Chowdhury, S., Zatsepin, N. A., Aquila, A., Coe, J., Gati, C., Hunter, M. S., Koglin, J. E., Kupitz, C., Nelson, G., Subramanian, G., White, T. A., Zhao, Y., Zook, J., Boutet, S., Cherezov, V., Spence, J. C. H., Fromme, R., Weierstall, U. & Fromme, P. (2015). *IUCrJ*. **2**, 421–430.

Coquelle, N., Brewster, A. S., Kapp, U., Shilova, A., Weinhausen, B., Burghammer, M. & Colletier, J. P. (2015). *Acta Cryst. D*. **71**, 1184–1196.

Coquelle, N., Sliwa, M., Woodhouse, J., Schirò, G., Adam, V., Aquila, A., Barends, T. R. M., Boutet, S., Byrdin, M., Carbajo, S., Mora, E. D. La, Doak, R. B., Feliks, M., Fieschi, F., Foucar, L., Guillon, V., Hilpert, M., Hunter, M. S., Jakobs, S., Koglin, J. E., Kovacsova, G., Lane, T. J., Lévy, B., Liang, M., Nass, K., Ridard, J., Robinson, J. S., Roome, C. M., Ruckebusch, C., Seaberg, M., Thepaut, M., Cammarata, M., Demachy, I., Field, M., Shoeman, R. L., Bourgeois, D., Colletier, J. P., Schlichting, I. & Weik, M. (2018). *Nat. Chem.* **10**, 31–37.

Crosson, S. & Moffat, K. (2001). *Proc. Natl Acad. Sci. USA*. **98**, 2995–3000.

Crosson, S. & Moffat, K. (2002). *Plant Cell*. **14**, 1067–1075.

DePonte, D. P., Weierstall, U., Schmidt, K., Warner, J., Starodub, D., Spence, J. C. H. & Doak, R. B. (2008). *J. Phys. D. Appl. Phys.* **41**, 195505.

Edlund, P., Takala, H., Claesson, E., Henry, L., Dods, R., Lehtivuori, H., Panman, M., Pande, K., White, T., Nakane, T., Berntsson, O., Gustavsson, E., Båth, P., Modi, V., Roy-Chowdhury, S., Zook, J., Berntsen, P., Pandey, S., Poudyal, I., Tenboer, J., Kupitz, C., Barty, A., Fromme, P., Koralek, J. D., Tanaka, T., Spence, J., Liang, M., Hunter, M. S., Boutet, S., Nango, E., Moffat, K., Groenhof, G., Ihalainen, J., Stojković, E. A., Schmidt, M. & Westenhoff, S. (2016). *Sci. Rep.* **6**, 35279.

Fedorov, R., Schlichting, I., Hartmann, E., Domratcheva, T., Fuhrmann, M. & Hegemann, P. (2003). *Biophys. J.* **84**, 2474–2482.

Feliks, M., Lafaye, C., Shu, X., Royant, A. & Field, M. (2016). *Biochemistry*. **55**, 4263–4274.

Gabruk, M. & Mysliwa-Kurdziel, B. (2015). *Biochemistry*. **54**, 5255–5262.

Gotthard, G., Von Stetten, D., Clavel, D., Noirclerc-Savoye, M. & Royant, A. (2017). *Biochemistry*. **56**, 6418–6422.

Heymann, M., Opthalage, A., Wierman, J. L., Akella, S., Szebenyi, D. M. E., Gruner, S. M. & Fraden, S. (2014). *IUCrJ*. **1**, 349–360.

Hunter, M. S., Segelke, B., Messerschmidt, M., Williams, G. J., Zatsepin, N. A., Barty, A.,

Benner, W. H., Carlson, D. B., Coleman, M., Graf, A., Hau-Riege, S. P., Pardini, T., Seibert, M. M., Evans, J., Boutet, S. & Frank, M. (2014). *Nat. Sci. Reports*. **4**, 6026.

Inoue, S. ichiro, Takemiya, A. & Shimazaki, K. ichiro (2010). *Curr. Opin. Plant Biol.* **13**, 587–593.

Kasahara, M., Swartz, T. E., Olney, M. A., Onodera, A., Mochizuki, N., Fukuzawa, H., Asamizu, E., Tabata, S., Kanegae, H., Takano, M., Christie, J. M., Nagatani, A. & Briggs, W. R. (2002). *Plant Physiol.* **129**, 762–773.

Kern, J., Alonso-Mori, R., Hellmich, J., Tran, R., Hattne, J., Laksmono, H., Glöckner, C., Echols, N., Sierra, R. G., Sellberg, J., Lassalle-Kaiser, B., Gildea, R. J., Glatzel, P., Grosse-Kunstleve, R. W., Latimer, M. J., McQueen, T. A., DiFiore, D., Fry, A. R., Messerschmidt, M., Miahnahri, A., Schafer, D. W., Seibert, M. M., Sokaras, D., Weng, T.-C., Zwart, P. H., White, W. E., Adams, P. D., Bogan, M. J., Boutet, S., Williams, G. J., Messinger, J., Sauter, N. K., Zouni, A., Bergmann, U., Yano, J. & Yachandra, V. K. (2012). *Proc. Natl Acad. Sci. USA.* **109**, 9721–9726.

Key, J., Šrajer, V., Pahl, R. & Moffat, K. (2007). *Biochemistry*. **46**, 4706–4715.

Khvostichenko, D. S., Schieferstein, J. M., Pawate, A. S., Laible, P. D. & Kenis, P. J. A. (2014). *Cryst. Growth Des.* **14**, 4886–4890.

Kirian, R. A., Wang, X., Weierstall, U., Schmidt, K. E., Spence, J. C. H., Hunter, M., Fromme, P., White, T., Chapman, H. N. & Holton, J. (2010). *Opt. Express*. **18**, 5713.

Kovácsová, G., Grünbein, M. L., Kloos, M., Barends, T. R. M., Schlesinger, R., Heberle, J., Kabsch, W., Shoeman, R. L., Doak, R. B. & Schlichting, I. (2017). *IUCrJ*. **4**, 400–410.

Kupitz, C., Basu, S., Grotjohann, I., Fromme, R., Zatsepin, N. A., Rendek, K. N., Hunter, M. S., Shoeman, R. L., White, T. A., Wang, D., James, D., Yang, J.-H., Cobb, D. E., Reeder, B., Sierra, R. G., Liu, H., Barty, A., Aquila, A. L., Deponte, D., Kirian, R. A., Bari, S., Bergkamp, J. J., Beyerlein, K. R., Bogan, M. J., Caleman, C., Chao, T.-C., Conrad, C. E., Davis, K. M., Fleckenstein, H., Galli, L., Hau-Riege, S. P., Kassemeyer, S., Laksmono, H., Liang, M., Lomb, L., Marchesini, S., Martin, A. V., Messerschmidt, M., Milathianaki, D., Nass, K., Ros, A., Roy-Chowdhury, S., Schmidt, K., Seibert, M., Steinbrener, J., Stellato, F., Yan, L., Yoon, C., Moore, T. A., Moore, A. L., Pushkar, Y., Williams, G. J., Boutet, S., Doak, R. B., Weierstall, U., Frank, M., Chapman, H. N., Spence, J. C. H. & Fromme, P. (2014). *Nature (London)*. **513**, 261–265.

Lelimousin, M., Noirclerc-Savoye, M., Lazareno-Saez, C., Paetzold, B., Le Vot, S., Chazal, R., Macheboeuf, P., Field, M. J., Bourgeois, D. & Royant, A. (2009). *Biochemistry*. **48**, 10038–10046.

Liscum, E., Askinosie, S. K., Leuchtman, D. L., Morrow, J., Willenburg, K. T. & Coats, D. R. (2014). *Plant Cell*. **26**, 38–55.

Lyubimov, A. Y., Murray, T. D., Koehl, A., Araci, I. E., Uervirojnangkoorn, M., Zeldin, O. B., Cohen, A. E., Soltis, S. M., Baxter, E. L., Brewster, A. S., Sauter, N. K., Brunger, A. T. & Berger, J. M. (2015). *Acta Cryst. D*. **71**, 928–940.

Martin-Garcia, J. M., Conrad, C. E., Nelson, G., Stander, N., Zatsepin, N. A., Zook, J., Zhu, L., Geiger, J., Chun, E., Kissick, D., Hilgart, M. C., Ogata, C., Ishchenko, A., Nagaratnam, N., Roy-Chowdhury, S., Coe, J., Subramanian, G., Schaffer, A., James, D., Ketwala, G., Venugopalan, N., Xu, S., Corcoran, S., Ferguson, D., Weierstall, U., Spence, J. C. H.,

Cherezov, V., Fromme, P., Fischetti, R. F. & Liu, W. (2017). *IUCrJ*. **4**, 439–454.

Matz, M. V, Fradkov, a F., Labas, Y. a, Savitsky, a P., Zaraisky, a G., Markelov, M. L. & Lukyanov, S. a (1999). *Nat. Biotechnol.* **17**, 969–973.

Meents, A., Wiedorn, M. O., Srajer, V., Henning, R., Sarrou, I., Bergtholdt, J., Barthelmess, M., Reinke, P. Y. A., Dierksmeyer, D., Tolstikova, A., Schaible, S., Messerschmidt, M., Ogata, C. M., Kissick, D. J., Taft, M. H., Manstein, D. J., Lieske, J., Oberthuer, D., Fischetti, R. F. & Chapman, H. N. (2017). *Nat. Commun.* **8**, 1281.

Nakamura, T., Zhao, Y., Yamagata, Y., Hua, Y. & Yang, W. (2012). *Nature (London)*. **487**, 196–201.

Nango, E., Royant, A., Kubo, M., Nakane, T., Wickstrand, C., Kimura, T., Tanaka, T., Tono, K., Song, C., Tanaka, R., Arima, T., Yamashita, A., Kobayashi, J., Hosaka, T., Mizohata, E., Nogly, P., Sugahara, M., Nam, D., Nomura, T., Shimamura, T., Im, D., Fujiwara, T., Yamanaka, Y., Jeon, B., Nishizawa, T., Oda, K., Fukuda, M., Andersson, R., Båth, P., Dods, R., Davidsson, J., Matsuoka, S., Kawatake, S., Murata, M., Nureki, O., Owada, S., Kameshima, T., Hatsui, T., Joti, Y., Schertler, G., Yabashi, M., Bondar, A.-N., Standfuss, J., Neutze, R. & Iwata, S. (2016). *Science*. **354**, 1552 LP – 1557.

Neutze, R., Wouts, R., van der Spoel, D., Weckert, E. & Hajdu, J. (2000). *Nature (London)*. **406**, 752–757.

Nogly, P., James, D., Wang, D., White, T. A., Zatsepin, N., Shilova, A., Nelson, G., Liu, H., Johansson, L., Heymann, M., Jaeger, K., Metz, M., Wickstrand, C., Wu, W., B??th, P., Berntsen, P., Oberthuer, D., Panneels, V., Cherezov, V., Chapman, H., Schertler, G., Neutze, R., Spence, J., Moraes, I., Burghammer, M., Standfuss, J. & Weierstall, U. (2015). *IUCrJ*. **2**, 168–176.

Nogly, P., Panneels, V., Nelson, G., Gati, C., Kimura, T., Milne, C., Milathianaki, D., Kubo, M., Wu, W., Conrad, C., Coe, J., Bean, R., Zhao, Y., Båth, P., Dods, R., Harimoorthy, R., Beyerlein, K. R., Rheinberger, J., James, D., DePonte, D., Li, C., Sala, L., Williams, G. J., Hunter, M. S., Koglin, J. E., Berntsen, P., Nango, E., Iwata, S., Chapman, H. N., Fromme, P., Frank, M., Abela, R., Boutet, S., Barty, A., White, T. A., Weierstall, U., Spence, J., Neutze, R., Schertler, G. & Standfuss, J. (2016). *Nat. Commun.* **7**, 12314.

Nogly, P., Weinert, T., James, D., Carbajo, S., Ozerov, D., Furrer, A., Gashi, D., Borin, V., Skopintsev, P., Jaeger, K., Nass, K., Båth, P., Bosman, R., Koglin, J., Seaberg, M., Lane, T., Kekilli, D., Brünle, S., Tanaka, T., Wu, W., Milne, C., White, T., Barty, A., Weierstall, U., Panneels, V., Nango, E., Iwata, S., Hunter, M., Schapiro, I., Schertler, G., Neutze, R. & Standfuss, J. (2018). *Science*. **361**,.

Oberthuer, D., Knoška, J., Wiedorn, M. O., Beyerlein, K. R., Bushnell, D. A., Kovaleva, E. G., Heymann, M., Gumprecht, L., Kirian, R. A., Barty, A., Mariani, V., Tolstikova, A., Adriano, L., Awel, S., Barthelmess, M., Dörner, K., Xavier, P. L., Yefanov, O., James, D. R., Nelson, G., Wang, D., Calvey, G., Chen, Y., Schmidt, A., Szczepek, M., Frielingsdorf, S., Lenz, O., Snell, E., Robinson, P. J., Šarler, B., Belšak, G., Maček, M., Wilde, F., Aquila, A., Boutet, S., Liang, M., Hunter, M. S., Scheerer, P., Lipscomb, J. D., Weierstall, U., Kornberg, R. D., Spence, J. C. H., Pollack, L., Chapman, H. N. & Bajt, S. (2017). *Sci. Rep.*

Oghbaey, S., Sarracini, A., Ginn, H. M., Pare-Labrosse, O., Kuo, A., Marx, A., Epp, S. W., Sherrell, D. A., Eger, B. T., Zhong, Y., Loch, R., Mariani, V., Alonso-Mori, R., Nelson,

S., Lemke, H. T., Owen, R. L., Pearson, A. R., Stuart, D. I., Ernst, O. P., Mueller-Werkmeister, H. M. & Miller, R. J. D. (2016). *Acta Cryst. D*. **72**, 944–955.

Pande, K., Hutchison, C. D. M., Groenhof, G., Aquila, A., Robinson, J. S., Tenboer, J., Basu, S., Boutet, S., DePonte, D. P., Liang, M., White, T. A., Zatsepin, N. A., Yefanov, O., Morozov, D., Oberthuer, D., Gati, C., Subramanian, G., James, D., Zhao, Y., Koralek, J., Brayshaw, J., Kupitz, C., Conrad, C., Roy-Chowdhury, S., Coe, J. D., Metz, M., Xavier, P. L., Grant, T. D., Koglin, J. E., Ketawala, G., Fromme, R., Šrajer, V., Henning, R., Spence, J. C. H., Ourmazd, A., Schwander, P., Weierstall, U., Frank, M., Fromme, P., Barty, A., Chapman, H. N., Moffat, K., van Thor, J. J. & Schmidt, M. (2016). *Science*.

Pawate, A. S., Srajer, V., Schieferstein, J., Guha, S., Henning, R., Kosheleva, I., Schmidt, M., Ren, Z., Kenis, P. J. A. & Perry, S. L. (2015). *Acta Cryst. F* . **71**, 823–830.

Prasher, D. C., Eckenrode, V. K., Ward, W. W., Prendergast, F. G. & Cormier, M. J. (1992). *Gene*. **111**, 229–233.

Rizzo, M. a, Springer, G. H., Granada, B. & Piston, D. W. (2004). *Nat. Biotechnol.* **22**, 445–449.

Roedig, P., Duman, R., Sanchez-Weatherby, J., Vartiainen, I., Burkhardt, A., Warmer, M., David, C., Wagner, A. & Meents, A. (2016). *J. Appl. Cryst.* **49**, 968–975.

Roedig, P., Ginn, H. M., Pakendorf, T., Sutton, G., Harlos, K., Walter, T. S., Meyer, J., Fischer, P., Duman, R., Vartiainen, I., Reime, B., Warmer, M., Brewster, A. S., Young, I. D., Michels-Clark, T., Sauter, N. K., Kotecha, A., Kelly, J., Rowlands, D. J., Sikorsky, M., Nelson, S., Damiani, D. S., Alonso-Mori, R., Ren, J., Fry, E. E., David, C., Stuart, D. I., Wagner, A. & Meents, A. (2017). *Nat. Methods*.

Roedig, P., Vartiainen, I., Duman, R., Panneerselvam, S., Stübe, N., Lorbeer, O., Warmer, M., Sutton, G., Stuart, D. I., Weckert, E., David, C., Wagner, A. & Meents, A. (2015). *Sci. Rep.* **5**, 10451.

Roessler, C. G., Agarwal, R., Allaire, M., Alonso-Mori, R., Andi, B., Bachega, J. F. R., Bommer, M., Brewster, A. S., Browne, M. C., Chatterjee, R., Cho, E., Cohen, A. E., Cowan, M., Datwani, S., Davidson, V. L., Defever, J., Eaton, B., Ellson, R., Feng, Y., Ghislain, L. P., Glownia, J. M., Han, G., Hattne, J., Hellmich, J., Héroux, A., Ibrahim, M., Kern, J., Kuczewski, A., Lemke, H. T., Liu, P., Majlof, L., McClintock, W. M., Myers, S., Nelsen, S., Olechno, J., Orville, A. M., Sauter, N. K., Soares, A. S., Soltis, S. M., Song, H., Stearns, R. G., Tran, R., Tsai, Y., Uervirojnangkoorn, M., Wilmot, C. M., Yachandra, V., Yano, J., Yukl, E. T., Zhu, D. & Zouni, A. (2016). *Structure*. **24**, 631–640.

Sancar, A. (2016). *Angew. Chemie Int. Ed.* **55**, 8502–8527.

Sanchez-Weatherby, J., Bowler, M. W., Huet, J., Gobbo, A., Felisaz, F., Lavault, B., Moya, R., Kadlec, J., Ravelli, R. B. G. & Cipriani, F. (2009). *Acta Cryst. D*. **65**, 1237–1246.

Schlichting, I. (2015). *IUCrJ*. **2**, 246–255.

Schotte, F., Cho, H. S., Kaila, V. R. I., Kamikubo, H., Dashdorj, N., Henry, E. R., Graber, T. J., Henning, R., Wulff, M., Hummer, G., Kataoka, M. & Anfinrud, P. A. (2012). *Proc. Natl Acad. Sci. USA*. **109**, 19256–19261.

Schotte, F., Lim, M., Jackson, T. A., Smirnov, A. V, Soman, J., Olson, J. S., Phillips, G. N., Wulff, M. & Anfinrud, P. a (2003). *Science*. **300**, 1944–1947.

Schubert, R., Kapis, S., Gicquel, Y., Bourenkov, G., Schneider, T. R., Heymann, M., Betzel, C. & Perbandt, M. (2016). *IUCrJ*. **3**, 393–401.

Schulz, E. C., Mehrabi, P., Müller-Werkmeister, H. M., Tellkamp, F., Jha, A., Stuart, W., Persch, E., De Gasparo, R., Diederich, F., Pai, E. F. & Miller, R. J. D. (2018). *Nat. Methods*. **15**, 901–904.

Seibert, M. M., Ekeberg, T., Maia, F. R. N. C., Svenda, M., Andreasson, J., Jönsson, O., Odić, D., Iwan, B., Rocker, A., Westphal, D., Hantke, M., Deponte, D. P., Barty, A., Schulz, J., Gumprecht, L., Coppola, N., Aquila, A., Liang, M., White, T. A., Martin, A., Caleman, C., Stern, S., Abergel, C., Seltzer, V., Claverie, J. M., Bostedt, C., Bozek, J. D., Boutet, S., Miahnahri, A. A., Messerschmidt, M., Krzywinski, J., Williams, G., Hodgson, K. O., Bogan, M. J., Hampton, C. Y., Sierra, R. G., Starodub, D., Andersson, I., Bajt, S., Barthelmess, M., Spence, J. C. H., Fromme, P., Weierstall, U., Kirian, R., Hunter, M., Doak, R. B., Marchesini, S., Hau-Riege, S. P., Frank, M., Shoeman, R. L., Lomb, L., Epp, S. W., Hartmann, R., Rolles, D., Rudenko, A., Schmidt, C., Foucar, L., Kimmel, N., Holl, P., Rudek, B., Erk, B., Hömke, A., Reich, C., Pietschner, D., Weidenspointner, G., Strüder, L., Hauser, G., Gorke, H., Ullrich, J., Schlichting, I., Herrmann, S., Schaller, G., Schopper, F., Soltau, H., Kühnel, K. U., Andritschke, R., Schröter, C. D., Krasniqi, F., Bott, M., Schorb, S., Rupp, D., Adolph, M., Gorkhover, T., Hirsemann, H., Potdevin, G., Graafsma, H., Nilsson, B., Chapman, H. N. & Hajdu, J. (2011). *Nature (London)*. **470**, 78–82.

Shimomura, O., Johnson, F. H. & Saiga, Y. (1962). *J. Cell. Comp. Physiol.* **59**, 223–239.

Shu, X., Lev-Ram, V., Deerinck, T. J., Qi, Y., Ramko, E. B., Davidson, M. W., Jin, Y., Ellisman, M. H. & Tsien, R. Y. (2011). *PLoS Biol.* **9**, e1001041.

Shu, X., Royant, A., Lin, M. Z., Aguilera, T. a, Lev-Ram, V., Steinbach, P. a & Tsien, R. Y. (2009). *Science*. **324**, 804–807.

Sierra, R. G., Laksmono, H., Kern, J., Tran, R., Hattne, J., Alonso-Mori, R., Lassalle-Kaiser, B., Glöckner, C., Hellmich, J., Schafer, D. W., Echols, N., Gildea, R. J., Grosse-Kunstleve, R. W., Sellberg, J., McQueen, T. A., Fry, A. R., Messerschmidt, M. M., Miahnahri, A., Seibert, M. M., Hampton, C. Y., Starodub, D., Loh, N. D., Sokaras, D., Weng, T. C., Zwart, P. H., Glatzel, P., Milathianaki, D., White, W. E., Adams, P. D., Williams, G. J., Boutet, S., Zouni, A., Messinger, J., Sauter, N. K., Bergmann, U., Yano, J., Yachandra, V. K. & Bogan, M. J. (2012). *Acta Cryst. D*. **68**, 1584–1587.

Sorigué, D., Légeret, B., Cuiné, S., Blangy, S., Moulin, S., Billon, E., Richaud, P., Brugière, S., Couté, Y., Nurizzo, D., Müller, P., Brettel, K., Pignol, D., Arnoux, P., Li-Beisson, Y., Peltier, G. & Beisson, F. (2017). *Science*. **357**, 903–907.

Stellato, F., Oberthur, D., Liang, M., Bean, R., Gati, C., Yefanov, O., Barty, A., Burkhardt, A., Fischer, P., Galli, L., Kirian, R. A., Meyer, J., Panneerselvam, S., Yoon, C. H., Chervinskii, F., Speller, E., White, T. A., Betzel, C., Meents, A. & Chapman, H. N. (2014). *IUCrJ*. **1**, 204–212.

von Stetten, D., Noirclerc-Savoye, M., Goedhart, J., Gadella, T. W. J. & Royant, A. (2012). *Acta Cryst. F* . **68**, 878–882.

Sugahara, M., Mizohata, E., Nango, E., Suzuki, M., Tanaka, T., Masuda, T., Tanaka, R., Shimamura, T., Tanaka, Y., Suno, C., Ihara, K., Pan, D., Kakinouchi, K., Sugiyama, S., Murata, M., Inoue, T., Tono, K., Song, C., Park, J., Kameshima, T., Hatsui, T., Joti, Y.,

Yabashi, M. & Iwata, S. (2015). *Nat. Methods*. **12**, 61–63.

Sugahara, M., Song, C., Suzuki, M., Masuda, T., Inoue, S., Nakane, T., Yumoto, F., Nango, E., Tanaka, R., Tono, K., Joti, Y., Kameshima, T., Hatsui, T., Yabashi, M., Nureki, O., Numata, K. & Iwata, S. (2016). *Nat. Sci. Reports*. **6**, 24484.

Sui, S., Wang, Y., Kolewe, K., Srajer, V., Henning, R., Schiffman, J. D., Dimitrakopoulos, C. & Perry, S. L. (2016). *Lab Chip*. **16**, 3082–3096.

Tenboer, J. (2015). *Science*. **346**, 1242–1246.

Wagner, J. R., Brunzelle, J. S., Forest, K. T. & Vierstra, R. D. (2005). *Nature (London)*. **438**, 325–331.

Wagner, J. R., Zhang, J., Von Stetten, D., Günther, M., Murgida, D. H., Mroginski, M. A., Walker, J. M., Forest, K. T., Hildebrandt, P. & Vierstra, R. D. (2008). *J. Biol. Chem.* **283**, 12212–12226.

Weierstall, U., James, D., Wang, C., White, T. A., Wang, D., Liu, W., Spence, J. C. H., Bruce Doak, R., Nelson, G., Fromme, P., Fromme, R., Grotjohann, I., Kupitz, C., Zatsepin, N. A., Liu, H., Basu, S., Wacker, D., Won Han, G., Katritch, V., Boutet, S., Messerschmidt, M., Williams, G. J., Koglin, J. E., Marvin Seibert, M., Klinker, M., Gati, C., Shoeman, R. L., Barty, A., Chapman, H. N., Kirian, R. A., Beyerlein, K. R., Stevens, R. C., Li, D., Shah, S. T. A., Howe, N., Caffrey, M. & Cherezov, V. (2014). *Nat. Commun.* **5**,.

Weinert, T., Skopintsev, P., James, D., Dworkowski, F., Panepucci, E., Kekilli, D., Furrer, A., Brünle, S., Mous, S., Ozerov, D., Wang, M. & Standfuss, J. (2019). *BioRxiv*.

White, T. A., Kirian, R. A., Martin, A. V, Aquila, A., Nass, K., Barty, A. & Chapman, H. N. (2012). *J. Appl. Cryst.* **45**, 335–341.

Wöhri, A. B., Katona, G., Johansson, L. C., Fritz, E., Malmerberg, E., Andersson, M., Vincent, J., Eklund, M., Cammarata, M., Wulff, M., Davidsson, J., Groenhof, G. & Neutze, R. (2010). *Science*. **328**, 630–633.

Wulff, M., Schotte, F., Naylora, G., Bourgeois, D., Moffat, K. & Mourou, G. (1997). *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.* **398**, 69–84.

Zarrine-Afsar, A., Barends, T. R. M., Muller, C., Fuchs, M. R., Lomb, L., Schlichting, I. & Miller, R. J. D. (2012). *Acta Cryst. D*. **68**, 321–323.

Zayner, J. P. & Sosnick, T. R. (2014). *PLoS One*. **9**, e87074.

Chapter 2

CRYSTALLOGENESIS:

METHODS & RESULTS

Coloured protein crystallization and crystal size control tuning for time-resolved study purposes

Scientific publications and non-published work presented in this PhD thesis manuscript share common strategies for obtaining protein crystals. These strategies have been developed at the early stages of the PhD project, by first exploring the crystallization properties of various Red Fluorescent Proteins (RFPs) and chromoproteins (coloured, but non-fluorescent protein). This includes in particular mScarlet, the brightest RFP engineered to this date, which we investigated in collaboration with the group of Prof. Gadella at the University of Amsterdam. We conducted an extensive structural study of this protein and six mutants. The crystallization strategies developed for this study were then adapted for each of the three proteins envisaged in the time-resolved crystallography project (*Dr*CBD, Twist-Cerulean and *At*Phot2LOV2) with methodological variations where necessary. The first prerequisite of any crystallization experiment is to work with a protein solution sample of highest purity. Molecular biology, protein expression and purification protocols are not detailed here, but one can refer to the method sections of **Chapters 3, 5 and 6**.

2.1 Crystallisation strategies

2.1.1 High throughput crystallization conditions screening

Concentrated protein solutions were sent to the high-throughput crystallization platform of the EMBL Grenoble outstation (HTX lab, https://htxlab.embl.fr/) in order to screen a large number of crystallization conditions. Screens were usually chosen from the *JCSG+* screen (Molecular Dimensions), the *Wizard Classic I & Classic II* screens (Rigaku Reagents), and the *Classics Suite* screen (Qiagen). These three screens constitute some of the most complete sparse matrix screens available, representing statistically favourable conditions to protein crystal formation. The proteins were provided to crystallization robots at given, protein-dependent concentrations in order to optimize the level of precipitation in all drops (not too much, not too little). Three [mother liquor / protein] ratios were usually screened [1/1], [1/2] and [1/3]. All crystallization plates were stored at 291 K (18°C). All conditions were monitored *via* an online automated imaging system (CRIMS, for CRystallization Information Management System) providing drop photographs on a pseudo exponential time scale (1, 3, 7, 15, 33, 61 and 87 days after drop preparation). Conditions yielding promising hits (crystals, microcrystals, spherulites, precipitates) were selected for manual enhancement depending on the complexity of the mother liquor, pH and the nature of the hit.

2.1.2 Crystal optimization

If the condition identified from automated screening has not resulted in a well-diffracting crystal, additional screening is performed using commercially available salt, pH, precipitant or additive screens (Hampton Research). These were performed by the hanging drop vapour diffusion method (**Fig. 2.1a**) at 18°C in 48-well crystallization plates with 12 mm cover slides holding drops of final volume typically less than 1 µL.

Once more favourable crystallization conditions had been obtained, crystals were manually enhanced by the hanging drop vapour diffusion method, using 24-wells plates (HR3-172, Hampton Research) at 18°C with 22 mm cover slides, holding drops of larger volumes (2 µL or more). Manual enhancing of the crystals consisted first in modifying the mother liquor composition using precipitant concentration and pH gradients. In parallel, several [mother liquor / protein] ratios were tested. Finally, a wide range of protein concentrations was investigated, as it is one of the critical parameters that control the size and number of the growing crystals. This strategy eventually leads to crystals that are classically suitable for diffraction experiments in terms of size (thus of diffraction power) and ease of handling (*i.e.* they can be manually fished as individual crystals).

However, this strategy is not suited for the less classical approaches that have been sought in (time-resolved) serial crystallography experiments. In the latter cases, we wished to accurately control the size and shape of the crystals, in particular to be able to perform experiments that could be reproduced many times. In order to do so, I developed a strategy that combines limited proteolysis with the micro-seeding technique. This eventually led me to obtain a rather precise control of the nucleation rate, size and shape of the crystals. Most particularly for the media injection-based SSX technique which required large amounts of microcrystals (less than 10 µm), the batch crystallization method (**Fig. 2.1c**, **paragraph 2.2.3**, *Dr*CBD) and the sitting drop vapour diffusion method (**Fig. 2.1b**, **paragraphs 2.2.4** (Twist-Cerulean) and **2.2.5** (*At*Phot2LOV2)) had to be investigated.

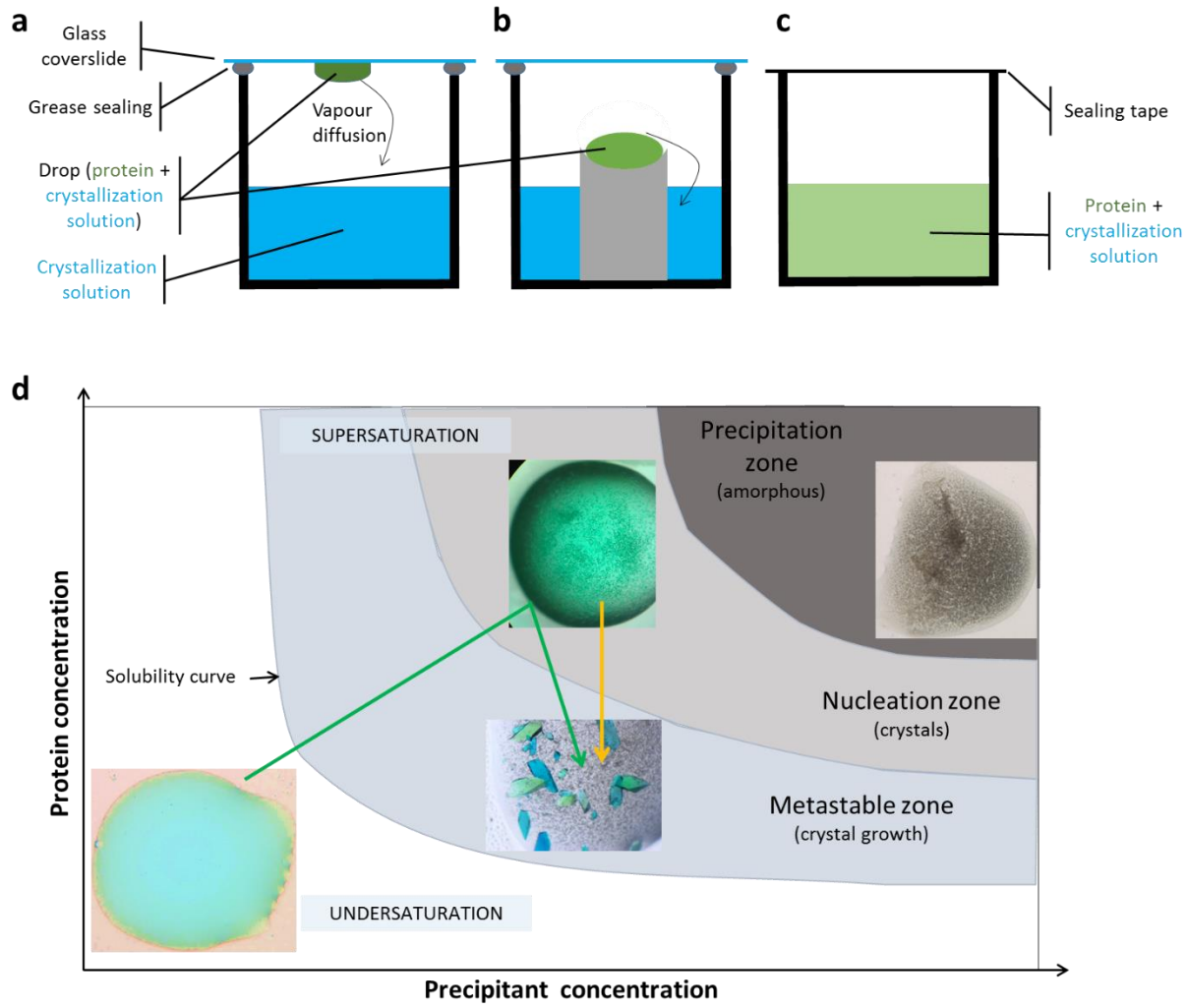


Figure 2.1: Schematic representation of the (**a**) hanging drop vapour diffusion, (**b**) sitting drop vapour diffusion, and (**c**) batch method crystallization setups. (**d**) Schematic of the protein crystallization phase diagram using photos of *Dr*CBD drops along the crystallisation process. The green arrow represents the path from the undersaturation, to the nucleation, then the metastable zone, using the vapour diffusion method. The orange arrow represents the path from the nucleation to the metastable zone in the method, respectively. See **paragraph 2.2.3** for a detailed protocol of batch crystallization of *Dr*CBD.

2.1.3 Limited proteolysis

Prior to crystallization, purified aliquots of protein were submitted to limited proteolysis. This step has proved to improve crystallization, by removing flexible terminal parts of the protein, which impairs the crystallization process. Depending on the target protein, the initial strategy was to test a protease screen (Proti-Ace kit 1 (HR2-429) and kit 2 (HR2-432), Hampton Research). These protease kits were originally designed to generate smaller domains of the target protein, which would be easier to crystallize or would diffract better than the full-

length protein. In the case of our coloured proteins, we tweaked the screen to limit proteolysis to achieve the removal of potentially flexible N- and C-terminal parts, leaving the protein core intact. Mass spectrometry experiments were performed to check the exact effect of the most efficient proteases (data not shown). Most often, the best protease happened to be trypsin, which was then used in our proteolysis protocol (see thereafter), which we adapted from the Proti-Ace kit guidelines.

In order to determine the optimal trypsin concentration, successive dilutions of a trypsin stock solution (50 $\mu g.\mu L^{-1}$) were used (**Table 2.1**). 2 µL of each trypsin dilution were mixed with 20 µL of protein sample. The protein concentration had to be limited to values around 0.5 $mg.mL^{-1}$ in order to limit the width of protein bands on a SDS-PAGE gel and thus to be able to identify small molecular weight variations. [Trypsin / protein] samples were incubated at 30°C under gentle agitation, and aliquots were collected after 1 and 2 h and analysed on a 15% SDS-PAGE gel (**Fig. 2.2**). Identification of small decreases in molecular weight indicate the successful removal of short terminal parts of the protein. Once the right protease dilution had been determined, it was then used on the whole, concentrated protein sample. This step could be performed extemporaneously before crystallization or proteolysed protein samples could be flash-frozen and stored at – 80 °C for later use.

Table 2.1: Limited proteolysis experiment. Trypsin concentration in each reaction dilution.

SOLUTION	DILUTION
A	No trypsin
B	1/10
C	1/100
D	1/500
E	1/1000
F	1/2000

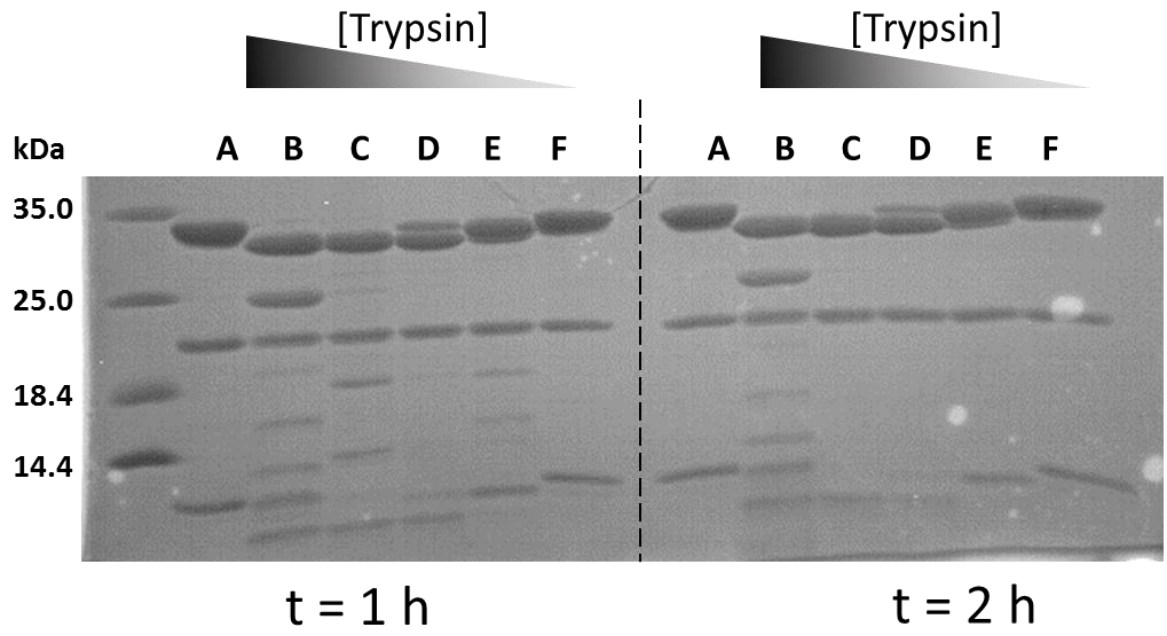


Figure 2.2: SDS-PAGE gel from a limited proteolysis screening on mScarlet of ~ 30 kDa molecular weight (**paragraph 2.2.2**), with dilutions listed in **Table 2.1**. The B dilution appears to digest the protein too much, while dilutions D and above appear not to be enough. A 2 h incubation time gives a cleaner lane for dilution C. Note: the ~20 kDa band (already present in lane A) comes from a partially autocleaved mScarlet at the location of the chromophore main chain, unrelated to trypsin proteolysis.

2.1.4 Micro-seeding protocol

I have used the micro-seeding technique to control the nucleation of crystals, hence their shape and size. Crystals were harvested from one, or several drops of a crystallization plate with the goal of producing a seed stock of highest concentration possible (Bergfors, 2003). Only fresh, small crystals were used for seeding, as defects accumulated in old crystals may impair the quality of the seeds. Crystals were crushed using a 0.1 mL glass tissue homogenizer containing the stabilizing solution in which the seeds will be eventually diluted and stored, either at 4°C for extemporaneous use (up to a year) or at – 80°C for long-term storage. The stabilizing solution is constituted of the crystallization condition with a 10 to 20% increase in precipitant concentration in order to avoid the crystal seeds dissolving. This glass homogenizer-based method was adapted from Luft and DeTitta (Luft & DeTitta, 1999) and used for the mScarlet crystals. I set up an improved seeding protocol (described on **Fig. 2.3**) for the crystallization of *Dr*CBD, Twist-Cerulean and *At*Phot2LOV2 using information from the Hampton Research crystallization manual "Crystal Growth 101" (https://hamptonresearch.com/documents/growth_101/39.pdf) and D'Arcy *et al* (D'Arcy *et al.*,

2014). The evolution of the protocol was necessary because even if the glass homogenizer method gave successful results, it still lacked a proper control of seed size homogeneity, which could be visualized under a microscope. The crushing step of the first method constitutes the first step of the second method and results in a macro-seed stock. In this first step, depending on the accessibility of the crystals (size of the crystallisation well, 'pipettability' of the crystals), one can either used the glass homogenizer or a dedicated glass tool (a "crystal crusher", HR4-216, Hampton Research, picture #2 on **Fig. 2.3**). In the second step, the seeds are vortexed at high speed for 3 to 5 minutes in microcentrifuge tubes containing plastic or metal beads (HR2-230, Hampton Research). The thorough crushing results in a homogenous preparation of micro-seeds. In the third and final step, the micro-seed stock solution is serially diluted, producing various seed solutions, which are then mixed with the protein of interest in a 1 to 10 ratio just before setting up the drops, allowing for a precise control of crystal size.

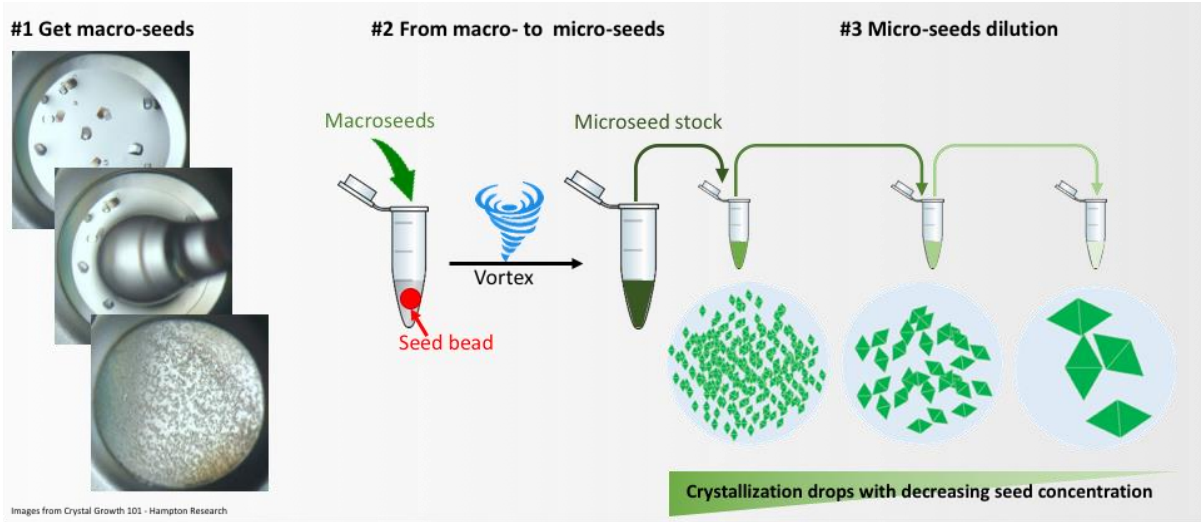


Figure 2.3: Micro-seeding production protocol. See main text for detailed protocol. Images from step #1 are from Hampton Research. An equivalent to the green crystal series can be found in **Figures 2.7** and **2.10**.

While careful usage of the micro-seeding technique has proved crucial in optimizing the size and number of my crystals, I have observed that in one particular case (*At*Phot2LOV2), the combination of micro-seeding and limited proteolysis allowed me to accurately control the number of crystals in each drop between 1 and 1000, when the use of micro-seeding only led to between 100 to 1000 crystals per drop (see **paragraph 2.2.5**).

2.2 Application to coloured proteins

2.2.1 The chromoproteins Pumpkin and Aubergine

During my Master 2 internship, I started crystallizing two chromoproteins: an orange one, Pumpkin, and a purple one, Aubergine. This project was a collaboration with Nathan Shaner from the Scintillon Institute in San Diego (California) with the goal of providing a structure that could serve to analyse the oligomeric interfaces and chromophore environment, in order to develop monomeric, fluorescent proteins by rational design. Pumpkin gave me the opportunity to obtain initial hits by automated screening (**Fig. 2.4a**), which I could improve in small crystals, with which I was able to collect data sets to medium resolution (3.0 Å) with a classical approach (data collection on one small crystal) and to higher resolution (2.5 Å) (data collected with the *MeshAndCollect* approach (Zander *et al.*, 2015) on several small crystals). Optimization of a hit that that appeared later led to crystals diffracting at 2.0 Å, which led to the successful structure determination. The process was more complicated for Aubergine, as initially the protein could not be concentrated above 1 $mg.mL^{-1}$ and only crystallized as very small needles, even after limited proteolysis and micro-seeding steps (**Fig. 2.4b**). I then used the Thermal Shift Assay technique (**Fig. 2.4c**) (Boivin *et al.*, 2013) to identify a buffer in which Aubergine was more stable. Consequently, Aubergine could be concentrated up to 8 $mg.mL^{-1}$ and crystallized. A 3.1 Å data set could be collected on one of the resulting crystals, and the structure of Aubergine could be solved. These two structures have not been published yet.

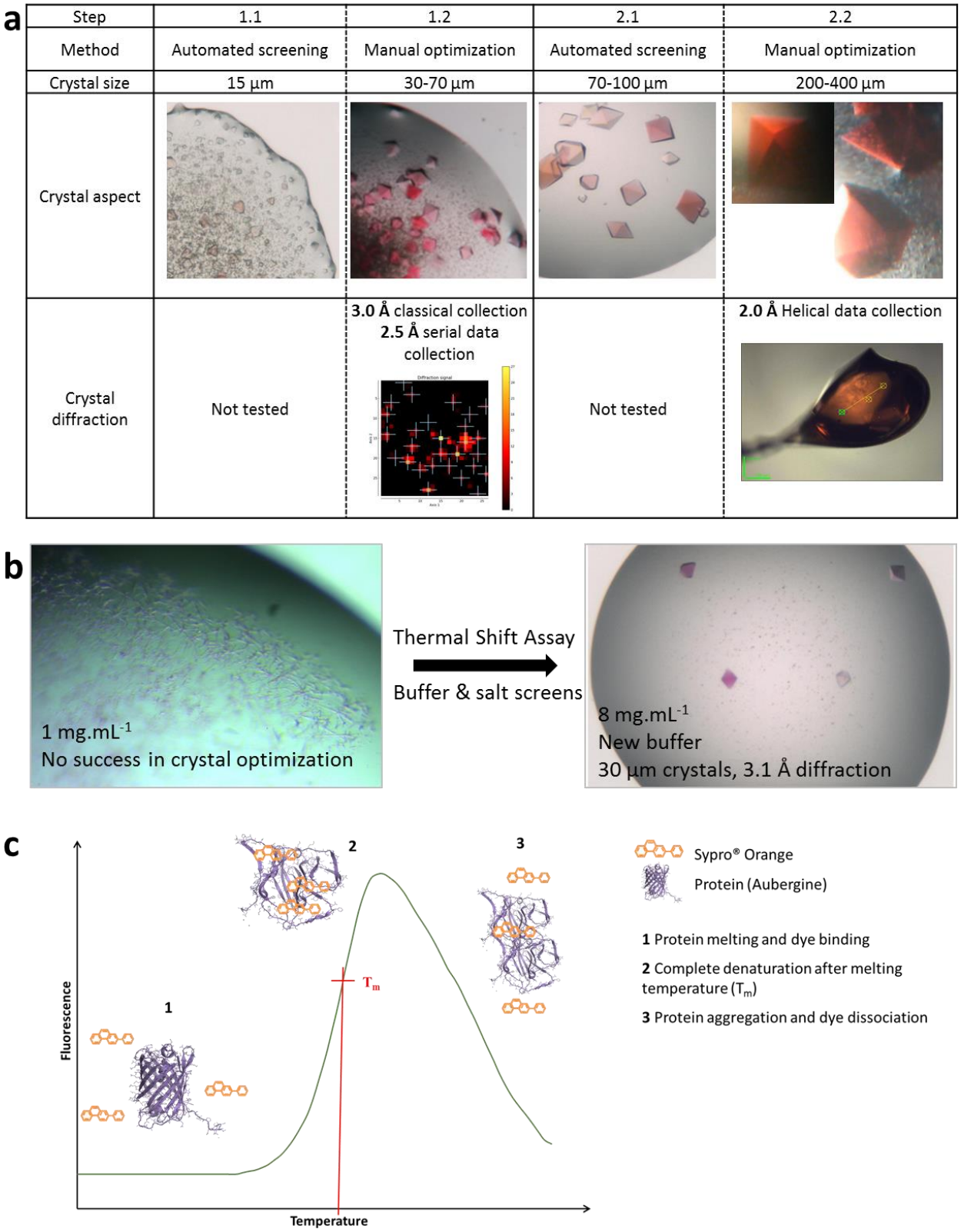


Figure 2.4: (**a**) Crystallization improvement of Pumpkin crystals. (**b**) Aubergine crystals before and after protein stability optimization using a thermal shift assay on various buffers. (**c**) Principle of a thermal shift assay.

2.2.2 The Red Fluorescent Protein mScarlet

As a second step in my Master 2 internship, and during the first months of my PhD, we conducted a project in collaboration with Prof. Gadella from the University of Amsterdam on the structural characterization of the fluorescence enhancement of the new GFP-type Red Fluorescent Protein mScarlet. This work was particularly important for me to develop my experience in crystallogenesis. I managed to crystallized mScarlet, as well as six of its mutants (unpublished data). As shown on **Fig. 2.5**, I could drastically improve the quality of mScarlet crystals from the initial hits obtained *via* automated screening, eventually leading to well-diffracting crystals, and I could thus solve the structure of the protein at 1.47 Å resolution, which allowed us to describe the planar conformation of the chromophore and its added stabilization by neighbouring residues when compared to other RFPs. This work was published as a brief communication in *Nature Methods*: Bindels DS, Haarbosch L, van Weeren L, Postma M, Wiese KE, Mastop M, Aumonier S, Gotthard G, Royant A, Hink MA, Gadella TW Jr (2017) 'mScarlet: a bright monomeric red fluorescent protein for cellular imaging' *Nat. Methods* **14**, 53-56 (**Fig. 2.6)**.

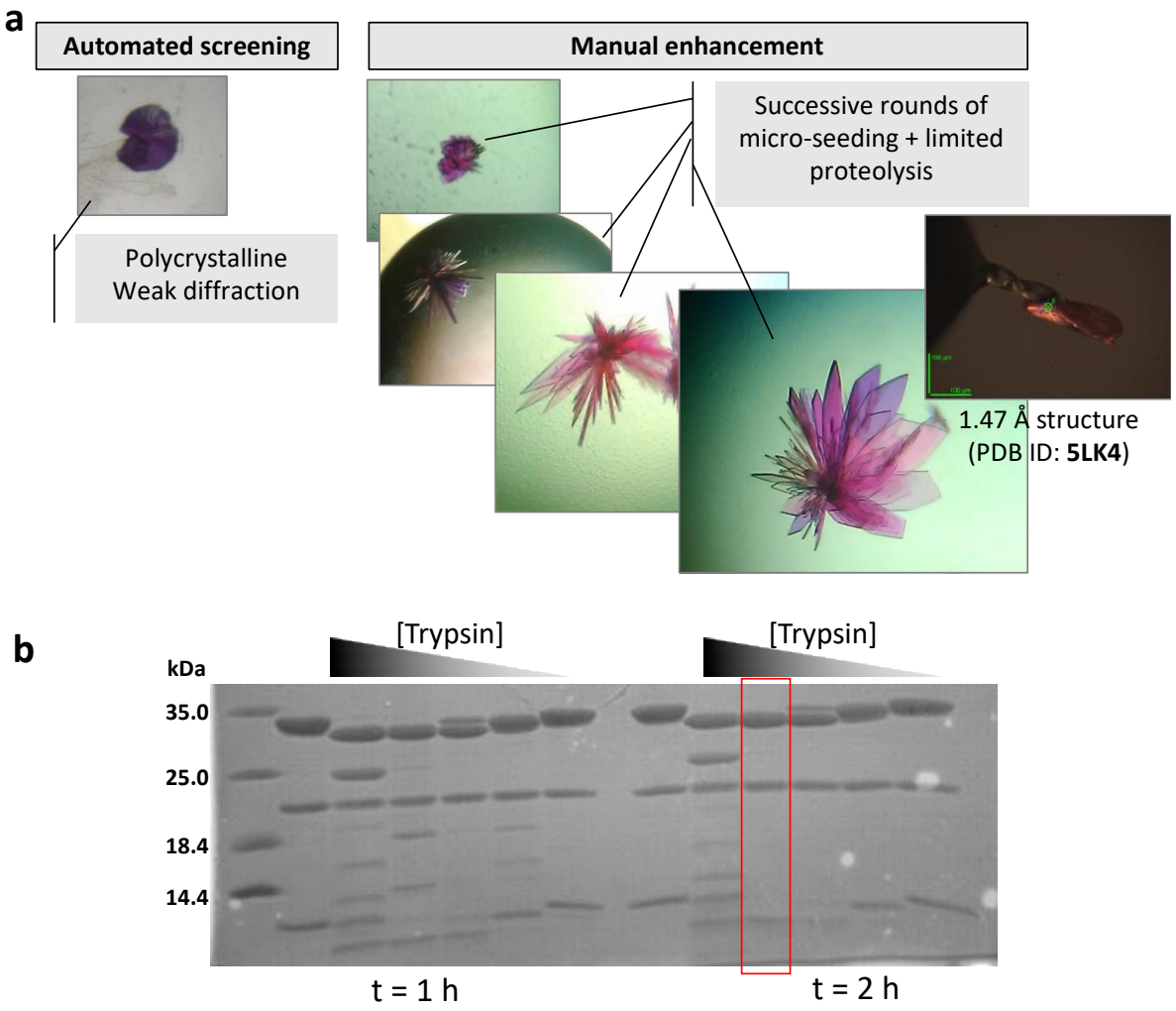


Figure 2.5: (**a**) Iterative improvement of mScarlet crystals after several cycles of manual enhancement. The last picture is that of the crystal used to record diffraction data on the ESRF beamline ID29. (**b**) Limited proteolysis experiment, from which the optimal trypsin concentration has been determined (red rectangle) and applied to the protein stock solution prior to crystallization.

mScarlet: a bright monomeric red fluorescent protein for cellular imaging

Daphne S Bindels[1,4], Lindsay Haarbosch[1,4], Laura van Weeren[1], Marten Postma[1], Katrin E Wiese[1], Marieke Mastop[1], Sylvain Aumonier[2,3], Guillaume Gotthard[2,3], Antoine Royant[2,3], Mark A Hink[1] & Theodorus W J Gadella Jr[1]

We report the engineering of mScarlet, a truly monomeric red fluorescent protein with record brightness, quantum yield (70%) and fluorescence lifetime (3.9 ns). We developed mScarlet starting with a consensus synthetic template and using improved spectroscopic screening techniques; mScarlet's crystal structure reveals a planar and rigidified chromophore. mScarlet outperforms existing red fluorescent proteins as a fusion tag, and it is especially useful as a Förster resonance energy transfer (FRET) acceptor in ratiometric imaging.

Fluorescent proteins (FPs) have become indispensable in biological research[1]. After the cloning of GFP from the jellyfish *Aequorea victoria*[2], several GFP spectral variants were developed, including blue, cyan and yellow FPs[3]. The palette of FPs was greatly expanded after the cloning of red FP (RFP) homologs from corals and other *Anthozoa* species[4,5]. However, all *Anthozoa* RFPs form obligate tetramers, which can seriously interfere with localization and functioning of RFP-fusion proteins. Monomerization of tetrameric RFPs was accompanied by a serious deterioration of the brightness and incomplete and/or partial green maturation of the resulting monomer[6]. After the development of the first monomeric RFP (mRFP), mRFP1 (ref. 6), several improved mRFPs have been reported: mCherry[7], mApple[8], TagRFP(-T)[8,9], mKate2 (ref. 10), mRuby2 (ref. 11), mRuby3 (ref. 12) and FusionRed[13] (reviewed in ref. 14). But all these mRFPs are dimmer than their tetrameric ancestors; they have quantum yields below 50%, and several still harbor additional problems due to incomplete or partial green maturation and a residual tendency to dimerize[15].

Spectral variants of FPs can be applied in FRET-based biosensors to probe molecular interactions, conformational changes and metabolite concentrations within living cells[16]. While good FP-based FRET pairs are available with cyan FPs (CFPs) as donors and yellow FPs (YFPs) as acceptors, it can be highly advantageous to use (additional) FRET pairs with an RFP as the donor or acceptor. This allows excitation at longer wavelengths, which is less harmful or toxic to biological samples, induces less autofluorescence[11,17] and enables multiplexing of several FRET sensors. Unfortunately, the aforementioned drawbacks of current mRFPs seriously limit their use in FRET experiments. Therefore, we focused on engineering a novel bright monomeric RFP with a high quantum yield and complete maturation.

Rather than evolving from a tetrameric natural ancestor, we designed a synthetic gene template to start the development of a new monomeric RFP. We based the template on mCherry and multiple other naturally occurring RFPs and chromo proteins; and we used knowledge about residues at the outer barrel surface to break the dimerization interfaces. We dubbed this template 'mRed7' (**Supplementary Note 1** and **Supplementary Fig. 1**).

Remarkably, mRed7 was fluorescent when expressed in bacteria. However, it exhibited a very low fluorescence lifetime and quantum yield. We performed two rounds of multiple site-directed mutagenesis using the OmniChange method[18], targeting the simultaneous change of multiple amino acid residues in the interior of the beta-barrel in an unbiased manner (**Supplementary Note 2** and **Supplementary Table 1**). The resulting libraries of new RFP variants were initially screened mainly for increased fluorescence lifetime in bacteria. Because the fluorescence lifetime is proportional to the quantum yield but not dependent on expression level, sample thickness or maturation efficiency, it allows for direct screening of the intrinsic brightness of FPs as shown previously for the evolution of mTurquoise and mTurquoise2 (refs. 19,20).

Following the multiple site-directed mutagenesis strategy, the RFPs were subjected to several rounds of random mutagenesis and screened for both increased brightness and proper maturation (i.e., minimal dead-end green chromophore formation[21]) by ratiometric screening using a cotranslated mTurquoise2 protein as reference (**Supplementary Fig. 2**). Eventually, we obtained three monomeric RFPs with distinct properties: a very bright mRFP with a high quantum yield that we dubbed 'mScarlet' and two variants with a single amino acid substitution, called 'mScarlet-I' and 'mScarlet-H' (**Supplementary Fig. 1**).

The normalized absorbance emission spectra indicate that the mScarlet variants are genuine RFPs (**Fig. 1a**) with mScarlet absorbance and emission maxima of 569 nm and 594 nm, respectively (**Supplementary Fig. 3**). Of note, the fluorescence lifetime of mScarlet is 3.9 ns, the highest value recorded to date for mRFPs (**Fig. 1b** and **Table 1**), and it shows monoexponential decay (**Supplementary Fig. 4** and **Supplementary Table 2**). The quantum yield of mScarlet is 0.70; much higher than the quantum

[1]Section of Molecular Cytology and van Leeuwenhoek Centre for Advanced Microscopy, Swammerdam Institute for Life Sciences, University of Amsterdam, Amsterdam, the Netherlands. [2]European Synchrotron Radiation Facility, Grenoble, France. [3]Institut de Biologie Structurale, Université Grenoble Alpes, CNRS, CEA, Grenoble, F-38044, France. [4]These authors contributed equally to this work. Correspondence should be addressed to T.W.J.G. (th.w.j.gadella@uva.nl).

RECEIVED 29 JULY; ACCEPTED 20 OCTOBER; PUBLISHED ONLINE 21 NOVEMBER 2016; CORRECTED ONLINE 12 DECEMBER 2016; DOI:10.1038/NMETH.4074

Figure 2.6: Article published in *Nature Methods* featuring the structure determination of mScarlet at 1.47 Å resolution.

2.2.3 The Cyan Fluorescent Protein Twist-Cerulean

Twist-Cerulean is a single-point mutant from the Cyan Fluorescent Protein Cerulean, which has been well studied in the group and for which crystallization conditions were already known (Lelimousin *et al.*, 2009; Gotthard *et al.*, 2017). Twist-Cerulean could be readily crystallized at a concentration of 13 $mg.mL^{-1}$ in 10 – 20% PEG8000, 100 mM $MgCl_2$, and 100 mM HEPES pH 6.75 – 7.5, at a [mother liquor / protein] ratio of [1/1]. These conditions lead to well-diffracting needle-shaped crystals of length ranging from 100 µm to 1 mm (**Fig. 2.7a)**. Crystals of these shapes and sizes are not suitable for (time-resolved) serial crystallography experiments because of the variable penetration depth of the actinic light (depending on the orientation) and the likely clogging of the HVE injector.

I first tried micro-seeding from the needle form to alter both the size and shape of crystals, which yielded either very small (**Fig. 2.7b,** seeding 1/10), or very big crystals (up to 500 µm long, **Fig. 2.7b,** seeding 1/1000), but still of the needle shape, although the ratio length over thickness had significantly reduced. No improvement in the crystal shape was observed by applying a second cycle of micro-seeding (**Fig. 2.7c**). However, by combining limited proteolysis (**Fig. 2.8**) and pH screening, bi-pyramidal shaped crystals of various sizes (at pH 7.0) and thicker needles (at pH 7.5) crystals could be obtained (**Fig. 2.7d)**. Besides, I tried a 'feeding' strategy on these crystals, *i.e.* I added protein solution to the crystallization drop after crystal formation (details not shown), which resulted in millimeter-sized crystals, which could find application in neutron crystallography (**Fig. 2.7f**).

The best condition to produce bi-pyramidal-shaped microcrystals was reproduced with the sitting drop technique (**Fig. 2.1b**) (HR3-158 plate, Hampton Research), with a drop volume up to 20 µL and a [1/1] ratio, in order to get thousands of microcrystals per drop (**Fig. 2.7e**). One such plate provided enough crystals for one to two injection rounds of the HVE injector (see **Chapter 4**).

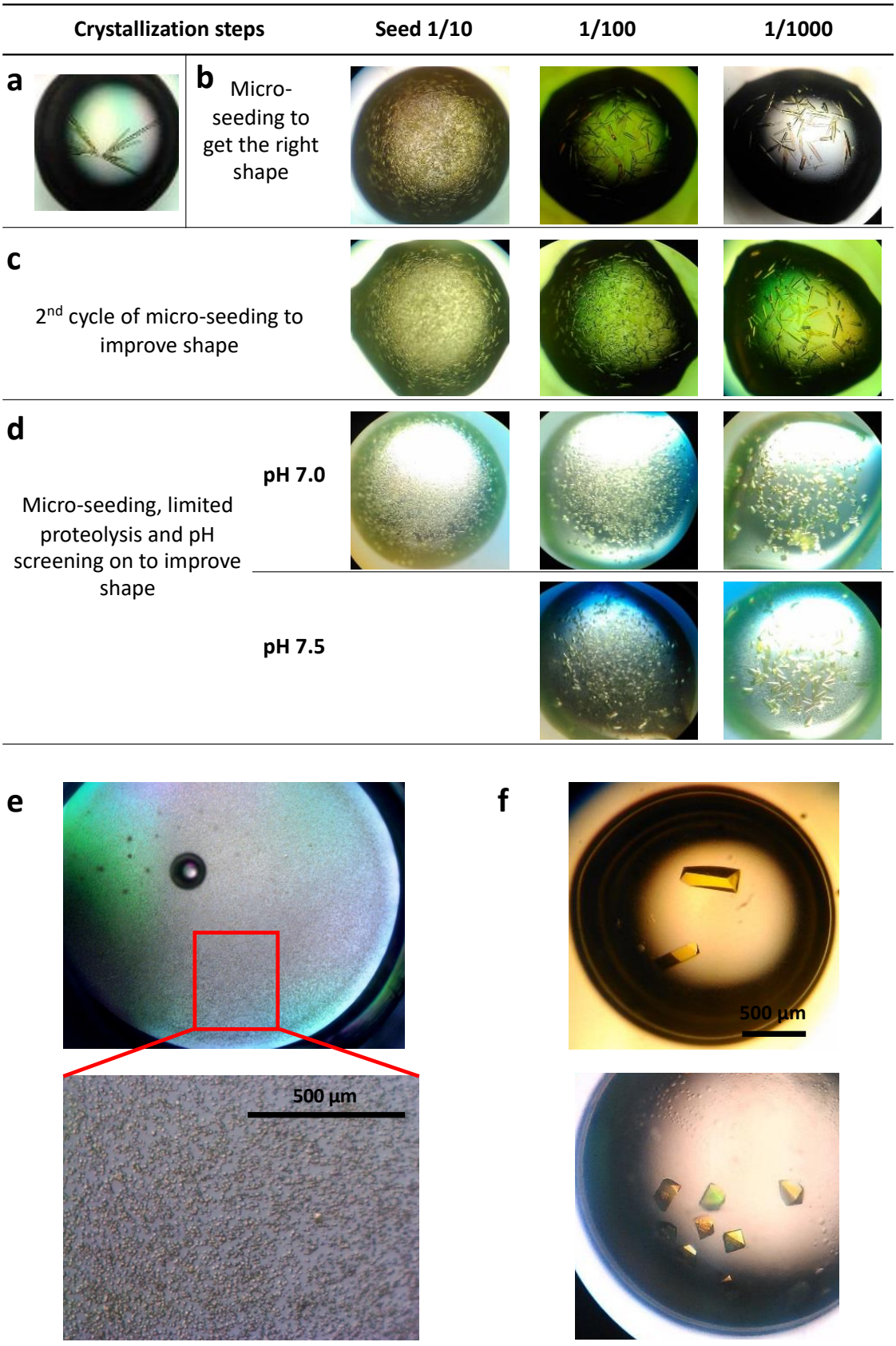


Figure 2.7: (**a**) to (**d**) Various crystallization attempts of Twist-Cerulean resulting in crystals of various sizes and shapes. (**e**) Best condition for injection-based SSX. (**f**) Best condition for large beam X-ray, or neutron crystallography.

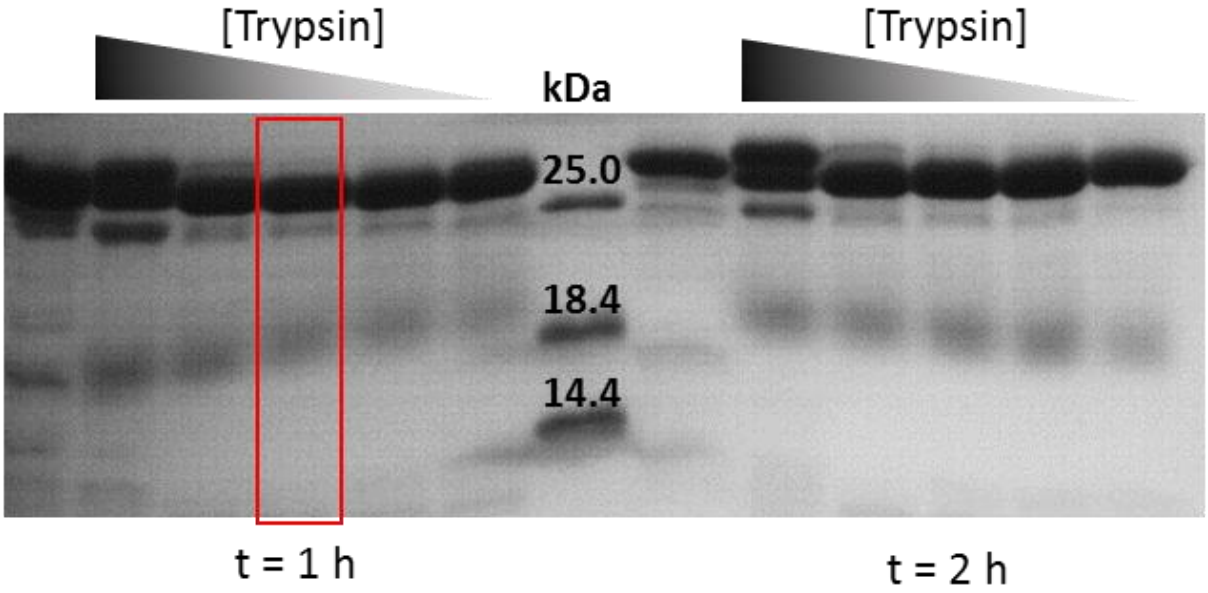


Figure 2.8: Limited proteolysis experiment on Twist-Cerulean, from which the optimal trypsin concentration has been determined (red rectangle) and applied to the protein stock solution prior to crystallization.

Achieving the extensive control of Twist-Cerulean crystallization has made it an attractive model protein to test methodological SSX developments. First, it was used to test the HVE injector (see **Chapter 4**). Second, the microcrystal condition was applied to the parent protein Cerulean to demonstrate the potential of the *MeshAndCollect* technique to determine a protein structure from a number of flash-cooled microcrystals, which was the topic of a video article in *Journal of Visualized Experiments*: Hutin S, Santoni G, Zander U, Foos N, Aumonier S, Gotthard G, Royant A, Mueller-Dieckmann C, Leonard G. (2019) 'Structure Solution of the Fluorescent Protein Cerulean Using *MeshAndCollect*.' *J. Vis. Exp.* **145**, e58594 (**Fig. 2.9)**. It was also used as a model in the investigation of the possibility of phasing a protein structure from microcrystals using a genetic algorithm (Foos *et al.*, 2019).

jove Journal of Visualized Experiments www.jove.com

Video Article

Structure Solution of the Fluorescent Protein Cerulean Using MeshAndCollect

Stephanie Hutin*[1], Gianluca Santoni*[1], Ulrich Zander[2], Nicolas Foos[1], Sylvain Aumonier[1], Guillaume Gotthard[1], Antoine Royant[1,3], Christoph Mueller-Dieckmann[1], Gordon Leonard[1]

[1]European Synchrotron Radiation Facility, Structural Biology Group

[2]European Molecular Biology Laboratory

[3]Univ. Grenoble Alpes, CNRS, CEA, IBS (Institut de Biologie Structurale)

*These authors contributed equally

Correspondence to: Stephanie Hutin at stephanie.hutin@esrf.fr

URL: https://www.jove.com/video/58594
DOI: doi:10.3791/58594

Keywords: Biochemistry, Issue 145, Crystal growth and mounting, macromolecular X-ray crystallography, beamline control software (MXCuBE2), information management system for macromolecular crystallography X-ray experiments (ISPyB), serial crystallography, MeshAndCollect, synchrotron radiation.

Date Published: 3/19/2019

Citation: Hutin, S., Santoni, G., Zander, U., Foos, N., Aumonier, S., Gotthard, G., Royant, A., Mueller-Dieckmann, C., Leonard, G. Structure Solution of the Fluorescent Protein Cerulean Using MeshAndCollect. *J. Vis. Exp.* (145), e58594, doi:10.3791/58594 (2019).

Abstract

X-ray crystallography is the major technique used to obtain high resolution information concerning the 3-dimensional structures of biological macromolecules. Until recently, a major requirement has been the availability of relatively large, well diffracting crystals, which are often challenging to obtain. However, the advent of serial crystallography and a renaissance in multi-crystal data collection methods has meant that the availability of large crystals need no longer be a limiting factor. Here, we illustrate the use of the automated MeshAndCollect protocol, which first identifies the positions of many small crystals mounted on the same sample holder and then directs the collection from the crystals of a series of partial diffraction data sets for subsequent merging and use in structure determination. MeshAndCollect can be applied to any type of micro-crystals, even if weakly diffracting. As an example, we present here the use of the technique to solve the crystal structure of the Cyan Fluorescent Protein (CFP) Cerulean.

Video Link

The video component of this article can be found at https://www.jove.com/video/58594/

Introduction

Macromolecular X-ray crystallography (MX) is, by far, the most used method for gaining atomic resolution insight into the three-dimensional structures of biological macromolecules. However, a major bottle necks is the requirement for relatively large, well diffracting crystals.

Often, and particularly when crystallizing membrane proteins, only very small crystals of a few microns in the largest dimension can be obtained. Radiation damage effects limit the resolution of a complete diffraction data set that can be collected from a single micro crystal[2], and very often, it is necessary to improve the signal to noise ratio and hence data set resolution, by merging several partial diffraction data sets from different, but isomorphic crystals. The increases in flux density of X-ray beams at synchrotron sources and elsewhere (*e.g.* X-ray free-electron lasers (X-FELs)), have meant that useful partial diffraction data sets can be collected from even very small crystals of biological macromolecules. This, in turn, has led to the development of new techniques for the collection and merging of partial diffraction data sets collected from many different crystals in order to produce a complete data set for structure solution. Such techniques are commonly referred to as serial crystallography (SX)[3,4,5,6,7,8]. A prototypical example of SX is the use of injector devices to introduce a narrow stream of a crystal slurry into the X-ray beam[3,4,5]. A diffraction pattern is recorded every time a crystal is exposed to X-rays leading to the collection, from many thousands of individual crystals, of 'still' diffraction images, information which is then merged to produce a complete data set. However, a considerable disadvantage of this type of serial data collection is that the processing of still images can be problematic. The data quality is considerably improved if crystals can be rotated and/or several diffraction images are collected from the same crystal during serial crystallography experiments[6].

MeshAndCollect[1] was developed with the aim of combining SX with 'standard' MX rotation data collection and allows, in an automatic fashion, experimenters to collect partial diffraction data sets from numerous crystals of the same macromolecular target mounted on the same or different sample holders. A complete diffraction data set is then obtained by merging the most isomorphous of the partial data sets collected. MeshAndCollect is compatible with any state-of-the-art synchrotron X-ray beamline for MX (ideally an insertion device facility with a relatively small (20 µm or less) beam size at the sample position). In addition to the compilation of complete data sets from a series of small, well-diffracting crystals, the method is also very suitable for the initial experimental assessment of the diffraction quality of micro-crystals and for the processing of opaque samples, *e.g., in meso* grown microcrystals of membrane proteins[9].

At the start of a MeshAndCollect experiment, the positions, in two dimensions, of each of the many crystal contained in a single sample holder are determined using a low dose X-ray scan. The diffraction images collected during this scan are automatically analyzed by the program

Figure 2.9: Article published in *Journal of Visualized Experiments* featuring the structure determination of Cerulean by a SSX approach from crystals grown using my protocol.

2.2.4 *At*Phot2LOV2: the LOV2 domain of phototropin 2 from *Arabidopsis thaliana*

Automated screening for *At*Phot2LOV2 led to one interesting hit, which was manually reproduced and improved. The best condition was found to consist of 8 $mg.mL^{-1}$ protein and 12 – 17% PEG8000, 200 mM calcium acetate and 100 mM MES pH 6, with a [mother liquor / protein] ratio of [1/1]. The size of the cube-shaped crystals never surpassed 80x80x80 μm^3, which was good enough for all our experiments. The major factor contributing to the improvement was the combination of limited proteolysis (**Fig. 2.10a**) and micro-seeding to control the nucleation rate, from hundreds of microcrystals (**Fig. 2.10b**, no dilution of the seed stock) to one single crystal (**Fig. 2.10b,** 1/1000 seed dilution) in a single drop. This accurate control of nucleation has allowed me to generate large amounts of crystals of ~50x50x50 μm^3 in a very reproducible way, which were used in the time-resolved crystallography experiments described in **Chapters 5 and 6**. Noteworthy, micro-seeding alone was not sufficient to control nucleation accurately (**Fig. 2.10c**), and needed to be coupled with limited proteolysis.

The amount of microcrystals obtained in hanging drops was too little for an SSX experiment. Large scale production was achieved using large volume sitting drops (HR3-158 plate, Hampton Research) with highly concentrated seed stocks obtained after harvesting more than 50 fresh crystals instead of less than 10 for previous seed stocks. 20 µL drops obtained with a [1:1] ratio resulted in thousands of microcrystals per drop. Pooling of all drops from a 24-well plate was enough for one to two injection rounds of the HVE injector (see **Chapter 4**).

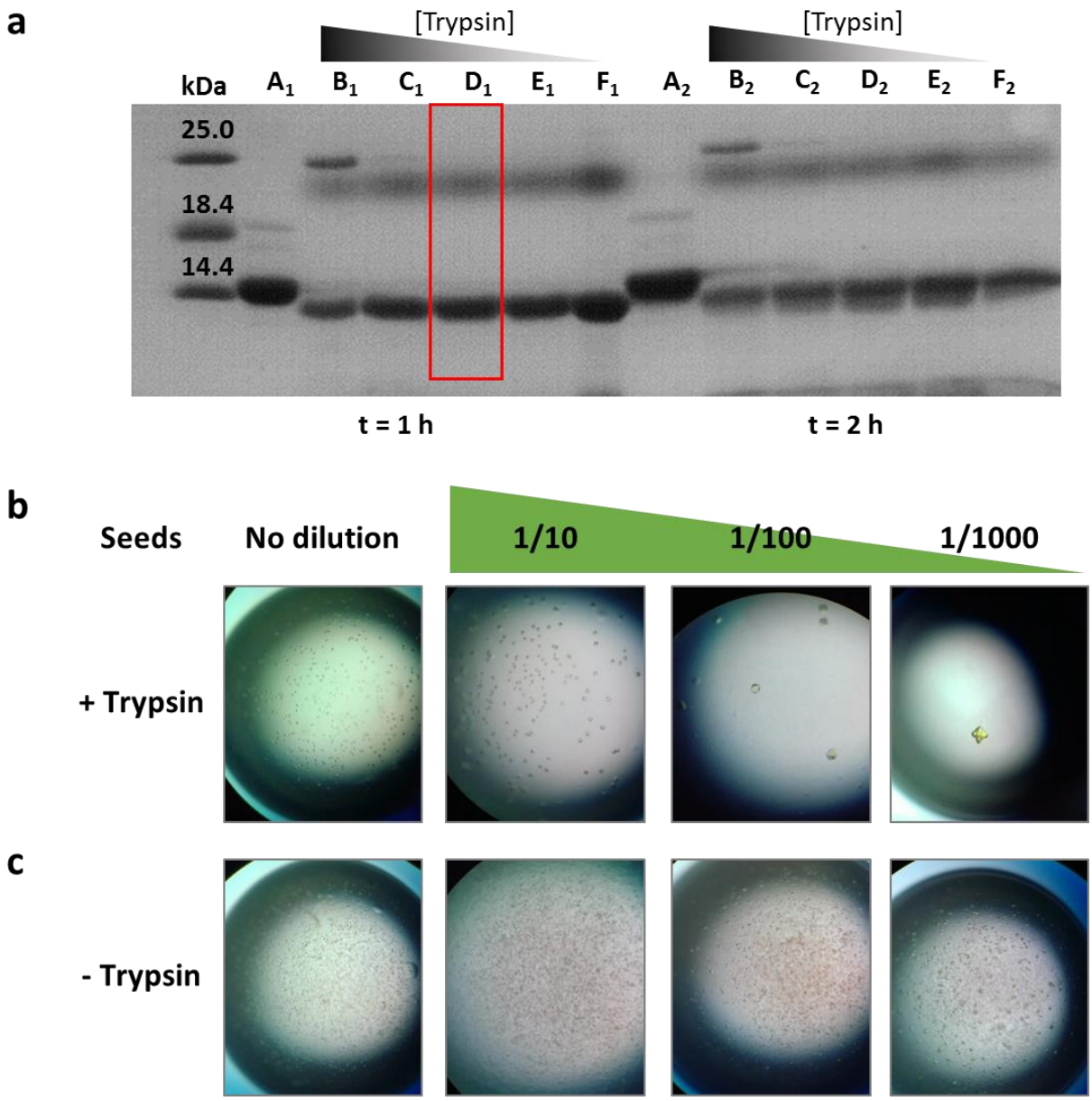


Figure 2.10: (**a**) Limited proteolysis experiment on *At*Phot2LOV2, from which the optimal trypsin concentration has been determined (red rectangle) and applied to the protein stock solution prior to crystallization. (**b**) and (**c**) Crystallization control of *At*Phot2LOV2 crystals by combining micro-seeding dilution and limited proteolysis. For comparison purposes, and in order to remove any artefact stemming from condition variability, each pair of +/- Trypsin drops (upper and lower photos) comes from the same crystallization well. Noteworthy, control of the nucleation is more accurately achieved with trypsin (upper line) than without (lower line).

2.2.5 *Dr*CBD: the chromophore-binding domain of bacteriophytochrome from *Deinoccocus radiodurans*

Automated crystallization screening led to many hits. Manual optimization of one of them eventually yielded well-diffracting crystals of a size ranging from 100 µm to 1 mm. The best condition was found to consist of 15 mg.mL^{-1} protein in 26 – 32% (v/v) 2-methyl-2,4-pentanediol (MPD), 20 mM $CaCl_2$ and 100 mM sodium acetate pH 4.0 – 5.0, with a [mother liquor / protein ratio] of [1/1]. The manual optimization included micro-seeding with the protein mixed with seeds at a 1/100 dilution. Interestingly, small changes in precipitant (MPD) concentration has a drastic effect on the localization of crystal appearance within the drop (**Fig. 2.11**). At lower concentrations, many crystals appear concentrated at the edge of the drop, while at higher concentrations, crystals are evenly distributed. This observation is a clear demonstration of the fact that the supersaturation phase (**Fig. 2.1d**) builds up faster at the edge of a drop (Terese Bergfors, personal communication). We concluded that a higher MPD concentration leads to *Dr*CBD crystals of better size homogeneity within the same drop. Using the millimetre-sized crystals thus obtained, classical diffraction experiments were successfully conducted at cryogenic and room temperature, which led to the structure determination of *Dr*CBD at 1.07 Å resolution at 100 K and at 2.0 Å resolution at 293 K (data not shown).

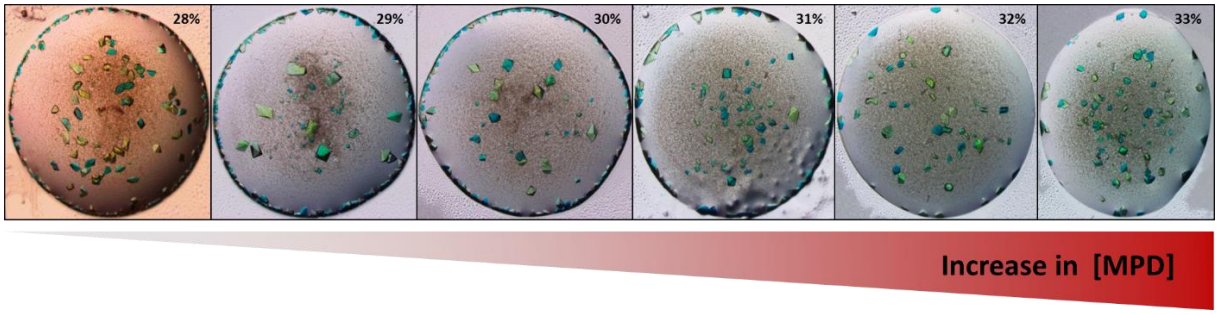


Figure 2.11: Effect of the concentration of the precipitant MPD on the crystal distribution within the 2 µL crystallization drop, which highlights that the supersaturation phase initially occurs at the edge.

In order to prepare samples for SSX experiments, I attempted to crystallize *Dr*CBD as hundreds of thousands of microcrystals ideally smaller than 10 µm. My protocol combining micro-seeding and limited proteolysis did not work for this protein. Limited proteolysis experiments resulted in a fraction of the crystals looking polycrystalline (**Fig. 2.12)**. We thus discarded these crystals for both structure determination and seed stock production. Moreover,

micro-seeding only led to the formation of a few microcrystals per drop, which cast a doubt on the stability of the seeds.

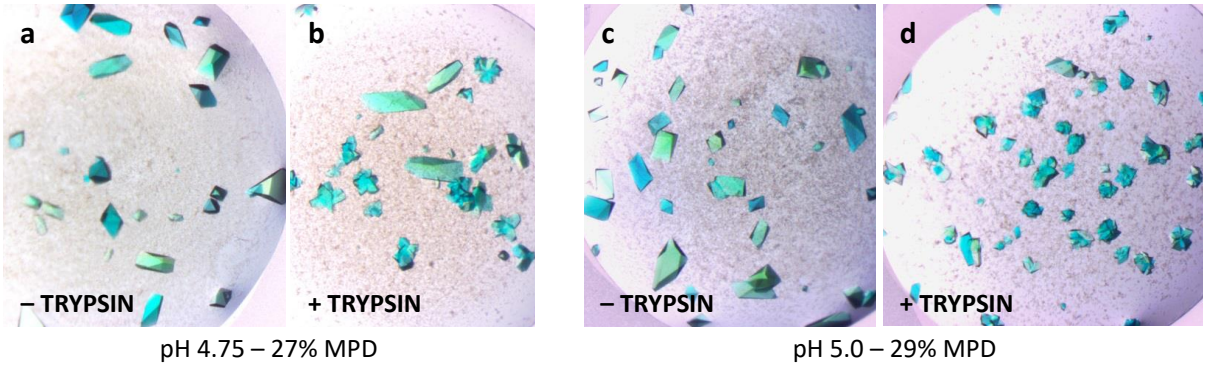


Figure 2.12: Effect of limited proteolysis on the shape of *Dr*CBD crystals in two similar crystallization conditions. (**a**) and (**b**) are two drops from the same crystallization well. *Idem* for (**c**) and (**d**). Some of the crystals in (**b**) and (**d**) are clearly polycrystalline.

I then decided to switch to the batch crystallization method (**Figure 2.1c**). Crystallization batches were first set up in microcentrifuge tubes, in a [mother liquor / protein] ratio of [1/1], with a final volume of 10 to 20 µL. Tubes were vortexed for 30 s and then stored in the dark at 293 K. Crystals reached their final size after 2 to 5 days. Various vortexing times, vortexing speeds and [mother liquor / protein] ratios were tested, resulting in many different crystal forms. For instance, the influence of vortexing can be visualized on **Figure 2.13**. The size of the obtained crystals was still heterogeneous, ranging from 1 to 500 µm (**Figure 2.14a**). Crystal filtration was considered in order to select only those with a suitable size for SSX, but as it appeared to break the crystals, it appeared difficult to optimize (by choosing from different types of filters) and was thus rejected.

I then browsed the literature to identify various crystallization conditions for *Dr*CBD, and I identified the condition of (Edlund *et al.*, 2016) as being very close to mine. I decided to add one, then two additives of the Edlund condition to mine (**Figure 2.14b**). This composite condition was then tested using the batch method in 500 µL microcentrifuge tubes. Using only PEG400 led to spherulites, but the addition of the reducing agent dithiothreitol (DTT) led to nice-looking crystals, but slightly heterogeneous in size (**Figure 2.14c**). These microcrystals were then harvested in a mesh loop at room temperature and showed diffraction spots (data not

shown), which constituted a validation step of the sample for an injection-based SSX experiment.

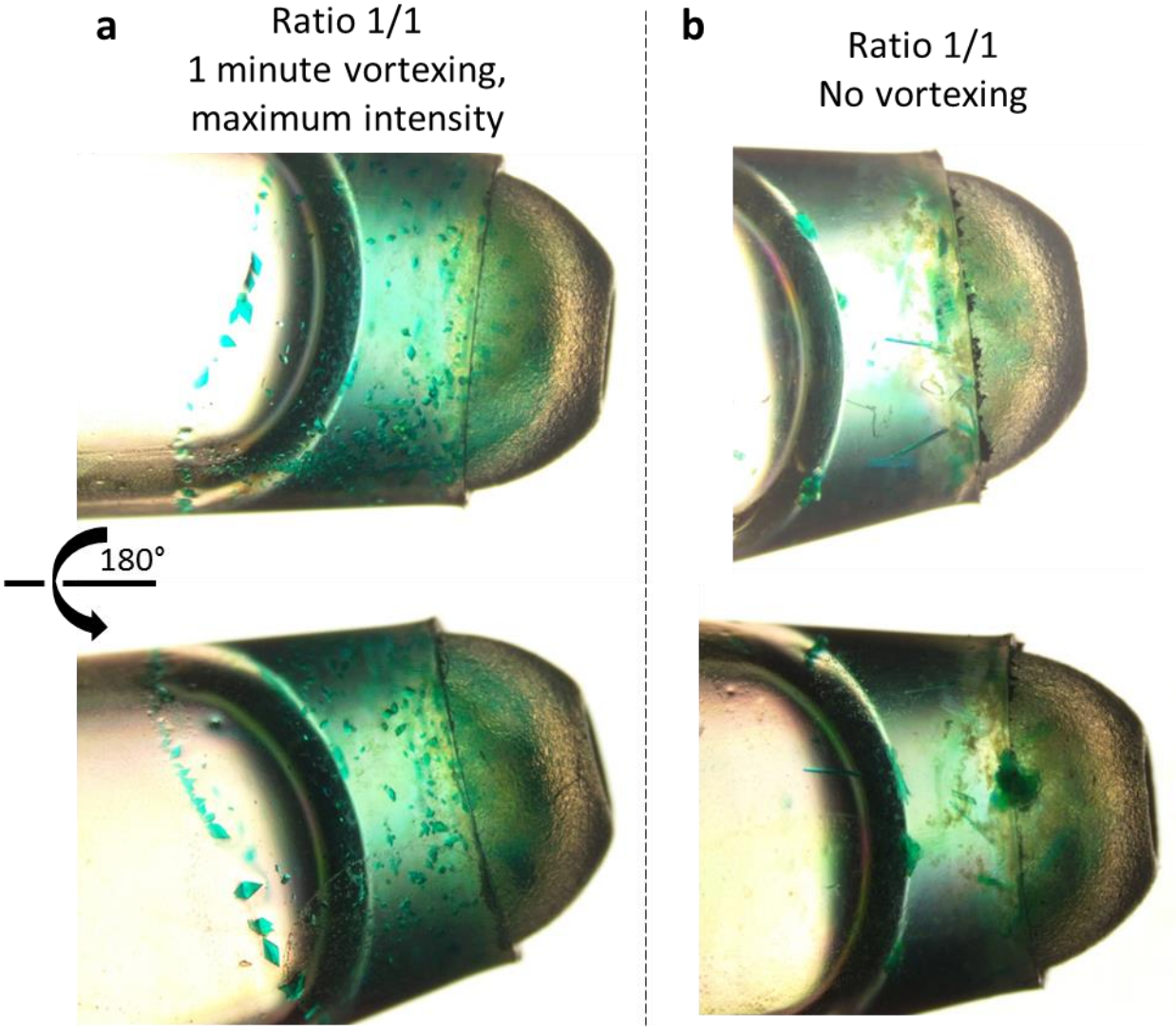


Figure 2.13: (**a**) Crystals obtained a few days after protein solution and mother liquor mixing right before vortexing, leading to bi-pyramidal shaped crystals. (**b**) *Idem*, but without vortexing, leading to needle-shaped crystals.

The homogeneity of microcrystal size still needed to be improved. The batch method was optimized by using 96-well microplates, which were placed on a slowly-rotating agitator for more than a day, in order to contrast with the high-speed, minute time scale vortexing in microcentrifuge tubes. This method was adapted from the microbatch mixing (MBM) method described by B.P. Mahon (Mahon *et al.*, 2016). Crystals thus obtained are shown on

Figure 2.14d. In the absence of agitation, large crystals grew (left panel). With a very slow agitation (~10 rpm), crystals grew smaller, but still not enough (centre panel). Finally, with a gentle agitation (~50 rpm), we obtained a homogeneous distribution of microcrystals, which were used in SSX experiments (see **Chapter 4**).

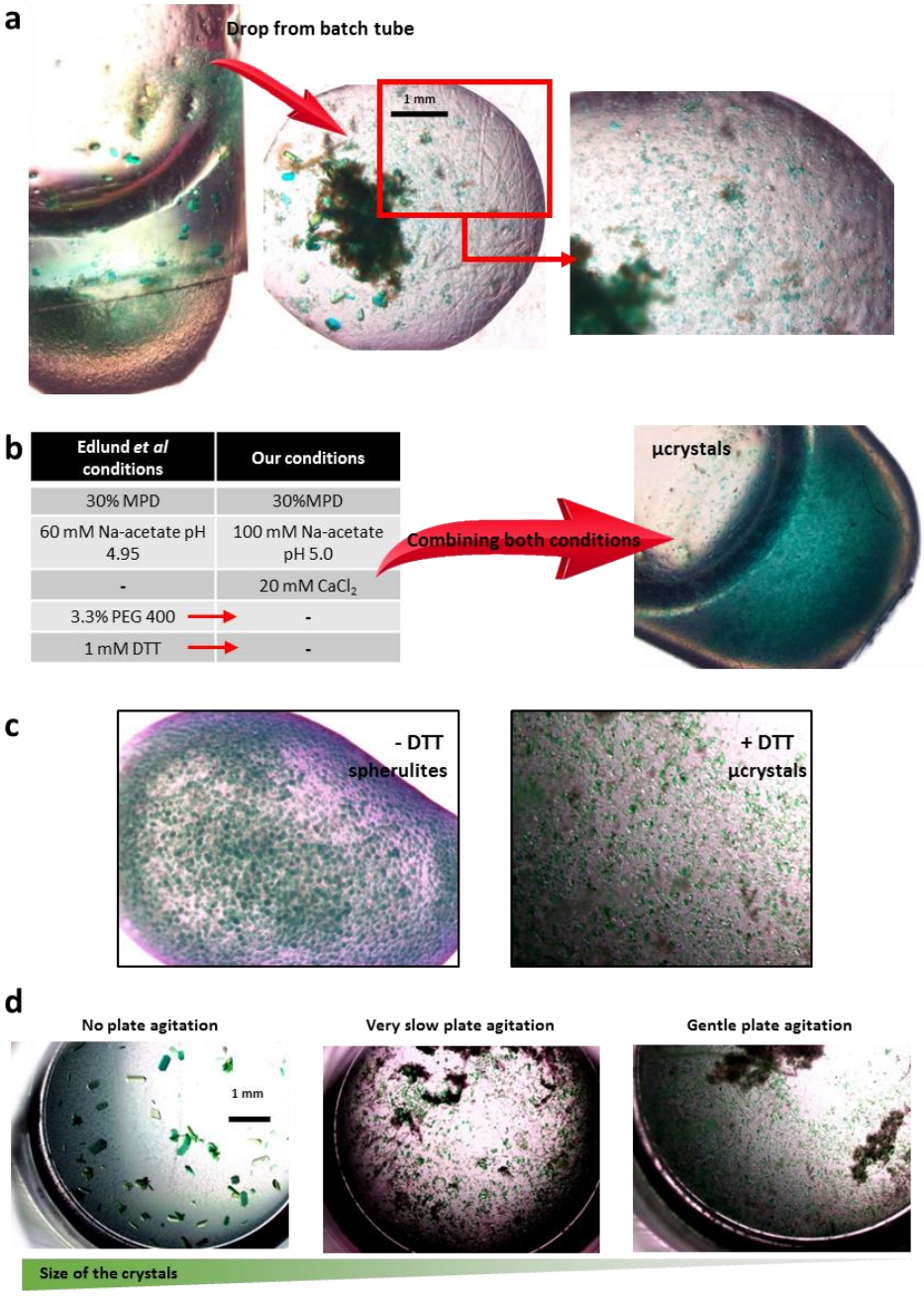


Figure 2.14: (**a**) Batch crystallization of *Dr*CBD in a microcentrifuge tube, producing crystals heterogeneous in size. (**b**) Crystallization condition optimisation by inclusion of additives identified in a literature search of *Dr*CBD crystallization conditions. (**c**) Effect of the reducing agent DTT on *Dr*CBD microcrystal shape. (**d**) Influence of plate agitation speed on crystal size using the microbatch mixing method in 96-well crystallization plates.

2.3 Crystal preservation in diffraction or spectroscopic data collection

2.3.1 Cryogenic temperature data collection

When needed, and as reference experiments, 100 K X-ray data were recorded at ESRF beamlines using a Cryostream 700 (Oxford Cryosystems, Oxford, UK), either with a Pilatus 6M detector on ID29 (De Sanctis *et al.*, 2012) or an Eiger 4M detector (Dectris) on ID30A-3 (Theveneau *et al.*, 2013). *At*Phot2LOV2 and Twist-Cerulean crystals were cryoprotected by soaking into a solution consisting of 1/5 ratio of 100% glycerol and the mother liquor. No additional cryoprotectant was needed in the case of *Dr*CBD, which already contains 26 – 32% (v/v) MPD acting as the cryoprotective condition.

2.3.2 Room temperature data collection

Room temperature data collections could be achieved thanks to an HC1 humidity controller (Sanchez-Weatherby *et al.*, 2009; Bowler *et al.*, 2015; Wheeler *et al.*, 2012) by setting the relative humidity (RH) of the wet stream to a value calculated from the mother liquor composition using Raoult's law and the Flory-Huggins model for the entropy of mixing of polymers (http://www.esrf.fr/UsersAndScience/Experiments/MX/How_to_use_our_beamlines/forms).

The RH for *Dr*CBD crystals, for which the precipitant is MPD (2-Methyl-2,4-pentanediol), a species which neither dissociates into multiple species nor is a polymer can be calculated using Raoult's law equation, where x corresponds to the mass fraction of the precipitant molecule and n is its molecular weight:

$$RH = \frac{1}{1 + \frac{x}{1-x}\frac{18}{n}}$$

*Dr*CBD crystals grew in 30% MPD (n = 118.18 Da), resulting in a RH of 93.9%.

For both Twist-Cerulean and *At*Phot2LOV2 crystals for which PEG8000 is the precipitant, the RH has been determined to be around 99.7% using a modified equation taking into account the Flory-Huggins model for the entropy of mixing of polymers, since Raoult's law starts to break down for PEGs of molecular weight above 1000 Da:

$$RH = \frac{1}{1+\frac{18x}{(1-x)m}} \exp\{(1-m/n)/(1+\frac{(1-x)m}{18x})\}$$

where *x* corresponds to the PEG mass fraction (12-20%) and *m* is the PEG molecular weight (8000 Da).

The theoretical RH values need to be checked in practice, which is done by live video monitoring of the evolution of the mother liquor volume surrounding the crystal in the loop. Shrinking or swelling indicate imperfect tuning of the humidity level, and small variations of the RH value are then manually tested.

2.4 References

Bergfors, T. (2003). *J. Struct. Biol.* **142**, 66–76.

Boivin, S., Kozak, S. & Meijers, R. (2013). *Protein Expr. Purif.* **91**, 192–206.

Bowler, M. W., Mueller, U., Weiss, M. S., Sanchez-Weatherby, J., Sorensen, T. L.-M., Thunnissen, M. M. G. M., Ursby, T., Gobbo, A., Russi, S., Bowler, M. G., Brockhauser, S., Svensson, O. & Cipriani, F. (2015). *Cryst. Growth Des.* **15**, 1043–1054.

D'Arcy, A., Bergfors, T., Cowan-Jacob, S. W. & Marsh, M. (2014). *Acta Cryst. F* . **70**, 1117–1126.

Edlund, P., Takala, H., Claesson, E., Henry, L., Dods, R., Lehtivuori, H., Panman, M., Pande, K., White, T., Nakane, T., Berntsson, O., Gustavsson, E., Båth, P., Modi, V., Roy-Chowdhury, S., Zook, J., Berntsen, P., Pandey, S., Poudyal, I., Tenboer, J., Kupitz, C., Barty, A., Fromme, P., Koralek, J. D., Tanaka, T., Spence, J., Liang, M., Hunter, M. S., Boutet, S., Nango, E., Moffat, K., Groenhof, G., Ihalainen, J., Stojković, E. A., Schmidt, M. & Westenhoff, S. (2016). *Sci. Rep.* **6**, 35279.

Foos, N., Cianci, M. & Nanao, M. H. (2019). *Acta Cryst. D.* **75**, 200–210.

Gotthard, G., von Stetten, D., Clavel, D., Noirclerc-Savoye, M. & Royant, A. (2017). *Biochemistry*. **56**, 6418–6422.

Lelimousin, M., Noirclerc-Savoye, M., Lazareno-Saez, C., Paetzold, B., Le Vot, S., Chazal,

R., Macheboeuf, P., Field, M. J., Bourgeois, D. & Royant, A. (2009). *Biochemistry*. **48**, 10038–10046.

Luft, J. R. & DeTitta, G. T. (1999). *Acta Cryst. D*. **55**, 988–993.

Mahon, B. P., Kurian, J. J., Lomelino, C. L., Smith, I. R., Socorro, L., Bennett, A., Hendon, A. M., Chipman, P., Savin, D. A., Agbandje-McKenna, M. & McKenna, R. (2016). *Cryst. Growth Des.* **16**, 6214–6221.

Sanchez-Weatherby, J., Bowler, M. W., Huet, J., Gobbo, A., Felisaz, F., Lavault, B., Moya, R., Kadlec, J., Ravelli, R. B. G. & Cipriani, F. (2009). *Acta Cryst. D*. **65**, 1237–1246.

De Sanctis, D., Beteva, A., Caserotto, H., Dobias, F., Gabadinho, J., Giraud, T., Gobbo, A., Guijarro, M., Lentini, M., Lavault, B., Mairs, T., McSweeney, S., Petitdemange, S., Rey-Bakaikoa, V., Surr, J., Theveneau, P., Leonard, G. A. & Mueller-Dieckmann, C. (2012). *J. Synchrotron Rad.* **19**, 455–461.

Theveneau, P., Baker, R., Barrett, R., Beteva, A., Bowler, M. W., Carpentier, P., Caserotto, H., De Sanctis, D., Dobias, F., Flot, D., Guijarro, M., Giraud, T., Lentini, M., Leonard, G. A., Mattenet, M., McCarthy, A. A., McSweeney, S. M., Morawe, C., Nanao, M., Nurizzo, D., Ohlsson, S., Pernot, P., Popov, A. N., Round, A., Royant, A., Schmid, W., Snigirev, A., Surr, J. & Mueller-Dieckmann, C. (2013). *J. Phys. Conf. Ser*. **425**, 012001.

Wheeler, M. J., Russi, S., Bowler, M. G. & Bowler, M. W. (2012). *Acta Cryst. F* . **68**, 111–114.

Zander, U., Bourenkov, G., Popov, A. N., De Sanctis, D., Svensson, O., Mccarthy, A. A., Round, E., Gordeliy, V., Mueller-Dieckmann, C. & Leonard, G. A. (2015). *Acta Cryst. D*. **71**, 2328–2343.

Chapter 3

RESULTS – ARTICLE 1

Comparison of the rate of specific and global radiation damage at room temperature
–
Implications for time-resolved crystallography

Presentation of Article 1

In order to envisage time-resolved crystallography experiments, we first practiced room temperature data collection, in particular by getting used to the operation of the humidity controller HC1 and how to best estimate the relative humidity for a given crystallisation condition (Paragraph 2.3.2). Because reaction intermediate states have been often identified as particularly radiosensitive at cryogenic temperature, we wondered how to monitor and limit specific damage in room temperature. Intuitively, we reasoned that since specific damage was much faster than global damage at cryogenic temperature, and since global damage at room temperature is significantly faster than at cryogenic temperature, there would be a chance that specific damage could be particularly fast at room temperature. To this end, we first attempted to characterize the rate of specific radiation damage at room temperature in proteins, for which specific radiation damage had been well characterized at cryogenic temperature: the Cyan Fluorescent Protein Cerulean and hen egg white lysozyme. Our inability to visualize any specific damage for these two proteins prompted us to derive a metric to compare the rate of specific and global damage at cryogenic and room temperature. We eventually concluded that the rates of both types of damage, which can be very different at cryogenic temperature, are close to each other at room temperature. This suggests that specific radiation damage may be a lesser problem at room than at cryogenic temperature, provided one manages to record full oscillation data sets from a single crystal. We then demonstrated that indeed, we could determine the room temperature structure of a reaction intermediate state in the photocycle of a LOV domain from a plant phototropin, yet shown to be particularly radiosensitive at cryogenic temperature. The following manuscript, which describes these results, is presented thereafter. It has been published in the July 2019 issue of *IUCrJ*.

Specific radiation damage is a lesser concern at room temperature

Guillaume Gotthard[a#], Sylvain Aumonier[a#], Daniele De Sanctis[a], Gordon Leonard[a], David von Stetten[a$], Antoine Royant[ab]*

[a]European Synchrotron Radiation Facility, Grenoble, F-38043, France

[b]Univ. Grenoble Alpes, CNRS, CEA, IBS, Grenoble, F-38000, France

[#]These authors contributed equally to the work

*Correspondence email: antoine.royant@ibs.fr

[$]Present address: European Molecular Biology Laboratory (EMBL) c/o DESY, Notkestrasse 85, D-22607 Hamburg, Germany

Synopsis We have investigated both global and specific radiation damage for three different proteins at cryogenic and room temperature. The large decoupling between global and specific radiation damage at cryogenic temperature appears to be practically abolished at room temperature, which has positive implications on time-resolved protein crystallography.

Abstract Carrying out macromolecular crystallography (MX) experiments at cryogenic temperatures significantly slows down the rate of global radiation thus facilitating the solution of high resolution crystal structures of macromolecular molecules. However, cryo-MX experiments suffer from the early onset of so-called specific radiation damage that affects certain amino acid residues and, in particular, the active sites of many proteins. Here we describe a series of MX experiments which suggest that at room temperature, specific and global radiation damage are much less decoupled than they are at cryogenic temperatures. The results reported here demonstrate the interest in reviving the practice of collecting MX diffraction data at room temperature and allow structural biologists to favourably envisage the development of time-resolved MX experiments at synchrotron sources.

3.1 Introduction

Radiation damage in macromolecular crystallography (MX) experiments is an unavoidable phenomenon. In the early days of MX this meant that it was necessary to compile a complete diffraction data set from partial data sets collected from several crystals, making

structure solution a very lengthy and arduous process (Perutz *et al.*, 1960; Kendrew *et al.*, 1960; Blake *et al.*, 1965). The situation changed dramatically in the 1990s when the advent of cryo-cooling (Hope, 1990; Garman & Schneider, 1997) greatly increased the absorbed dose a crystal could tolerate before radiation damage destroyed a crystal's diffraction properties. This allowed the collection of complete, high resolution diffraction data sets from a single crystal and paved the way for the explosion in the number and type of macromolecular crystal structures that have been determined in the intervening quarter of a century. However, in the early 2000s a series of seminal papers (Ravelli & McSweeney, 2000; Burmeister, 2000; Weik *et al.*, 2000) showed that it was not only global radiation damage that was the enemy of the crystallographer. Indeed, two types of radiation damage could be defined: (*i*) global damage, observed mainly in reciprocal space and corresponding to a degradation of the diffraction properties of a crystal and, (*ii*) specific damage, mainly observed in the real space and corresponding to the photoreduction of metal centres, the photoreduction of photoactive protein chromophores, the breakage of covalent bonds (e.g. disulphide bridges) and/or the loss of electron density for the side-chains of some amino acids. These two phenomena are largely decoupled at cryogenic temperature. While acceptable absorbed dose limits for global radiation damage have been proposed to be in the range 20 – 30 MGy (Henderson, 1995; Owen *et al.*, 2006), specific radiation damage occurs at much lower absorbed doses (e.g. ~7 kGy for the peroxo group in the active site of a reaction intermediate state of urate oxidase (Bui *et al.*, 2014), ~ 60 kGy for the alteration of the chromophore in bacteriorhodopsin (Borshchevskiy *et al.*, 2014) and ~ 1.6 MGy for the breakage of disulphide bonds in lysozyme (Carpentier *et al.*, 2010)). This decoupling of global and specific radiation damage often means that extreme care should be taken when interpreting the results of MX–determined crystal structures and experimenters must be sure that the electron density maps produced do not contain artefacts due to specific radiation damage. Most often this means that MX should be complemented with *in crystallo* optical spectroscopy (*ic*OS) (von Stetten *et al.*, 2015) or other measurements which provide information on the changes in the chemical state of a macromolecule that can be induced by exposure to X-rays.

In addition to the potential artefacts induced by specific radiation damage, a second potential pitfall of cryo-crystallography is the fact that the cryo-cooling process can artificially trap biologically inactive conformations of amino acid side chains. This can also lead to misinterpretation of enzyme mechanisms and of the roles of particular amino acids in specific biological processes (Fraser *et al.*, 2011). For this reason, room-temperature (RT) MX

experiments are experiencing something of a renaissance. However, at RT the absorbed doses that induce global radiation damage can be 2 orders of magnitude lower than those at cryogenic temperatures (Nave & Garman, 2005; Southworth-Davies *et al.*, 2007) and a thorough characterization of the radiation sensitivity of the crystals under study should usually be carried out if a complete data set is to be collected from a single crystal at room temperature. The question of whether dose-rate has an effect on radiation sensitivity at room temperature has been extensively investigated but did not receive a clear answer for dose-rates covering 50 $Gy.s^{-1}$ to 680 $kGy.s^{-1}$ (Southworth-Davies *et al.*, 2007; Rajendran *et al.*, 2011; Owen *et al.*, 2012; Warkentin *et al.*, 2012; Leal *et al.*, 2013). These works suggest however that the total absorbed dose had to be limited to the hundreds of kGy dose scale in order to record a complete data set, depending on the solvent content of the crystal (Leal *et al.*, 2013).

In the course of studying the chromophore photoswitching behaviour of the fluorescent protein Cerulean (Gotthard *et al.*, 2017), we attempted to solve its room temperature structure. Cerulean possesses a radiosensitive glutamate residue next to the chromophore, whose side chain is severely affected by specific radiation damage (decarboxylation) during cryo-MX experiments. We thus expected this residue to be severely affected by specific radiation damage at RT. However, and to our surprise, the RT crystal structure we obtained from a diffraction data set collected from a single crystal showed no sign of specific radiation damage. This prompted us to perform a more extensive study aimed at investigating specific radiation damage in room temperature MX experiments. This topic has first been considered by Southworth-Davies and coworkers when investigating the effects of various dose rates at room temperature (Southworth-Davies *et al.*, 2007) and then by Russi and coworkers in the course of comparing structural heterogeneity at room and cryogenic temperature (Russi *et al.*, 2017). In both studies, they observed that specific radiation was less obvious at room than at cryogenic temperature. In this study, we recorded both X-ray diffraction data and, in order to obtain X-ray independent estimates of specific radiation damage, complementary data using Raman or UV-vis absorption spectroscopy. In addition to Cerulean, two other systems were studied: lysozyme, in order to examine the photoreduction of disulphide bonds at RT, and the photoadduct of the LOV2 domain of phototropin 2 from *Arabidopsis thaliana*, which upon light activation forms a covalent bond between the protein and its co-factor that is particularly sensitive to specific radiation damage at cryogenic temperatures. The results reported here suggest that specific and global radiation damage are much less decoupled at RT than they are at cryogenic temperature, thus confirming the interest of reviving the practice of collecting

diffraction data at RT. They also allow structural biologists to favourably envisage the development of time-resolved MX experiments at synchrotron sources.

3.2. Methods

3.2.1 Protein expression and purification

Cerulean was overexpressed and purified using a previously described protocol (Lelimousin *et al.*, 2009; Gotthard *et al.*, 2017). HEWL was purchased from Roche Applied Science (catalogue No. 10837059001) and was dissolved in distilled water to the concentration of 40 $mg.ml^{-1}$. The gene coding for the LOV2 domain of the phototropin 2 from *Arabidopsis thaliana* was synthetized (Genecust, Ellange, Luxemburg) and inserted in a pBad plasmid. The plasmid was transformed into a *Escherichia coli* BL21 strain and cells were grown in 2 L of ZYP-5052 auto-inducible medium (Studier, 2005) at 37 °C until OD_{600nm} reached 1.25. Protein expression was then induced with 0.2% L-arabinose for 14 hours at 17 °C. Cells were harvested by centrifugation (20 min at 4000*g*) and pellets were resuspended in 25 mL of a lysis buffer comprising 50 mM Tris pH 8.0, 300 mM NaCl, 10 mM imidazole, 0.25 $mg.mL^{-1}$ lysozyme, 400 $\mu g.mL^{-1}$ DNAse I, 20 mM $MgSO_4$, and anti-protease cocktail (Complete EDTA-free, Roche), and frozen at -80 °C. Resuspended pellets were sonicated 4 times for 30 s at 35 W power (VC-750 ultrasonic processor, Bioblock Scientific) and cell debris was harvested by centrifugation (40 min at 15000*g* at 4 °C). The protein was purified from the clarified lysate using a nickel affinity column (HisTrap HP 5 mL, GE Healthcare, Little Chalfont, UK) followed by size exclusion chromatography (Superdex 75 10/300 GL, GE Healthcare, Little Chalfont, UK). Purified *At*Phot2LOV2 was concentrated to 5 $mg.mL^{-1}$ and submitted to digestion with trypsin (1 h, ratio of 1:100) prior to crystallization.

3.2.2 Protein crystallization

Cerulean was crystallized as previously described (Lelimousin *et al.*, 2009; Gotthard *et al.*, 2017) using the hanging drop vapour diffusion method (1:1 ratio in 2 µL drops) at 293 K with a protein concentration of 13 $mg.mL^{-1}$ in a condition consisting of 10 – 20% PEG8000, 100 mM $MgCl_2$, and 100 mM HEPES pH 6.75 – 7.5. Needle shaped crystals grew in 5 days and were used to seed subsequent optimized crystallization conditions (10 – 12 % PEG8000) by mixing the protein solution with the seed solution in a 1:10 or 1:100 ratio. Three dimensional

rod-type crystals then appeared after incubation for one week at 293 K. HEWL was crystallized using the sitting-drop vapour diffusion method (1:1 ratio in 2 μL drops) in a crystallization condition consisting of 250 – 400 mM NaCl, 100 mM sodium acetate pH 4.8. Crystals belonging to the tetragonal space group $P4_32_12$ grew at 293 K within one week. *At*Phot2LOV2 was crystallized using the hanging drop vapour diffusion method (1:1 ratio in 2 μL drops) at 293 K with a protein at the concentration of 5 $mg.mL^{-1}$. Crystals appeared after 2 days in a condition consisting of 12 – 17% PEG8000, 200 mM calcium acetate, 100 mM MES pH 6.0.

3.2.3 X-ray data collection

Crystals were cryoprotected by transfer into a solution consisting of the reservoir solution diluted with 20% (v/v) glycerol (99.5% grade). X-ray data sets were recorded either at 100 K using an Oxford Cryostream 700 cryogenic system (Oxford Cryosystems, Oxford, UK) or at room temperature using an HC1 humidity controller with a humidity level calculated from the composition of the mother liquor (Sanchez-Weatherby *et al.*, 2009). Crystal sizes and data collection parameters are summarized in **Table 1**. X-ray data collections were carried out on the ESRF beamlines ID29 (De Sanctis *et al.*, 2012), ID30B (McCarthy *et al.*, 2018) and ID30A-3 (Theveneau *et al.*, 2013), which all have a Gaussian beam profile with an ellipsoid or circular shape. Data collection statistics are reported in **Table 3**.

In order to perform the RT diffraction experiment on the unstable (in time, and potentially in dose) photoadduct of *At*Phot2LOV2, we mounted a single crystal on the goniometer of ESRF beamline ID30A-3 under the air flux of an HC1 humidity controller, then illuminated it with a 470 nm blue LED for 10 minutes in order to populate the crystal with the covalent intermediate. Upon starting data collection (**Table 3**) the LED was turned off (process triggered by the trigger of the beamline's Eiger 4M detector) and 1650° of oscillation data (corresponding to 4 complete data sets) were collected in a contiguous manner in a total of 10 seconds. Diffraction images corresponding to each of the 4 data sets were then processed as described below.

Data collection parameters

Protein	Temp. (K)	Crystal dimensions (µm)			Relative humidity level (%)	Unit cell (Å)	Beamline	Beam size† (µm x µm)	Energy (keV)	Initial flux (ph.s^{-1})	Wedge (°)	Exposure time (s)
		X	Y	Z		a, b, c						
Cerulean	100	92	194	106	-	51.1 62.7 70.3	ID29	30x40	12.7	2.3e+11	90 - 290	40
	293	170	230	170	99	51.9 63.0 71.3	ID29	30x40	12.7	2.7e+10	90 - 290	40
Lysozyme	100	75	93	75	-	77.6 77.6 37.1	ID30B	40x40	12.7	3.4e+10	12 - 102	18
	293	218	379	202	98	79.2 79.2 38.1	ID29	30x40	11.5	4.1e+10	0 - 100	20
***At*Phot2LOV2 (dark state)**	100	50	50	50	-	40.1 40.1 131.5	ID30A-3	15x15	12.8	3.5e+11	0 - 107	43
	293	50	50	50	98	40.9 40.9 132.7	ID30A-3	15x15	12.8	1.4e+11	0 - 120	14
***At*Phot2LOV2 (light state)**	100	75	76	82	-	40.3 40.3 131.3	ID29	30x40	12.7	1.1e+10	87 - 177	36
	293	70	70	70	98	41.5 41.5 133.5	ID30A-3	15x15	12.8	2.0e+11	0 - 360	2.2

† The size is defined as the vertical and horizontal full-width at half maximum (FWHM) of the elliptical-shape beam. Collimation parameters were set at 3 times the FWHM in each direction.

3.2.4 Data reduction and structure refinement

X-ray data sets were integrated, scaled and merged using the XDS program suite (Kabsch, 2010). Absorbed doses corresponding to the average dose of the exposed region were calculated using *RADDOSE-3D* (Zeldin *et al.*, 2013). The 1.15 Å structure of Cerulean (PDB ID: 2wso, Lelimousin *et al.*, 2009), the 1.2 Å structure of lysozyme (PDB ID: 5ebh, Zander *et al.*, 2016) or the 1.7 Å structure of *At*Phot2LOV2 (PDB ID: 4eep, Christie *et al.*, 2012) were used as starting models for molecular replacement or refinement. Models were manually rebuilt and refined using *Coot* (Emsley *et al.*, 2010) and *REFMAC*5 (Murshudov *et al.*, 2011), respectively. In order to quantify the extent of specific damage from crystallographic data, the default refinement protocol of *REFMAC5* was modified by removing the restraint on *B*-factor continuity along covalent chains of atoms so that the *B*-factor of an atom for which electron density disappears completely can refine to a value as high as 999 Å^2 (Murshudov *et al.*, 2011). Data collection and refinement statistics are shown in **Table 3**. Structural coordinates and structure factors were deposited in the PDB (http://www.pdb.org/) with the following accession

codes 6QQ8, 6QQ9, 6QQA, 6QQB, 6QQC, 6QQD, 6QQE, 6QQF, 6QQH, 6QQI, 6QQJ, 6QQK and 6QSA. Decays of the B_0 / B_n ratios as a function of dose were modelled with a monoexponential decay behaviour A + B * exp[-dose / τ] where τ is the life-dose, or, when necessary, as a biexponential decay behaviour A + B * exp[-dose / τ_1] + C * exp[-dose / τ_2].

3.2.5 Online Raman spectroscopy

Online Raman spectroscopy was performed on beamline ID29 as previously described using a setup specifically designed for the collection of X-ray and Raman data in an interleaved manner (von Stetten *et al.*, 2017). In brief, Raman spectra were recorded using an inVia Raman instrument (Renishaw plc, Wotton-under-Edge, Gloucestershire, UK) equipped with a near-infrared (785 nm) 300 mW diode laser source. Raman spectra were measured from the X-ray exposed region of a static lysozyme crystal with a composite acquisition time of 10 x 10 s for the 300 to 1800 cm^{-1} spectral window. Spectra were corrected for background using the *WiRE*^{TM} software version 3.4 (Renishaw plc, Wotton-under-Edge, Gloucestershire, UK). X-ray burn cycles were performed in-between Raman data sets by opening the shutter for increasing amounts of time, but no diffraction data were recorded. Absorbed doses were between 39 kGy to 1.21 MGy and 1.2 kGy to 283 kGy at 100 and 293 K, respectively. The damage on disulphide bonds was evaluated by integrating the peak height for the 510 cm^{-1} band over the course of incremental dose absorption, whose evolution with dose was modelled with a monoexponential decay behaviour A + B * exp[-dose / τ].

3.2.6 *In crystallo* UV-visible absorption spectroscopy of LOV2

Offline UV-vis absorption spectra of an *At*Phot2LOV2 crystal were recorded at room temperature on the ESRF ID29S-Cryobench microspectrophotometer (von Stetten *et al.*, 2015) using a high-sensitivity fixed-grating QE65Pro spectrophotometer with a back-thinned CCD detector (Ocean Optics, Dunedin, FL, USA), a balanced deuterium-halogen DH2000-BAL light source (Ocean Optics, Dunedin, FL, USA) and the humidity controller HC1 (Sanchez-Weatherby *et al.*, 2009), before and after illumination with a 470 nm LED at 28 $W.cm^{-2}$. Spectra were averaged from ten 400 ms acquisitions on ~50 µm thick crystals.

Online UV-vis absorption spectra were recorded on beamline ID30A-3 of the ESRF using a dedicated microspectrophotometer (McGeehan *et al.*, 2009) and the same

spectrophotometer and lamp on a *At*Phot2LOV2 crystal illuminated for 5 min with a 470 nm LED at 0.7 $W.cm^{-2}$ before being flash-cooled at 100 K, and subsequent X-ray irradiation. Spectra were recorded at 5 Hz (200 ms acquisition time) on a ~50 µm thick crystal maintained in a fixed position. Each spectrum recorded under X-ray irradiation corresponded to an absorbed dose of 14 kGy.

3.3 Results

3.3.1 Cerulean

Fluorescent proteins of the Green Fluorescent Protein family are beta-barrel-shaped proteins containing a fluorescent chromophore formed from the autocatalytic cyclisation of three consecutive amino acid residues and, as such, they provide convenient genetically-encoded fluorescent reporters of localization or interaction *in cellulo* (Tsien, 1998). Tuning the properties of fluorescent proteins (colour, brightness, pH sensitivity, *etc*.) by mutagenesis has been facilitated by the availability of crystal structures of fluorescent proteins, all determined at cryogenic temperature. Nevertheless, RT crystal structures have also become available based on diffraction data collected either at synchrotrons (Kaucikas *et al.*, 2015) or at XFELs (Colletier *et al.*, 2016; Coquelle *et al.*, 2018; Hutchison *et al.*, 2017). The sensitivity of fluorescent proteins to specific X-ray induced radiation damage at cryogenic temperature has been well documented showing that the conserved glutamate residue close to the chromophore (Glu222 according to the GFP sequence) appears specifically affected at relatively low absorbed doses (*i.e.* ~ 0.1, 0.8 and 0.1 MGy for IrisFP (Adam *et al.*, 2009), EGFP (Royant & Noirclerc-Savoye, 2011) and mNeonGreen (Clavel *et al.*, 2016), respectively). Given the sensitivity of this residue in fluorescent proteins to specific radiation damage, the Cyan Fluorescent Protein Cerulean (Rizzo *et al.*, 2004), which produces well-diffracting crystals (Lelimousin *et al.*, 2009; Gotthard *et al.*, 2017), was chosen as the first target in the work described here.

3.3.1*a* Radiation damage to Cerulean at cryogenic temperature (100 K)

In order to investigate specific damage in crystals of Cerulean at 100 K, 20 consecutive data sets, corresponding to accumulated doses between 290 kGy and 5.8 MGy were recorded from the same position of a single crystal (**Table 3**). As expected from studies of other

fluorescent proteins (Adam *et al.*, 2009; Royant & Noirclerc-Savoye, 2011), Cerulean is sensitive to specific radiation damage at cryogenic temperatures and the progressive decarboxylation of Glu222 was observed in ($2mF_{obs(i)}$-$DF_{calc(i)}$, $\alpha_{calc(i)}$) electron density maps (**Fig. 1*a***) and in the ($F_{obs(i)} - F_{obs(1)}$, $\alpha_{calc(1)}$) Fourier difference map (**Fig. 1*b***). The progressive loss of this group results in the rotation of the side chain of Thr65, whose O_γ is originally engaged in a hydrogen bond with the carboxylate group of Glu222 (**Fig. 1*c***, left) and becomes engaged in a hydrogen bond with the carbonyl group of Thr61 (**Fig. 1*c***, right). Two water molecules next to the Thr-Trp-Gly chromophore are also affected by the decarboxylation of Glu222 as evidenced by the negative peaks in the ($F_{obs(20)} - F_{obs(1)}$, $\alpha_{calc(1)}$) map (**Fig. 1*b***). Examination of the ($2mF_{obs(20)} - DF_{calc(20)}$, $\alpha_{calc(20)}$) electron density maps led us to replace these two water molecules by a linear carbon dioxide molecule, the obvious result of the photodecarboxylation process (**Fig. 1*d***). In order to compare the speed of global and specific radiation damage, we derived doses characteristic of each phenomenon from diffraction data and refined structure analysis.

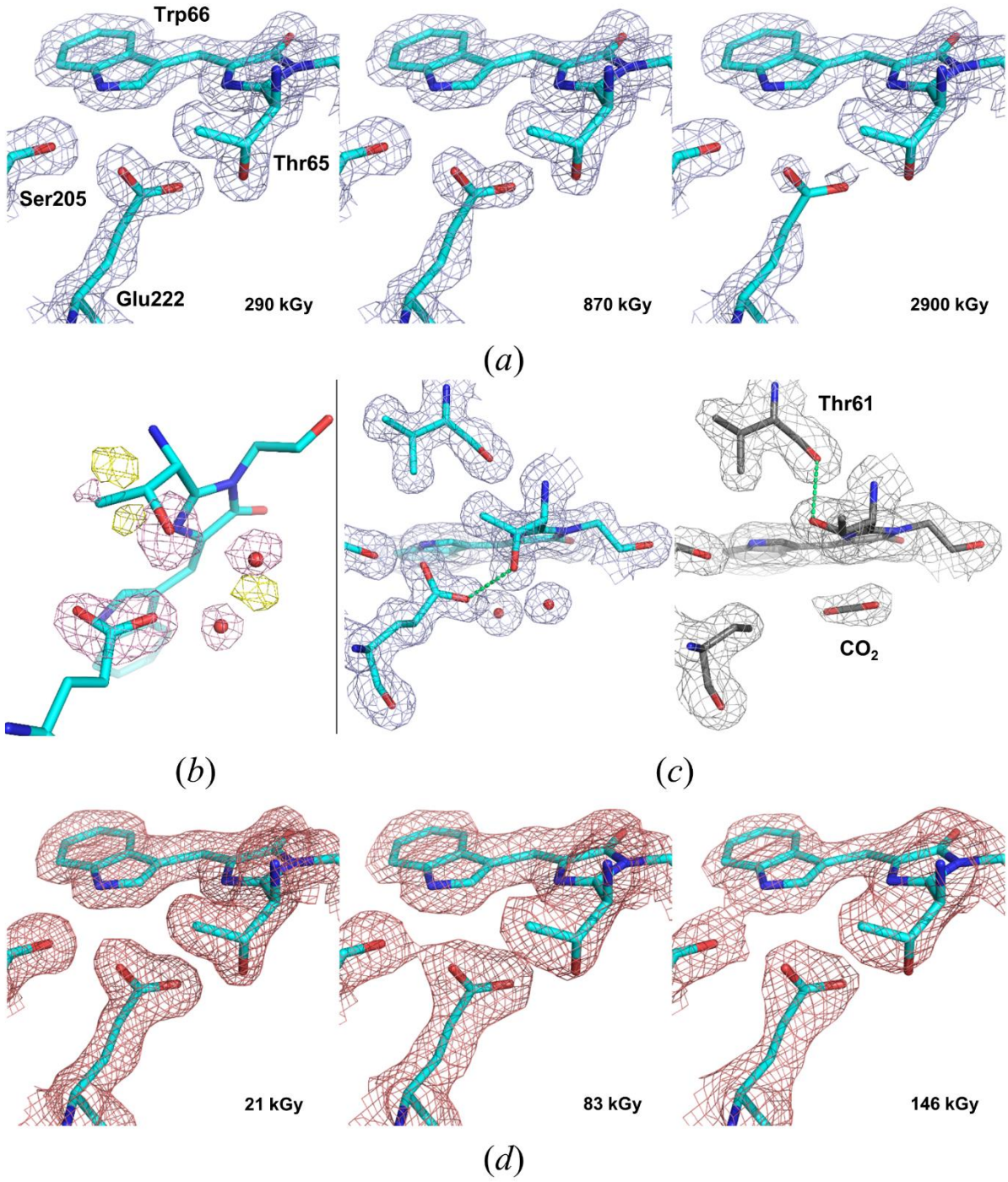


Figure 1: Evolution of the (2mF_{obs} – DF_{calc}, α_{calc}) electron density map in data sets from a Cerulean crystal recorded at increasing absorbed doses at cryogenic and room temperature. (*a*) Series recorded at 100 K (maps contoured at a 1.5 σ level). (*b*) ($F_{obs(20)} - F_{obs(1)}$, $\alpha_{calc(1)}$) Fourier difference map calculated between the final and the initial 100 K data sets highlighting the specific radiation damage to Glu222 and its structural consequences (maps contoured at a ±5.0 σ level). (*c*) (2mF_{obs} – DF_{calc}, α_{calc}) electron density map for the first (left) and the last (right) 100 K data sets illustrating the decarboxylation process of Glu222 and the rotation of Thr65 (maps contoured at a 1.0 σ level). (*d*) Series recorded at 293 K (maps contoured at a 1.5 σ level).

Diffraction resolution and B_{Wilson} are common markers of global radiation damage (Ravelli & McSweeney, 2000). Indeed, at 100 K the resolution of the data sets progressively decreased from d_{min} = 1.46 Å at an absorbed dose of 290 kGy to d_{min} = 1.82Å at an absorbed dose of 5800 kGy, which amounts to a decrease in the diffraction power of the crystal. Owen and coworkers derived a half-dose $D_{1/2}$ characteristic of global radiation damage by modelling a linear decay of the total diffracted intensities as a function of dose (Owen *et al.*, 2006), which we also derived for our crystals (**Table 2**). In addition to this diffracted-intensity based metric, we derived a second dose constant from the evolution of $B_{Wilson}(n=1)/B_{Wilson}(n)$ that we modelled as a monoexponential decay (blue trace in **Fig. 2**). We obtained a B_{Wilson}-derived life-dose of $\tau_{Glob\text{-}CT}$ = 17.6 MGy (CT standing for 'cryogenic temperature'), which we chose as a metric of global radiation damage of Cerulean at 100 K. Both values (**Table 2**) are consistent with the Henderson and Garman absorbed dose limits (20 and 30 MGy, respectively) for crystals of biological macromolecules (Henderson 1995; Owen *et al*. 2006).

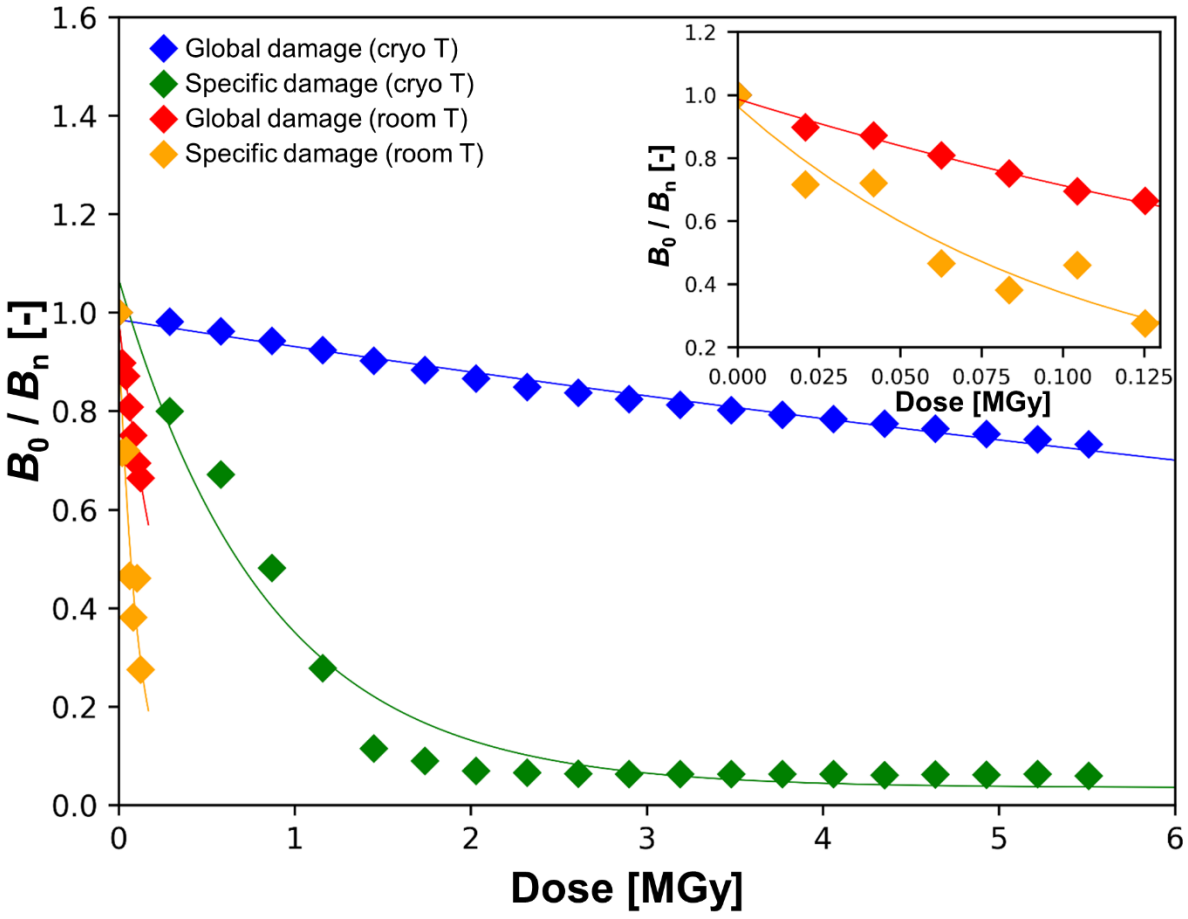


Figure 2*:* Evolution of B-factors as a function of dose for the irradiation series of Cerulean crystals at cryogenic and room temperature. The evolution of the Wilson B-factor (blue: 100 K, red: 293 K) represents the global radiation damage and the evolution of the atomic B-factor of Glu222-C_δ (green: 100 K, orange: 293 K) illustrates the specific radiation damage to the carboxylate group of Glu222. A close-up on the low dose range is presented in the inset.

To evaluate the characteristic dose at which specific radiation damage occurs at 100 K in crystals of Cerulean, we analysed the evolution of atomic *B*-factors of the progressively radiolysed carboxylate group of Glu222. C_δ is the only atom in the carboxylate group for which electron density eventually fully disappears (**Fig. 1*a***). All 20 structures were refined as if Glu222 had not been affected, with the smooth variation restraint on *B*-factors of atoms implicated in consecutive covalent bonds (Murshudov *et al.*, 2011) specifically removed. The evolution of $B_{\text{Glu222-O}\delta}(n=1)/B_{\text{Glu222-O}\delta}(n)$ was then modelled as a monoexponential decay resulting in a life-dose for specific radiation damage in crystals of Cerulean at 100 K ($\tau_{\text{Spec-CT}}$) of 843 kGy (green trace in **Fig. 2**).

The life-doses derived above indicate a strong decoupling between global and specific radiation damage in crystals of Cerulean at cryogenic temperature. In order to quantify this decoupling effect, we introduce the ratio between the global, and the specific radiation damage life-doses ($\Delta_{\text{G/S-CT}} = \frac{\tau_{\text{Glob-CT}}}{\tau_{\text{Spec-CT}}}$), which we estimated at 20.8 for Cerulean at cryogenic temperature. $\Delta_{\text{G/S}}$ is thus the decoupling factor between global and specific radiation damage for a given protein at a given temperature. Life-doses and decoupling factors are summarized in **Table 2**.

Table 2 Global and specific radiation damage dose constants

Protein	**Cerulean**		**Lysozyme**		***At*Phot2LOV2 photoadduct**	
Temperature (K)	100	293	100	293	100	293
$\boldsymbol{D_{1/2}}$[†] **(kGy)**	8790	134	9190	105	32700	169
$\boldsymbol{\tau_{Glob}}$[‡] **(kGy)**	17600	308	18200	251	34300	389
$\boldsymbol{\tau_{Spec}}$[‡] **(kGy)**	843	105	1475, 10500	198	22, 387	49
$\boldsymbol{\Delta_{G/S}}$	20.8	2.9	12.4	1.3	1590	8.0

[†] The half-dose $D_{1/2}$ was calculated as in Owen *et al*., 2006, by modelling the total diffracting power of a crystal as a function of absorbed dose as a linear decay.

[‡] The life-dose τ_{Glob} was calculated by modelling the decay of the Wilson *B*-factor of the crystal as a function of absorbed dose as a monoexponential decay. The life-dose τ_{Spec} was calculated by modelling the decay of the atomic *B*-factor of displaced atom at the position of the specific damage as a function of absorbed dose as a monoexponential decay. The specific damage curve for HEWL at 100 K is best fit by two exponential decays, as already observed (Carpentier *et al.*, 2010), which suggested the existence of a X-ray induced repair mechanism. The same is likely to apply for the specific damage to the covalent bond in the *At*Phot2LOV2 photoadduct.

3.3.1*b* Radiation damage to Cerulean at room temperature (293 K)

We performed a similar cumulative absorbed dose experiment on a Cerulean crystal maintained at room temperature during data collection (**Table 3, Figs. 1*d*** and **2**). We collected 7 data sets with accumulated doses ranging between 21 kGy and 146 kGy. As expected, there is a rapid drop-off of diffraction resolution (from d_{min} = 1.66 Å for the first, to d_{min} = 2.45 Å for the last data set). The life-dose for global radiation damage at room temperature $\tau_{Glob\text{-}RT}$ was calculated to be 308 kGy (red trace in **Fig. 2**), 57 times lower than observed at 100 K and consistent with previous estimations concerning global radiation damage rates at room and cryogenic temperatures (decrease factor between 26 and 113) (Nave & Garman, 2005; Southworth-Davies *et al.*, 2007). Intriguingly however, ($2mF_{obs} - DF_{calc}$, α_{calc}) electron density maps calculated from the successive cumulative dose data sets showed little, if any, sign of specific radiation damage (*i.e.* decarboxylation of Glu222) (**Fig. 1*d***). More surprisingly, calculation of successive ($F_{obs(n)} - F_{obs(1)}$, $\alpha_{calc(1)}$) Fourier difference maps did not show the build-up of strong peaks with increasing values as is generally observed in the analysis of a cryogenic radiation damage data set series. This prevented us from comparing the rate of specific damage at both temperatures using a Fourier difference map-based approach (Carpentier *et al.*, 2010; Bury & Garman, 2018). Based on the evolution of the normalized atomic *B*-factors of the atoms comprising the carboxyl group of Glu222, we derived the radiation damage life-dose for this moiety at room temperature ($\tau_{Spec\text{-}RT}$) (orange trace in **Fig. 2**) to be 105 kGy, a value relatively close to that of the global damage. This leads to a decoupling factor of $\Delta_{G/S\text{-}RT}$ = 2.9 for Cerulean at room temperature, which represents a 7-fold reduction of the radiation damage decoupling between cryogenic and room temperature. This observation is consistent with the lack of obvious visible evidence of specific radiation damage in crystals of Cerulean at room temperature. To investigate whether this observation is an isolated phenomenon, similar experiments were carried out on crystals of Hen Egg-White lysozyme, an archetypal radiation damage test case.

3.3.2 Hen Egg-White Lysozyme

Hen Egg-White Lysozyme (HEWL) is one of the first proteins in which specific radiation damage was identified (Weik *et al.*, 2000; Ravelli & McSweeney, 2000). The most prominent damage occurs on its four disulphide bonds, which lengthen due to the formation of an anionic radical and eventually break (Weik *et al.*, 2002; Sutton *et al.*, 2013). Raman

spectroscopy provides an X-ray-independent metric of damage to the disulphide bonds through the monitoring of the Raman S-S stretching mode at 510 cm^{-1} (Carpentier *et al.*, 2010). The bi-exponential evolution of its decay has revealed the presence of an X-ray-induced repair mechanism, in which the anionic radical can revert back to the oxidized state upon further irradiation.

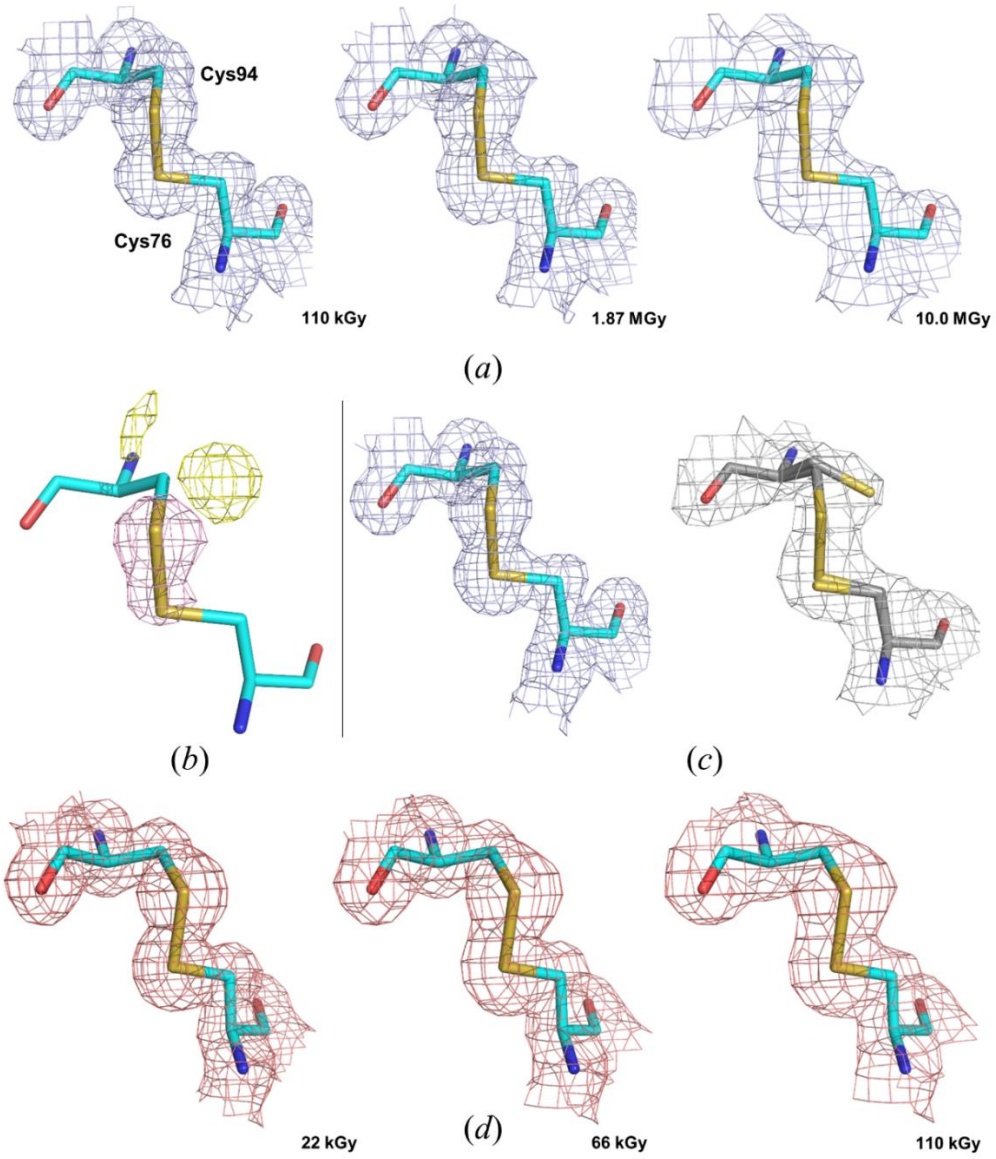


Figure 3: Specific radiation damage in crystals of HEWL. Evolution of the ($2mF_{obs} - DF_{calc}$, α_{calc}) electron density map in data sets of lysozyme recorded with increasing doses at cryogenic and room temperature. (*a*) Series recorded at 100 K (maps contoured at a 0.6 σ level). (*b*) ($F_{obs(6)} - F_{obs(1)}$, $\alpha_{calc(1)}$) Fourier difference map calculated between the final and the initial data sets recorded at 100 K highlighting the specific radiation damage to the disulphide bond Cys76-Cys94 (map contoured at a ±4.0 σ level). (*c*) ($2mF_{obs} - DF_{calc}$, α_{calc}) electron density map for the first (left) and the last (right) data sets recorded at 100 K illustrating the breakage of the disulphide bond leading to the reorientation of the side chain of Cys94 (maps contoured at a 1.0 σ level). (*d*) Series recorded at 293 K (maps contoured at a 1.0 σ level).

3.3.2*a* Radiation damage to Hen Egg-White Lysozyme at cryogenic temperature (100 K)

Using the same approach as for Cerulean, 9 consecutive data sets, corresponding to accumulated absorbed doses between 110 kGy and 10.0 MGy, were collected at 100 K from the same position of a single crystal. Over the 9 data sets collected, the resolution of the diffraction data decreased from d_{min} = 1.42 Å for the first, to d_{min} = 1.92 Å for the last data set. Examination of the ($2mF_{obs(i)}$-$DF_{calc(i)}$, $\alpha_{calc(i)}$) maps showed the expected progressive reduction of disulphide bonds, in particular for the Cys76-Cys94 disulphide bridge (**Fig. 3*a***). Based on the evolutions of the Wilson *B*-factors, $\tau_{Glob\text{-}CT}$ was estimated to be 18.2 MGy (blue trace in **Fig. 4**), consistent with the Henderson and Garman dose limits.

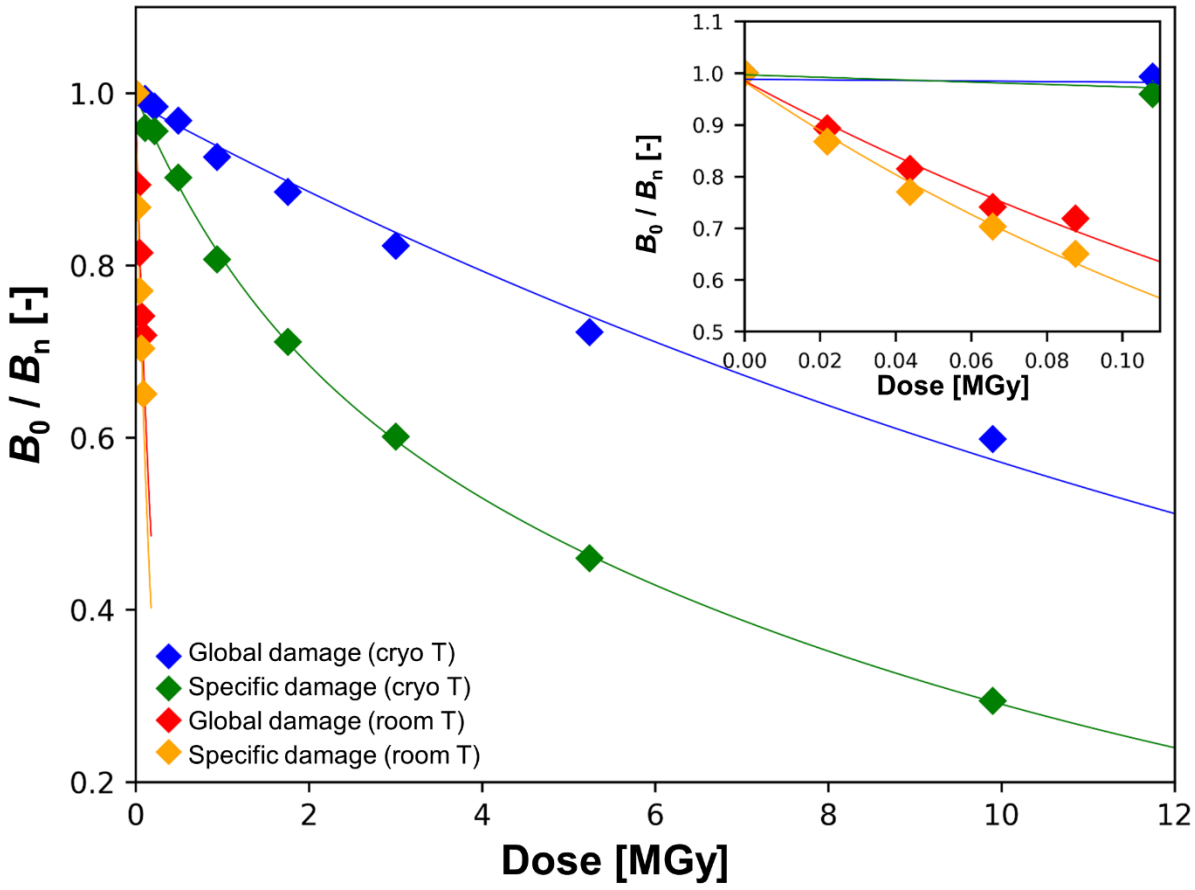


Figure 4: Evolution of *B*-factors as a function of dose for the irradiation series of HEWL crystals at cryogenic and room temperature. The evolution of the Wilson *B*-factor (blue: 100 K, red: 293 K) illustrates the global radiation damage and the evolution of the atomic *B*-factor of Cys94-S_γ (green: 100 K, orange: 293 K) illustrates the specific radiation damage to the disulphide bond Cys76-Cys94. A close-up on the low dose range is presented in the inset.

In order to calculate the life-dose of specific radiation damage at 100 K, we concentrated on the disulphide bond between Cys76 and Cys94, as its reduction leads to the

largest observable structural change, which is the reorientation of the side chain of Cys94 upon disulphide bond breakage (**Figs. 3*b*** and **3*c***). The evolution of the normalized atomic *B*-factor of S_γ of Cys94 is, as expected (Carpentier *et al.*, 2010), best modelled by a bi-exponential decay, with a fast specific radiation damage life-dose constant ($\tau_{Spec\text{-}CT}$) of 1.48 MGy and a slow one ($\tau'_{Spec\text{-}CT}$) of 10.5 MGy (green trace in **Fig. 4**). Thus, as observed for Cerulean, there is a clear decoupling between specific and global radiation damage in crystals of lysozyme at 100 K ($\Delta_{G/S\text{-}CT}$ = 12.4).

3.3.2*b* Radiation damage to Hen Egg-White Lysozyme at room temperature (293 K)

To investigate radiation damage in crystals of HEWL at room temperature, 5 consecutive data sets, corresponding to accumulated absorbed doses between 22 kGy and 110 kGy, were collected from the same position of a single crystal (**Table 3**). Over the course of the 5 data sets collected, the resolution of the diffraction data decreased from d_{min} = 1.37 Å for the first, to d_{min} = 1.89 Å for the last data set and we derived a life-dose constant for global radiation damage at room temperature ($\tau_{Glob\text{-}RT}$) of 251 kGy (red trace in **Fig. 4**), approximately 73 times lower than seen in our experiment at 100 K, again consistent with the [26 - 113] range of increase (Nave & Garman, 2005; Southworth-Davies *et al.*, 2007). However, careful inspection of electron density maps calculated for each of the 5 successive data sets showed no obvious sign of specific radiation damage to the Cys76-Cys94 disulphide bond (**Fig. 3*d***), in accordance with previous observations (Southworth-Davies *et al.*, 2007; Russi *et al.*, 2017). Indeed, the evolution of the atomic *B*-factor of S_γ of Cys94 can be best modelled with a life-dose of $\tau_{Spec\text{-}RT}$ = 198 kGy (orange trace in **Fig. 4**), which gives a decoupling factor $\Delta_{G/S\text{-}RT}$ for lysozyme at room temperature of only 1.3, which constitutes an almost perfect concurrency between the two types of radiation damage.

3.3.2*c* Spectroscopic investigations of specific radiation damage to Hen Egg-White Lysozyme

We were intrigued by the apparent absence of specific radiation damage in our room temperature diffraction based investigations of both Cerulean and HEWL. To further investigate this, a diffraction independent method, Raman spectroscopy, was used to investigate the photoreduction of disulphide bonds in crystals of HEWL at room temperature

by monitoring the Raman peak at 510 cm^{-1}, which is assigned to disulphide bonds (Carpentier *et al.*, 2010).

On-line *in crystallo* non-resonant Raman spectra were recorded on beamline ID29 (von Stetten *et al.*, 2017) in an interleaved manner, before and after X-ray burning cycles at room temperature (293 K). A decrease of the 510 cm^{-1} peak height is observed demonstrating that specific radiation damage to disulphide bonds does occur at 293 K as well (**Figs. 5*a*** and **5*b***). Using the evolution of the 510 cm^{-1} peak height as a metric, the specific radiation damage life-dose in crystals of HEWL at 293 K was calculated to be 89 kGy (**Fig. 5*b***), which is of the same order as the life-dose derived from the diffraction based experiments.

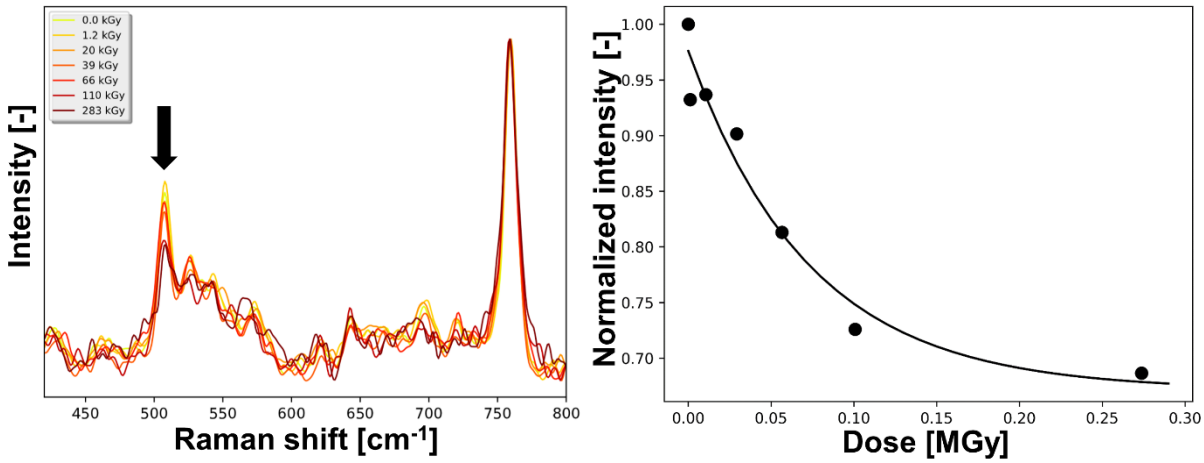


Figure 5: Specific X-ray damage in hen white-egg lysozyme probed by Raman spectroscopy. (*a*) Evolution of Raman spectra recorded with increasing X-ray doses on a lysozyme crystal at 293 K. The arrow indicates the only band whose intensity decreases with increasing dose. (*b*) Decay of the disulphide bond stretching mode at 510 cm^{-1}.

3.3.3 Photoadduct of the LOV2 domain from *Arabidopsis thaliana* phototropin 2

Light-, Oxygen-, or Voltage-sensing (LOV) domains are protein modules found in higher plants, unicellular algae, fungi and bacteria that allow for the sensing of environmental conditions (Christie *et al.*, 2015). In particular, they are found in the blue-light photoreceptor phototropin used by higher plants to mediate positive or negative growth towards or away from a light source. LOV domains contain a light-sensing chromophore, the flavin FMN, which forms a covalent adduct with a cysteine upon absorption of a blue light photon, while exhibiting a blue-shift of its absorption maximum (**Fig. 6*a***). The crystal structures of various LOV domain photoadducts have been determined both at room temperature under photostationary

illumination (LOV2 domain from the chimeric phytochrome/phototropin phy3 of *Adiantum capillus-veneris* (Crosson & Moffat, 2002)) and at cryogenic temperature (LOV1 domain from phototropin 1 of *Chlamydomonas reinhardtii* (Fedorov *et al.*, 2003)). In the latter case, the radiosensitivity of the photoadduct required a composite data set approach (i.e. the merging of partial data sets from several crystals). In the former, the continuous illumination ensured a maximal population of the photoadduct in the crystal.

3.3.4*a* High X-ray sensitivity of the *At*Phot2LOV2 photoadduct at 100 K

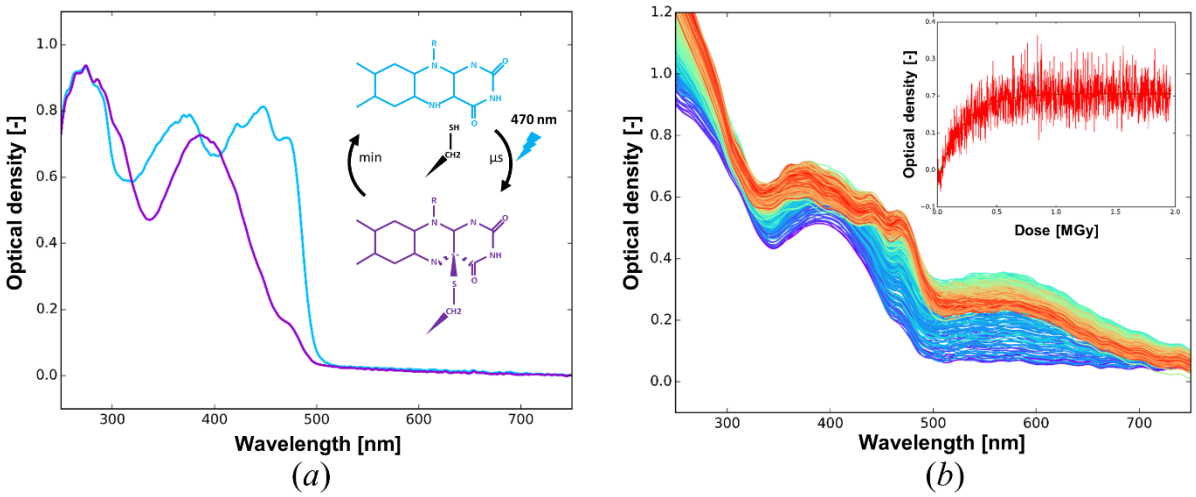


Figure 6: Spectroscopic characterization of the dark and light states of *At*Phot2LOV2. (*a*) Absorption spectra of the dark state (blue) and the blue-light-induced photoadduct (purple) recorded from crystals at room temperature (*b*) Evolution of UV-vis absorption spectra recorded with increasing X-ray doses from an *At*Phot2LOV2 crystal at 100 K (inset: dose-dependent evolution of the absorbance at 490 nm illustrating the X-ray induced relaxation to the dark state).

We chose the LOV2 domain of phototropin 2 from *Arabidopsis thaliana* (*At*Phot2LOV2) and investigated the global and specific radiation damage sensitivity of its photoadduct. We first determined the crystal structure of *At*Phot2LOV2 in its dark state at 1.38 Å resolution at cryogenic temperature. As can be seen in **Fig. 7a**, the S_γ atom of the cysteine residue involved in photoadduct formation presents two alternate conformations: one in which the atom is at 3.6 Å from $C_{4\alpha}$ (65% occupancy), and one in which the $S_\gamma - C_{4\alpha}$ distance is 4.3 Å (35% occupancy) (**Fig. 7*a***).

The photoadduct was generated in the crystal by illumination with a 470 nm LED just before flash-cooling. As for Cerulean and HEWL, we then collected 15 successive data sets from the same position of a single crystal, corresponding to accumulated absorbed doses between 24 kGy and 360 kGy, which constitutes a low dose range (**Table 3**). Over the course of the data set collection, the diffraction limits hardly decreased from d_{min} = 1.70 Å for the first, to d_{min} = 1.71 Å for the last data set and we derived a life-dose $\tau_{Glob\text{-}CT}$ = 34.3 MGy **(Fig. 8)**. As for the experiments described above for Cerulean and HEWL, this value is consistent with the Henderson and Garman limits.

Analysis of the crystal structure from the first data set (absorbed dose 24 kGy) shows the presence of a mixture of the light state structure with 50% occupancy (presence of a S_γ – $C_{4\alpha}$ covalent bond) and of the dark state with 50% occupancy (**Fig. 7*b***), which results from the photostationary equilibrium obtained under continuous illumination before flash-cooling. The second data set (absorbed dose 48 kGy) already shows significant photoreduction of the S_γ – $C_{4\alpha}$ bond as evidenced by ($2mF_{obs}$ - DF_{calc}, α_{calc}) electron density maps and the Fourier difference map (F_{obs2} - F_{obs1}, α_{calc1}) calculated between data sets 1 and 2 (**Figs. 7c** and **7*d***). The S_γ atom of Cys426 implicated in the covalent adduct in the light state structure was chosen to derive the specific radiation damage life-dose $\tau_{Spec\text{-}CT}$. The evolution of its *B*-factor is best modelled by a bi-exponential decay, with a fast specific radiation damage life-dose constant ($\tau_{Spec\text{-}CT}$) of 22 kGy and a slow one ($\tau'_{Spec\text{-}CT}$) of 387 kGy. The resulting decoupling factor $\Delta_{G/S\text{-}CT}$ of 1590 demonstrates that at 100 K this bond is two orders of magnitude more radiosensitive than the carboxylate group of Glu222 in Cerulean and the disulphide bond Cys94 in HEWL (**Fig. 8**).

In order to confirm these diffraction-based estimates, we used the complementary technique UV-vis absorption microspectrophotometry (McGeehan *et al.*, 2009) on beamline ID30A-3 (MASSIF-3) of the ESRF (Theveneau *et al.*, 2013). Upon exposure to X-rays, the 390 nm peak characteristic of the light state changes into 450 and 470 nm peaks characteristic of the dark state (**Fig. 6*b***). This conversion can be interpreted as a result of the X-ray induced decay of the photoadduct species to the dark state (**Fig. 7*a***). We modelled the intensity increase of the 475 nm peak (inset of **Fig. 6b**) with the monoexponential behaviour A – B * exp[-dose / τ], which gives a life-dose of 207 kGy for the phenomenon. This value compares well with the fast and slow life-doses of 22 and 387 kGy derived from the diffraction data.

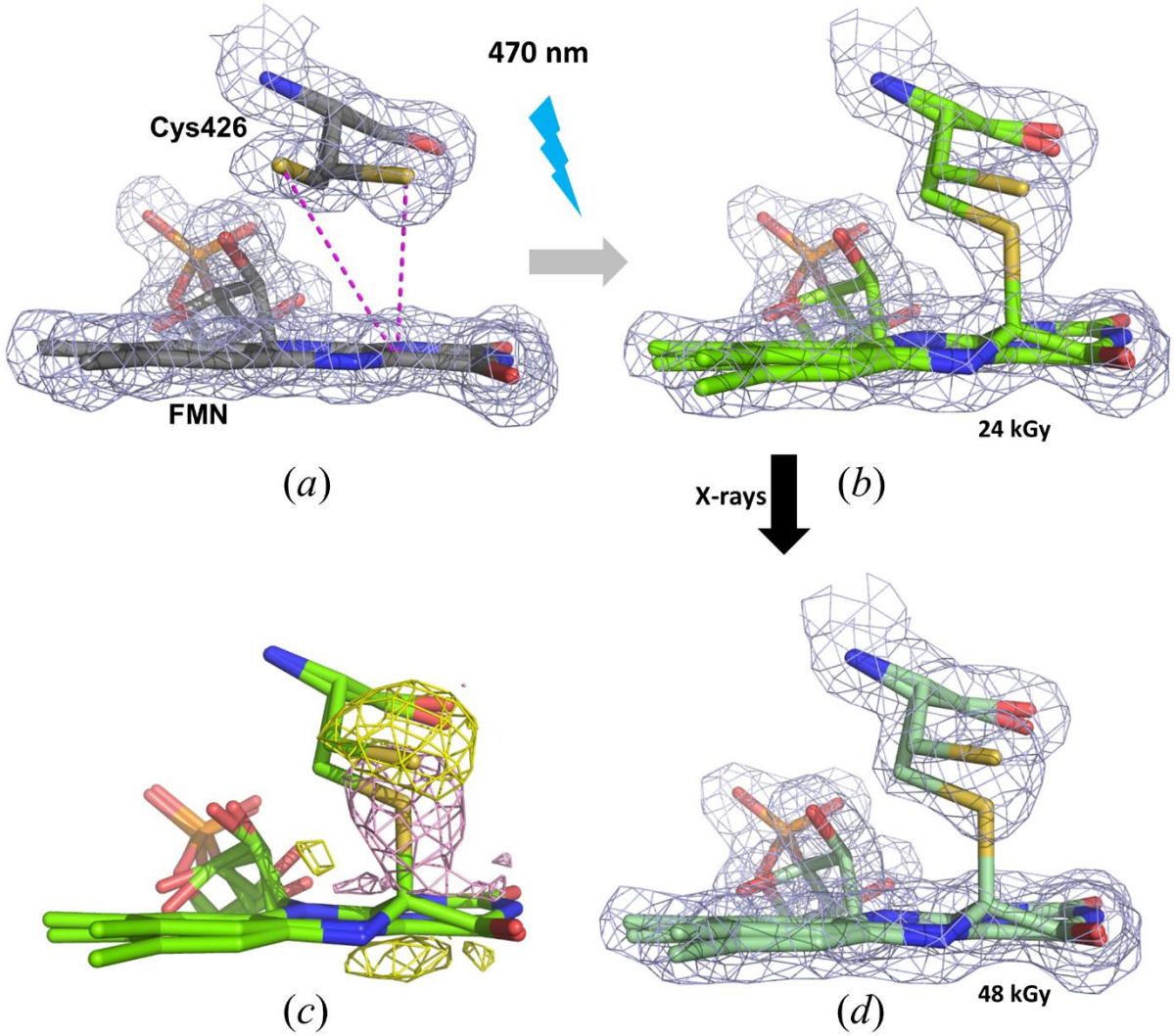


Figure 7: (2mF_{obs} – DF_{calc}, α_{calc}) electron density map in data sets of the dark and light states of *At*Phot2LOV2 at 100 K. (*a*) Structure of the dark state at the position of the chromophore FMN (maps contoured at a 1.5 σ level). (*b*) Structure of the light state recorded with an accumulated dose of 24 kGy (maps contoured at a 1.5 σ level). (*c*) ($F_{obs(2)} - F_{obs(1)}$, $\alpha_{calc(1)}$) Fourier difference map calculated between the second and the first data sets of the irradiation series on the *At*Phot2LOV2 light state highlighting the very fast specific radiation damage to the Cys426-FMN covalent bond (map contoured at a ±4.0 σ level). (*d*) Structure of the light state recorded with a total accumulated dose of 48 kGy (maps contoured at a 1.5 σ level).

3.3.4*b* The *At*Phot2LOV2 photoadduct at room temperature

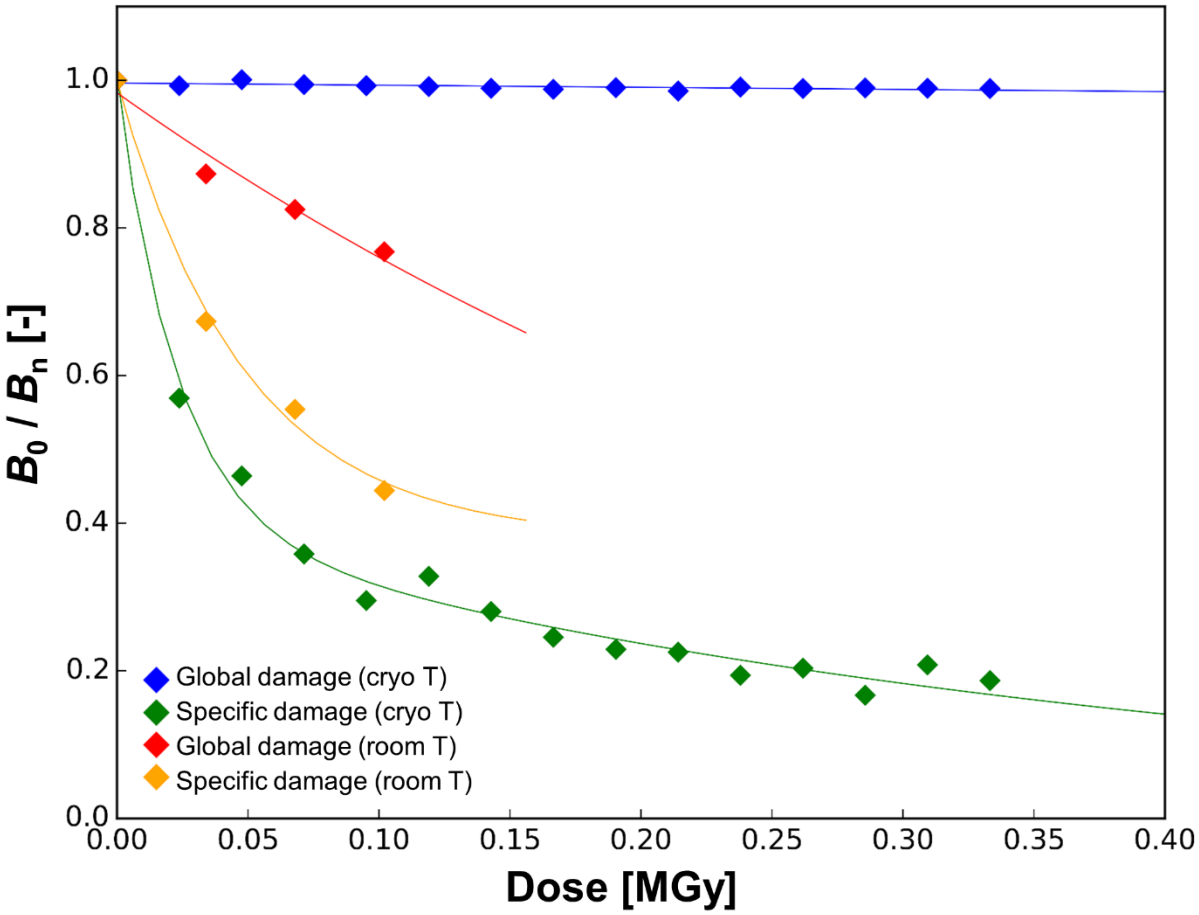


Figure 8: Evolution of *B*-factors as a function of dose for the irradiation series of *At*Phot2LOV2 crystals at cryogenic and room temperatures. The evolution of the Wilson *B*-factor (blue: 100 K, red: 293 K) illustrates global radiation damage and the evolution of the atomic *B*-factor of Cys426-S_γ (green: 100 K, orange: 293 K) illustrates the specific radiation damage to the covalent bond Cys426-FMN that occurs concurrently with the time-dependent relaxation of the photoadduct.

While the *At*Phot2LOV2 photoadduct builds up on the microsecond time scale, its decay back to the dark state occurs on the second to minute time scale (Kasahara, 2002), which opens up the possibility of determining its crystal structure at room temperature. We devised a strategy to record diffraction data as fast as possible in order to probe both global and specific radiation damage while minimizing the extent of intermediate state relaxation. To this end, we used the EIGER-4M detector (Dectris, Switzerland) on beamline ID30A-3 (Theveneau *et al.*, 2013) to record diffraction data from a crystal of *At*Phot2LOV2 at RT immediately after blue light irradiation. A continuous wedge of 1440° was recorded in 8.64 s, resulting in 4 successive

360°-data sets, each of them for an absorbed dose of 34 kGy. Over the course of the data collection, the resolution of the diffraction data as defined by the $CC_{1/2}$ of the outer shell above 0.7 decreased from d_{min} = 2.40 Å to d_{min} = 2.78 Å. Representation of the 4 successive ($2mF_{obs(i)}$-$DF_{calc(i)}$, $\alpha_{calc(i)}$) maps at a 1.5 σ level does not show a major disappearance of the electron density corresponding to the S_γ – $C_{4\alpha}$ covalent bond (**Fig. 9**). Based on the evolution of B_{Wilson} and of the atomic *B*-factor of Cys426-S_γ, we calculate RT life-doses $\tau_{Glob\text{-}RT}$ and $\tau_{Spec\text{-}RT}$ of 389 kGy and 49 kGy, respectively (**Fig. 8**). The resulting decoupling factor $\Delta_{G/S\text{-}RT}$ of 8.0 constitutes a 200-fold reduction compared to the value at cryogenic temperature. Given that a significant part of the intermediate state population has already relaxed even in 10 s, the 'true' value of $\Delta_{G/S\text{-}RT}$ has to be closer to those observed for Cerulean ($\Delta_{G/S\text{-}RT}$ = 2.9) and lysozyme ($\Delta_{G/S\text{-}RT}$ = 1.3) (**Table 2**). This means that in all our three cases, the decoupling factor between global and radiation damage is less than 10 and probably close to 1, posing the question of the severity of specific radiation damage at room temperature.

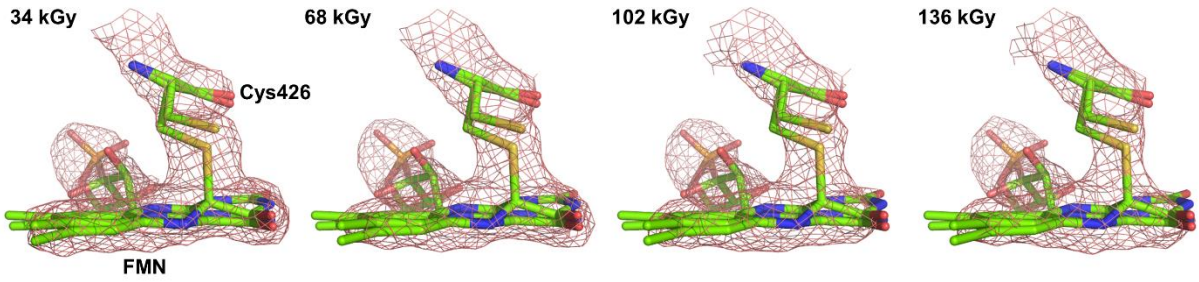


Figure 9: Evolution of the ($2mF_{obs} – DF_{calc}$, α_{calc}) electron density map map (maps contoured at a 2.0 σ level) in data sets of *At*Phot2LOV2 recorded with increasing doses at room temperature.

3.4 Discussion

Our initial inability to detect traces of specific radiation damage in room temperature diffraction data sets of single crystals of the fluorescent protein Cerulean prompted us to define a diffraction-based metric that would allow for an easy comparison of the appearance of specific damage in different systems at various temperatures. To this end, the decoupling factor $\Delta_{G/S\text{-}T}$ which uses the onset of global damage for normalization at a given temperature T, was introduced. This ratio was calculated for both Cerulean and the well-studied protein lysozyme at cryogenic and room temperatures. At cryogenic temperature, the decoupling factors of 12 and 21, respectively illustrates that partial decarboxylation of a key glutamate residue in Cerulean and partial reduction of a disulphide bond in lysozyme occur at moderate doses, well

before the effects of global damage are apparent. At room temperature however, the decoupling factors drop to much lower values (1 and 3, respectively), suggesting a 'recoupling' of specific and global damage. In order to investigate whether this phenomenon also holds true for more radiosensitive systems, the same analysis was performed on the photoadduct of a phototropin LOV2 domain. Here a huge decoupling factor of 1590 at cryogenic temperature reduces to a mere 8 at room temperature. In other words, and in contrast to the situation at cryogenic temperature, specific and global radiation damage evolve on similar dose scales at room temperature for these three systems that all involve covalent bond breakage. Specific damage thus may appear to be as random as global damage, which would explain for instance why it does not show in Fourier difference maps.

Further studies, including simulations, are required to understand the mechanisms causing the difference in behaviour at cryogenic *vs*. room temperature. A potential reason may reside in the free energy landscape offered at the respective temperatures to the solvated electrons and free radicals that are generated by the interaction of X-rays with bulk solvent molecules. Escape lanes for solvated electrons and free radicals terminate either on random groups, resulting in global damage, or on a few groups that are particularly reactive with electrons and radicals, resulting in specific damage. The key difference between free radicals and solvated electrons is that the former are trapped, and therefore mostly inactive at cryogenic temperature, while the latter can still diffuse (Kmetko *et al.*, 2011) and be funnelled towards electron-avid groups within the protein. At room temperature however, all free radicals created in the bulk solvent region can diffuse and impair crystalline order through various mechanisms such as the perturbation of crystalline contacts through direct damage to the protein or the generation of gas molecules (Garman, 2010). The recoupling of specific and global damage at room temperature suggests that the specificity of certain X-ray induced damage to proteins may only arise at cryogenic temperature.

These results are of considerable interest for genuinely time-resolved protein crystallography, which is performed at room temperature (in contrast to methods relying on the cryo-trapping of intermediate states), and which requires determining the structure of particularly X-ray sensitive intermediate states. Indeed, structure determination of reaction intermediate states trapped at cryogenic temperature has often required great care in minimizing, or controlling, the deposited dose (Berglund *et al.*, 2002; Matsui *et al.*, 2002; Adam *et al.*, 2004; Bui *et al.*, 2014). The 'recoupling' of specific and global radiation damage at room temperature will make specific damage much less of a problem, provided a full data

set can been recorded on a single crystal, which can be tested by monitoring parameters indicative of global damage, an option that is now readily accessible on synchrotron beamlines by software such as *DOZOR* (Zander *et al.*, 2015). An added advantage is that achieving the determination of a room temperature structure of a given protein at a sufficiently high resolution will allow, by comparison, for the identification of potential pitfalls in mechanism interpretation due to specific radiation damage occurring in a cryogenic structure of the same protein. In summary, the development of easy-to-use humidity-controlled crystal environments and of fast and noise-free X-ray detectors is triggering a rebirth of room temperature crystallography, which should be favoured in projects where getting a structure close to the physiological state structure is more important than reaching the highest resolution possible.

In conclusion, if one wishes to determine the structure of a protein whose active site (or another part) is particularly sensitive to X-rays, one can either work at cryogenic temperature and perform a thorough radiation damage study by recording positive and negative control data sets aimed at closely monitoring the geometry of the active site (with the aim of deriving the maximum dose below which one should accumulate a complete data set) or work at room temperature and only focus on adjusting the X-ray flux to get a complete data set at the price of a reduced diffraction resolution.

Table 3 Data collection and structure refinement statistics

Values in parentheses are for the outer shell.

	Cerulean									
Data collection	**D1-100K**	**D2-100K**	**D3-100K**	**D4-100K**	**D5-100K**	**D6-100K**	**D7-100K**	**D8-100K**	**D9-100K**	**D10-100K**
Temperature (K)	100									
Accumulated dose (MGy)	0.29	0.58	0.87	1.16	1.45	1.74	2.03	2.32	2.61	2.9
ESRF Beamline	ID29									
Wavelength (Å)	0.976									
Space group	$P2_12_12_1$									
Cell dimensions										
a, *b*, *c* (Å)	51.11 62.72 70.34	51.13 62.74 70.37	51.15 62.67 70.39	51.16 62.78 70.41	51.18 62.80 70.42	51.20 62.82 70.43	51.22 62.84 70.45	51.21 62.82 70.41	51.25 62.87 70.47	51.27 62.88 70.48
Resolution range† (Å)	46.82 - 1.46 (1.50 - 1.46)	46.83 - 1.47 (1.51 - 1.47)	46.84 - 1.48 (1.52 - 1.48)	46.86 - 1.50 (1.54 - 1.50)	46.87 - 1.52 (1.56 - 1.52)	46.88 - 1.55 (1.59 - 1.55)	46.89 - 1.57 (1.61 - 1.57)	46.90 - 1.60 (1.64 - 1.60)	46.91 - 1.63 (1.67 - 1.63)	46.93 - 1.66 (1.70 - 1.66)
Wilson *B*-factor (Å^2)	28.3	28.9	29.5	30.1	30.7	31.4	32.1	32.7	33.4	33.9
Unique reflections	39910 (2939)	39156 (2843)	38421 (2778)	36921 (2648)	35579 (2618)	3393 (2445)	32406 (2390)	30647 (2218)	29061 (2114)	27543 (2002)
Multiplicity	7.1 (7.3)	7.0 (7.3)	7.0 (7.3)	7.0 (7.2)	7.0 (7.2)	7.0 (7.1)	7.0 (7.0)	7.0 (6.7)	7.0 (6.5)	7.0 (7.2)
Completeness (%)	99.9 (99.9)	99.9 (99.8)	99.9 (99.9)	99.8 (99.9)	99.8 (99.9)	99.8 (99.8)	99.8 (99.7)	99.8 (99.6)	99.8 (99.3)	99.9 (100.0)
Mean I/sigma(I)	24.07 (1.56)	24.42 (1.51)	24.93 (1.45)	25.52 (1.58)	26.21 (1.56)	27.72 (1.66)	28.31 (1.73)	29.55 (1.83)	30.32 (1.82)	31.42 (2.04)
R_{meas}‡	0.042 (1.387)	0.041 (1.437)	0.040 (1.478)	0.040 (1.379)	0.039 (1.370)	0.038 (1.272)	0.037 (1.216)	0.036 (1.130)	0.036 (1.144)	0.036 (1.103)
$CC_{1/2}$	0.999 (0.724)	1.0 (0.697)	1.0 (0.709)	1.0 (0.750)	1.0 (0.705)	1.0 (0.703)	1.0 (0.712)	1.0 (0.728)	1.0 (0.709)	1.0 (0.717)
Refinement statistics										
Resolution (Å)	46.81 - 1.46 (1.50 – 1.46)									
R_{work}	0.167 (0.42)									
R_{free}	0.197 (0.38)									
N. of atoms	2,161									
Average atomic *B*-factor (Å^2)	23.47									
R.m.s deviations										
Bond lengths (Å)	0.012									
Bond angles (°)	1.63									
PDB entry code	6QQ8									

† The resolution cutoff is based on $CC_{1/2}$.‡ Rmeas = Rmerge × $[N/(N-1)]^{1/2}$, where N is the data multiplicity.

	Cerulean									
Data collection	**D11-100K**	**D12-100K**	**D13-100K**	**D14-100K**	**D15-100K**	**D16-100K**	**D17-100K**	**D18-100K**	**D19-100K**	**D20-100K**
Temperature (K)	100									
Accumulated dose (MGy)	3.19	3.48	3.77	4.06	4.35	4.64	4.93	5.22	5.51	5.80
ESRF Beamline	ID29									
Wavelength (Å)	0.976									
Space group	$P2_12_12_1$									
Cell dimensions										
a, *b*, *c* (Å)	51.28 62.90 70.48	51.30 62.91 70.49	51.32 62.93 70.50	51.33 62.94 70.51	51.35 62.96 70.51	51.36 62.97 70.52	51.33 62.94 70.51	51.39 62.99 70.53	51.40 63.00 70.53	51.41 63.00 70.53
Resolution range (Å)	46.93 - 1.68 (1.72 - 1.68)	46.94 - 1.70 (1.74 - 1.70)	46.95 - 1.72 (1.76 - 1.72)	46.95 - 1.74 (1.79 - 1.74)	46.97 - 1.76 (1.81 - 1.76)	46.97 - 1.78 (1.83 - 1.78)	46.96 - 1.79 (1.84 - 1.79)	46.99 - 1.80 (1.85 - 1.80)	46.99 - 1.81 (1.86 - 1.81)	46.99 - 1.82 (1.87 - 1.82)
Wilson *B*-factor ($Å^2$)	34.4	34.9	35.4	35.8	36.2	36.6	37.1	37.6	38.2	38.7
Unique reflections	26603 (1935)	25703 (1848)	24856 (1814)	24055 (1760)	23260 (1697)	22518 (1644)	22113 (1613)	21797 (1588)	21457 (1562)	21133 (1544)
Multiplicity	7.0 (7.4)	6.9 (7.4)	6.9 (7.4)	6.9 (7.4)	6.8 (7.4)	6.8 (7.4)	6.8 (7.4)	6.8 (7.4)	6.8 (7.3)	6.8 (7.3)
Completeness (%)	99.9 (100.0)	99.9 (99.9)	99.9 (99.9)	99.9 (100.0)	99.8 (99.9)	99.9 (100.0)	99.8 (100.0)	99.8 (99.9)	99.9 (100.0)	99.9 (100.00)
Mean I/sigma(I)	30.87 (2.09)	31.11 (2.06)	30.77 (2.20)	31.25 (2.19)	31.10 (2.24)	31.16 (2.27)	30.41 (2.17)	29.48 (2.12)	28.85 (2.08)	28.23 (2.05)
R_{meas}	0.037 (1.079)	0.037 (1.078)	0.037 (1.068)	0.037 (1.060)	0.038 (1.030)	0.038 (1.018)	0.039 (1.066)	0.040 (1.073)	0.041 (1.072)	0.042 (1.088)
$CC_{1/2}$	1.0 (0.721)	1.0 (0.760)	1.0 (0.758)	1.0 (0.745)	1.0 (0.719)	1.0 (0.775)	1.0 (0.744)	1.0 (0.767)	1.0 (0.774)	1 (0.726)
Refinement statistics										
Resolution (Å)										46.99 - 1.82 (1.87 – 1.82)
R_{work}										0.154 (0.281)
R_{free}										0.195 (0.345)
N. of atoms										2,220
Average atomic *B*-factor ($Å^2$)										33.9
R.m.s deviations										
Bond lengths (Å)										0.009
Bond angles (°)										1.38
PDB entry code										6QQ9

	Cerulean						
Data collection	**D1-293K**	**D2-293K**	**D3-293K**	**D4-293K**	**D5-293K**	**D6-293K**	**D7-293K**
Temperature (K)				293			
Accumulated dose (MGy)	0.021	0.042	0.063	0.084	0.105	0.125	0.146
ESRF Beamline				ID29			
Wavelength (Å)				0.976			
Space group							
Cell dimensions	51.93 63.04 71.32	51.99 62.91 71.45	52.08 62.94 71.58	52.11 62.91 71.64	52.12 62.87 71.68	52.09 62.77 71.63	52.04 62.67 71.56
a, *b*, *c* (Å)							
Resolution range (Å)	47.24 - 1.66 (1.70 - 1.66)	47.22 - 1.81 (1.86 - 1.81)	47.27 - 1.95 (2.00 - 1.95)	47.28 - 2.07 (2.12 - 2.07)	47.27 - 2.18 (2.24 - 2.18)	47.21 - 2.32 (2.38 - 2.32)	47.15 - 2.45 (2.51 - 2.45)
Wilson *B*-factor (Å^2)	33.1	36.9	38.1	41.0	44.1	47.7	49.9
Unique reflections	28306 (2062)	21964 (1592)	17737 (1296)	14881 (1060)	12788 (912)	10628 (773)	9048 (670)
Multiplicity	7.3 (7.4)	7.3 (7.4)	7.1 (7.1)	7.1 (6.8)	7.0 (7.7)	6.8 (7.6)	6.6 (7.4)
Completeness (%)	100.0 (100.0)	100.0 (99.9)	99.9 (100.0)	99.9 (99.6)	100.0 (100.0)	99.9 (100.0)	99.9 (100.00)
Mean I/sigma(I)	15.43 (2.09)	21.87 (2.13)	12.01 (2.06)	10.65 (2.06)	9.71 (2.03)	9.65 (2.09)	8.90 (2.08)
R_{merge}							
R_{meas}	0.07 (1.053)	0.06 (1.190)	0.132 (1.534)	0.155 (1.501)	0.168 (1.610)	0.174 (1.562)	0.185 (1.425)
$CC_{1/2}$	0.998 (0.753)	1.00 (0.788)	0.998 (0.731)	0.997 (0.641)	0.997 (0.774)	0.997 (0.815)	0.997 (0.749)
Refinement statistics							
Resolution (Å)	47.23 - 1.66 (1.70 – 1.66)						47.15 - 2.45 (2.51 – 2.45)
R_{work}	0.144 (0.254)						0.256 (0.386)
R_{free}	0.171 (0.285)						0.281 (0.417)
N. of atoms	2,161						2,161
Average atomic *B*-factor (Å^2)	31.3						50.3
R.m.s deviations							
Bond lengths (Å)	0.01						0.006
Bond angles (°)	1.65						1.31
PDB entry code	6QQA						6QQB

HEWL									
Data collection	**D1-100K**	**D2-100K**	**D3-100K**	**D4-100K**	**D5-100K**	**D6-100K**	**D7-100K**	**D8-100K**	**D9-100K**
Temperature (K)							100		
Accumulated dose (MGy)	0.11	0.22	0.33	0.60	1.05	1.87	3.12	5.35	10.01
ESRF Beamline							ID29		
Wavelength (Å)							0.976		
Space group							$P4_32_12$		
Cell dimensions									
a, *b*, *c* (Å)	77.56 77.56 37.13	77.56 77.56 37.14	77.58 77.58 37.15	77.60 77.60 37.16	77.64 77.64 37.18	77.69 77.69 37.21	77.77 77.77 37.25	77.87 77.87 37.31	77.96 77.96 37.38
Resolution range (Å)	33.49 - 1.42 (1.46 - 1.42)	33.50 - 1.43 (1.47 - 1.43)	33.51 - 1.43 (1.47 - 1.43)	33.52 - 1.44 (1.48 - 1.44)	33.54 - 1.49 (1.53 - 1.49)	33.56 - 1.52 (1.56 - 1.52)	33.60 - 1.58 (1.62 - 1.58)	33.65 - 1.72 (1.76 - 1.72)	33.71 - 1.92 (1.97 - 1.92)
Wilson *B*-factor ($Å^2$)	23.1	23.3	23.5	23.9	25.0	26.7	28.1	32.0	38.7
Unique reflections	20446 (1472)	20032 (1407)	20027 (1419)	19659 (1405)	17827 (1291)	16874 (1242)	15081 (1124)	11760 (893)	8490 (640)
Multiplicity	6.5 (6.8)	6.5 (6.8)	6.5 (6.8)	6.5 (6.9)	6.5 (6.7)	6.5 (6.6)	6.5 (6.1)	6.5 (6.4)	6.4 (6.6)
Completeness (%)	93.1 (92.5)	93.2 (91.8)	93.1 (92.0)	93.1 (92.1)	93.2 (94.2)	93.3 (94.1)	93.1 (94.6)	92.6 (96.3)	91.8 (95.7)
Mean I/sigma(I)	17.49 (1.99)	17.28 (2.03)	17.51 (1.96)	17.07 (1.96)	17.87 (2.12)	17.48 (1.96)	17.15 (1.86)	16.93 (1.98)	15.34 (1.88)
R_{meas}	0.059 (0.836)	0.060 (0.816)	0.059 (0.842)	0.060 (0.837)	0.058 (0.791)	0.059 (0.835)	0.062 (0.909)	0.065 (0.878)	0.078 (1.015)
$CC_{1/2}$	1.000 (0.703)	0.999 (0.726)	1.000 (0.710)	1.000 (0.726)	1.000 (0.724)	1.000 (0.713)	0.999 (0.718)	0.999 (0.711)	0.999 (0.706)
Refinement statistics									
Resolution (Å)	33.49 - 1.42 (1.46 – 1.42)								33.71 - 1.92 (19.7 - 1.92)
R_{work}	0.169 (0.262)								0.156 (0.258)
R_{free}	0.203 (0.288)								0.216 (0.267)
N. of atoms	1 212								1217
Average atomic *B*-factor ($Å^2$)	22.3								40.9
R.m.s deviations									
Bond lengths (Å)	0,011								0.009
Bond angles (°)	1,78								1.572
PDB entry code	6QQC								6QQD

	HEWL				
Data collection	**D1-293K**	**D2-293K**	**D3-293K**	**D4-293K**	**D5-293K**
Temperature (K)	293				
Accumulated dose (MGy)	0.020	0.040	0.060	0.080	0.100
ESRF Beamline	ID29				
Wavelength (Å)	1.07				
Space group	$P4_32_12$				
Cell dimensions					
a, *b*, *c* (Å)	79.24 79.24 38.06	79.15 79.15 38.10	79.07 79.07 38.12	78.88 78.88 38.09	79.01 79.01 38.18
Resolution range (Å)	39.62 – 1.37 (1.41 – 1.37)	39.58 - 1.49 (1.53 - 1.49)	39.54 - 1.59 (1.63 - 1.59)	39.44 - 1.80 (1.85 - 1.80)	39.51 - 1.95 (2.00 - 1.95)
Wilson *B*-factor (Å^2)	25.0	28.0	30.7	33.7	34.7
Unique reflections	25762 (1828)	20176 (1450)	16682 (1201)	11545 (813)	9218 (665)
Multiplicity	7.0 (7.0)	7.0 (7.3)	7.0 (7.2)	6.6 (7.1)	6.7 (7.2)
Completeness (%)	98.9 (96.9)	99.2 (97.8)	99.4 (99.8)	99.3 (98.1)	99.6 (99.7)
Mean I/sigma(I)	17.10 (2.24)	16.75 (2.43)	15.84 (2.31)	13.62 (2.54)	10.45 (2.73)
R_{meas}	0.055 (0.810)	0.059 (0.816)	0.066 (0.863)	0.108 (1.122)	0.152 (1.138)
$CC_{1/2}$	0.999 (0.70)	0.999 (0.725)	0.999 (0.715)	0.998 (0.725)	0.997 (0.708)
Refinement statistics					
Resolution (Å)	39.61 - 1.37 (1.41 – 1.37)				39.51 - 1.95 (2.00 – 1.95)
R_{work}	0.165 (0.271)				0.167 (0.223)
R_{free}	0.195 (0.255)				0.217 (0.230)
N. of atoms	1,221				1,143
Average atomic *B*-factor (Å^2)	20.4				28.0
R.m.s deviations					
Bond lengths (Å)	0.05				0.009
Bond angles (°)	1.10				1.27
PDB entry code	6QQE				6QQF

	*At*Phot2LOV2							
Data collection	**Dark-100K**	**Light-100K D1**	**Light-100K D2**	**Light-100K D3**	**Light-100K D4**	**Light-100K D5**	**Light-100K D6**	**Light-100K D7**
Temperature (K)	100				100			
Accumulated dose (MGy)	2.68	0.024	0.048	0.071	0.095	0.119	0.143	0.167
ESRF Beamline	ID30A-3				ID29			
Wavelength (Å)	0.968				0.976			
Space group					$P4_32_12$			
Cell dimensions								
a, *b*, *c* (Å)	40.15 40.15 131.57	40.32 40.32 131.28	40.33 40.33 131.28	40.32 40.32 131.28	40.38 40.38 131.40	40.38 40.38 131.41	40.38 40.38 131.41	40.38 40.38 131.41
Resolution range (Å)	38.40 - 1.38 (1.43 - 1.38)	38.55 - 1.70 (1.76 - 1.70)	38.55 - 1.70 (1.76 - 1.70)	38.55 - 1.70 (1.76 - 1.70)	38.60 - 1.71 (1.77 - 1.71)	38.60 - 1.71 (1.77 - 1.71)	38.61 - 1.71 (1.77 - 1.71)	38.61 - 1.71 (1.77 - 1.71)
Wilson *B*-factor ($Å^2$)	18.8	29.1	29.1	29.1	29.3	29.4	29.4	29.5
Unique reflections	23065 (2265)	12655 (1202)	12659 (1199)	12667 (1207)	12393 (1204)	12383 (1201)	12396 (1200)	12382 (1197)
Multiplicity	7.18 (6.32)	6.24 (6.47)	6.24 (6.45)	6.24 (6.47)	6.29 (6.45)	6.29 (6.43)	6.28 (6.43)	6.29 (6.41)
Completeness (%)	99.8 (99.8)	99.7 (99.0)	99.8 (98.8)	99.8 (99.4)	98.4 (99.5)	98.3 (99.1)	98.4 (98.7)	98.3 (98.4)
Mean I/sigma(I)	16.84 (1.68)	10.18 (1.74)	10.39 (1.73)	10.24 (1.71)	11.46 (1.75)	11.46 (1.78)	11.34 (1.78)	11.37 (1.67)
R_{meas}	0.061 (0.981)	0.103 (0.826)	0.101 (0.846)	0.102 (0.830)	0.096 (0.945)	0.095 (0.933)	0.096 (0.937)	0.097 (0.980)
$CC_{1/2}$	0.999 (0.568)	0.997 (0.783)	0.998 (0.791)	0.998 (0.775)	0.998 (0.804)	0.996 (0.762)	0.996 (0.782)	0.998 (0.757)
Refinement statistics								
Resolution (Å)	38.40 - 1.38 (1.42 - 1.38)	38.55 - 1.70 (1.74 - 1.70)	38.55 - 1.7 (1.74 - 1.70)					
R_{work}	0.140 (0.268)	0.194 (0.289)	0.190 (0.287)					
R_{free}	0.170 (0.288)	0.239 (0.335)	0.216 (0.391)					
N. of atoms	1,230	1,208	1,212					
Average atomic *B*-factor ($Å^2$)	16.6	23.5	23.7					
R.m.s deviations								
Bond lengths (Å)	0.006	0.004	0.005					
Bond angles (°)	1.35	1.60	1.64					
PDB entry code	6QQH	6QQI	6QSA					

	*At*Phot2LOV2							
Data collection	**Light-100K D8**	**Light-100K D9**	**Light-100K D10**	**Light-100K D11**	**Light-100K D12**	**Light-100K D13**	**Light-100K D14**	**Light-100K D15**
Temperature (K)	100							
Accumulated dose (MGy)	0.190	0.214	0.238	0.262	0.286	0.309	0.333	0.357
ESRF Beamline	ID29							
Wavelength (Å)	0.976							
Space group	$P4_32_12$							
Cell dimensions								
a, *b*, *c* (Å)	40.38 40.38 131.42	40.39 40.39 131.43	40.39 40.39 131.43	40.39 40.39 131.43	40.33 40.33 131.31	40.34 40.34 131.32	40.34 40.34 131.33	40.35 40.35 131.35
Resolution range (Å)	38.61 - 1.71 (1.77 - 1.71)	38.61 - 1.72 (1.78 - 1.72)	38.61 - 1.71 (1.77 - 1.71)	38.56 - 1.70 (1.76 - 1.70)	38.56 - 1.71 (1.77 - 1.71)	38.56 - 1.71 (1.77 - 1.71)	38.56 - 1.71 (1.77 - 1.71)	38.57 - 1.71 (1.77 - 1.71)
Wilson *B*-factor ($Å^2$)	29.5	29.4	29.6	29.4	29.5	29.4	29.5	29.5
Unique reflections	12371 (1186)	12182 (1160)	12390 (1205)	12664 (1199)	12453 (1188)	12463 (1200)	12467 (1204)	12475 (1203)
Multiplicity	6.29 (6.44)	6.28 (6.39)	6.29 (6.44)	6.24 (6.46)	6.23 (6.38)	6.24 (6.40)	6.23 (6.40)	6.24 (6.45)
Completeness (%)	98.2 (97.5)	98.3 (99.1)	98.3 (99.3)	99.8 (99.2)	99.8 (98.3)	99.8 (99.3)	99.8 (99.3)	99.8 (99.3)
Mean I/sigma(I)	11.37 (1.70)	11.60 (1.77)	11.18 (1.67)	10.06 (1.66)	10.27 (1.67)	10.23 (1.68)	10.12 (1.70)	10.06 (1.65)
Refinement Statistics								
R_{meas}	0.096 (0.975)	0.095 (0.937)	0.098 (0.983)	0.103 (0.879)	0.101 (0.875)	0.102 (0.855)	0.103 (0.837)	0.104 (0.875)
$CC_{1/2}$	0.997 (0.741)	0.998 (0.783)	0.998 (0.760)	0.997 (0.705)	0.997 (0.727)	0.996 (0.775)	0.997 (0.758)	0.997 (0.743)
Refinement statistics								
Resolution (Å)								
R_{work}								
R_{free}								
N. of atoms								
Average atomic *B*-factor ($Å^2$)								
R.m.s deviations								
Bond lengths (Å)								
Bond angles (°)								

	*At*Phot2LOV2				
Data collection	**Dark 293K**	**Light-293K D1**	**Light-293K D2**	**Light-293K D3**	**Light-293K D4**
Temperature (K)	293				
Accumulated dose (MGy)	0.354	0.034	0.068	0.102	0.136
ESRF Beamline	ID30A-3				
Wavelength (Å)	0.968				
Space group	$P4_32_12$				
Cell dimensions					
a, *b*, *c* (Å)	40.886 40.886 132.691	41.45 41.45 133.53	41.45 41.45 133.53	41.45 41.45 133.53	41.45 41.45 133.53
Resolution range (Å)	39.07 - 2.08 (2.15 - 2.08)	39.59 - 2.40 (2.50 - 2.40)	39.59 - 2.60 (2.70 - 2.60)	39.59 - 2.76 (2.86 - 2.76)	39.59 - 2.78 (2.88 - 2.78)
Wilson *B*-factor (Å^2)	43.1	52.0	59.6	63.0	67.7
Unique reflections	6924 (594)	4926 (558)	3897 (394)	3275 (278)	3217 (267)
Multiplicity	8.56 (8.73)	23.1 (23.5)	23.1 (22.0)	23.4 (21.1)	23.4 (21.1)
Completeness (%)	93.6 (87.4)	98.0 (99.1)	97.6 (99.0)	97.4 (88.8)	97.5 (86.4)
Mean I/sigma(I)	10.72 (1.73)	17.64 (1.52)	19.14 (1.69)	19.69 (2.47)	17.12 (1.72)
R_{meas}	0.145 (1.169)	0.177 (2.108)	0.160 (1.749)	0.154 (1.208)	0.184 (1.639)
$CC_{1/2}$	0.999 (0.501)	0.999 (0.726)	0.999 (0.727)	0.999 (0.760)	0.999 (0.712)
Refinement Statistics					
Resolution (Å)	24.21 - 2.08 (2.14 - 2.08)	39.59 - 2.40 (2.46 - 2.40)			
R_{work}	0.247 (0.348)	0.239 (0.322)			
R_{free}	0.292 (0.338)	0.286 (0.314)			
N. of atoms	956	1043			
Average atomic *B*-factor (Å^2)	38.7	46.9			
R.m.s deviations					
Bond lengths (Å)	0.002	0.002			
Bond angles (°)	1.24	1.47			
PDB entry code	6QQJ	6QQK			

Acknowledgements

SA acknowledges a PhD fellowship from the ESRF. We thank Ulrike Kapp and Melissa Saidi for their help with the preparation of lysozyme crystals. The ESRF is acknowledged for access to beamlines and facilities for molecular biology *via* its in-house research programme. This work used the icOS Lab, which is a platform of the Grenoble Instruct Centre (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) with support from the French Infrastructure for Integrated Structural Biology (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology.

3.5 References

Adam, V., Carpentier, P., Violot, S., Lelimousin, M. L., Darnault, C., Nienhaus, G. U. & Bourgeois, D. (2009). *J. Am. Chem. Soc.* **131**, 18063–18065.

Adam, V., Royant, A., Nivière, V., Molina-Heredia, F. P. & Bourgeois, D. (2004). *Structure*, **12**, 1729–1740.

Arvai, A. S., Hartfield, K. A., Christie, J. M., Hitomi, K., Pratt, A. J., Tainer, J. A., Getzoff, E. D. & Mettlen, M. (2012). *J. Biol. Chem.* **287**, 22295–22304.

Berglund, G. I., Carlsson, G. H., Smith, A. T., Szöke, H., Henriksen, A. & Hajdu, J. (2002). *Nature (London)*, **417**, 463–468.

Blake, C. C. F., Koenig, D. F., Mair, G. A., North, A. C. T., Phillips, D. C. & Sarma, V. R. (1965). *Nature (London)*, **206**, 757–761.

Borshchevskiy, V., Round, E., Erofeev, I., Weik, M., Ishchenko, A., Gushchin, I., Mishin, A., Willbold, D., Büldt, G. & Gordeliy, V. (2014). *Acta Cryst.* D**70**, 2675–2685.

Bui, S., Von Stetten, D., Jambrina, P. G., Prangé, T., Colloc'H, N., De Sanctis, D., Royant, A., Rosta, E. & Steiner, R. A. (2014). *Angew. Chem. Int. Ed. Engl.* **53**, 13710–13714.

Burmeister, W. P. (2000). *Acta Cryst.* D**56**, 328–341.

Bury, C. S. & Garman, E. F. (2018). *J. Appl. Cryst.* **51**, 952–962.

Carpentier, P., Royant, A., Weik, M. & Bourgeois, D. (2010). *Structure*, **18**, 1410–1419.

Christie, J. M., Blackwood, L., Petersen, J. & Sullivan, S. (2015). *Plant Cell Physiol.* **56**, 401–

413.

Clavel, D., Gotthard, G., von Stetten, D., De Sanctis, D., Pasquier, H., Lambert, G. G., Shaner, N. C. & Royant, A. (2016). *Acta Cryst.* D**72**, 1298–1307.

Colletier, J. P., Sliwa, M., Gallat, F. X., Sugahara, M., Guillon, V., Schirò, G., Coquelle, N., Woodhouse, J., Roux, L., Gotthard, G., Royant, A., Uriarte, L. M., Ruckebusch, C., Joti, Y., Byrdin, M., Mizohata, E., Nango, E., Tanaka, T., Tono, K., Yabashi, M., Adam, V., Cammarata, M., Schlichting, I., Bourgeois, D. & Weik, M. (2016). *J. Phys. Chem. Lett.* **7**, 882–887.

Coquelle, N., Sliwa, M., Woodhouse, J., Schirò, G., Adam, V., Aquila, A., Barends, T. R. M., Boutet, S., Byrdin, M., Carbajo, S., Mora, E. D. La, Doak, R. B., Feliks, M., Fieschi, F., Foucar, L., Guillon, V., Hilpert, M., Hunter, M. S., Jakobs, S., Koglin, J. E., Kovacsova, G., Lane, T. J., Lévy, B., Liang, M., Nass, K., Ridard, J., Robinson, J. S., Roome, C. M., Ruckebusch, C., Seaberg, M., Thepaut, M., Cammarata, M., Demachy, I., Field, M., Shoeman, R. L., Bourgeois, D., Colletier, J. P., Schlichting, I. & Weik, M. (2018). *Nat. Chem.* **10**, 31–37.

Crosson, S. & Moffat, K. (2002). *Plant Cell*, **14**, 1067–1075.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.

Fedorov, R., Schlichting, I., Hartmann, E., Domratcheva, T., Fuhrmann, M. & Hegemann, P. (2003). *Biophys. J.* **84**, 2474–2482.

Fraser, J. S., van den Bedem, H., Samelson, A. J., Lang, P. T., Holton, J. M., Echols, N. & Alber, T. (2011). *Proc. Natl Acad. Sci. USA*, **108**, 16247–16252.

Garman, E. F. (2010). *Acta Cryst.* D**66**, 339–351.

Garman, E. F. & Schneider, T. R. (1997). *J. Appl. Cryst.* **30**, 211–237.

Gotthard, G., von Stetten, D., Clavel, D., Noirclerc-Savoye, M. & Royant, A. (2017). *Biochemistry*, **56**, 6418–6422.

Henderson, R. (1995). *Q. Rev. Biophys.* **28**, 171–193.

Hope, H. (1990). *Annu. Rev. Biophys. Biophys. Chem.* **19**, 107–126.

Hutchison, C. D. M., Cordon-Preciado, V., Morgan, R. M. L., Nakane, T., Ferreira, J., Dorlhiac, G., Sanchez-Gonzalez, A., Johnson, A. S., Fitzpatrick, A., Fare, C., Marangos,

J. P., Yoon, C. H., Hunter, M. S., Deponte, D. P., Boutet, S., Owada, S., Tanaka, R., Tono, K., Iwata, S. & Van Thor, J. J. (2017). *Int. J. Mol. Sci.* **18**,.

Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.

Kasahara, M. (2002). *Plant Physiol.* **129**, 762–773.

Kaucikas, M., Fitzpatrick, A., Bryan, E., Struve, A., Henning, R., Kosheleva, I., Srajer, V., Groenhof, G. & Van Thor, J. J. (2015). *Proteins*. **83**, 397–402.

Kendrew, J. C., Dickerson, R. E., Strandberg, B. E., Hart, R. G., Davies, D. R., Phillips, D. C. & Shore, V. C. (1960). *Nature (London)*, **185**, 422–427.

Kmetko, J., Warkentin, M., Englich, U. & Thorne, R. E. (2011). *Acta Cryst.* D**67**, 881–893.

Leal, R. M. F., Bourenkov, G., Russi, S. & Popov, A. N. (2013). *J. Synchrotron Rad.* **20**, 14–22.

Lelimousin, M., Noirclerc-Savoye, M., Lazareno-Saez, C., Paetzold, B., Le Vot, S., Chazal, R., Macheboeuf, P., Field, M. J., Bourgeois, D. & Royant, A. (2009). *Biochemistry*, **48**, 10038–10046.

Matsui, Y., Sakai, K., Murakami, M., Shiro, Y., Adachi, S. I., Okumura, H. & Kouyama, T. (2002). *J. Mol. Biol.* **324**, 469–481.

McCarthy, A. A., Barrett, R., Beteva, A., Caserotto, H., Dobias, F., Felisaz, F., Giraud, T., Guijarro, M., Janocha, R., Khadrouche, A., Lentini, M., Leonard, G. A., Lopez Marrero, M., Malbet-Monaco, S., McSweeney, S., Nurizzo, D., Papp, G., Rossi, C., Sinoir, J., Sorez, C., Surr, J., Svensson, O., Zander, U., Cipriani, F., Theveneau, P. & Mueller-Dieckmann, C. (2018). *J. Synchrotron Rad.* **25**, 1249–1260.

McGeehan, J., Ravelli, R. B. G., Murray, J. W., Owen, R. L., Cipriani, F., McSweeney, S., Weik, M. & Garman, E. F. (2009). *J. Synchrotron Rad.* **16**, 163–172.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.

Nave, C. & Garman, E. F. (2005). *J. Synchrotron Rad.* **12**, 257–260.

Owen, R. L., Axford, D., Nettleship, J. E., Owens, R. J., Robinson, J. I., Morgan, A. W., Doré, A. S., Lebon, G., Tate, C. G., Fry, E. E., Ren, J., Stuart, D. I. & Evans, G. (2012). *Acta Cryst.* D**68**, 810–818.

Owen, R. L., Rudiño-Piñera, E. & Garman, E. F. (2006). *Proc. Natl Acad. Sci. USA.* **103**, 4912–4917.

Perutz, M. F., Rossmann, M. G., Cullis, A. F., Muirhead, H., Will, G. & North, A. C. T. (1960). *Nature (London)*, **185**, 416–422.

Rajendran, C., Dworkowski, F. S. N., Wang, M. & Schulze-Briese, C. (2011). *J. Synchrotron Rad.* **18**, 318–328.

Ravelli, R. B. & McSweeney, S. M. (2000). *Structure*, **8**, 315–328.

Rizzo, M. a, Springer, G. H., Granada, B. & Piston, D. W. (2004). *Nat. Biotechnol.* **22**, 445–449.

Royant, A. & Noirclerc-Savoye, M. (2011). *J. Struct. Biol.* **174**, 385–390.

Russi, S., González, A., Kenner, L. R., Keedy, D. A., Fraser, J. S. & Van Den Bedem, H. (2017). *J. Synchrotron Rad.* **24**, 73–82.

Sanchez-Weatherby, J., Bowler, M. W., Huet, J., Gobbo, A., Felisaz, F., Lavault, B., Moya, R., Kadlec, J., Ravelli, R. B. G. & Cipriani, F. (2009). *Acta Cryst.* D**65**, 1237–1246.

De Sanctis, D., Beteva, A., Caserotto, H., Dobias, F., Gabadinho, J., Giraud, T., Gobbo, A., Guijarro, M., Lentini, M., Lavault, B., Mairs, T., McSweeney, S., Petitdemange, S., Rey-Bakaikoa, V., Surr, J., Theveneau, P., Leonard, G. A. & Mueller-Dieckmann, C. (2012). *J. Synchrotron Rad.* **19**, 455–461.

Southworth-Davies, R. J., Medina, M. A., Carmichael, I. & Garman, E. F. (2007). *Structure*, **15**, 1531–1541.

von Stetten, D., Giraud, T., Bui, S., Steiner, R. A., Fihman, F., de Sanctis, D. & Royant, A. (2017). *J. Struct. Biol.* **200**, 124–127.

von Stetten, D., Giraud, T., Carpentier, P., Sever, F., Terrien, M., Dobias, F., Juers, D. H., Flot, D., Mueller-Dieckmann, C., Leonard, G. a., de Sanctis, D. & Royant, A. (2015). *Acta Cryst.* D**71**, 15–26.

Studier, F. W. (2005). *Protein Expr. Purif.* **41**, 207–234.

Sutton, K. A., Black, P. J., Mercer, K. R., Garman, E. F., Owen, R. L., Snell, E. H. & Bernhard, W. A. (2013). *Acta Cryst.* D**69**, 2381–2394.

Theveneau, P., Baker, R., Barrett, R., Beteva, A., Bowler, M. W., Carpentier, P., Caserotto, H., De Sanctis, D., Dobias, F., Flot, D., Guijarro, M., Giraud, T., Lentini, M., Leonard, G. A., Mattenet, M., McCarthy, A. A., McSweeney, S. M., Morawe, C., Nanao, M., Nurizzo, D., Ohlsson, S., Pernot, P., Popov, A. N., Round, A., Royant, A., Schmid, W., Snigirev, A., Surr, J. & Mueller-Dieckmann, C. (2013). *J. Phys. Conf. Ser.* **425**, 012001.

Tsien, R. Y. (1998). *Annu. Rev. Biochem.* **67**, 509–544.

Warkentin, M., Badeau, R., Hopkins, J. B., Mulichak, A. M., Keefe, L. J. & Thorne, R. E. (2012). *Acta Cryst.* D**68**, 124–133.

Weik, M., Bergès, J., Raves, M. L., Gros, P., McSweeney, S., Silman, I., Sussman, J. L., Houée-Levin, C. & Ravelli, R. B. G. (2002). *J. Synchrotron Rad.* **9**, 342–346.

Weik, M., Ravelli, R. B. G., Kryger, G., McSweeney, S., Raves, M. L., Harel, M., Gros, P., Silman, I., Kroon, J. & Sussman, J. L. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 623–628.

Zander, U., Bourenkov, G., Popov, A. N., De Sanctis, D., Svensson, O., Mccarthy, A. A., Round, E., Gordeliy, V., Mueller-Dieckmann, C. & Leonard, G. A. (2015). *Acta Cryst.* D**71**, 2328–2343.

Zander, U., Hoffmann, G., Cornaciu, I., Marquette, J.-P., Papp, G., Landret, C., Seroul, G., Sinoir, J., Röwer, M., Felisaz, F., Rodriguez-Puente, S., Mariaule, V., Murphy, P., Mathieu, M., Cipriani, F. & Márquez, J. A. (2016). *Acta Cryst.* D**72**, 454–466.

Zeldin, O. B., Gerstel, M. & Garman, E. F. (2013). *J. Appl. Cryst.* **46**, 1225–1230.

Chapter 4

METHODS & UNPUBLISHED RESULTS
–
TIME-RESOLVED CRYSTALLOGRAPHY

Crystal injection-based Serial Synchrotron Crystallography (SSX)

–

Early developments of the Time-Resolved Serial Oscillation Crystallography (TR-SOX) method on *At*Phot2LOV2

This chapter contains unpublished results of my attempts to develop a time-resolved crystallography by first, using an injector-based SSX approach, and second, developing an approach based on the recording of oscillation data sets for which the initial time resolution was ~30 s, decreased to ~1 s and finally decreased to 63 ms. The shortest time resolution will be described in the manuscript presented in **Chapter 6**.

4.1 Crystal injection-based Serial Synchrotron Crystallography (SSX)

4.1.1 Sample preparation protocol

I have optimized the sample preparation protocol for my three proteins to the specificity of the High Viscosity Extrusion (HVE) sample injector (Botha *et al.*, 2015) that was installed on beamline ID30A-3 of the ESRF thanks to the help of Prof. Ilme Schlichting (Max Planck Institute, Heidelberg, Germany). As injection devices require a huge amount of crystalline material, 12 to 24 drops of a sitting drop crystallization plate (10 µL of volume per drop) were harvested by pipetting and then pooled in a microcentrifuge tube. Choosing the centrifugation speed and duration parameters proved to be crucial to maintain the integrity of the crystals. Besides, beamtime usage optimization put limits to the sample preparation time. In brief, crystals were centrifuged for 30 to 60 minutes at 20°C at a low spinning speed (100 to 500*g*). Depending on the buffer nature and crystal mechanical resistance, the centrifugation cycle may have to be repeated several times in order to properly pellet the crystals. The supernatant was removed between each centrifugation cycle in order to obtain a highly concentrated crystalline solution. The volume of crystal sample used in one injection run had to be 3 µL in order to be mixed with 30 µL of grease (Super Lube, No. 21030, Synco Chemical Corporation, http://www.super-lube.com/), thus in a [1/10] ratio. As a consequence, the final crystal concentration had to be adjusted depending on the hit rate results of initial runs, and if too concentrated, or not concentrated enough, crystals had to be diluted or concentrated in additional steps. A necessary step is to test that a crystal preparation of a given protein is not damaged by the injection medium, which can be achieved by visual inspection under a microscope or by eventual inspection of diffraction images on the beamline. Fortunately, in my case two out of my three proteins readily worked with the aforementioned grease. Initially, we mixed crystals with the grease using a spatula on a glass cover as described in (Sugahara *et al.*, 2015), and then transferred the resulting mixture into a glass syringe (Gastight type, Hamilton, Reno, Nevada). However, this method potentially led to a significant loss of sample. We then

modified our procedure by mixing the crystal sample with the grease using two distinct syringes. We first filled a 100 µL syringe with 30 µL of grease (**Fig. 4.1a**, green double arrow). Particular care was taken to prevent the formation of air bubbles within the medium. Secondly, the crystalline sample was collected by pipetting from the centrifuged tube and transferred into a second 100 µL syringe. Finally, both syringes were connected via a threaded coupling piece, and a thorough mixing step was performed in order to obtain a homogeneous medium evenly embedding the microcrystals (**Fig. 4.1b**, green double arrow).

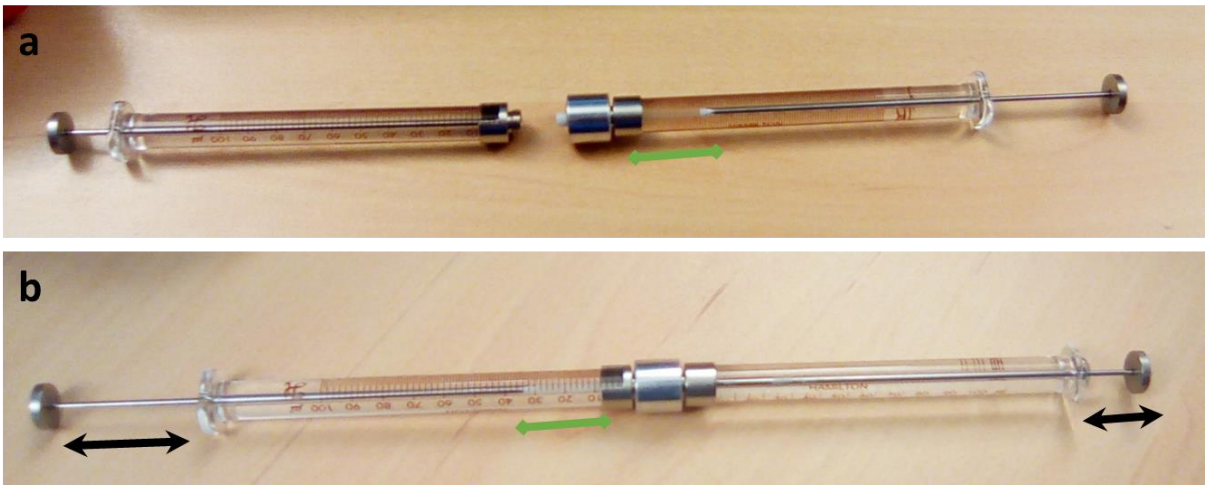


Figure 4.1: Gastight syringes and syringe coupler for media mixing. (**a**) 100 µL syringe filled with Super Lube grease (right syringe, green double arrow). The crystalline sample is transferred into the left syringe after coupling of the two syringes. (**b**) Efficient mixing is achieved by alternatively pushing on each syringe plunger (black double arrows) and visual inspection of the homogeneity of the resulting mixture (green double arrow).

4.1.2 Sample loading and operation of the HVE injector at the beamline

An exhaustive protocol for the operation of the original HVE sample injector can be found in the Supplementary Information of Botha *et al*. (2015). We adapted it to the specifics of the Microdiffractometer of beamline ID30A-3.

After disconnection of the two syringes, the one containing the [grease / crystals] mixture is screwed on the injector head and the material is loaded by pressing the plunger (**Fig. 4.2a)**. The syringe is then unscrewed and the injector is assembled as follows. The silica capillary (**Fig. 4.2b),** by which the sample will be extruded during the experiment, is screwed on the tip region of the injector using a dedicated tool, which helps to prevent breakage of the capillary (**Fig. 4.2c**). The next step consists of carefully positioning the front end of the injector

(**Fig. 4.2d**), bringing the borosilicate capillary around the silica capillary. Fine-tuning of their respective position (**Fig. 4.2e**, yellow double arrow) is critical so that the gas flux coming from the borosilicate capillary contributes to the straightening of the grease jet coming from the silica capillary. The fully assembled injector is then mounted on the Microdiffractometer MD2 of the beamline via a dedicated motorized holder (**Fig. 4.2f**).

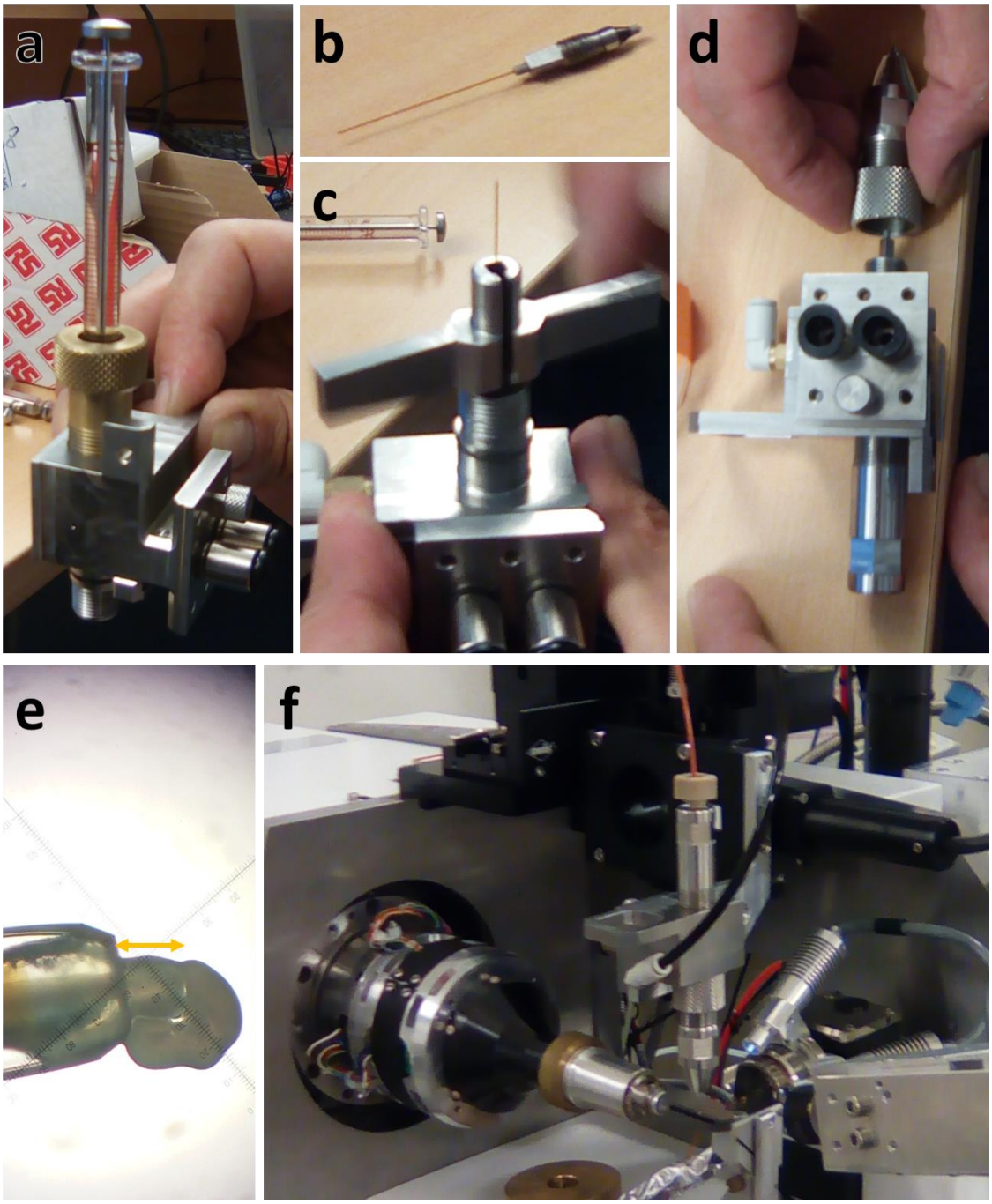


Figure 4.2: (**a**) Screwing of the syringe on the injector head and loading of the sample. (**b**), (**c**) Mounting of the silica capillary conducting the crystal-containing viscous medium. (**d**), (**e**) Positioning of the gas-flowing borosilicate capillary, which serves to stabilize the grease jet. (**f**) Fully-assembled HVE sample injector mounted on the Microdiffractometer of beamline ID30A-3 of the ESRF.

The injector is connected at the top to a water tubing (**Fig. 4.3a**, blue arrow) coming from an HPLC system, which applies a constant flow rate (10 µL.min^{-1}) necessary to generate a continuous and straight grease jet of 100 µm diameter (micro-stream), which corresponds to ~ 3 µL.min^{-1} volumetric speed at the output of the capillary. The injector is also connected on its side to a gas (N_2 or He) tubing (**Fig. 4.3a**, green arrow), which applies a constant pressure of typically 0.3 bar. An air booster can transiently increase the pressure to 4.0 bar in order to clear any accumulation of grease at the nozzle. The injector is supported by a motorized stage on the beamline. Those motors are used to align the jet in the three dimensions in front of the X-ray beam. The alignment is constantly monitored *via* the MD2 video camera. The injector was then vertically adjusted so that the X-ray beam position is located ~100 µm down from the end of the silica capillary (**Figs. 4.3b** and **4.3c**), in order to probe the jet in its most stable zone (further down, the ambient air and/or static electricity affect the shape of the jet).

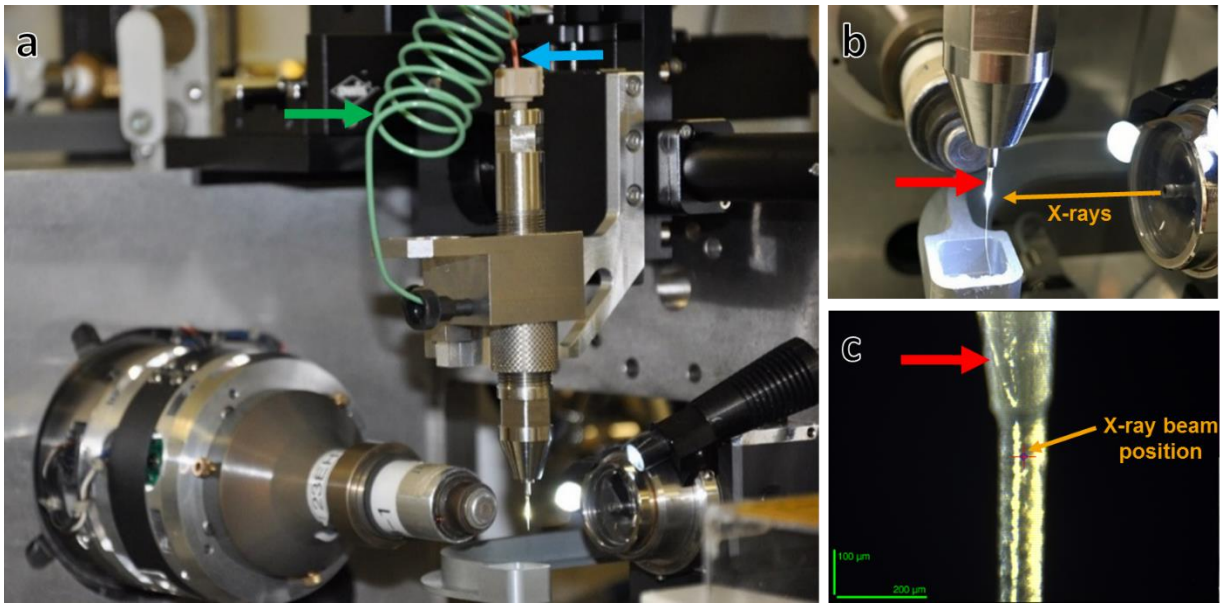


Figure 4.3: (**a**) HVE sample injector mounted on the beamline Microdiffractometer, with connections to the water flux coming from an HPLC (blue arrow) and to a gas flux (green arrow). (**b**), (**c**) Close-up on the grease extrusion area with the capillary indicated with the red arrow.

4.1.3 Data collection, data integration and structure refinement

Successive data sets of 100,000 still images were recorded using an Eiger 4M pixel detector operated at a 200 Hz acquisition rate, with the unattenuated X-ray beam, until the grease jet stopped. Typically, samples prepared as described above could be delivered for 30 to 60 minutes, depending on stochastic events such as formation of air bubbles or curling of the grease jet requiring transient pause in the injection process. The key parameter in an SSX experiment is the fraction of the 100,000 images which contain diffraction spots from a microcrystal. This fraction is called the 'hit rate' and should not be too high (less than 30%, or the chances of getting diffraction from two or more crystals become significant) or too little (a few percent or lower, since the experimental time and the quantity of injected crystals would need to increase dramatically). We aimed to obtain several thousands of hits for a given sample. The software *NanoPeakCell* (Coquelle *et al.*, 2015) was used to identify and select hits, that is images with enough diffraction spots. *CrystFEL* (White *et al.*, 2016) was then used to integrate the selected images. Once a data set could be constituted, we performed molecular replacement using *PHASER* (McCoy *et al.*, 2007) and refined using *REFMAC5* (Murshudov *et al.*, 1997).

4.1.4 Results obtained from Twist-Cerulean microcrystals

Our SSX experiment on Twist-Cerulean microcrystals was the first one that we performed. We experienced issues during this first experiment. In particular, the borosilicate capillary was damaged, which led to perturbation of the jet and gradual accumulation of grease at the top of capillary. First, it is possible that it may have resulted in multiple crystal lattices on the same image for certain hits, but most importantly it led to an increase in grease scattering (**Figs. 4.4a** and **4.4b**), which decreased the signal-to-noise ratio of reflections in this area of the detector, hampering hit detection in the hit sorting software.

From the 876,571 collected images, 21,939 hits could be indexed and integrated (~2.5% hit rate), leading to a 2.2 Å data set with acceptable statistics (**Table 4.1**). We could refine the structure (**Fig. 4.4c**) and compare it with structures determined from cryogenic and room temperature oscillation data sets (**Table 4.2**). This experiment constituted a proof of principle that we could solve a room temperature structure on beamline ID30A-3 using an HVE sample injector.

Table 4.1: Data-collection statistics for SSX data sets obtained from Twist-Cerulean microcrystals. Values in parentheses are for the highest resolution shell.

Twist-Cerulean	SSX
Space group	$P2_12_12_1$
Resolution [Å]	47.3 – 2.20 (2.28 – 2.20)
Number of images	876,571
Number of crystals	21,939
Hit rate [%]	2.5
Multiplicity	191 (128)
R_{split} [%]	15.3 (66.7)
CC* [%]	98.9 (89.4)
Signal-to-noise	4.5 (1.4)
Completeness [%]	100.0 (100.0)

Table 4.2: Refinement statistics for Twist-Cerulean data sets obtained with oscillation data collections on two ~500 µm crystals at 100 K and 298 K, respectively, and obtained by SSX on ~10 µm microcrystals.

Twist-Cerulean	100 K	298 K	SSX
R_{work} [%]	13.1	12.9	18.4
R_{free} [%]	15.0	17.1	23.2
Rmsd bond length [Å]	0.010	0.007	0.011

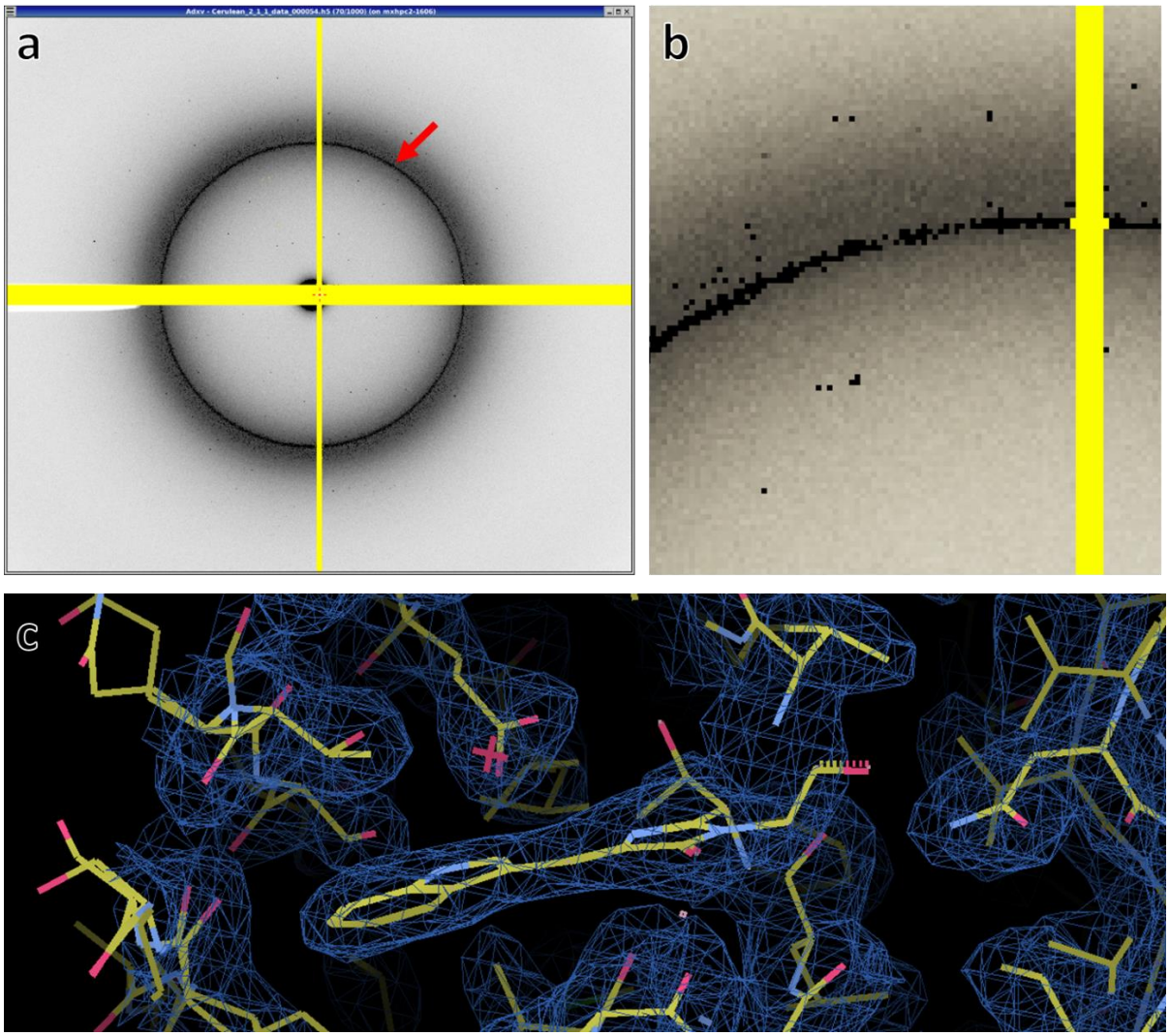


Figure 4.4: (**a**) Diffraction pattern from a Twist-Cerulean microcrystal collected in an injector-based SSX experiment. (**b**) Close-up on the grease ring. (**c**) 2*Fo*-*Fc* electron density map (at a 1 σ level) superimposed on the model of the Twist-Cerulean chromophore region.

4.1.5 Results obtained from *Dr*CBD crystals

My *Dr*CBD crystals illustrate one of the main difficulties of performing medium injection SSX, which is the necessity of preserving diffraction properties throughout all the preparative steps, from the crystallization condition all the way to the embedding in the carrying medium. While our SSX experimental conditions did not appear to affect much the diffraction power of crystals of Twist-Cerulean and *At*Phot2LOV2 (see thereafter), *Dr*CBD crystals never yielded any image with a detectable diffraction pattern. We could only hypothesize that the crystals got damaged in the pipetting or centrifugation steps, or that the grease medium perturbed the crystalline order.

4.1.6 Results obtained from *At*Phot2LOV2 crystals

The goal of our SSX experiments with *At*Phot2LOV2 microcrystals was to first check if we could determine the structure of the dark state, and then to determine the structure under blue light steady-state irradiation. Illumination was ensured by the addition of a 470 nm LED focused on the sample through a 5x magnifying objective, which could be precisely positioned thanks to X, Y, Z manual translation stages (**Fig. 4.5b)**. We recorded 250,000 images on non-illuminated crystals. ~5000 images could be indexed and integrated, leading to the structure determination at 2.1 Å resolution (data reduction and refinement statistics not shown). While we expected a dark structure, inspection of the $2F_{obs} - F_{calc}$ electron density map revealed that a significant proportion of the light state had built up (**Fig. 4.6a**). The explanation resides in the fact that in order to monitor the behaviour of the grease jet, we had left the front light of the MD2 switched on, which has a blue component thus contributing to the build-up of the light state. As a consequence, we modified our protocol in the next experiment by switching off the MD2 front light.

In our second experiment, we attempted to determine the structure of both states. The data collection timeline is depicted on **Fig. 4.5a**, with a total of 870,000 images collected from a single injection. The first 502,000 images were collected in the dark (**Fig. 4.5c** showing the jet with the back light before the start of the data collection), and the next 368,000 under LED illumination (**Fig. 4.5d** and **Table 4.3**).

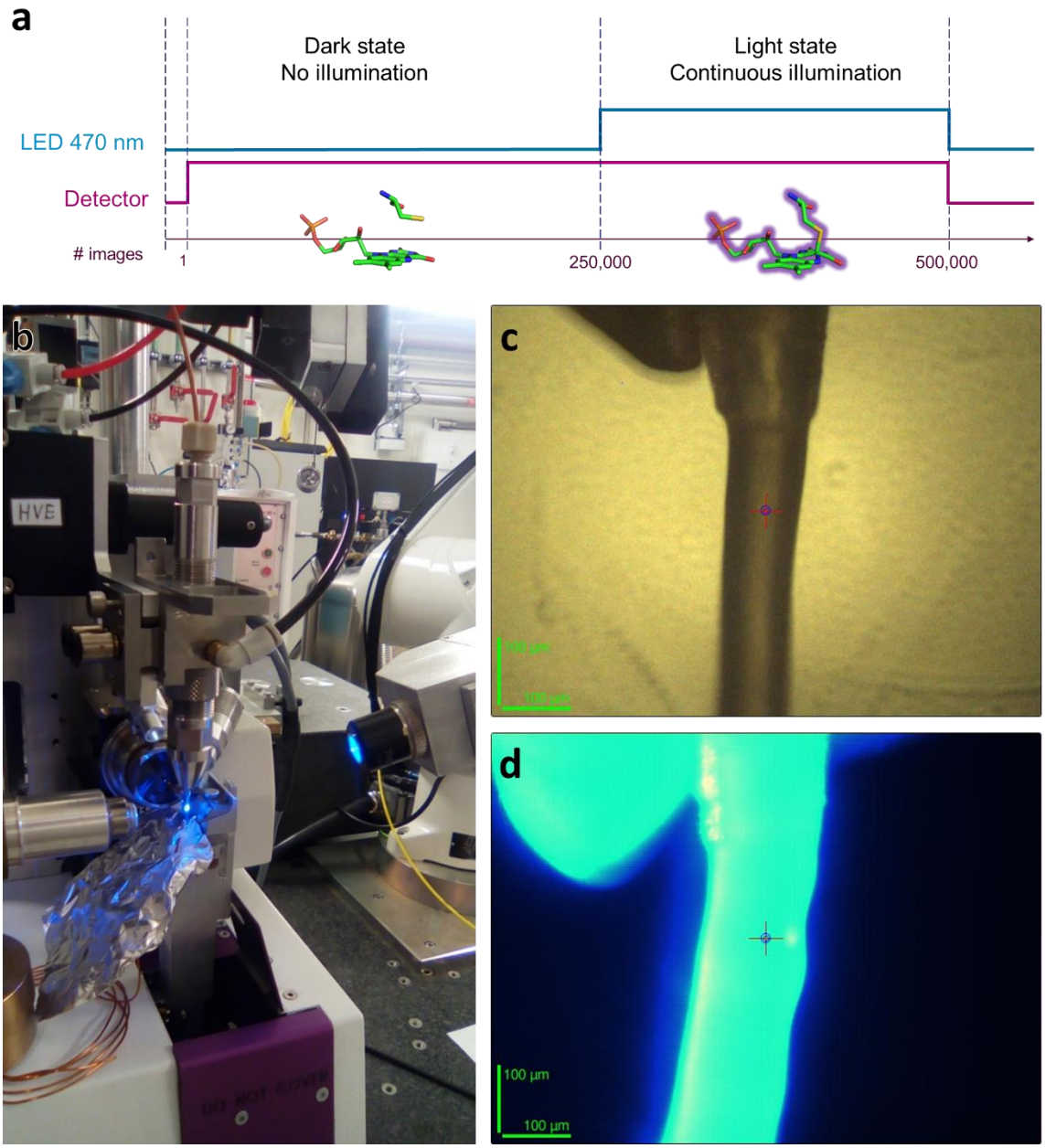


Figure 4.5: (**a**) Principle of the SSX experiment on *At*Phot2LOV2 microcrystals. (**b**) SSX setup on beamline ID30A-3 under LED illumination. (**c**) Grease jet visualized with the back light before data collection. (**d**) Grease jet under LED illumination.

We aimed to collect diffraction data for the two states from a single injection run in order to minimize artefacts that may arise from any of the sample preparation steps, *e.g.* the quantity of harvested crystals or the efficiency of embedding microcrystals in the grease. 857 diffraction patterns could be indexed and integrated at 2.7 Å for the dark state data set ($2F_{obs} - F_{calc}$ electron density map shown on **Fig. 4.6b**). However, only 416 diffraction patterns were successfully indexed and integrated for the light state data set, resulting in a lower resolution of 3.0 Å ($2F_{obs} - F_{calc}$ electron density map shown on **Fig. 4.6c**). The lower hit rates of the

second experiment compared to the first one (separated by several months) has to be explained by protein batch and crystallization variability, which led to smaller (average size of ~1 µm *vs.* ~10 µm) and fewer crystals. The two maps, contoured at the same σ level (**Figs. 4.6b** and **4.6c**), evidently show two different situations, although the phases are given by the same dark state model. From the dark state data, there is clearly no electron density between the cysteine and the chromophore. However, from the light state data, there is electron density continuity, strongly suggesting that the light state has been indeed built up.

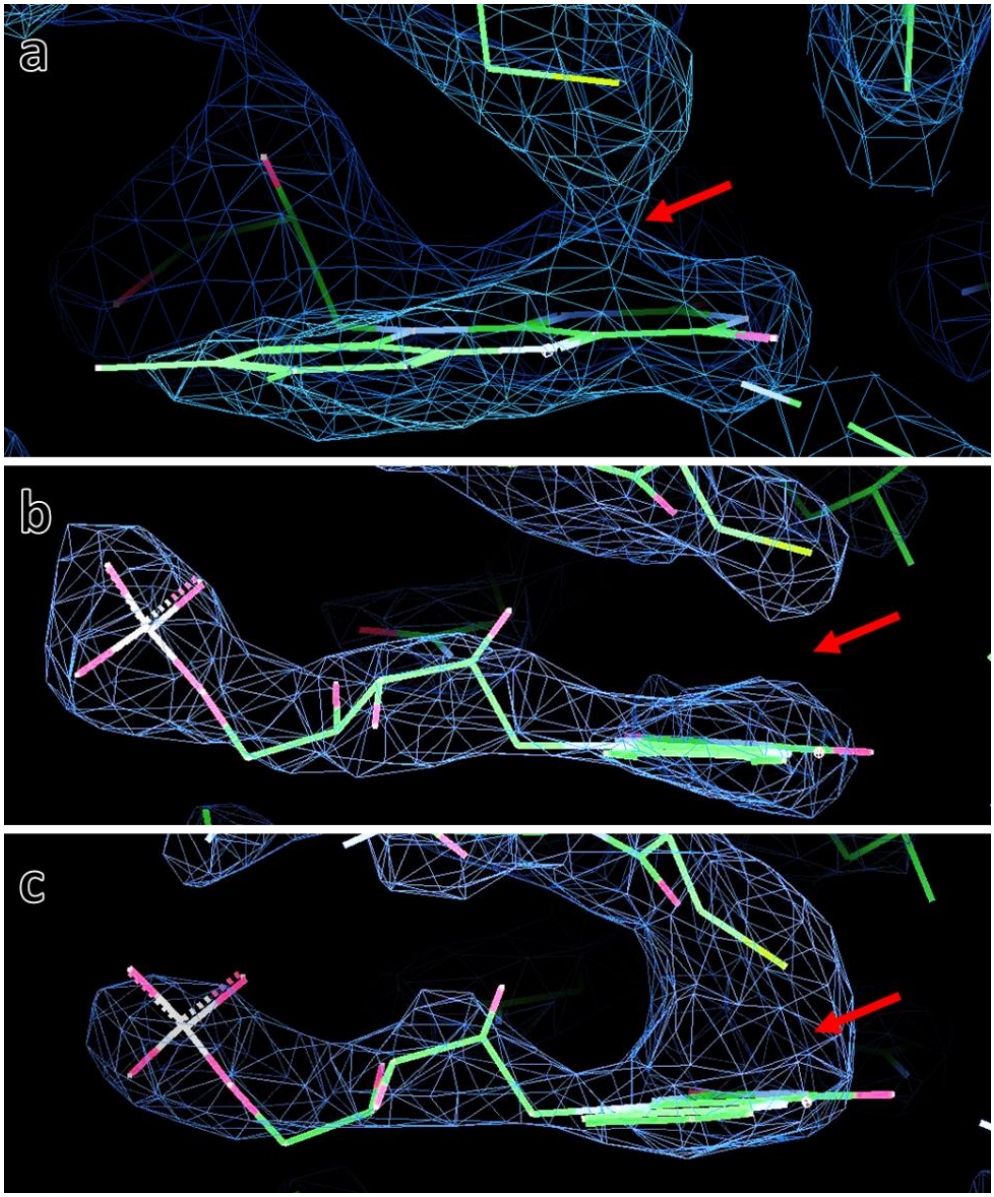


Figure 4.6: $2F_{obs} - F_{calc}$ electron density maps around the chromophore FMN and Cys426 showing the potential build-up of a covalent bond (red arrow). (**a**) Results of the first experiment (2.1 Å electron density map contoured at a 1.0 σ level). (**b**) Results of the second experiment – dark state data collection (2.7 Å electron density map contoured at a 1.5 σ level). (**c**) Results of the second experiment – light state data collection (3.0 Å electron density map contoured at a 1.5 σ level).

Table 4.3: Data-collection statistics for SSX data sets obtained from *At*Phot2LOV2 microcrystals. Values in parentheses are for the highest resolution shell.

***At*Phot2LOV2**	**SSX[dark]**	**SSX[light]**
Space group	$P4_32_12$	$P4_32_12$
Resolution [Å]	39.5 – 2.7 (2.80 – 2.70)	39.5 – 3.0 (3.10 – 3.00)
Number of images	502,000	368,000
Number of crystals	857	416
Hit rate [%]	0.17	0.11
Multiplicity	14 (9)	7.3 (4.5)
Rsplit [%]	31.6 (157.8)	31.1 (150.4)
CC* [%]	97.8 (55.2)	96.1 (22.9)
Signal-to-noise	2.5 (0.8)	3.1 (0.7)
Completeness [%]	99.7 (99.5)	90.5 (73.9)

Our SSX results on *At*Phot2LOV2 are very encouraging since they show that we were capable of resolving the formation of the covalent adduct between the protein and its chromophore, even with data sets coming from a rather small number of crystals. With persistence, in particular in reproducing the quality of the microcrystals from the first experiment, we would have eventually managed to determine the structure of the dark, and of the light state in photostationary conditions at a resolution close to 2.0 Å. These results could have paved the way to envisage pump-probe experiments using the HVE sample injector, but we decided to focus on a different approach to perform time-resolved crystallography experiments (**Paragraph 4.2** and **Chapter 5 and 6**).

.

4.2 Early developments of the Time-Resolved Synchrotron Oscillation Crystallography (TR-SOX) method on *At*Phot2LOV2

The experiments aimed at identifying specific radiation damage at room temperature on the Cyan Fluorescent Protein Cerulean and on Hen Egg White Lysozyme (**Chapter 3**) suggested that it should be possible to visualize at room temperature the presence of the chromophore-cysteine covalent bond in a LOV domain, which had been shown to be highly sensitive to X-rays at cryogenic temperatures (Fedorov *et al.*, 2003). Consequently, if one was able to control the absorbed dose to be able to record partial or complete data sets, one should be able to determine its structure in a time-resolved manner.

From our knowledge of SSX, we were hopeful to investigate if recording full, or partial oscillation data sets could provide enough information from only a relatively small number of crystals, in stark contrast with the SSX- or SFX-base methods requiring millions of microcrystals. Upon successful development of our method (**Chapter 6**), we named it TR-SOX for time-resolved synchrotron oscillation crystallography.

4.2.1 Spectroscopic characterization of the dark and light states of *At*Phot2LOV2

Before any diffraction experiment, we first characterized how to build-up the light state in *At*Phot2LOV2 crystals, using *in crystallo* UV-visible light absorption spectroscopy at the ID29S-Cryobench laboratory (von Stetten *et al.*, 2015), which has been rebranded as the *ic*OS Lab in year 2019 (*ic*OS standing for *in crystallo* Optical Spectroscopy). Crystals were put in the wet stream of an HC1 humidity controller and illuminated for several minutes. We initially used a 440 nm laser as the actinic source, and were surprised to see the build-up of a species absorbing between 550 and 650 nm (**Fig. 4.7a**, green arrow), which resembled the triplet state (rise time ~ns, lifetime ~µs) (Swartz *et al.*, 2001). We were puzzled by the possibility that the triplet state could be stabilized at room temperature for at least several minutes. We then switched to a 470 nm LED and observed the build-up of the expected 390 nm peak characteristic of the light state (**Fig. 4.7b**, red arrow). We recorded diffraction data from a 440 nm-irradiated crystal of LOV2 in a 'trigger-freeze' approach (Bourgeois & Royant, 2005), but did not manage to identify any structural change. As a consequence, we did not investigate further this red-shifted species.

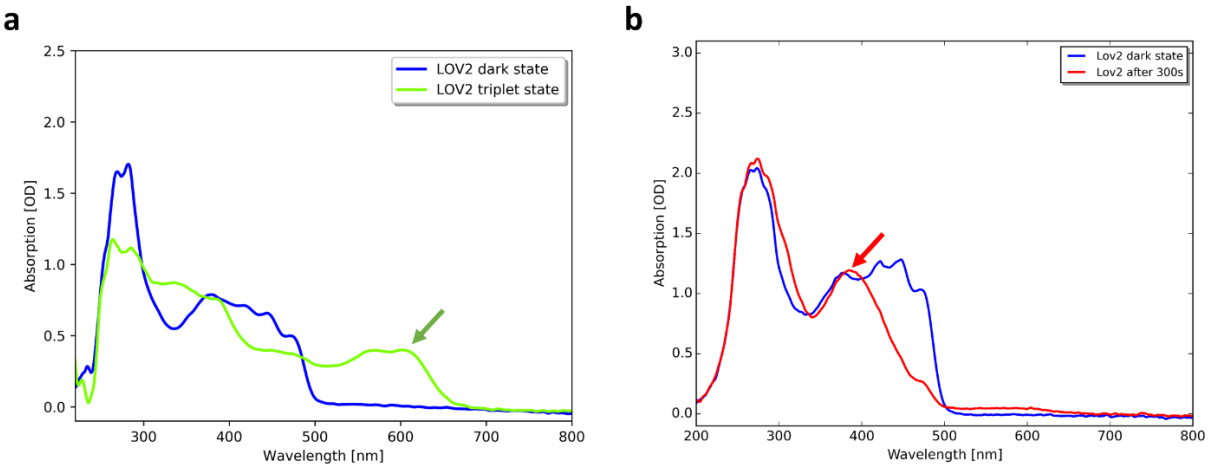


Figure 4.7: Room temperature UV-visible absorption spectra of *At*Phot2LOV2 crystals upon (**a**) 440 nm laser excitation showing the signature of the triplet state (green arrow) and (**b**) 470 nm actinic LED light showing the signature of the light state (red arrow). Note: the dark state spectrum on the left contains a contribution of the light state, which must have been triggered by the visualisation light.

4.2.2 Proof of principle experiment: 30 s time resolution

The room temperature specific radiation damage study led us to hypothesize that if we could record data fast enough from a photo-irradiated crystal of *At*Phot2LOV2, we should be able to visualize the structure of the light state. The data collection needed to be fast enough compared to the relaxation time of the light state, known to be between 60 to 180 s (Swartz *et al.*, 2001). The Pilatus 6M of beamline ID29 could record a full data set in less than 30 s. We thus mounted an *At*Phot2LOV2 crystal on the MD2 diffractometer put under the wet air stream of the de-humidifier HC1b (Sanchez-Weatherby *et al.*, 2009), thus at room temperature, and taped an optical fibre close to the sample position, connected to a 470 nm LED connected *via* the laser PSS (Personal Safety System) of the beamline (**Fig. 4.8**, top). We then programmed two consecutive 30 s data collections in the control software *MxCuBE* (Gabadinho *et al.*, 2010), and switched off the laser PSS key one to two seconds before opening of the X-ray shutter, effectively shutting down the LED illumination, and initiating relaxation of the light state in the crystal (**Fig. 4.8**, bottom). We did that in order to prevent LED illumination from re-populating the light state fraction which could have been quickly destroyed by X-rays. As hoped, the structure determined from the first data set showed a significant proportion of light state (~50%) and that from the second data set showed a lower proportion (~30%), effectively demonstrating that X-rays at room temperature did not induce specific damage comparably quickly as at cryogenic temperature, once normalized by global damage (**Chapter 3**). In

addition, this experiment constitutes our first genuine time-resolved crystallography experiment, with a time resolution of 30 s, that is the duration of a complete data collection with a Pilatus 6M. To study the same phenomenon (the relaxation of the light state) at a faster rate, we took advantage of the Eiger 4M detector that became available later on beamline ID30A-3, and decreased the time resolution to 3 s, which is more than enough to study a phenomenon that is completed after 200 s (**Chapter 5**). In order to study a faster phenomenon (the build-up of the light state population), we could easily decrease the duration of a full data collection to 1.68 s close to the 1 s time resolution, a psychological barrier in time-resolved protein crystallography (**Paragraph 4.2.3**).

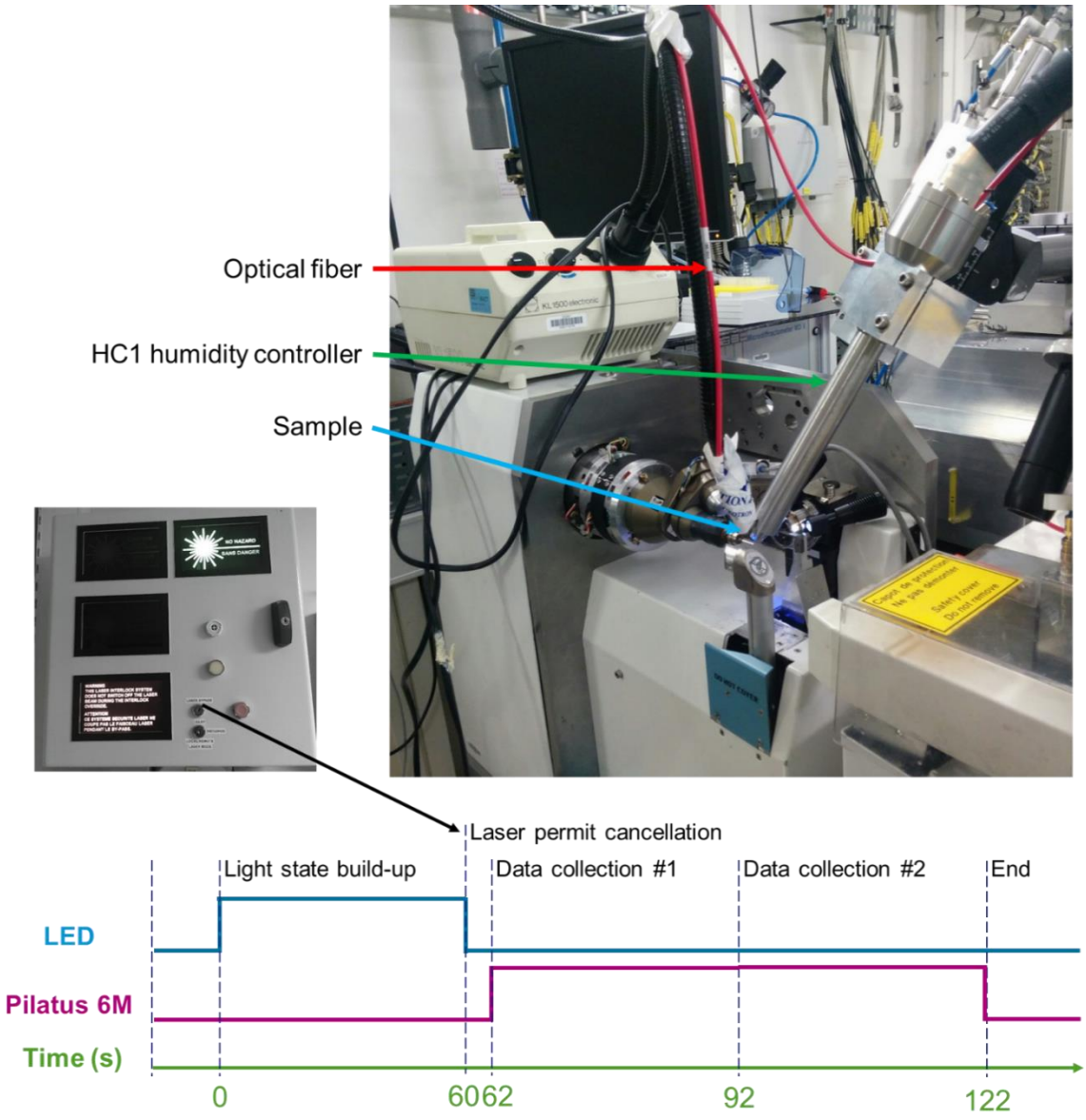


Figure 4.8: Experimental setup and timeline on beamline ID29 showing the optical fibre conducting the light from the LED. Turning the key on the Laser Personal Safety System (PSS) outside the hutch, effectively cancels the laser permit, which turns off the LED power supply.

4.2.3 Sub-second time-resolved crystallography

4.2.3a Monitored phenomenon

In order to test the capabilities of the Eiger 4M detector to study the light state relaxation in detail, we illuminated crystals on the diffractometer of beamline ID30A-3. We were initially surprised to see on the video camera that the crystal did not instantly changed colour upon remotely-triggered turning on of the LED (**Fig. 4.9a**), since the light state builds up on the microsecond time scale. We eventually reasoned that the LED was not providing enough photons to convert all molecules within microseconds, and that this was the limiting factor for photoconverting the whole crystal from the blue-light absorbing (thus yellow) dark state species to the UV-light (thus colourless) light state species (**Fig. 4.9b**). Noteworthy, the dark state is fluorescent under blue light illumination, while the light state is not. The full spectroscopic demonstration of the effect of the LED irradiation power is presented in **Chapter 6**. As a consequence, we were intrigued to see if we could monitor by time-resolved crystallography the build-up of the light state population within an *At*Phot2LOV2 crystal, which seemed to occur on the timescale of a few seconds.

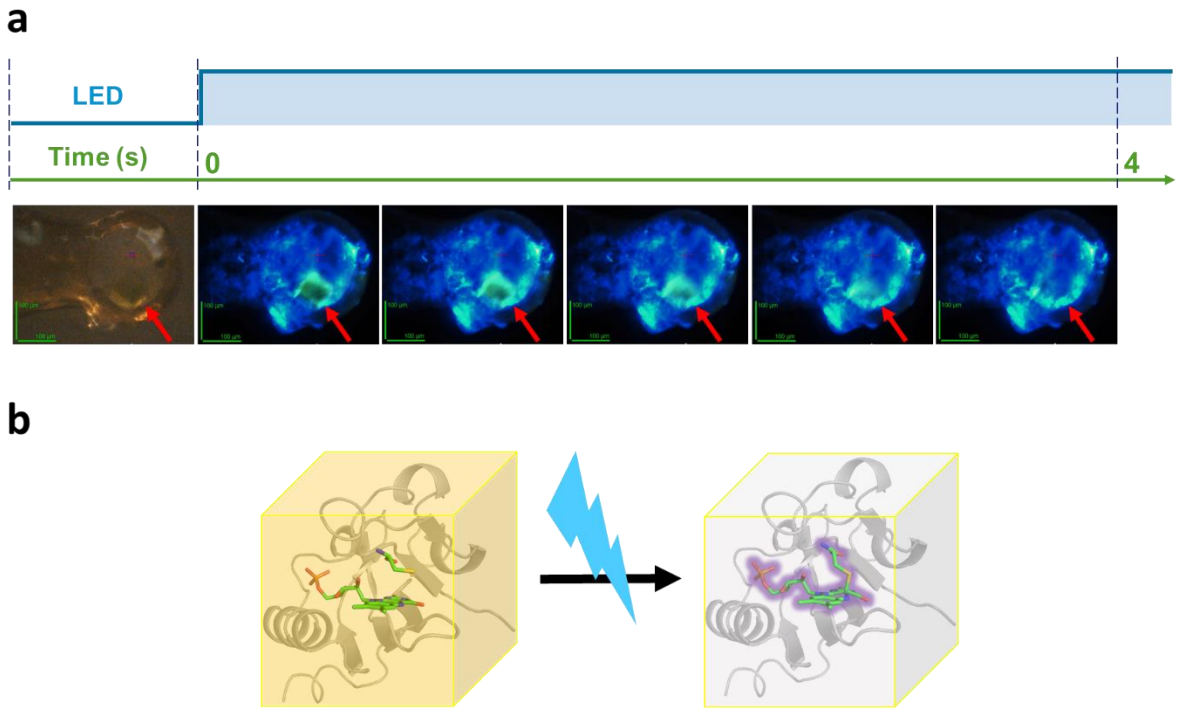


Figure 4.9: (**a**) Slow crystal discolouration upon turning on of the LED. The crystal position is indicated with a red arrow. (**b**) Schematic representation of the discolouration of an *At*Phot2LOV2 crystal due to the dark to light state photoconversion.

4.2.3b Experimental setup

The illumination setup is identical to that used for the SSX experiment (**Paragraph 4.1.6**) and is fully depicted on (**Fig. 4.10**). Crystals were maintained at room temperature with the HC1 humidity controller set at a relative humidity of 98% (see **Chapter 2, Paragraph 2.3.2**).

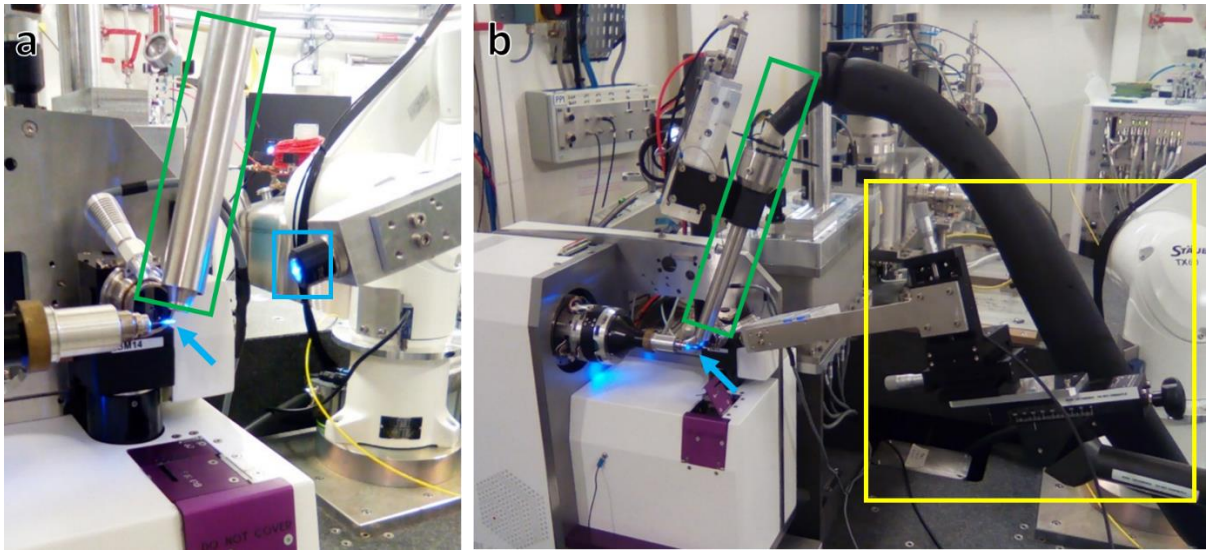


Figure 4.10: (**a**) and (**b**) Overview of the experimental setup for TR-SOX experiments on beamline ID30A-3. A loop containing an *At*Phot2LOV2 crystal (blue arrow) is placed in the wet stream of a HC1 humidity controller (green rectangle), under the illumination of a 470 nm LED light focused thanks to a 5x magnifying objective (blue rectangle). The LED is precisely positioned by X, Y, Z translation stages (yellow rectangle).

4.2.3c Data collection and data analysis strategy

The first step was to record a dark state structure to be used as a reference from a 100° wedge of data recorded on one crystal, which was chosen. In a second step, a 1,000° wedge of data was collected at the same location, with the turning on of the LED synchronised with the TTL output of the detector signalling the start of the data collection. The X-ray flux was minimized in order to limit diffraction decay over the whole course of the data collection. Data collection parameters are summarized in **Table 4.4**.

Table 4.4: Data collection parameters.

Experiment goal	Experiment / beamline parameters												
	ESRF Beamline	Temperature (K)	Flux (photons /sec)	Transmission (%)	Wavelength (Å)	Energy (KeV)	Phi start (total) (°)	oscillation range (°)	n° images	Exposure time (s)	Total exposure time (s)	Detector Resolution (Å)	Humidifier set at (%)
Dark state	ID30A-3	293	3,87E+10	0,25	0,9677	12,812	0 (100)	0,5	200	0,00417	0,834	1,53	98
Light state	ID30A-3	293	3,17E+10	0,25	0,9677	12,812	0 (1000)	0,5	2000	0,00417	8,34	1,53	98

We restricted our data analysis to the first 500° (1000 images) to study the first 4 seconds in detail. In order to provide enough time points, we decided to use a 'sliding' 400-image data set, starting from image 1, 101, […] and 601 (**Fig. 4.11**). Each of the 7 resulting data sets correspond to time intervals [0 s – 1.68 s], [0.84 – 2.10 s], […] and [2.52 s – 4.20 s]. For simplicity we named each of the time point by the centre of the time interval: t_1 = 0.84 s, t_2 = 1.26 s, […] and t_7 = 3.36 s. These time points are not truly independent, but since we are attempting to monitor an increase in population, we are only adding a time average to a population average description, and are still producing a molecular movie of a phenomenon. The pseudo time resolution could be improved by using a 50-frame or less sliding window. The genuine time resolution could be improved by increasing the speed of the goniometer, which was achieved at a later date (implementation of a Power PMAC electronics on the MD2 diffractometer, increasing the maximum speed from 120°/s to 500°/s).

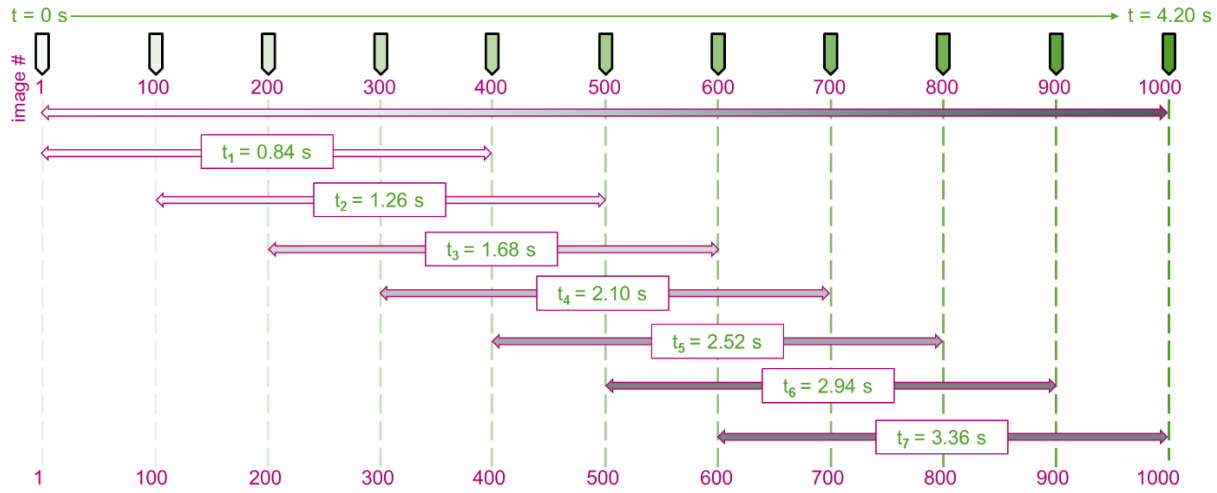


Figure 4.11: Data analysis strategy to produce a time-dependent evolution in a single *At*Phot2LOV2 crystal upon blue light illumination.

Data sets were integrated and scaled with *XDS* (Kabsch, 2010) and *XSCALE*, using I/σ(I) and $CC_{1/2}$ cutoffs of 2.0 and 0.75, respectively, resulting in resolutions ranging from 2.68 Å to 2.87 Å. The variation in resolution can be most attributed to a less well-diffracting zone of the crystal likely to arise from a suboptimal crystal centring (for example for the region corresponding to images 200 to 600) (**Fig. 4.12)**.

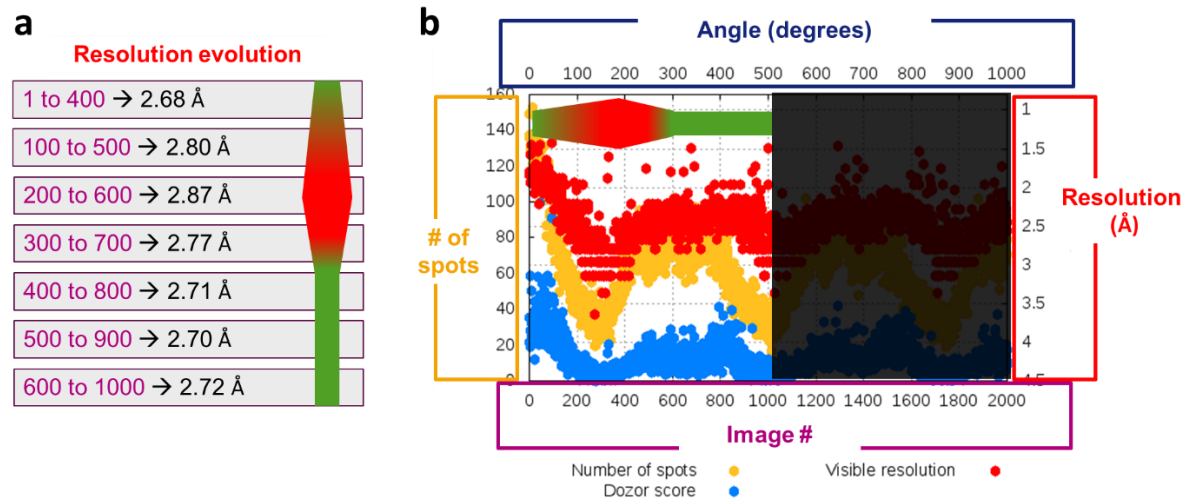


Figure 4.12: (**a**) Evolution of diffraction resolution for the seven pseudo time points (1000 first images). The lowest resolution is observed for the three data sets containing images ~250 to ~450 (red zone). (**b**) *DOZOR* graph plotting the evolution of data collection indicators: number of detected diffraction spots, apparent diffraction resolution and *DOZOR* score.

Structures were refined with a script using the dark state structure as a starting model. $2F_{obs} - F_{calc}$ electron density maps were calculated for each time point t_1 to t_7 and represented at various σ levels (from top to bottom of **Fig. 4.13**: 1.5, 2.0 and 2.5 σ, respectively). The various σ levels allow the reader to quickly visualize the progressive build-up of the covalent bond formed between Cys426 and the chromophore FMN. Red arrows indicate the time points for which the bonding is not yet obvious.

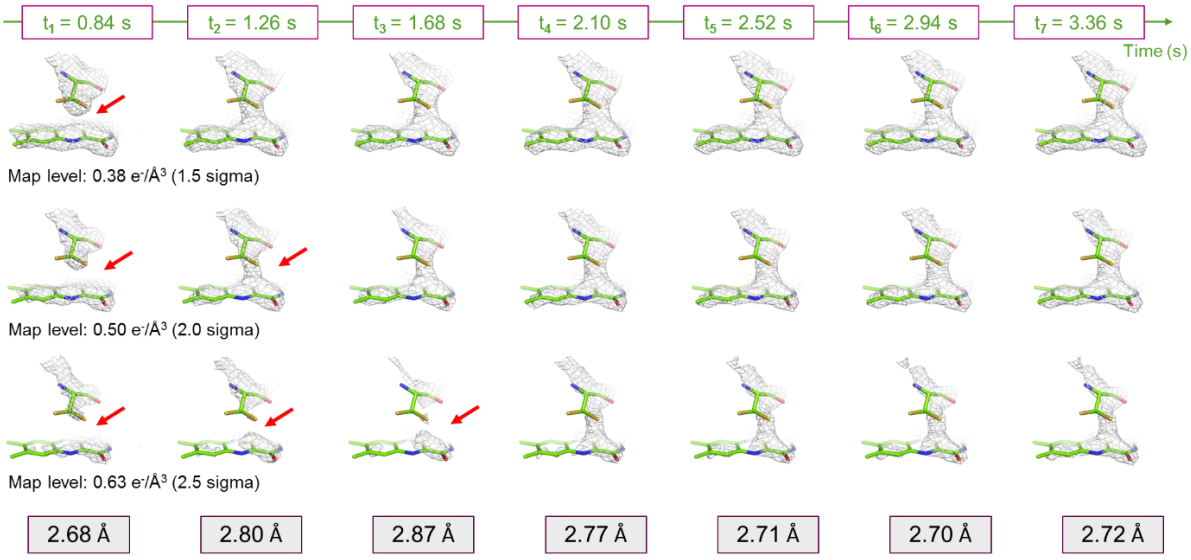


Figure 4.13: $2F_{obs} - F_{calc}$ electron density maps between Cys426 and the chromophore FMN at the 7 various time points, contoured from top to bottom at a 1.5, 2.0 and 2.5 σ level, respectively. The red indicate time points for which the bonding is not yet obvious.

In our first attempt to quantify the speed of light state population build-up, we monitored the maximum σ level of each $2F_{obs} - F_{calc}$ map for which the electron density is discontinuous, and plotted it as a function of time (**Fig. 4.14**). We also plotted the peak height of $2F_{obs[Light(n)]} - F_{obs[Dark]}$ as a function of time, which confirmed the general trend of the first curve (data not shown). While this clearly appears as a naive, oversimplified way of describing at a molecular level the phenomenon that we witnessed on the video camera, this constitutes our first attempt to generate a molecular movie of the phenomenon.

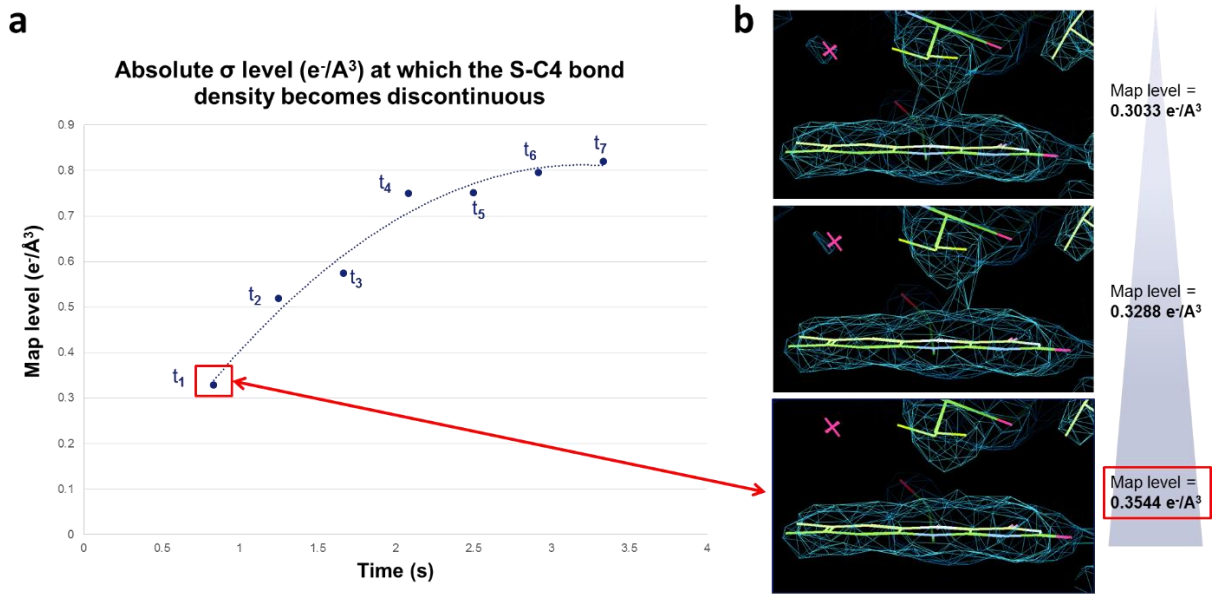


Figure 4.14: (**a**) Representation of the maximum σ level (on the absolute scale in $e^-/Å^3$) for which the electron density between the chromophore and the cysteine is discontinuous, as (**b**) determined from σ scrolling in *COOT*.

4.2.3d Conclusion and perspectives

This experiment based on a sliding complete oscillation data set recorded on one given crystal gave us the idea to envisage a serial crystallography approach based on the repetition of the same experiment on single crystals, in which complete oscillation data sets would be separated into partial data sets, used to reconstruct full, non-overlapping data sets representative of short time ranges. We were eventually able to decrease the resolution of our time-resolved experiment down to 63 ms (**Chapter 6**).

4.3 References

Botha, S., Nass, K., Barends, T. R. M., Kabsch, W., Dworkowski, F., Foucar, L., Panepucci, E., Wang, M., Shoeman, R. L., Schlichting, I. & Doak, R. B. (2015). *Acta Cryst. D*. **71**, 387–397.

Bourgeois, D. & Royant, A. (2005). *Curr. Opin. Struct. Biol.* **15**, 538–547.

Coquelle, N., Brewster, A. S., Kapp, U., Shilova, A., Weinhausen, B., Burghammer, M. & Colletier, J. P. (2015). *Acta Cryst. D*. **71**, 1184–1196.

Fedorov, R., Schlichting, I., Hartmann, E., Domratcheva, T., Fuhrmann, M. & Hegemann, P. (2003). *Biophys. J.* **84**, 2474–2482.

Gabadinho, J., Beteva, A., Guijarro, M., Rey-Bakaikoa, V., Spruce, D., Bowler, M. W., Brockhauser, S., Flot, D., Gordon, E. J., Hall, D. R., Lavault, B., McCarthy, A. A., McCarthy, J., Mitchell, E., Monaco, S., Mueller-Dieckmann, C., Nurizzo, D., Ravelli, R. B. G., Thibault, X., Walsh, M. A., Leonard, G. A. & McSweeney, S. M. (2010). *J. Synchrotron Rad.* **17**, 700–707.

Kabsch, W. (2010). *Acta Cryst. D*. **66**, 125–132.

Mccoy, A. J., Grosse-kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst*. **D53**, 240–255.

Sanchez-Weatherby, J., Bowler, M. W., Huet, J., Gobbo, A., Felisaz, F., Lavault, B., Moya, R., Kadlec, J., Ravelli, R. B. G. & Cipriani, F. (2009). *Acta Cryst. D*. **65**, 1237–1246.

von Stetten, D., Giraud, T., Carpentier, P., Sever, F., Terrien, M., Dobias, F., Juers, D. H., Flot, D., Mueller-Dieckmann, C., Leonard, G. a., de Sanctis, D. & Royant, A. (2015). *Acta Cryst. D*. **71**, 15–26.

Sugahara, M., Mizohata, E., Nango, E., Suzuki, M., Tanaka, T., Masuda, T., Tanaka, R., Shimamura, T., Tanaka, Y., Suno, C., Ihara, K., Pan, D., Kakinouchi, K., Sugiyama, S., Murata, M., Inoue, T., Tono, K., Song, C., Park, J., Kameshima, T., Hatsui, T., Joti, Y., Yabashi, M. & Iwata, S. (2015). *Nat. Methods*. **12**, 61–63.

Swartz, T. E., Corchnoy, S. B., Christie, J. M., Lewis, J. W., Szundi, I., Briggs, W. R. & Bogomolni, R. A. (2001). *J. Biol. Chem.* **276**, 36493–36500.

White, T. A., Mariani, V., Brehm, W., Yefanov, O., Barty, A., Beyerlein, K. R., Chervinskii, F., Galli, L., Gati, C., Nakane, T., Tolstikova, A., Yamashita, K., Yoon, C. H., Diederichs, K. & Chapman, H. N. (2016). *J. Appl. Cryst.* **49**, 680–689.

Chapter 5

RESULTS – ARTICLE 2

Room temperature time-resolved crystallography on a plant photoreceptor domain with a 3 s time resolution

–

Monitoring of the slow relaxation of the light state of *At*Phot2LOV2

Presentation of Article 2

In **Chapter 3**, we showed that it is possible to determine at room temperature the structure of the light state of the plant photoreceptor domain *At*Phot2LOV2, without significantly affecting the covalent bond between the FMN chromophore and a nearby cysteine residue, a bond that we showed was particularly sensitive to X-rays at cryogenic temperature. This gave us the idea to try to follow the relaxation of the light state of *At*Phot2LOV2 by time-resolved crystallography on a slow time scale. Our initial experiment, described in **Chapter 4**, **Paragraph 4.2.2** used a Pilatus 6M as X-ray detector, effectively providing us with a ~30 s time resolution. The comparison of the two data sets successively recorded at [0-30 s] and [30-60 s] showed that the light state was still significantly present in the second data set, yet in a lower proportion as expected. The availability of the faster Eiger 4M detector, operated at a 238 Hz frame rate, gave us the opportunity to monitor the relaxation of the light state with a 3 s time resolution. The following manuscript describes our initial findings, and additional structure refinement of all 19 time points is necessary before considering submission to a journal.

Room temperature crystallography visualisation of LOV domain photoadduct relaxation on the second to minute time scale

Sylvain Aumonier[a], David von Stetten[a,$], Guillaume Gotthard[a], Gordon Leonard[a], Antoine Royant[ab]*

[a]European Synchrotron Radiation Facility, Grenoble, F-38043, France

[b]Univ. Grenoble Alpes, CNRS, CEA, IBS (Institut de Biologie Structurale), Grenoble, F-38000, France

*Correspondence email: antoine.royant@ibs.fr

[$]Present address: European Molecular Biology Laboratory (EMBL) c/o DESY, Notkestrasse 85, D-22607 Hamburg, Germany

Abstract The LOV2 domain of phototropin 2 from *Arabidopsis thaliana* undergoes a photocycle in which the signalling state $LOV2_{390}$ builds up on the microsecond time scale under blue light illumination but decays on the minute time scale, *i.e*. 8 orders of magnitude slower. We have monitored the relaxation of $LOV2_{390}$ in the dark by X-ray crystallography on the 30 minute time scale with a 3 s time resolution. Using structures from 19 different time points at resolutions between 2.0 and 2.9 Å, we derived a relaxation constant of ~200 s, which was confirmed by *in crystallo* UV-visible absorption spectroscopy. At t = 0 s, the key residue Gln489 is completely disordered and progressively reorders to re-establish a hydrogen bond with the chromophore FMN, while the nearby multi-conformer Phe470 relaxes to its single resting conformation. Unexpectedly, the 4-fold symmetry is broken after 100 to 200 s which is due to the reordering of the C-terminus occurring in only 50% of the molecules even after 30 minutes. Our work provides a basis for performing routine time-resolved protein crystallography experiments at synchrotrons for slow phenomena occurring on the 10- to 100-second time scale with a 1-second time resolution.

5.1 Introduction

Time-resolved crystallography can be performed in two ways. The first one is the pump-probe approach, in which the pump consists of, usually, an optical laser pulse and the probe is a (short) X-ray pulse leading to a still diffraction image. To a first approximation, the time resolution is the shortest time between the two pulses. In the second approach, a partial, or full oscillation data set is recorded instead of the still image, and in this case the time resolution is usually the duration of the data set. In the late 1990s, a full macromolecular crystallography (MX) data set typically took 2 h at a 3rd-generation synchrotron using a Mar 345 imaging plate detector or a MAR133 CCD detector. In the early 2000s, the data set collection time went down to approximately 20 min thanks to the ADSC series of CCD detectors. A revolution occurred in 2009 with the availability of the first pixel detector for macromolecular crystallography (Pilatus 6M, Dectris), with which data sets could be recorded in 20 s. Currently, the new generation of pixel detector Eiger 4M can operate at a 750 Hz acquisition rate, thus permitting data collections in typically 2 s, making the speed of the goniometer an additional limiting factor for the speed of data acquisition in oscillation data sets.

Experiments that are resolved in time cannot be performed at cryogenic temperatures, at which non-harmonic protein movements are usually frozen out. The availability of the humidity controller HC1 in an MX sample environment has greatly eased the possibility of performing room temperature crystallography (Bowler *et al.*, 2015), in comparison with the traditional, tedious quartz capillary method. As a consequence, room temperature crystallography experiments have become routine at synchrotron MX beamlines, provided sufficient care is taken to limit the absorbed X-ray doses to a few hundreds of kGy (Leal *et al.*, 2013; Gotthard *et al.*, 2019).

In a previous study, we determined the structure of the signalling state (or light state, $LOV2_{390}$) of the LOV2 domain of the blue-light photoreceptor phototropin 2 from *Arabidopsis thaliana* (*At*Phot2LOV2) (Gotthard *et al.*, 2019). The signalling state builds up in microseconds, but decays in tens to hundreds of seconds (Kasahara *et al.*, 2002). The most notable features of the signalling state structure is the formation of a covalent bond between the FMN chromophore and a cysteine residue, as well as displacement, or disordering of the N- and C-termini of the protein (Crosson & Moffat, 2002; Gotthard *et al.*, 2019). A steady-state equilibrium between the dark and the light states was generated in an *At*Phot2LOV2

crystal upon constant illumination from a blue LED source, and a 2.2 s data collection was initiated immediately after turning off the LED power, through a synchronization of the detector and LED TTL signals. The success of this experiment, in which we could observe an occupancy of the light state close to 50%, prompted us to devise an experimental protocol to visualize the progressive relaxation of the light to the dark state over several hundreds of seconds.

5.2 Methods

5.2.1 Protein expression, purification and crystallisation

The gene coding for the phototropin 2 LOV2 domain from *Arabidopsis thaliana* (*At*Phot2LOV2) with a C-terminal hexahistidine tag was synthesized as previously described (Gotthard *et al.*, 2019). The protein was expressed as previously described (Gotthard *et al.*, 2019), except for the growing medium, which was composed of 16 $g.L^{-1}$ tryptone, 10 $g.L^{-1}$ yeast extract and 5 $g.L^{-1}$ NaCl (2YT medium). Cells were centrifuged and frozen at -80 °C. Thawed pellets were sonicated for 4 x 30 s at 35 W power and cell debris were centrifuged at 15000 g, for 45 min at 4 °C. The protein was purified on a nickel affinity column (His-Trap HP 5 mL, GE HealthCare) using an imidazole gradient (50 mM Tris pH 8.0, 300 mM NaCl, 10 to 500 mM imidazole). The final purification step consisted of size exclusion chromatography (Superdex 75 - 10/300 GL, GE HealthCare), after which the protein solution was concentrated to 8 $mg.mL^{-1}$. Protein aliquots were submitted to limited proteolysis using 0.5 $\mu g.\mu L^{-1}$ trypsin. An initial batch of crystals was produced by the hanging drop vapour diffusion method (2 µL drops of 1:1 mother liquor: protein ratio) in a condition consisting of 12 – 17% PEG8000 and 200 mM calcium acetate in 100 mM MES buffer pH 6.0. The resulting square-shaped crystals of heterogeneous size were then used for micro-seeding with a 1:100 dilution of the seed solution with the protein solution, leading to a homogenous production of crystals of 50 x 50 x 50 μm^3.

5.2.2 X-ray data collection

Diffraction data were collected on the ESRF beamline ID30A-3 (Theveneau *et al.*, 2013). Each crystal was placed in the wet stream of an HC1 humidity controller (Arinax, Voreppe, France) on the MD2 Microdiffractometer goniometer head (Arinax, Voreppe,

France). The relative humidity value was initially set at 99.65%, as determined from the composition of the mother liquor (Wheeler *et al.*, 2012), and refined visually to 98-99%. The light from a fibre-coupled 27.5 mW blue LED (λ_{max} = 470 nm, model M470F3, Thorlabs, Newton, New Jersey) was collimated at the entrance of a 5-x magnifying objective positioned at 75 mm from the sample position so as to deliver 3.5 µW on the crystal with a 2.2 mm focal spot. Crystals were illuminated for 60 s prior to data collection, in order to maximize the light state population. The LED was then switched off, which corresponded to t = 0 and the start of a data collection was subsequently delayed for increasing amounts of time (from 0 ± 1 s to 1225 ± 1 s) (**Figure 5.1a**, **Table 5.1**). The duration of each data set was either 2, 2.4 or 3 s.

5.2.3 Data reduction and structure refinement

18 of the 19 diffraction data sets were integrated and reduced using the GrenADES suite chosen from the series of automated data reduction pipelines (Monaco *et al.*, 2013), whose results are available through the web information management system ISPyB (Delagenière *et al.*, 2011). The data set for time point 222 s was integrated and reduced with the XDSAPP pipeline (Sparta *et al.*, 2016), which gave better results. The data set resolution was then fixed at the value for which $CC_{1/2}$ was above 0.7 in the automated scaling log file. For each data set, a molecular replacement step was performed using *PHASER* (Mccoy *et al.*, 2007) with the model corresponding to time point 66 of the TR-SOX experiment on *At*Phot2LOV2 (see **Chapter 6**) with occupancies reset at 0.50/0.50 or 0.33/0.33/0.33 depending on the number of alternate conformations. In our preliminary approach, the only structure refinement step consisted in occupancy refinement of the 7 groups defined in Chapter 6 using *PHENIX* (Adams *et al.*, 2010).

5.2.4 *In crystallo* UV-visible absorption spectroscopy

UV-visible light absorption spectra of *At*Phot2LOV2 crystals were obtained using the *ic*OS Lab located at the ESRF (von Stetten *et al.*, 2015). The reference white light was provided by a DH2000-BAL lamp (Ocean Optics, Largo, Florida) connected to the incoming objective *via* two 200 µm fibres separated by a TTL-controllable shutter. The objective collecting the transmitted light was connected to a grating-based HR2000+ spectrophotometer (Ocean Optics, Largo, Florida) *via* a 400 µm fibre. The actinic light was provided by the same fibre-

coupled LED as in **paragraph 5.2.3**, which was connected to the setup *via* a 1000 µm fibre. A given *At*Phot2LOV2 crystal was initially irradiated for 60 s until the LED was switched off. In order to minimize artefactual white-light-induced regeneration of the light state because of reference light, the shutter only opened for 50 ms every second, for a total experiment duration of 500 s (**Figure 5.1a**).

5.3 Results

We first determined the rate of light state relaxation in crystals using *in crystallo* UV-visible absorption spectroscopy. One crystal of *At*Pht2LOV2 was put in the flow of a humidity controller HC1 on the Minidiffractometer MD2M of the *ic*OS Lab, irradiated for 60 s and its absorbance spectrum was monitored as function of time (**Figure 5.1b, left**). Plotting the evolution of the optical density at 475 nm (wavelength corresponding to one of the absorption maxima of the dark state), which can be fitted as a mono-exponential decay a – b*exp ($-t/\tau_{spectr.}$) with $\tau_{spectr.}$ = 53 s (**Figure 5.1b, right**).

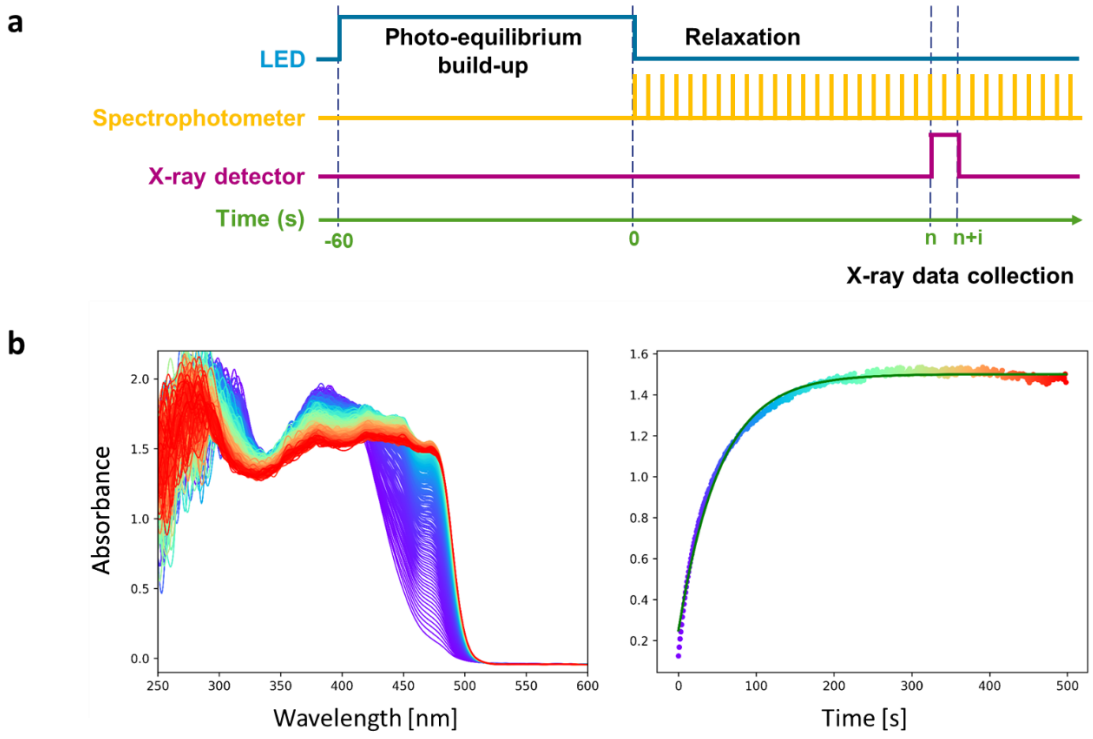


Figure 5.1: (**a**) Schematic of the time-resolved spectroscopic and crystallographic experiment. The duration of the X-ray data collection **i** varies between 2 and 3 s. (**b**) Monitoring of light state relaxation by *in crystallo* UV-visible light absorption spectroscopy. Left: series of spectra recorded for 5000 s. Right: Evolution of the optical density at 475 nm, representative of the dark state replenishment (colour code: rainbow from start (purple) to end (red)).

We collected 19 data sets from 11 crystals, corresponding to 19 time points from 0 to 1225 s, which evenly samples 20 min of light state relaxation on a logarithmic scale (**Tables 5.1 & 5.2**). Surprisingly, a number of data sets could not be reduced in the original tetragonal space group $P4_32_12$, but in the lower symmetry orthorhombic space group $P2_12_12_1$. The first orthorhombic data set appears at 96 s, while the last tetragonal data set occurs at 222 s, meaning that both space groups coexist in-between these two time points depending on the specific experiment. That all final data sets were in the orthorhombic space group posed the question whether a crystal could revert to the initial space group. We also recorded a 20^{th} data set at time point 1802 s (diffraction resolution of 3.3 Å, data not shown), which could only be reduced in $P2_12_12_1$. For three crystals, multiple data sets were recorded on different locations of each crystal (coloured lines in **Table 5.1**). In particular, 5 and 4 cycles of illumination/data collection were performed for crystals 14 (blue lines) and 18 (green lines) and showed alternate occurrences of the tetragonal and orthorhombic space groups, strongly suggesting that the illumination step was sufficient, and probably necessary to restore the initial state of the crystal at $t = 0$.

In order to understand the rationale for the loss of crystal symmetry, we examined the structural data obtained from the orthorhombic data sets. Inspection of the electron density maps indicated that the covalent adduct has mostly disappeared for both monomers in the asymmetric unit. Both N- and C-termini have reordered, but the conformation of the C-terminus is conspicuously different for the two monomers. For monomer 1 (**Figure 5.2a**), the conformation is very close to that in the dark state structure obtained from a tetragonal space group (**Figure 5.2c**). For monomer 2 (**Figure 5.2b**), the orientation of the C-terminus clearly deviates after residue 493, forming a hairpin-like structure distinct from the other conformations (**Figure 5.2d**).

Table 5.1: Data collection parameters. Colours indicate data sets recorded on different locations of the same crystal.

Time point	Crystal #[†]	Run #	X-ray flux (ph/s)	# of images	Oscillation range / image (°)	Total oscillation range (°)	Exposure time per frame (s)	Total exposure time (s)	Space group	Diffraction resolution (Å)
0 s	14	5	9.88e+10	200	0.5	100	0.01	2.0	$P4_32_12$	2.66
3 s	14	4	1.03e+11	200	0.5	100	0.01	2.0	$P4_32_12$	2.76
4 s	5	2	1.07e+11	200	0.5	100	0.01	2.0	$P4_32_12$	2.58
13 s	18	3	1.04e+11	300	0.5	150	0.01	3.0	$P4_32_12$	2.57
15 s	15	1	1.03e+11	240	0.5	120	0.01	2.4	$P4_32_12$	2.63
31 s	21	1	1.05e+11	300	0.5	150	0.01	3.0	$P4_32_12$	2.59
45 s	10	1	1.01e+11	200	0.5	100	0.01	2.0	$P4_32_12$	2.85
60 s	14	1	1.07e+11	200	0.5	100	0.01	2.0	$P4_32_12$	2.54
77 s	17	1	1.05e+11	300	0.5	150	0.01	3.0	$P4_32_12$	2.80
82 s	19	1	9.85e+10	300	0.5	150	0.01	3.0	$P4_32_12$	2.30
96 s	20	1	1.08e+11	300	0.5	150	0.01	3.0	$P2_12_12_1$	2.91
122 s	18	2	1.03e+11	300	0.5	150	0.01	3.0	$P2_12_12_1$	2.48
140 s	13	1	1.05e+11	200	0.5	100	0.01	2.0	$P4_32_12$	2.63
162 s	14	3	1.10e+11	200	0.5	100	0.01	2.0	$P2_12_12_1$	2.39
181 s	19	2	1.05e+11	300	0.5	150	0.01	3.0	$P4_32_12$	2.80
222 s	11	1	1.02e+11	200	0.5	100	0.01	2.0	$P4_32_12$	2.79
270 s	14	2	1.02e+11	200	0.5	100	0.01	2.0	$P2_12_12_1$	2.05
600 s	18	4	1.06e+11	300	0.5	150	0.01	3.0	$P2_12_12_1$	2.65
1225 s	18	5	1.03e+11	300	0.5	150	0.01	3.0	$P2_12_12_1$	2.74

[†]Crystals 5 to 14 were measured on April 25th, 2018 and crystals 15 to 21 were measured on April 27th, 2018.

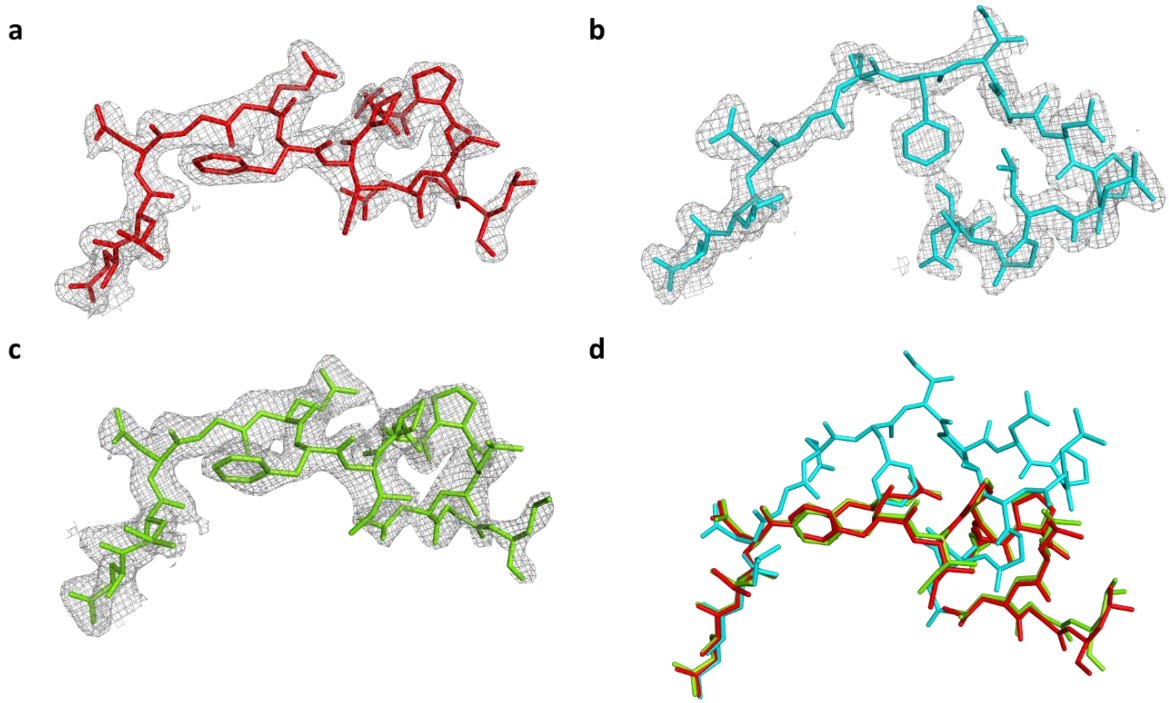


Figure 5.2: Conformation of the C-terminus in the relaxed state and comparison with the dark state. **a.** $2F_{obs}$-F_{calc} electron density map (contoured at a 1.0 σ level) superimposed on residues 490 to 503 of monomer 1 at time point t = 270 s (space group $P2_12_12_1$). **b.** $2F_{obs}$-F_{calc} electron density map (contoured at a 1.0 σ level) superimposed on residues 490 to 503 of monomer 2 at time point t = 270 s (space group $P2_12_12_1$). **c.** $2F_{obs}$-F_{calc} electron density map (contoured at a 1.0 σ level) superimposed on residues 490 to 502 (green) of the dark state structure (time point < 0 s, **Chapter 6**, space group $P4_32_12$). **d.** Superposition of the three structures, showing that the C-ter in monomer 1 is close to the conformation adopted in the dark state and that the C-ter in monomer 2 has a significantly different orientation

Table 5.2: Data reduction and structure refinement statistics.

Values in parentheses are for the outer shell.

Time point	0 s	3 s	4 s	13 s	15 s	31 s	45 s	60 s	77 s	82 s
Data reduction[¥]										
Space group	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$
Cell dimensions *a*, *b*, *c* (Å)	41.58 – 41.58 – 133.56	41.26 – 41.26 – 132.86	41.30 – 41.30 – 133.27	41.61 – 41.61 – 134.00	41.36 – 41.36 – 133.28	41.15 – 41.15 – 133.21	41.38 – 41.38 – 133.74	41.08 – 41.08 – 133.67	41.26 – 41.26 – 133.71	41.06 – 41.06 – 133.08
Resolution range (Å)	39.7 – 2.66 (2.73 – 2.66)	132.9 – 2.69 (2.76 – 2.69)	39.4 – 2.58 (2.65 – 2.58)	134.0 – 2.57 (2.64 – 2.57)	39.5 – 2.56 (2.63 – 2.56)	133.2 – 2.59 (2.66 - 2.59)	39.5 – 2.70 (2.77 – 2.70)	39.3 – 2.41 (2.47 – 2.41)	39.4 – 2.66 (2.73 – 2.66)	39.2 – 2.30 (2.36 – 2.30)
Wilson *B*-factor (Å^2)	63.6	62.9	57.0	60.4	53.7	58.8	60.4	61.6	55.7	53.2
Unique reflections	3753 (269)	3512 (254)	4006 (269)	4144 (281)	4123 (284)	3972 (263)	3556 (265)	4866 (325)	3702 (260)	5547 (377)
Multiplicity	6.7 (7.3)	6.8 (7.0)	6.7 (6.9)	10.2 (9.6)	6.8 (6.7)	10.0 (9.2)	6.7 (6.9)	6.9 (7.1)	10.0 (9.8)	10.3 (10.5)
Completeness (%)	99.6 (99.3)	96.9 (95.5)	99.1 (93.4)	98.3 (94.9)	99.9 (99.6)	98.8 (94.6)	99.2 (100.0)	99.8 (99.4)	99.7 (100.0)	99.8 (97.9)
Mean I/sigma(I)	8.3 (0.9)	8.7 (1.0)	7.6 (1.2)	8.3 (1.1)	7.2 (1.2)	7.2 (1.1)	7.5 (1.1)	9.8 (1.2)	7.6 (1.0)	9.6 (1.0)
R_{meas}[‡]	0.151 (1.374)	0.146 (1.662)	0.174 (1.313)	0.177 (1.824)	0.179 (1.365)	0.211 (2.035)	0.166 (1.399)	0.113 (1.402)	0.221 (1.843)	0.150 (1.697)
$CC_{1/2}$	0.998 (0.675)	0.998 (0.617)	0.996 (0.762)	0.994 (0.753)	0.995 (0.646)	0.995 (0.530)	0.996 (0.500)	0.998 (0.685)	0.988 (0.626)	0.999 (0.658)
Structure refinement										
Resolution[†] (Å)	33.4 – 2.66 (2.76 – 2.66)	35.0 – 2.76 (2.86 – 2.76)	39.4 – 2.58 (2.67 – 2.58)	35.3 – 2.57 (2.66 – 2.57)	39.5 – 2.63 (2.73 – 2.63)	24.3 – 2.59 (2.68 – 2.59)	39.5 – 2.85 (2.95 – 2.85)	35.0 – 2.54 (2.63 – 2.54)	39.4 – 2.80 (2.90 – 2.80)	34.9 – 2.30 (2.38 – 2.30)
Unique reflections	3748 (362)	3248 (300)	4001 (361)	4126 (378)	3840 (370)	3937 (354)	3038 (301)	4199 (403)	3193 (301)	5536 (517)
R_{work}	0.292 (0.378)	0.282 (0.357)	0.288 (0.398)	0.288 (0.371)	0.279 (0.376)	0.255 (0.367)	0.260 (0.356)	0.248 (0.341)	0.249 (0.318)	0.258 (0.371)
R_{free}	0.261 (0.518)	0.306 (0.552)	0.320 (0.389)	0.309 (0.378)	0.298 (0.370)	0.276 (0.451)	0.242 (0.498)	0.263 (0.303)	0.250 (0.501)	0.256 (0.318)
N. of atoms	1536	1536	1536	1536	1536	1536	1536	1536	1536	1536
Average atomic *B*-factor (Å^2)	51.0	49.8	48.8	46.9	40.7	44.5	44.6	47.1	39.0	43.5
R.m.s deviations										
Bond lengths (Å)	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003
Bond angles (°)	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66

[¥] Statistics obtained from automatic data processing using the GrenADES pipeline (Monaco *et al.*, 2013), except for time point 222 s, for which statistics were obtained from automatic data processing using the XDSAPP pipeline (Sparta *et al.*, 2016). [†]The resolution cutoff is based on $CC_{1/2}$.= 0.7 in the outer shell. [‡]$R_{meas} = R_{merge} \times [N/(N-1)]^{1/2}$, where N is the data multiplicity.

Time point	**96 s**	**122 s**	**140 s**	**162 s**	**181 s**	**222 s**	**270 s**	**600 s**	**1225 s**
Data reduction¥									
Space group	$P2_12_12_1$	$P2_12_12_1$	$P4_32_12$	$P2_12_12_1$	$P4_32_12$	$P4_32_12$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions *a*, *b*, *c* (Å)	43.02 – 133.21 – 40.20	40 21 – 43.08 – 132.87	40.96 – 40.96 – 133.66	40.30 – 43.20 – 132.81	40.95 – 40.95 – 132.54	40.95 – 40.95 – 133.26	40.27 – 43.24 – 132.89	40.33 – 43.23 – 133.04	40.31 – 43.20 – 133.00
Resolution range† (Å)	66.6 – 2.68 (2.75 – 2.68)	66.4 – 2.48 (2.55 – 2.48)	39.2 – 2.63 (2.70 – 2.63)	132.8 – 2.27 (2.33 – 2.27)	39.1 – 2.80 (2.87 – 2.80)	39.1 – 2.50 (2.56 – 2.50)	132.9 – 1.94 (1.99 – 1.94)	133.0 – 2.44 (2.50 – 2.44)	133.0 – 2.60 (2.67 – 2.60)
Wilson *B*-factor (Å^2)	47.7	47.4	60.3	40.7	68.9	53.4	35.9	45.7	47.6
Unique reflections	6866 (423)	8539 (587)	3773 (273)	10793 (784)	3121 (211)	4361 (319)	17096 (1214)	8875 (624)	7318 (504)
Multiplicity	5.3 (5.0)	5.3 (5.2)	6.7 (7.2)	3.8 (4.0)	10.1 (7.3)	6.8 (7.1)	3.9 (3.9)	5.5 (5.4)	5.5 (5.1)
Completeness (%)	99.1 (92.2)	98.2 (92.3)	99.8 (100.0)	95.2 (98.9)	99.2 (91.7)	99.7 (99.1)	95.3 (94.4)	96.1 (94.5)	95.8 (91.0)
Mean I/sigma(I)	5.0 (1.0)	6.7 (1.0)	6.5 (1.1)	6.3 (1.3)	6.9 (0.8)	4.0 (0.25)	7.8 (1.0)	6.6 (1.2)	5.6 (1.1)
R_{meas}‡	0.250 (1.615)	0.196 (1.205)	0.186 (1.442)	0.157 (0.978)	0.231 (2.095)	0.315 (2.810)	0.113 (0.962)	0.192 (1.292)	0.230 (1.426)
$CC_{1/2}$	0.988 (0.372)	0.994 (0.698)	0.993 (0.664)	0.994 (0.645)	0.997 (0.592)	0.993 (0.189)	0.997 (0.524)	0.995 (0.500)	0.984 (0.492)
Structure refinement									
Resolution† (Å)	40.9 – 2.91 (3.01 – 2.91)	41.0 – 2.48 (2.57 – 2.48)	39.2 – 2.63 (2.72 – 2.63)	29.8 – 2.39 (2.48 – 2.39)	39.1 – 2.80 (2.90 – 2.80)	39.1 – 2.79 (2.89 – 2.79)	36.2 – 2.05 (2.12 – 2.05)	41.1 – 2.65 (2.75 – 2.65)	41.1 – 2.74 (2.84 – 2.74)
Unique reflections	5453 (512)	8535 (803)	3770 (363)	9213 (928)	3099 (276)	3163 (298)	14519 (1450)	6953	6298 (623)
R_{work}	0.302 (0.392)	0.329 (0.432)	0.261 (0.358)	0.332 (0.407)	0.278 (0.422)	0.267 (0.332)	0.347 (0.411)	0.319 (0.385)	0.321 (0.378)
R_{free}	0.328 (0.313)	0.327 (0.452)	0.264 (0.351)	0.330 (0.408)	0.307 (0.435)	0.260 (0.318)	0.325 (0.411)	0.304 (0.351)	0.342 (0.336)
N. of atoms	3072	3072	1536	3072	1536	1536	3072	3072	3072
Average atomic *B*-factor (Å^2)	34.0	46.6	41.1	31.7	52.0	46.1	29.8	33.6	34.5
R.m.s deviations									
Bond lengths (Å)	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003
Bond angles (°)	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66

In our preliminary approach, and in order to assess the quality of our data, we decided not to refine structures independently for the 19 different data sets, and adapted the refinement strategy adopted in **Chapter 6**. We imposed a model with alternate conformations in fixed positions, corresponding to the dark and light states, and to refine the alternate occupancies of 7 groups, comprised of 6 stretches of residues and of the FMN chromophore. Each of these 6 stretches of residues flank one given residue which has been shown to move unambiguously in the build-up of the light state: Val392, Ile403, Cys426, Leu456, Phe470 and Gln489 (**Chapter 6**). For clarity, the resulting occupancies were plotted as a function of time on both a logarithmic and on a linear scales (**Figures 5.3 & 5.4**).

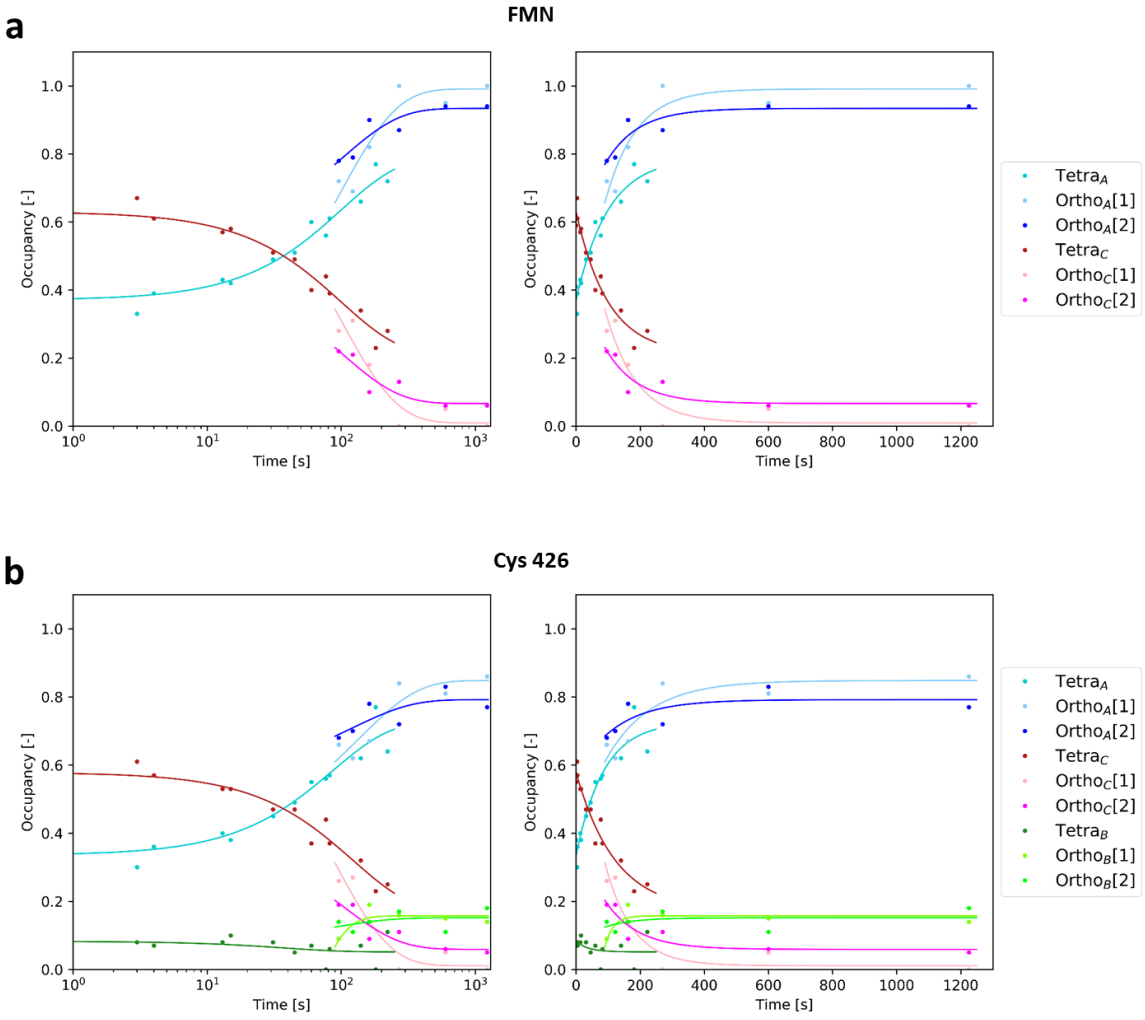


Figure 5.3: Time evolution of light- and dark state populations for (**a**) FMN and (**b**) Cys426. Left panel: logarithmic time scale, right panel: linear time scale. Tetra$_{A/B/C}$ indicate the occupancy of the various conformer populations in the tetragonal space group $P4_32_12$. Ortho$_{A/B/C}$[1] and Ortho$_{A/B/C}$[2] indicate the various conformer populations of the two molecules of the asymmetric unit in the orthorhombic space group $P2_12_12_1$. Each set of data point have been fitted by mono-exponential decay which has been represented as a continuous line.

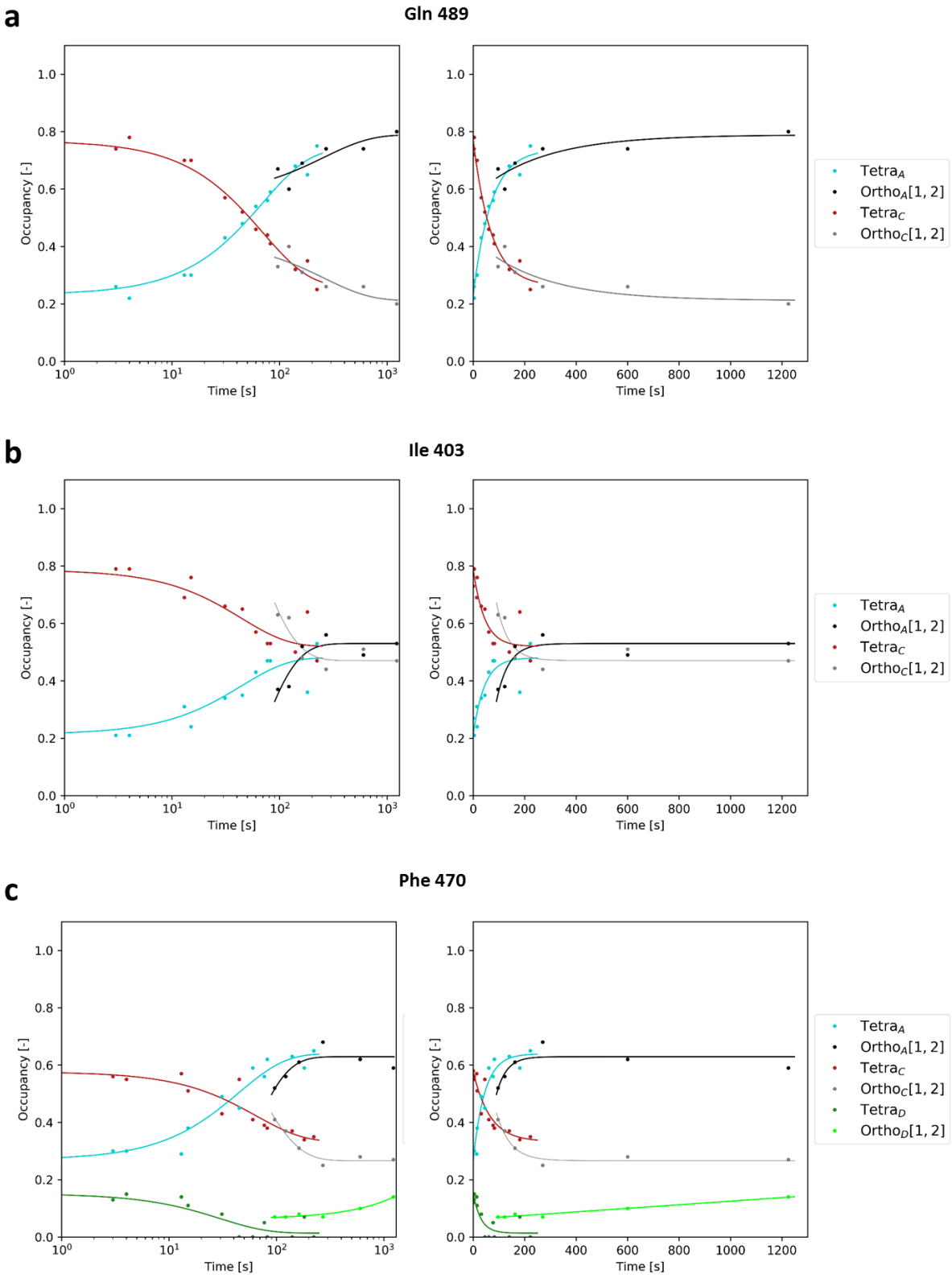


Figure 5.4: Time evolution of light- and dark state populations for (**a**) Gln489, (**b**) Phe470 and (**c**) Ile403. Left panel: logarithmic time scale, right panel: linear time scale. Tetra$_{A/C/D}$ indicate the occupancy of the various conformer populations in the tetragonal space group $P4_32_12$. Ortho$_{A/C/D}$[1,2] indicate the various conformer populations averaged over the two molecules of the asymmetric unit in

the orthorhombic space group $P2_12_12_1$. Each set of data point have been fitted by mono-exponential decay which has been represented as a continuous line.

The best results were obtained for the occupancy refinement of FMN and Cys426, for which the starting occupancy values are ~40% for the dark state (conformations A and B) and ~60% for the light state (conformations C and D), which progressively evolve towards ~100% and ~0%, respectively. The experimental data points were treated as three separate groups: those from the tetragonal space group ($Tetra_A$ and $Tetra_C$), those from one monomer of the orthorhombic space group ($Ortho_A[1]$ and $Ortho_C[1]$) and those from the second monomer of the orthorhombic space group ($Ortho_A[2]$ and $Ortho_C[2]$). The fitted mono-exponential curves of the three different set of data points overlap well on the [96 s – 122 s] time range, and the time constants (**Table 5.3**) are in good agreement (99 s for FMN, 116, 103 and 50 s for the three conformations of Cys426) if one takes into account the low occupancy of the B conformer of Cys426.

Table 5.3: Time constants (in seconds) of occupancy evolutions for the chromophore various key residues of *At*Phot2LOV2 modelled as mono-exponential decays.

Conformation	**FMN**	**Cys426**	**Gln489**	**Ile403**	**Phe470**
$Tetra_A$	100	86	70	44	45
$Ortho_A[1]$	96	138	191	63	60
$Ortho_A[2]$	100	123	992	25	35
$Ortho_A[1,2]$	-	-	257	47	40
Average	*99*	*116*	*164*	*46*	*43*
$Tetra_C$	100	122	70	44	65
$Ortho_C[1]$	96	85	191	78	125
$Ortho_C[2]$	100	101	992	25	25
$Ortho_C[1,2]$	-	-	257	47	55
Average	*99*	*103*	*164*	*46*	*60*
$Tetra_B$		38			
$Ortho_B[1]$		26			
$Ortho_B[2]$		85			
Average		*50*			
$Tetra_D$					30
$Ortho_D[1]$					179
$Ortho_D[2]$					80
$Ortho_D[1,2]$					226589

The same approach does not work as well for Gln489, Ile403 and Phe470 (**Figure 5.4**), where an averaging in-between monomers 1 and 2 is required to get sensible time constants (70 s and 164 s for Gln489, 44 s and 46 s for Ile403, 45 s, 43 s, 65 s and 60 s for Phe470). However, only the occupancies of Gln489 refine to sensible values (that is close to 100%/0%), while those of Ile403 and Phe470 refine to ~50%/~50% and ~60%/~25%/~15%. Finally, refined occupancies for Val392 and Leu456 did not make sense, and are not shown.

5.4. Discussion

In this work, we have attempted to monitor the relaxation of the signalling state of *At*Phot2LOV2 using a time-resolved crystallography approach with a 3 s time resolution, well suitable for a phenomenon occurring with a time constant between 53 s (derived from spectroscopic data) and 108 s (derived from the occupancy refinement of the FMN and Cys426 from diffraction data). Clearly, our approach of refining the occupancies of fixed models in the various time points does not give a consistent and comprehensive description of the phenomenon, and a more detailed refinement protocol needs to be applied to each time point, which will constitute a time-consuming task. It is likely that obtaining data sets of higher resolution (for instance by being less conservative on the absorbed X-ray dose on each crystal) would facilitate structural refinement of complete alternate structures.

Nevertheless, we were able to visualize the surprising phenomenon that the relaxation of the light state in *At*Phot2LOV2 crystals induced a phase transition of crystals that can be reversed with light. To some extent, this is reminiscent of the photo-induced martensitic transformation observed in crystals of the reversibly switchable fluorescent protein Tetdron (Dajnowicz *et al.*, 2019). In this study, light irradiation does not lead to a change in crystallographic symmetry, only to a lengthening of two cell parameters which is provoked by a reorganization of the initial asymmetric octamer (with small interaction areas) to two tetramers (with extensive interaction areas within each tetramer). This phase transition appears to be stable for months. However, the retraction of this paper shortly thereafter questions this interpretation.

Our interpretation of the whole phenomenon is that irradiation of a crystal of *At*Phot2LOV2 in the dark state in the tetragonal space group first induces the formation of the covalent bond between FMN and Cys426 leading to a disordering of the N-ter and C-ter parts

of each molecule of the crystal (**Figure 5.5**). Once the illumination is stopped, each molecule slowly relaxes to the dark state, and in order for one molecule to see its C-ter reorder in its original position, a neighbouring molecule has to allow its own C-ter to order in a more compact, hairpin-like conformation, thus breaking the tetragonal symmetry. Obviously, there is a significant energy barrier for the crystal to cross, so that the C-ter of the second molecule relaxes as well, which can, apparently, be only achieved by re-illumination.

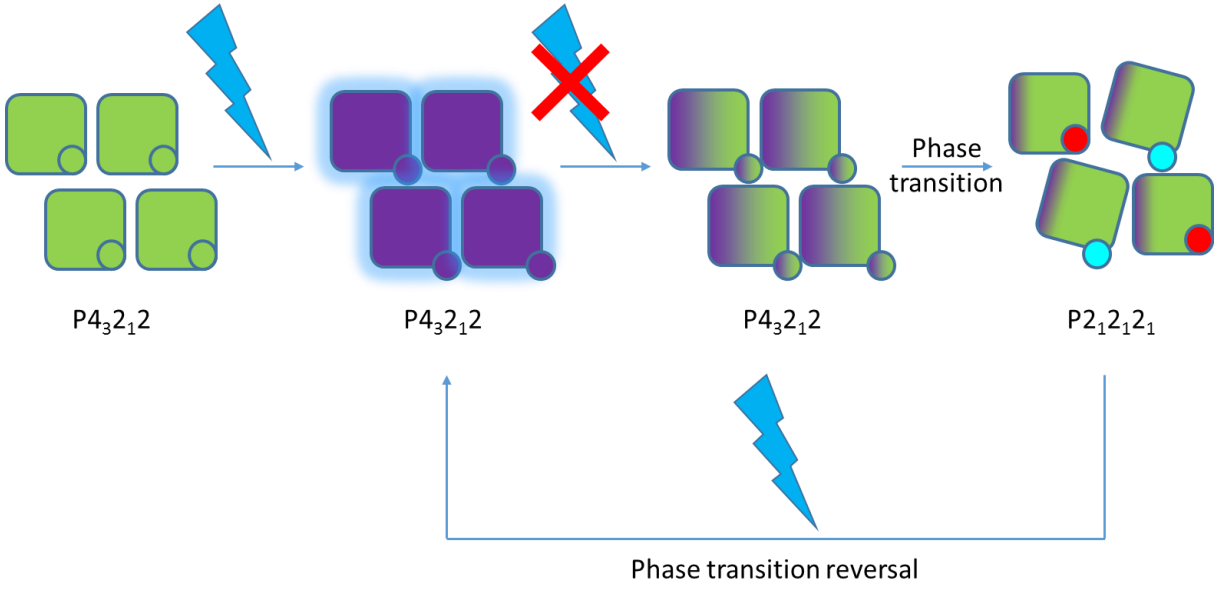


Figure 5.5: Schematic of space group conversion.

5.5 References

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010). *Acta Cryst. D*. **66**, 213–221.

Bowler, M. W., Mueller, U., Weiss, M. S., Sanchez-Weatherby, J., Sorensen, T. L.-M., Thunnissen, M. M. G. M., Ursby, T., Gobbo, A., Russi, S., Bowler, M. G., Brockhauser, S., Svensson, O. & Cipriani, F. (2015). *Cryst. Growth Des.* **15**, 1043–1054.

Crosson, S. & Moffat, K. (2002). *Plant Cell*. **14**, 1067–1075.

Dajnowicz, S., Langan, P. S., Weiss, K. L., Ivanov, I. N. & Kovalevsky, A. (2019). *IUCrJ*. **6**, 1–11.

Delagenière, S., Brenchereau, P., Launer, L., Ashton, A. W., Leal, R., Veyrier, S., Gabadinho, J., Gordon, E. J., Jones, S. D., Levik, K. E., Mcsweeney, S. M., Monaco, S., Nanao, M., Spruce, D., Svensson, O., Walsh, M. A. & Leonard, G. A. (2011). *Bioinformatics*. **27**, 3186–3192.

Gotthard, G., Aumonier, S., de Sanctis, D., Leonard, G., von Stetten, D. & Royant, A. (2019). *IUCrJ*. in press.

Kasahara, M., Swartz, T. E., Olney, M. A., Onodera, A., Mochizuki, N., Fukuzawa, H., Asamizu, E., Tabata, S., Kanegae, H., Takano, M., Christie, J. M., Nagatani, A. & Briggs, W. R. (2002). *Plant Physiol.* **129**, 762–773.

Leal, R. M. F., Bourenkov, G., Russi, S. & Popov, A. N. (2013). *J. Synchrotron Rad.* **20**, 14–22.

Mccoy, A. J., Grosse-kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.

Monaco, S., Gordon, E., Bowler, M. W., Delagenière, S., Guijarro, M., Spruce, D., Svensson, O., McSweeney, S. M., McCarthy, A. A., Leonard, G. & Nanao, M. H. (2013). *J. Appl. Cryst.* **46**, 804–810.

Sparta, K. M., Krug, M., Heinemann, U., Mueller, U. & Weiss, M. S. (2016). *J. Appl. Cryst.* **49**, 1085–1092.

von Stetten, D., Giraud, T., Carpentier, P., Sever, F., Terrien, M., Dobias, F., Juers, D. H., Flot, D., Mueller-Dieckmann, C., Leonard, G. a., de Sanctis, D. & Royant, A. (2015). *Acta Cryst. D*. **71**, 15–26.

Theveneau, P., Baker, R., Barrett, R., Beteva, A., Bowler, M. W., Carpentier, P., Caserotto, H., De Sanctis, D., Dobias, F., Flot, D., Guijarro, M., Giraud, T., Lentini, M., Leonard, G. A., Mattenet, M., McCarthy, A. A., McSweeney, S. M., Morawe, C., Nanao, M., Nurizzo, D., Ohlsson, S., Pernot, P., Popov, A. N., Round, A., Royant, A., Schmid, W., Snigirev, A., Surr, J. & Mueller-Dieckmann, C. (2013). *J. Phys. Conf. Ser.* **425**, 012001.

Wheeler, M. J., Russi, S., Bowler, M. G. & Bowler, M. W. (2012). *Acta Cryst. F* . **68**, 111–114.

Chapter 6

RESULTS – ARTICLE 3

Room temperature time-resolved crystallography on a plant photoreceptor domain with a 63 ms time resolution

–

Development of the Time-Resolved Serial Oscillation Crystallography (TR-SOX) method

Presentation of Article 3

In **Chapter 5**, we investigated by time-resolved crystallography the relaxation of the light state of *At*Phot2LOV2 on the [0 – 30 min] time range with a 3 s time resolution. While characterizing the experimental conditions of this experiment, we observed the transient loss of fluorescence of our crystals when initiating the build-up of the light state by turning on of the actinic LED. Since the phenomenon occurred within a couple of seconds, and given that the light state builds up on the microsecond time scale, we reasoned that the observed phenomenon is that of the progressive photo-transformation of the crystal under moderate light level. In brief, the LED does not provide enough photons to transform all molecules at once. However, this constitutes a phenomenon much faster than light state relaxation, and we were curious to see how much we could increase the time resolution of our TRX experiments and aimed to visualize with a 10- to 100-ms time resolution.

Millisecond time-resolved serial oscillation crystallography at synchrotrons

Sylvain Aumonier[1,2], Gianluca Santoni[1], Guillaume Gotthard[1], David von Stetten[1$], Gordon Leonard[1], Antoine Royant[1,2*]

[1]European Synchrotron Radiation Facility, F-38043 Grenoble, France.

[2]Univ. Grenoble Alpes, CNRS, CEA, IBS (Institut de Biologie Structurale), F-38000 Grenoble, France.

*Correspondence email: antoine.royant@ibs.fr

[$]Present address: European Molecular Biology Laboratory (EMBL) c/o DESY, Notkestrasse 85, D-22607 Hamburg, Germany

Abstract We report a method, using monochromatic synchrotron radiation, for the room temperature collection, processing and merging of X-ray oscillation diffraction data from a finite series of samples in order to observe intermediate species build-up in protein crystals. Using this method, we monitored the build-up of the light state of a blue-light photoreceptor domain with a time resolution of 63 ms, visualizing the progressive rearrangement of protein and chromophore during this process.

6.1 Introduction

Synchrotron-based X-ray crystallography has been largely responsible for the explosion in macromolecular crystallography (MX) structure deposition, which has occurred since the early 2000s (http://biosync.sbkb.org) with the overwhelming fraction of the structural information obtained exploiting crystals in which the conformation and biological state of the system under study is static, whether held at cryogenic or room temperature. However, gaining insight into a molecular mechanism of a protein greatly benefits from watching it in action (*i.e.* while it is performing its function (Hajdu *et al.*, 2000)) and for this time-resolved crystallography (TRX) experiments are necessary. Room temperature Laue crystallography pioneered TRX at synchrotrons, providing a time resolution down to the 100-picosecond level (Schotte *et al.*, 2003). However, the number of systems to which the method could be applied

is limited, and although the technique produced some stunning successes (Schotte *et al.*, 2003; Wöhri *et al.*, 2010; Schotte *et al.*, 2012), TRX gradually fell out of favour. Recently, the development of serial crystallography (SX, (Chapman *et al.*, 2011)) and advances in sample delivery (Grünbein & Kovacs, 2019) and X-ray detectors (Broennimann *et al.*, 2006; Casanas *et al.*, 2016) have resulted in a renewed interest in TRX both at free electron laser (TR-SFX, (Johansson *et al.*, 2017) and synchrotron sources (TR-SSX, (Schulz *et al.*, 2018)). However, both TR-SFX and TR-SSX techniques are, in the main, currently based on the recording of still diffraction images from many thousands of crystals for each time point and their requirement in sample quantity is thus very high. We therefore developed a method which we called time-resolved serial oscillation crystallography (TR-SOX), to perform TR-SSX using diffraction data collected from only tens to hundreds of crystals and which has the potential to produce millisecond time resolution movies of biomolecular processes occurring in crystals.

TR-SOX was used to probe the build-up of the light state in the photocycle of the LOV2 domain of *PHOT2* (phototropin 2) from *Arabidopsis thaliana* (*At*Phot2LOV2) with a time resolution of 63 milliseconds. *PHOT2* is a high-intensity blue light photoreceptor mediating phototropism in higher plants (Christie, 2007). The protein comprises two successive Light-Oxygen-Voltage (LOV) domains (LOV1 and LOV2), a bridging helix J_α and a kinase domain (**Fig. 1a**). Upon blue light irradiation, J_α unfolds and triggers autophosphorylation of the Serine/Threonine kinase domain, which activates the signalling casacade (Harper *et al.*, 2003). The photoactivation of LOV domains has been intensively studied by various biophysical methods including fast spectroscopic techniques (Swartz *et al.*, 2001; Kottke *et al.*, 2003; Kennis *et al.*, 2003). The dark state of *At*Phot2LOV2 absorbs light maximally at λ = 447 nm *via* its flavin mononucleotide (FMN) chromophore. Upon photon absorption, a triplet state absorbing at λ = 660 nm forms in nanoseconds and converts on the microsecond time-scale, to the light state, a flavin-cysteine adduct absorbing at λ = 390 nm (**Fig. 1b**). Finally, the system relaxes back to the dark state on the 100-second level.

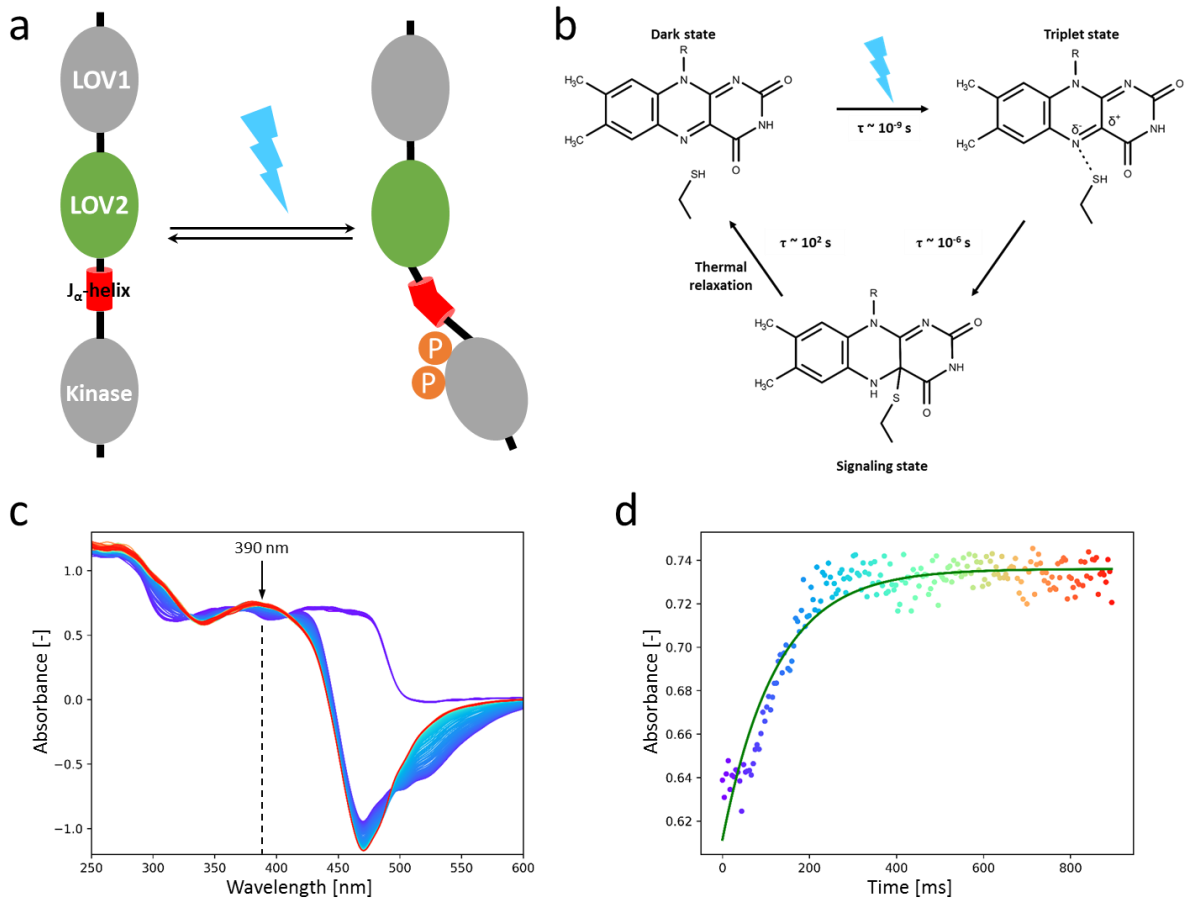


Figure 1 | Photoactivation of the LOV2 domain of phototropin 2 from *Arabidopsis thaliana*. (**a**) Schematic of phototropin 2 domain reorganization upon blue light illumination. (**b**) The photocycle of *At*Phot2LOV2. (**c**) Time-resolved series of UV-visible absorption spectra of a *At*Phot2LOV2 crystal under 3.2 µW irradiation of a 470 nm LED, recorded at 23 Hz (colour range: violet (t < 0 s) to red (t = 9 s)). The absorption maximum of the light state is indicated with an arrow. (**d**) Evolution of the optical density at λ = 390 nm characteristic of the *At*Phot2LOV2 light state build-up, modelled as a mono-exponential rise (rise time τ = 1.23 s).

6.2 Development of the TR-SOX method

In order to observe a phenomenon on the millisecond to second time scale, we slowed down the rate of light state build-up within a crystal of *At*Phot2LOV2 by using an attenuated 470 nm light-emitting diode (LED) as actinic light. The limited number of photons photo-activate all protein molecules within the crystal in a few seconds (**Supplementary Movie 1**) or tens of seconds. The probed phenomenon is thus not the build-up of the light state (which occurs in microseconds at the single molecule level), but that of its population within the crystal. We used *in crystallo* UV-visible absorption spectroscopy at the IBS/ESRF *ic*OS Laboratory to quantify the phenomenon as a function of actinic light power. Absorption spectra

were continuously recorded at a 23 Hz rate providing a 43-millisecond time resolution. Build-up rates were observed to be linearly correlated with LED power over a 50-fold power range (**Supplementary Fig. 2**), proving that light state build-up in the crystal is solely accounted for by the number of available photons. We thus determined that a light power of 3.2 µW at the crystal surface would result in a build-up time constant just above 1 s (1,235 ms) (**Fig. 1c & 1d**).

For TR-SOX experiments on the ESRF beamline ID30A-3, 88 crystals of *At*Phot2LOV2 were sequentially mounted in a loop under a wet air stream at ambient temperature, under low light conditions and synchronized LED illumination with X-ray shutter opening, goniometer axis rotation and X-ray detector acquisition. The X-ray beam flux was attenuated so that the total absorbed dose for each data set was between 115-162 kGy (see Methods), thus minimizing resolution loss and thio-ether bond reduction (Gotthard *et al.*, in press). For each crystal 990 images of 0.5° oscillation range were recorded at an acquisition rate of 250 Hz. Each of the 88 resulting data sets was then separated in 15-frame sub-wedges assigned to 66 different time periods T_1=[0–63 ms] to T_{66}=[4,099–4,158 ms], which were ascribed to 66 time points from t_1=63 ms to t_{66}=4,158 ms (**Fig. 2a**).

For each time point of the diffraction experiment, 36 to 52 sub-wedges were successfully integrated and optimally merged together using *ccCluster* (**Figs. 2b & 2c**). The clustering threshold $CC_{threshold,}$ was chosen such that the diffraction resolution of each final data set was maximised (see **6.5 Methods**). 66 complete data sets corresponding to time points t_1 to t_{66} were thus obtained at a diffraction resolution starting at 2.45 Å for t_1 and rapidly settling at 2.75 Å after t_{11} = 693 ms (**Supplementary Table 1**, **Supplementary Fig. 4a**). Two other data integration and reduction parameters also evolve with time in a similar manner. First, the value $CC_{threshold}$ used to determine the maximum resolution diffraction increases steadily from 0.39 at t_1 and stabilizes around 0.68 after t_{24} = 1,512 ms (**Supplementary Fig. 4b)**; the cell parameter ***a*** follows a comparable pattern by increasing from 40.73 Å at t_1 to an average of 40.97 Å after t_{34} = 2,142 ms, while the cell parameter ***c*** appears to remain constant within the noise (**Supplementary Fig. 4c & d**). That these three parameters settle to a plateau may indicate that their evolution is not primarily driven by X-ray radiation damage and is possibly an effect of photo-activation of the protein leading to subtle changes in the unit cell parameters and crystal contacts and thus to a decrease in the degree of isomorphism.

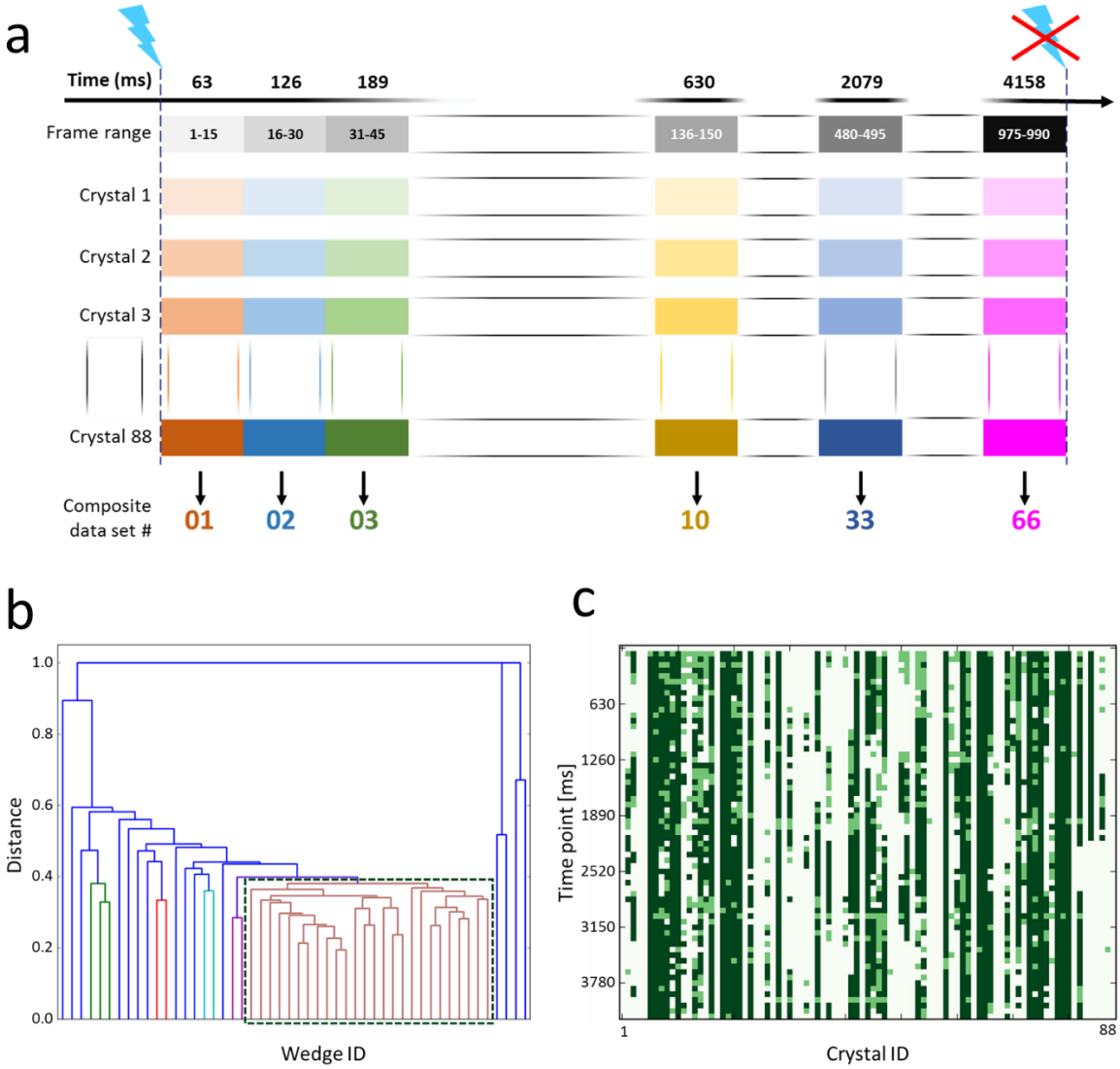


Figure 2. | Data collection strategy and data set composition. (**a**) Complete data sets were collected from 88 different crystals. For each crystal the diffraction images were separated into 15 frame wedges corresponding to 66 time points (from 63 to 4158 ms). (**b**) Dendrogram of wedge clustering for the initial time point (t = 0 – 63 ms). The 26 wedges forming the cluster highlighted in dark green group at a linkage distance of 0.39 and were subsequently merged together to produce a 2.45 Å data set with 100% completeness and 12.6 redundancy. (**c**) Visualisation of the data set wedges integrated (light and dark green), and of the wedges retained in the composite data sets corresponding to the 66 different time points (dark green only).

6.3 Structural analysis and discussion

To determine the structure of the dark state of *At*Phot2LOV2, representative of the time series before t = 0 ms, we carried out the same diffraction experiment, but without illumination, on a series of 32 crystals. The 500 first images of 15 data sets could be successfully integrated

and were subjected to clustering with *ccCluster* which retained 5 sub-wedges using a $CC_{threshold}$ of 0.43 and resulted in a complete, highly multiplicitous data set reduced and merged in space group $P4_32_12$ at 2.2 Å resolution (**Supplementary Table 1**). Refinement of the crystal state structure of the dark state was then carried out using a previously-determined structure of *At*Phot2LOV2 (PDB entry 6QQJ, (Gotthard *et al.*, in press)) as a starting model. Most residues were modelled in a single conformation A, while 10 of them were modelled with alternate conformations A and B, including Cys426 which forms a covalent bond with FMN in the light state. The hydrophilic side of the FMN isoalloxazine ring is stabilized by two hydrogen bonds to Asn458 and one key hydrogen bond to Gln489 (**Fig. 3a**).

Modelling of the crystal structure of the light state of *At*Phot2LOV2 (denoted C and D in **Fig. 3b**) was guided by inspection of $F_{obs} - F_{calc}$ and $2F_{obs} - F_{calc}$ electron density maps for the latest time points (t_{60} to t_{66}) in our diffraction data series. Here, in addition to the formation of the covalent bond between Cys426 and the FMN and the rotation of the FMN moiety, several residues surrounding the chromophore show clear signs of side chain translation (Gln489, Ile403, Val392), rotation and reorientation (Phe470), but also displacement of their main chain (**Figs. 3d** & **3e**, **Supplementary Fig. 5**). Phe470 is shown to adopt two new conformations C and D, the latter of which is made possible by the displacement and reorientation of Leu456 (**Supplementary Fig. 6**). We modelled continuous sequence stretches around those residues and refined a conformation C of the protein for each of the resulting segments **1** to **6,** all composed of rigid secondary structure elements (**Supplementary Figs 1 & 7**).

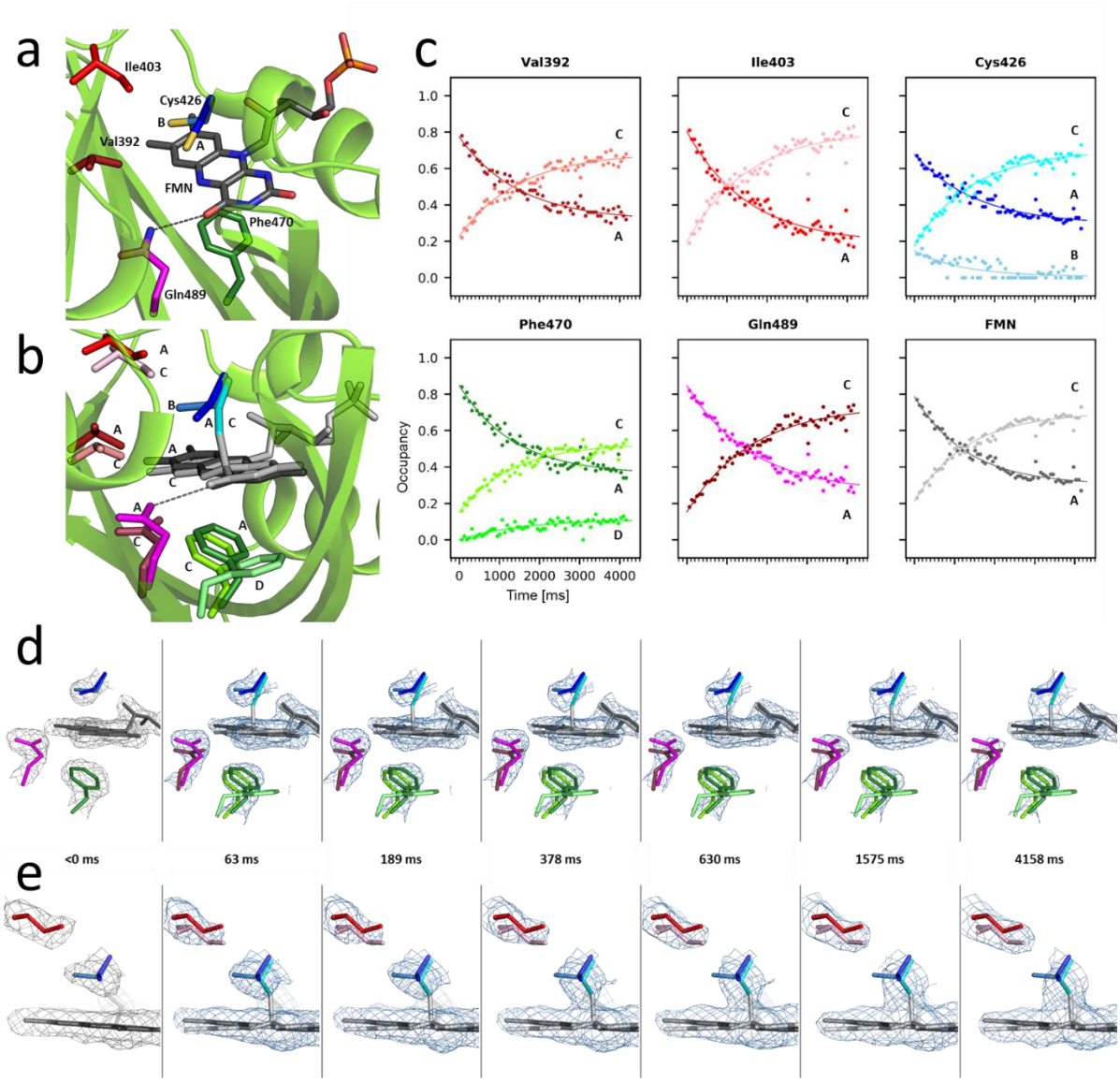


Figure 3 | (**a**) Structure of *At*Phot2LOV2 in the dark state showing key residues surrounding the FMN chromophore. (**b**) Structure of *At*Phot2LOV2 in the light state (C and D conformers, light shade colours) superimposed on that of the dark state (A and B conformers, darker shade colours). (**c**) Time evolution of the respective occupancies of each key residue conformations modelled as monoexponential rises or decays. (**d**) Snapshots at time points 0, 1, 3, 6, 10, 25 and 66 showing $2F_{obs} - F_{calc}$ electron density maps contoured at a 1.2 σ level superimposed on the structural models of the dark and light states highlighting covalent bond formation between Cys426 and the FMN, disordering of the Gln489 side chain and displacement of the Phe470 side chain. (**e**) A different view showing $2F_{obs} - F_{calc}$ electron density maps contoured at a 0.7 σ level illustrating rotation of the isoalloxazine ring of the FMN, translation of Ile403 and progressive disappearance of the B conformation of Cys426.

In order to follow the time course of the dark to the light state conversion, we evaluated their respective contribution to each time point. Occupancies of the 6 segments were refined independently, while atomic coordinates and B-factors were fixed. The evolution of the respective occupancies of each of the prominent residues and cofactor are represented on **Fig. 3c**. All decays (conformations A and B) and rises (conformations C and D) could be modelled by a mono-exponential behaviour (**Fig. 3c)**, whose time constants are very similar, with an average value of 1445 ms ± 135 ms (**Supplementary Table 2**). This is in very good agreement with the spectroscopy-derived rise time of 1235 ms for light state build-up (**Fig. 1d**), demonstrating that we have been able to visualize at the molecular level the progressive conversion from the dark to the light state of *At*Phot2LOV2 molecules within the crystal on a ~50 millisecond time scale.

The evolution of $2F_{obs} - F_{calc}$ electron density maps and structural models obtained (**Figs. 3d & 3e**, **Supplementary Movies 2 & 3**) allow an atomic description of the dark- to light state conversion of *At*Phot2LOV2. Absorption of a photon by FMN leads to the recruitment of both the A and B conformers of Cys426 in segment **3** to establish a covalent bond between its S_γ atom and the C_{4a} atom of FMN. This results in a ~ 6° rotation of the chromophore isoalloxazine ring around the axis of its ribityl tail and in the disruption of a stabilizing hydrogen to Gln489. The disappearance of conformer B of Cys426 added to the rotation of the isoalloxazine ring creates a void that is readily occupied by both Val392 and Ile403 whose side chains translate by 0.6–1.1 Å towards the FMN, displacing both protein segments **1** and **2**. The rotation of the chromophore pushes Phe470 into two new conformers, thus displacing segment **5**. The minor conformer D of Phe470 forces a rearrangement of Leu456, thus displacing segment **4**. Finally, the change in electronic configuration of the FMN upon covalent bond formation induces the loss of the stabilizing hydrogen bond between the FMN and Gln489, whose main and side chains move away, displacing segment **6**. Segments **1** and **6**, constituting the N- and C-termini of the protein, respectively, have been already shown to be involved in signal transduction through the modification of their interactions with the J_α helix (Halavaty & Moffat, 2007). In particular, the structure of the LOV2-J_α domain of *Avena sativa* (PDB entry 2V1A (Halavaty & Moffat, 2007)) reveals that the equivalent of segments **1**, **5** and **6** defined here form a groove accommodating the J_α helix, and support the notion that their concerted displacement would destabilize this helix, which in turn would activate the kinase domain.

6.4 Conclusion

Here we have used room temperature TR-SOX coupled with complementary optical spectroscopy to follow, and provide structural insight into, the activation mechanism of a light-activated phototropin. One noteworthy advantage of the method based on oscillation crystallography is the significant reduction of required sample material compared to still images-based crystallography. Here, the time resolution obtained was limited by the maximum goniometer rotation speed (180°/second) and detector readout rate (250 Hz). Increasing both these parameters will likely lead to recording molecular movies on the sub-millisecond time scale and below. This will allow researchers to characterize large scale structural changes that are often associated with photoreceptor or photoenzyme activation, providing an alternative to XFEL experiments for microsecond to millisecond protein dynamics studies.

6.5 Methods

6.5.1 Protein expression and purification

The gene coding for the phototropin 2 LOV2 domain from *Arabidopsis thaliana* (*At*Phot2LOV2) with a C-terminal hexahistidine tag was synthetized (Genecust, Ellange, Luxemburg) and inserted in a pBad vector using BamHI/PmeI cloning sites. The vector was transformed into the *Escherichia coli* host strain BL21 (Invitrogen) and grown on ampicillin (100 $\mu g.mL^{-1}$) selective agar plates. Cells were grown in 2YT medium supplemented with ampicillin (100 $\mu g.mL^{-1}$) at 37 °C until an optical density at 600 nm of 0.6 was reached. Protein expression was then induced with 0.02% arabinose and cells were grown overnight at 17 °C. Cells were harvested and centrifuged at 4000 g for 20 min at 4 °C. Pellets were resuspended in 25 mL of lysis buffer (50 mM Tris pH 8.0, 300 mM NaCl, 10 mM imidazole, 0.25 $mg.mL^{-1}$ lysozyme, 400 $\mu g.mL^{-1}$ DNAse I, 20 mM $MgSO_4$ and Anti-protease Complete EDTA-free) per litre of centrifuged medium and frozen at -80 °C. Thawed pellets were sonicated and cell debris were centrifuged at 15000 g, for 45 min at 4 °C. The protein was purified from the clarified lysate using a nickel affinity column (His-Trap HP 5 mL, GE HealthCare) and eluted against an imidazole gradient (50 mM Tris pH 8.0, 300 mM NaCl, 10 to 500 mM imidazole). A second purification step consisted of a size exclusion chromatography (Superdex 75 - 10/300 GL, GE HealthCare), after which the purified proteins was concentrated to 8 $mg.mL^{-1}$.

6.5.2 Protein crystallization

Prior to crystallization, *At*Phot2LOV2 aliquots were submitted to limited proteolysis using 0.5 µg.µL^{-1} trypsin, a step which had been found to be crucial in controlling crystal quality. A first batch of crystals were produced by the hanging drop vapour diffusion method (2 µL drops of 1:1 mother liquor: protein ratio) in a condition consisting of 12 – 17% PEG8000 and 200 mM calcium acetate in 100 mM MES buffer pH 6.0. This resulted in square-shaped crystals of heterogeneous size, which were then used for micro-seeding with a 1:100 dilution of the seed solution with the protein solution. The resulting crystals are of a much more homogeneous size of 50 x 50 x 50 µm^3.

6.5.3 *In crystallo* UV-visible absorption spectroscopy

UV-vis absorption spectra of *At*Phot2LOV2 were obtained at the *ic*OS Lab at the ESRF (von Stetten *et al.*, 2015). The reference light is provided by a DH2000-BAL lamp (Ocean Optics, Largo, Florida) and connected to the setup by a 200 µm fibre. The transmitted light is connected to a grating-based QE65Pro spectrophotometer (Ocean Optics, Largo, Florida) *via* a 400 µm fibre to the output detector. The actinic light is output by a LED of peak wavelength at 470 nm (Model M470F3, Thorlabs, Newton, New Jersey) connected to the setup *via* a 1000 µm fibre. The tuneable power of the LED was checked with an energy meter (Nova II, Ophir, Jerusalem, Israel) in the range 0 to 27.5 mW. The resulting power on the crystal was thus varied between 0 and 3.2 µW.

6.5.4 Diffraction data collection and sorting

All diffraction data were collected on beamline ID30A-3 of the ESRF (Theveneau *et al.*, 2013). Each crystal was maintained at room temperature using a humidity controller (HC1, Arinax, Voreppe, France) with a relative humidity set at 99.65% (value determined from the composition of the mother liquor (Wheeler *et al.*, 2012). The LED output was collimated at the entrance of a 5x magnifying objective positioned at 75 mm from the sample position so as to deliver 3.5 µW on the crystal with a 2.2 mm focal spot (**Supplementary Figure S3**). The limited illumination power ensures a gradual, slow photo-conversion of the whole crystal from the top right corner (source of the illumination) to the bottom left corner (**Supplementary Video 1**).

Crystals were manually mounted one by one on the MD2 microdiffractometer (Arinax, Voreppe, France) with the top right corner of the crystals centred, to our best efforts, in the 15 µm diameter X-ray beam. Each data collection comprised 1,000 images of 0.5° oscillation images collected with an exposure time of 4.2 ms. For each data set the first 990 images were binned into for 66 x 15-image sub data collections, corresponding to 66 time points. In order to prevent significant resolution loss during the data collection due to global radiation damage, the X-ray flux was limited to $4.4 – 5.7 \times 10^{10}$ ph s^{-1}, corresponding to an absorbed dose of 115-162 kGy per full data collection (calculated using Raddose3D (Zeldin *et al.*, 2013)). We performed 88 such data collections to monitor the build-up of the light state. The LED was triggered by the rising front of the master TTL signal of the Eiger 4M detector, which signals the start of the data collection synchronized with X-ray shutter opening. In order to generate our reference dark state, we repeated this experiment on 32 different crystals without using the LED.

6.5.5 Diffraction data set processing

For each of the 88 light state data sets each 15 image sub-wedge was separately integrated using *XDS* (Kabsch, 2010) in space group $P4_32_12$. The 66-sub-wedges corresponding to the 66 time points were then subject to analysis using *ccCluster* (Santoni *et al.*, 2017), which clusters isomorphous sub-data sets and eliminates strong outliers. Based on the output of the program, suitable clusters of sub-wedges were carried forward to the next stage. Here, reflections were combined into a single file using *Pointless* (Evans, 2011) and symmetry-related reflections then scaled and merged using *Aimless* (Evans & Murshudov, 2013). Final resolution limits were assigned using a $CC_{1/2}$ cutoff of 0.70 in the highest resolution shell with $\langle I/\sigma(I) \rangle$ greater than 1.0. Inclusion of clustering nodes with increasing distance cutoffs was iteratively tested until the diffraction resolution, as defined above, could be maximized. The dark state data set was composed by merging the 500 first images of the best 5 sub-data set cluster obtained from the 32 data sets recorded in the absence of LED illumination, in order to obtain a high quality reference data set.

6.5.6 Structure determination and refinement of the dark and light states

A model for the dark state data set (time point t < 0 ms) was constructed from a previously-determined room temperature structure of *At*Phot2LOV2 (PDB entry code 6QQJ, (Gotthard *et al.*, in press)). Structure refinement (including the 4 parameters x, y, z, B for each atom and group occupancy for each residue in alternate conformation) was performed using *Phenix* (Adams *et al.*, 2010) with iterative inspection of electron density maps in *Coot* (Emsley *et al.*, 2010), which led to the identification of 10 residues in alternative conformations A and B, including Cys426.

A model for the light state was built using the data set at time point t = 4,158 ms. Initial maps were calculated by inserting the fixed dark state structure determined at t < 0 ms. Residual densities in $F_{obs} - F_{calc}$ and $2F_{obs} - F_{calc}$ maps were investigated, and checked in maps at earlier time points for consistency. A number of residues showed clear displacements and were modelled as alternate conformation C. In order to allow for the accommodation of significant movements of the main chain, residues flanking moving residues were also modelled in alternate conformations. Overall, six continuous stretches of the protein sequence were modelled in alternate conformations A and C (**Supplementary Fig. 1**). The two key residues Cys426 and Phe470 were modelled in three conformations: $Cys426_A$, $Cys426_B$ and $Phe470_A$ in the dark state, and $Cys426_C$, $Phe470_C$ and $Phe470_D$ in the light state. For consistency of occupancy refinement, the 10 residues having a conformer B in those stretches have been restrained to their A conformer in the fixed dark state. Once a satisfying atomic model of the light-state in the latest time point 66 had been obtained, occupancies for each of the 6 strands were refined using this model for each time point.

6.5.7 Occupancy refinement for all time points

Structure refinement of each time point between t = 63 and t = 4095 ms was performed by inserting a combination of the dark and light states with occupancies 0.5/0.5 for conformation A and C, and 0.33/0.33/0.33 for conformations A/B/C or A/C/D. The only parameters that have been refined are the occupancy of protein segments **1** to **6**.

6.6 References

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010). *Acta Cryst. D.* **66**, 213–221.

Broennimann, C., Eikenberry, E. F., Henrich, B., Horisberger, R., Huelsen, G., Pohl, E., Schmitt, B., Schulze-Briese, C., Suzuki, M., Tomizaki, T., Toyokawa, H. & Wagner, A. (2006). *J. Synchrotron Rad.* **13**, 120–130.

Casanas, A., Warshamanage, R., Finke, A. D., Panepucci, E., Olieric, V., Nöll, A., Tampé, R., Brandstetter, S., Förster, A., Mueller, M., Schulze-Briese, C., Bunk, O. & Wang, M. (2016). *Acta Cryst. D.* **72**, 1036–1048.

Chapman, H. N., Fromme, P., Barty, A., White, T. A., Kirian, R. A., Aquila, A., Hunter, M. S., Schulz, J., Deponte, D. P., Weierstall, U., Doak, R. B., Maia, F. R. N. C., Martin, A. V, Schlichting, I., Lomb, L., Coppola, N., Shoeman, R. L., Epp, S. W., Hartmann, R., Rolles, D., Rudenko, A., Foucar, L., Kimmel, N., Weidenspointner, G., Holl, P., Liang, M., Barthelmess, M., Caleman, C., Boutet, S., Bogan, M. J., Krzywinski, J., Bostedt, C., Bajt, S., Gumprecht, L., Rudek, B., Erk, B., Schmidt, C., Hömke, A., Reich, C., Pietschner, D., Ströder, L., Hauser, G., Gorke, H., Ullrich, J., Herrmann, S., Schaller, G., Schopper, F., Soltau, H., Kühnel, K. U., Messerschmidt, M., Bozek, J. D., Hau-Riege, S. P., Frank, M., Hampton, C. Y., Sierra, R. G., Starodub, D., Williams, G. J., Hajdu, J., Timneanu, N., Seibert, M. M., Andreasson, J., Rocker, A., Jönsson, O., Svenda, M., Stern, S., Nass, K., Andritschke, R., Schröter, C. D., Krasniqi, F., Bott, M., Schmidt, K. E., Wang, X., Grotjohann, I., Holton, J. M., Barends, T. R. M., Neutze, R., Marchesini, S., Fromme, R., Schorb, S., Rupp, D., Adolph, M., Gorkhover, T., Andersson, I., Hirsemann, H., Potdevin, G., Graafsma, H., Nilsson, B. & Spence, J. C. H. (2011). *Nature (London).* **470**, 73–78.

Christie, J. M. (2007). *Annu. Rev. Plant Biol.* **58**, 21–45.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst. D.* **66**, 486–501.

Evans, P. R. (2011). *Acta Cryst. D.* **67**, 282–292.

Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst. D.* **69**, 1204–1214.

Gotthard, G., Aumonier, S., de Sanctis, D., Leonard, G., von Stetten, D. & Royant, A. (2019).

IUCrJ. **6**, 665–680.

Grünbein, M. L. & Kovacs, G. N. (2019). *Acta Cryst. D*. **75**, 178–191.

Hajdu, J., Wilmot, C. M., Neutze, R., Sjögren, T., Edman, K., Szöke, A. & Wilmouth, R. C. (2000). *Nat. Struct. Biol.* **7**, 1006–1012.

Halavaty, A. S. & Moffat, K. (2007). *Biochemistry*. **46**, 14001–14009.

Harper, S. M., Neil, L. C. & Gardner, K. H. (2003). *Science*. **301**, 1541–1544.

Johansson, L. C., Stauch, B., Ishchenko, A. & Cherezov, V. (2017). *Trends Biochem. Sci.* **42**, 749–762.

Kabsch, W. (2010). *Acta Cryst. D*. **66**, 125–132.

Kennis, J. T. M., Crosson, S., Gauden, M., Van Stokkum, I. H. M., Moffat, K. & Van Grondelle, R. (2003). *Biochemistry*. **42**, 3385–3392.

Kottke, T., Heberle, J., Hehn, D., Dick, B. & Hegemannt, P. (2003). *Biophys. J.* **84**, 1192–1201.

Santoni, G., Zander, U., Mueller-Dieckmann, C., Leonard, G. & Popov, A. (2017). *J. Appl. Cryst.* **50**, 1844–1851.

Schotte, F., Cho, H. S., Kaila, V. R. I., Kamikubo, H., Dashdorj, N., Henry, E. R., Graber, T. J., Henning, R., Wulff, M., Hummer, G., Kataoka, M. & Anfinrud, P. A. (2012). *Proc. Natl Acad. Sci. USA*. **109**, 19256–19261.

Schotte, F., Lim, M., Jackson, T. A., Smirnov, A. V, Soman, J., Olson, J. S., Phillips, G. N., Wulff, M. & Anfinrud, P. a (2003). *Science*. **300**, 1944–1947.

Schulz, E. C., Mehrabi, P., Müller-Werkmeister, H. M., Tellkamp, F., Jha, A., Stuart, W., Persch, E., De Gasparo, R., Diederich, F., Pai, E. F. & Miller, R. J. D. (2018). *Nat. Methods*. **15**, 901–904.

von Stetten, D., Giraud, T., Carpentier, P., Sever, F., Terrien, M., Dobias, F., Juers, D. H., Flot, D., Mueller-Dieckmann, C., Leonard, G. a., de Sanctis, D. & Royant, A. (2015). *Acta Cryst. D*. **71**, 15–26.

Swartz, T. E., Corchnoy, S. B., Christie, J. M., Lewis, J. W., Szundi, I., Briggs, W. R. & Bogomolni, R. A. (2001). *J. Biol. Chem.* **276**, 36493–36500.

Theveneau, P., Baker, R., Barrett, R., Beteva, A., Bowler, M. W., Carpentier, P., Caserotto, H., De Sanctis, D., Dobias, F., Flot, D., Guijarro, M., Giraud, T., Lentini, M., Leonard, G. A., Mattenet, M., McCarthy, A. A., McSweeney, S. M., Morawe, C., Nanao, M., Nurizzo, D., Ohlsson, S., Pernot, P., Popov, A. N., Round, A., Royant, A., Schmid, W., Snigirev, A., Surr, J. & Mueller-Dieckmann, C. (2013). *J. Phys. Conf. Ser*. **425**, 012001.

Wheeler, M. J., Russi, S., Bowler, M. G. & Bowler, M. W. (2012). *Acta Cryst. F* . **68**, 111–114.

Wöhri, A. B., Katona, G., Johansson, L. C., Fritz, E., Malmerberg, E., Andersson, M., Vincent, J., Eklund, M., Cammarata, M., Wulff, M., Davidsson, J., Groenhof, G. & Neutze, R. (2010). *Science*. **328**, 630–633.

Zeldin, O. B., Gerstel, M. & Garman, E. F. (2013). *J. Appl. Cryst.* **46**, 1225–1230.

6.7 Supplementary information

Pages 6-19 Supplementary Tables 1 and 2

Page 6-24 Supplementary Figures 1 to 7

Page 6-31 Supplementary Videos 1 to 3 (legends)

Supplementary Table 1 | Data reduction and refinement statistics for all 67 data sets.

Time point	Time [ms]	Integr. data sets	Retained data sets	Distance cutoff	Unit cell a [Å]	Unit cell c [Å]	Resolution range [Å]	Unique reflections	$CC_{1/2}$	I/σ(I)	Multiplicity	R_{pim}	Completeness [%]	R_{cryst} [%]	R_{free} [%]	rmsd on bonds [Å]	rmsd on angles [°]
0	< 0	32	5	0.43	40.83	133.64	44.55-2.20 (2.27-2.20)	553457 (6340)	0.999 (0.921)	14.0 (1.9)	87.3 (93.9)	0.051 (0.557)	100.0 (100.0)	20.2 (33.3)	24.9 (36.7)	0.002	0.55
1	63	50	26	0.39	40.73	133.28	44.43-2.45 (2.55-2.45)	4613 (500)	0.964 (0.775)	6.0 (2.8)	12.6 (13.0)	0.161 (0.795)	100.0 (100.0)	24.8 (38.2)	27.4 (0.3176)	0.003	0.65
2	126	52	38	0.41	40.77	133.45	44.48-2.50 (2.60-2.50)	4370 (478)	0.981 (0.812)	6.1 (2.9)	18.4 (19.1)	0.151 (0.813)	100.0 (100.0)	0.22.9 (29.7)	26.4 (32.4)	0.003	0.65
3	189	50	29	0.40	40.73	133.34	40.73-2.50 (2.60-2.50)	4365 (474)	0.979 (0.773)	6.7 (2.9)	14.0 (14.4)	0.151 (0.779)	100.0 (100.0)	23.0 (32.3)	25.0 (28.3)	0.003	0.65
4	252	48	39	0.45	40.77	133.38	40.77-2.60 (2.72-2.60)	3895 (461)	0.972 (0.739)	6.3 (2.7)	19.0 (19.9)	0.155 (0.899)	100.0 (100.0)	22.5 (30.8)	25.2 (29.7)	0.003	0.65
5	315	47	28	0.40	40.76	133.43	38.98-2.50 (2.60-2.50)	4369 (480)	0.967 (0.701)	5.4 (2.4)	13.6 (14.1)	0.189 (1.144)	100.0 (100.0)	23.4 (33.6)	26.6 (26.4)	0.003	0.65
6	378	44	27	0.45	40.76	133.48	44.49-2.55 (2.66-2.55)	4128 (471)	0.973 (0.720)	6.1 (2.6)	13.1 (13.7)	0.179 (0.912)	100.0 (100.0)	22.7 (31.9)	25.2 (28.9)	0.003	0.65
7	441	44	33	0.50	40.79	133.42	44.47-2.45 (2.55-2.45)	4639 (507)	0.965 (0.700)	5.1 (1.8)	16.0 (16.8)	0.242 (1.729)	100.0 (100.0)	22.9 (33.7)	25.6 (30.3)	0.003	0.65
8	504	42	33	0.52	40.74	133.16	44.39-2.45 (2.55-2.45)	4620 (502)	0.955 (0.718)	5.1 (1.8)	16.0 (16.5)	0.284 (2.092)	100.0 (100.0)	23.2 (36.6)	26.3 (36.5)	0.003	0.65
9	567	40	33	0.52	40.78	133.38	40.78-2.65 (2.78-2.65)	3708 (465)	0.970 (0.749)	5.7 (2.4)	15.9 (16.6)	0.167 (0.852)	100.0 (100.0)	21.7 (31.6)	24.7 (32.8)	0.003	0.65
10	630	41	36	0.51	40.80	133.46	39.02-2.65 (2.78-2.65)	3707 (461)	0.975 (0.776)	5.0 (1.8)	17.5 (18.5)	0.172 (1.091)	100.0 (100.0)	22.0 (33.7)	24.3 (36.8)	0.003	0.65
11	693	42	29	0.44	40.79	133.38	39.00-2.75 (2.90-2.75)	3320 (465)	0.971 (0.887)	6.2 (3.0)	13.9 (14.5)	0.138 (0.526)	99.9 (100.0)	21.5 (27.3)	23.7 (31.9)	0.003	0.65
12	756	44	26	0.44	40.78	133.35	39.00-2.75 (2.90-2.75)	3320 (465)	0.966 (0.770)	5.5 (2.6)	12.6 (13.1)	0.153 (0.652)	99.9 (99.9)	22.1 (30.0)	25.7 (32.7)	0.003	0.65
13	819	45	43	0.56	40.84	133.41	44.47-2.45 (2.55-2.45)	4642 (498)	0.982 (0.743)	7.0 (1.8)	41.8 (43.2)	0.238 (2.386)	100.0 (100.0)	24.3 (40.3)	28.6 (42.2)	0.003	0.65
14	882	44	38	0.50	40.82	133.32	44.44-2.70 (2.83-2.70)	3518 (433)	0.968 (0.734)	4.9 (1.5)	18.3 (19.1)	0.165 (1.012)	100.0 (100.0)	22.4 (34.4)	25.3 (35.0)	0.003	0.65
15	945	42	34	0.53	40.81	133.36	40.81-2.75 (2.90-2.75)	3310 (469)	0.969 (0.764)	5.4 (2.3)	16.5 (17.2)	0.164 (0.756)	100.0 (100.0)	22.5 (31.9)	27.3 (43.9)	0.003	0.65

Time point	Time [ms]	Integr. data sets	Retained data sets	Distance cutoff	Unit cell a [Å]	Unit cell c [Å]	Resolution range [Å]	Unique reflections	$CC_{1/2}$	I/σ(I)	Multiplicity	R_{pim}	Completeness [%]	R_{cryst} [%]	R_{free} [%]	rmsd on bonds [Å]	rmsd on angles [°]
16	1008	42	35	0.53	40.84	133.36	44.45-2.75 (2.90-2.75)	3328 (468)	0.972 (0.863)	5.5 (2.1)	16.8 (17.8)	0.170 (0.770)	100.0 (100.0)	22.6 (30.9)	25.1 (28.9)	0.003	0.65
17	1071	43	36	0.60	40.85	133.28	44.43-2.75 (2.90-2.75)	3333 (467)	0.965 (0.841)	5.1 (2.1)	17.3 (18.0)	0.178 (0.878)	100.0 (100.0)	21.7 (33.1)	26.3 (30.1)	0.003	0.65
18	1134	41	31	0.53	40.80	133.18	40.80-2.75 (2.90-2.75)	3321 (466)	0.963 (0.811)	5.3 (2.3)	14.9 (15.5)	0.185 (0.875)	100.0 (100.0)	22.6 (31.6)	26.2 (33.3)	0.003	0.65
19	1197	46	34	0.55	40.87	133.31	44.44-2.65 (2.78-2.65)	3724 (466)	0.966 (0.710)	4.6 (1.4)	16.4 (17.0)	0.202 (1.392)	100.0 (100.0)	22.9 (35.5)	26.7 (23.1)	0.003	0.65
20	1260	41	30	0.52	40.83	133.24	44.41-2.75 (2.90-2.75)	3334 (469)	0.962 (0.724)	5.4 (2.4)	14.3 (14.9)	0.187 (0.787)	100.0 (100.0)	22.3 (30.2)	26.8 (35.0)	0.003	0.65
21	1323	42	32	0.51	40.84	133.19	44.40-2.65 (2.78-2.65)	3708 (464)	0.964 (0.790)	4.9 (1.5)	15.5 (16.2)	0.239 (1.697)	100.0 (100.0)	22.7 (32.2)	26.4 (32.2)	0.003	0.65
22	1386	44	33	0.58	40.84	133.33	44.44-2.75 (2.90-2.75)	3340 (469)	0.970 (0.709)	4.8 (1.9)	15.9 (16.5)	0.177 (0.814)	100.0 (100.0)	22.7 (29.8)	26.4 (36.0)	0.003	0.65
23	1449	46	38	0.61	40.91	133.40	44.47-2.75 (2.90-2.75)	3347 (466)	0.970 (0.837)	4.8 (1.8)	18.2 (19.0)	0.186 (0.875)	100.0 (100.0)	22.6 (32.0)	26.2 (42.7)	0.003	0.65
24	1512	46	36	0.66	40.85	133.35	40.85-2.90 (3.08-2.90)	2870 (423)	0.971 (0.707)	5.4 (2.5)	17.0 (17.3)	0.156 (0.680)	100.0 (100.0)	21.5 (26.2)	23.6 (23.0)	0.003	0.65
25	1575	44	38	0.67	40.90	133.31	44.44-2.75 (2.90-2.75)	3348 (469)	0.968 (0.754)	5.2 (2.1)	18.1 (18.9)	0.194 (0.878)	100.0 (100.0)	22.8 (30.2)	26.4 (37.6)	0.003	0.65
26	1638	43	37	0.70	40.91	133.37	44.46-2.65 (2.78-2.65)	3726 (466)	0.957 (0.716)	4.5 (1.5)	17.7 (18.4)	0.248 (1.655)	100.0 (100.0)	22.8 (31.6)	25.6 (26.1)	0.003	0.65
27	1701	45	34	0.71	40.91	133.30	44.43-2.75 (2.90-2.75)	335 (471)	0.957 (0.811)	4.9 (2.0)	16.3 (16.9)	0.204 (0.844)	100.0 (100.0)	22.9 (29.7)	26.7 (31.0)	0.003	0.65
28	1764	47	35	0.61	40.91	133.41	40.91-2.70 (2.83-2.70)	3515 (434)	0.948 (0.708)	4.7 (1.6)	17.0 (17.5)	0.247 (1.537)	100.0 (100.0)	23.5 (30.4)	26.4 (34.0)	0.003	0.65
29	1827	42	37	0.66	40.88	133.46	40.88-2.70 (2.83-2.70)	3521 (435)	0.962 (0.702)	4.5 (1.5)	17.9 (18.8)	0.217 (1.301)	100.0 (100.0)	23.5 (33.0)	27.9 (39.9)	0.003	0.65
30	1890	45	37	0.71	40.91	133.33	44.44-2.60 (2.72-2.60)	3940 (461)	0.959 (0.704)	4.3 (1.2)	18.0 (19.0)	0.311 (2.235)	100.0 (100.0)	23.6 (30.2)	27.5 (43.3)	0.003	0.65
31	1953	45	25	0.59	40.88	133.48	44.49-2.75 (2.90-2.75)	3344 (460)	0.946 (0.774)	4.5 (1.8)	11.9 (12.4)	0.208 (0.976)	100.0 (100.0)	23.3 (30.5)	26.6 (30.1)	0.003	0.65
32	2016	42	32	0.55	40.89	133.46	40.89-2.75 (2.90-2.75)	3346 (463)	0.940 (0.810)	5.1 (2.0)	15.3 (15.7)	0.213 (0.871)	100.0 (100.0)	23.0 (33.5)	28.2 (40.6)	0.003	0.65

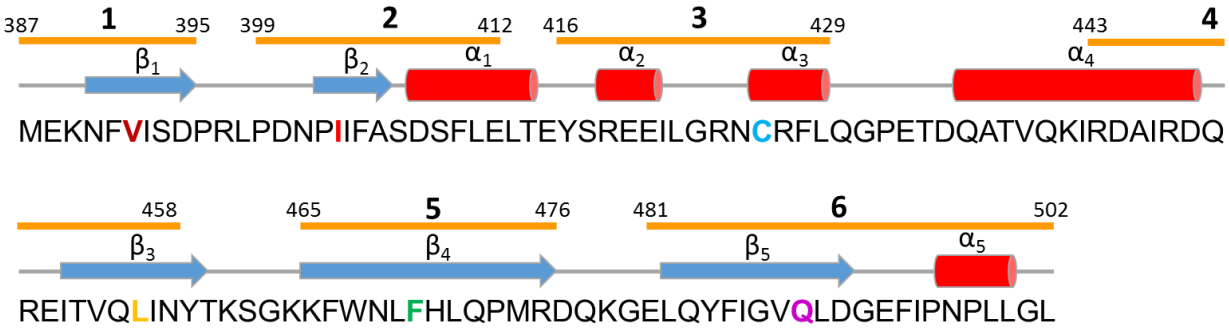
Time point	Time [ms]	Integr. data sets	Retained data sets	Distance cutoff	Unit cell a [Å]	Unit cell c [Å]	Resolution range [Å]	Unique reflections	$CC_{1/2}$	I/σ(I)	Multiplicity	R_{pim}	Completeness [%]	R_{cryst} [%]	R_{free} [%]	rmsd on bonds [Å]	rmsd on angles [°]
33	2079	40	34	0.63	40.92	133.35	44.45-2.75 (2.90-2.75)	3358 (472)	0.949 (0.720)	6.9 (1.7)	15.8 (16.5)	0.232 (1.164)	100.0 (100.0)	23.3 (30.5)	26.1 (47.2)	0.003	0.65
34	2142	45	40	0.67	40.95	133.37	44.46-2.75 (2.90-2.75)	3361 (468)	0.956 (0.772)	5.1 (1.8)	18.0 (18.8)	0.210 (1.095)	100.0 (100.0)	23.3 (31.3)	27.0 (32.9)	0.003	0.65
35	2205	39	37	0.62	40.98	133.45	44.48-2.75 (2.90-2.75)	3366 (469)	0.973 (0.846)	4.5 (1.1)	17.3 (18.0)	0.159 (0.927)	100.0 (100.0)	23.4 (32.8)	26.7 (38.2)	0.003	0.65
36	2268	43	36	0.72	40.99	133.49	44.50-2.75 (2.90-2.75)	3366 (469)	0.976 (0.798)	4.4 (1.0)	16.9 (17.8)	0.162 (0.959)	100.0 (100.0)	24.0 (37.6)	26.2 (48.4)	0.003	0.65
37	2331	41	40	0.69	40.96	133.38	44.46-2.75 (2.90-2.75)	3357 (468)	0.967 (0.698)	4.6 (1.2)	19.0 (19.6)	0.160 (0.961)	100.0 (100.0)	24.0 (34.1)	28.8 (53.0)	0.003	0.65
38	2394	42	37	0.77	40.98	133.38	40.98-2.75 (2.90-2.75)	3361(469)	0.979 (0.802)	4.5 (1.2)	17.4 (18.2)	0.151 (0.885)	100.0 (100.0)	24.1 (32.8)	27.0 (43.4)	0.003	0.65
39	2457	42	40	0.78	40.99	133.40	40.99-2.75 (2.90-2.75)	3355 (469)	0.976 (0.806)	4.5 (1.2)	18.8 (19.5)	0.159 (0.885)	100.0 (100.0)	24.1 (34.2)	28.0 (41.8)	0.003	0.65
40	2520	42	38	0.66	40.96	133.29	44.43-2.75 (2.90-2.75)	3346 (471)	0.979 (0.809)	4.5 (1.2)	18.0 (18.8)	0.154 (0.850)	100.0 (100.0)	23.9 (35.7)	27.4 (37.6)	0.003	0.65
41	2583	41	36	0.72	40.98	133.40	40.98-2.75 (2.90-2.75)	3364 (469)	0.910 (0.744)	4.5 (1.7)	16.8 (17.4)	0.308 (0.971)	100.0 (100.0)	25.3 (30.8)	28.6 (37.4)	0.003	0.65
42	2646	39	33	0.74	41.01	133.41	41.01-2.75 (2.90-2.75)	3365 (467)	0.970 (0.762)	4.3 (1.3)	15.4 (16.1)	0.167 (0.879)	100.0 (100.0)	24.4 (33.5)	29.1 (44.4)	0.003	0.65
43	2709	40	36	0.71	41.01	133.44	41.01-2.75 (2.90-2.75)	3365 (469)	0.976 (0.940)	4.4 (1.2)	17.4 (16.9)	0.163 (0.905)	100.0 (100.0)	24.0 (33.4)	28.4 (39.1)	0.003	0.65
44	2772	43	40	0.92	41.00	133.37	44.46-2.75 (2.90-2.75)	3366 (467)	0.818 (0.819)	4.5 (1.3)	18.8 (19.5)	0.205 (1.314)	100.0 (100.0)	24.1 (31.8)	30.1 (39.1)	0.003	0.65
45	2835	47	39	0.73	41.01	133.35	44.45-2.80 (2.95-2.80)	3208 (454)	0.975 (0.764)	4.6 (1.5)	18.2 (18.9)	0.179 (1.009)	100.0 (100.0)	24.0 (32.8)	26.7 (28.8)	0.003	0.65
46	2898	46	37	0.67	40.97	133.34	44.45-2.75 (2.90-2.75)	3361 (472)	0.969 (0.734)	4.5 (1.2)	17.2 (17.7)	0.165 (0.954)	100.0 (100.0)	23.7 (33.6)	27.1 (36.1)	0.003	0.65
47	2961	46	39	0.71	40.96	133.22	44.41-2.75 (2.90-2.75)	3348 (474)	0.971 (0.831)	4.5 (1.3)	18.8 (19.3)	0.181 (1.095)	100.0 (100.0)	24.2 (31.9)	26.4 (40.0)	0.003	0.65
48	3024	43	23	0.55	40.85	133.31	40.85-2.80 (2.95-2.80)	3179 (448)	0.963 (0.714)	5.1 (1.9)	10.5 (10.7)	0.173 (0.894)	100.0 (100.0)	23.2 (29.9)	27.0 (34.8)	0.003	0.65

Time point	Time [ms]	Integr. data sets	Retained data sets	Distance cutoff	Unit cell a [Å]	Unit cell c [Å]	Resolution range [Å]	Unique reflections	$CC_{1/2}$	I/σ(I)	Multiplicity	R_{pim}	Completeness [%]	R_{cryst} [%]	R_{free} [%]	rmsd on bonds [Å]	rmsd on angles [°]
49	3087	43	36	0.73	40.99	133.36	40.99-2.75 (2.90-2.75)	3364 (472)	0.799 (0.747)	5.5 (1.9)	16.6 (17.1)	0.558 (1.088)	100.0 (100.0)	26.7 (29.4)	26.8 (38.7)	0.003	0.65
50	3150	43	37	0.68	40.94	133.36	44.45-2.80 (2.95-2.80)	3195 (451)	0.969 (0.760)	4.7 (1.4)	17.2 (18.0)	0.172 (1.073)	100.0 (100.0)	23.5 (29.7)	27.4 (33.6)	0.003	0.65
51	3213	43	34	0.69	40.93	133.27	44.42-2.75 (2.90-2.75)	3353 (474)	0.846 (0.798)	4.1 (1.4)	15.9 (16.6)	0.401 (0.930)	100.0 (100.0)	25.3 (31.7)	29.2 (29.6)	0.003	0.65
52	3276	41	32	0.60	40.93	133.36	39.13-2.95 (3.13-2.95)	2729 (413)	0.930 (0.716)	5.3 (2.3)	14.8 (15.0)	0.207 (0.709)	99.9 (100.0)	23.6 (29.7)	26.2 (30.0)	0.003	0.65
53	3339	40	35	0.77	40.95	133.32	40.95-2.75 (2.90-2.75)	3352 (473)	0.938 (0.738)	4.5 (1.5)	16.4 (16.8)	0.213 (0.976)	100.0 (100.0)	24.3 (31.9)	29.3 (38.8)	0.003	0.65
54	3402	36	28	0.64	41.01	133.35	44.45-2.75 (2.90-2.75)	3362 (467)	0.955 (0.700)	4.7 (1.7)	13.0 (13.6)	0.203 (0.959)	100.0 (100.0)	24.4 (32.1)	27.7 (39.2)	0.003	0.65
55	3456	37	33	0.74	40.98	133.30	44.43-2.75 (2.90-2.75)	3363 (475)	0.938 (0.803)	4.7 (1.6)	15.4 (16.0)	0.189 (0.976)	100.0 (100.0)	24.4 (33.2)	27.4 (37.8)	0.003	0.65
56	3528	40	30	0.64	40.97	133.29	44.43-2.75 (2.90-2.75)	3363 (475)	0.969 (0.775)	4.9 (1.8)	14.0 (14.4)	0.203 (0.979)	100.0 (100.0)	24.8 (32.8)	28.3 (38.3)	0.003	0.65
57	3591	38	31	0.64	40.98	133.33	44.44-2.75 (2.90-2.75)	3365 (472)	0.872 (0.734)	4.8 (1.5)	14.4 (15.0)	0.347 (1.118)	100.0 (100.0)	24.7 (31.1)	29.1 (43.7)	0.003	0.65
58	3654	41	36	0.74	40.99	133.31	39-18-2.75 (2.90-2.75)	3359 (472)	0.975 (0.769)	4.5 (1.1)	17.0 (17.8)	0.165 (0.982)	100.0 (100.0)	24.3 (34.6)	30.1 (39.0)	0.003	0.65
59	3717	40	36	0.66	40.97	133.26	44.42-2.75 (2.90-2.75)	3351 (474)	0.958 (0.750)	5.0 (1.6)	16.9 (17.8)	0.200 (1.068)	100.0 (100.0)	24.4 (33.4)	28.3 (37.2)	0.003	0.65
60	3780	40	35	0.69	40.98	133.23	34.90-2.75 (2.90-2.75)	3357 (475)	0.946 (0.719)	4.5 (1.2)	16.5 (17.3)	0.174 (0.969)	100.0 (100.0)	24.4 (37.8)	28.9 (42.3)	0.003	0.65
61	3843	39	35	0.67	40.98	133.31	44.44-2.85 (3.00-2.85)	3043 (412)	0.970 (0.749)	4.7 (1.3)	16.3 (16.9)	0.142 (0.710)	100.0 (100.0)	24.0 (32.1)	27.5 (45.4)	0.003	0.65
62	3906	41	38	0.70	40.97	133.22	44.41-2.75 (2.90-2.75)	3361 (475)	0.975 (0.753)	4.5 (1.3)	17.8 (18.4)	0.182 (1.070)	100.0 (100.0)	24.6 (34.0)	28.4 (38.9)	0.003	0.65
63	3969	38	25	0.64	40.87	133.21	39.07-2.90 (3.08-2.90)	2853 (425)	0.953 (0.708)	5.0 (1.6)	11.6 (12.0)	0.158 (0.672)	99.9 (100.0)	23.4 (28.9)	28.4 (36.7)	0.003	0.65
64	4032	44	41	0.70	41.01	133.28	44.43-2.80 (2.95-2.80)	3195 (454)	0.979 (0.747)	4.9 (1.3)	19.2 (19.9)	0.147 (0.848)	100.0 (100.0)	24.1 (31.1)	27.4 (28.6)	0.003	0.65
65	4095	40	35	0.69	41.03	133.32	44.44-2.75 (2.90-2.75)	3364 (467)	0.972 (0.791)	4.5 (1.1)	16.4 (17.2)	0.163 (0.954)	100.0 (100.0)	24.7 (34.1)	28.3 (35.3)	0.003	0.65
66	4158	36	32	0.65	41.01	133.15	41.01-2.75 (2.90-2.75)	3355 (468)	0.974 (0.766)	4.5 (1.2)	15.0 (15.8)	0.146 (0.766)	100.0 (100.0)	24.2 (32.9)	29.2 (37.9)	0.003	0.65

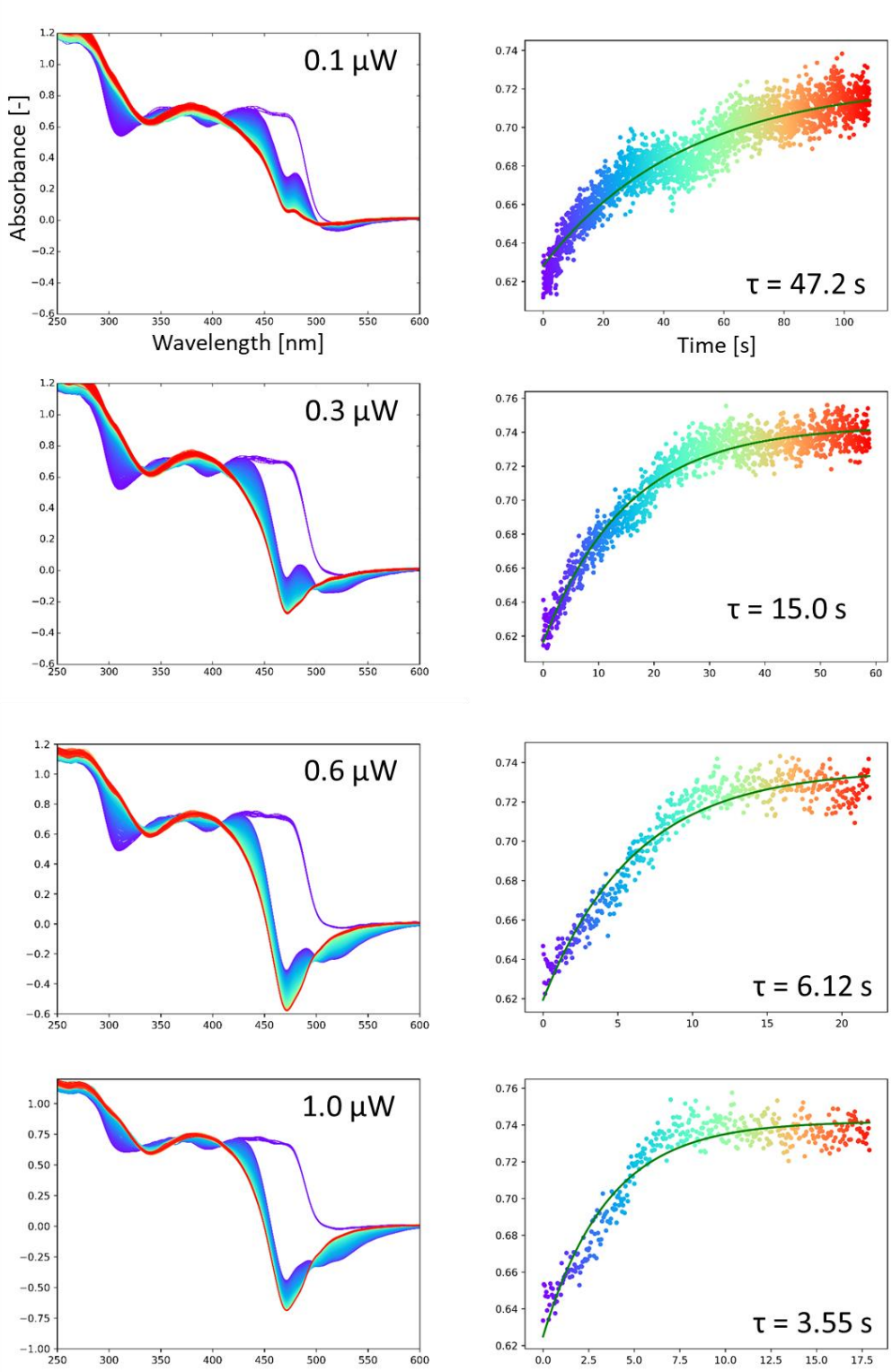
Supplementary Table 2 | Rise and decay time constants (obtained from mono-exponential fits) of key residue population evolutions derived from structural analysis.

Chapter 7 Residue (conformation)	**Chapter 8 Time constant [ms]**
Chapter 9 FMN (A,C)	Chapter 10 1325
Chapter 11 Val392 (A,C)	Chapter 12 1490
Chapter 13 Ile403 (A,C)	Chapter 14 1527
Chapter 15 Cys426 (A)	Chapter 16 1546
Chapter 17 Cys426 (B)	Chapter 18 1185
Chapter 19 Cys 426 (C)	Chapter 20 1434
Chapter 21 Leu456 (A, C)	Chapter 22 1527
Chapter 23 Phe470 (A)	Chapter 24 1360
Chapter 25 Phe470 (C)	Chapter 26 1304
Chapter 27 Phe470 (D)	Chapter 28 1576
Chapter 29 Gln489 (A, C)	Chapter 30 1621
Chapter 31 Average	**Chapter 32 1445 +/- 135**

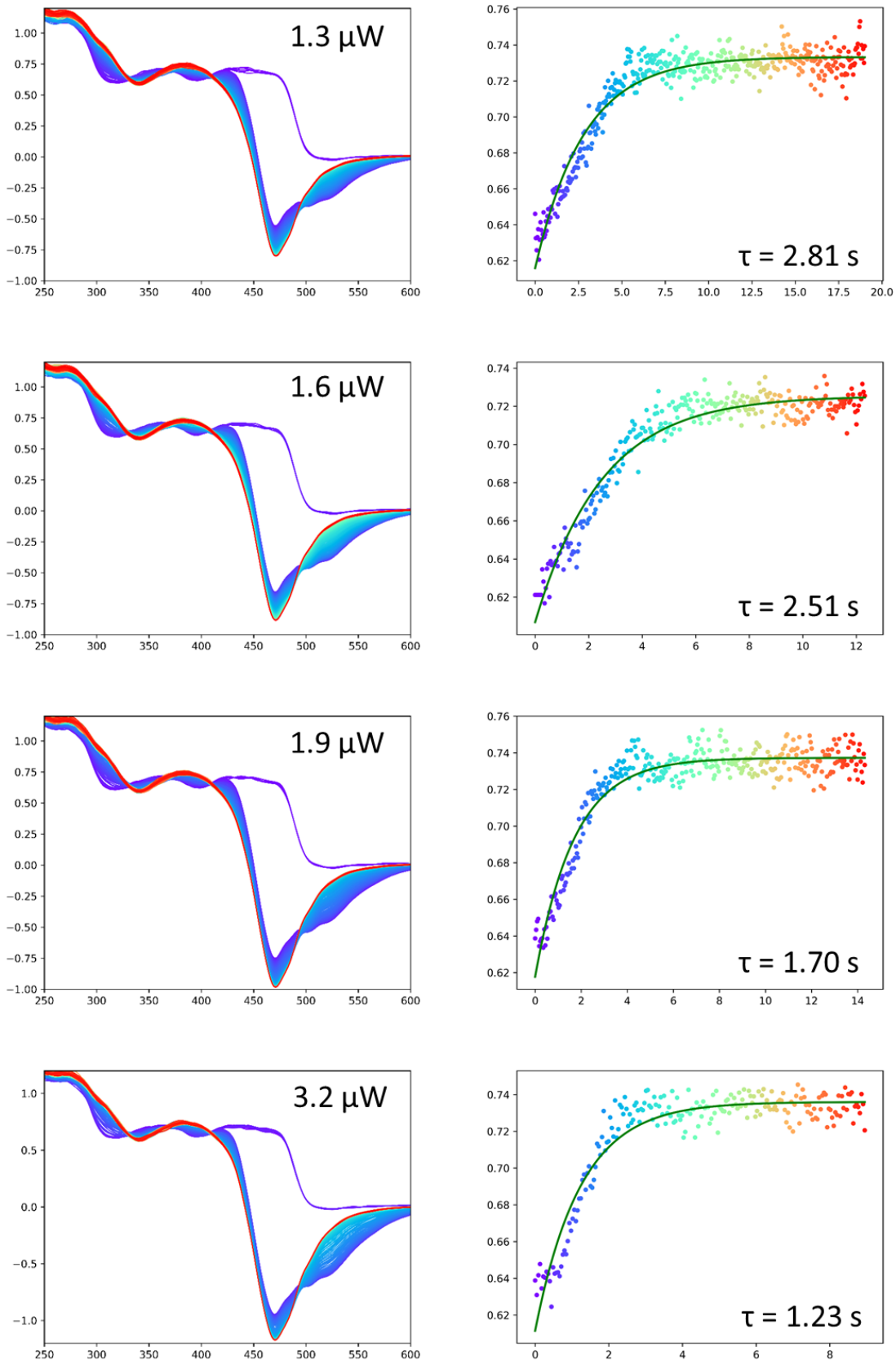
Supplementary Figure 1 | Sequence of the LOV2 domain from *Arabidopsis thaliana* featuring key residues in colour, secondary structure elements (α-helices in red, β-strands in blue) and protein segments 1 to 6 that move between the dark and the light state (represented in orange).



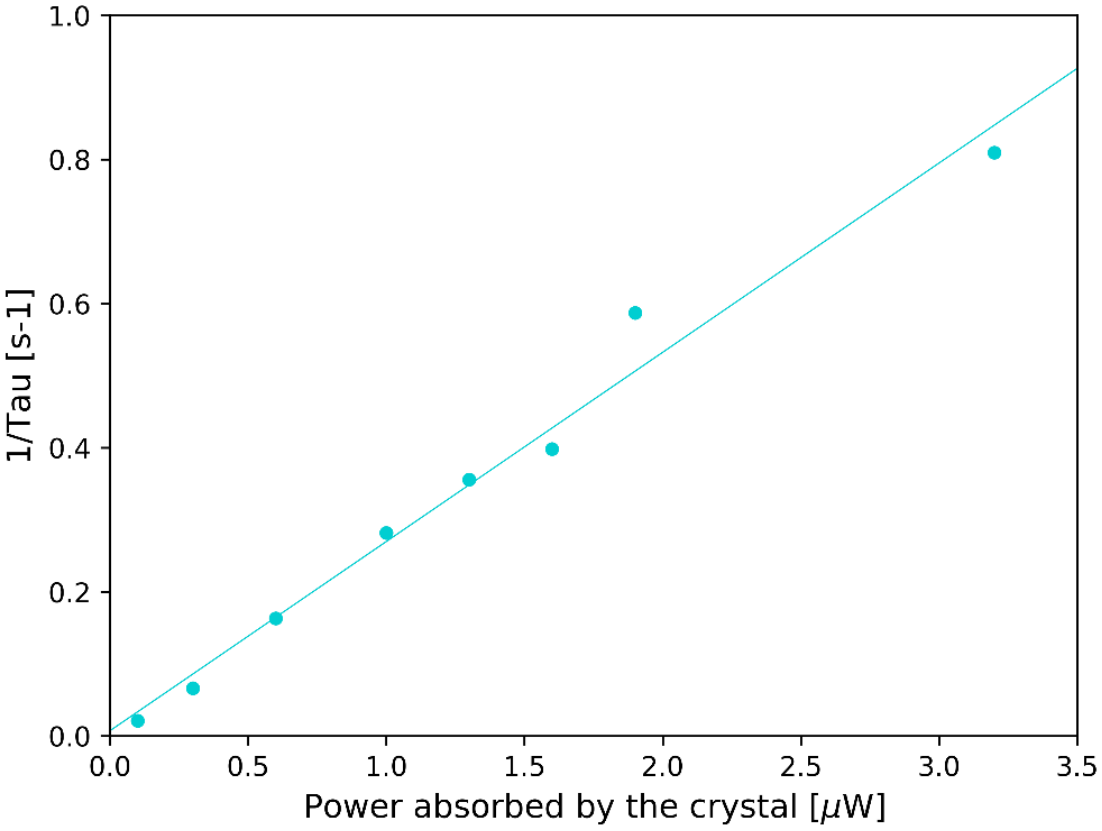
Supplementary Figure 2 | Series of UV-vis spectra of *At*phot2LOV2 crystals under illumination of a 470 nm LED with a nominal power of 0.1 to 5.0 mW. The evolution of the 390 nm peak is characteristic of light state build-up and can be modelled by a mono-exponential rise a+b*exp(-t/τ). The 1/τ *vs.* light power graph shows that the speed of light state build-up is linearly dependent on the number of photons per time unit.



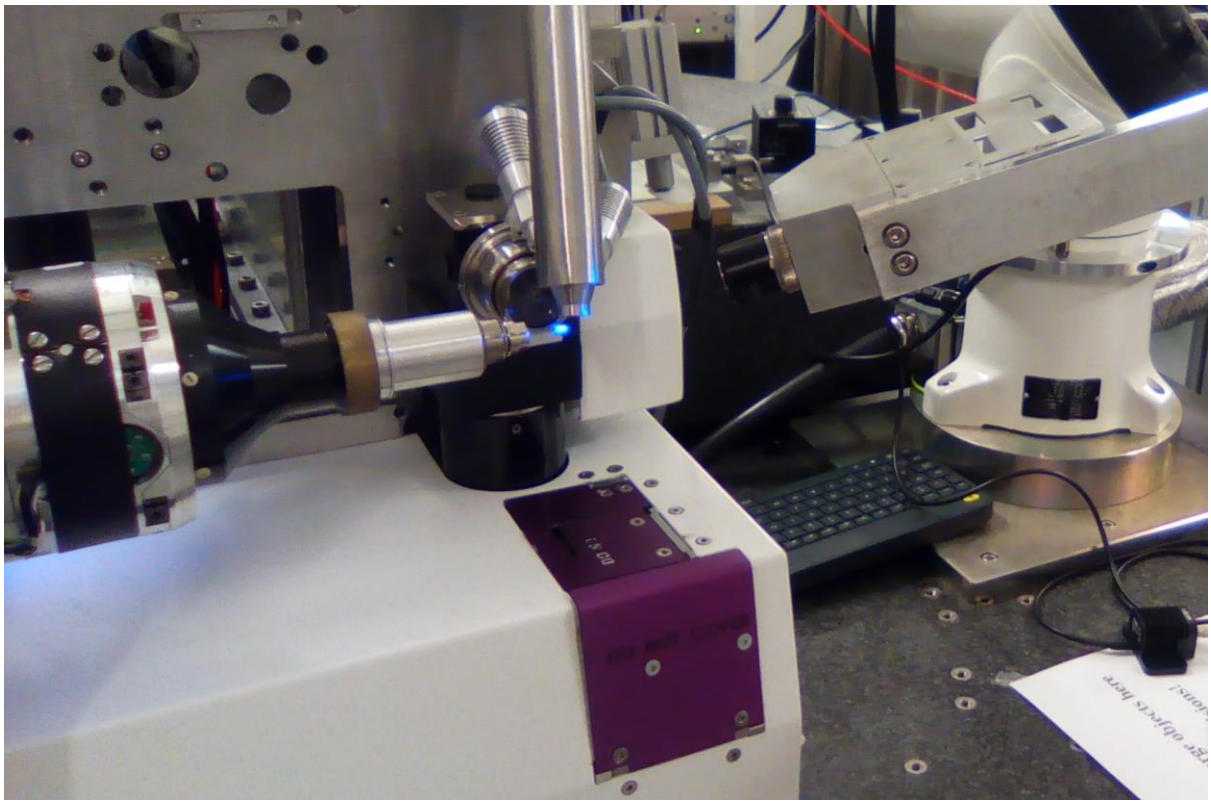
Supplementary Figure 2 (continued)



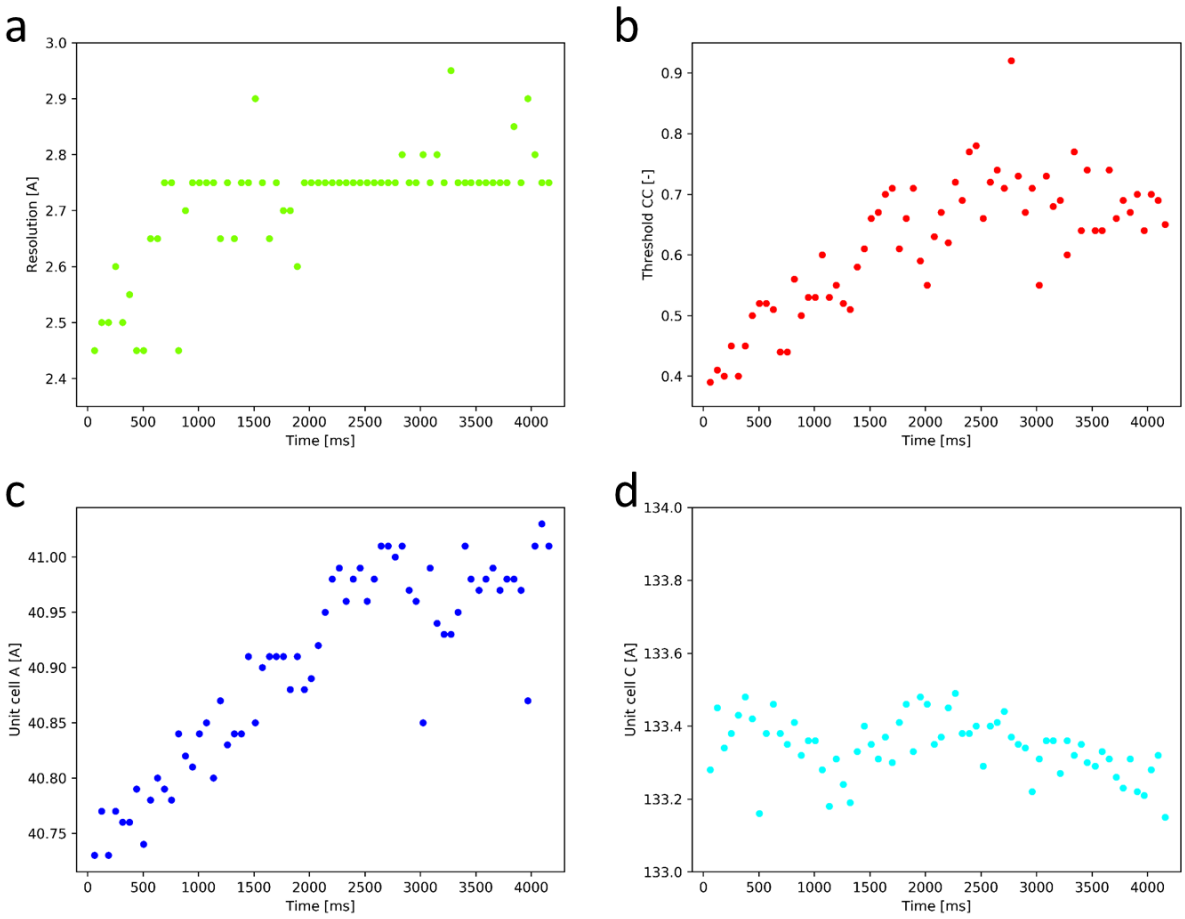
Supplementary Figure 2 (continued)



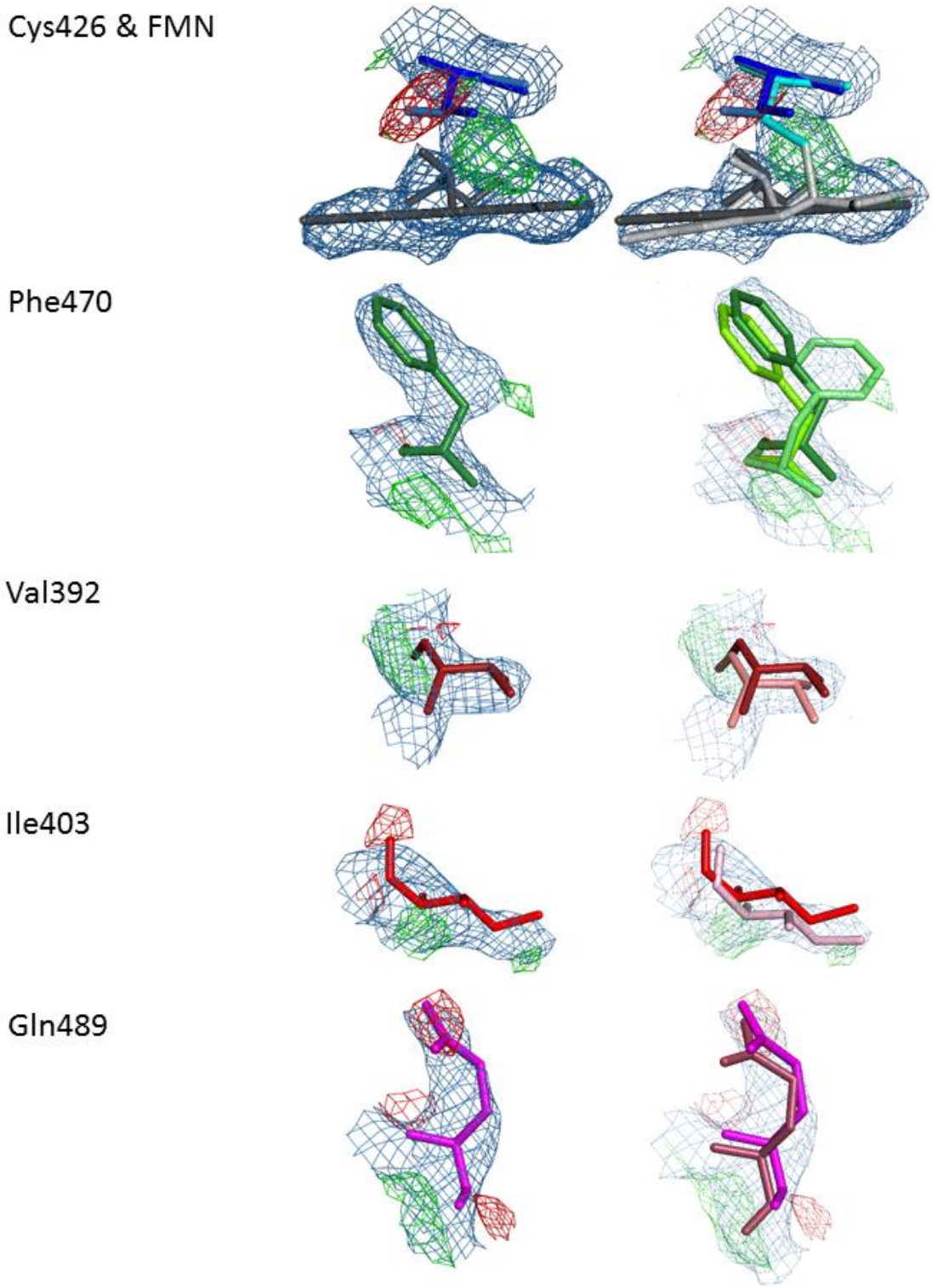
Supplementary Figure 3 | Photograph of the experimental setup on beamline ID30A-3. The crystal is maintained under the wet flux of an HC1 humidity controller (top) and is illuminated with a 470 nm LED (right).



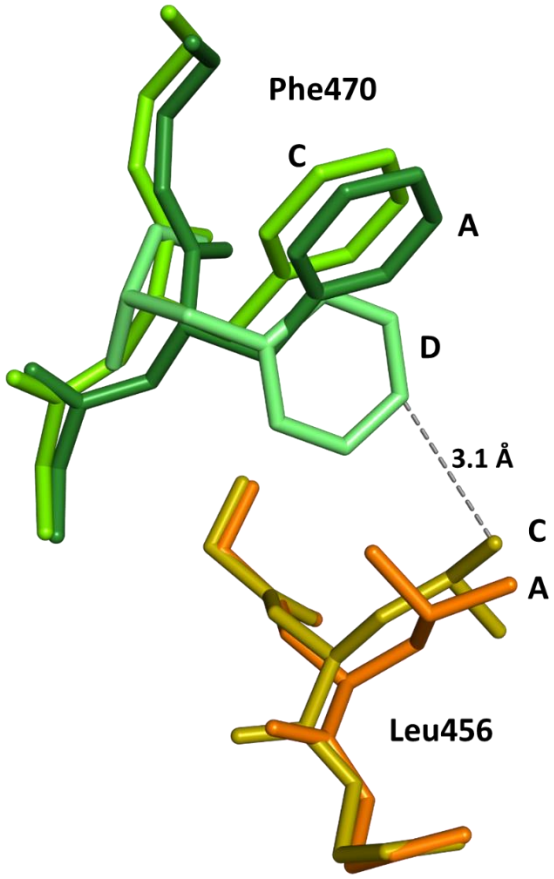
Supplementary Figure 4 | Evolution of various crystallographic data reduction parameters as a function of time: (**a**) diffraction resolution, (**b**) distance threshold for the clustering of sub data, (**c**) unit cell parameter ***a***, (**d**) unit cell parameter ***c***.



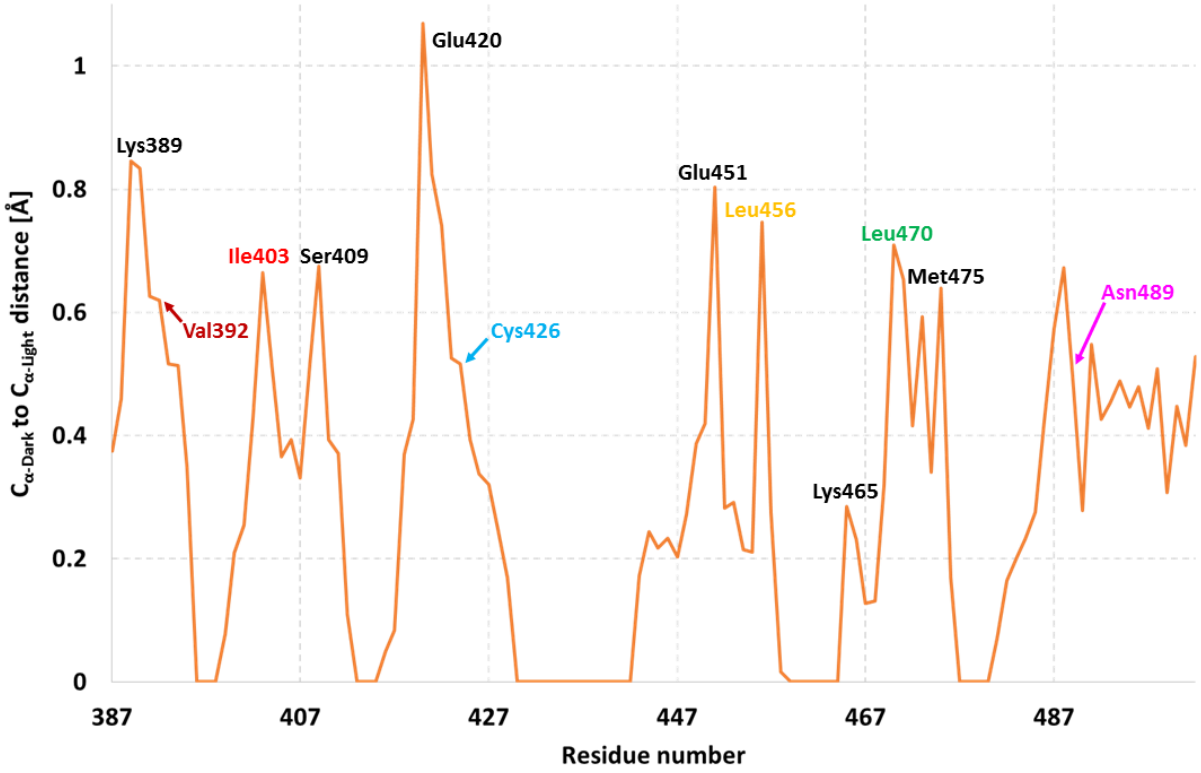
Supplementary Figure 5 | $F_{obs} - F_{calc}$ and $2F_{obs} - F_{calc}$ electron density maps for all key residues and the FMN at time point 66 (data set at 2.75 Å) before modelling of the light state, contoured at a 2.5 and 1.0 σ level, respectively, superimposed on the dark state model (left) and both dark and light state models (right).



Supplementary Figure 6 | Leu456 reorientation and displacement triggered by Phe470 side chain reorientation, which introduces a close van der Waals interaction between both side chains, as evidenced by the 3.1 Å distance between the two closest atoms.



Supplementary Figure 7 | C_α to C_α distance plot between the dark and the light state showing which region has moved by how much.



Supplementary material

Supplementary video 1 | Slowed-down build-up of the light state in a *At*Phot2LOV2 crystal upon LED irradiation coming from the top-right corner. The dark state absorbs until 490 nm and is thus fluorescent upon 470 nm excitation, while the light state absorbs significantly only until 440-450 nm, and is thus not fluorescent.

Supplementary video 2 | Time evolution from 0 to 4158 ms of the 2*F*o-*F*c electron density map (contoured at a 1.2 σ level) showing the covalent bond formation between Cys426 and the FMN, disordering of the Gln489 side chain and displacement of the Phe470 side chain. The speed of the movie has been halved for visualisation comfort (126 ms/frame). The dark state frame is paused for 1 s to clearly indicate the start of the movie.

Supplementary video 3 | Time evolution from 0 to 4158 ms of the 2*F*o-*F*c electron density map (contoured at a 0.7 σ level) showing the translation of Ile403 upon progressive disappearance of the B conformation of Cys426. The speed of the movie has been halved for visualisation comfort (126 ms/frame). The dark state frame is paused for 1 s to clearly indicate the start of the movie.

Chapter 7

CONCLUSION

7.1 Leads for improvement of the TR-SOX method

In the experiments reported in **Chapter 6**, we were particularly careful to limit the absorbed dose to 162 kGy during the course of the whole 4,158 ms data collection and we could infer from our results that neither global nor specific damage was affecting the structure of the light state. Consequently, one can imagine delivering 162 kGy or more being absorbed (the maximum dose is a parameter which needs to be investigated), for each time point. In a first approach, these 162 kGy could be delivered on a single 15-image of partial oscillation data sets.

Let us take time point #10 as an example (**Fig. 7.1**). With the actual protocol, it corresponds to images 150 to 165 over the 990 images of the full dataset, leading to a snapshot structure of the 630-693 ms point over the 0-4,158 ms timespan. One needs to introduce a precise delay between the turning on of the LED, and the opening of the X-ray shutter. For time point #10, this means starting LED illumination at t = 0 ms and trigger X-ray data collection at for 630-693 ms. With this method, we would not need to collect full datasets anymore, 15 images (7.5-degree oscillation) per crystal would be enough.

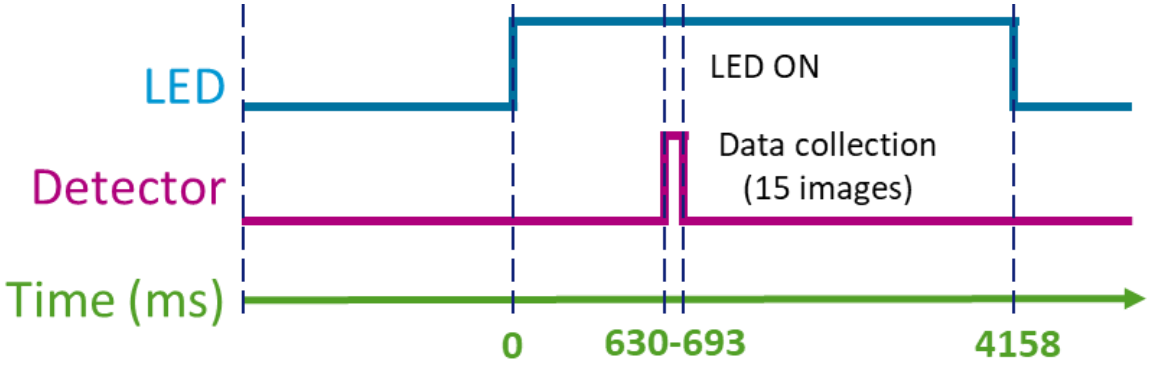


Figure 7.1: Timeline of the enhanced protocol. Example for time-point #10.

In a second step, one should probe what is the maximum dose that could tolerated to be absorbed in the 15-image wedge, which could be as high as several hundreds of kGy. The gain in diffraction power from using a higher flux density beam, thus a higher dose, may also be beneficial to the time resolution, as shorter image wedges could be successfully integrated and merged. For instance, going from 15 to 10-image wedges, would increase the time resolution from 63 to 42 ms.

However, these improvements would require the recording of data from many more crystals, which is not reasonable with our current approach (manual fishing, loop mounting, hutch interlocking). We thus need to find a compromise between our TR-SOX method and fixed-target SSX methods, using larger meshed loops or chips containing several (ten to hundreds) macro-crystals in order to reduce at minimum the preparation steps carried out manually. Another limitation is that if two samples are illuminated at the same time, once the first diffraction experiment is performed on the first one, one needs to wait for the full relaxation process to occur on the second one before performing the second diffraction experiment, and a solution to limit illumination to one single sample is desirable.

7.2 Fixed-target SSX – eliminating the relaxation waiting time using in plate diffraction

A way of getting rid of the waiting time for self-relaxation of the targeted reaction could be to resort to fixed-target diffraction experiment using crystallization plates. These have now been more frequently used thanks to the development of automated pipelines from crystallization screening to in-plate diffraction. In the case of a light-triggered reaction, the walls of each well containing a crystal should be totally light-tight to their neighbors, and with the LED or laser beam focused precisely on the well, one might be able to shoot each crystal in the plate without the need of waiting for the unavoidable time for the biological reaction to be fully reversed.

This approach of collecting data could also be supported by the development and setting up of a room temperature sample changer. This fixed-target SSX method could be challenging in both hardware and software development, especially in terms of synchronization and automation of every step of the experiment. One major issue that can be anticipated is the focusing of the illumination device into a well of a given size, even if we have plate with light-tight walls. This focusing issue can be solved by using a moveable cover for all wells at the top and the bottom of the plate, with only one hole above the well that is currently under investigation. This should ensure full independence between one well and the surrounding ones, thereby allowing a straightforward dataset collection without waiting for the incompressible self-relaxation time of a biological system such as LOV2. Of course, such an experiment also requires crystals that should be of a given size and single per well. This implies precise in-plate crystallization, preferably using automation for performing all steps of making

the crystallization solution and drops, and potentially involving automated seeding methods to achieve better control of the crystallization.

7.3 Use of microfluidics in time-resolved experiments

An important application of SSX methods describing how to compose time-resolved molecular movies may find its place in a receptor-ligand interaction study. It could be of prime importance for developing new ligands that can target specifically any intermediate state of a reaction instead of the few states – mostly initial and final states – that are usually visible using conventional X-ray crystallography. Knowing the sequence of events occurring when a ligand binds to its receptor can also allow the engineering of new receptors using, for example, site-directed mutagenesis, on the binding pocket area and or its surrounding residues.

To achieve this, there is the need of developing microfluidic devices that are suitable and highly efficient for the mixing of a protein crystal and its potential ligand, performed in a time-resolved manner, and preferably adaptable directly on synchrotron beamlines for immediate dataset collection after a relevant mixing time. Setting up such an experiment leads to a large number of experimental parameters that need to be controlled, as the time of diffusion of the ligand into the protein crystal, or how to choose the suitable strategy for data collection, using either oscillation or still image serial crystallography. One other major issue will be how to deliver the crystalline sample to the X-ray beam. There are several possibilities that might be investigated, as an assembly between a microfluidic device where the mixing occurs and an extrusion injector for liquid/viscous media sample delivery, or more ideally to obtain direct dataset collection on the microfluidic chips.

When such methods as direct mixing-and-diffraction become available in the next few years, it should be achievable to obtain true molecular movie of how an interaction between a target and its ligand occurs.

Annex 1

RESULTS – ARTICLE 4

Beyond the photosensory function of *At*Phot2LOV2: turning it into a ROS generator

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' (2019) *Sci. Rep.* **9**:2428, doi: 0.1038/s41598-019-38955-3

SCIENTIFIC REPORTS

OPEN

Tailing miniSOG: structural bases of the complex photophysics of a flavin-binding singlet oxygen photosensitizing protein

Received: 2 August 2018
Accepted: 11 January 2019
Published online: 20 February 2019

Joaquim Torra[1], Céline Lafaye[2], Luca Signor[2], Sylvain Aumonier[3], Cristina Flors[4,5], Xiaokun Shu[6,7], Santi Nonell[1], Guillaume Gotthard[3] & Antoine Royant[2,3]

miniSOG is the first flavin-binding protein that has been developed with the specific aim of serving as a genetically-encodable light-induced source of singlet oxygen ($^1O_2$). We have determined its 1.17 Å resolution structure, which has allowed us to investigate its mechanism of photosensitization using an integrated approach combining spectroscopic and structural methods. Our results provide a structural framework to explain the ability of miniSOG to produce $^1O_2$ as a competition between oxygen- and protein quenching of its triplet state. In addition, a third excited-state decay pathway has been identified that is pivotal for the performance of miniSOG as $^1O_2$ photosensitizer, namely the photo-induced transformation of flavin mononucleotide (FMN) into lumichrome, which increases the accessibility of oxygen to the flavin FMN chromophore and makes protein quenching less favourable. The combination of the two effects explains the increase in the $^1O_2$ quantum yield by one order of magnitude upon exposure to blue light. Besides, we have identified several surface electron-rich residues that are progressively photo-oxidized, further contributing to facilitate the production of $^1O_2$. Our results help reconcile the apparent poor level of $^1O_2$ generation by miniSOG and its excellent performance in correlative light and electron microscopy experiments.

miniSOG (for mini Singlet Oxygen Generator)[1] is a 106 amino acid flavin-binding protein that generates $^1O_2$ under exposure to blue light. It was originally developed by Shu and coworkers for correlative light and electron microscopy (CLEM) as it both fluoresces and catalyzes the photo-oxidation of diaminobenzidine (DAB), providing high-resolution images[1]. Novel applications are being actively developed since[2–5]. miniSOG was engineered from the LOV2 (Light, Oxygen and Voltage) domain of *Arabidopsis thaliana* phototropin 2[1]. Proteins based on LOV domains are blue-light photoreceptors that form a light-induced and reversible flavin-cysteine covalent adduct that consumes the energy of the excited state[6]. Replacement of the cysteine residue by an alanine or glycine avoids the formation of the covalent bond and leads to a fluorescent protein[7,8]. miniSOG contains six mutations as compared to its precursor, two of them involving residues surrounding the chromophore. Its cofactor FMN is ubiquitously found in nature[9,10] and generates $^1O_2$ with high quantum yield (Φ_Δ)[11], but also other reactive oxygen species (ROS)[12].

Close inspection of the photophysical and photosensitizing properties of miniSOG reveals a number of striking observations: (1) its Φ_Δ is much lower than that of FMN (0.03 vs. 0.51)[11,13,14]; (2) the lifetime (τ_T) of triplet miniSOG (^{3}miniSOG*) is much shorter than that of FMN in nitrogen-saturated solutions (33.6 µs[15] vs 200 µs[16]); (3) oxygen quenching is less efficient than for FMN ($\tau_T^{air} = 31.3$ µs[15] vs 3.1 µs in air-saturated solutions); (4) in

[1]Institut Químic de Sarrià, Universitat Ramon Llull, Via Augusta 390, Barcelona, 08017, Spain. [2]Univ. Grenoble Alpes, CNRS, CEA, IBS (Institut de Biologie Structurale), F-38000, Grenoble, France. [3]European Synchrotron Radiation Facility, F-38043, Grenoble, France. [4]Madrid Institute for Advanced Studies in Nanoscience (IMDEA Nanoscience), Ciudad Universitaria de Cantoblanco, C/Faraday 9, 28049, Madrid, Spain. [5]Nanobiotechnology Unit Associated to the National Center for Biotechnology (CNB-CSIC-IMDEA), Ciudad Universitaria de Cantoblanco, 28049, Madrid, Spain. [6]Department of Pharmaceutical Chemistry, University of California-San Francisco, San Francisco, California, 94158-9001, United States. [7]Cardiovascular Research Institute, University of California-San Francisco, San Francisco, California, 94158-9001, United States. Correspondence and requests for materials should be addressed to S.N. (email: santi.nonell@iqs.url.edu) or G.G. (email: guillaume.gotthard@esrf.fr) or A.R. (email: antoine.royant@ibs.fr)

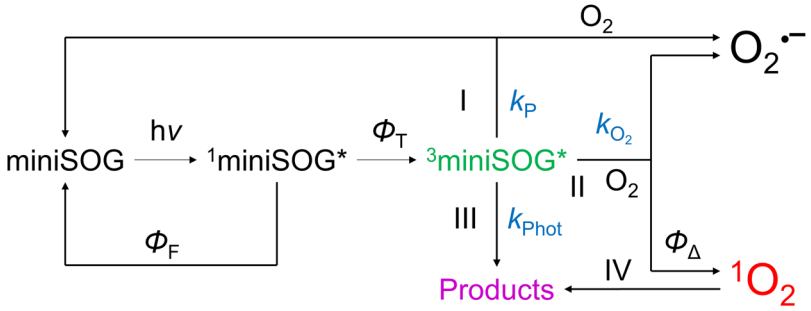


Figure 1. Model of oxygen photosensitization by miniSOG. Φ_F: fluorescence quantum yield, Φ_T: triplet state quantum yield, k_P: protein quenching rate constant, k_{Phot}: photoproduct formation rate constant, k_{O_2}: oxygen quenching rate constant, Φ_Δ: singlet oxygen quantum yield. I, II, III, and IV indicate the three deactivation and one oxidation pathways discussed in the main text.

addition to $^1O_2$ it also produces superoxide ($O_2^{\bullet-}$)[12,14]; (5) it undergoes a remarkable transformation upon exposure to light, whereby Φ_Δ increases 10-fold (to ~0.3) and τ_T^{air} shortens by 10-fold (to ~3 µs)[13,14]. The absence of a structure of miniSOG so far had prevented to rationalize these observations, which we have attempted here using a combined structural and photophysical approach.

Based on the extensive data present in the literature and the photophysical and structural results presented herein, a mechanism of excited-state deactivation of miniSOG can be proposed that involves three main pathways (Fig. 1). The shorter lifetime of ^{3}miniSOG* compared to ^{3}FMN* indicates that protein quenching is a major mechanism of triplet decay (pathway I). Its rate constant k_P is largely determined by electron transfer with nearby electron-rich residues[17]. Quenching of the singlet state can be safely ruled out since no shortening of the fluorescence lifetime or decrease in the fluorescence quantum yield are observed relative to free FMN. In the presence of oxygen, a second decay pathway (pathway II) is possible, namely oxygen quenching to produce $^1O_2$ (energy transfer) or $O_2^{\bullet-}$ (electron transfer), as observed for FMN in solution[12]. It is also possible to produce $O_2^{\bullet-}$ by reaction of oxygen with a radical anion formed during protein quenching in pathway I. Finally, miniSOG undergoes a photoinduced transformation (pathway III, rate constant k_{Phot}), for which we provide here a detailed description for the first time.

Results and Discussion

High resolution crystal structure of miniSOG. We have solved the structure of miniSOG at 1.17 Å resolution (Fig. 2a and Supporting Information), which shows an increase in rigidity of the environment of the chromophore compared to that in the LOV2 domain, the location of potential quenchers of the excited states of FMN, and the phosphoribityl tail of FMN lying in a tunnel bridging the bulk solvent and the chromophore encased in the core of the protein (Fig. 2b). The latter hinders oxygen access to the isoalloxazine ring. The presence halfway through the tunnel of a chloride ion, which can be a good mimic of molecular oxygen[18,19], suggests that oxygen diffusion can occur.

Deactivation mechanism of miniSOG triplet excited state (Pathways I and II). The values of the relevant rate constants for pathways I and II can be inferred from the ^{3}miniSOG* lifetime measurements. Comparison of the decay rate constant ($1/\tau_T$) of miniSOG and SOPP3, the miniSOG mutant with the longest τ_T reported so far (3.3 ms in nitrogen-saturated solutions)[17] allows us to estimate the rate constant for protein quenching ($k_P = k_T^{N2} - k_T^{N2}{}_{,SOPP3}$, Table 1). SOPP3 is a miniSOG variant, which encases the same chromophore FMN and, most importantly, lacks most of the electron-rich residues present in the vicinity of the flavin in miniSOG. Hence, protein quenching of the triplet chromophore in SOPP3 is essentially suppressed, which makes SOPP3 a convenient model for the study of the contribution of protein quenching in miniSOG. Likewise, the pseudo-first order rate constant for oxygen quenching ($k_{O_2} = k_T^{Air} - k_T^{N_2}$) can be estimated from τ_T data in air- and nitrogen-saturated solutions (Table 1).

Comparison of k_p and k_{O_2} in Table 1 reveals that protein quenching (pathway I) is the main triplet deactivation pathway, removing 93% of the triplets in air-saturated solutions $k_P/(k_P + k_{O_2})$. Oxygen only quenches 7% of the triplets, which limits Φ_Δ to $0.6 \times 0.07 = 0.042$ (Eq. 1), in excellent agreement with the experimental value. It can therefore be concluded that the modest Φ_Δ of miniSOG is due to an unfavorable combination of low k_{O_2} and high k_P values, as proposed previously[17].

$$\Phi_\Delta = \Phi_T \times \frac{k_{O_2}}{k_p + k_{Phot} + k_{O_2}} \approx \Phi_T \times \frac{k_{O_2}}{k_p + k_{O_2}} \quad (1)$$

Our structural results above suggest that the low value of k_{O_2} is due to the steric hindrance of the ribityl tail within the tunnel which provides oxygen access to the FMN. Regarding k_P, the miniSOG structure shows that six electron-rich residues are positioned within 8.2 to 10.2 Å from the isoalloxazine ring, namely Tyr30, Tyr73, Trp81, His85, Met89 and Tyr98, and are thus close enough to the chromophore to act as electron-transfer quenchers of ^{3}miniSOG*[20]. In addition, four hydrophilic residues, Glu44, Asp72, Asp82 and Glu103, form hydrogen bonds

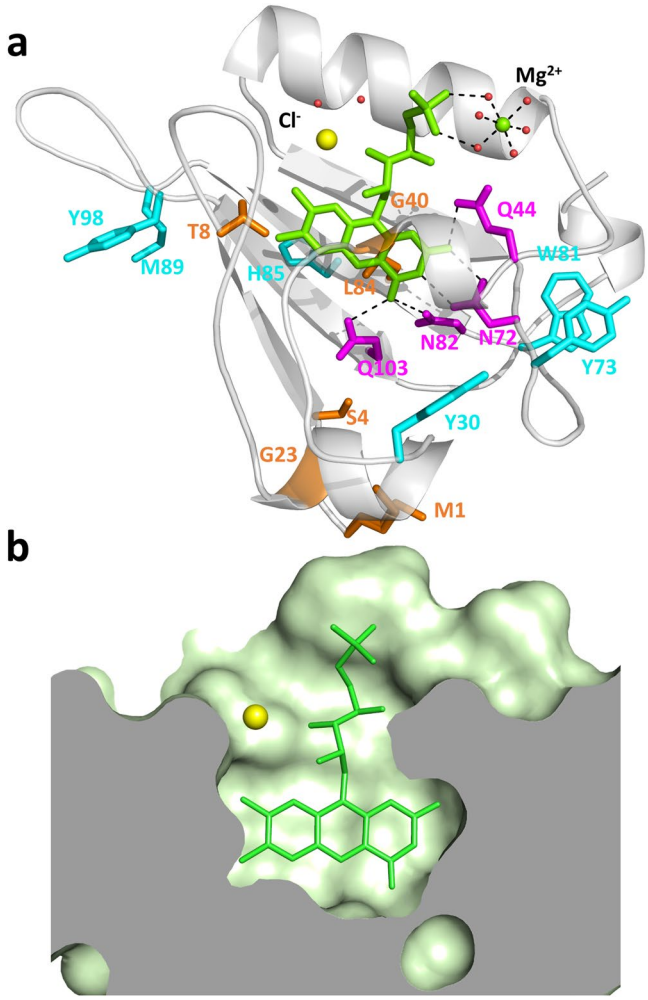


Figure 2. High-resolution crystallographic structure of miniSOG. (**a**) Secondary structure (white) represented with FMN (green), chloride (yellow), magnesium (green) and water molecules (red). Represented residues: mutations from LOV2 (orange), residues hydrogen-bonded to the FMN ring (magenta), and potential quenchers of ^{3}FMN* (cyan). (**b**) Topology of the FMN-binding site.

Parameter	in D_2O	in H_2O	Pathway	References
τ_S	5.0 ns	4.9–5.5 ns		13,15,43
Φ_F	0.43	0.37–0.44		1,14,15,43,44
Φ_T	0.6	0.6		15
Φ_Δ	0.03–0.04	0.03–0.05		13,14,23,31,44
k_T^{N2}	$2.41 \times 10^4\,s^{-1}$	$2.98 \times 10^4\,s^{-1}$		15
$\boldsymbol{k_P}$	$\mathbf{2.38 \times 10^4\,s^{-1}}$	$\mathbf{2.95 \times 10^4\,s^{-1}}$[b]	**I**	This work
k_T^{Air}	$2.59 \times 10^4\,s^{-1}$	$3.19 \times 10^4\,s^{-1}$		15
$\boldsymbol{k_{O_2}}$	$\mathbf{1.8 \times 10^3\,s^{-1}}$	$\mathbf{2.3 \times 10^3\,s^{-1}}$	**II**	This work
k_{Phot}	$\mathbf{6.0\,s^{-1}}$	—	**III**	This work

Table 1. Photophysical properties of miniSOG in D_2O- and H_2O-based phosphate buffer. D_2O was used to increase the singlet oxygen lifetime, thus boosting the reactions and processes in which singlet oxygen is involved and facilitating its detection[45]. [a]Assuming the same value of $k_T^{N2}{}_{,SOPP3}$ in H_2O and D_2O.

with FMN, and may thus enhance protein quenching and $O_2^{\bullet-}$ formation[21]. Replacing selectively these residues should lead to a lengthening of the triplet lifetime of miniSOG[22] and hence to a higher fraction of triplets being trapped by oxygen, thus to a higher Φ_Δ value. In fact, some of these positions have already been mutated in light of their capacity of direct electron transfer from the FMN: such miniSOG mutants show considerably longer τ_T values (e.g., 196 µs for miniSOG Q103L (SOPP)[23], 1.1 ms for miniSOG W81F (Supplementary Fig. S2), and 3.3 ms for SOPP3[17] in oxygen-free solutions) and larger Φ_Δ values (0.25, 0.33 and 0.6, respectively), in agreement with

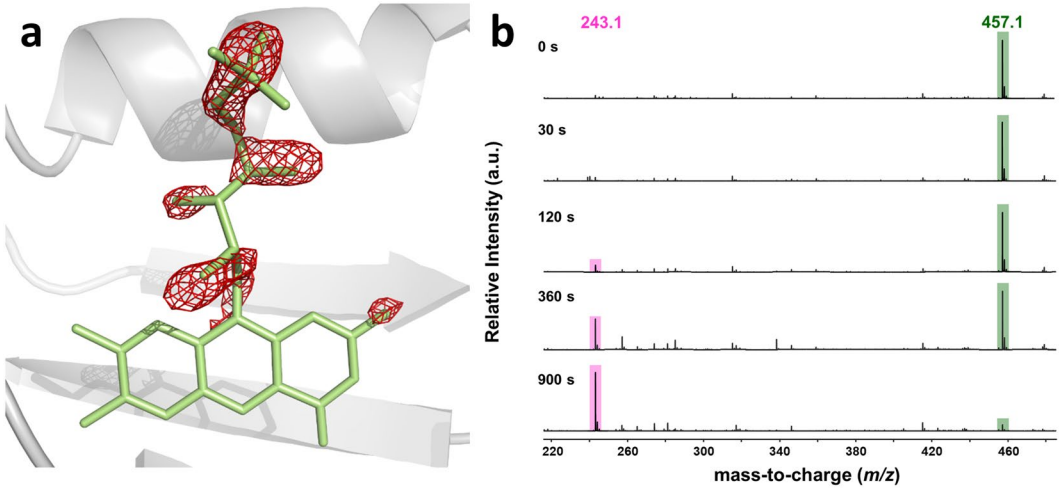


Figure 3. Blue-light induced structural changes on the chromophore of miniSOG. (**a**) 2.0 Å resolution difference Fourier map calculated between non-irradiated and irradiated parts of a miniSOG crystal contoured at a −3.0 σ level (magenta) superimposed on the FMN molecule (green). (**b**) ESI-TOF mass spectra acquired in the low mass range ($m/z < 500$) of miniSOG progressively irradiated with blue-light.

Eq. 1. It is worth noting also that miniSOG produces more $O_2^{\bullet-}$ than free FMN[12], which indicates that the radical anion pathway contributes to the production of $O_2^{\bullet-}$. Indeed, SOPP shows an 8-fold higher Φ_Δ value than miniSOG but only a 1.3 higher yield of $O_2^{\bullet-}$ [23]. Thus, removal of hydrophilic side chains in the vicinity of the chromophore should strongly reduce the relative formation of $O_2^{\bullet-}$ *vs.* $^1O_2$.

Consequences of blue-light irradiation of miniSOG on its FMN chromophore (Pathway III). In light of Eq. 1, the observed 10-fold decrease in τ_T and similar increase in Φ_Δ upon extended photolysis suggest severe changes in both k_P and k_{O_2}. Blue-light (440 nm) irradiation of a miniSOG crystal at 10 W·cm^{-2} led to a five-fold decrease of the fluorescence signal over a 30 min course (Supplementary Fig. S3) and was gentle enough to keep diffraction around 2.0 Å resolution while affecting a sufficient fraction of molecules so that structural alterations could be visualized in electron density maps. A difference Fourier map calculated from non-irradiated and irradiated parts of a crystal revealed the loss of electron density all along the ribityl tail of the FMN (Fig. 3a), strongly suggesting its cleavage. Besides, Electrospray ionisation time-of-flight (ESI-TOF) mass spectrometry performed on irradiated protein samples show (Fig. 3b) the progressive disappearance of the FMN peak at $m/z = 457.1$ in favor of a peak at $m/z = 243.1$.

To get further insights into the photoconversion, we performed additional photophysical investigations. Besides the already-known shortening of τ_T and increase in Φ_Δ, exposure of miniSOG samples to light induces photobleaching of the FMN chromophore and appearance of new absorption and fluorescence bands (Fig. 4a,b). The leaching out of FMN from miniSOG was routinely checked and could be safely ruled out. The quantum yield and rate constant of pathway III could be estimated (Table 1, Supplementary Fig. S1). Noteworthy, the Φ_Δ value increases when the photoconverted miniSOG is excited at 355 nm, but remains essentially constant when probed at 473 nm (Fig. 4c,d).

Phototransformation of FMN to lumichrome (LC) is consistent with all of the above observations: (1) LC is a photodegradation product of flavins in aqueous solutions[24]; (2) the observed mass loss upon irradiation matches the molar mass difference between FMN (456.3 Da) and LC (242.2 Da); (3) LC absorbs and fluoresces at shorter wavelengths than FMN, (Fig. 5); (4) LC lacks the phosphoribityl tail of FMN, which facilitates the access of molecular oxygen to the isoalloxazine ring, resulting in the increase of k_{O_2} and the decrease of τ_T; (5) LC is a worse electron acceptor than FMN, hence protein quenching is less favored. The $\Delta_r G°$ value for quenching of ^{3}riboflavin* by tryptophan is −86.5 kJ·mol^{-1} (riboflavin is analogous to FMN except for the phosphate group) while is more positive for ^{3}LC*, −67.2 kJ·mol^{-1} [25]; (6) finally, LC is also an excellent $^1O_2$ photosensitizer[25–27], hence the combination of a higher k_{O_2} and a lower k_P yield a higher Φ_Δ value (Eq. 1) when excited at 355 nm but not at 473 nm, where LC barely absorbs.

Consequences of blue-light irradiation of miniSOG on its amino acid residues (Pathway IV). We investigated if our structural data could also support a decrease in k_P. Indeed, the $2F_{obs} - F_{calc}$ electron density map of blue-light irradiated miniSOG reveals the unambiguous oxidation of three surface residues during irradiation (Fig. 6a, Supplementary Fig. S4). Tyr73 has been partially converted to a γ-peroxotyrosine. The loss of electron density on Trp81 is compatible with the formation of *N*-formylkynurenine (NFK), a well-known tryptophan oxidation product[28,29]. Finally, His85 can be modeled by either a singly, or a doubly oxidized histidine, namely 2-oxo-histidine and 2,4-dioxo-histidine. Mass spectrometry analysis of blue-light irradiated miniSOG samples reveals sequential additions of +16 mass units to the native protein mass of 13882.0 Da, consistent with increasing oxidation steps of the protein (Fig. 6b). All three structural modifications account for six of the eight additional

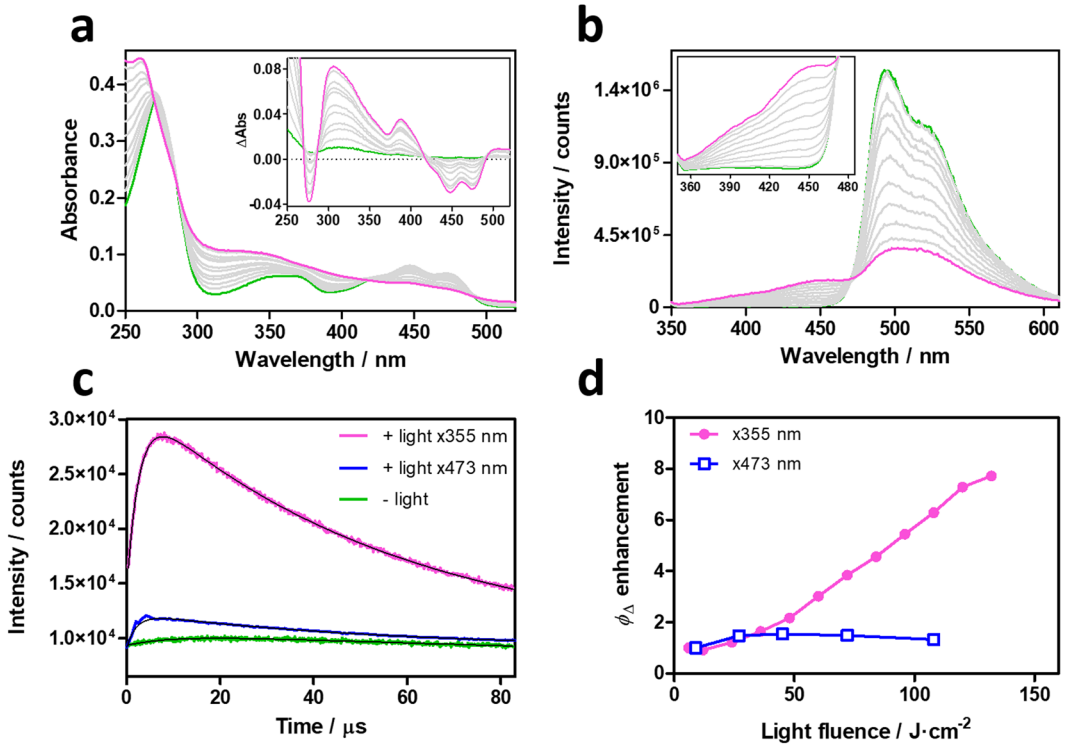


Figure 4. Spectroscopic characterization of photoconverted miniSOG. Evolution of miniSOG's absorbance (**a**) and fluorescence (**b**) upon laser irradiation at 355 nm. Insets show the difference absorption spectra before and after irradiation, and a zoomed-in image of the new fluorescence bands. (**c**) Time-resolved NIR $^1O_2$ phosphorescence decays of native (green) and photoconverted miniSOG excited at 473 nm (blue) or 355 nm (magenta). (**d**) Observed Φ_Δ enhancement at 473 nm (blue) and 355 nm (magenta).

oxygen atoms evidenced in the mass spectrometry analysis. The two non-assigned additions could correspond to oxidation of Tyr30, Met89 or Tyr98, although we did not observe unambiguous oxidation of these residues. Oxidation of Tyr73, His85 and Trp81 eliminates potential quenchers of ^{3}miniSOG*, thereby decreasing the value of k_p. According to Eq. 1, this should contribute to an increase in Φ_Δ. However, since protein oxidation (pathway IV) occurs simultaneously to FMN → LC transformation, which also increases Φ_Δ, it is not possible to ascertain the individual contribution of both effects.

Finally, oxidation of tryptophan into NFK could contribute to the increased Φ_Δ value observed at 355 nm since NFK is a potent singlet oxygen photosensitizer ($\Phi_\Delta = 0.17$)[30]. However, the W81F mutant shows a doubled Φ_Δ (=0.33) already before photolysis on account of its lower k_p value (Eq. 1), indicating that the potential benefits of producing NFK as secondary photosensitizer are of minor value as compared to the effect of eliminating a protein quencher.

Conclusion

We have performed an extensive structural characterization of miniSOG in the dark and its photoproduct formed in the presence of molecular oxygen, which led us to explain in structural terms the details of its complex photophysical behavior. miniSOG is initially moderately efficient towards $^1O_2$ generation because of a combination of limited oxygen accessibility and ^{3}FMN quenching by electron-rich side chains. Prolonged irradiation to blue light leads to several structural alterations of miniSOG, which include photodegradation of FMN into LC and oxidation of the quenching side chains. All this results in an increase of Φ_Δ when photoconverted miniSOG is excited at the wavelengths where the formed LC absorbs. Formation of LC liberates the access of molecular oxygen to the alloxazine ring and reduces protein quenching of the triplet state, while oxidized electron-rich side chains cannot quench the triplet state of the chromophore. The competition between oxygen quenching and protein quenching of flavin triplet state seems to be a general feature of flavin-binding proteins[31], hence our results will be useful to guide the evolution of such a protein towards retaining or gaining a specific function. Finally, our results explain the apparent discrepancy between the poor level of singlet oxygen generation by miniSOG, which had been consistently measured at low light fluences, and its efficiency in CLEM experiments, in which the singlet oxygen generation capability of miniSOG is exploited over its whole lifetime.

Methods

Chemical compounds. Riboflavin-5′-monophosphate sodium salt hydrate (FMN) (Chemochroma), Lumichrome (Santa Cruz Biotechnology), tris(hydroxymethyl)aminomethane (Merck), sodium chloride (Panreac Applichem), imidazole (Merck), L-arabinose (Sigma Aldrich) and ampicillin (Sigma Aldrich) were used as received. Phosphate-buffered saline (PBS) or deuterated dPBS solutions were prepared by dissolving the required amount of a PBS tablet (Sigma Aldrich) in milliQ water or deuterium oxide (Sigma-Aldrich).

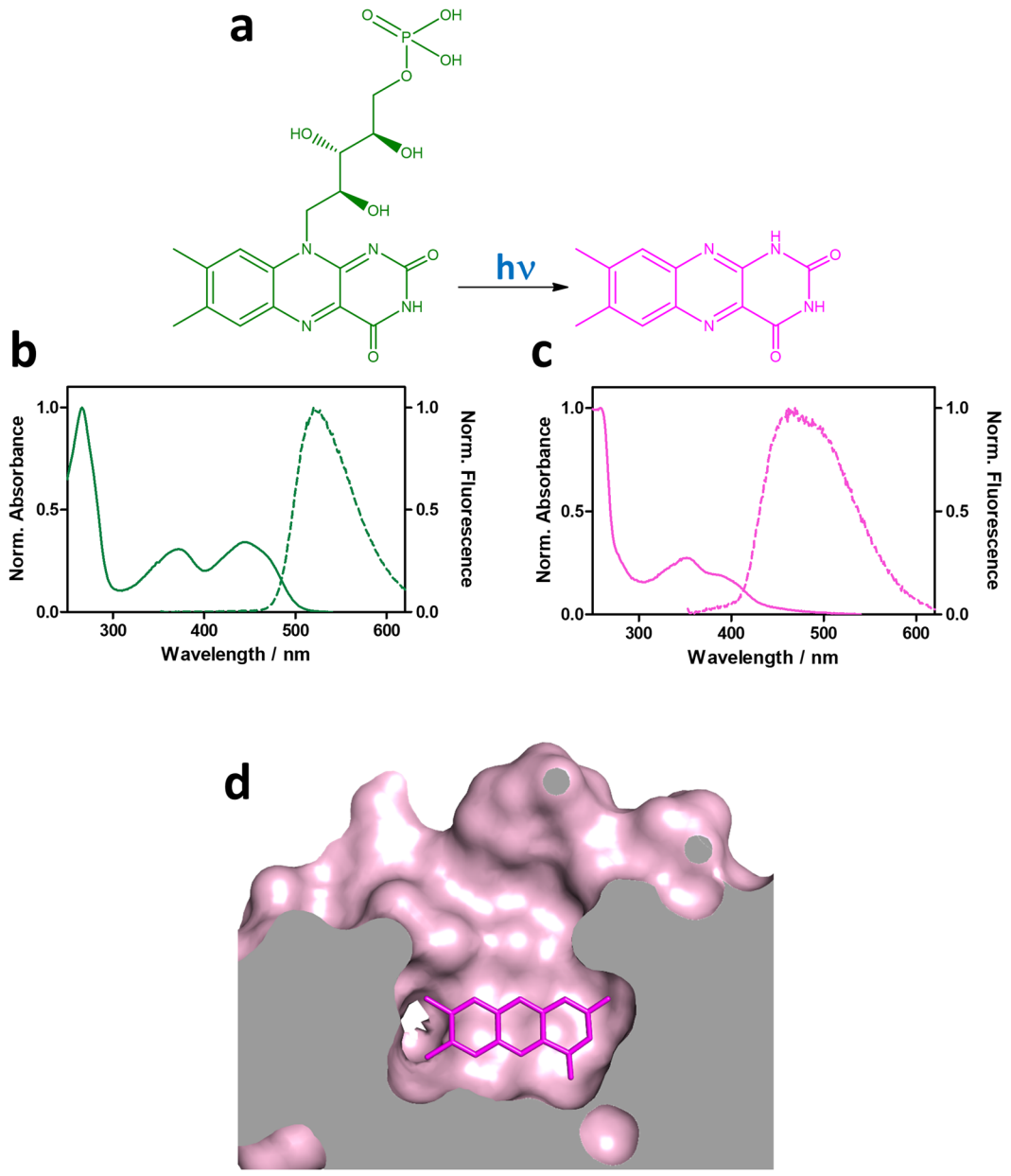


Figure 5. FMN is photoconverted to lumichrome. (**a**) Photoconversion of FMN (green) to LC (magenta). Normalized absorption and fluorescence spectra of (**b**) FMN and (**c**) LC. (**d**) Topology of blue-light irradiated miniSOG showing the increased access to the alloxazine ring.

Expression and purification. Genes coding for a C-terminal 6xHis-tagged recombinant miniSOG and miniSOG W81F were inserted in a pBad expression vector and over-expressed in *Escherichia coli* CodonPlus (DE3) RIL Cells (Stratagene) or in TOP10 cells (Invitrogen). Bacterial cells were grown in LB broth medium containing 1 mM Ampicillin. At an OD_{600} of approximately 0.6, expression of recombinant protein was induced by the addition of L-arabinose and cells were grown for an additional 24 h at 25 °C. Cells were pelleted by centrifugation (4000 g, 4 °C, 30 min), re-suspended in *buffer A* (20 mM Tris-Hcl pH 8.0, 500 mM NaCl), complemented with complete protease inhibitors-EDTA (Roche) and disrupted using a micro-fluidizer. The soluble fraction was recovered by centrifugation (40,000 g, 4 °C, 30 min), and loaded on a 1 mL Ni-NTA superflow column (Qiagen) pre-equilibrated with *buffer A*. The His-tagged protein was eluted with 150 mM imidazole in *buffer A*. Fractions containing purified proteins were pooled and concentrated to a volume of 0.5 mL using Centricon devices (Amicon 10 kDa cut-off) and loaded onto a size-exclusion chromatography column (Hiload Superdex75 10/300, GE Healthcare) for the final step of the purification procedure. The column was equilibrated with 20 mM Tris-HCl pH 8.0 and the pooled peak fractions were concentrated to 4 $mg{\cdot}mL^{-1}$. Protein expression and purification was always performed in the dark or under red light. The purity of the protein solutions was confirmed by SDS-PAGE. The final concentration was determined by UV-vis absorption spectroscopy using a molar absorption coefficient of 14 $mM^{-1}{\cdot}cm^{-1}$ at 448 nm.

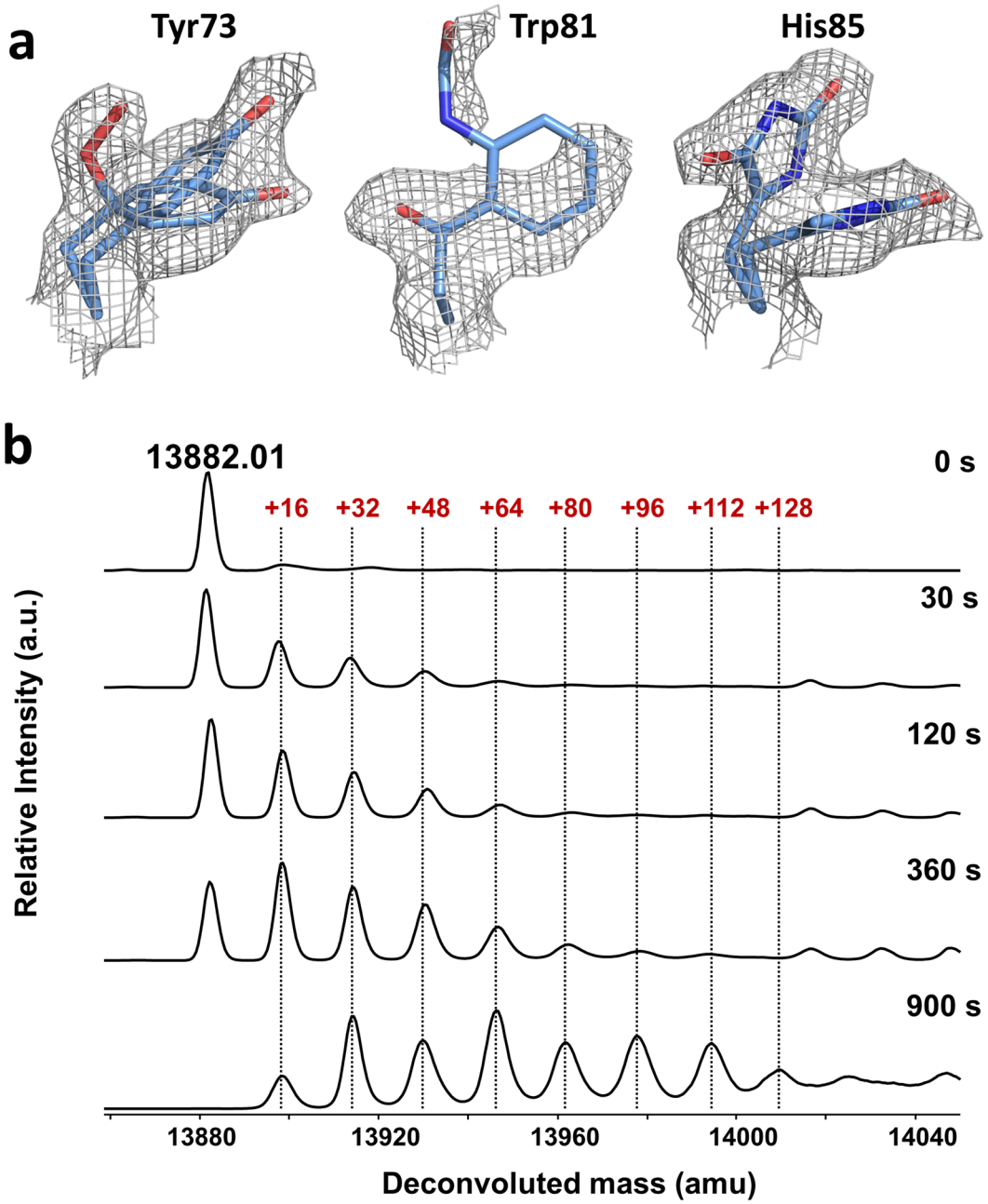


Figure 6. Identification of protein residue oxidation in blue-light irradiated miniSOG. (**a**) $2F_{obs} - F_{calc}$ electron density maps contoured at a 1.0, 0.6 and 1.0 σ level superimposed on the refined model of residues Tyr73, Trp81, and His85, respectively. (**b**) Deconvoluted ESI-TOF mass spectra of miniSOG progressively irradiated with blue-light. The peak at 13882.0 Da corresponds to the native protein.

Spectroscopic measurements. All spectroscopic measurements were performed using quartz cuvettes (Hellma) under magnetic stirring and at room temperature. Absorption spectra were recorded on a double beam Cary 6000i spectrophotometer (Varian). Fluorescence spectra were measured on Fluoromax-4 spectrofluorometer (Horiba). Time-resolved near-infrared (NIR) phosphorescence signals at 1275 nm were measured using a customized PicoQuant Fluotime 200 lifetime system. Briefly, an AO-Z-473 solid state AOM Q-switched laser (Changchun New Industries Optoelectronics Technology Co., China) was used for excitation at 473 nm, working at 1.0 kHz repetition rate at 473 nm. The average power that reached the sample was conveniently modulated by neutral density filters. For excitation at 355 nm, the frequency-tripled output of a diode-pumped pulsed Nd:YAG laser (FTSS355-Q, Crystal Laser, Berlin, Germany) was used, working at 1 kHz repetition (0.5 mW, or 5 mW, 1 ns pulse width). An uncoated SKG-5 filter (CVI Laser Corporation, Albuquerque, U.S.A.) was placed at the exit port of the laser to remove any NIR component. The luminescence exiting from the sample was filtered by a 1100 nm long-pass filter (Edmund Optics, York, U.K.) and a narrow bandpass filter at 1275 nm (bk-1270-70-B, bk Interfernzoptik, Germany) to remove any scattered laser radiation and isolate the $^1O_2$ emission. A

TE-cooled near-IR sensitive photo multiplier tube assembly (H9170-45, Hamamatsu Photonics Hamamatsu City, Japan) in combination with a multichannel scaler (NanoHarp 250, PicoQuant Gmbh, Germany) was used as photon-counting detector. The time-resolved $^1O_2$ emission decays were analyzed by fitting Eq. 2 [32] to the data using GraphPad Prism 5.

$$S_{(t)} = S_{(0)}\frac{\tau_\Delta}{\tau_\Delta - \tau_T}\left(e^{\frac{-t}{\tau_\Delta}} - e^{\frac{-t}{\tau_T}}\right) \quad (2)$$

τ_T and τ_Δ are the lifetimes of the photosensitizer triplet state and of $^1O_2$, respectively, and $S_{(0)}$ is a quantity proportional to Φ_Δ. Φ_Δ values were determined by comparing the $S_{(0)}$ values of optically-matched solutions of the corresponding flavoprotein and FMN at 473 nm (Eq. 3)[32].

$$\phi_{\Delta,protein} = \frac{S_{(0)protein}}{S_{(0)FMN}}\phi_{\Delta,FMN} \quad (3)$$

FMN was taken as reference photosensitizer with $\Phi_\Delta = 0.51$ in PBS[11] and 0.57 in dPBS[33].

Transient absorption spectra were monitored by nanosecond laser flash photolysis using a Q-switched Nd-YAG laser (Surelite I-10, Continuum) operating at the 3rd harmonic. The luminescence exiting from the sample was filtered by a 610 nm long-pass filter (CVI Laser Corporation, NM, USA). Changes in the sample absorbance were detected at 715 nm using a Hamamatsu R928 photomultiplier to monitor the intensity variations of an analysis beam produced by a 75 W short arc Xe lamp (USHIO) and spectral discrimination was obtained using a PTI 101 monochromator. The signal was fed to a Lecroy Wavesurfer 454 oscilloscope for digitizing and averaging (typically 10 shots) and finally transferred to a PC for data storage and analysis. The system was controlled using the in-house-developed LKS software (LabView, National Instruments).

Determination of k_{Phot}. The rate constant for the photoproduct formation has been determined measuring the progressive loss of miniSOG in solution as a function of the absorbed light dose at 473 nm using Eq. 4. The slope of the resulting plot yielded the photobleaching quantum yield Φ_{Phot}, from which the rate constant for photobleaching was calculated as:

$$k_{Phot} = \frac{\Phi_{Phot}}{\tau_T^{Air}} \quad (4)$$

X-ray crystallography. *Crystallization procedures.* miniSOG was concentrated to 4 mg·mL^{-1}. The crystallization condition consisted of 100 mM Tris-HCl pH 8.0, 20 mM $MgCl_2$, 28% PEG 4000, 0 or 15 mM $CoCl_2$ at 20 °C. Crystals appeared and grew to final size after 1–2 days.

Data collection and processing. X-ray data were collected on beamlines ID23-1[34] and ID29[35] of the ESRF and were indexed, integrated, merged and scaled using the *XDS* software package[36]. Molecular replacement was carried out using the model structure of LOV2 (PDB ID: 4eep) with the program *Phaser MR*[37]. Structure refinement was performed using *Refmac5*[38] and manual improvement of the model with *Coot*[39]. The native structure of miniSOG was used as a starting model for model building of bleached-miniSOG. Data collection and refinement statistics are presented in Supplementary Table S1. Structure analysis and representation were performed with *Pymol*[40].

Preparation of photobleached miniSOG samples. *miniSOG crystals.* A single miniSOG crystal was soaked in a cryoprotectant solution containing 20% of glycerol then harvested with a nylon loop. The crystal was exposed to 440 nm laser (10 W·cm^{-2}) on the ID29S-Cryobench setup[41] at room temperature using a HC1 humidity control device[42]. Spectra were recorded at a 1 Hz rate. After 30 min of total exposure, the crystal was flashcooled in liquid nitrogen.

miniSOG solutions. Fresh miniSOG or miniSOG W81F solutions in air-saturated deuterated PBS were illuminated at 355 nm (~5 mW·cm^{-2}) or 473 nm (~15 mW·cm^{-2}) for elapsed irradiation times. Absorption and fluorescence spectra as well as time-resolved $^1O_2$ phosphorescence decays were recorded at different time intervals of cumulative irradiation.

Liquid chromatography-mass spectrometry (LC-MS). Liquid Chromatography Electrospray Ionization Mass Spectrometry (LC/ESI-MS) was carried out on a 6210 LC/ESI-TOF mass spectrometer interfaced with a binary HPLC pump system (Agilent Technologies). The mass spectrometer was calibrated in the positive ion mode with ESI-L (low concentration tuning mix, Agilent Technologies) before each series of measurements, the calibration providing mass accuracy <1 ppm in the 100–3200 *m/z* range. All solvents used were HPLC grade: water and acetonitrile (LC-MS Chromasolv, Sigma-Aldrich); formic acid was from Acros Organics (puriss., p.a.).

Data acquisition was carried out in the positive ion mode with spectra in the profile mode and mass spectra were recorded in the 130–2000 *m/z* range. The mass spectrometer was operated with the following experimental settings: ESI source temperature was set at 325 °C; nitrogen was used as drying gas (5 L/min) and as nebulizer gas (30 psi); the capillary needle voltage was set at 3500 V. Fragmentor value was of 250 V and skimmer of 65 V. The instrument was operated in the 2 GHz (extended dynamic range) mode and spectra acquisition rate was of 1 spectrum/s.

Before analysis, miniSOG samples were diluted to a final concentration of 10 μM in acetonitrile/water/formic acid (50:50:0.1, *v*/*v*/*v*) and infused directly in the mass spectrometer by a syringe pump at a flow rate of 10 μl/min. A blank run was carried out infusing only protein buffer diluted at the same ratio as the protein sample in the same solvent system.

The MS data were acquired and processed with the MassHunter workstation software (Data acquisition v.B.04.00, Qualitative analysis with Bioconfirm v.B.07.00, Agilent Technologies).

Protein Data Bank accession codes. The structures of non-irradiated and blue-light irradiated miniSOG have been deposited in the Protein Data Bank under entry codes 6GPU and 6GPV, respectively.

References

1. Shu, X. *et al*. A genetically encoded tag for correlated light and electron microscopy of intact cells, tissues, and organisms. *PLoS Biol.* **9**, e1001041 (2011).
2. Ryumina, A. P. *et al*. Flavoprotein miniSOG as a genetically encoded photosensitizer for cancer cells. *Biochim. Biophys. Acta* **1830**, 5059–67 (2013).
3. Buckley, A. M., Petersen, J., Roe, A. J., Douce, G. R. & Christie, J. M. LOV-based reporters for fluorescence imaging. *Curr. Opin. Chem. Biol.* **27**, 39–45 (2015).
4. Souslova, E. A., Mironova, K. E. & Deyev, S. M. Applications of genetically encoded photosensitizer miniSOG: From correlative light electron microscopy to immunophotosensitizing. *J. Biophotonics* **10**, 338–352 (2017).
5. Makhijani, K. *et al*. Precision optogenetic tool for selective single- and multiple-cell ablation in a live animal model system. *Cell Chem. Biol.* **24**, 110–119 (2017).
6. Salomon, M., Christie, J. M., Knieb, E., Lempert, U. & Briggs, W. R. Photochemical and mutational analysis of the FMN-binding domains of the plant blue light receptor, phototropin. *Biochemistry* **39**, 9401–9410 (2000).
7. Swartz, T. E. *et al*. The photocycle of a flavin-binding domain of the blue light photoreceptor phototropin. *J. Biol. Chem.* **276**, 36493–36500 (2001).
8. Drepper, T. *et al*. Reporter proteins for *in vivo* fluorescence without oxygen. *Nat. Biotechnol.* **25**, 443–445 (2007).
9. Massey, V. The chemical and biological versatility of riboflavin. *Biochem. Soc. Trans.* **28**, 283–296 (2000).
10. Losi, A. & Gärtner, W. The evolution of flavin-binding photoreceptors: an ancient chromophore serving trendy blue-light sensors. *Annu. Rev. Plant Biol.* **63**, 49–72 (2012).
11. Baier, J. *et al*. Singlet oxygen generation by UVA light exposure of endogenous photosensitizers. *Biophys. J.* **91**, 1452–1459 (2006).
12. Barnett, M. E., Baran, T. M., Foster, T. H. & Wojtovich, A. P. Quantification of light-induced miniSOG superoxide production using the selective marker, 2-hydroxyethidium. *Free Radic. Biol. Med.* **116**, 134–140 (2018).
13. Ruiz-González, R. *et al*. Singlet oxygen generation by the genetically encoded tag miniSOG. *J. Am. Chem. Soc.* **135**, 9564–9567 (2013).
14. Pimenta, F. M., Jensen, R. L., Breitenbach, T., Etzerodt, M. & Ogilby, P. R. Oxygen-dependent photochemistry and photophysics of "miniSOG," a protein-encased flavin. *Photochem. Photobiol.* **89**, 1116–1126 (2013).
15. Westberg, M., Bregnhøj, M., Etzerodt, M. & Ogilby, P. R. Temperature sensitive singlet oxygen photosensitization by LOV-derived fluorescent flavoproteins. *J. Phys. Chem. B* **121**, 2561–2574 (2017).
16. Fritz, B. J., Matsui, K., Kasai, S. & Yoshimura, A. Triplet lifetimes of some flavins. *Photochem. Photobiol.* **45**, 539–541 (1987).
17. Westberg, M., Bregnhøj, M., Etzerodt, M. & Ogilby, P. R. No photon wasted: an efficient and selective singlet oxygen photosensitizing protein. *J. Phys. Chem. B* **121**, 9366–9371 (2017).
18. Roeser, D., Schmidt, B., Preusser-Kunze, A. & Rudolph, M. G. Probing the oxygen-binding site of the human formylglycine-generating enzyme using halide ions. *Acta Crystallogr. D Biol. Crystallogr.* **63**, 621–627 (2007).
19. Colloc'h, N. *et al*. Oxygen pressurized X-ray crystallography: probing the dioxygen binding site in cofactorless urate oxidase and implications for its catalytic mechanism. *Biophys. J.* **95**, 2415–22 (2008).
20. Tanaka, F. *et al*. Donor−acceptor distance-dependence of photoinduced electron-transfer rate in flavoproteins. *J. Phys. Chem. B* **111**, 5694–5699 (2007).
21. Yagi, K., Ohishi, N., Nishimoto, K., Choi, J. D. & Song, P.-S. Effect of hydrogen bonding on electronic spectra and reactivity of flavins. *Biochemistry* **19**, 1553–1557 (1980).
22. Raffelberg, S., Mansurova, M., Gärtner, W. & Losi, A. Modulation of the photocycle of a LOV domain photoreceptor by the hydrogen-bonding network. *J. Am. Chem. Soc.* **133**, 5346–5356 (2011).
23. Westberg, M., Holmegaard, L., Pimenta, F. M., Etzerodt, M. & Ogilby, P. R. Rational design of an efficient, genetically-encodable, protein-encased singlet oxygen photosensitizer. *J. Am. Chem. Soc.* **137**, 1632–1642 (2015).
24. Ahmad, I. & Vaid, F. H. M. Photochemistry of flavins in aqueous and organic solvents. In *Flavins Photochemistry and Photobiology* (eds Silva, E. & Edwards, A. M.) 13–40 (Cambridge: Royal Society of Chemistry, 2006).
25. Remucal, C. K. & McNeill, K. Photosensitized amino acid degradation in the presence of riboflavin and its derivatives. *Environ. Sci. Technol.* **45**, 5230–5237 (2011).
26. Sikorska, E. *et al*. Spectroscopy and photophysics of lumiflavins and lumichromes. *J. Phys. Chem. A* **108**, 1501–1508 (2004).
27. Valerón Bergh, V. J., Bruzell, E., Hegge, A. B. & Tønnesen, H. H. Influence of formulation on photoinactivation of bacteria by lumichrome. *Pharmazie* **70**, 574–580 (2015).
28. Walrant, P. & Santus, R. N-formyl-kynurenine, a tryptophan photooxidation product, as a photodynamic sensitizer. *Photochem. Photobiol.* **19**, 411–417 (1974).
29. Fukunaga, Y., Katsuragi, Y., Izumi, T. & Sakiyama, F. Fluorescence characteristics of kynurenine and N'-formylkynurenine. Their use as reporters of the environment of tryptophan 62 in hen egg-white lysozyme. *J. Biochem.* **92**, 129–141 (1982).
30. Krishna, C. M. *et al*. A Study of the photodynamic efficiencies of some eye lens constituents. *Photochem. Photobiol.* **54**, 51–58 (1991).
31. Endres, S. *et al*. An optogenetic toolbox of LOV-based photosensitizers for light-driven killing of bacteria. *Sci. Rep.* **8**, 1–14 (2018).
32. Nonell, S. & Braslavsky, S. E. Time-resolved singlet oxygen detection. *Methods Enzymol.* **319**, 37–49 (2000).
33. Rodríguez-Pulido, A. *et al*. Assessing the potential of photosensitizing flavoproteins as tags for correlative microscopy. *Chem. Commun.* **52**, 8405–8408 (2016).
34. Nurizzo, D. *et al*. The ID23-1 structural biology beamline at the ESRF. *J. Synchrotron Radiat.* **13**, 227–238 (2006).
35. de Sanctis, D. *et al*. ID29: a high-intensity highly automated ESRF beamline for macromolecular crystallography experiments exploiting anomalous scattering. *J. Synchrotron Radiat.* **19**, 455–461 (2012).
36. Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–32 (2010).
37. McCoy, A. J. *et al*. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
38. Vagin, A. A. *et al*. REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2184–95 (2004).
39. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).

40. http://pymol.org/. The PyMOL Molecular Graphics System, Version 1.8.7 Schrödinger, LLC.
41. von Stetten, D. *et al*. In crystallo optical spectroscopy (icOS) as a complementary tool on the macromolecular crystallography beamlines of the ESRF. *Acta Crystallogr. D Biol. Crystallogr.* **71**, 15–26 (2015).
42. Sanchez-Weatherby, J. *et al*. Improving diffraction by humidity control: A novel device compatible with X-ray beamlines. *Acta Crystallogr. D Biol. Crystallogr.* **65**, 1237–1246 (2009).
43. Wingen, M. *et al*. The photophysics of LOV-based fluorescent proteins – new tools for cell biology. *Photochem. Photobiol. Sci.* **13**, 875–883 (2014).
44. Rodríguez-Pulido, A. *et al*. Fluorescent flavoprotein heterodimers: combining photostability with singlet oxygen generation. *ChemPhotoChem* **2**, 571–574 (2018).
45. Foote, C. S. & Clennan, E. L. Properties and reactions of singlet dioxygen. In *Active Oxygen in Chemistry* (eds Foote, C.S., Valentine, J., Greenberg, A. & Liebman, J.F.) 105-140 (Springer Netherlands, 1995).

Acknowledgements

We thank Rubén Ruiz-González, Marjolaine Noirclerc-Savoye and David von Stetten for their contribution at an early stage of the project. The ESRF is acknowledged for access to beamlines and facilities for molecular biology *via* its in-house research program. AR acknowledges funding from the French Agence Nationale de la Recherche (project SOxygen, ANR-11-JSV5-0009 and project CrystalBall, ANR-14-CE06-0005-02), from the Spanish Ministerio de Economía y Competitividad (CTQ2016-78454-C2-1-R, MAT2015-66605-P and SEV-2016-0686) and the Fundació la Marató de TV3 (grant No. 20133133). This work used the platforms of the Grenoble Instruct-ERIC Center (ISBG: UMS 3518 CNRS-CEA-UGA-EMBL) with support from FRISBI (ANR-10-INBS-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB).

Author Contributions

J.T. performed and analyzed the spectroscopic and photophysical investigations. C.L., S.A., G.G. and A.R. performed and analyzed the structural work. L.S. performed and analyzed the mass spectrometry experiments. C.F. and X.S. participated to the discussion of the data. S.N., G.G. and A.R. designed the study, discussed the data and wrote the manuscript with J.T. and C.F.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-019-38955-3.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary information to "Tailing miniSOG: structural bases of the complex photophysics of a flavin-binding singlet oxygen photosensitizing protein"

Supplementary results

Analysis of the 1.17 Å structure of miniSOG

The structure of miniSOG was solved with the molecular replacement method using a previously determined structure of the LOV2 domain of *A. thaliana* phototropin 2 (PDB 4eep).[1] The structures are very close from each other (r.m.s. deviation on Cα atoms of 0.39 Å) and we have identified the two mutations C40/426G (miniSOG/phototropin 2 numbering) and F84/470L to affect significantly the FMN-binding cavity. C40G is the photoreaction-hindering mutation, and the removal of a side chain at position 40, added to the smaller size of the side chain at position 84, induce the displacement of three strands towards the isoalloxazine ring, providing the chromophore with a smaller binding pocket, thus with a more rigid environment. A chloride ion is located next to the ribityl chain positioned halfway through the tunnel. The presence of magnesium and chloride is favored by the composition of the crystallization mother liquor. We note that the chloride ion can be a good mimic of molecular oxygen,[2,3] which suggests that this location may serve as a transient affinity site for oxygen when diffusing from the solvent to the chromophore via the tunnel. Finally, the chromophore is stabilized by the protein through hydrogen bonds on one half of the isoalloxazine ring, and van der Waals interactions on the other half.

Mass spectrometry analysis of irradiated miniSOG

In order to identify the precise location of the cleavage, we performed ESI TOF mass spectrometry analysis on protein samples that had been exposed to increasing blue-light doses, acquiring the spectral data in the low mass range region (m/z < 500). We could observe the progressive disappearance of the FMN signal, corresponding to the $[M+H]^+$ ion at m/z = 457.1 in favor of a $[M+H]^+$ ion at m/z = 243.1, which corroborates the interpretation that lumichrome (LC, 242.2 Da) is the main photoproduct.

Supplementary tables and figures

Table S1: Data collection and refinement statistics

Data collection	**miniSOG**	**Blue-light irradiated miniSOG**
PDB code	6GPU	6GPV
ESRF Beamline	ESRF ID23-1	ESRF ID29
Wavelength (Å)	0.9763	0.9763
Space group	P 4_3 2_1 2	
Cell dimensions		
a, b, c (Å)	40.02, 40.02, 134.29	40.49 40.49 133.74
Resolution range (Å)	80 – 1.17 (1.20 – 1.17)	80 – 2.0 (2.1 – 2.0)
Wilson B-factor (Å^2)	16.235	31.821
Unique reflections	37913 (2747)	8116 (1064)
Multiplicity	8.2 (8.0)	8.2 (8.7)
Completeness (%)	99.8 (99.9)	99.8 (100.0)
Mean I/sigma(I)	22.1 (2.0)	12.7 (2.5)
R_{meas}	0.062 (1.205)	0.176 (1.06)
$CC_{1/2}$	1.0 (0.704)	0.997 (0.735)
Resolution (Å)	38.35 - 1.17 (1.20 – 1.17)	38.76 – 2.0 (2.05 – 2.0)
R_{work} / R_{free}	0.132 (0.234) / 0.164 (0.266)	0.180 (0.254) / 0.240 (0.278)
N. of atoms	1253	1172
Average atomic *B*-factor (Å^2)	15.170	34.467
R.m.s deviations		
Bond lengths (Å)	0.012	0.010
Bond angles (°)	1.70	1.55

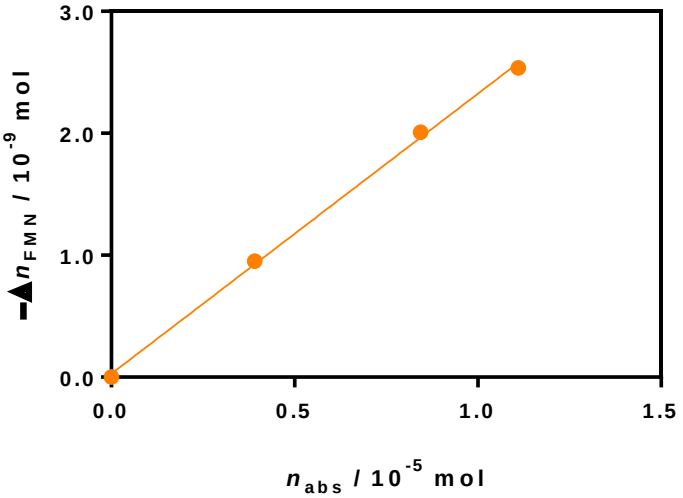


Figure S1. Photobleaching of miniSOG as a function of the absorbed light dose at 473 nm. The intercept yields the photobleaching quantum yield (Φ_{Phot} = 2.3 x 10^{-4}), which, combined with a triplet lifetime of 38.6 µs, yields (equation S3) the rate constant for photobleaching (k_{Phot} = 6.0 s^{-1}).

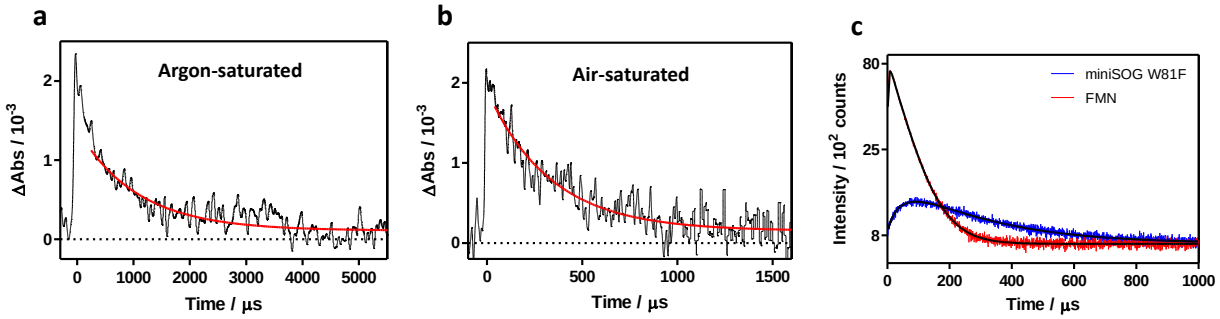


Table S2. $^1O_2$ signal fitting results

	τ_T / µs	τ_Δ / µs	Φ_Δ
miniSOG W81F	265	45	0.33
FMN	3.2	62	0.57[4]

Figure S2. Transient absorption decays of miniSOG W81F in (**a**) argon-saturated and (**b**) air-saturated dPBS. Fitting lines are in red. (λ_{ex} = 355 nm, λ_{obs} = 715 nm). (**c**) Time-resolved $^1O_2$ NIR phosphorescence of optically matched solutions of miniSOG W81F (blue) and FMN (red) in air-saturated dPBS-solutions. Fitting lines are in black. Y axis is in log scale.

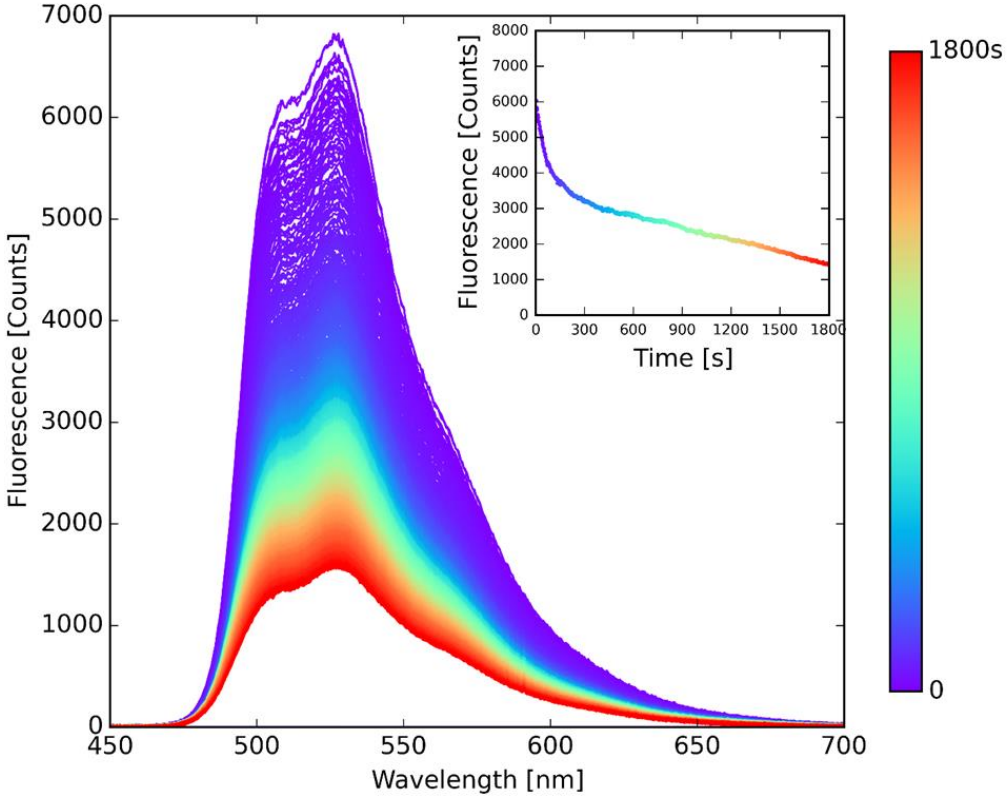


Figure S3: Time evolution of the fluorescence emission spectrum of a miniSOG crystal under sequential blue light irradiation of a 440 nm laser. Inset shows the decay of the integrated fluorescence signal.

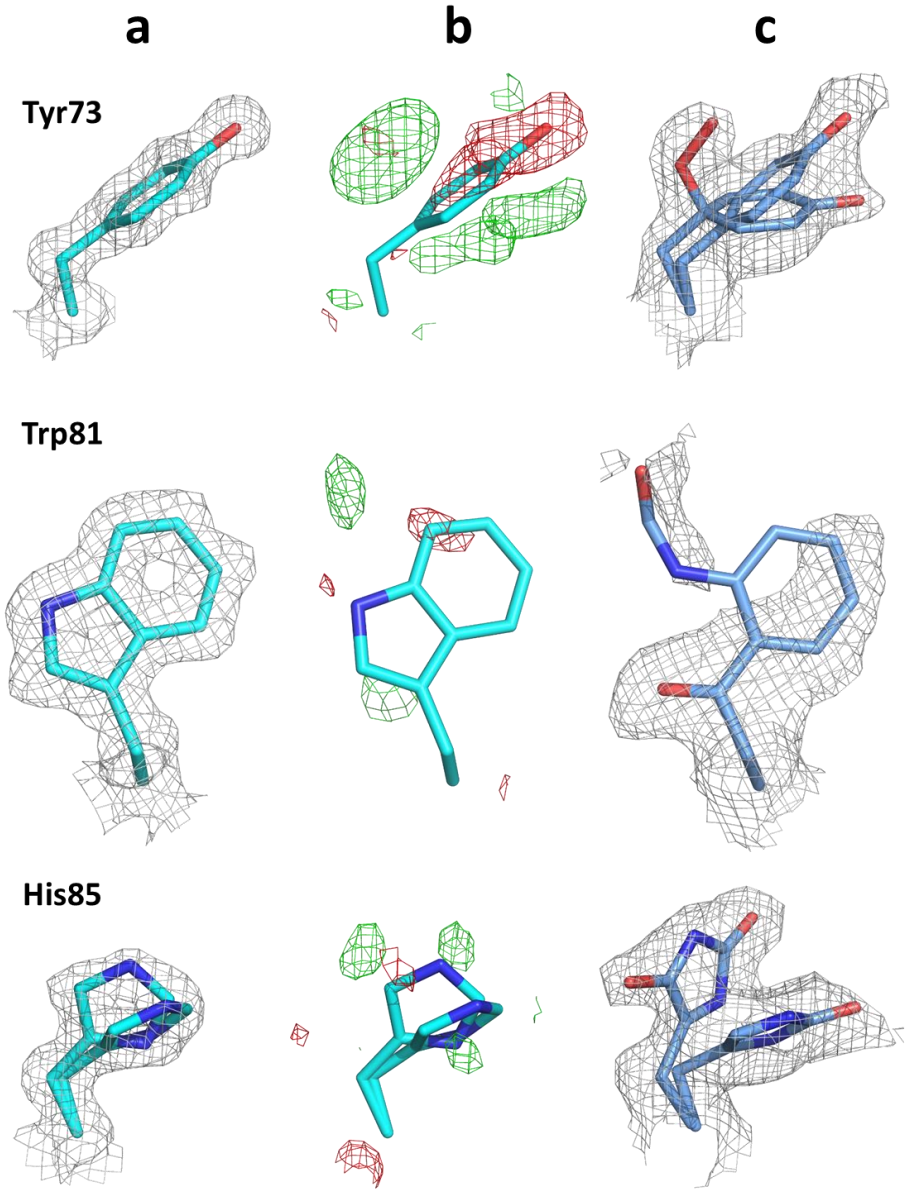


Figure S4: a. $2F_{obs} - F_{calc}$ electron density map contoured at a 1.0 σ level in the 1.17 Å native structure of miniSOG superimposed on the protein model for Tyr73, Trp81 and His85 (from top to bottom). **b.** Difference Fourier electron density map contoured at a 2.5, 2.0 and 2.5 σ level calculated between data sets recorded on a non-irradiated part, and a blue-light irradiated part of a crystal for Tyr73, Trp81 and His85, respectively. **c.** $2F_{obs} - F_{calc}$ electron density map contoured at a 1.0, 0.6 and 1.0 σ level in the 2.0 Å structure of blue-light irradiated miniSOG superimposed on the protein model for Tyr73, Trp81 and His85, respectively. Tyr73 is partially modelled as a γ-peroxotyrosine, Trp81 is modelled as a N-formylkynurenine and His85 is modelled partially as a singly oxidized histidine (2-oxo-histidine) and as a doubly oxidized histidine (2,4 dioxo-histidine).

Supplementary references

1. Christie, J. M. *et al.* Structural tuning of the fluorescent protein iLOV for improved photostability. *J. Biol. Chem.* **287,** 22295–22304 (2012).
2. Roeser, D., Schmidt, B., Preusser-Kunze, A. & Rudolph, M. G. Probing the oxygen-binding site of the human formylglycine-generating enzyme using halide ions. *Acta Crystallogr. D Biol. Crystallogr.* **63,** 621–627 (2007).
3. Colloc'h, N. *et al.* Oxygen pressurized X-ray crystallography: probing the dioxygen binding site in cofactorless urate oxidase and implications for its catalytic mechanism. *Biophys. J.* **95,** 2415–22 (2008).
4. Rodríguez-Pulido, A. *et al.* Assessing the potential of photosensitizing flavoproteins as tags for correlative microscopy. *Chem. Commun.* **52,** 8405–8408 (2016).

Time-resolved crystallography (TRX) enables the identification at a near atomic detail of the progressive structural changes of a protein fulfilling its function, thus leading to the establishment of genuine molecular movies. Laue (polychromatic) diffraction at 3rd-generation synchrotrons first managed to obtain such movies at a 100 ps time scale, and more recently, the 4th generation X-ray sources XFELs (X-ray Free Electron Lasers) achieved a 100 fs time scale using a monochromatic beam. The goal of this thesis was to investigate the possibility of performing TRX on a macromolecular monochromatic crystallography beamline, by taking advantage of the latest technological developments. We envisaged three different photosensitive proteins, for which I first developed various crystallogenesis strategies in order to control the size and shape of the crystals. I also considered various ways of presenting the crystals to the beam at room temperature, either a single crystal mounted in a loop and kept in a wet air stream, or microcrystals moving in a grease jet passing through the beam. We then investigated how specific radiation damage could affect the structures of intermediate states at room temperature, given how sensitive they can be at cryogenic temperature. We concluded that it constitutes a much lesser concern at room than at cryogenic temperature. After initial characterization of the various proteins and sample environments, we focused on a time-resolved crystallography approach on crystals of the plant photoreceptor domain *At*Phot2LOV2 (the LOV2 domain of the blue-light photoreceptor photoropin 2 from *Arabidopsis thaliana*). The dark state *At*Phot2LOV2 converts in microseconds into the signalling state, or light state, which relaxes in hundreds of seconds. Using a fast pixel X-ray detector, we characterized the structural decay of the light state, and showed that it proceeds *via* a space group conversion over a 20 min time course. We then slowed down the rate of the light state in crystals by limiting the quantity of photons, and we were able to monitor the process of light state population build-up with a 63 ms time resolution using an approach recording full data sets on less than 100 crystals, which we called TR-SOX for time-resolved serial oscillation crystallography, which takes advantage of merging partial data sets. Overall, our work paves the way for time-resolved crystallography on MX monochromatic synchrotron beamlines on regular-sized crystals of proteins that undergo structural rearrangements on the millisecond time scale.

La cristallographie résolue dans le temps (TRX) permet l'identification à l'échelle quasi atomique des changements structuraux progressifs au sein d'une protéine lorsqu'elle réalise sa fonction, permettant la constitution de véritables films moléculaires. La diffraction Laue (polychromatique) dans les synchrotrons de 3ème génération a d'abord permis d'obtenir de tels films à une échelle de temps de 100 ps, et plus récemment les sources de rayons X de 4ème génération, les lasers à électrons libres (XFELs, pour X-Ray Free Electron Lasers), ont permis d'atteindre une résolution temporelle de 100 fs en utilisant un faisceau monochromatique. L'objectif de cette thèse était d'explorer la possibilité d'effectuer des expériences de TRX au synchrotron sur les lignes de lumière monochromatiques dédiées à la cristallographie macromoléculaire en tirant parti des tout derniers développements technologiques. Nous avons envisagé de travailler sur trois protéines photosensibles différentes, pour lesquelles j'ai dans un premier temps développé différentes stratégies de cristallogenèse afin de contrôler la taille et la forme des cristaux. J'ai également testé différentes façons d'exposer les cristaux au faisceau de rayons X à température ambiante, soit des cristaux uniques montés sur une boucle et maintenu dans un flux d'humidité contrôlée, soit des microcristaux incorporés dans un flux de graisse passant à travers le faisceau de rayons X. Nous avons ensuite évalué comment les dommages radiatifs spécifiques pouvaient affecter les structures d'état intermédiaires de protéines à température ambiante, espèces particulièrement sensibles à température cryogénique. Nous avons conclu de cette étude que le dommage radiatif spécifique constitue un moindre problème à température ambiante qu'à température cryogénique. Après une caractérisation initiale des différentes protéines et des moyens de présentation au faisceau, nous avons concentré notre approche de cristallographie résolue dans le temps sur des cristaux d'un domaine de photorécepteur de plante, *At*Phot2LOV2 (le domaine LOV2 du récepteur à la lumière bleue phototropine 2 de la plante *Arabidopsis thaliana*). L'état initial de *At*Phot2LOV2 se convertit en quelques microsecondes en un état de signalisation, ou état photo-activé, qui se relaxe en quelques centaines de secondes. En utilisant un détecteur rapide de rayons X, nous avons caractérisé structuralement le phénomène de relaxation, et montré qu'elle se développe en une vingtaine de minutes *via* une conversion de groupe d'espace. Nous avons ensuite diminué la vitesse de formation de l'état photo-activé dans les cristaux en limitant la quantité de photons nécessaires à la photoconversion, et nous avons pu visualiser l'augmentation progressive de la population en l'état photo-activé dans le cristal avec une résolution temporelle de 63 ms, en utilisant une approche basée sur l'enregistrement de données de diffraction sur moins de 100 cristaux et tirant parti de la combinaison de jeux de données partiels, approche que nous avons nommée TR-SOX (Time-Resolved Serial Oscillation Crystallography). En résumé, notre travail ouvre la voie aux expériences de cristallographie résolue dans le temps sur les lignes de lumières monochromatiques de cristallographie macromoléculaire au synchrotron, sur des cristaux de taille usuelle pour des protéines subissant des réarrangements structuraux à l'échelle de la milliseconde.