wwPDB X-ray Structure Validation Summary Report (i)

Nov 11, 2023 – 11:18 am GMT

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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

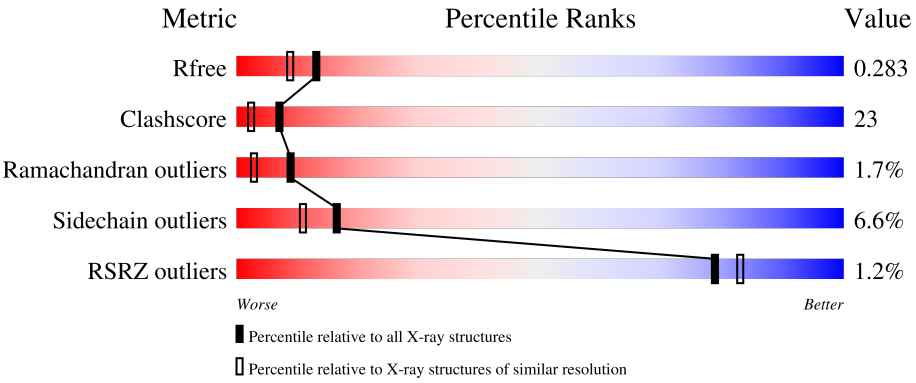
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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.



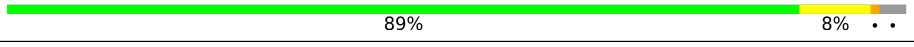
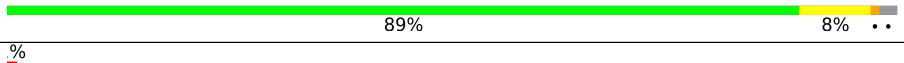
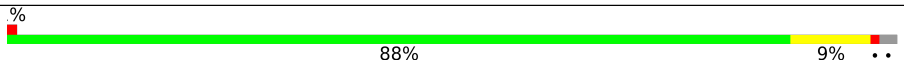
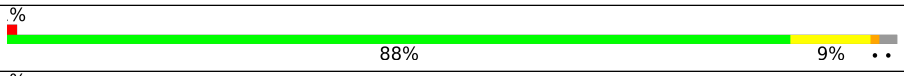
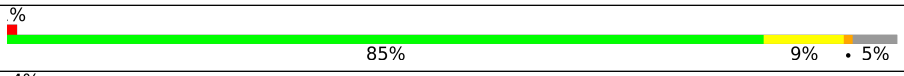
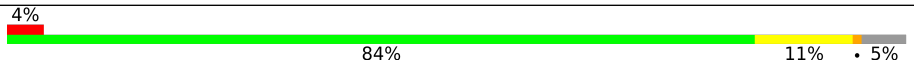
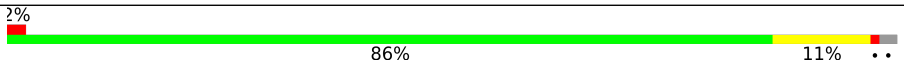
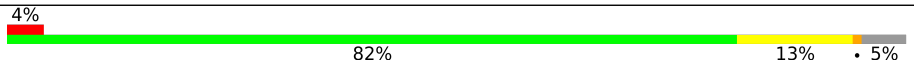
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2523 (2.16-2.12)
Clashscore	141614	2653 (2.16-2.12)
Ramachandran outliers	138981	2618 (2.16-2.12)
Sidechain outliers	138945	2617 (2.16-2.12)
RSRZ outliers	127900	2485 (2.16-2.12)

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 of the lower bar indicate the fraction of residues that contain outliers for >=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 <=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	71% 23% • •
1	B	62	77% 15% 5% •
1	C	62	76% 16% • 5%
1	D	62	65% 26% 5% 5%
1	E	62	58% 34% • 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	62	
1	G	62	
1	H	62	
2	aa	170	
2	bb	170	
2	cc	170	
2	dd	170	
2	ee	170	
2	ff	170	
2	gg	170	
2	hh	170	

2 Entry composition (i)

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	aa	165	Total	C	N	O	S	0	0	0
			1368	879	231	251	7			
2	bb	167	Total	C	N	O	S	1	1	0
			1399	897	241	254	7			
2	cc	166	Total	C	N	O	S	0	1	0
			1384	888	235	254	7			
2	dd	166	Total	C	N	O	S	1	1	0
			1386	889	236	254	7			
2	ee	162	Total	C	N	O	S	1	2	0
			1367	881	229	250	7			
2	ff	162	Total	C	N	O	S	1	1	0
			1356	872	228	249	7			
2	gg	166	Total	C	N	O	S	0	0	0
			1378	885	234	252	7			
2	hh	162	Total	C	N	O	S	0	0	0
			1348	868	226	247	7			

There are 88 discrepancies between the modelled and reference sequences:

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

Continued on next page...

Continued from previous page...

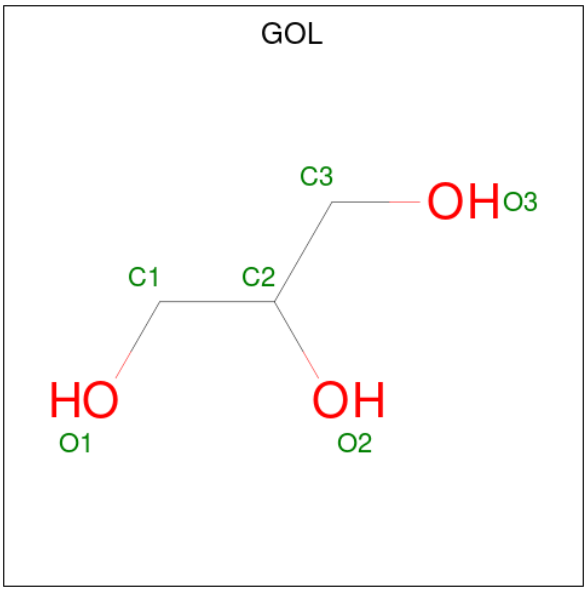
Chain	Residue	Modelled	Actual	Comment	Reference
bb	227	SER	-	expression tag	UNP Q8T6U0
bb	228	HIS	-	expression tag	UNP Q8T6U0
bb	229	HIS	-	expression tag	UNP Q8T6U0
bb	230	HIS	-	expression tag	UNP Q8T6U0
bb	231	HIS	-	expression tag	UNP Q8T6U0
bb	232	HIS	-	expression tag	UNP Q8T6U0
bb	233	HIS	-	expression tag	UNP Q8T6U0
cc	64	RC7	HIS	chromophore	UNP Q8T6U0
cc	64	RC7	TYR	chromophore	UNP Q8T6U0
cc	64	RC7	GLY	chromophore	UNP Q8T6U0
cc	226	GLY	-	expression tag	UNP Q8T6U0
cc	227	SER	-	expression tag	UNP Q8T6U0
cc	228	HIS	-	expression tag	UNP Q8T6U0
cc	229	HIS	-	expression tag	UNP Q8T6U0
cc	230	HIS	-	expression tag	UNP Q8T6U0
cc	231	HIS	-	expression tag	UNP Q8T6U0
cc	232	HIS	-	expression tag	UNP Q8T6U0
cc	233	HIS	-	expression tag	UNP Q8T6U0
dd	64	RC7	HIS	chromophore	UNP Q8T6U0
dd	64	RC7	TYR	chromophore	UNP Q8T6U0
dd	64	RC7	GLY	chromophore	UNP Q8T6U0
dd	226	GLY	-	expression tag	UNP Q8T6U0
dd	227	SER	-	expression tag	UNP Q8T6U0
dd	228	HIS	-	expression tag	UNP Q8T6U0
dd	229	HIS	-	expression tag	UNP Q8T6U0
dd	230	HIS	-	expression tag	UNP Q8T6U0
dd	231	HIS	-	expression tag	UNP Q8T6U0
dd	232	HIS	-	expression tag	UNP Q8T6U0
dd	233	HIS	-	expression tag	UNP Q8T6U0
ee	64	RC7	HIS	chromophore	UNP Q8T6U0
ee	64	RC7	TYR	chromophore	UNP Q8T6U0
ee	64	RC7	GLY	chromophore	UNP Q8T6U0
ee	226	GLY	-	expression tag	UNP Q8T6U0
ee	227	SER	-	expression tag	UNP Q8T6U0
ee	228	HIS	-	expression tag	UNP Q8T6U0
ee	229	HIS	-	expression tag	UNP Q8T6U0
ee	230	HIS	-	expression tag	UNP Q8T6U0
ee	231	HIS	-	expression tag	UNP Q8T6U0
ee	232	HIS	-	expression tag	UNP Q8T6U0
ee	233	HIS	-	expression tag	UNP Q8T6U0
ff	64	RC7	HIS	chromophore	UNP Q8T6U0
ff	64	RC7	TYR	chromophore	UNP Q8T6U0

Continued on next page...

Continued from previous page...

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

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

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

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	E	2	Total Mg 2 2	0	0
4	F	1	Total Mg 1 1	0	0
4	ff	1	Total Mg 1 1	0	0
4	G	1	Total Mg 1 1	0	0
4	gg	1	Total Mg 1 1	0	0
4	H	2	Total Mg 2 2	0	0
4	hh	3	Total Mg 3 3	0	0

- Molecule 5 is water.

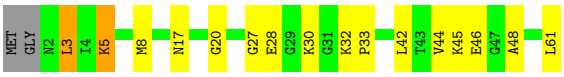
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	18	Total O 18 18	0	0
5	aa	62	Total O 62 62	0	0
5	B	21	Total O 21 21	0	0
5	bb	62	Total O 62 62	0	0
5	C	16	Total O 16 16	0	0
5	cc	59	Total O 59 59	0	0
5	D	17	Total O 17 17	0	0
5	dd	54	Total O 54 54	0	0
5	E	22	Total O 22 22	0	0
5	ee	71	Total O 71 71	0	0
5	F	19	Total O 19 19	0	0
5	ff	47	Total O 47 47	0	0
5	G	20	Total O 20 20	0	0
5	gg	57	Total O 57 57	0	0
5	H	19	Total O 19 19	0	0
5	hh	79	Total O 79 79	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: Green fluorescent protein

Chain A:

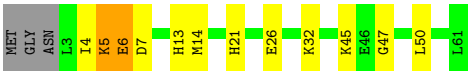


• Molecule 1: Green fluorescent protein

Chain B:

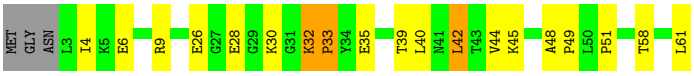
• Molecule 1: Green fluorescent protein

Chain C:



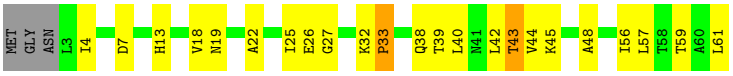
• Molecule 1: Green fluorescent protein

Chain D:



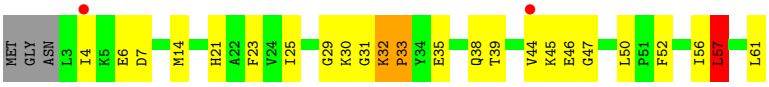
• Molecule 1: Green fluorescent protein

Chain E:



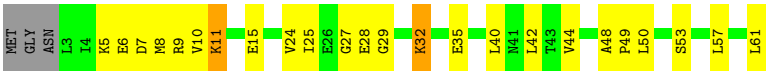
• Molecule 1: Green fluorescent protein

Chain F:



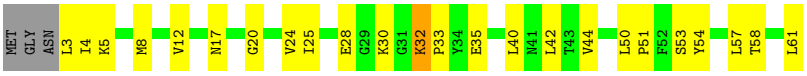
• Molecule 1: Green fluorescent protein

Chain G:



• Molecule 1: Green fluorescent protein

Chain H:

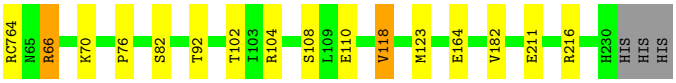


• Molecule 2: Green fluorescent protein

Chain aa:

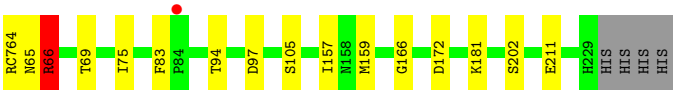
• Molecule 2: Green fluorescent protein

Chain bb:



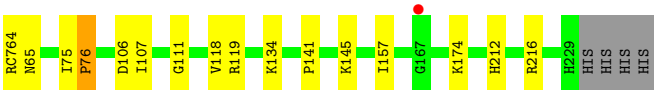
• Molecule 2: Green fluorescent protein

Chain cc:



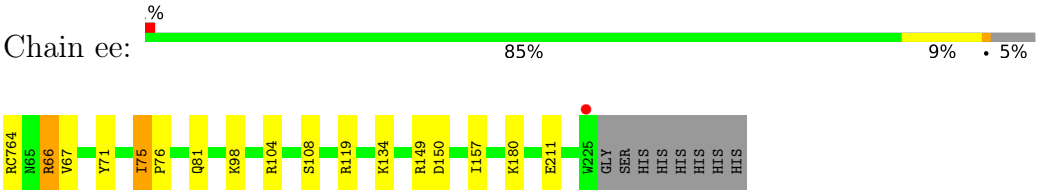
• Molecule 2: Green fluorescent protein

Chain dd:



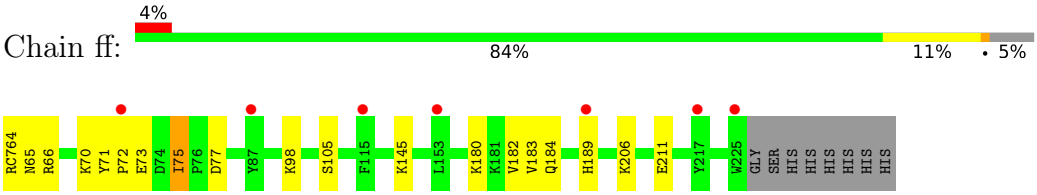
• Molecule 2: Green fluorescent protein

Chain ee:



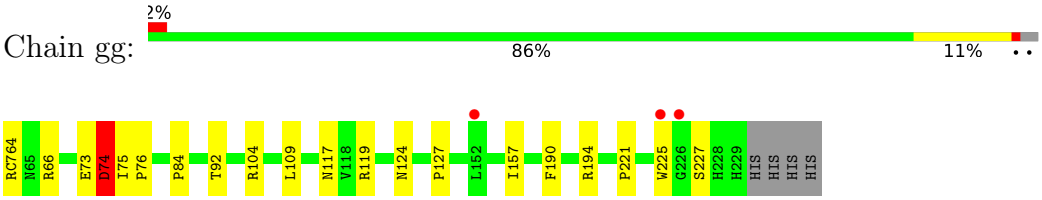
• Molecule 2: Green fluorescent protein

Chain ff:



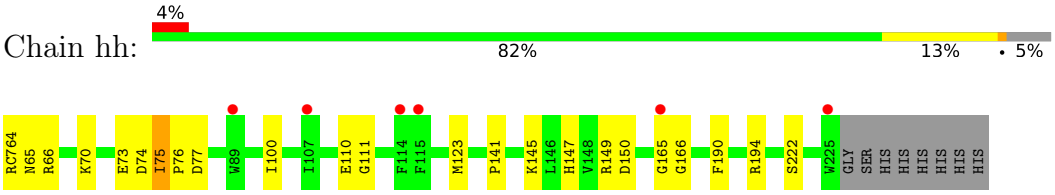
• Molecule 2: Green fluorescent protein

Chain gg:



• Molecule 2: Green fluorescent protein

Chain hh:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	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
$< I/\sigma(I) >$ [1]	2.07 (at 2.14Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.195 , 0.279 0.196 , 0.283	Depositor DCC
R_{free} test set	1817 reflections (2.01%)	wwPDB-VP
Wilson B-factor (Å^2)	24.6	Xtriage
Anisotropy	0.098	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.26 , 14.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	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^2)	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.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: RC7, MG, NLW, GOL

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	aa	0.99	2/1384 (0.1%)	1.03	8/1873 (0.4%)
2	bb	2.45	5/1417 (0.4%)	1.05	8/1917 (0.4%)
2	cc	0.99	2/1401 (0.1%)	1.02	6/1896 (0.3%)
2	dd	0.97	3/1403 (0.2%)	0.99	6/1899 (0.3%)
2	ee	0.89	2/1383 (0.1%)	0.99	5/1872 (0.3%)
2	ff	0.81	1/1371 (0.1%)	1.03	2/1856 (0.1%)
2	gg	0.78	2/1395 (0.1%)	0.95	3/1888 (0.2%)
2	hh	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	bb	0	3
2	cc	0	2
2	ff	0	2
All	All	0	8

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	bb	110	GLU	CD-OE1	62.29	1.94	1.25
2	bb	110	GLU	CD-OE2	54.77	1.85	1.25
2	bb	110	GLU	CG-CD	15.68	1.75	1.51
2	cc	65	ASN	C-N	-12.10	1.06	1.34
2	hh	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(^{o})	Ideal(^{o})
2	cc	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	bb	110	GLU	CG-CD-OE1	9.15	136.61	118.30
2	bb	118	VAL	O-C-N	-9.11	108.12	122.70
2	cc	65	ASN	CA-C-N	8.86	136.69	117.20

There are no chirality outliers.

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	bb	118	VAL	Mainchain
2	bb	66	ARG	Peptide,Mainchain
2	cc	66	ARG	Mainchain
2	cc	83	PHE	Peptide
2	ff	66	ARG	Mainchain

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	471	0	456	5	0
1	B	462	0	451	6	0
1	C	454	0	445	5	0
1	D	454	0	445	16	0
1	E	463	0	457	18	0
1	F	454	0	445	19	0
1	G	454	0	445	17	0
1	H	463	0	454	19	0
2	aa	1368	0	1304	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	bb	1399	0	1329	0	0
2	cc	1384	0	1313	0	0
2	dd	1386	0	1315	0	0
2	ee	1367	0	1302	0	0
2	ff	1356	0	1294	0	0
2	gg	1378	0	1311	0	0
2	hh	1348	0	1288	0	0
3	E	6	0	8	1	0
3	G	6	0	8	0	0
3	dd	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	2	0	0	0	0
4	ff	1	0	0	0	0
4	gg	1	0	0	0	0
4	hh	3	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	3	0
5	H	19	0	0	5	0
5	aa	62	0	0	0	0
5	bb	62	0	0	0	0
5	cc	59	0	0	0	0
5	dd	54	0	0	0	0
5	ee	71	0	0	0	0
5	ff	47	0	0	0	0
5	gg	57	0	0	0	0
5	hh	79	0	0	0	0
All	All	15333	0	14078	104	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 104 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:E:57:LEU:HG	5:E:413:HOH:O	1.56	1.06
1:F:7:ASP:OD1	1:F:32:LYS:HE2	1.56	1.06
1:F:52:PHE:HE1	1:F:57:LEU:HD11	1.38	0.87
1:G:57:LEU:HG	5:G:404:HOH:O	1.74	0.86
1:F:25:ILE:HG12	1:F:44:VAL:HG22	1.59	0.82

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%)	9 3
1	B	58/62 (94%)	52 (90%)	6 (10%)	0	100 100
1	C	57/62 (92%)	51 (90%)	5 (9%)	1 (2%)	8 2
1	D	57/62 (92%)	53 (93%)	4 (7%)	0	100 100
1	E	58/62 (94%)	55 (95%)	3 (5%)	0	100 100
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	aa	162/170 (95%)	151 (93%)	10 (6%)	1 (1%)	25 17
2	bb	165/170 (97%)	150 (91%)	15 (9%)	0	100 100
2	cc	164/170 (96%)	151 (92%)	10 (6%)	3 (2%)	8 2
2	dd	164/170 (96%)	148 (90%)	13 (8%)	3 (2%)	8 2
2	ee	161/170 (95%)	142 (88%)	16 (10%)	3 (2%)	8 2
2	ff	160/170 (94%)	137 (86%)	19 (12%)	4 (2%)	5 1
2	gg	163/170 (96%)	141 (86%)	16 (10%)	6 (4%)	3 0
2	hh	159/170 (94%)	145 (91%)	7 (4%)	7 (4%)	2 0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	1759/1856 (95%)	1587 (90%)	143 (8%)	29 (2%)	9 3

5 of 29 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	aa	201	ASP
1	C	6	GLU
2	cc	66	ARG
2	dd	111	GLY
2	dd	134	LYS

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	50/50 (100%)	44 (88%)	6 (12%)	5 2
1	B	49/50 (98%)	44 (90%)	5 (10%)	7 3
1	C	48/50 (96%)	44 (92%)	4 (8%)	11 6
1	D	48/50 (96%)	44 (92%)	4 (8%)	11 6
1	E	49/50 (98%)	44 (90%)	5 (10%)	7 3
1	F	48/50 (96%)	43 (90%)	5 (10%)	7 3
1	G	48/50 (96%)	46 (96%)	2 (4%)	30 26
1	H	49/50 (98%)	47 (96%)	2 (4%)	30 27
2	aa	148/153 (97%)	141 (95%)	7 (5%)	26 21
2	bb	151/153 (99%)	142 (94%)	9 (6%)	19 14
2	cc	150/153 (98%)	142 (95%)	8 (5%)	22 18
2	dd	150/153 (98%)	142 (95%)	8 (5%)	22 18
2	ee	148/153 (97%)	138 (93%)	10 (7%)	16 10
2	ff	147/153 (96%)	137 (93%)	10 (7%)	16 10
2	gg	149/153 (97%)	139 (93%)	10 (7%)	16 10
2	hh	146/153 (95%)	136 (93%)	10 (7%)	16 10

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1578/1624 (97%)	1473 (93%)	105 (7%)	16 10

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

Mol	Chain	Res	Type
2	ee	98	LYS
2	ff	65	ASN
2	hh	123	MET
2	ee	108	SER
2	ee	211	GLU

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

Mol	Chain	Res	Type
2	hh	156	ASN
2	gg	223	GLN
2	dd	158	ASN
2	gg	212	HIS
2	dd	133	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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	$\#\|Z\| > 2$
2	RC7	gg	64	2	24,26,27	3.34	8 (33%)	27,35,37	4.09	11 (40%)
2	RC7	dd	64	2	24,26,27	1.76	7 (29%)	27,35,37	2.41	7 (25%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
2	RC7	cc	64	2	24,26,27	1.52	6 (25%)	27,35,37	2.59	9 (33%)
2	RC7	ff	64	2	24,26,27	3.31	7 (29%)	27,35,37	3.67	17 (62%)
2	RC7	ee	64	2	24,26,27	2.97	5 (20%)	27,35,37	3.57	11 (40%)
1	NLW	D	61	1	8,8,8	0.65	0	10,10,10	1.16	1 (10%)
2	RC7	bb	64	2	24,26,27	1.94	5 (20%)	27,35,37	2.52	10 (37%)
1	NLW	C	61	1	8,8,8	0.76	0	10,10,10	1.01	0
1	NLW	F	61	1	8,8,8	0.38	0	10,10,10	1.05	1 (10%)
2	RC7	aa	64	2	24,26,27	3.02	9 (37%)	27,35,37	2.44	8 (29%)
1	NLW	G	61	1	8,8,8	0.38	0	10,10,10	1.00	1 (10%)
1	NLW	E	61	1	8,8,8	1.03	0	10,10,10	1.24	1 (10%)
1	NLW	B	61	1	8,8,8	0.40	0	10,10,10	1.23	2 (20%)
2	RC7	hh	64	2	24,26,27	3.22	5 (20%)	27,35,37	2.50	8 (29%)
1	NLW	H	61	1	8,8,8	0.36	0	10,10,10	1.09	1 (10%)
1	NLW	A	61	1	8,8,8	0.41	0	10,10,10	0.96	1 (10%)

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
2	RC7	gg	64	2	-	0/9/28/29	0/3/3/3
2	RC7	dd	64	2	-	1/9/28/29	0/3/3/3
2	RC7	cc	64	2	-	0/9/28/29	0/3/3/3
2	RC7	ff	64	2	-	0/9/28/29	0/3/3/3
2	RC7	ee	64	2	-	1/9/28/29	0/3/3/3
1	NLW	D	61	1	-	2/8/8/8	-
2	RC7	bb	64	2	-	2/9/28/29	0/3/3/3
1	NLW	C	61	1	-	4/8/8/8	-
1	NLW	F	61	1	-	0/8/8/8	-
2	RC7	aa	64	2	-	1/9/28/29	0/3/3/3
1	NLW	G	61	1	-	2/8/8/8	-
1	NLW	E	61	1	-	6/8/8/8	-
1	NLW	B	61	1	-	1/8/8/8	-
2	RC7	hh	64	2	-	0/9/28/29	0/3/3/3
1	NLW	H	61	1	-	3/8/8/8	-
1	NLW	A	61	1	-	3/8/8/8	-

The worst 5 of 52 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	hh	64	RC7	CA1-C1	-14.56	1.26	1.45
2	gg	64	RC7	CA1-C1	-13.79	1.27	1.45
2	ff	64	RC7	CA1-C1	-12.52	1.29	1.45
2	ee	64	RC7	CA1-C1	-12.23	1.29	1.45
2	aa	64	RC7	CA1-C1	9.50	1.58	1.45

The worst 5 of 89 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	gg	64	RC7	C2-CA2-N2	-11.88	100.61	108.93
2	gg	64	RC7	CA2-C2-N3	11.56	108.83	103.37
2	ee	64	RC7	CA2-C2-N3	10.25	108.22	103.37
2	ee	64	RC7	C2-CA2-N2	-9.19	102.50	108.93
2	ff	64	RC7	C2-CA2-N2	-8.28	103.13	108.93

There are no chirality outliers.

5 of 26 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	61	NLW	NH2-C-CA-N
1	C	61	NLW	NH2-C-CA-N
1	A	61	NLW	NH2-C-CA-CB
2	dd	64	RC7	C3-CA3-N3-C2
1	B	61	NLW	O-C-CA-N

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides 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	dd	301	-	5,5,5	0.62	0	5,5,5	0.40	0
3	GOL	G	301	-	5,5,5	0.45	0	5,5,5	0.72	0
3	GOL	E	301	-	5,5,5	0.35	0	5,5,5	0.78	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	dd	301	-	-	4/4/4/4	-
3	GOL	G	301	-	-	4/4/4/4	-
3	GOL	E	301	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	dd	301	GOL	O1-C1-C2-C3
3	dd	301	GOL	C1-C2-C3-O3
3	E	301	GOL	O1-C1-C2-O2
3	E	301	GOL	O1-C1-C2-C3
3	G	301	GOL	O1-C1-C2-C3

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	E	301	GOL	1	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	hh	2
2	bb	2
2	dd	2
2	cc	2
2	gg	1
2	ee	1

The worst 5 of 10 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	hh	64:RC7	C3	65:ASN	N	1.87
1	gg	64:RC7	C3	65:ASN	N	1.77
1	bb	64:RC7	C3	65:ASN	N	1.68
1	ee	64:RC7	C3	65:ASN	N	1.68
1	dd	64:RC7	C3	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^{th} 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^2)	Q<0.9
1	A	59/62 (95%)	-0.23	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%) 45 52	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	aa	164/170 (96%)	-0.20	0 100 100	8, 24, 43, 60	2 (1%)
2	bb	166/170 (97%)	-0.22	0 100 100	9, 26, 41, 57	1 (0%)
2	cc	165/170 (97%)	-0.27	1 (0%) 89 91	11, 25, 40, 48	2 (1%)
2	dd	165/170 (97%)	-0.15	1 (0%) 89 91	10, 26, 42, 71	3 (1%)
2	ee	161/170 (94%)	-0.00	1 (0%) 89 91	10, 26, 42, 65	2 (1%)
2	ff	161/170 (94%)	0.23	7 (4%) 35 43	14, 29, 52, 96	3 (1%)
2	gg	165/170 (97%)	0.07	3 (1%) 68 74	16, 31, 48, 73	2 (1%)
2	hh	161/170 (94%)	0.21	6 (3%) 41 49	13, 30, 46, 75	2 (1%)
All	All	1774/1856 (95%)	-0.06	21 (1%) 79 83	8, 27, 45, 96	25 (1%)

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	ff	225	TRP	6.4
2	ff	115	PHE	6.1
2	hh	225	TRP	4.1
2	ff	217	TYR	3.5
2	ee	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. The B-factors column lists the minimum, median, 95^{th} 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^2)	Q<0.9
1	NLW	E	61	9/9	0.85	0.21	17,21,27,29	0
1	NLW	H	61	9/9	0.87	0.17	17,18,20,24	0
1	NLW	F	61	9/9	0.88	0.13	21,22,24,24	0
2	RC7	gg	64	24/25	0.89	0.14	21,26,33,37	0
1	NLW	G	61	9/9	0.89	0.19	21,22,26,33	0
2	RC7	ff	64	24/25	0.91	0.12	17,21,24,29	0
2	RC7	hh	64	24/25	0.91	0.13	16,23,25,27	0
2	RC7	cc	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
2	RC7	ee	64	24/25	0.93	0.11	14,18,20,25	0
2	RC7	bb	64	24/25	0.94	0.10	12,17,19,21	0
2	RC7	dd	64	24/25	0.94	0.10	18,21,24,28	0
1	NLW	D	61	9/9	0.95	0.10	18,18,19,21	0
1	NLW	C	61	9/9	0.95	0.12	13,15,17,20	0
1	NLW	A	61	9/9	0.96	0.12	18,19,20,22	0
2	RC7	aa	64	24/25	0.96	0.09	13,15,17,18	0

6.3 Carbohydrates (i)

There are no monosaccharides 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^{th} 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^2)	Q<0.9
3	GOL	dd	301	6/6	0.94	0.09	18,19,20,20	0
3	GOL	E	301	6/6	0.94	0.10	20,21,22,24	0
3	GOL	G	301	6/6	0.94	0.10	21,26,26,26	0
4	MG	hh	302	1/1	0.94	0.10	18,18,18,18	0
4	MG	G	302	1/1	0.95	0.08	21,21,21,21	0
4	MG	E	303	1/1	0.95	0.09	28,28,28,28	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	MG	E	302	1/1	0.96	0.05	22,22,22,22	0
4	MG	ff	301	1/1	0.96	0.11	19,19,19,19	0
4	MG	H	302	1/1	0.97	0.09	19,19,19,19	0
4	MG	H	301	1/1	0.97	0.13	20,20,20,20	0
4	MG	hh	303	1/1	0.97	0.07	25,25,25,25	0
4	MG	hh	301	1/1	0.98	0.08	19,19,19,19	0
4	MG	F	101	1/1	0.98	0.12	19,19,19,19	0
4	MG	gg	301	1/1	0.98	0.07	14,14,14,14	0

6.5 Other polymers (i)

There are no such residues in this entry.