



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 16, 2020 – 07:31 pm BST

PDB ID : 3RWT  
Title : Crystal structure of circular permuted Red Fluorescent Protein mKate(cp 154-153)  
Authors : Wang, Q.; Sondermann, H.  
Deposited on : 2011-05-09  
Resolution : 3.00 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

<https://www.wwpdb.org/validation/2017/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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

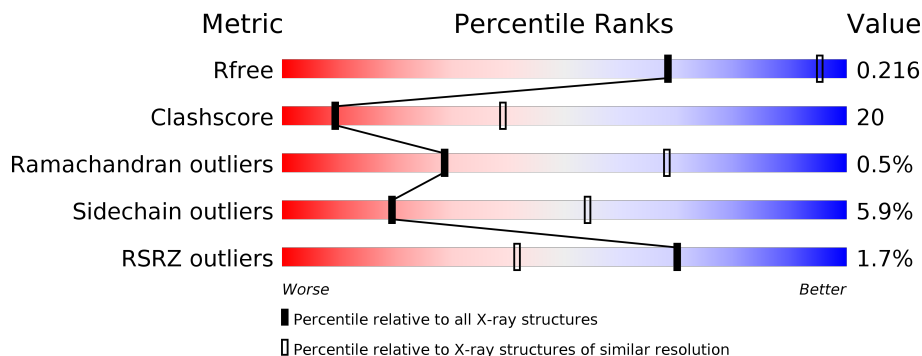
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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.



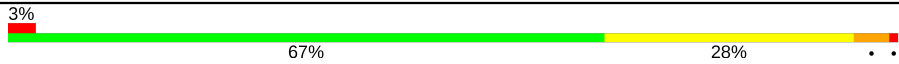

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 respectively. 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	235	 2% 65% 31% 5%
1	B	235	 % 60% 36% .
1	C	235	 % 72% 24% .
1	D	235	 2% 69% 26% .
1	E	235	 % 57% 38% .
1	F	235	 % 68% 28% .

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	G	235	 <p>3% 67% 28% . .</p>
1	H	235	 <p>2% 65% 31% .</p>

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 14977 atoms, of which 0 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,Fluorescent protein FP480.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	F	235	1863	1181	315	353	14	0	0	0
1	A	235	1873	1188	316	354	15	0	2	0
1	B	235	1873	1188	316	354	15	0	2	0
1	C	235	1873	1188	316	354	15	0	2	0
1	D	235	1873	1188	316	354	15	0	2	0
1	E	235	1873	1188	316	354	15	0	2	0
1	G	235	1873	1188	316	354	15	0	2	0
1	H	235	1873	1188	316	354	15	0	2	0

There are 152 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	0	GLU	GLY	engineered mutation	UNP D0VX33
F	1	PHE	GLY	engineered mutation	UNP D0VX33
F	6	SER	ALA	conflict	UNP D0VX33
F	8	MET	LEU	conflict	UNP D0VX33
F	22	LEU	PHE	conflict	UNP D0VX33
F	45	ARG	TYR	conflict	UNP D0VX33
F	80	GLY	-	linker	UNP D0VX33
F	81	GLY	-	linker	UNP D0VX33
F	82	THR	-	linker	UNP D0VX33
F	83	GLY	-	linker	UNP D0VX33
F	84	GLY	-	linker	UNP D0VX33
F	85	SER	-	linker	UNP D0VX33
F	126	MET	GLN	conflict	UNP D0VX33

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
F	?	-	MET	chromophore	UNP D0VX33
F	?	-	TYR	chromophore	UNP D0VX33
F	148	NRQ	GLY	chromophore	UNP D0VX33
F	152	LYS	HIS	conflict	UNP D0VX33
F	165	PHE	TRP	conflict	UNP D0VX33
F	228	SER	HIS	conflict	UNP D0VX33
A	0	GLU	GLY	engineered mutation	UNP D0VX33
A	1	PHE	GLY	engineered mutation	UNP D0VX33
A	6	SER	ALA	conflict	UNP D0VX33
A	8	MET	LEU	conflict	UNP D0VX33
A	22	LEU	PHE	conflict	UNP D0VX33
A	45	ARG	TYR	conflict	UNP D0VX33
A	80	GLY	-	linker	UNP D0VX33
A	81	GLY	-	linker	UNP D0VX33
A	82	THR	-	linker	UNP D0VX33
A	83	GLY	-	linker	UNP D0VX33
A	84	GLY	-	linker	UNP D0VX33
A	85	SER	-	linker	UNP D0VX33
A	126	MET	GLN	conflict	UNP D0VX33
A	?	-	MET	chromophore	UNP D0VX33
A	?	-	TYR	chromophore	UNP D0VX33
A	148	NRQ	GLY	chromophore	UNP D0VX33
A	152	LYS	HIS	conflict	UNP D0VX33
A	165	PHE	TRP	conflict	UNP D0VX33
A	228	SER	HIS	conflict	UNP D0VX33
B	0	GLU	GLY	engineered mutation	UNP D0VX33
B	1	PHE	GLY	engineered mutation	UNP D0VX33
B	6	SER	ALA	conflict	UNP D0VX33
B	8	MET	LEU	conflict	UNP D0VX33
B	22	LEU	PHE	conflict	UNP D0VX33
B	45	ARG	TYR	conflict	UNP D0VX33
B	80	GLY	-	linker	UNP D0VX33
B	81	GLY	-	linker	UNP D0VX33
B	82	THR	-	linker	UNP D0VX33
B	83	GLY	-	linker	UNP D0VX33
B	84	GLY	-	linker	UNP D0VX33
B	85	SER	-	linker	UNP D0VX33
B	126	MET	GLN	conflict	UNP D0VX33
B	?	-	MET	chromophore	UNP D0VX33
B	?	-	TYR	chromophore	UNP D0VX33
B	148	NRQ	GLY	chromophore	UNP D0VX33
B	152	LYS	HIS	conflict	UNP D0VX33

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	165	PHE	TRP	conflict	UNP D0VX33
B	228	SER	HIS	conflict	UNP D0VX33
C	0	GLU	GLY	engineered mutation	UNP D0VX33
C	1	PHE	GLY	engineered mutation	UNP D0VX33
C	6	SER	ALA	conflict	UNP D0VX33
C	8	MET	LEU	conflict	UNP D0VX33
C	22	LEU	PHE	conflict	UNP D0VX33
C	45	ARG	TYR	conflict	UNP D0VX33
C	80	GLY	-	linker	UNP D0VX33
C	81	GLY	-	linker	UNP D0VX33
C	82	THR	-	linker	UNP D0VX33
C	83	GLY	-	linker	UNP D0VX33
C	84	GLY	-	linker	UNP D0VX33
C	85	SER	-	linker	UNP D0VX33
C	126	MET	GLN	conflict	UNP D0VX33
C	?	-	MET	chromophore	UNP D0VX33
C	?	-	TYR	chromophore	UNP D0VX33
C	148	NRQ	GLY	chromophore	UNP D0VX33
C	152	LYS	HIS	conflict	UNP D0VX33
C	165	PHE	TRP	conflict	UNP D0VX33
C	228	SER	HIS	conflict	UNP D0VX33
D	0	GLU	GLY	engineered mutation	UNP D0VX33
D	1	PHE	GLY	engineered mutation	UNP D0VX33
D	6	SER	ALA	conflict	UNP D0VX33
D	8	MET	LEU	conflict	UNP D0VX33
D	22	LEU	PHE	conflict	UNP D0VX33
D	45	ARG	TYR	conflict	UNP D0VX33
D	80	GLY	-	linker	UNP D0VX33
D	81	GLY	-	linker	UNP D0VX33
D	82	THR	-	linker	UNP D0VX33
D	83	GLY	-	linker	UNP D0VX33
D	84	GLY	-	linker	UNP D0VX33
D	85	SER	-	linker	UNP D0VX33
D	126	MET	GLN	conflict	UNP D0VX33
D	?	-	MET	chromophore	UNP D0VX33
D	?	-	TYR	chromophore	UNP D0VX33
D	148	NRQ	GLY	chromophore	UNP D0VX33
D	152	LYS	HIS	conflict	UNP D0VX33
D	165	PHE	TRP	conflict	UNP D0VX33
D	228	SER	HIS	conflict	UNP D0VX33
E	0	GLU	GLY	engineered mutation	UNP D0VX33
E	1	PHE	GLY	engineered mutation	UNP D0VX33

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	6	SER	ALA	conflict	UNP D0VX33
E	8	MET	LEU	conflict	UNP D0VX33
E	22	LEU	PHE	conflict	UNP D0VX33
E	45	ARG	TYR	conflict	UNP D0VX33
E	80	GLY	-	linker	UNP D0VX33
E	81	GLY	-	linker	UNP D0VX33
E	82	THR	-	linker	UNP D0VX33
E	83	GLY	-	linker	UNP D0VX33
E	84	GLY	-	linker	UNP D0VX33
E	85	SER	-	linker	UNP D0VX33
E	126	MET	GLN	conflict	UNP D0VX33
E	?	-	MET	chromophore	UNP D0VX33
E	?	-	TYR	chromophore	UNP D0VX33
E	148	NRQ	GLY	chromophore	UNP D0VX33
E	152	LYS	HIS	conflict	UNP D0VX33
E	165	PHE	TRP	conflict	UNP D0VX33
E	228	SER	HIS	conflict	UNP D0VX33
G	0	GLU	GLY	engineered mutation	UNP D0VX33
G	1	PHE	GLY	engineered mutation	UNP D0VX33
G	6	SER	ALA	conflict	UNP D0VX33
G	8	MET	LEU	conflict	UNP D0VX33
G	22	LEU	PHE	conflict	UNP D0VX33
G	45	ARG	TYR	conflict	UNP D0VX33
G	80	GLY	-	linker	UNP D0VX33
G	81	GLY	-	linker	UNP D0VX33
G	82	THR	-	linker	UNP D0VX33
G	83	GLY	-	linker	UNP D0VX33
G	84	GLY	-	linker	UNP D0VX33
G	85	SER	-	linker	UNP D0VX33
G	126	MET	GLN	conflict	UNP D0VX33
G	?	-	MET	chromophore	UNP D0VX33
G	?	-	TYR	chromophore	UNP D0VX33
G	148	NRQ	GLY	chromophore	UNP D0VX33
G	152	LYS	HIS	conflict	UNP D0VX33
G	165	PHE	TRP	conflict	UNP D0VX33
G	228	SER	HIS	conflict	UNP D0VX33
H	0	GLU	GLY	engineered mutation	UNP D0VX33
H	1	PHE	GLY	engineered mutation	UNP D0VX33
H	6	SER	ALA	conflict	UNP D0VX33
H	8	MET	LEU	conflict	UNP D0VX33
H	22	LEU	PHE	conflict	UNP D0VX33
H	45	ARG	TYR	conflict	UNP D0VX33

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
H	80	GLY	-	linker	UNP D0VX33
H	81	GLY	-	linker	UNP D0VX33
H	82	THR	-	linker	UNP D0VX33
H	83	GLY	-	linker	UNP D0VX33
H	84	GLY	-	linker	UNP D0VX33
H	85	SER	-	linker	UNP D0VX33
H	126	MET	GLN	conflict	UNP D0VX33
H	?	-	MET	chromophore	UNP D0VX33
H	?	-	TYR	chromophore	UNP D0VX33
H	148	NRQ	GLY	chromophore	UNP D0VX33
H	152	LYS	HIS	conflict	UNP D0VX33
H	165	PHE	TRP	conflict	UNP D0VX33
H	228	SER	HIS	conflict	UNP D0VX33

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Mg 1 1	0	0

- Molecule 3 is water.

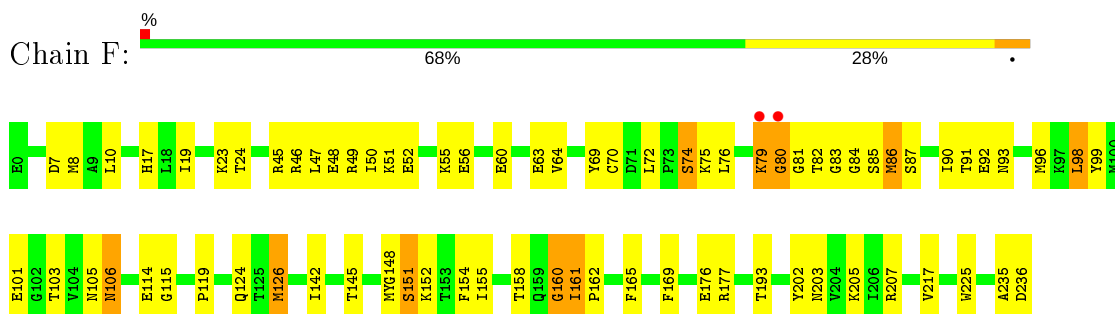
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	F	1	Total O 1 1	0	0
3	A	1	Total O 1 1	0	0



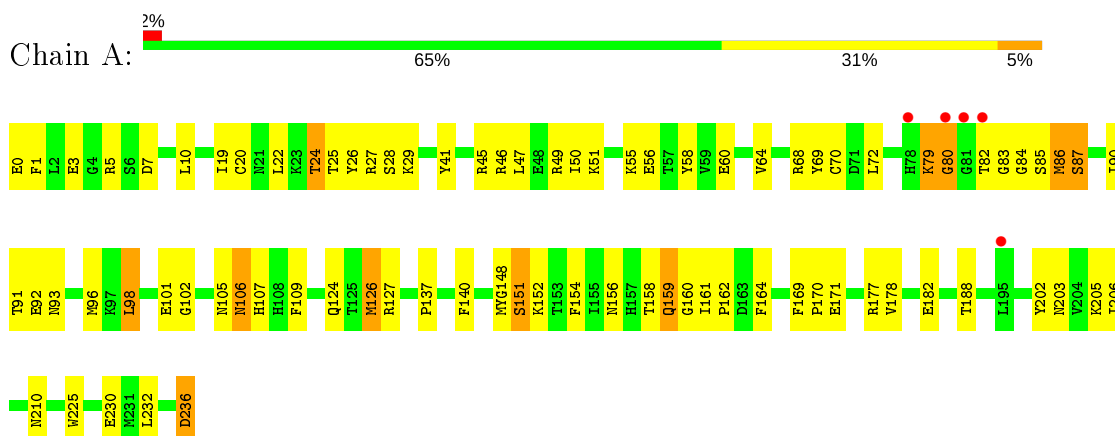
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes 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.

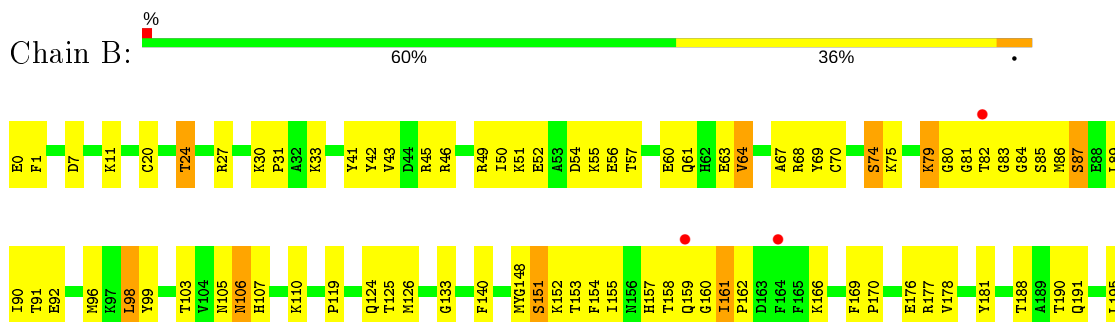
- Molecule 1: Fluorescent protein FP480,Fluorescent protein FP480



- Molecule 1: Fluorescent protein FP480,Fluorescent protein FP480



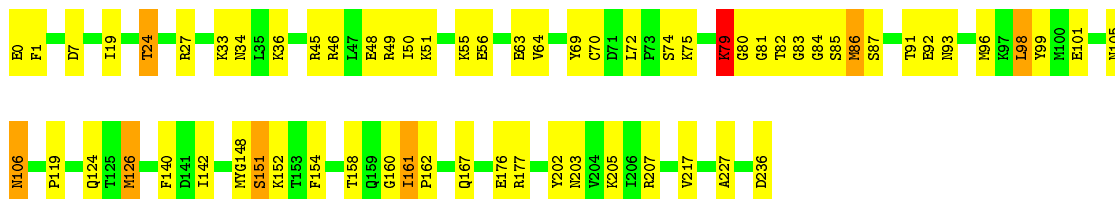
- Molecule 1: Fluorescent protein FP480,Fluorescent protein FP480





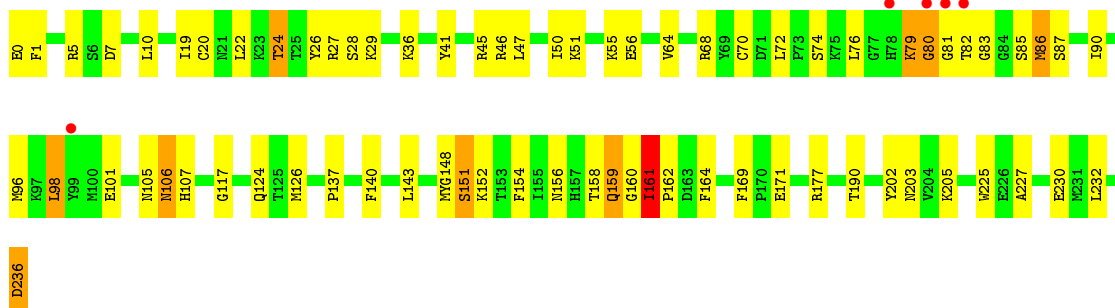
- Molecule 1: Fluorescent protein FP480, Fluorescent protein FP480

Chain C: 72% 24%



- Molecule 1: Fluorescent protein FP480, Fluorescent protein FP480

Chain D: 69% 26%



- Molecule 1: Fluorescent protein FP480, Fluorescent protein FP480

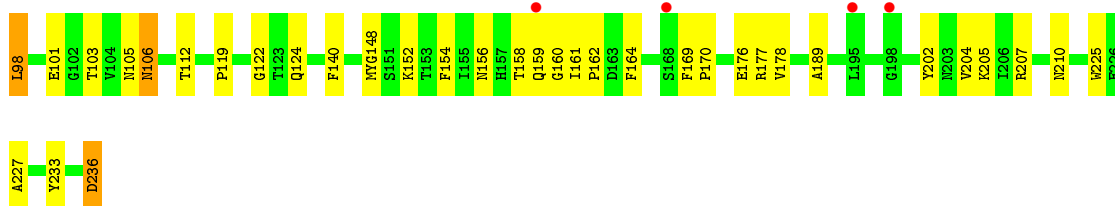
Chain E: 57% 38%



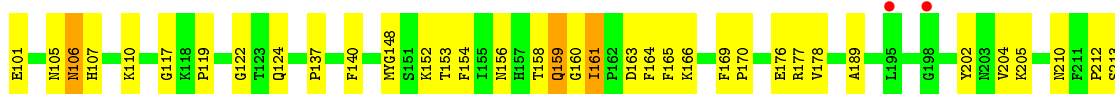
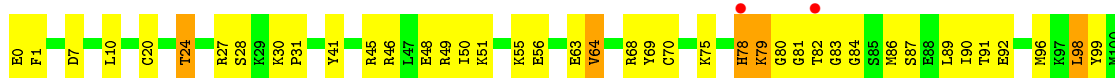
- Molecule 1: Fluorescent protein FP480, Fluorescent protein FP480

Chain G: 67% 28%





- Molecule 1: Fluorescent protein FP480, Fluorescent protein FP480



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	71.46Å 71.44Å 367.73Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.90 – 3.00 38.98 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.7 (38.90-3.00) 99.7 (38.98-3.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.36 (at 3.01Å)	Xtrriage
Refinement program	PHENIX 1.6.4_486	Depositor
R, $R_{free}$	0.217 , 0.278 0.209 , 0.216	Depositor DCC
$R_{free}$ test set	1951 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	71.8	Xtrriage
Anisotropy	0.180	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 39.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.478 for k,h,-l	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	14977	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	95.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.94% 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, 0.333 respectively for untwinned datasets, and 0.375, 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: MG, NRQ

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	0.54	0/1897	0.75	0/2555
1	B	0.49	0/1897	0.75	0/2555
1	C	0.52	0/1897	0.76	1/2555 (0.0%)
1	D	0.54	0/1897	0.77	0/2555
1	E	0.51	0/1897	0.73	0/2555
1	F	0.54	0/1881	0.79	3/2534 (0.1%)
1	G	0.50	1/1897 (0.1%)	0.77	2/2555 (0.1%)
1	H	0.50	1/1897 (0.1%)	0.74	0/2555
All	All	0.52	2/15160 (0.0%)	0.76	6/20419 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	F	0	1
1	G	0	1
All	All	0	2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	H	20	CYS	CB-SG	-6.39	1.71	1.82
1	G	20	CYS	CB-SG	-6.25	1.71	1.82

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	81	GLY	N-CA-C	-7.72	93.80	113.10

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	81	GLY	N-CA-C	-7.17	95.18	113.10
1	F	79	LYS	N-CA-C	-6.22	94.21	111.00
1	G	80	GLY	N-CA-C	-6.17	97.69	113.10
1	F	80	GLY	N-CA-C	-5.08	100.39	113.10

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	F	160	GLY	Mainchain
1	G	80	GLY	Peptide

## 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	1873	0	1840	85	0
1	B	1873	0	1840	91	0
1	C	1873	0	1840	68	0
1	D	1873	0	1840	67	0
1	E	1873	0	1840	98	0
1	F	1863	0	1825	61	0
1	G	1873	0	1840	72	0
1	H	1873	0	1840	82	0
2	A	1	0	0	0	0
3	A	1	0	0	1	0
3	F	1	0	0	0	0
All	All	14977	0	14705	586	0

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 20.

The worst 5 of 586 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:B:79:LYS:HG2	1:B:80:GLY:N	1.57	1.17

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:79:LYS:HG2	1:C:80:GLY:H	0.98	1.15
1:B:79:LYS:HG2	1:B:80:GLY:H	1.10	1.11
1:E:79:LYS:HG2	1:E:80:GLY:H	1.10	1.07
1:E:79:LYS:CG	1:E:80:GLY:H	1.66	1.07

There are no symmetry-related clashes.

## 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	232/235 (99%)	213 (92%)	17 (7%)	2 (1%)	17	55
1	B	232/235 (99%)	218 (94%)	14 (6%)	0	100	100
1	C	232/235 (99%)	217 (94%)	14 (6%)	1 (0%)	34	72
1	D	232/235 (99%)	215 (93%)	14 (6%)	3 (1%)	12	45
1	E	232/235 (99%)	217 (94%)	14 (6%)	1 (0%)	34	72
1	F	230/235 (98%)	213 (93%)	16 (7%)	1 (0%)	34	72
1	G	232/235 (99%)	217 (94%)	14 (6%)	1 (0%)	34	72
1	H	232/235 (99%)	219 (94%)	12 (5%)	1 (0%)	34	72
All	All	1854/1880 (99%)	1729 (93%)	115 (6%)	10 (0%)	29	68

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	159	GLN
1	C	79	LYS
1	D	80	GLY
1	A	80	GLY
1	D	161	ILE

### 5.3.2 Protein sidechains [i](#)

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	202/200 (101%)	190 (94%)	12 (6%)	19	54
1	B	202/200 (101%)	188 (93%)	14 (7%)	15	48
1	C	202/200 (101%)	189 (94%)	13 (6%)	17	51
1	D	202/200 (101%)	189 (94%)	13 (6%)	17	51
1	E	202/200 (101%)	187 (93%)	15 (7%)	13	44
1	F	200/200 (100%)	190 (95%)	10 (5%)	24	60
1	G	202/200 (101%)	189 (94%)	13 (6%)	17	51
1	H	202/200 (101%)	190 (94%)	12 (6%)	19	54
All	All	1614/1600 (101%)	1512 (94%)	102 (6%)	19	51

5 of 102 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	151	SER
1	D	106	ASN
1	H	86	MET
1	C	161	ILE
1	D	74	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 34 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	34	ASN
1	C	203	ASN
1	H	34	ASN
1	C	108	HIS
1	A	62	HIS

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.



## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

8 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	NRQ	A	148	1	23,24,25	4.69	6 (26%)	23,32,34	3.36	8 (34%)
1	NRQ	G	148	1	23,24,25	4.29	4 (17%)	23,32,34	3.75	7 (30%)
1	NRQ	F	148	1	23,24,25	4.45	4 (17%)	23,32,34	3.18	5 (21%)
1	NRQ	D	148	1	23,24,25	4.63	6 (26%)	23,32,34	3.38	9 (39%)
1	NRQ	B	148	1	23,24,25	4.49	6 (26%)	23,32,34	3.33	8 (34%)
1	NRQ	C	148	1	23,24,25	4.50	4 (17%)	23,32,34	3.43	6 (26%)
1	NRQ	H	148	1	23,24,25	4.27	4 (17%)	23,32,34	3.65	7 (30%)
1	NRQ	E	148	1	23,24,25	4.48	6 (26%)	23,32,34	3.27	8 (34%)

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	NRQ	A	148	1	-	1/9/31/32	0/2/2/2
1	NRQ	G	148	1	-	3/9/31/32	0/2/2/2
1	NRQ	F	148	1	-	2/9/31/32	0/2/2/2
1	NRQ	D	148	1	-	1/9/31/32	0/2/2/2
1	NRQ	B	148	1	-	2/9/31/32	0/2/2/2
1	NRQ	C	148	1	-	2/9/31/32	0/2/2/2
1	NRQ	H	148	1	-	3/9/31/32	0/2/2/2
1	NRQ	E	148	1	-	3/9/31/32	0/2/2/2

The worst 5 of 40 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	148	NRQ	CB2-CA2	19.93	1.51	1.35

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	148	NRQ	CB2-CA2	19.35	1.51	1.35
1	C	148	NRQ	CB2-CA2	19.13	1.51	1.35
1	F	148	NRQ	CB2-CA2	19.00	1.51	1.35
1	E	148	NRQ	CB2-CA2	18.27	1.50	1.35

The worst 5 of 58 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	H	148	NRQ	O2-C2-CA2	-12.41	123.99	130.96
1	G	148	NRQ	O2-C2-CA2	-12.12	124.15	130.96
1	C	148	NRQ	O2-C2-CA2	-10.13	125.27	130.96
1	G	148	NRQ	CA2-C2-N3	10.09	108.14	103.37
1	D	148	NRQ	O2-C2-CA2	-9.92	125.39	130.96

There are no chirality outliers.

5 of 17 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	148	NRQ	C1-CA1-CB1-CG1
1	G	148	NRQ	CA1-CB1-CG1-SD
1	F	148	NRQ	CA1-CB1-CG1-SD
1	D	148	NRQ	C1-CA1-CB1-CG1
1	C	148	NRQ	CA1-CB1-CG1-SD

There are no ring outliers.

8 monomers are involved in 21 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	148	NRQ	2	0
1	G	148	NRQ	1	0
1	F	148	NRQ	5	0
1	D	148	NRQ	2	0
1	B	148	NRQ	3	0
1	C	148	NRQ	4	0
1	H	148	NRQ	1	0
1	E	148	NRQ	3	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	234/235 (99%)	-0.14	5 (2%) 63 34	43, 94, 134, 150	0
1	B	234/235 (99%)	-0.12	3 (1%) 77 51	44, 95, 145, 209	0
1	C	234/235 (99%)	-0.24	0 100 100	46, 87, 120, 155	0
1	D	234/235 (99%)	-0.16	5 (2%) 63 34	42, 94, 132, 154	0
1	E	234/235 (99%)	-0.12	3 (1%) 77 51	47, 94, 140, 206	0
1	F	234/235 (99%)	-0.23	2 (0%) 84 63	46, 87, 122, 179	0
1	G	234/235 (99%)	-0.05	8 (3%) 45 19	47, 107, 155, 199	0
1	H	234/235 (99%)	-0.08	5 (2%) 63 34	46, 108, 153, 170	0
All	All	1872/1880 (99%)	-0.14	31 (1%) 70 41	42, 94, 144, 209	0

The worst 5 of 31 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	81	GLY	7.2
1	A	81	GLY	6.1
1	H	82	THR	3.7
1	H	195	LEU	3.6
1	A	195	LEU	3.5

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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. 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	B-factors(Å <sup>2</sup> )	Q<0.9
1	NRQ	G	148	23/24	0.93	0.23	54,74,90,103	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
1	NRQ	A	148	23/24	0.94	0.18	44,59,96,110	0
1	NRQ	D	148	23/24	0.94	0.24	49,64,91,122	0
1	NRQ	H	148	23/24	0.95	0.26	55,72,91,108	0
1	NRQ	B	148	23/24	0.96	0.18	61,71,84,135	0
1	NRQ	C	148	23/24	0.96	0.20	53,65,80,128	0
1	NRQ	F	148	23/24	0.96	0.18	51,66,76,118	0
1	NRQ	E	148	23/24	0.96	0.20	62,70,85,138	0

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

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. 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	B-factors( $\text{\AA}^2$ )	Q<0.9
2	MG	A	237	1/1	0.92	0.51	68,68,68,68	0

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

There are no such residues in this entry.