Electronic Supplementary Information

Structural basis of the radical pair state in photolyases and cryptochromes.

Andrea Cellini,^{*a*} Madan Kumar Shankar,^{*a*} Weixiao Yuan Wahlgren^{*a*}, Amke Nimmrich^{*a*}, Antonia Furrer^{*b*}, Daniel James^{*b*}, Maximilian Wranik^{*b*}, Sylvain Aumonier^{*c*}, Emma V. Beale^{*d*}, Florian Dworkowski^{*c*}, Jörg Standfuss^{*b*}, Tobias Weinert^{*b*} and Sebastian Westenhoff,^{**a*,*e*}

^{*a*} Department of Chemistry and Molecular Biology, University of Gothenburg Box 462, 40530 Gothenburg, Sweden

^b Division of Biology and Chemistry-Laboratory for Biomolecular Research, Paul Scherrer Institut, 5232 Villigen, Switzerland

^c Photon Science Division - Laboratory for Macromolecules and Bioimaging (LSB), Paul Scherrer Institut, 5232 Villigen, Switzerland

^{*d*} Photon Science Division - Laboratory for Synchrotron Radiation and Femtochemistry (LSF), Paul Scherrer Institut, 5232 Villigen, Switzerland

^e Department of Chemistry-BMC, University of Uppsala, Husargatan 3, 752 37 Uppsala

* To whom correspondence should be addressed; E-mail: westenho@cmb.gu.se

Experimental details

Protein expression and crystallyzation

Dm(6-4) photolyase gene was codon optimized and ordered at GenScript. The gene was inserted between NcoI and XhoI in pET21d (+) plasmid with a stop-codon before the C-terminal His-tag sequence. The plasmid was transformed in BL21(DE3) cells and grew in a Studier medium¹with additional 50 μ g/mL carbenicillin for 2-3 hours at 37 ° and then overnight at 20 °. Cells were lysated with sonication (Q700 sonicator, Qsonica) with the settings: 30 % amplitude, pulse-on time 10 second, pulse-off time 30 second and process time of 10 min. The protein was purified through Hi-trap Heparin Column purification (Ge Healthcare) followed by size exclusion chromatography (Hiload Superdex 16/600 200 pg, Ge Healthcare)². The protein concentration was estimated by determining the FAD concentration and the photoconversion of the protein was evaluated by illuminating the sample with a 445 nm LED³. All the experimental procedures were performed under safe red light.

Batch crystallization

The crystals were first grown at 4 $^{\circ}$ C in hanging drop plates with a reservoir of 100 mM bis-tris pH=6.5, 200 mM lithium sulphate monohydrate, 22 % PEG 3350, 0.5 % Ethyl acetate. After two days, the macrocrystals were crushed and used as seeds for batch crystallization³. Prior to collecting data, the crystals were spun down, 90 % of the mother liquor was removed and the content of the vial was mixed in 1:3 ratio with hydroxy ethyl cellulose (HEC) matrix. HEC matrix was produced by dissolving 22 % of HEC in water (w/w).

Data collection

The data were collected at beamline X06SA at the Swiss Light Source Synchrotron (SLS). The sample was extruded through a high viscosity injector with a 75 μ m diameter nozzle⁴. The sample flow rate was set to 0.08 μ l/min (300 μ m/s) for both dark and light steady state data sets. For the light data set, a CW 473 nm laser(Crystalaser CL-473-150) with a fluence of 612 mW/mm2 was employed. From the flow speed and distance between nozzle tip and x-ray interaction point (90 μ m) we estimate that the sample was illuminated for approximately 0.3s before being probed by the x-ray. The sample was prepared and data were collected at room temperature.

Data analysis

The diffraction images were indexed, integrated and merged with Crystfel 0.9.1⁵. Indexamajig was used for indexing patterns with xgandalf⁶. These patterns had indexing ambiguity issues related to the tetragonal Bravais lattice of the crystals. This issue was addressed and corrected with ambigator by using the operator h,k,-l. The resulting stream file was merged, scaled and post-refined with partialator. The partial reflection were treated using the unity model in partialator. The dark state structure is phased using MR (PDB ID: 7AYV), The structure was further refined using Phenix⁷ and subsequently modelled with COOT⁸.

Difference maps (Fobsdark-Fobslight) were computed using PHENIX's Isomorphous difference map module⁷ and loaded in Coot⁸ as reference for residues position refinement for the light structure. The extrapolated structure factors were calculated as follows: $(Fe = Fo(dark) + 1/r\Delta Fo^{9,10})$.

The extrapolated maps were used to estimate the percentage of activation (the percentage of the protein that get excited by the laser). The mean electron density of the residues that we see changes occurring at and the mean negative features at 3 rmsd were plotted against different level of activa-

tion (Fig.S1)¹¹. The point at which the two plots intersects gives us the percentage of protein that is activated by light. The activation factor was deemed to be 14 %. The extrapolated map was used to refine the light activated structure in COOT⁸. Further, the calculated difference maps were plotted using the reported methods¹².

The dark structure has been deposited in PDB with the 7QUT ID, whereas the light structure can be found in the CXIDB (ID 202) .

UV/vis spectroscopy UV/VIS absorption measurements of Dm(6-4)photolyase in solution were performed to investigate which state is reached by illumination. The sample was excited for 100 ms with light from an LED at 453 nm and spectra were recorded 5 ms after the end of the illumination. The results are presented in Fig. 3, panel A shows the raw spectra for illuminated and dark measurement. Panel B shows the background corrected spectra, calculated from the spectra in panel A. This corresponds to the absorption spectra of Dm(6-4)photolyase before and after illumination. For the dark measurement (blue) we observe a good agreement with the expected absorption profile for a photolyase in the oxidized FAD state and a reduced state in light (red). Panel C shows absorption spectra of the different oxidation states of the FAD taken from literature.¹³ Panel D displays the difference absorption spectrum between light and dark (blue) which was computed by subtracting the two spectra in panel A. A matrix division was performed using the spectra in C and this yielded the contribution of the states to the difference scattering (see inset). The resulting difference spectrum is shown in red. This verifies that after illumination for 100ms the photolyase is in the semi-reduced FAD⁻ state. The minor negative contribution of FADH is attributed to uncertainties in the measurement. Table 1 Data and refinement statistics.

	dark	light
PDB code	7QUT	
Space group	P 41	P41
Cell constants		
a, b, c (Å)	103.60 103.60 52.08	103.60 103.60 52.08
$lpha,eta,\gamma$ (°)	90.0 90.0 90.0	90.0 90.0 90.0
Resolution (Å) †	19.30-2.24 (2.24-2.26)	19.30-2.50 (2.50-2.52)
Data completeness $(\%)^{\dagger}$	100 (100)	100 (100)
$R_{split}(\%)$ [†] *	9.53 (151.63)	7.87 (72.15)
$CC^{* \dagger 1}$	0.99(0.58)	0.99 (0.34)
CC1/2 ²	0.96 (0.20)	0.96 (0.06)
$<$ I/ σ (I) $>$ †	8.58 (0.71)	13.95 (1.41)
Multiplicity [†]	1948	2902
Number of hits	461638	829692
Number of indexed hits	64785	82895
Number of total reflection	52396029	56202479
Number of unique reflections	26853	19365
Refinement		
R_{work}/R_{free}	0.15/0.19	0.41/0.44 ³
Wilson B-factor (Å ²)	41.6	n/a
Total number of atoms	4277	4277
Average B, all atoms ($Å^2$)	59.72	60.05
R.m.s deviations		
Bond lengths (Å)	0.007	0.014
Bond angle (°)	1.56	1.53
[†] Highest resolution shell is shown in parenthesis.		

[†] Highest resolution shell is shown in parenthesis. * Rsplit= $1/\sqrt{2 \frac{\sum kkl/leven-Iodd|}{1/2 \sum hkl/leven+Iodd|}}$ 1 CC^{*} = $\sqrt{\frac{2CCl/2}{1+CCl/2}}$ 2 CC1/2= $\frac{\sigma t^2}{\sigma t^2 + \sigma t^2}$ where σt^2 is the variance of the difference between the intensities and their average and σe^2 is the variance of random error of merged half datasets. 3 The R values are not meaningful because the structure was refined only in real space and only in selected areas.



Fig.S 1 Negative features plotted in function of the percentage of activation. The point of intersection determines the level of activation. In our case, the activation level is about 14%

				258T		268G		278K		288K		294-		301№	l .	311A		321-		330W	
D.melanogaster(6-4)PL/1-540	250 A	PNS-	LEPS		PYLKF	GCLSA	RLENC	KLKE		ОРКН S	OPP-		VSLIG		EFYYT	VAAAE	PNFDR	1L - GN	v ү <mark>с</mark> мо	IPWOE-H	HPD 335
X.laevis CRY1/1-616	240 N	ANS-	LLAS	TTGLS	PYLRF	GCLSC	RLFYF	KL TD		VKKNS	SPP-		LSLYG	OLLWR	EFFYT	AATNN	PREDK	1D - GN	PICVO	IPWDR-M	NPE 325
E.rubecola CRY1/1-620	240 N	ANS-	LLAS	PTGLS	PYLRF	GCLSC	RLFYF	KL TD		VKKNS	SPP-		LSLYG		EFFYT	AATNN	PREDKI	1E - GN	PICVO	IPWDK-M	NPE 325
C.livia CRY1/1-620	240 N	ANS-	LLAS	PTGLS	PYLRE	GCLSC	RLFYF	KL TD		VKKNS	SPP-		LSLYG		EFFYT	AATNN	PREDK	1E - GN	PICVO	IPWDK-M	NPE 325
D.rerio(6-4)PL/1-519	240 S	PNS-	LIPS	TTVLS	PYVRF	GCLSA	RTFWV	RL AD	VYRG	K - TH S	DPP-		VSLHG	OLLWR	EFFYT	TAVGI	PNFNK	1E - GN	SACVO	VDWDN-M	NPE 324
X.laevis(6-4)PL/1-526	240 E	PNS-	LTPS	TTVLS	ΡΥνκε	GCLSA	RTFW	K I AD	IYOG	к - кн s	DPP-		VSLHG		EFYYT	TGAGI	PNFNK	1E - GN	PVCVO	VDWDN-M	NKE 324
A.thaliana (6-4)PL/1-556	267 D	PSAF	LKPA	TTVMS	PYLKE	GCLSS	RYFYC	CLON		V К К Н Т	SPP-		VSLLG		EFFYT	TAFGT	PNFDK	1K - GN	RICKO	IPWNE-D	DHA 353
E.coli CPD PL/1-472	229 F	PAV-	E G	TSRLS	ASLAT	GGLSP	ROCLH	RLLA	EOPO	ALDGG	AGS-		- VWLN	ELIWR	EFYRH	LITYH	PSLCK	RPFL	AWTDR	VOWOS-N	NPA 312
A.thaliana CRY2/1-612	238 K	VVG-	N S	TSLLS	PYLHF	GEISV	RHVFC	CARM	IKOIII	WARDK	NSEG	EE SA		GIGLR	EYSRY	ICFNF	PFTHE) S - L L	SHLRF	FPWDA-D	DVD 326
A.thaliana CRY1/1-681	241 K	ADS -	A T	TSFLS	PHLHF	GEVSV	RKVFH	LVRI	KOVA	WANEG	NEAG	EE S \	/NLFLK	SIGLR	EYSRY	ISFNH	PYSHE	RP - L L (GHLKF	FPWAV-D	DEN 329
Dmelanogaster CRY1/1-542	255 A	LPN-	IHDS	PKSMS	AHLRE	GCLSV	RREYV	/S VHD	LFKN	VOLRA	CVRG	VOMTGO	GAHITG		EYFYT	MSVNN	PNYDR	1E - GN	DICLS	IPWAKPM	NEN 348
C.livia CRY4/1-525	238 I	PNS-	LLPS	TTGLS	PYFSM	GCLSV	RTFFH	RLSN	IIYAO	AKHHS	LPP-		VSLOG	OLLWR	EFFYT	VASAT	ONFTO	1A - GN	PICLO	IHWYE-D	DAE 323
_							_														
		33	9A	3	49F	3	59Q		369Н		379D		389R		399D		409W		419Y	4	127A
D.melanogaster(6-4)PL/1-540	336 H	LEAV	VTH G R	TGYPF	IDAIM	R <mark>Q L</mark> RQ	EGWIF	HLAR	HAVA	CFLTR	GDLŴ	I SWEE	QRVFE	QLLLD	Q <mark>D</mark> WAL	NAGNW	MWL S A S	SAFFHO	QYFR-	- VYS <mark>P</mark> VA	4 F <mark>G</mark> 429
X.laevis_CRY1/1-616	326 A	LAKV	VA E <mark>G R</mark>	TGFPW	/	T <mark>Q L</mark> RQ	EGWIH	HLAR	HAVA	CFLTR	GDLW	I SWEE	MKVFE	ELLLD	A <mark>D</mark> WS V	'N A <mark>G</mark> S W	MWL SC :	SFFQ	QFFH-	- CYCPVC	GF <mark>G</mark> 419
E.rubecola_CRY1/1-620	326 A	LAKV	VA E <mark>G</mark> R	TG F P W	/	T <mark>Q L</mark> RQ	EGWIH	HLAR	HAVA	CFLTR	GDLW	ISWEE	MKVFE	ELLLD	A <mark>D</mark> WS V	'N A <mark>G</mark> S W	MWL SC :	SFFQ	QFFH-	- CYC PVC	GF <mark>G</mark> 419
C.livia_CRY1/1-620	326 A	LAKV	VA E <mark>G</mark> R	TGFPW	/	T <mark>Q L</mark> RQ	EGWIH	HLAR	HAVA	CFLTR	GDLW	I SWEE	S V K V F E	ELLLD	A <mark>D</mark> WS V	'N A <mark>G</mark> S <mark>W</mark>	MWL SC :	SFFQ	QFFH-	- CYCPVC	GF <mark>G</mark> 419
D.rerio(6-4)PL/1-519	325 H	LAAV	VREA R	TGFPF	IDTIM	T <mark>Q L</mark> RQ	EGWI	HLAR	HAVA	CFLTR	GDLW	I SWEE	QKVFE	ELLLD	S <mark>D</mark> WS L	NAGNW	QWL SA	STFFHO	QYFR-	- VYS <mark>P</mark> IA	4F <mark>G</mark> 418
X.laevis(6-4)PL/1-526	325 H	LEAV	VS E <mark>G</mark> R	TGYPF	I <mark>D</mark> A I M	T <mark>QL</mark> RT	EGWIH	HLAR	HAVA	CFLTR	GDLW	I SWEE	QKVFE	ELLLD	A <mark>D</mark> WS L	NAGNW	'L WL <mark>S</mark> A <mark>!</mark>	SAFFHO	QFFR-	- VYS <mark>P</mark> VA	4 F <mark>G</mark> 418
A.thaliana_(6-4)PL/1-556	354 M	LAAV	VRD <mark>G</mark> K	TGYPW	/	V <mark>Q L</mark> L K	W <mark>G WM</mark> H	HLAR	HCVA	CFLTR	GDLF	IHWEQ	RDVFE	R L L I D	S <mark>D</mark> WA I	NNGNW	MWL SC :	SFFY	QFNR-	- IYSPIS	5 F <mark>G</mark> 447
E.coli_CPD_PL/1-472	313 H	LQAV	VQ E <mark>G</mark> K	TGYPI	VDAAM	R <mark>QL</mark> NS	TGWMH	NRLR	MITA	SFLVK	-DLL	IDWRE	E R Y F M	SQLID	GDLAA	NNGGW	Q <mark>W</mark> A A S '	T <mark>G T</mark> D A A	A P Y F R	- IFNPT	TQ <mark>G</mark> 406
A.thaliana_CRY2/1-612	327 K	FKAV	VRQ <mark>G</mark> R	TGYPL	V D A G M	RELWA	TGWMH	NRIR	VIVS	SFAVK	- F L L	L P WKW	SMKYFW	DTLLD	ADLEC	DILGW	QYISG	S I PDGI	HELDR	- LDNPAL	LQ <mark>G</mark> 420
A.thaliana_CRY1/1-681	330 Y	FKAV	VRQ <mark>G</mark> R	TGYPL	V <mark>D</mark> A G M	RELWA	TGWLH	DRIR	VVVS	SFFVK	- V L Q		SMKYFW	D T L L <mark>D</mark>	A <mark>D</mark> LES	DALGW	Q <mark>Y</mark> ITG	T L P D S I	REFDR	- I D N P Q F	FE <mark>G</mark> 423
Dmelanogaster_CRY1/1-542	349 L	LQSV	VR L <mark>G</mark> Q	TG F P L	I <mark>D</mark> G AM	R <mark>Q L</mark> L A	EGWLH	HTLR	NTVA	TFLTR	GGLW	Q S <mark>W</mark> E H <mark>C</mark>	SLQHFL	KYLLD	A <mark>D</mark> WS V	CAGNW	MWV S S S	SAFER	LLDSS	LVTCPVA	4L <u>A</u> 444
C.livia_CRY4/1-525	324 R	LHKV	<mark>ν</mark> κ τ Α Q	TG F P W	/	T <mark>QL</mark> RQ	EGWI	HLAR	HAVA	CFLTR	GDLW	I S WE E	MKVFE	E L L L D	ADYSI	NAGNW	MWL SA	AFFHI	HYTR-	- F C <mark>P</mark> V F	RF <mark>G</mark> 417
			437	н	447	s	45	7P	4	467A		477P		487H	ale ale	497A	ţ	07G		510-	
D.melanogaster(6-4)PL/1-540	430 K	KTDF	OGHY	IRKYV	PELSK	YPAGC	IYEPV	KASL	VDQR	AYGCV	LGTD	YPHRIN	KHEVV	HKENI	K RMG A	AYKVN	REVRTO	G K E			509
X.laevis CRY1/1-616	420 K	R TD P	PNGDY	IRRYL	PILKG	F <mark>P</mark> P K Y	IYDPV	N A P E	туок	ΑΑΚΟΙ	IGVN	<mark>Ү Р</mark> К Р М \	/NHAEA	SRLNI	ERMKQ	IYQQL	SRYRGI	GL			499
E.rubecola CRY1/1-620	420 R	R TD P	NGDY	IRRYL	PVL RG	FPAKY	IYDPV	N A P E	SIQK	ΑΑΚΟΙ	IGVN	<mark>Ү Р</mark> К Р М \	/NHAEA	SRLNI	ERMKQ	IYQQL	SRYRGI	GL			499
C.livia CRY1/1-620	420 R	R TD F	NGDY	IRRYL	PVL RG	F <mark>P</mark> A K Y	IYDPV	N A P E	SIQK	ΑΑΚΟΙ	IGVN	<mark>Ү Р</mark> К Р М \	/NHAEA	SRLNI	ERMKQ	IYQQL	SRYRGI	GL			499
D.rerio(6-4)PL/1-519	419 K	KTDK	(H <mark>G</mark> DY	IKKYL	PVLKK	F <mark>P</mark> TE Y	IYE <mark>P</mark> V	K A P R	SVQE	RAGCI	VGKD	Y P R P I \	/DHEVV	нккиі	LRMKA	AYAK-	R S I	PED			494
X.laevis(6-4)PL/1-526	419 K	KTDK	(N <mark>G</mark> DY	IKKYL	PILKK	F <mark>P</mark> A E Y	IYEPV	K S P R	SLQE	RAGCI	IGKD	ΥΡΚΡΙΛ	/EHNVV	SKQNI	QRMKA	AYAR-	R S (GST			494
A.thaliana (6-4)PL/1-556	448 K	KYDF	D <mark>G</mark> KY	IRHFL	PVLKD	м <mark>р</mark> ко ү	IYE <mark>P</mark> V	TAPL	SVQT	KANCI	VGKD	<mark>Ү Р</mark> КРМ\	/LHDSA	SKECK	RKMGE	AYALN	KKMDGI	(VD			527
E.coli CPD PL/1-472	407 E	KFDH	IE <mark>G</mark> EF	IRQWL	PELRD	V <mark>P</mark> G K V	VH E <mark>P</mark> V	KWAQ	KAG-		VTLD	ΥΡΟΡΙΛ	/ЕНКЕА	RVQTL	AAYEA	ARKGK					472
A.thaliana_CRY2/1-612	421 A	KYDF	P E <mark>G</mark> E <mark>Y</mark>	IRQWL	PELAR	L <mark>P</mark> T E W	THH P V	DAPL	TVLK	ASGVE	LG TN	ΥΑΚΡΙΝ	/DIDTA	RELLA	KAISR	TREAQ	IMIGA	APD		E I \	AD 505
A.thaliana CRY1/1-681	424 Y	K F D F	N G E Y	VRRWL	PELSR	L P TDW	тннри	N A P E	SVLQ	AAGIE	LGSN	YPLPIN	/GLDEA	KARLH	EALSQ	MWQLE	AASRA	AIENG	SEEGL	GDSAEVE	EEA 519
Dmelanogaster CRY1/1-542	445 K	R L D F	DG TY	IKQYV	PELMN	V <mark>P</mark> K E F	VHEPV	RMSA	EQQE	QYECL	IGVH	YPERII	DLSMA	VKRNM	LAMKS	LRNSL	ITPPH	IC R			524
C.livia CRY4/1-525	418 K	RTDF	PEGOY	IRKYL	PVLKN	F P T K Y	IYEPV	TASE	EEOR	O A G C I	IGRD	YPFPM		SDRNL	OLMRR	VREEO	RG	TAO			494

Fig.S 2 Alignment of the C-terminal region of photolyases and cryptocrome. The alignment was performed in Jalview 2.11.1.4 with Clustal omega¹⁴. Residues conserved more than 25% are coloured according to clustal color scheme. The asterisks are placed on top of residues which show difference signal in our data.



Fig.S 3 Results from the UV/Vis spectroscopy measurements. A: Raw spectra, B: absorption spectra before and after illumination, C: absorption spectra of different oxidation states of FAD from ¹³, D: difference spectrum from experiment (blue) and result from matrix division (red). For details see text.

Notes and references

- 1 F. W. Studier, Protein expression and purification, 2005, 41, 207-234.
- 2 M. J. Maul, T. R. Barends, A. F. Glas, M. J. Cryle, T. Domratcheva, S. Schneider, I. Schlichting and T. Carell, Angewandte Chemie International Edition, 2008, 47, 10076–10080.
- 3 A. Cellini, W. Yuan Wahlgren, L. Henry, S. Pandey, S. Ghosh, L. Castillon, E. Claesson, H. Takala, J. Kübel, A. Nimmrich et al., Acta Crystallographica Section D: Structural Biology, 2021, 77, 1001–1009.
- U. Weierstall, D. James, C. Wang, T. A. White, D. Wang, W. Liu, J. C. Spence, R. B. Doak, G. Nelson, P. Fromme, R. Fromme, I. Grotjohann, C. Kupitz, N. A. Zatsepin, H. Liu, S. Basu, D. Wacker, G. Won Han, V. Katritch, M. Boutet, Sébastien Messerschmidt, G. J. Williams, J. E. Koglin, M. M. Seibert, M. Klinker, C. Gati, R. L. Shoeman, A. Barty, H. N. Chapman, R. A. Kirian, K. R. Beyerlein, D. Stevens, Raymond C. Li, S. T. A. Shah, N. Howe, M. Caffrey and V. Cherezov, *Nature communications*, 2014, 5, 1–6.
- 5 T. A. White, R. A. Kirian, A. V. Martin, A. Aquila, K. Nass, A. Barty and H. N. Chapman, Journal of applied crystallography, 2012, 45, 335–341.
- 6 Y. Gevorkov, O. Yefanov, A. Barty, T. A. White, V. Mariani, W. Brehm, A. Tolstikova, R.-R. Grigat and H. N. Chapman, Acta Crystallographica Section A: Foundations and Advances, 2019, 75, 694–704.
- 7 D. Liebschner, P. V. Afonine, M. L. Baker, G. Bunkóczi, V. B. Chen, T. I. Croll, B. Hintze, L.-W. Hung, S. Jain, A. J. McCoy et al., Acta Crystallographica Section D: Structural Biology, 2019, 75, 861–877.
- 8 P. Emsley, B. Lohkamp, W. G. Scott and K. Cowtan, Acta Crystallographica Section D: Biological Crystallography, 2010, 66, 486–501.
- 9 P. Nogly, T. Weinert, D. James, S. Carbajo, D. Ozerov, A. Furrer, D. Gashi, V. Borin, P. Skopintsev, K. Jaeger, K. Nass, P. Båth, R. Bosman, J. Koglin, M. Seaberg, T. Lane, D. Kekilli, S. Brünle, T. Tanaka, W. Wu, C. Milne, T. White, A. Barty, U. Weierstall, V. Panneels, E. Nango, S. Iwata, M. Hunter, I. Schapiro, G. Schertler, R. Neutze and S. Jörg, *Science*, 2018, 361, eaat0094.
- 10 K. Pande, C. D. Hutchison, G. Groenhof, A. Aquila, J. S. Robinson, J. Tenboer, S. Basu, S. Boutet, D. P. DePonte, M. Liang et al., Science, 2016, 352, 725–729.

- 11 S. Pandey, R. Bean, T. Sato, I. Poudyal, J. Bielecki, J. C. Villarreal, O. Yefanov, V. Mariani, T. A. White, C. Kupitz et al., Nature methods, 2020, 17, 73–78.
- E. Claesson, W. Y. Wahlgren, H. Takala, S. Pandey, L. Castillon, V. Kuznetsova, L. Henry, M. Panman, M. Carrillo, J. Kübel, R. Nanekar, L. Isaksson, A. Nimmrich, A. Cellini, D. Morozov, M. Maj, M. Kurttila, R. Bosman, E. Nango, R. Tanaka, T. Tanaka, L. Fangjia, S. Iwata, S. Owada, K. Moffat, G. Groenhof, E. A. Stojković, S. M. Ihalainen, Janne A. and S. Westenhoff, *Elife*, 2020, 9, e53514.
- 13 P. Müller, E. Ignatz, S. Kiontke, K. Brettel and L. O. Essen, Chemical Science, 2018, 9, 1200–1212.
- 14 A. M. Waterhouse, J. B. Procter, D. M. Martin, M. Clamp and G. J. Barton, *Bioinformatics*, 2009, 25, 1189–1191.