



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 07:24 PM GMT

PDB ID : 4OQW  
Title : Crystal structure of mCardinal far-red fluorescent protein  
Authors : Burg, J.S.; Chu, J.; Lam, A.J.; Lin, M.Z.; Garcia, K.C.  
Deposited on : 2014-02-10  
Resolution : 2.21 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

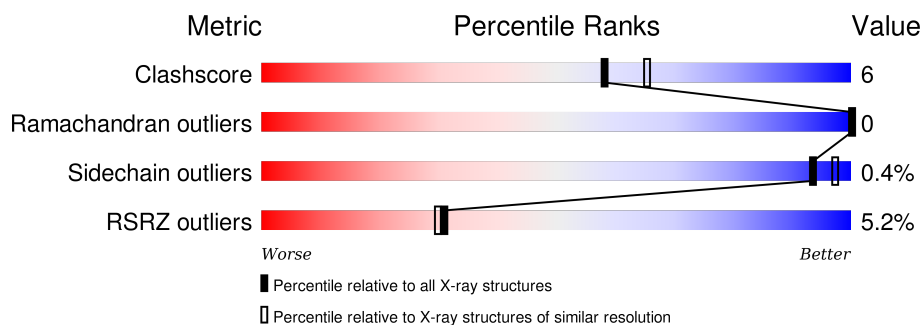
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.21 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	102246	5146 (2.24-2.20)
Ramachandran outliers	100387	5065 (2.24-2.20)
Sidechain outliers	100360	5066 (2.24-2.20)
RSRZ outliers	91569	4414 (2.24-2.20)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	228	
1	B	228	
1	C	228	
1	D	228	
1	E	228	
1	F	228	
1	G	228	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	H	228	<div><div></div><div>12%</div><div></div><div>78%</div><div></div><div>8%</div><div></div><div>14%</div></div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 27289 atoms, of which 12883 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Fluorescent protein FP480.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	226	Total	C	H	N	O	S	0	0	0
			3539	1148	1731	307	340	13			
1	B	224	Total	C	H	N	O	S	0	0	0
			3498	1136	1707	304	338	13			
1	C	223	Total	C	H	N	O	S	0	0	0
			3489	1133	1707	303	333	13			
1	D	226	Total	C	H	N	O	S	0	0	0
			3517	1143	1715	306	340	13			
1	E	214	Total	C	H	N	O	S	0	0	0
			3224	1064	1553	285	309	13			
1	F	215	Total	C	H	N	O	S	0	0	0
			3313	1085	1608	293	315	12			
1	G	212	Total	C	H	N	O	S	0	0	0
			3193	1060	1529	281	311	12			
1	H	196	Total	C	H	N	O	S	0	0	0
			2845	961	1333	250	288	13			

There are 184 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	SER	-	EXPRESSION TAG	UNP D0VX33
A	0	LYS	-	EXPRESSION TAG	UNP D0VX33
A	1	GLY	-	EXPRESSION TAG	UNP D0VX33
A	2	GLU	-	EXPRESSION TAG	UNP D0VX33
A	6	LYS	THR	CONFLICT	UNP D0VX33
A	28	THR	SER	CONFLICT	UNP D0VX33
A	61	CYS	SER	CONFLICT	UNP D0VX33
A	63	NRQ	MET	CHROMOPHORE	UNP D0VX33
A	63	NRQ	TYR	CHROMOPHORE	UNP D0VX33
A	63	NRQ	GLY	CHROMOPHORE	UNP D0VX33
A	67	LYS	HIS	CONFLICT	UNP D0VX33
A	80	PHE	TRP	CONFLICT	UNP D0VX33
A	104	VAL	ALA	CONFLICT	UNP D0VX33

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	121	LEU	ILE	CONFLICT	UNP D0VX33
A	143	THR	HIS	CONFLICT	UNP D0VX33
A	146	THR	MET	CONFLICT	UNP D0VX33
A	158	CYS	ALA	CONFLICT	UNP D0VX33
A	160	MET	LEU	CONFLICT	UNP D0VX33
A	171	HIS	ILE	CONFLICT	UNP D0VX33
A	174	LEU	PHE	CONFLICT	UNP D0VX33
A	194	PHE	TYR	CONFLICT	UNP D0VX33
A	197	ARG	TYR	CONFLICT	UNP D0VX33
A	207	ASN	LYS	CONFLICT	UNP D0VX33
B	-1	SER	-	EXPRESSION TAG	UNP D0VX33
B	0	LYS	-	EXPRESSION TAG	UNP D0VX33
B	1	GLY	-	EXPRESSION TAG	UNP D0VX33
B	2	GLU	-	EXPRESSION TAG	UNP D0VX33
B	6	LYS	THR	CONFLICT	UNP D0VX33
B	28	THR	SER	CONFLICT	UNP D0VX33
B	61	CYS	SER	CONFLICT	UNP D0VX33
B	63	NRQ	MET	CHROMOPHORE	UNP D0VX33
B	63	NRQ	TYR	CHROMOPHORE	UNP D0VX33
B	63	NRQ	GLY	CHROMOPHORE	UNP D0VX33
B	67	LYS	HIS	CONFLICT	UNP D0VX33
B	80	PHE	TRP	CONFLICT	UNP D0VX33
B	104	VAL	ALA	CONFLICT	UNP D0VX33
B	121	LEU	ILE	CONFLICT	UNP D0VX33
B	143	THR	HIS	CONFLICT	UNP D0VX33
B	146	THR	MET	CONFLICT	UNP D0VX33
B	158	CYS	ALA	CONFLICT	UNP D0VX33
B	160	MET	LEU	CONFLICT	UNP D0VX33
B	171	HIS	ILE	CONFLICT	UNP D0VX33
B	174	LEU	PHE	CONFLICT	UNP D0VX33
B	194	PHE	TYR	CONFLICT	UNP D0VX33
B	197	ARG	TYR	CONFLICT	UNP D0VX33
B	207	ASN	LYS	CONFLICT	UNP D0VX33
C	-1	SER	-	EXPRESSION TAG	UNP D0VX33
C	0	LYS	-	EXPRESSION TAG	UNP D0VX33
C	1	GLY	-	EXPRESSION TAG	UNP D0VX33
C	2	GLU	-	EXPRESSION TAG	UNP D0VX33
C	6	LYS	THR	CONFLICT	UNP D0VX33
C	28	THR	SER	CONFLICT	UNP D0VX33
C	61	CYS	SER	CONFLICT	UNP D0VX33
C	63	NRQ	MET	CHROMOPHORE	UNP D0VX33
C	63	NRQ	TYR	CHROMOPHORE	UNP D0VX33

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	63	NRQ	GLY	CHROMOPHORE	UNP D0VX33
C	67	LYS	HIS	CONFLICT	UNP D0VX33
C	80	PHE	TRP	CONFLICT	UNP D0VX33
C	104	VAL	ALA	CONFLICT	UNP D0VX33
C	121	LEU	ILE	CONFLICT	UNP D0VX33
C	143	THR	HIS	CONFLICT	UNP D0VX33
C	146	THR	MET	CONFLICT	UNP D0VX33
C	158	CYS	ALA	CONFLICT	UNP D0VX33
C	160	MET	LEU	CONFLICT	UNP D0VX33
C	171	HIS	ILE	CONFLICT	UNP D0VX33
C	174	LEU	PHE	CONFLICT	UNP D0VX33
C	194	PHE	TYR	CONFLICT	UNP D0VX33
C	197	ARG	TYR	CONFLICT	UNP D0VX33
C	207	ASN	LYS	CONFLICT	UNP D0VX33
D	-1	SER	-	EXPRESSION TAG	UNP D0VX33
D	0	LYS	-	EXPRESSION TAG	UNP D0VX33
D	1	GLY	-	EXPRESSION TAG	UNP D0VX33
D	2	GLU	-	EXPRESSION TAG	UNP D0VX33
D	6	LYS	THR	CONFLICT	UNP D0VX33
D	28	THR	SER	CONFLICT	UNP D0VX33
D	61	CYS	SER	CONFLICT	UNP D0VX33
D	63	NRQ	MET	CHROMOPHORE	UNP D0VX33
D	63	NRQ	TYR	CHROMOPHORE	UNP D0VX33
D	63	NRQ	GLY	CHROMOPHORE	UNP D0VX33
D	67	LYS	HIS	CONFLICT	UNP D0VX33
D	80	PHE	TRP	CONFLICT	UNP D0VX33
D	104	VAL	ALA	CONFLICT	UNP D0VX33
D	121	LEU	ILE	CONFLICT	UNP D0VX33
D	143	THR	HIS	CONFLICT	UNP D0VX33
D	146	THR	MET	CONFLICT	UNP D0VX33
D	158	CYS	ALA	CONFLICT	UNP D0VX33
D	160	MET	LEU	CONFLICT	UNP D0VX33
D	171	HIS	ILE	CONFLICT	UNP D0VX33
D	174	LEU	PHE	CONFLICT	UNP D0VX33
D	194	PHE	TYR	CONFLICT	UNP D0VX33
D	197	ARG	TYR	CONFLICT	UNP D0VX33
D	207	ASN	LYS	CONFLICT	UNP D0VX33
E	-1	SER	-	EXPRESSION TAG	UNP D0VX33
E	0	LYS	-	EXPRESSION TAG	UNP D0VX33
E	1	GLY	-	EXPRESSION TAG	UNP D0VX33
E	2	GLU	-	EXPRESSION TAG	UNP D0VX33
E	6	LYS	THR	CONFLICT	UNP D0VX33

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	28	THR	SER	CONFLICT	UNP D0VX33
E	61	CYS	SER	CONFLICT	UNP D0VX33
E	63	NRQ	MET	CHROMOPHORE	UNP D0VX33
E	63	NRQ	TYR	CHROMOPHORE	UNP D0VX33
E	63	NRQ	GLY	CHROMOPHORE	UNP D0VX33
E	67	LYS	HIS	CONFLICT	UNP D0VX33
E	80	PHE	TRP	CONFLICT	UNP D0VX33
E	104	VAL	ALA	CONFLICT	UNP D0VX33
E	121	LEU	ILE	CONFLICT	UNP D0VX33
E	143	THR	HIS	CONFLICT	UNP D0VX33
E	146	THR	MET	CONFLICT	UNP D0VX33
E	158	CYS	ALA	CONFLICT	UNP D0VX33
E	160	MET	LEU	CONFLICT	UNP D0VX33
E	171	HIS	ILE	CONFLICT	UNP D0VX33
E	174	LEU	PHE	CONFLICT	UNP D0VX33
E	194	PHE	TYR	CONFLICT	UNP D0VX33
E	197	ARG	TYR	CONFLICT	UNP D0VX33
E	207	ASN	LYS	CONFLICT	UNP D0VX33
F	-1	SER	-	EXPRESSION TAG	UNP D0VX33
F	0	LYS	-	EXPRESSION TAG	UNP D0VX33
F	1	GLY	-	EXPRESSION TAG	UNP D0VX33
F	2	GLU	-	EXPRESSION TAG	UNP D0VX33
F	6	LYS	THR	CONFLICT	UNP D0VX33
F	28	THR	SER	CONFLICT	UNP D0VX33
F	61	CYS	SER	CONFLICT	UNP D0VX33
F	63	NRQ	MET	CHROMOPHORE	UNP D0VX33
F	63	NRQ	TYR	CHROMOPHORE	UNP D0VX33
F	63	NRQ	GLY	CHROMOPHORE	UNP D0VX33
F	67	LYS	HIS	CONFLICT	UNP D0VX33
F	80	PHE	TRP	CONFLICT	UNP D0VX33
F	104	VAL	ALA	CONFLICT	UNP D0VX33
F	121	LEU	ILE	CONFLICT	UNP D0VX33
F	143	THR	HIS	CONFLICT	UNP D0VX33
F	146	THR	MET	CONFLICT	UNP D0VX33
F	158	CYS	ALA	CONFLICT	UNP D0VX33
F	160	MET	LEU	CONFLICT	UNP D0VX33
F	171	HIS	ILE	CONFLICT	UNP D0VX33
F	174	LEU	PHE	CONFLICT	UNP D0VX33
F	194	PHE	TYR	CONFLICT	UNP D0VX33
F	197	ARG	TYR	CONFLICT	UNP D0VX33
F	207	ASN	LYS	CONFLICT	UNP D0VX33
G	-1	SER	-	EXPRESSION TAG	UNP D0VX33

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
G	0	LYS	-	EXPRESSION TAG	UNP D0VX33
G	1	GLY	-	EXPRESSION TAG	UNP D0VX33
G	2	GLU	-	EXPRESSION TAG	UNP D0VX33
G	6	LYS	THR	CONFLICT	UNP D0VX33
G	28	THR	SER	CONFLICT	UNP D0VX33
G	61	CYS	SER	CONFLICT	UNP D0VX33
G	63	NRQ	MET	CHROMOPHORE	UNP D0VX33
G	63	NRQ	TYR	CHROMOPHORE	UNP D0VX33
G	63	NRQ	GLY	CHROMOPHORE	UNP D0VX33
G	67	LYS	HIS	CONFLICT	UNP D0VX33
G	80	PHE	TRP	CONFLICT	UNP D0VX33
G	104	VAL	ALA	CONFLICT	UNP D0VX33
G	121	LEU	ILE	CONFLICT	UNP D0VX33
G	143	THR	HIS	CONFLICT	UNP D0VX33
G	146	THR	MET	CONFLICT	UNP D0VX33
G	158	CYS	ALA	CONFLICT	UNP D0VX33
G	160	MET	LEU	CONFLICT	UNP D0VX33
G	171	HIS	ILE	CONFLICT	UNP D0VX33
G	174	LEU	PHE	CONFLICT	UNP D0VX33
G	194	PHE	TYR	CONFLICT	UNP D0VX33
G	197	ARG	TYR	CONFLICT	UNP D0VX33
G	207	ASN	LYS	CONFLICT	UNP D0VX33
H	-1	SER	-	EXPRESSION TAG	UNP D0VX33
H	0	LYS	-	EXPRESSION TAG	UNP D0VX33
H	1	GLY	-	EXPRESSION TAG	UNP D0VX33
H	2	GLU	-	EXPRESSION TAG	UNP D0VX33
H	6	LYS	THR	CONFLICT	UNP D0VX33
H	28	THR	SER	CONFLICT	UNP D0VX33
H	61	CYS	SER	CONFLICT	UNP D0VX33
H	63	NRQ	MET	CHROMOPHORE	UNP D0VX33
H	63	NRQ	TYR	CHROMOPHORE	UNP D0VX33
H	63	NRQ	GLY	CHROMOPHORE	UNP D0VX33
H	67	LYS	HIS	CONFLICT	UNP D0VX33
H	80	PHE	TRP	CONFLICT	UNP D0VX33
H	104	VAL	ALA	CONFLICT	UNP D0VX33
H	121	LEU	ILE	CONFLICT	UNP D0VX33
H	143	THR	HIS	CONFLICT	UNP D0VX33
H	146	THR	MET	CONFLICT	UNP D0VX33
H	158	CYS	ALA	CONFLICT	UNP D0VX33
H	160	MET	LEU	CONFLICT	UNP D0VX33
H	171	HIS	ILE	CONFLICT	UNP D0VX33
H	174	LEU	PHE	CONFLICT	UNP D0VX33

*Continued on next page...*



*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
H	194	PHE	TYR	CONFLICT	UNP D0VX33
H	197	ARG	TYR	CONFLICT	UNP D0VX33
H	207	ASN	LYS	CONFLICT	UNP D0VX33

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	132	Total	O	0	0
			132	132		
2	B	103	Total	O	0	0
			103	103		
2	C	106	Total	O	0	0
			106	106		
2	D	91	Total	O	0	0
			91	91		
2	E	61	Total	O	0	0
			61	61		
2	F	57	Total	O	0	0
			57	57		
2	G	53	Total	O	0	0
			53	53		
2	H	68	Total	O	0	0
			68	68		

### 3 Residue-property plots [i](#)


These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Fluorescent protein FP480

Chain A: 




- Molecule 1: Fluorescent protein FP480

Chain B: 




- Molecule 1: Fluorescent protein FP480

Chain C: 




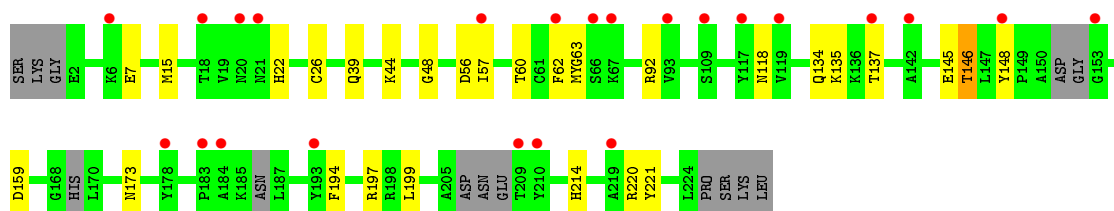
- Molecule 1: Fluorescent protein FP480

Chain D: 



- Molecule 1: Fluorescent protein FP480

Chain E: 



- | SER | VAL  | SER   | VAL |
|-----|------|-------|-----|
| LYS | LYS  | LYS   | LYS |
| LEU | LEU  | LYS   | LYS |
|     | ARG  | G17   | G1  |
|     | GLY  | T18   |     |
|     | VAL  | VAL   |     |
|     | M125 | ASN   |     |
|     | F126 | ASN   |     |
|     |      | HIS   |     |
|     | G130 | HIS   |     |
|     | PRO  | F24   |     |
|     | VAL  | R25   |     |
|     | M133 | C26   |     |
|     |      |       |     |
|     | T137 | G37   |     |
|     | L138 | T38   |     |
|     | G139 | Q39   |     |
|     | W140 |       |     |
|     |      |       |     |
|     | T146 | P50   |     |
|     | L147 | LEU   |     |
|     | Y148 | PRO   |     |
|     | P149 | F53   |     |
|     | A150 | A54   |     |
|     | D151 | F55   |     |
|     |      |       |     |
|     | E155 | L58   |     |
|     | G156 | ALA   |     |
|     | R157 | T60   |     |
|     |      | C61   |     |
|     | A161 | F62   |     |
|     |      | MYG63 |     |
|     | G166 | S66   |     |
|     | GLY  | K67   |     |
|     | GLY  |       |     |
|     | H169 | Q74   |     |
|     |      |       |     |
|     | K182 | F79   |     |
|     | PRO  |       |     |
|     | ALA  | S83   |     |
|     | K185 | P84   |     |
|     |      | P85   |     |
|     | R197 | R92   |     |
|     | ARG  | VAL   |     |
|     | LEU  | T94   |     |
|     | E200 |       |     |
|     | R201 | G99   |     |
|     | ILE  | G100  |     |
|     | E204 | V101  |     |
|     |      |       |     |
|     | E212 | V104  |     |
|     |      | THR   |     |
|     | V216 | GLN   |     |
|     |      | D107  |     |
|     | V221 |       |     |
|     | C222 | L110  |     |
|     | D223 |       |     |
|     | L224 | Y117  |     |
|     |      |       |     |

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	53.20 Å 136.70 Å 167.60 Å 90.00° 90.10° 90.00°	Depositor
Resolution (Å)	47.56 – 2.21 47.56 – 2.21	Depositor EDS
% Data completeness (in resolution range)	94.8 (47.56-2.21) 94.4 (47.56-2.21)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.48 (at 2.20 Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.4_1496)	Depositor
R, $R_{free}$	0.223 , 0.251 (Not available) , (Not available)	Depositor DCC
$R_{free}$ test set	NotAvailable	DCC
Wilson B-factor (Å <sup>2</sup> )	26.1	Xtriage
Anisotropy	0.517	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 27.7	EDS
Estimated twinning fraction	0.500 for h,-k,-l 0.387 for h,-k,-l	Xtriage
Reported twinning fraction	0.500 for h,-k,-l	Depositor
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.43$ , $\langle L^2 \rangle = 0.26$	Xtriage
Outliers	1 of 113418 reflections (0.001%)	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	27289	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.25% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup> Intensities estimated from amplitudes.

<sup>2</sup> Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: NRQ, MLY

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z  > 5$	RMSZ	$\# Z  > 5$
1	A	0.27	0/1803	0.47	0/2436
1	B	0.26	0/1786	0.46	0/2414
1	C	0.28	0/1777	0.49	0/2403
1	D	0.26	0/1797	0.45	0/2428
1	E	0.22	0/1659	0.41	0/2241
1	F	0.24	0/1697	0.44	0/2291
1	G	0.24	0/1649	0.44	0/2219
1	H	0.27	0/1490	0.43	0/2000
All	All	0.26	0/13658	0.45	0/18432

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1808	1731	1769	20	1
1	B	1791	1707	1745	24	0
1	C	1782	1707	1741	23	0
1	D	1802	1715	1758	30	0
1	E	1671	1553	1585	23	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	F	1705	1608	1642	19	1
1	G	1664	1529	1560	14	1
1	H	1512	1333	1363	13	0
2	A	132	0	0	8	0
2	B	103	0	0	7	1
2	C	106	0	0	11	0
2	D	91	0	0	10	0
2	E	61	0	0	7	0
2	F	57	0	0	8	0
2	G	53	0	0	3	0
2	H	68	0	0	4	0
All	All	14406	12883	13163	158	2

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 6.

The worst 5 of 158 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:201:ARG:NH1	1:D:204:GLU:OE1	2.05	0.88
1:C:29:GLU:OE2	2:C:341:HOH:O	1.96	0.84
1:A:56:ASP:OD2	1:A:137:THR:OG1	1.98	0.81
1:H:107:ASP:OD2	2:H:313:HOH:O	2.03	0.77
1:C:42:ARG:NH1	2:C:402:HOH:O	2.18	0.77

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:71:ASN:OD1	2:B:320:HOH:O[2_647]	2.11	0.09
1:F:159:ASP:OD2	1:G:173:ASN:ND2[2_656]	2.14	0.06

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	219/228 (96%)	217 (99%)	2 (1%)	0	100	100
1	B	217/228 (95%)	214 (99%)	3 (1%)	0	100	100
1	C	216/228 (95%)	213 (99%)	3 (1%)	0	100	100
1	D	219/228 (96%)	213 (97%)	6 (3%)	0	100	100
1	E	199/228 (87%)	197 (99%)	2 (1%)	0	100	100
1	F	204/228 (90%)	198 (97%)	6 (3%)	0	100	100
1	G	193/228 (85%)	189 (98%)	4 (2%)	0	100	100
1	H	167/228 (73%)	163 (98%)	4 (2%)	0	100	100
All	All	1634/1824 (90%)	1604 (98%)	30 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	192/194 (99%)	192 (100%)	0	100	100
1	B	190/194 (98%)	189 (100%)	1 (0%)	92	96
1	C	189/194 (97%)	188 (100%)	1 (0%)	92	96
1	D	191/194 (98%)	191 (100%)	0	100	100
1	E	168/194 (87%)	167 (99%)	1 (1%)	90	95
1	F	175/194 (90%)	175 (100%)	0	100	100
1	G	165/194 (85%)	164 (99%)	1 (1%)	90	95
1	H	145/194 (75%)	144 (99%)	1 (1%)	88	94
All	All	1415/1552 (91%)	1410 (100%)	5 (0%)	93	97

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	148	TYR
1	C	55	PHE
1	E	146	THR
1	G	210	TYR
1	H	148	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	106	GLN
1	D	171	HIS
1	F	21	ASN
1	B	171	HIS
1	E	134	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

24 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
1	MLY	A	163	1	9,10,11	0.53	0	9,11,13	0.95	0
1	MLY	A	181	1	9,10,11	0.52	0	9,11,13	0.88	0
1	NRQ	A	63	1	23,24,25	4.89	11 (47%)	23,32,34	4.13	8 (34%)
1	MLY	B	163	1	9,10,11	0.54	0	9,11,13	1.02	0
1	MLY	B	181	1	9,10,11	0.47	0	9,11,13	0.91	0
1	NRQ	B	63	1	23,24,25	4.85	11 (47%)	23,32,34	3.93	7 (30%)
1	MLY	C	163	1	9,10,11	0.50	0	9,11,13	0.96	0
1	MLY	C	181	1	9,10,11	0.49	0	9,11,13	0.89	0
1	NRQ	C	63	1	23,24,25	4.84	11 (47%)	23,32,34	4.01	8 (34%)



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
1	MLY	D	163	1	9,10,11	0.52	0	9,11,13	0.89	0
1	MLY	D	181	1	9,10,11	0.47	0	9,11,13	0.92	0
1	NRQ	D	63	1	23,24,25	4.89	11 (47%)	23,32,34	3.94	8 (34%)
1	MLY	E	163	1	9,10,11	0.49	0	9,11,13	0.90	0
1	MLY	E	181	1	9,10,11	0.50	0	9,11,13	0.92	0
1	NRQ	E	63	1	23,24,25	4.96	10 (43%)	23,32,34	4.11	7 (30%)
1	MLY	F	163	1	9,10,11	0.49	0	9,11,13	0.88	0
1	MLY	F	181	1	9,10,11	0.69	0	9,11,13	0.76	0
1	NRQ	F	63	1	23,24,25	4.91	11 (47%)	23,32,34	4.14	8 (34%)
1	MLY	G	163	1	9,10,11	0.48	0	9,11,13	0.92	0
1	MLY	G	181	1	9,10,11	0.64	0	9,11,13	0.95	1 (11%)
1	NRQ	G	63	1	23,24,25	4.93	12 (52%)	23,32,34	4.12	9 (39%)
1	MLY	H	163	1	9,10,11	0.50	0	9,11,13	0.91	0
1	MLY	H	181	1	9,10,11	0.56	0	9,11,13	0.90	0
1	NRQ	H	63	1	23,24,25	4.93	11 (47%)	23,32,34	4.12	7 (30%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	MLY	A	163	1	-	0/7/9/11	0/0/0/0
1	MLY	A	181	1	-	0/7/9/11	0/0/0/0
1	NRQ	A	63	1	-	0/9/31/32	0/2/2/2
1	MLY	B	163	1	-	0/7/9/11	0/0/0/0
1	MLY	B	181	1	-	0/7/9/11	0/0/0/0
1	NRQ	B	63	1	-	0/9/31/32	0/2/2/2
1	MLY	C	163	1	-	0/7/9/11	0/0/0/0
1	MLY	C	181	1	-	0/7/9/11	0/0/0/0
1	NRQ	C	63	1	-	0/9/31/32	0/2/2/2
1	MLY	D	163	1	-	0/7/9/11	0/0/0/0
1	MLY	D	181	1	-	0/7/9/11	0/0/0/0
1	NRQ	D	63	1	-	0/9/31/32	0/2/2/2
1	MLY	E	163	1	-	0/7/9/11	0/0/0/0
1	MLY	E	181	1	-	0/7/9/11	0/0/0/0
1	NRQ	E	63	1	-	0/9/31/32	0/2/2/2
1	MLY	F	163	1	-	0/7/9/11	0/0/0/0
1	MLY	F	181	1	-	0/7/9/11	0/0/0/0
1	NRQ	F	63	1	-	0/9/31/32	0/2/2/2
1	MLY	G	163	1	-	0/7/9/11	0/0/0/0

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	MLY	G	181	1	-	0/7/9/11	0/0/0/0
1	NRQ	G	63	1	-	0/9/31/32	0/2/2/2
1	MLY	H	163	1	-	0/7/9/11	0/0/0/0
1	MLY	H	181	1	-	0/7/9/11	0/0/0/0
1	NRQ	H	63	1	-	0/9/31/32	0/2/2/2

The worst 5 of 88 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	63	NRQ	CA2-C2	-9.17	1.38	1.48
1	D	63	NRQ	CA2-C2	-9.16	1.38	1.48
1	C	63	NRQ	CA2-C2	-9.13	1.38	1.48
1	G	63	NRQ	CA2-C2	-9.09	1.38	1.48
1	A	63	NRQ	CA2-C2	-9.07	1.38	1.48

The worst 5 of 63 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	63	NRQ	CG2-CB2-CA2	-11.42	115.39	130.22
1	F	63	NRQ	CG2-CB2-CA2	-11.19	115.69	130.22
1	A	63	NRQ	O2-C2-CA2	-11.08	124.97	130.95
1	H	63	NRQ	CG2-CB2-CA2	-11.06	115.86	130.22
1	G	63	NRQ	CG2-CB2-CA2	-10.63	116.42	130.22

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

9 monomers are involved in 45 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	63	NRQ	8	0
1	B	63	NRQ	6	0
1	C	63	NRQ	7	0
1	D	63	NRQ	7	0
1	E	63	NRQ	5	0
1	F	63	NRQ	5	0
1	G	163	MLY	1	0
1	G	63	NRQ	4	0
1	H	63	NRQ	2	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	223/228 (97%)	-0.43	0 100 100	7, 16, 38, 69	0
1	B	221/228 (96%)	-0.44	0 100 100	8, 19, 35, 51	0
1	C	220/228 (96%)	-0.42	1 (0%) 91 91	9, 21, 43, 76	0
1	D	223/228 (97%)	-0.43	0 100 100	11, 21, 51, 78	0
1	E	211/228 (92%)	0.60	23 (10%) 7 7	33, 63, 86, 157	0
1	F	212/228 (92%)	0.54	15 (7%) 19 18	31, 62, 86, 127	0
1	G	209/228 (91%)	0.70	22 (10%) 8 7	35, 62, 89, 106	0
1	H	193/228 (84%)	0.99	28 (14%) 3 3	39, 68, 97, 120	0
All	All	1712/1824 (93%)	0.12	89 (5%) 31 30	7, 41, 84, 157	0

The worst 5 of 89 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	151	ASP	7.5
1	H	140	TRP	5.8
1	H	99	GLY	5.7
1	H	62	PHE	5.6
1	G	100	GLY	5.4

### 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
1	MLY	A	181	11/12	0.97	0.11	-	10,17,32,32	0
1	MLY	G	181	11/12	0.89	0.14	-	40,52,76,79	0
1	NRQ	F	63	23/24	0.77	0.31	-	85,112,126,143	0
1	NRQ	B	63	23/24	0.95	0.14	-	15,19,31,43	0
1	NRQ	C	63	23/24	0.95	0.14	-	14,18,22,28	0
1	MLY	G	163	11/12	0.90	0.17	-	33,49,58,61	0
1	NRQ	G	63	23/24	0.92	0.17	-	42,59,67,98	0
1	MLY	C	163	11/12	0.98	0.12	-	9,15,35,35	0
1	MLY	E	181	11/12	0.90	0.26	-	49,62,72,73	0
1	MLY	D	181	11/12	0.96	0.12	-	14,18,48,48	0
1	MLY	H	181	11/12	0.91	0.17	-	51,65,77,81	0
1	MLY	A	163	11/12	0.95	0.12	-	8,11,25,25	0
1	MLY	E	163	11/12	0.88	0.16	-	52,65,73,73	0
1	MLY	B	163	11/12	0.95	0.14	-	15,23,33,33	0
1	NRQ	D	63	23/24	0.98	0.11	-	12,18,32,43	0
1	NRQ	H	63	23/24	0.85	0.24	-	52,87,96,104	0
1	NRQ	A	63	23/24	0.97	0.13	-	9,14,22,30	0
1	MLY	F	181	11/12	0.94	0.17	-	37,45,50,54	0
1	MLY	H	163	11/12	0.87	0.19	-	48,61,73,73	0
1	MLY	B	181	11/12	0.96	0.12	-	16,21,41,45	0
1	MLY	F	163	11/12	0.71	0.28	-	77,93,111,113	0
1	MLY	D	163	11/12	0.95	0.15	-	14,23,57,57	0
1	NRQ	E	63	23/24	0.72	0.36	-	132,143,153,161	0
1	MLY	C	181	11/12	0.96	0.14	-	19,32,59,59	0

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.