wwPDB X-ray Structure Validation Summary Report (i)

May 14, 2020 – 01:57 pm BST

PDB ID	:	1WPB
Title	:	Structure of Escherichia coli yfbU gene product
Authors	:	Borek, D.; Chen, Y.; Zheng, M.; Skarina, T.; Savchenko, A.; Edwards, A.; Otwinowski, Z.; Midwest Center for Structural Genomics (MCSG)
Deposited on	:	2004-09-01
Resolution	:	2.00 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

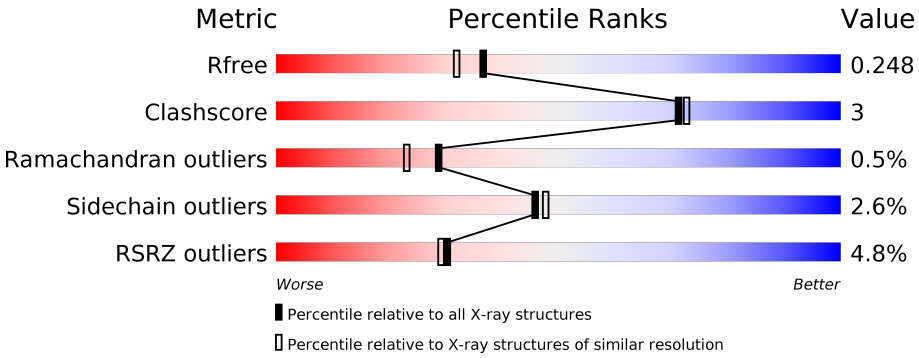
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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 >=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	172	7% 86% 10%
1	B	172	4% 89% 8%
1	C	172	6% 89% 8%
1	D	172	4% 88% 8%
1	E	172	7% 84% 10%
1	F	172	6% 88% 9%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	172	5% 88% 9%
1	H	172	3% 91% 6%
1	I	172	2% 92% 5%
1	J	172	4% 88% 9%
1	K	172	3% 88% 8%
1	L	172	8% 86% 11%
1	M	172	6% 88% 8%
1	N	172	3% 85% 12%
1	O	172	5% 90% 7%
1	P	172	3% 87% 9%

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	CL	C	2011	-	-	-	X
2	CL	J	2016	-	-	-	X
3	GOL	B	3003	-	X	-	-
3	GOL	C	3036	-	-	X	-
3	GOL	D	3005	-	-	X	-
3	GOL	E	3004	-	-	X	-
3	GOL	G	3001	-	-	X	-
3	GOL	H	3006	-	-	X	-
3	GOL	K	3011	-	-	X	-
3	GOL	M	3043	-	-	X	-
3	GOL	N	3012	-	X	-	-

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 24406 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.

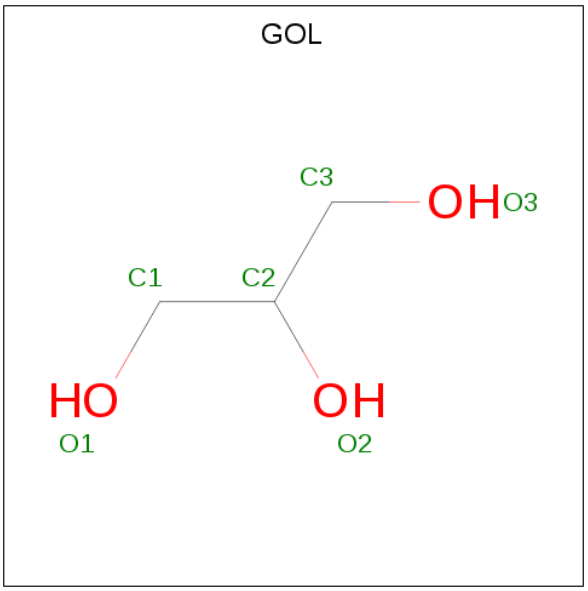
- Molecule 1 is a protein called hypothetical protein yfbU.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	168	Total	C	N	O	S	0	2	0
			1410	879	257	261	13			
1	B	168	Total	C	N	O	S	0	2	0
			1408	878	254	263	13			
1	C	168	Total	C	N	O	S	0	1	0
			1403	875	254	261	13			
1	D	168	Total	C	N	O	S	0	1	0
			1405	876	256	260	13			
1	E	168	Total	C	N	O	S	0	5	0
			1425	887	260	265	13			
1	F	168	Total	C	N	O	S	0	2	0
			1405	876	254	262	13			
1	G	168	Total	C	N	O	S	0	2	0
			1407	877	254	263	13			
1	H	168	Total	C	N	O	S	0	1	0
			1403	875	254	261	13			
1	I	168	Total	C	N	O	S	0	1	0
			1403	875	254	261	13			
1	J	168	Total	C	N	O	S	0	1	0
			1403	875	254	261	13			
1	K	168	Total	C	N	O	S	0	1	0
			1406	877	256	260	13			
1	L	168	Total	C	N	O	S	0	2	0
			1407	877	255	262	13			
1	M	167	Total	C	N	O	S	0	1	0
			1394	870	252	259	13			
1	N	168	Total	C	N	O	S	0	1	0
			1403	875	254	261	13			
1	O	168	Total	C	N	O	S	0	1	0
			1403	875	254	261	13			
1	P	168	Total	C	N	O	S	0	1	0
			1403	875	254	261	13			

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	P	1	Total Cl 1 1	0	0
2	G	1	Total Cl 1 1	0	0
2	J	3	Total Cl 3 3	0	0
2	D	3	Total Cl 3 3	0	0
2	K	2	Total Cl 2 2	0	0
2	E	2	Total Cl 2 2	0	0
2	H	3	Total Cl 3 3	0	0
2	B	1	Total Cl 1 1	0	0
2	I	4	Total Cl 4 4	0	0
2	C	5	Total Cl 5 5	0	0
2	A	1	Total Cl 1 1	0	0
2	N	1	Total Cl 1 1	0	0
2	O	2	Total Cl 2 2	0	0
2	L	2	Total Cl 2 2	0	0
2	F	3	Total Cl 3 3	0	0
2	M	2	Total Cl 2 2	0	0

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



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 6 3 3	0	0
3	A	1	Total C O 6 3 3	0	0
3	B	1	Total C O 6 3 3	0	0
3	B	1	Total C O 6 3 3	0	0
3	B	1	Total C O 6 3 3	0	0
3	B	1	Total C O 6 3 3	0	0
3	B	1	Total C O 6 3 3	0	0
3	C	1	Total C O 6 3 3	0	0
3	C	1	Total C O 6 3 3	0	0
3	D	1	Total C O 6 3 3	0	0
3	D	1	Total C O 6 3 3	0	0
3	D	1	Total C O 6 3 3	0	0
3	D	1	Total C O 6 3 3	0	0
3	E	1	Total C O 6 3 3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	E	1	Total	C	O	0	0
			6	3	3		
3	E	1	Total	C	O	0	0
			6	3	3		
3	E	1	Total	C	O	0	0
			6	3	3		
3	F	1	Total	C	O	0	0
			6	3	3		
3	G	1	Total	C	O	0	0
			6	3	3		
3	G	1	Total	C	O	0	0
			6	3	3		
3	H	1	Total	C	O	0	0
			6	3	3		
3	H	1	Total	C	O	0	0
			6	3	3		
3	H	1	Total	C	O	0	0
			6	3	3		
3	I	1	Total	C	O	0	0
			6	3	3		
3	I	1	Total	C	O	0	0
			6	3	3		
3	J	1	Total	C	O	0	0
			6	3	3		
3	J	1	Total	C	O	0	0
			6	3	3		
3	K	1	Total	C	O	0	0
			6	3	3		
3	K	1	Total	C	O	0	0
			6	3	3		
3	K	1	Total	C	O	0	0
			6	3	3		
3	K	1	Total	C	O	0	0
			6	3	3		
3	L	1	Total	C	O	0	0
			6	3	3		
3	L	1	Total	C	O	0	0
			6	3	3		
3	M	1	Total	C	O	0	0
			6	3	3		
3	M	1	Total	C	O	0	0
			6	3	3		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	M	1	Total C O 6 3 3	0	0
3	N	1	Total C O 6 3 3	0	0
3	N	1	Total C O 6 3 3	0	0
3	N	1	Total C O 6 3 3	0	0
3	N	1	Total C O 6 3 3	0	0
3	O	1	Total C O 6 3 3	0	0
3	O	1	Total C O 6 3 3	0	0
3	P	1	Total C O 6 3 3	0	0
3	P	1	Total C O 6 3 3	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	96	Total O 96 96	0	0
4	B	96	Total O 96 96	0	0
4	C	103	Total O 103 103	0	0
4	D	110	Total O 110 110	0	0
4	E	88	Total O 88 88	0	0
4	F	85	Total O 85 85	0	0
4	G	89	Total O 89 89	0	0
4	H	104	Total O 104 104	0	0
4	I	130	Total O 130 130	0	0
4	J	121	Total O 121 121	0	0

Continued on next page...

Continued from previous page...

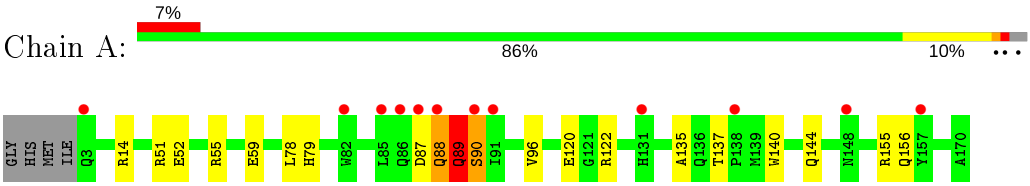
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	K	117	Total O 117 117	0	0
4	L	102	Total O 102 102	0	0
4	M	92	Total O 92 92	0	0
4	N	92	Total O 92 92	0	0
4	O	85	Total O 85 85	0	0
4	P	108	Total O 108 108	0	0

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 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.

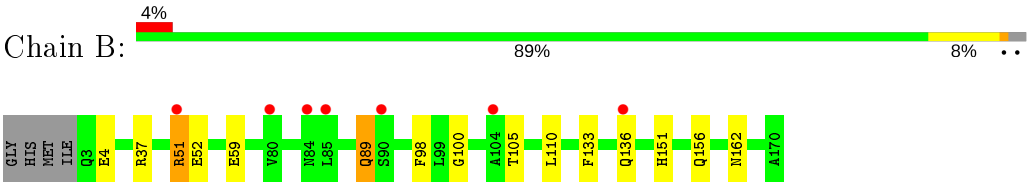
- Molecule 1: hypothetical protein yfbU

Chain A:



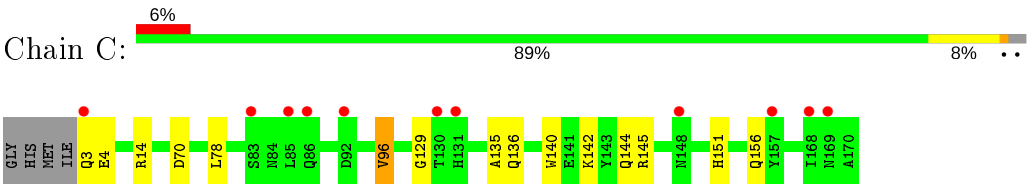
- Molecule 1: hypothetical protein yfbU

Chain B:



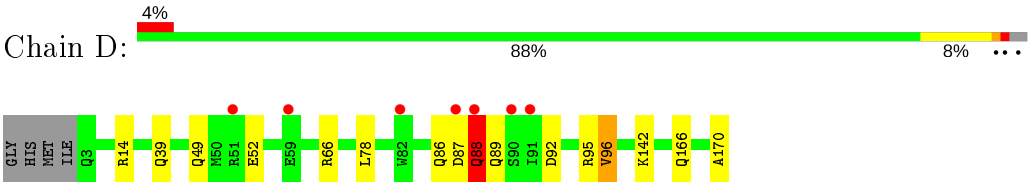
- Molecule 1: hypothetical protein yfbU

Chain C:



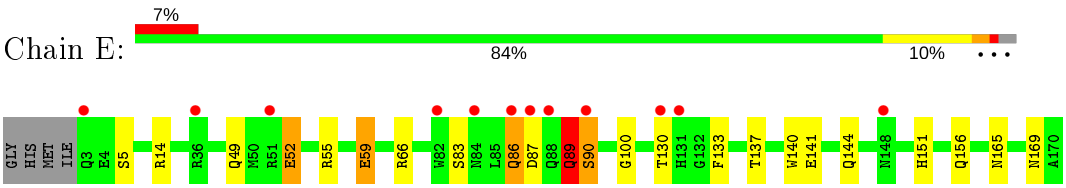
- Molecule 1: hypothetical protein yfbU

Chain D:



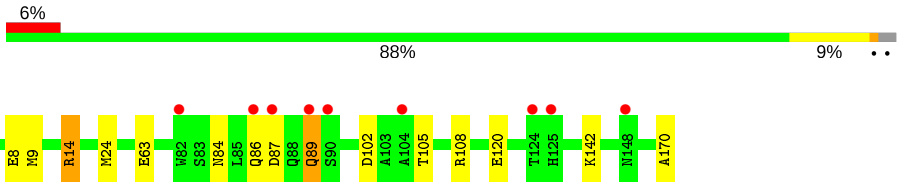
- Molecule 1: hypothetical protein yfbU

Chain E:



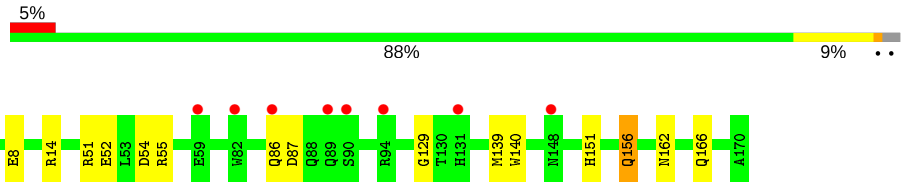
- Molecule 1: hypothetical protein yfbU

Chain F:



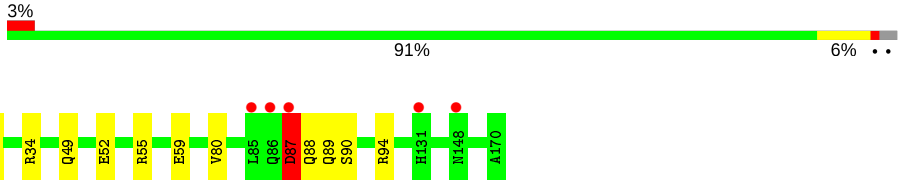
• Molecule 1: hypothetical protein yfbU

Chain G:



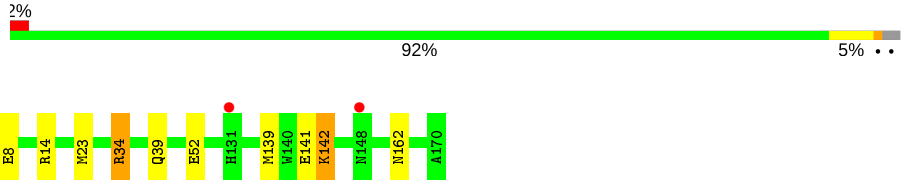
• Molecule 1: hypothetical protein yfbU

Chain H:



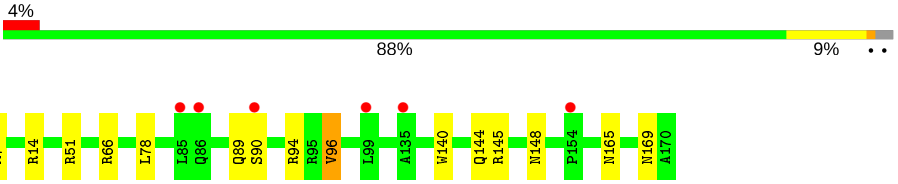
• Molecule 1: hypothetical protein yfbU

Chain I:



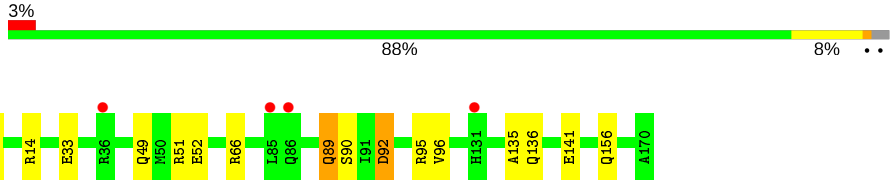
• Molecule 1: hypothetical protein yfbU

Chain J:



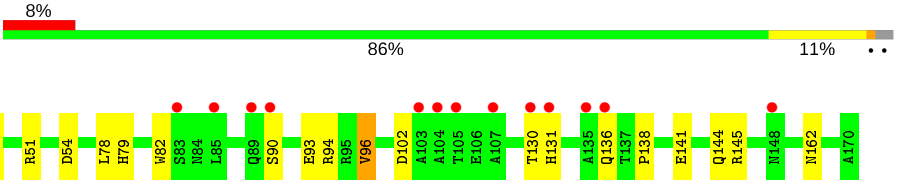
• Molecule 1: hypothetical protein yfbU

Chain K:



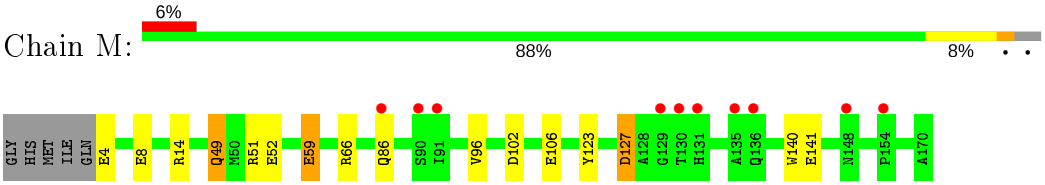
• Molecule 1: hypothetical protein yfbU

Chain L:



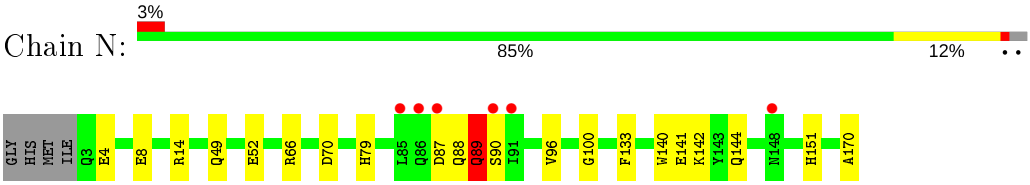
• Molecule 1: hypothetical protein yfbU

Chain M:



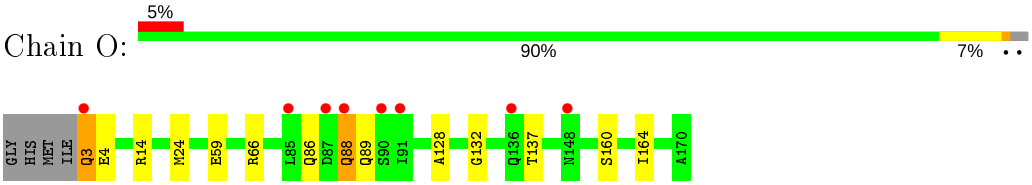
• Molecule 1: hypothetical protein yfbU

Chain N:



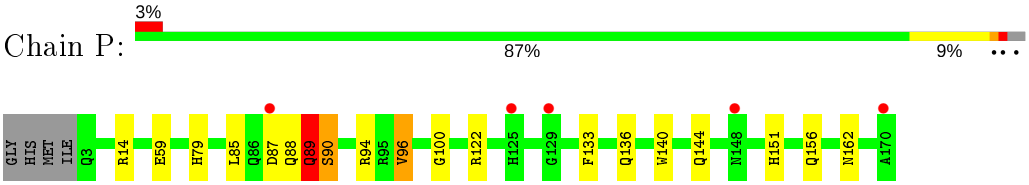
• Molecule 1: hypothetical protein yfbU

Chain O:



• Molecule 1: hypothetical protein yfbU

Chain P:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 2 3	Depositor
Cell constants a, b, c, α, β, γ	230.52Å 230.52Å 230.52Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	15.00 – 2.00 230.52 – 1.95	Depositor EDS
% Data completeness (in resolution range)	100.0 (15.00-2.00) 99.8 (230.52-1.95)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	1.45 (at 1.95Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.190 , 0.227 0.211 , 0.248	Depositor DCC
R_{free} test set	1481 reflections (0.50%)	wwPDB-VP
Wilson B-factor (Å^2)	31.4	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 61.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.015 for l,-k,h	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	24406	wwPDB-VP
Average B, all atoms (Å^2)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 16.61% 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

5.1 Standard geometry

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

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	0.88	1/1447 (0.1%)	0.84	2/1950 (0.1%)
1	B	0.86	0/1445	0.80	0/1948
1	C	0.90	0/1436	0.85	2/1936 (0.1%)
1	D	0.91	0/1439	0.86	4/1939 (0.2%)
1	E	0.87	2/1476 (0.1%)	0.91	4/1988 (0.2%)
1	F	0.80	2/1443 (0.1%)	0.82	3/1946 (0.2%)
1	G	0.77	0/1445	0.81	2/1948 (0.1%)
1	H	0.90	0/1436	0.94	4/1936 (0.2%)
1	I	0.95	3/1436 (0.2%)	0.87	3/1936 (0.2%)
1	J	0.87	0/1436	0.87	5/1936 (0.3%)
1	K	0.90	2/1439 (0.1%)	0.90	4/1939 (0.2%)
1	L	0.92	0/1444	0.88	2/1947 (0.1%)
1	M	0.87	2/1427 (0.1%)	0.89	4/1924 (0.2%)
1	N	0.82	1/1436 (0.1%)	0.82	3/1936 (0.2%)
1	O	0.85	0/1436	0.86	4/1936 (0.2%)
1	P	0.88	0/1436	0.85	2/1936 (0.1%)
All	All	0.87	13/23057 (0.1%)	0.86	48/31081 (0.2%)

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	A	0	1
1	C	0	1
1	D	1	3
1	E	0	2
1	G	0	2
1	H	0	1
1	K	0	1

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
1	L	0	1
1	O	0	1
1	P	0	2
All	All	1	15

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	59	GLU	CB-CG	7.46	1.66	1.52
1	I	141	GLU	CG-CD	5.80	1.60	1.51
1	M	140	TRP	CB-CG	-5.79	1.39	1.50
1	N	141	GLU	CG-CD	5.77	1.60	1.51
1	K	141	GLU	CD-OE1	5.66	1.31	1.25

The worst 5 of 48 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	H	14	ARG	NE-CZ-NH2	-12.21	114.19	120.30
1	H	14	ARG	NE-CZ-NH1	11.90	126.25	120.30
1	M	14	ARG	NE-CZ-NH2	11.02	125.81	120.30
1	G	14	ARG	NE-CZ-NH2	11.01	125.81	120.30
1	G	14	ARG	NE-CZ-NH1	-10.88	114.86	120.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	D	88	GLN	CA

5 of 15 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	89	GLN	Peptide
1	C	129	GLY	Peptide
1	D	86	GLN	Peptide
1	D	87	ASP	Peptide
1	D	88	GLN	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	1410	0	1353	12	2
1	B	1408	0	1346	6	0
1	C	1403	0	1344	11	0
1	D	1405	0	1351	10	0
1	E	1425	0	1366	15	0
1	F	1405	0	1347	9	0
1	G	1407	0	1346	9	0
1	H	1403	0	1344	7	0
1	I	1403	0	1344	5	0
1	J	1403	0	1344	6	0
1	K	1406	0	1351	11	0
1	L	1407	0	1346	7	0
1	M	1394	0	1336	10	0
1	N	1403	0	1344	13	0
1	O	1403	0	1344	5	0
1	P	1403	0	1344	14	0
2	A	1	0	0	1	0
2	B	1	0	0	0	0
2	C	5	0	0	0	0
2	D	3	0	0	0	0
2	E	2	0	0	0	0
2	F	3	0	0	1	0
2	G	1	0	0	0	0
2	H	3	0	0	1	0
2	I	4	0	0	0	0
2	J	3	0	0	0	0
2	K	2	0	0	0	0
2	L	2	0	0	0	0
2	M	2	0	0	0	0
2	N	1	0	0	0	0
2	O	2	0	0	0	0
2	P	1	0	0	0	0
3	A	12	0	16	3	0
3	B	30	0	39	3	0
3	C	12	0	16	4	1
3	D	24	0	32	4	0
3	E	24	0	32	7	1
3	F	6	0	8	0	0
3	G	12	0	16	5	0
3	H	18	0	24	6	0
3	I	12	0	16	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	J	12	0	16	0	0
3	K	24	0	32	7	0
3	L	12	0	16	1	0
3	M	18	0	23	5	1
3	N	24	0	31	4	0
3	O	12	0	16	0	0
3	P	12	0	15	3	0
4	A	96	0	0	0	0
4	B	96	0	0	1	0
4	C	103	0	0	1	1
4	D	110	0	0	2	0
4	E	88	0	0	2	0
4	F	85	0	0	0	0
4	G	89	0	0	2	0
4	H	104	0	0	1	0
4	I	130	0	0	3	0
4	J	121	0	0	0	0
4	K	117	0	0	0	0
4	L	102	0	0	1	0
4	M	92	0	0	3	0
4	N	92	0	0	0	1
4	O	85	0	0	0	0
4	P	108	0	0	0	1
All	All	24406	0	21898	154	4

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 3.

The worst 5 of 154 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:N:88:GLN:O	1:N:90:SER:N	2.01	0.92
1:E:156:GLN:NE2	3:E:3002:GOL:O1	2.09	0.84
1:C:156:GLN:HE22	3:C:3036:GOL:H32	1.44	0.82
1:K:49:GLN:HE21	3:K:3011:GOL:C1	1.92	0.81
1:M:123:TYR:H	3:M:3043:GOL:H11	1.44	0.79

All (4) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:M:3043:GOL:O1	4:N:3085:HOH:O[5_555]	1.84	0.36
1:A:156:GLN:NE2	3:E:3004:GOL:O1[9_555]	2.16	0.04
1:A:52:GLU:OE1	3:C:3036:GOL:O2[5_555]	2.16	0.04
4:C:3070:HOH:O	4:P:3141:HOH:O[10_646]	2.18	0.02

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	168/172 (98%)	162 (96%)	4 (2%)	2 (1%)	13 7
1	B	168/172 (98%)	165 (98%)	2 (1%)	1 (1%)	25 19
1	C	167/172 (97%)	163 (98%)	4 (2%)	0	100 100
1	D	167/172 (97%)	162 (97%)	4 (2%)	1 (1%)	25 19
1	E	171/172 (99%)	164 (96%)	6 (4%)	1 (1%)	25 19
1	F	168/172 (98%)	162 (96%)	5 (3%)	1 (1%)	25 19
1	G	168/172 (98%)	167 (99%)	1 (1%)	0	100 100
1	H	167/172 (97%)	161 (96%)	4 (2%)	2 (1%)	13 7
1	I	167/172 (97%)	166 (99%)	1 (1%)	0	100 100
1	J	167/172 (97%)	164 (98%)	2 (1%)	1 (1%)	25 19
1	K	167/172 (97%)	165 (99%)	1 (1%)	1 (1%)	25 19
1	L	168/172 (98%)	163 (97%)	4 (2%)	1 (1%)	25 19
1	M	166/172 (96%)	162 (98%)	4 (2%)	0	100 100
1	N	167/172 (97%)	163 (98%)	3 (2%)	1 (1%)	25 19
1	O	167/172 (97%)	161 (96%)	6 (4%)	0	100 100
1	P	167/172 (97%)	161 (96%)	4 (2%)	2 (1%)	13 7
All	All	2680/2752 (97%)	2611 (97%)	55 (2%)	14 (0%)	29 23

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	90	SER
1	D	88	GLN
1	E	90	SER
1	H	87	ASP
1	H	88	GLN

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	150/151 (99%)	147 (98%)	3 (2%)	55	58
1	B	150/151 (99%)	143 (95%)	7 (5%)	26	22
1	C	149/151 (99%)	146 (98%)	3 (2%)	55	58
1	D	149/151 (99%)	147 (99%)	2 (1%)	69	74
1	E	153/151 (101%)	146 (95%)	7 (5%)	27	23
1	F	150/151 (99%)	150 (100%)	0	100	100
1	G	150/151 (99%)	147 (98%)	3 (2%)	55	58
1	H	149/151 (99%)	145 (97%)	4 (3%)	44	46
1	I	149/151 (99%)	147 (99%)	2 (1%)	69	74
1	J	149/151 (99%)	144 (97%)	5 (3%)	37	36
1	K	149/151 (99%)	146 (98%)	3 (2%)	55	58
1	L	150/151 (99%)	144 (96%)	6 (4%)	31	29
1	M	148/151 (98%)	143 (97%)	5 (3%)	37	36
1	N	149/151 (99%)	148 (99%)	1 (1%)	84	88
1	O	149/151 (99%)	145 (97%)	4 (3%)	44	46
1	P	149/151 (99%)	142 (95%)	7 (5%)	26	22
All	All	2392/2416 (99%)	2330 (97%)	62 (3%)	46	48

5 of 62 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	94	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	J	96	VAL
1	P	89	GLN
1	J	3	GLN
1	K	89	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 36 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	166	GLN
1	K	49	GLN
1	P	88	GLN
1	I	88	GLN
1	K	89	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)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 80 ligands modelled in this entry, 36 are monoatomic - leaving 44 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	P	3034	-	5,5,5	1.24	1 (20%)	5,5,5	0.82	0
3	GOL	N	3008	-	5,5,5	0.84	0	5,5,5	0.37	0
3	GOL	N	3044	-	5,5,5	0.53	0	5,5,5	0.71	0
3	GOL	K	3041	-	5,5,5	0.77	0	5,5,5	1.07	0
3	GOL	D	3018	-	5,5,5	0.78	0	5,5,5	1.32	0
3	GOL	D	3005	-	5,5,5	0.48	0	5,5,5	1.15	1 (20%)
3	GOL	E	3002	-	5,5,5	0.89	0	5,5,5	0.59	0
3	GOL	B	3015	1	5,5,5	0.84	0	5,5,5	1.29	1 (20%)
3	GOL	O	3026	-	5,5,5	0.83	0	5,5,5	1.29	1 (20%)
3	GOL	I	3031	-	5,5,5	0.70	0	5,5,5	0.96	0
3	GOL	C	3021	-	5,5,5	0.54	0	5,5,5	0.50	0
3	GOL	E	3022	-	5,5,5	0.76	0	5,5,5	1.36	0
3	GOL	E	3035	-	5,5,5	0.67	0	5,5,5	0.99	0
3	GOL	F	3039	-	5,5,5	1.02	1 (20%)	5,5,5	1.03	0
3	GOL	M	3007	-	5,5,5	1.31	1 (20%)	5,5,5	1.70	2 (40%)
3	GOL	B	3003	-	5,5,5	1.35	1 (20%)	5,5,5	1.61	1 (20%)
3	GOL	H	3016	-	5,5,5	0.81	0	5,5,5	1.19	1 (20%)
3	GOL	M	3042	-	5,5,5	0.80	0	5,5,5	0.90	0
3	GOL	N	3029	-	5,5,5	0.71	0	5,5,5	0.59	0
3	GOL	D	3019	-	5,5,5	0.52	0	5,5,5	0.58	0
3	GOL	D	3020	-	5,5,5	0.51	0	5,5,5	0.75	0
3	GOL	L	3028	-	5,5,5	0.54	0	5,5,5	0.81	0
3	GOL	M	3043	-	5,5,5	0.49	0	5,5,5	0.75	0
3	GOL	K	3010	-	5,5,5	0.61	0	5,5,5	1.04	0
3	GOL	J	3033	-	5,5,5	0.67	0	5,5,5	1.26	1 (20%)
3	GOL	I	3032	-	5,5,5	0.58	0	5,5,5	0.64	0
3	GOL	N	3012	-	5,5,5	1.21	1 (20%)	5,5,5	1.54	2 (40%)
3	GOL	B	3024	-	5,5,5	0.52	0	5,5,5	0.46	0
3	GOL	L	3027	-	5,5,5	0.36	0	5,5,5	1.11	0
3	GOL	K	3009	-	5,5,5	0.62	0	5,5,5	1.71	2 (40%)
3	GOL	O	3013	-	5,5,5	0.42	0	5,5,5	0.91	0
3	GOL	C	3036	-	5,5,5	1.02	0	5,5,5	2.11	2 (40%)
3	GOL	H	3025	-	5,5,5	0.63	0	5,5,5	0.64	0
3	GOL	G	3038	-	5,5,5	0.63	0	5,5,5	0.73	0
3	GOL	K	3011	-	5,5,5	0.78	0	5,5,5	0.37	0
3	GOL	E	3004	1	5,5,5	0.52	0	5,5,5	0.57	0
3	GOL	A	3037	-	5,5,5	0.96	0	5,5,5	1.57	2 (40%)
3	GOL	A	3040	-	5,5,5	0.92	0	5,5,5	1.34	0
3	GOL	G	3001	-	5,5,5	0.61	0	5,5,5	0.89	0
3	GOL	J	3030	-	5,5,5	0.45	0	5,5,5	1.13	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
3	GOL	P	3014	-	5,5,5	0.26	0	5,5,5	0.83	0
3	GOL	B	3023	-	5,5,5	0.85	0	5,5,5	1.17	0
3	GOL	H	3006	-	5,5,5	0.92	0	5,5,5	1.13	0
3	GOL	B	3017	-	5,5,5	0.50	0	5,5,5	0.85	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	P	3034	-	-	2/4/4/4	-
3	GOL	N	3008	-	-	0/4/4/4	-
3	GOL	N	3044	-	-	4/4/4/4	-
3	GOL	K	3041	-	-	2/4/4/4	-
3	GOL	D	3018	-	-	3/4/4/4	-
3	GOL	D	3005	-	-	2/4/4/4	-
3	GOL	E	3002	-	-	2/4/4/4	-
3	GOL	B	3015	1	-	2/4/4/4	-
3	GOL	O	3026	-	-	2/4/4/4	-
3	GOL	I	3031	-	-	2/4/4/4	-
3	GOL	C	3021	-	-	4/4/4/4	-
3	GOL	E	3022	-	-	1/4/4/4	-
3	GOL	E	3035	-	-	4/4/4/4	-
3	GOL	F	3039	-	-	4/4/4/4	-
3	GOL	M	3007	-	-	2/4/4/4	-
3	GOL	B	3003	-	-	4/4/4/4	-
3	GOL	H	3016	-	-	1/4/4/4	-
3	GOL	M	3042	-	-	2/4/4/4	-
3	GOL	N	3029	-	-	0/4/4/4	-
3	GOL	D	3019	-	-	4/4/4/4	-
3	GOL	D	3020	-	-	3/4/4/4	-
3	GOL	L	3028	-	-	2/4/4/4	-
3	GOL	M	3043	-	-	0/4/4/4	-
3	GOL	K	3010	-	-	4/4/4/4	-
3	GOL	J	3033	-	-	3/4/4/4	-
3	GOL	I	3032	-	-	3/4/4/4	-
3	GOL	N	3012	-	-	3/4/4/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GOL	B	3024	-	-	0/4/4/4	-
3	GOL	L	3027	-	-	4/4/4/4	-
3	GOL	K	3009	-	-	0/4/4/4	-
3	GOL	O	3013	-	-	3/4/4/4	-
3	GOL	C	3036	-	-	2/4/4/4	-
3	GOL	H	3025	-	-	2/4/4/4	-
3	GOL	G	3038	-	-	0/4/4/4	-
3	GOL	K	3011	-	-	2/4/4/4	-
3	GOL	E	3004	1	-	2/4/4/4	-
3	GOL	A	3037	-	-	1/4/4/4	-
3	GOL	A	3040	-	-	2/4/4/4	-
3	GOL	G	3001	-	-	2/4/4/4	-
3	GOL	J	3030	-	-	4/4/4/4	-
3	GOL	P	3014	-	-	0/4/4/4	-
3	GOL	B	3023	-	-	2/4/4/4	-
3	GOL	H	3006	-	-	2/4/4/4	-
3	GOL	B	3017	-	-	1/4/4/4	-

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	M	3007	GOL	O2-C2	-2.88	1.34	1.43
3	N	3012	GOL	O2-C2	-2.59	1.35	1.43
3	P	3034	GOL	O2-C2	-2.43	1.36	1.43
3	B	3003	GOL	O2-C2	-2.23	1.36	1.43
3	F	3039	GOL	O2-C2	-2.08	1.37	1.43

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	3036	GOL	C3-C2-C1	3.57	125.59	111.70
3	B	3003	GOL	O3-C3-C2	-3.30	94.37	110.20
3	K	3009	GOL	O1-C1-C2	-3.11	95.29	110.20
3	M	3007	GOL	C3-C2-C1	2.68	122.14	111.70
3	A	3037	GOL	O3-C3-C2	2.62	122.78	110.20

There are no chirality outliers.

5 of 94 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	D	3020	GOL	C1-C2-C3-O3
3	C	3021	GOL	O1-C1-C2-C3
3	C	3021	GOL	C1-C2-C3-O3
3	K	3041	GOL	O1-C1-C2-C3
3	F	3039	GOL	O1-C1-C2-O2

There are no ring outliers.

21 monomers are involved in 55 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	P	3034	GOL	1	0
3	N	3008	GOL	2	0
3	D	3005	GOL	4	0
3	E	3002	GOL	3	0
3	B	3015	GOL	1	0
3	M	3007	GOL	2	0
3	B	3003	GOL	1	0
3	H	3016	GOL	1	0
3	N	3029	GOL	1	0
3	M	3043	GOL	3	1
3	N	3012	GOL	1	0
3	L	3027	GOL	1	0
3	K	3009	GOL	2	0
3	C	3036	GOL	4	1
3	K	3011	GOL	5	0
3	E	3004	GOL	4	1
3	A	3040	GOL	3	0
3	G	3001	GOL	5	0
3	P	3014	GOL	2	0
3	B	3023	GOL	1	0
3	H	3006	GOL	5	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{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	168/172 (97%)	0.36	12 (7%)	16	15	24, 36, 59, 73	11 (6%)
1	B	168/172 (97%)	0.34	7 (4%)	36	35	21, 36, 62, 73	10 (5%)
1	C	168/172 (97%)	0.45	11 (6%)	18	18	23, 35, 61, 71	9 (5%)
1	D	168/172 (97%)	0.32	7 (4%)	36	35	23, 35, 56, 66	10 (5%)
1	E	168/172 (97%)	0.34	12 (7%)	16	15	21, 36, 61, 73	13 (7%)
1	F	168/172 (97%)	0.23	10 (5%)	21	20	24, 36, 63, 71	14 (8%)
1	G	168/172 (97%)	0.28	8 (4%)	30	29	24, 38, 62, 70	13 (7%)
1	H	168/172 (97%)	0.25	5 (2%)	50	49	23, 35, 59, 69	14 (8%)
1	I	168/172 (97%)	0.15	3 (1%)	68	66	23, 34, 45, 59	12 (7%)
1	J	168/172 (97%)	0.30	7 (4%)	36	35	22, 35, 59, 67	9 (5%)
1	K	168/172 (97%)	0.23	5 (2%)	50	49	23, 34, 57, 68	11 (6%)
1	L	168/172 (97%)	0.44	14 (8%)	11	10	22, 35, 54, 64	13 (7%)
1	M	167/172 (97%)	0.38	10 (5%)	21	20	23, 36, 64, 76	10 (5%)
1	N	168/172 (97%)	0.28	6 (3%)	42	42	23, 35, 61, 72	10 (5%)
1	O	168/172 (97%)	0.36	8 (4%)	30	29	24, 36, 61, 73	9 (5%)
1	P	168/172 (97%)	0.25	5 (2%)	50	49	22, 35, 59, 72	12 (7%)
All	All	2687/2752 (97%)	0.31	130 (4%)	30	29	21, 35, 60, 76	180 (6%)

The worst 5 of 130 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	O	87	ASP	6.2
1	D	90	SER	5.3
1	C	86	GLN	5.1
1	K	85	LEU	4.7
1	C	130	THR	4.6

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

There are no non-standard protein/DNA/RNA residues in this entry.

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^{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	O	3026	6/6	0.49	0.25	60,62,63,64	0
3	GOL	D	3018	6/6	0.60	0.25	50,57,58,59	0
3	GOL	O	3013	6/6	0.67	0.26	58,60,62,66	0
2	CL	C	2011	1/1	0.67	0.45	70,70,70,70	1
3	GOL	B	3024	6/6	0.69	0.22	68,68,69,69	0
3	GOL	A	3040	6/6	0.72	0.24	41,54,57,58	0
2	CL	M	2017	1/1	0.72	0.39	61,61,61,61	1
3	GOL	C	3036	6/6	0.73	0.27	47,49,52,55	0
2	CL	C	2031	1/1	0.73	0.39	61,61,61,61	1
3	GOL	P	3014	6/6	0.74	0.19	42,47,52,55	0
3	GOL	A	3037	6/6	0.75	0.28	41,48,52,53	0
2	CL	J	2016	1/1	0.75	0.42	55,55,55,55	1
2	CL	O	2036	1/1	0.75	0.39	66,66,66,66	1
3	GOL	E	3035	6/6	0.76	0.17	53,57,59,60	0
3	GOL	N	3044	6/6	0.76	0.22	66,68,69,70	0
3	GOL	E	3004	6/6	0.77	0.24	51,54,58,60	0
2	CL	H	2014	1/1	0.77	0.38	61,61,61,61	1
2	CL	H	2024	1/1	0.77	0.32	53,53,53,53	1
3	GOL	P	3034	6/6	0.78	0.26	38,48,52,57	0
3	GOL	H	3016	6/6	0.78	0.18	49,53,56,58	0
3	GOL	D	3020	6/6	0.79	0.20	44,56,58,59	0
3	GOL	L	3028	6/6	0.79	0.22	56,57,59,60	0
3	GOL	B	3023	6/6	0.79	0.22	46,52,54,58	0
2	CL	F	2034	1/1	0.79	0.32	64,64,64,64	1
3	GOL	J	3033	6/6	0.80	0.19	47,58,59,60	0
3	GOL	G	3001	6/6	0.80	0.23	49,51,54,54	0
3	GOL	B	3017	6/6	0.80	0.20	55,57,60,61	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	CL	I	2015	1/1	0.80	0.39	59,59,59,59	1
3	GOL	B	3015	6/6	0.81	0.24	48,52,55,56	0
2	CL	D	2022	1/1	0.81	0.32	52,52,52,52	1
3	GOL	I	3032	6/6	0.81	0.16	50,55,56,57	0
2	CL	N	2030	1/1	0.81	0.35	62,62,62,62	1
3	GOL	F	3039	6/6	0.81	0.30	43,47,48,50	0
3	GOL	H	3025	6/6	0.82	0.18	62,63,64,65	0
2	CL	K	2027	1/1	0.82	0.29	52,52,52,52	1
3	GOL	M	3042	6/6	0.83	0.29	47,52,54,56	0
3	GOL	N	3029	6/6	0.83	0.17	55,56,58,58	0
2	CL	B	2010	1/1	0.83	0.36	54,54,54,54	1
3	GOL	L	3027	6/6	0.83	0.16	44,54,57,57	0
3	GOL	D	3019	6/6	0.83	0.20	52,54,55,56	0
2	CL	E	2032	1/1	0.83	0.36	58,58,58,58	1
3	GOL	D	3005	6/6	0.83	0.23	49,54,59,59	0
3	GOL	N	3008	6/6	0.84	0.22	54,55,56,60	0
3	GOL	K	3010	6/6	0.84	0.18	55,59,60,63	0
3	GOL	E	3022	6/6	0.84	0.26	54,55,56,57	0
3	GOL	E	3002	6/6	0.84	0.27	52,53,55,57	0
3	GOL	N	3012	6/6	0.85	0.19	49,50,53,54	0
2	CL	L	2028	1/1	0.85	0.25	58,58,58,58	1
3	GOL	J	3030	6/6	0.85	0.15	46,54,56,60	0
3	GOL	B	3003	6/6	0.85	0.25	35,38,44,47	0
2	CL	E	2013	1/1	0.86	0.46	62,62,62,62	1
3	GOL	K	3041	6/6	0.86	0.15	38,50,52,57	0
3	GOL	C	3021	6/6	0.86	0.17	54,55,56,56	0
2	CL	C	2021	1/1	0.86	0.25	53,53,53,53	1
3	GOL	M	3007	6/6	0.87	0.26	46,47,52,56	0
3	GOL	G	3038	6/6	0.87	0.17	51,57,58,60	0
3	GOL	M	3043	6/6	0.88	0.25	46,51,51,55	0
2	CL	J	2026	1/1	0.88	0.33	56,56,56,56	1
3	GOL	H	3006	6/6	0.88	0.23	45,49,53,53	0
2	CL	M	2029	1/1	0.89	0.19	50,50,50,50	1
3	GOL	K	3009	6/6	0.89	0.21	42,44,48,53	0
3	GOL	K	3011	6/6	0.90	0.26	50,52,55,57	0
3	GOL	I	3031	6/6	0.91	0.18	48,53,56,59	0
2	CL	D	2012	1/1	0.91	0.31	55,55,55,55	1
2	CL	A	2020	1/1	0.92	0.30	57,57,57,57	1
2	CL	L	2009	1/1	0.92	0.45	57,57,57,57	1
2	CL	G	2033	1/1	0.93	0.19	52,52,52,52	1
2	CL	I	2025	1/1	0.94	0.26	47,47,47,47	1
2	CL	O	2018	1/1	0.94	0.31	55,55,55,55	1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	CL	P	2019	1/1	0.94	0.42	53,53,53,53	1
2	CL	F	2023	1/1	0.96	0.36	53,53,53,53	1
2	CL	F	2035	1/1	0.97	0.25	42,42,42,42	1
2	CL	I	2001	1/1	0.99	0.14	19,19,19,19	0
2	CL	I	2007	1/1	0.99	0.12	24,24,24,24	0
2	CL	C	2006	1/1	0.99	0.14	27,27,27,27	0
2	CL	J	2002	1/1	0.99	0.13	21,21,21,21	0
2	CL	H	2005	1/1	1.00	0.12	20,20,20,20	0
2	CL	D	2008	1/1	1.00	0.11	24,24,24,24	0
2	CL	K	2003	1/1	1.00	0.13	25,25,25,25	0
2	CL	C	2004	1/1	1.00	0.16	23,23,23,23	0

6.5 Other polymers (i)

There are no such residues in this entry.