



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 8, 2016 – 08:39 PM EDT

PDB ID : 5EXC  
Title : Photoconverted red fluorescent protein DendRFP  
Authors : Pletnev, V.Z.; Pletneva, N.V.; Pletnev, S.V.  
Deposited on : 2015-11-23  
Resolution : 2.14 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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027939  
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) : rb-20027939

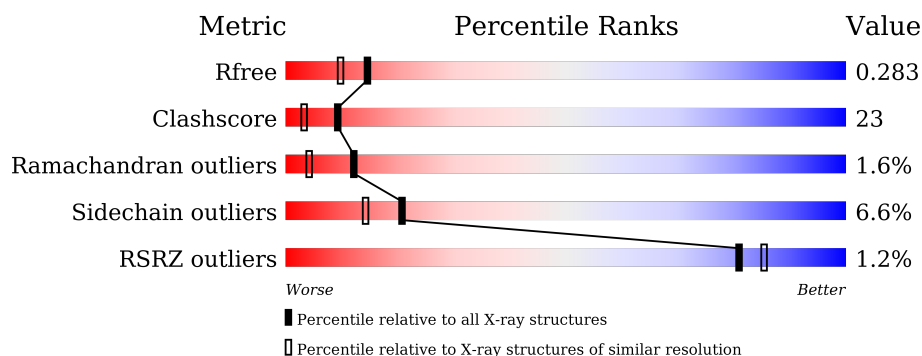
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.14 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.











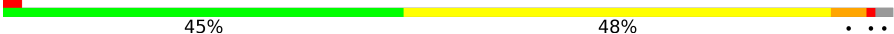

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	91344	1693 (2.16-2.12)
Clashscore	102246	1824 (2.16-2.12)
Ramachandran outliers	100387	1798 (2.16-2.12)
Sidechain outliers	100360	1798 (2.16-2.12)
RSRZ outliers	91569	1699 (2.16-2.12)

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	62	<div> <div>60%</div> <div>32%</div> <div>5%</div> <div>•</div> </div>
1	B	62	<div> <div>66%</div> <div>26%</div> <div>5%</div> <div>•</div> </div>
1	C	62	<div> <div>61%</div> <div>31%</div> <div>•</div> <div>5%</div> </div>
1	D	62	<div> <div>55%</div> <div>35%</div> <div>5%</div> <div>5%</div> </div>
1	E	62	<div> <div>55%</div> <div>37%</div> <div>•</div> <div>5%</div> </div>
1	F	62	<div> <div>3%</div> <div>48%</div> <div>40%</div> <div>5%</div> <div>•</div> <div>5%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	62	
1	H	62	
2	I	170	
2	J	170	
2	K	170	
2	L	170	
2	M	170	
2	N	170	
2	O	170	
2	P	170	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	RC7	I	64	X	-	-	-
2	RC7	J	64	X	-	X	-
2	RC7	K	64	X	-	-	-
2	RC7	L	64	X	-	-	-
2	RC7	M	64	X	-	-	-
2	RC7	N	64	X	-	-	-
2	RC7	O	64	X	-	X	-
2	RC7	P	64	X	-	-	-

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 15333 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 Green fluorescent protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	60	Total	C	N	O	S	0	1	0
			471	298	79	92	2			
1	B	60	Total	C	N	O	S	0	0	0
			462	293	78	89	2			
1	C	59	Total	C	N	O	S	0	0	0
			454	289	76	87	2			
1	D	59	Total	C	N	O	S	0	0	0
			454	289	76	87	2			
1	E	59	Total	C	N	O	S	0	1	0
			463	295	78	88	2			
1	F	59	Total	C	N	O	S	1	0	0
			454	289	76	87	2			
1	G	59	Total	C	N	O	S	1	0	0
			454	289	76	87	2			
1	H	59	Total	C	N	O	S	0	1	0
			463	297	76	88	2			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP Q8T6U0
A	1	GLY	-	expression tag	UNP Q8T6U0
B	0	MET	-	initiating methionine	UNP Q8T6U0
B	1	GLY	-	expression tag	UNP Q8T6U0
C	0	MET	-	initiating methionine	UNP Q8T6U0
C	1	GLY	-	expression tag	UNP Q8T6U0
D	0	MET	-	initiating methionine	UNP Q8T6U0
D	1	GLY	-	expression tag	UNP Q8T6U0
E	0	MET	-	initiating methionine	UNP Q8T6U0
E	1	GLY	-	expression tag	UNP Q8T6U0
F	0	MET	-	initiating methionine	UNP Q8T6U0
F	1	GLY	-	expression tag	UNP Q8T6U0
G	0	MET	-	initiating methionine	UNP Q8T6U0

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
G	1	GLY	-	expression tag	UNP Q8T6U0
H	0	MET	-	initiating methionine	UNP Q8T6U0
H	1	GLY	-	expression tag	UNP Q8T6U0

- Molecule 2 is a protein called Green fluorescent protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	I	165	Total	C	N	O	S	0	0	0
			1368	879	231	251	7			
2	J	167	Total	C	N	O	S	1	1	0
			1399	897	241	254	7			
2	K	166	Total	C	N	O	S	0	1	0
			1384	888	235	254	7			
2	L	166	Total	C	N	O	S	1	1	0
			1386	889	236	254	7			
2	M	162	Total	C	N	O	S	1	2	0
			1367	881	229	250	7			
2	N	162	Total	C	N	O	S	1	1	0
			1356	872	228	249	7			
2	O	166	Total	C	N	O	S	0	0	0
			1378	885	234	252	7			
2	P	162	Total	C	N	O	S	0	0	0
			1348	868	226	247	7			

There are 72 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
I	64	RC7	-	chromophore	UNP Q8T6U0
I	226	GLY	-	expression tag	UNP Q8T6U0
I	227	SER	-	expression tag	UNP Q8T6U0
I	228	HIS	-	expression tag	UNP Q8T6U0
I	229	HIS	-	expression tag	UNP Q8T6U0
I	230	HIS	-	expression tag	UNP Q8T6U0
I	231	HIS	-	expression tag	UNP Q8T6U0
I	232	HIS	-	expression tag	UNP Q8T6U0
I	233	HIS	-	expression tag	UNP Q8T6U0
J	64	RC7	-	chromophore	UNP Q8T6U0
J	226	GLY	-	expression tag	UNP Q8T6U0
J	227	SER	-	expression tag	UNP Q8T6U0
J	228	HIS	-	expression tag	UNP Q8T6U0
J	229	HIS	-	expression tag	UNP Q8T6U0
J	230	HIS	-	expression tag	UNP Q8T6U0

*Continued on next page...*

*Continued from previous page...*

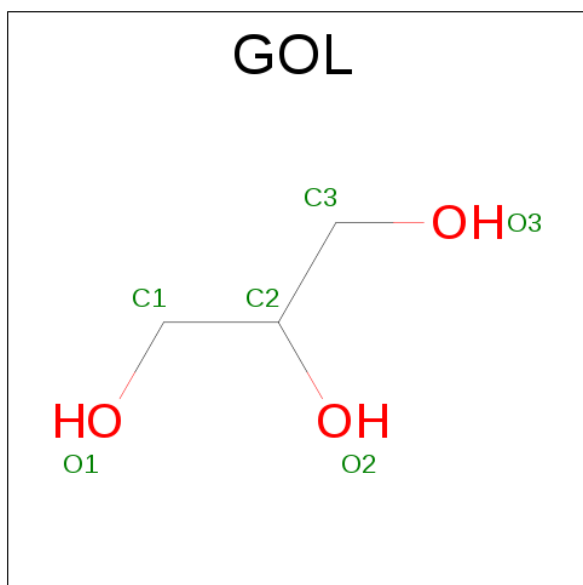
Chain	Residue	Modelled	Actual	Comment	Reference
J	231	HIS	-	expression tag	UNP Q8T6U0
J	232	HIS	-	expression tag	UNP Q8T6U0
J	233	HIS	-	expression tag	UNP Q8T6U0
K	64	RC7	-	chromophore	UNP Q8T6U0
K	226	GLY	-	expression tag	UNP Q8T6U0
K	227	SER	-	expression tag	UNP Q8T6U0
K	228	HIS	-	expression tag	UNP Q8T6U0
K	229	HIS	-	expression tag	UNP Q8T6U0
K	230	HIS	-	expression tag	UNP Q8T6U0
K	231	HIS	-	expression tag	UNP Q8T6U0
K	232	HIS	-	expression tag	UNP Q8T6U0
K	233	HIS	-	expression tag	UNP Q8T6U0
L	64	RC7	-	chromophore	UNP Q8T6U0
L	226	GLY	-	expression tag	UNP Q8T6U0
L	227	SER	-	expression tag	UNP Q8T6U0
L	228	HIS	-	expression tag	UNP Q8T6U0
L	229	HIS	-	expression tag	UNP Q8T6U0
L	230	HIS	-	expression tag	UNP Q8T6U0
L	231	HIS	-	expression tag	UNP Q8T6U0
L	232	HIS	-	expression tag	UNP Q8T6U0
L	233	HIS	-	expression tag	UNP Q8T6U0
M	64	RC7	-	chromophore	UNP Q8T6U0
M	226	GLY	-	expression tag	UNP Q8T6U0
M	227	SER	-	expression tag	UNP Q8T6U0
M	228	HIS	-	expression tag	UNP Q8T6U0
M	229	HIS	-	expression tag	UNP Q8T6U0
M	230	HIS	-	expression tag	UNP Q8T6U0
M	231	HIS	-	expression tag	UNP Q8T6U0
M	232	HIS	-	expression tag	UNP Q8T6U0
M	233	HIS	-	expression tag	UNP Q8T6U0
N	64	RC7	-	chromophore	UNP Q8T6U0
N	226	GLY	-	expression tag	UNP Q8T6U0
N	227	SER	-	expression tag	UNP Q8T6U0
N	228	HIS	-	expression tag	UNP Q8T6U0
N	229	HIS	-	expression tag	UNP Q8T6U0
N	230	HIS	-	expression tag	UNP Q8T6U0
N	231	HIS	-	expression tag	UNP Q8T6U0
N	232	HIS	-	expression tag	UNP Q8T6U0
N	233	HIS	-	expression tag	UNP Q8T6U0
O	64	RC7	-	chromophore	UNP Q8T6U0
O	226	GLY	-	expression tag	UNP Q8T6U0
O	227	SER	-	expression tag	UNP Q8T6U0

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
O	228	HIS	-	expression tag	UNP Q8T6U0
O	229	HIS	-	expression tag	UNP Q8T6U0
O	230	HIS	-	expression tag	UNP Q8T6U0
O	231	HIS	-	expression tag	UNP Q8T6U0
O	232	HIS	-	expression tag	UNP Q8T6U0
O	233	HIS	-	expression tag	UNP Q8T6U0
P	64	RC7	-	chromophore	UNP Q8T6U0
P	226	GLY	-	expression tag	UNP Q8T6U0
P	227	SER	-	expression tag	UNP Q8T6U0
P	228	HIS	-	expression tag	UNP Q8T6U0
P	229	HIS	-	expression tag	UNP Q8T6U0
P	230	HIS	-	expression tag	UNP Q8T6U0
P	231	HIS	-	expression tag	UNP Q8T6U0
P	232	HIS	-	expression tag	UNP Q8T6U0
P	233	HIS	-	expression tag	UNP Q8T6U0

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	L	1	Total	C	O	0	0
			6	3	3		
3	E	1	Total	C	O	0	0
			6	3	3		
3	G	1	Total	C	O	0	0
			6	3	3		

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	P	2	Total Mg 2 2	0	0
4	G	1	Total Mg 1 1	0	0
4	E	2	Total Mg 2 2	0	0
4	H	3	Total Mg 3 3	0	0
4	N	1	Total Mg 1 1	0	0
4	O	1	Total Mg 1 1	0	0
4	F	1	Total Mg 1 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	18	Total O 18 18	0	0
5	I	62	Total O 62 62	0	0
5	B	21	Total O 21 21	0	0
5	J	62	Total O 62 62	0	0
5	C	16	Total O 16 16	0	0
5	K	59	Total O 59 59	0	0
5	D	17	Total O 17 17	0	0
5	L	54	Total O 54 54	0	0
5	E	22	Total O 22 22	0	0
5	M	71	Total O 71 71	0	0
5	F	19	Total O 19 19	0	0
5	N	47	Total O 47 47	0	0
5	G	20	Total O 20 20	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	O	57	Total 57	O 57	0	0
5	H	19	Total 19	O 19	0	0
5	P	79	Total 79	O 79	0	0

### 3 Residue-property plots

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.

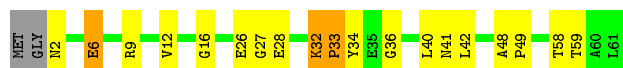
- Molecule 1: Green fluorescent protein

Chain A: 



- Molecule 1: Green fluorescent protein

Chain B: 



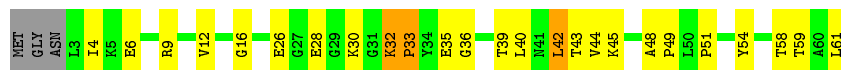
- Molecule 1: Green fluorescent protein

Chain C: 



- Molecule 1: Green fluorescent protein

Chain D: 



- Molecule 1: Green fluorescent protein

Chain E: 



- Molecule 1: Green fluorescent protein

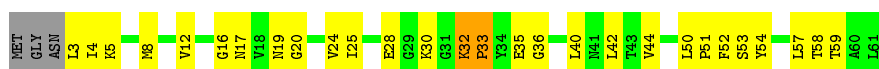
Chain F: 



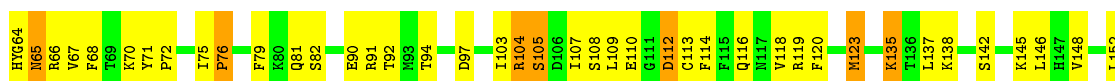
- Molecule 1: Green fluorescent protein



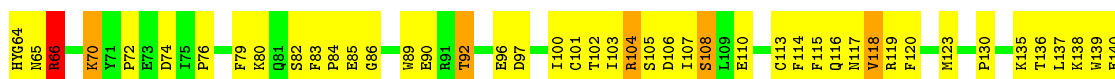
- Molecule 1: Green fluorescent protein



- Molecule 2: Green fluorescent protein



- Molecule 2: Green fluorescent protein

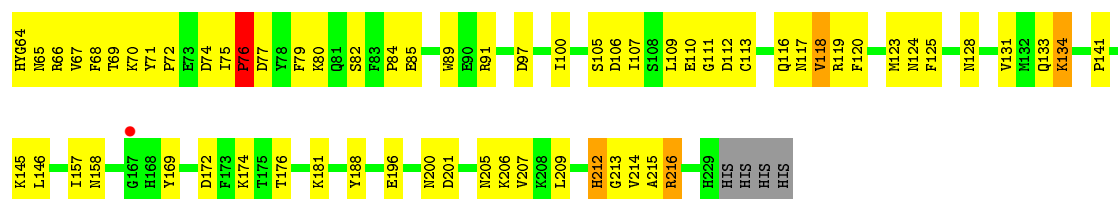


- Molecule 2: Green fluorescent protein

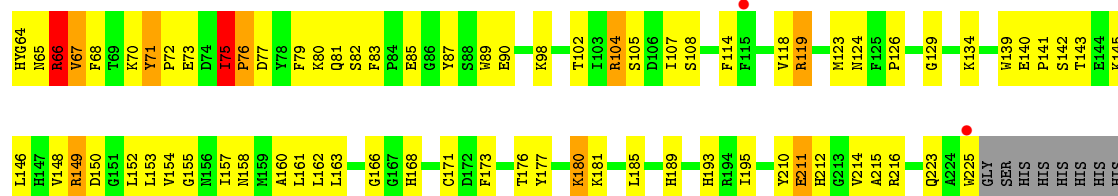


- Molecule 2: Green fluorescent protein

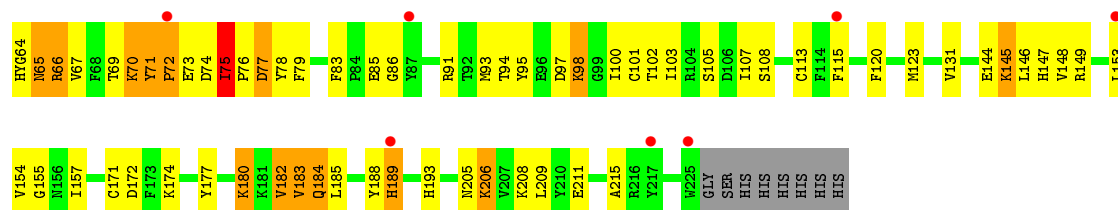




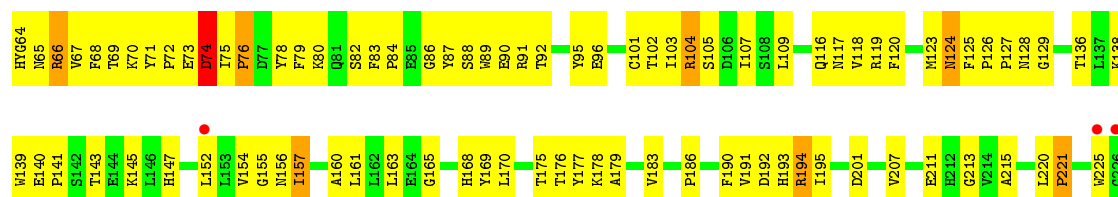
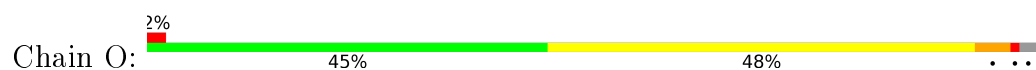
- Molecule 2: Green fluorescent protein



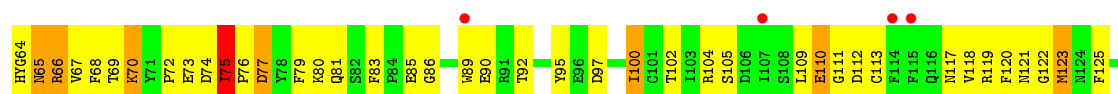
- Molecule 2: Green fluorescent protein

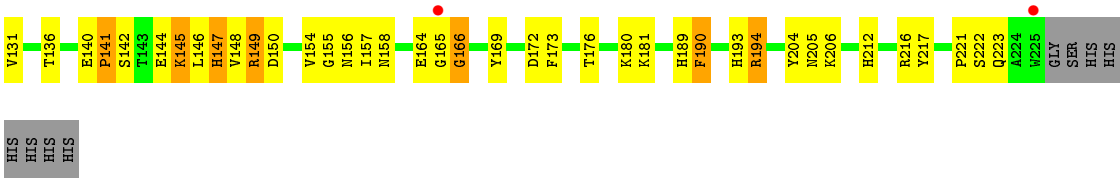


- Molecule 2: Green fluorescent protein



- Molecule 2: Green fluorescent protein





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	57.42Å 69.52Å 124.48Å 89.88° 89.94° 65.57°	Depositor
Resolution (Å)	29.87 – 2.14 29.87 – 2.14	Depositor EDS
% Data completeness (in resolution range)	93.2 (29.87-2.14) 93.3 (29.87-2.14)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.07 (at 2.14Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.195 , 0.279 0.196 , 0.283	Depositor DCC
$R_{free}$ test set	1817 reflections (2.05%)	DCC
Wilson B-factor (Å <sup>2</sup> )	24.6	Xtriage
Anisotropy	0.098	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.26 , 14.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	0.470 for h,h-k,-l 0.469 for -h,-k,l 0.467 for -h,-h+k,-l	Xtriage
Reported twinning fraction	0.285 for H, K, L 0.253 for H, H-K, -L 0.218 for -H, -H+K, -L 0.244 for -H, -K, L	Depositor
Outliers	0 of 90214 reflections	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	15333	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.08% 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 ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, MG, RC7, NLW

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	0.86	0/470	0.94	1/634 (0.2%)
1	B	0.88	1/461 (0.2%)	0.91	0/622
1	C	0.81	0/453	0.93	1/611 (0.2%)
1	D	0.76	0/453	0.96	2/611 (0.3%)
1	E	0.88	1/462 (0.2%)	1.05	0/622
1	F	0.94	1/453 (0.2%)	0.97	3/611 (0.5%)
1	G	1.06	2/453 (0.4%)	1.03	3/611 (0.5%)
1	H	0.80	1/466 (0.2%)	0.95	2/629 (0.3%)
2	I	0.99	2/1384 (0.1%)	1.03	8/1873 (0.4%)
2	J	2.45	5/1417 (0.4%)	1.05	8/1917 (0.4%)
2	K	0.99	2/1401 (0.1%)	1.02	6/1896 (0.3%)
2	L	0.97	3/1403 (0.2%)	0.99	6/1899 (0.3%)
2	M	0.89	2/1383 (0.1%)	0.99	5/1872 (0.3%)
2	N	0.81	1/1371 (0.1%)	1.03	2/1856 (0.1%)
2	O	0.78	2/1395 (0.1%)	0.95	3/1888 (0.2%)
2	P	0.89	2/1363 (0.1%)	0.98	4/1845 (0.2%)
All	All	1.14	25/14788 (0.2%)	1.00	54/19997 (0.3%)

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	G	0	1
2	I	1	0
2	J	1	3
2	K	1	2
2	L	1	0
2	M	1	0
2	N	1	2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	#Chirality outliers	#Planarity outliers
2	O	1	0
2	P	1	0
All	All	8	8

The worst 5 of 25 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	J	110	GLU	CD-OE1	62.29	1.94	1.25
2	J	110	GLU	CD-OE2	54.77	1.85	1.25
2	J	110	GLU	CG-CD	15.68	1.75	1.51
2	K	65	ASN	C-N	-12.10	1.06	1.34
2	P	65	ASN	C-N	-11.72	1.07	1.34

The worst 5 of 54 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	K	65	ASN	O-C-N	-12.19	103.20	122.70
1	G	11	LYS	CD-CE-NZ	-9.80	89.16	111.70
2	J	110	GLU	CG-CD-OE1	9.15	136.61	118.30
2	J	118	VAL	O-C-N	-9.11	108.12	122.70
2	K	65	ASN	CA-C-N	8.86	136.69	117.20

5 of 8 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	I	64	RC7	C1
2	J	64	RC7	C1
2	K	64	RC7	C1
2	L	64	RC7	C1
2	M	64	RC7	C1

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	J	118	VAL	Mainchain
2	J	66	ARG	Mainchain,Peptide
2	K	66	ARG	Mainchain
2	K	83	PHE	Peptide
2	N	66	ARG	Mainchain



## 5.2 Too-close contacts

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	471	0	456	20	0
1	B	462	0	451	17	0
1	C	454	0	445	16	0
1	D	454	0	445	27	0
1	E	463	0	457	27	0
1	F	454	0	445	30	0
1	G	454	0	445	31	0
1	H	463	0	454	34	0
2	I	1368	0	1307	65	0
2	J	1399	0	1332	58	0
2	K	1384	0	1316	60	0
2	L	1386	0	1317	53	0
2	M	1367	0	1305	88	0
2	N	1356	0	1297	69	0
2	O	1378	0	1313	88	0
2	P	1348	0	1290	93	0
3	E	6	0	8	3	0
3	G	6	0	8	0	0
3	L	6	0	8	0	0
4	E	2	0	0	0	0
4	F	1	0	0	0	0
4	G	1	0	0	0	0
4	H	3	0	0	0	0
4	N	1	0	0	0	0
4	O	1	0	0	0	0
4	P	2	0	0	0	0
5	A	18	0	0	0	0
5	B	21	0	0	2	0
5	C	16	0	0	1	0
5	D	17	0	0	5	0
5	E	22	0	0	8	0
5	F	19	0	0	2	0
5	G	20	0	0	4	0
5	H	19	0	0	5	0
5	I	62	0	0	5	0
5	J	62	0	0	2	0
5	K	59	0	0	7	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	L	54	0	0	3	0
5	M	71	0	0	5	0
5	N	47	0	0	6	0
5	O	57	0	0	13	0
5	P	79	0	0	10	0
All	All	15333	0	14099	651	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 23.

The worst 5 of 651 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)
2:N:66:ARG:HB2	2:N:79:PHE:CE2	1.67	1.30
2:N:66:ARG:HB2	2:N:79:PHE:CD2	1.79	1.16
2:N:72:PRO:HD2	2:N:75:ILE:HG21	1.19	1.14
2:J:90:GLU:OE2	2:J:104:ARG:NH1	1.84	1.10
2:I:64:RC7:OH	2:I:142:SER:OG	1.73	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	59/62 (95%)	51 (86%)	7 (12%)	1 (2%)	11	3
1	B	58/62 (94%)	52 (90%)	6 (10%)	0	100	100
1	C	57/62 (92%)	51 (90%)	5 (9%)	1 (2%)	11	3
1	D	57/62 (92%)	53 (93%)	4 (7%)	0	100	100
1	E	58/62 (94%)	55 (95%)	3 (5%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	F	57/62 (92%)	53 (93%)	4 (7%)	0	100	100
1	G	57/62 (92%)	53 (93%)	4 (7%)	0	100	100
1	H	58/62 (94%)	54 (93%)	4 (7%)	0	100	100
2	I	163/170 (96%)	151 (93%)	11 (7%)	1 (1%)	30	21
2	J	166/170 (98%)	150 (90%)	16 (10%)	0	100	100
2	K	165/170 (97%)	151 (92%)	11 (7%)	3 (2%)	11	3
2	L	165/170 (97%)	149 (90%)	13 (8%)	3 (2%)	11	3
2	M	162/170 (95%)	143 (88%)	16 (10%)	3 (2%)	10	3
2	N	161/170 (95%)	138 (86%)	19 (12%)	4 (2%)	7	1
2	O	164/170 (96%)	141 (86%)	17 (10%)	6 (4%)	4	0
2	P	160/170 (94%)	145 (91%)	8 (5%)	7 (4%)	3	0
All	All	1767/1856 (95%)	1590 (90%)	148 (8%)	29 (2%)	12	4

5 of 29 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	I	201	ASP
1	C	6	GLU
2	K	66	ARG
2	L	111	GLY
2	L	134	LYS

### 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	50/50 (100%)	44 (88%)	6 (12%)	6	2
1	B	49/50 (98%)	44 (90%)	5 (10%)	9	4
1	C	48/50 (96%)	44 (92%)	4 (8%)	14	8
1	D	48/50 (96%)	44 (92%)	4 (8%)	14	8
1	E	49/50 (98%)	44 (90%)	5 (10%)	9	4

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	F	48/50 (96%)	43 (90%)	5 (10%)	9	4
1	G	48/50 (96%)	46 (96%)	2 (4%)	36	32
1	H	49/50 (98%)	47 (96%)	2 (4%)	37	34
2	I	148/153 (97%)	141 (95%)	7 (5%)	32	27
2	J	151/153 (99%)	142 (94%)	9 (6%)	24	18
2	K	150/153 (98%)	142 (95%)	8 (5%)	28	23
2	L	150/153 (98%)	142 (95%)	8 (5%)	28	23
2	M	148/153 (97%)	138 (93%)	10 (7%)	20	13
2	N	147/153 (96%)	137 (93%)	10 (7%)	20	13
2	O	149/153 (97%)	139 (93%)	10 (7%)	20	14
2	P	146/153 (95%)	136 (93%)	10 (7%)	20	13
All	All	1578/1624 (97%)	1473 (93%)	105 (7%)	21	14

5 of 105 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	L	145	LYS
2	M	104	ARG
2	P	100	ILE
2	L	157	ILE
1	E	39	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
2	L	133	GLN
2	L	158	ASN
2	O	212	HIS
2	L	124	ASN
2	O	223	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

16 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	NLW	A	61	1	7,8,8	0.46	0	8,10,10	1.00	1 (12%)
1	NLW	B	61	1	7,8,8	0.52	0	8,10,10	1.27	2 (25%)
1	NLW	C	61	1	7,8,8	0.92	0	8,10,10	1.42	2 (25%)
1	NLW	D	61	1	7,8,8	0.68	0	8,10,10	1.31	1 (12%)
1	NLW	E	61	1	7,8,8	1.19	1 (14%)	8,10,10	1.58	2 (25%)
1	NLW	F	61	1	7,8,8	0.46	0	8,10,10	1.37	2 (25%)
1	NLW	G	61	1	7,8,8	0.38	0	8,10,10	1.13	1 (12%)
1	NLW	H	61	1	7,8,8	0.25	0	8,10,10	1.22	1 (12%)
2	RC7	I	64	2	22,26,27	4.33	10 (45%)	21,35,37	3.77	11 (52%)
2	RC7	J	64	2	22,26,27	4.45	9 (40%)	21,35,37	3.38	10 (47%)
2	RC7	K	64	2	22,26,27	4.18	7 (31%)	21,35,37	3.31	7 (33%)
2	RC7	L	64	2	22,26,27	4.24	6 (27%)	21,35,37	3.20	7 (33%)
2	RC7	M	64	2	22,26,27	4.10	6 (27%)	21,35,37	3.46	7 (33%)
2	RC7	N	64	2	22,26,27	4.73	11 (50%)	21,35,37	3.40	10 (47%)
2	RC7	O	64	2	22,26,27	4.60	10 (45%)	21,35,37	3.48	10 (47%)
2	RC7	P	64	2	22,26,27	4.57	10 (45%)	21,35,37	2.99	7 (33%)

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	NLW	A	61	1	-	0/7/8/8	0/0/0/0
1	NLW	B	61	1	-	0/7/8/8	0/0/0/0
1	NLW	C	61	1	-	0/7/8/8	0/0/0/0
1	NLW	D	61	1	-	0/7/8/8	0/0/0/0
1	NLW	E	61	1	-	0/7/8/8	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	NLW	F	61	1	-	0/7/8/8	0/0/0/0
1	NLW	G	61	1	-	0/7/8/8	0/0/0/0
1	NLW	H	61	1	-	0/7/8/8	0/0/0/0
2	RC7	I	64	2	1/1/4/5	0/10/28/29	0/3/3/3
2	RC7	J	64	2	1/1/4/5	0/10/28/29	0/3/3/3
2	RC7	K	64	2	1/1/4/5	0/10/28/29	0/3/3/3
2	RC7	L	64	2	1/1/4/5	0/10/28/29	0/3/3/3
2	RC7	M	64	2	1/1/4/5	0/10/28/29	0/3/3/3
2	RC7	N	64	2	1/1/4/5	0/10/28/29	0/3/3/3
2	RC7	O	64	2	1/1/4/5	0/10/28/29	0/3/3/3
2	RC7	P	64	2	1/1/4/5	0/10/28/29	0/3/3/3

The worst 5 of 70 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	J	64	RC7	CB2-CA2	-16.21	1.31	1.50
2	K	64	RC7	CB2-CA2	-14.97	1.32	1.50
2	L	64	RC7	CB2-CA2	-14.62	1.33	1.50
2	I	64	RC7	CB2-CA2	-13.36	1.34	1.50
2	P	64	RC7	CB2-CA2	-13.00	1.35	1.50

The worst 5 of 81 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	K	64	RC7	CB2-CA2-C2	-8.45	120.64	128.49
2	L	64	RC7	CB2-CA2-C2	-7.06	121.93	128.49
2	J	64	RC7	CB2-CA2-C2	-5.43	123.45	128.49
2	O	64	RC7	O2-C2-N3	-4.15	119.82	125.45
2	O	64	RC7	CB1-CG1-CD3	-4.08	123.24	129.55

5 of 8 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	J	64	RC7	C1
2	P	64	RC7	C1
2	L	64	RC7	C1
2	N	64	RC7	C1
2	I	64	RC7	C1

There are no torsion outliers.

There are no ring outliers.

11 monomers are involved in 52 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	C	61	NLW	1	0
1	D	61	NLW	1	0
1	G	61	NLW	1	0
2	I	64	RC7	6	0
2	J	64	RC7	10	0
2	K	64	RC7	5	0
2	L	64	RC7	4	0
2	M	64	RC7	7	0
2	N	64	RC7	3	0
2	O	64	RC7	13	0
2	P	64	RC7	4	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 14 ligands modelled in this entry, 11 are monoatomic - leaving 3 for Mogul analysis.

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	$\# Z  > 2$
3	GOL	E	301	-	5,5,5	0.31	0	5,5,5	0.75	0
3	GOL	G	301	-	5,5,5	0.43	0	5,5,5	0.70	0
3	GOL	L	301	-	5,5,5	0.58	0	5,5,5	0.43	0

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
3	GOL	E	301	-	-	0/4/4/4	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GOL	G	301	-	-	0/4/4/4	0/0/0/0
3	GOL	L	301	-	-	0/4/4/4	0/0/0/0

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.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	E	301	GOL	3	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

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

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	P	2
2	J	2
2	K	2
2	L	2
2	O	1
2	M	1

The worst 5 of 10 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	P	64:RC7	C	65:ASN	N	1.87
1	O	64:RC7	C	65:ASN	N	1.77
1	J	64:RC7	C	65:ASN	N	1.68
1	M	64:RC7	C	65:ASN	N	1.68
1	L	64:RC7	C	65:ASN	N	1.10



## 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	59/62 (95%)	-0.24	0 100 100	10, 28, 37, 49	0
1	B	59/62 (95%)	-0.22	0 100 100	9, 24, 44, 46	0
1	C	58/62 (93%)	-0.23	0 100 100	11, 28, 40, 47	1 (1%)
1	D	58/62 (93%)	-0.28	0 100 100	16, 26, 41, 44	0
1	E	58/62 (93%)	0.00	0 100 100	12, 26, 37, 40	1 (1%)
1	F	58/62 (93%)	0.07	2 (3%) 49 58	14, 26, 48, 53	1 (1%)
1	G	58/62 (93%)	-0.17	0 100 100	16, 27, 40, 44	3 (5%)
1	H	58/62 (93%)	0.15	0 100 100	15, 29, 50, 53	2 (3%)
2	I	164/170 (96%)	-0.20	0 100 100	8, 24, 43, 60	2 (1%)
2	J	166/170 (97%)	-0.22	0 100 100	9, 26, 41, 57	1 (0%)
2	K	165/170 (97%)	-0.27	1 (0%) 90 92	11, 25, 40, 48	2 (1%)
2	L	165/170 (97%)	-0.15	1 (0%) 90 92	10, 26, 42, 71	3 (1%)
2	M	161/170 (94%)	-0.00	2 (1%) 81 85	10, 26, 42, 65	2 (1%)
2	N	161/170 (94%)	0.23	7 (4%) 39 48	14, 29, 52, 96	3 (1%)
2	O	165/170 (97%)	0.07	3 (1%) 71 78	16, 31, 48, 73	2 (1%)
2	P	161/170 (94%)	0.20	6 (3%) 45 55	13, 30, 46, 75	2 (1%)
All	All	1774/1856 (95%)	-0.06	22 (1%) 81 85	8, 27, 45, 96	25 (1%)

The worst 5 of 22 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	N	225	TRP	6.4
2	N	115	PHE	6.0
2	P	225	TRP	4.1
2	N	217	TYR	3.5
2	M	225	TRP	2.9

## 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. 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(Å <sup>2</sup> )	Q<0.9
2	RC7	J	64	24/25	0.94	0.10	-	12,17,19,21	0
2	RC7	P	64	24/25	0.91	0.12	-	16,23,25,27	0
1	NLW	E	61	9/9	0.85	0.21	-	17,21,27,29	0
2	RC7	L	64	24/25	0.94	0.10	-	18,21,24,28	0
1	NLW	C	61	9/9	0.95	0.12	-	13,15,17,20	0
1	NLW	D	61	9/9	0.95	0.10	-	18,18,19,21	0
1	NLW	A	61	9/9	0.96	0.12	-	18,19,20,22	0
1	NLW	F	61	9/9	0.88	0.13	-	21,22,24,24	0
2	RC7	N	64	24/25	0.91	0.12	-	17,21,24,29	0
2	RC7	I	64	24/25	0.96	0.09	-	13,15,17,18	0
1	NLW	G	61	9/9	0.89	0.19	-	21,22,26,33	0
1	NLW	H	61	9/9	0.87	0.17	-	17,18,20,24	0
2	RC7	O	64	24/25	0.89	0.14	-	21,26,33,37	0
2	RC7	M	64	24/25	0.93	0.11	-	14,18,20,25	0
2	RC7	K	64	24/25	0.92	0.10	-	15,18,21,24	0
1	NLW	B	61	9/9	0.92	0.14	-	16,19,20,21	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. 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(Å <sup>2</sup> )	Q<0.9
4	MG	H	301	1/1	0.97	0.13	0.99	20,20,20,20	0
4	MG	N	301	1/1	0.96	0.11	-0.08	19,19,19,19	0
4	MG	F	101	1/1	0.98	0.12	-0.14	19,19,19,19	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	MG	P	301	1/1	0.94	0.10	-0.97	18,18,18,18	0
3	GOL	L	301	6/6	0.94	0.09	-1.10	18,19,20,20	0
4	MG	G	302	1/1	0.95	0.08	-1.40	21,21,21,21	0
3	GOL	E	301	6/6	0.94	0.10	-1.42	20,21,22,24	0
4	MG	P	302	1/1	0.97	0.07	-1.70	25,25,25,25	0
4	MG	O	301	1/1	0.98	0.07	-1.84	14,14,14,14	0
4	MG	H	303	1/1	0.97	0.09	-1.84	19,19,19,19	0
3	GOL	G	301	6/6	0.94	0.10	-1.88	21,26,26,26	0
4	MG	E	302	1/1	0.96	0.05	-2.40	22,22,22,22	0
4	MG	E	303	1/1	0.96	0.09	-2.49	28,28,28,28	0
4	MG	H	302	1/1	0.98	0.08	-2.60	19,19,19,19	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.