Full wwPDB X-ray Structure Validation Report (i)

Jun 17, 2024 – 06:20 PM EDT

PDB ID	:	2YEU
Title	:	Structural and functional insights of DR2231 protein, the MazG-like nucleoside triphosphate pyrophosphohydrolase from Deinococcus radiodurans, complex with Gd
Authors	:	Goncalves, A.M.D.; de Sanctis, D.; McSweeney, S.M.
Deposited on	:	2011-03-30
Resolution	:	2.00 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

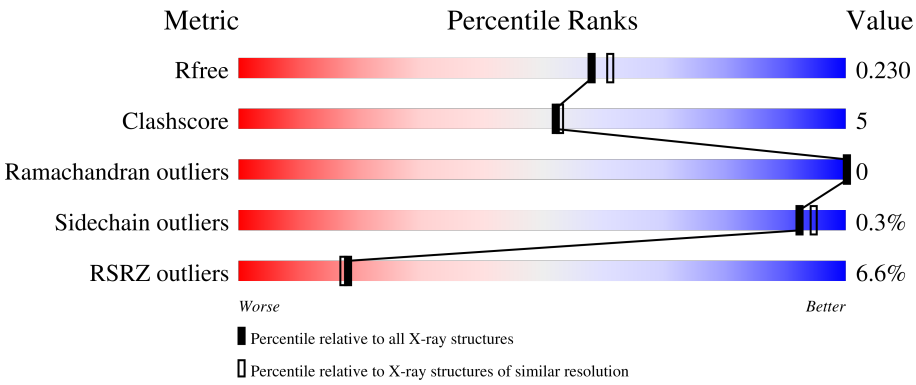
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 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	154	5% 91% 6% •
1	B	154	6% 88% 7% 5%
1	C	154	5% 86% 10% •
1	D	154	5% 90% 8% •
1	E	154	9% 87% 8% • •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	154	9% 84% 10% 5%

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
3	GOL	B	1148	-	-	X	-
3	GOL	B	1151	-	-	X	-
3	GOL	D	1147	-	X	-	-
3	GOL	D	1150	-	-	X	-

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 14606 atoms, of which 6898 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 DR2231.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	149	2272	713	1129	211	217	2	0	0	0
1	B	147	2240	703	1114	209	212	2	0	0	0
1	C	149	2275	714	1131	211	217	2	0	0	0
1	D	150	2283	716	1135	212	218	2	0	0	0
1	E	148	2245	705	1113	210	216	1	0	0	0
1	F	146	2229	700	1108	208	211	2	0	0	0

There are 36 discrepancies between the modelled and reference sequences:

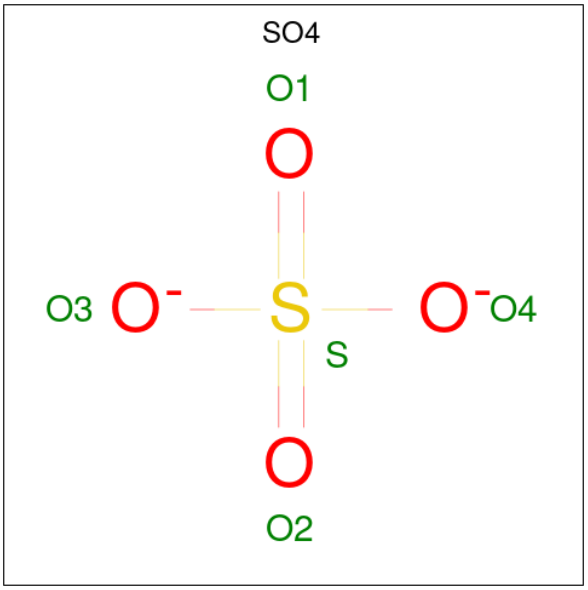
Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	GLY	-	expression tag	UNP Q9RS96
A	-4	ILE	-	expression tag	UNP Q9RS96
A	-3	ASP	-	expression tag	UNP Q9RS96
A	-2	PRO	-	expression tag	UNP Q9RS96
A	-1	PHE	-	expression tag	UNP Q9RS96
A	0	THR	-	expression tag	UNP Q9RS96
B	-5	GLY	-	expression tag	UNP Q9RS96
B	-4	ILE	-	expression tag	UNP Q9RS96
B	-3	ASP	-	expression tag	UNP Q9RS96
B	-2	PRO	-	expression tag	UNP Q9RS96
B	-1	PHE	-	expression tag	UNP Q9RS96
B	0	THR	-	expression tag	UNP Q9RS96
C	-5	GLY	-	expression tag	UNP Q9RS96
C	-4	ILE	-	expression tag	UNP Q9RS96
C	-3	ASP	-	expression tag	UNP Q9RS96
C	-2	PRO	-	expression tag	UNP Q9RS96
C	-1	PHE	-	expression tag	UNP Q9RS96

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	0	THR	-	expression tag	UNP Q9RS96
D	-5	GLY	-	expression tag	UNP Q9RS96
D	-4	ILE	-	expression tag	UNP Q9RS96
D	-3	ASP	-	expression tag	UNP Q9RS96
D	-2	PRO	-	expression tag	UNP Q9RS96
D	-1	PHE	-	expression tag	UNP Q9RS96
D	0	THR	-	expression tag	UNP Q9RS96
E	-5	GLY	-	expression tag	UNP Q9RS96
E	-4	ILE	-	expression tag	UNP Q9RS96
E	-3	ASP	-	expression tag	UNP Q9RS96
E	-2	PRO	-	expression tag	UNP Q9RS96
E	-1	PHE	-	expression tag	UNP Q9RS96
E	0	THR	-	expression tag	UNP Q9RS96
F	-5	GLY	-	expression tag	UNP Q9RS96
F	-4	ILE	-	expression tag	UNP Q9RS96
F	-3	ASP	-	expression tag	UNP Q9RS96
F	-2	PRO	-	expression tag	UNP Q9RS96
F	-1	PHE	-	expression tag	UNP Q9RS96
F	0	THR	-	expression tag	UNP Q9RS96

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



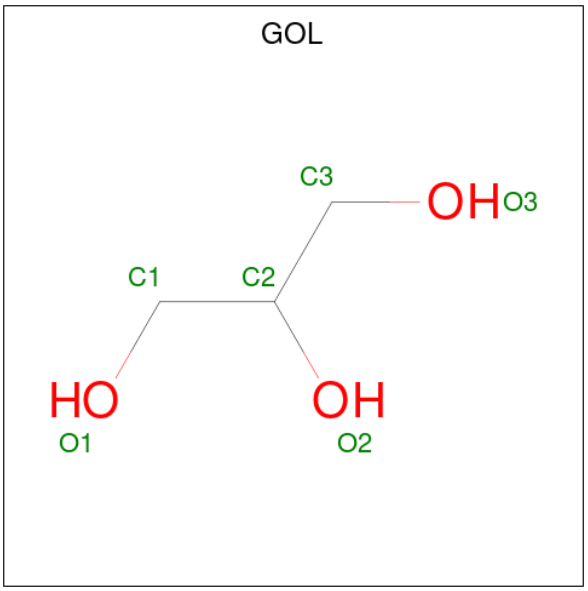
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
2	A	1	5	4	1	0	0
			Total	O	S		
2	A	1	5	4	1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
2	A	1	5	4	1	0	0
2	B	1	5	4	1	0	0
2	B	1	5	4	1	0	0
2	C	1	5	4	1	0	0
2	C	1	5	4	1	0	0
2	C	1	5	4	1	0	0
2	D	1	5	4	1	0	0
2	D	1	5	4	1	0	0
2	E	1	5	4	1	0	0
2	E	1	5	4	1	0	0
2	E	1	5	4	1	0	0
2	F	1	5	4	1	0	0
2	F	1	5	4	1	0	0
2	F	1	5	4	1	0	0

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



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
3	A	1	14	3	8	3	0	0
3	A	1	14	3	8	3	0	0
3	A	1	14	3	8	3	0	0
3	A	1	14	3	8	3	0	0
3	B	1	14	3	8	3	0	0
3	B	1	14	3	8	3	0	0
3	B	1	14	3	8	3	0	0
3	B	1	14	3	8	3	0	0
3	B	1	14	3	8	3	0	0
3	C	1	14	3	8	3	0	0
3	D	1	14	3	8	3	0	0
3	D	1	14	3	8	3	0	0
3	D	1	14	3	8	3	0	0
3	D	1	14	3	8	3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
3	D	1	14	3	8	3	0	0
			Total	C	H	O		
3	D	1	14	3	8	3	0	0
			Total	C	H	O		
3	E	1	14	3	8	3	0	0
			Total	C	H	O		
3	E	1	14	3	8	3	0	0
			Total	C	H	O		
3	F	1	14	3	8	3	0	0
			Total	C	H	O		
3	F	1	14	3	8	3	0	0
			Total	C	H	O		
3	F	1	14	3	8	3	0	0

- Molecule 4 is GADOLINIUM ATOM (three-letter code: GD) (formula: Gd).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Gd		
4	A	1	1	1	0	0
			Total	Gd		
4	B	1	1	1	0	0
			Total	Gd		
4	C	1	1	1	0	0
			Total	Gd		
4	D	1	1	1	0	0
			Total	Gd		
4	E	1	1	1	0	0
			Total	Gd		
4	F	1	1	1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	124	124	124	0	0
			Total	O		
5	B	114	114	114	0	0
			Total	O		
5	C	143	143	143	0	0
			Total	O		
5	D	123	123	123	0	0

Continued on next page...

Continued from previous page...

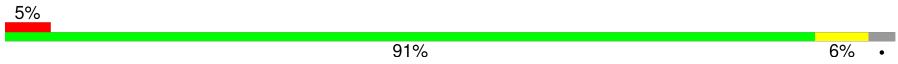
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	E	85	85	85	0	0
			Total	O		
5	F	93	93	93	0	0

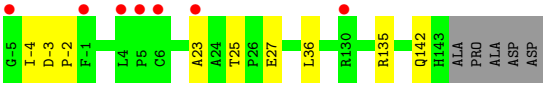
3 Residue-property plots (i)

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: DR2231

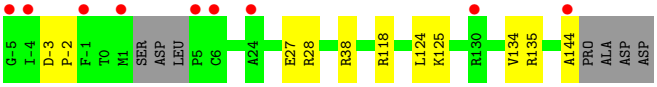
Chain A:





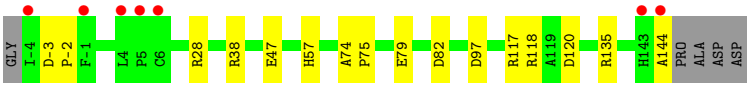
- Molecule 1: DR2231

Chain B:



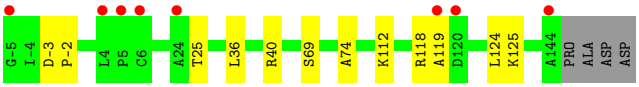
- Molecule 1: DR2231

Chain C:



- Molecule 1: DR2231

Chain D:

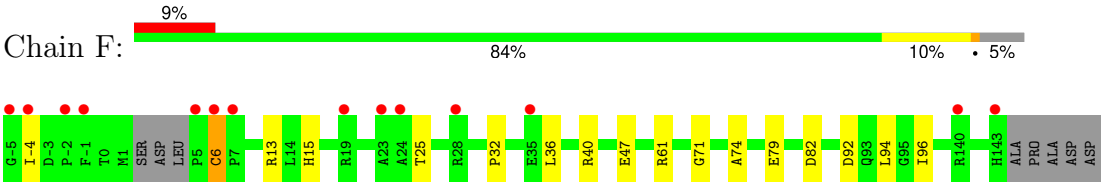


- Molecule 1: DR2231

Chain E:

- Molecule 1: DR2231

Chain F:


9%
84%
10%
5%
G-5
I-4
D-3
P-2
F-1
T0
M1
SER
ASP
LEU
P5
C6
P7
R13
L14
H15
R19
A23
A24
T25
R28
P32
E35
L36
R40
E47
R61
G71
A74
E79
D82
D92
Q93
L94
G95
I96
R140
H143
ALA
PRO
ALA
ASP
ASP

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	67.82Å 110.75Å 165.98Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	92.06 – 2.00 20.00 – 2.00	Depositor EDS
% Data completeness (in resolution range)	96.3 (92.06-2.00) 96.4 (20.00-2.00)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.19 (at 2.01Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, R_{free}	0.188 , 0.232 0.187 , 0.230	Depositor DCC
R_{free} test set	4090 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	19.9	Xtriage
Anisotropy	0.044	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.44 , 46.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	14606	wwPDB-VP
Average B, all atoms (Å^2)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.62% 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: SO4, GOL, GD

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	0.78	0/1169	0.77	0/1593
1	B	0.85	0/1151	0.75	0/1566
1	C	0.82	0/1170	0.77	2/1595 (0.1%)
1	D	0.85	0/1174	0.76	0/1600
1	E	0.70	0/1156	0.72	1/1572 (0.1%)
1	F	0.74	0/1146	0.76	1/1559 (0.1%)
All	All	0.79	0/6966	0.76	4/9485 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	52	ARG	NE-CZ-NH2	-6.43	117.09	120.30
1	F	6	CYS	CA-CB-SG	6.20	125.16	114.00
1	C	97	ASP	CB-CG-OD1	5.55	123.29	118.30
1	C	117	ARG	NE-CZ-NH1	5.38	122.99	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1143	1129	1128	11	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	1126	1114	1113	19	0
1	C	1144	1131	1130	12	0
1	D	1148	1135	1133	14	0
1	E	1132	1113	1112	12	0
1	F	1121	1108	1108	11	0
2	A	15	0	0	0	0
2	B	10	0	0	1	0
2	C	15	0	0	0	0
2	D	10	0	0	0	0
2	E	15	0	0	1	0
2	F	15	0	0	0	0
3	A	24	32	32	4	0
3	B	30	40	40	14	0
3	C	6	8	8	1	0
3	D	36	48	48	7	0
3	E	12	16	16	3	0
3	F	18	24	24	1	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
4	D	1	0	0	0	0
4	E	1	0	0	0	0
4	F	1	0	0	0	0
5	A	124	0	0	1	1
5	B	114	0	0	0	0
5	C	143	0	0	1	0
5	D	123	0	0	3	1
5	E	85	0	0	0	0
5	F	93	0	0	1	0
All	All	7708	6898	6892	75	1

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 5.

All (75) 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:A:135:ARG:HH21	3:A:1149:GOL:H31	1.24	1.03
1:B:118:ARG:HH22	3:B:1149:GOL:H2	1.32	0.92
1:D:69:SER:HB2	3:D:1147:GOL:O2	1.77	0.84
1:B:135:ARG:H	3:B:1148:GOL:H31	1.43	0.84

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:125:LYS:HE3	3:B:1151:GOL:C1	2.18	0.74
1:E:19:ARG:NH2	3:E:1149:GOL:H2	2.03	0.73
1:C:28:ARG:HH22	1:C:144:ALA:HB3	1.54	0.73
1:D:25:THR:HG22	5:D:2026:HOH:O	1.92	0.70
1:E:19:ARG:HH21	3:E:1149:GOL:H2	1.57	0.69
1:B:135:ARG:H	3:B:1148:GOL:C3	2.05	0.69
1:A:-4:ILE:HD12	1:A:-4:ILE:O	1.96	0.66
1:B:135:ARG:HD2	3:B:1148:GOL:H11	1.77	0.66
1:D:69:SER:CB	3:D:1147:GOL:O2	2.43	0.65
1:D:125:LYS:HD2	3:D:1150:GOL:H2	1.78	0.65
1:A:142:GLN:HG2	5:A:2030:HOH:O	1.97	0.64
1:B:134:VAL:N	3:B:1148:GOL:H32	2.13	0.63
1:A:27:GLU:OE2	3:A:1147:GOL:H12	1.98	0.63
1:B:125:LYS:HE3	3:B:1151:GOL:H11	1.82	0.62
1:B:134:VAL:HB	3:B:1148:GOL:H12	1.81	0.61
1:A:-3:ASP:HB2	1:A:-2:PRO:CD	2.31	0.59
1:E:106:HIS:CD2	1:F:13:ARG:HD2	2.39	0.58
1:A:-4:ILE:HD12	1:A:-4:ILE:C	2.24	0.57
1:D:112:LYS:HD3	1:E:-1:PHE:CZ	2.41	0.56
1:C:38:ARG:HD3	2:E:1147:SO4:O3	2.04	0.56
1:C:57:HIS:HB2	5:C:2065:HOH:O	2.06	0.56
1:E:28:ARG:HH12	1:E:144:ALA:HB2	1.70	0.56
1:B:125:LYS:HE3	3:B:1151:GOL:H12	1.89	0.55
1:B:134:VAL:H	3:B:1148:GOL:H32	1.72	0.54
1:A:-3:ASP:HB2	1:A:-2:PRO:HD2	1.89	0.53
1:F:71:GLY:HA3	3:F:1147:GOL:H31	1.90	0.53
1:D:124:LEU:HD22	3:D:1150:GOL:H31	1.89	0.53
1:A:135:ARG:NH2	3:A:1149:GOL:H31	2.07	0.53
1:C:-3:ASP:HB2	1:C:-2:PRO:CD	2.39	0.52
3:C:1148:GOL:H2	1:E:-2:PRO:HG2	1.92	0.52
1:F:15:HIS:CE1	1:F:25:THR:HG22	2.45	0.52
1:B:27:GLU:H	3:B:1150:GOL:C2	2.23	0.51
1:A:27:GLU:CD	3:A:1147:GOL:H12	2.31	0.51
3:D:1148:GOL:H31	5:D:2051:HOH:O	2.09	0.51
1:B:38:ARG:HD3	2:B:1145:SO4:O3	2.12	0.50
1:D:-3:ASP:HB2	1:D:-2:PRO:CD	2.42	0.49
1:D:25:THR:HG23	1:D:25:THR:O	2.12	0.49
1:E:52:ARG:HH22	3:E:1148:GOL:H2	1.77	0.49
1:C:118:ARG:HD2	1:C:120:ASP:OD1	2.13	0.49
1:C:-3:ASP:HB2	1:C:-2:PRO:HD2	1.94	0.49
1:F:36:LEU:O	1:F:40:ARG:HG2	2.13	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:47:GLU:CD	1:F:82:ASP:HB3	2.33	0.48
1:F:61:ARG:NH2	5:F:2037:HOH:O	2.46	0.48
1:B:-3:ASP:HB2	1:B:-2:PRO:HD2	1.95	0.47
1:D:118:ARG:O	1:D:119:ALA:C	2.53	0.47
1:E:3:ASP:OD2	1:E:5:PRO:N	2.49	0.46
1:C:28:ARG:HH12	1:C:144:ALA:N	2.13	0.46
1:D:124:LEU:CD2	3:D:1150:GOL:H31	2.46	0.46
1:B:-3:ASP:HB2	1:B:-2:PRO:CD	2.46	0.45
1:A:36:LEU:C	1:A:36:LEU:HD23	2.37	0.45
1:E:-3:ASP:HB2	1:E:-2:PRO:HD2	1.99	0.45
1:E:77:ALA:HB2	1:F:96:ILE:HD13	1.99	0.44
1:B:135:ARG:N	3:B:1148:GOL:H31	2.22	0.43
1:F:25:THR:HG21	1:F:92:ASP:OD2	2.18	0.43
1:D:25:THR:N	5:D:2026:HOH:O	2.52	0.43
1:D:36:LEU:HD23	1:D:40:ARG:HG2	2.00	0.43
1:C:135:ARG:HG2	1:D:74:ALA:CB	2.49	0.42
1:B:125:LYS:CE	3:B:1151:GOL:H11	2.47	0.42
1:B:118:ARG:HG3	1:B:124:LEU:HD21	2.00	0.42
1:D:124:LEU:HB3	3:D:1150:GOL:H31	2.02	0.42
1:F:32:PRO:HG3	1:F:94:LEU:CD2	2.49	0.42
1:C:74:ALA:HB3	1:C:75:PRO:HD3	2.01	0.42
1:C:79:GLU:OE1	1:C:79:GLU:HA	2.20	0.41
1:C:47:GLU:CD	1:C:82:ASP:HB3	2.40	0.41
1:A:23:ALA:O	1:A:25:THR:HG23	2.21	0.41
1:C:28:ARG:HH22	1:C:144:ALA:CB	2.30	0.41
1:E:138:ILE:HD12	1:F:74:ALA:HA	2.03	0.41
1:B:28:ARG:HH22	1:B:144:ALA:HB3	1.86	0.41
1:F:79:GLU:HA	1:F:79:GLU:OE1	2.21	0.40
1:B:27:GLU:H	3:B:1150:GOL:H2	1.85	0.40
1:E:130:ARG:HA	1:E:131:PRO:HD3	1.95	0.40

All (1) 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)
5:A:2071:HOH:O	5:D:2038:HOH:O[3_655]	1.71	0.49

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	147/154 (96%)	143 (97%)	4 (3%)	0	100	100
1	B	143/154 (93%)	140 (98%)	3 (2%)	0	100	100
1	C	147/154 (96%)	144 (98%)	3 (2%)	0	100	100
1	D	148/154 (96%)	145 (98%)	3 (2%)	0	100	100
1	E	142/154 (92%)	141 (99%)	1 (1%)	0	100	100
1	F	142/154 (92%)	141 (99%)	1 (1%)	0	100	100
All	All	869/924 (94%)	854 (98%)	15 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	117/120 (98%)	117 (100%)	0	100	100
1	B	114/120 (95%)	114 (100%)	0	100	100
1	C	117/120 (98%)	117 (100%)	0	100	100
1	D	117/120 (98%)	117 (100%)	0	100	100
1	E	115/120 (96%)	115 (100%)	0	100	100
1	F	114/120 (95%)	112 (98%)	2 (2%)	59	63
All	All	694/720 (96%)	692 (100%)	2 (0%)	92	95

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	-4	ILE
1	F	6	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 43 ligands modelled in this entry, 6 are monoatomic - leaving 37 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	$\#\|Z\| > 2$
2	SO4	D	1145	-	4,4,4	0.21	0	6,6,6	0.31	0
3	GOL	B	1147	-	5,5,5	0.95	0	5,5,5	0.59	0
2	SO4	A	1145	-	4,4,4	0.21	0	6,6,6	0.21	0
3	GOL	D	1152	-	5,5,5	0.82	0	5,5,5	1.13	0
3	GOL	F	1149	-	5,5,5	0.31	0	5,5,5	0.51	0
3	GOL	E	1149	-	5,5,5	0.34	0	5,5,5	0.42	0
3	GOL	D	1148	-	5,5,5	0.74	0	5,5,5	0.97	0
3	GOL	B	1148	-	5,5,5	0.44	0	5,5,5	0.38	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	GOL	D	1151	-	5,5,5	0.87	0	5,5,5	0.27	0
3	GOL	E	1148	-	5,5,5	1.31	0	5,5,5	0.86	0
3	GOL	F	1147	-	5,5,5	0.13	0	5,5,5	0.75	0
3	GOL	D	1149	-	5,5,5	0.57	0	5,5,5	0.45	0
2	SO4	C	1147	-	4,4,4	0.44	0	6,6,6	0.54	0
2	SO4	B	1145	-	4,4,4	0.40	0	6,6,6	0.40	0
3	GOL	D	1150	-	5,5,5	0.63	0	5,5,5	0.85	0
2	SO4	C	1146	-	4,4,4	0.23	0	6,6,6	0.39	0
2	SO4	A	1144	-	4,4,4	0.28	0	6,6,6	0.46	0
3	GOL	A	1147	-	5,5,5	0.47	0	5,5,5	0.16	0
2	SO4	D	1146	-	4,4,4	0.37	0	6,6,6	0.40	0
3	GOL	A	1148	-	5,5,5	0.40	0	5,5,5	0.62	0
3	GOL	B	1150	-	5,5,5	0.61	0	5,5,5	0.52	0
2	SO4	A	1146	-	4,4,4	0.33	0	6,6,6	0.27	0
2	SO4	F	1146	-	4,4,4	0.27	0	6,6,6	0.31	0
3	GOL	D	1147	-	5,5,5	0.82	0	5,5,5	2.89	4 (80%)
2	SO4	E	1147	-	4,4,4	0.24	0	6,6,6	0.30	0
3	GOL	B	1151	-	5,5,5	0.19	0	5,5,5	0.63	0
3	GOL	B	1149	-	5,5,5	0.71	0	5,5,5	0.55	0
3	GOL	A	1149	-	5,5,5	0.36	0	5,5,5	0.49	0
2	SO4	B	1146	-	4,4,4	0.21	0	6,6,6	0.20	0
2	SO4	E	1146	-	4,4,4	0.24	0	6,6,6	0.21	0
2	SO4	F	1144	-	4,4,4	0.16	0	6,6,6	0.28	0
3	GOL	C	1148	-	5,5,5	1.03	0	5,5,5	1.49	1 (20%)
3	GOL	F	1148	-	5,5,5	0.44	0	5,5,5	0.60	0
2	SO4	F	1145	-	4,4,4	0.35	0	6,6,6	0.39	0
2	SO4	C	1145	-	4,4,4	0.27	0	6,6,6	0.14	0
2	SO4	E	1145	-	4,4,4	0.19	0	6,6,6	0.51	0
3	GOL	A	1150	-	5,5,5	0.75	0	5,5,5	0.70	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	B	1147	-	-	2/4/4/4	-
3	GOL	D	1152	-	-	1/4/4/4	-
3	GOL	F	1149	-	-	0/4/4/4	-
3	GOL	E	1149	-	-	2/4/4/4	-
3	GOL	D	1148	-	-	1/4/4/4	-
3	GOL	B	1148	-	-	3/4/4/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GOL	D	1151	-	-	2/4/4/4	-
3	GOL	E	1148	-	-	3/4/4/4	-
3	GOL	F	1147	-	-	2/4/4/4	-
3	GOL	D	1149	-	-	4/4/4/4	-
3	GOL	D	1150	-	-	4/4/4/4	-
3	GOL	A	1147	-	-	0/4/4/4	-
3	GOL	A	1148	-	-	4/4/4/4	-
3	GOL	B	1150	-	-	2/4/4/4	-
3	GOL	D	1147	-	-	4/4/4/4	-
3	GOL	B	1151	-	-	2/4/4/4	-
3	GOL	B	1149	-	-	2/4/4/4	-
3	GOL	A	1149	-	-	4/4/4/4	-
3	GOL	C	1148	-	-	0/4/4/4	-
3	GOL	F	1148	-	-	1/4/4/4	-
3	GOL	A	1150	-	-	4/4/4/4	-

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	D	1147	GOL	O1-C1-C2	-4.08	92.03	110.38
3	D	1147	GOL	O3-C3-C2	-3.68	93.82	110.38
3	C	1148	GOL	C3-C2-C1	-2.43	102.88	111.80
3	D	1147	GOL	O2-C2-C3	2.30	118.69	109.18
3	D	1147	GOL	C3-C2-C1	2.01	119.18	111.80

There are no chirality outliers.

All (47) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1149	GOL	C1-C2-C3-O3
3	A	1150	GOL	O1-C1-C2-C3
3	A	1150	GOL	C1-C2-C3-O3
3	B	1147	GOL	O1-C1-C2-C3
3	B	1148	GOL	C1-C2-C3-O3
3	B	1148	GOL	O2-C2-C3-O3
3	B	1149	GOL	O1-C1-C2-O2
3	B	1151	GOL	O1-C1-C2-C3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
3	D	1147	GOL	O1-C1-C2-O2
3	D	1147	GOL	O2-C2-C3-O3
3	D	1149	GOL	O1-C1-C2-C3
3	D	1149	GOL	C1-C2-C3-O3
3	D	1150	GOL	O1-C1-C2-C3
3	D	1150	GOL	C1-C2-C3-O3
3	D	1151	GOL	O1-C1-C2-C3
3	E	1148	GOL	C1-C2-C3-O3
3	E	1149	GOL	C1-C2-C3-O3
3	F	1147	GOL	O1-C1-C2-C3
3	B	1150	GOL	O2-C2-C3-O3
3	A	1148	GOL	O1-C1-C2-C3
3	A	1148	GOL	C1-C2-C3-O3
3	A	1149	GOL	O1-C1-C2-C3
3	B	1149	GOL	O1-C1-C2-C3
3	B	1150	GOL	C1-C2-C3-O3
3	D	1147	GOL	O1-C1-C2-C3
3	D	1147	GOL	C1-C2-C3-O3
3	D	1152	GOL	O1-C1-C2-C3
3	F	1148	GOL	O1-C1-C2-C3
3	A	1148	GOL	O1-C1-C2-O2
3	A	1149	GOL	O1-C1-C2-O2
3	A	1149	GOL	O2-C2-C3-O3
3	A	1150	GOL	O1-C1-C2-O2
3	A	1150	GOL	O2-C2-C3-O3
3	B	1147	GOL	O1-C1-C2-O2
3	B	1151	GOL	O1-C1-C2-O2
3	D	1150	GOL	O1-C1-C2-O2
3	D	1150	GOL	O2-C2-C3-O3
3	D	1151	GOL	O1-C1-C2-O2
3	E	1149	GOL	O2-C2-C3-O3
3	D	1149	GOL	O2-C2-C3-O3
3	F	1147	GOL	O1-C1-C2-O2
3	D	1148	GOL	O1-C1-C2-O2
3	D	1149	GOL	O1-C1-C2-O2
3	E	1148	GOL	O2-C2-C3-O3
3	A	1148	GOL	O2-C2-C3-O3
3	B	1148	GOL	O1-C1-C2-O2
3	E	1148	GOL	O1-C1-C2-O2

There are no ring outliers.

15 monomers are involved in 32 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	E	1149	GOL	2	0
3	D	1148	GOL	1	0
3	B	1148	GOL	7	0
3	E	1148	GOL	1	0
3	F	1147	GOL	1	0
2	B	1145	SO4	1	0
3	D	1150	GOL	4	0
3	A	1147	GOL	2	0
3	B	1150	GOL	2	0
3	D	1147	GOL	2	0
2	E	1147	SO4	1	0
3	B	1151	GOL	4	0
3	B	1149	GOL	1	0
3	A	1149	GOL	2	0
3	C	1148	GOL	1	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 (i)

6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th 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	149/154 (96%)	-0.10	7 (4%)	31	30	11, 19, 35, 44	0
1	B	147/154 (95%)	-0.02	9 (6%)	21	20	11, 20, 52, 69	0
1	C	149/154 (96%)	-0.14	7 (4%)	31	30	10, 18, 35, 47	0
1	D	150/154 (97%)	-0.04	8 (5%)	26	25	11, 18, 36, 48	0
1	E	148/154 (96%)	0.25	14 (9%)	8	7	14, 27, 51, 71	0
1	F	146/154 (94%)	0.29	14 (9%)	8	7	15, 26, 60, 77	0
All	All	889/924 (96%)	0.04	59 (6%)	18	17	10, 20, 47, 77	0

All (59) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	24	ALA	12.0
1	F	-5	GLY	10.0
1	F	6	CYS	7.3
1	E	144	ALA	6.4
1	B	144	ALA	6.4
1	B	-5	GLY	6.1
1	E	5	PRO	5.8
1	D	144	ALA	5.6
1	E	3	ASP	5.4
1	D	-5	GLY	5.2
1	A	5	PRO	5.1
1	A	4	LEU	5.1
1	F	5	PRO	5.0
1	E	-5	GLY	4.6
1	C	144	ALA	4.4
1	A	6	CYS	4.4
1	F	-1	PHE	4.4
1	D	24	ALA	4.3
1	E	6	CYS	4.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	6	CYS	4.2
1	F	-4	ILE	4.1
1	E	-1	PHE	4.1
1	C	5	PRO	4.1
1	B	-4	ILE	4.1
1	E	130	ARG	4.1
1	A	-5	GLY	4.0
1	F	-2	PRO	4.0
1	B	5	PRO	3.8
1	C	4	LEU	3.8
1	F	23	ALA	3.8
1	B	1	MET	3.3
1	F	7	PRO	3.3
1	E	2	SER	3.2
1	A	130	ARG	3.1
1	E	63	ALA	3.1
1	A	23	ALA	3.0
1	D	4	LEU	2.9
1	B	24	ALA	2.8
1	C	-1	PHE	2.8
1	E	143	HIS	2.7
1	D	119	ALA	2.5
1	F	28	ARG	2.5
1	A	-1	PHE	2.5
1	C	-4	ILE	2.4
1	B	-1	PHE	2.4
1	E	140	ARG	2.3
1	C	143	HIS	2.3
1	E	7	PRO	2.3
1	F	19	ARG	2.3
1	D	5	PRO	2.3
1	B	130	ARG	2.2
1	E	119	ALA	2.2
1	C	6	CYS	2.2
1	F	140	ARG	2.2
1	D	120	ASP	2.1
1	E	0	THR	2.1
1	F	35	GLU	2.0
1	F	143	HIS	2.0
1	D	6	CYS	2.0

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 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, 95th 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	A	1150	6/6	0.54	0.34	40,42,46,46	0
3	GOL	B	1149	6/6	0.55	0.26	37,41,44,44	0
3	GOL	B	1147	6/6	0.69	0.28	42,44,46,46	0
3	GOL	D	1147	6/6	0.70	0.35	41,51,57,57	0
3	GOL	B	1150	6/6	0.79	0.31	36,38,41,41	0
3	GOL	E	1148	6/6	0.79	0.22	33,37,38,38	0
3	GOL	D	1149	6/6	0.80	0.25	39,43,47,47	0
3	GOL	A	1149	6/6	0.80	0.30	37,38,43,43	0
3	GOL	A	1147	6/6	0.82	0.44	47,49,51,51	0
3	GOL	C	1148	6/6	0.83	0.24	19,26,32,32	0
3	GOL	F	1148	6/6	0.83	0.26	36,40,41,41	0
2	SO4	F	1146	5/5	0.84	0.29	33,34,35,35	5
3	GOL	E	1149	6/6	0.84	0.26	42,45,47,47	0
3	GOL	F	1147	6/6	0.84	0.34	36,40,46,46	0
3	GOL	D	1150	6/6	0.84	0.28	28,33,39,39	0
3	GOL	D	1152	6/6	0.85	0.32	24,34,41,41	0
3	GOL	D	1148	6/6	0.86	0.20	19,30,32,32	0
2	SO4	A	1145	5/5	0.87	0.15	28,29,32,32	5
3	GOL	B	1151	6/6	0.87	0.29	39,40,41,41	0
2	SO4	A	1146	5/5	0.87	0.21	21,24,26,27	5
3	GOL	F	1149	6/6	0.88	0.17	34,37,42,42	0
3	GOL	D	1151	6/6	0.89	0.23	32,35,35,35	0
3	GOL	B	1148	6/6	0.90	0.19	40,42,44,44	0
3	GOL	A	1148	6/6	0.91	0.20	19,31,38,38	0
2	SO4	C	1147	5/5	0.92	0.39	56,56,57,58	0
2	SO4	E	1146	5/5	0.92	0.14	37,38,38,38	5
2	SO4	E	1147	5/5	0.93	0.20	30,30,32,32	5

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	SO4	F	1144	5/5	0.93	0.16	19,24,25,25	5
2	SO4	B	1145	5/5	0.94	0.24	39,41,44,44	0
2	SO4	B	1146	5/5	0.94	0.19	15,16,19,21	5
2	SO4	C	1146	5/5	0.95	0.19	25,25,26,27	5
2	SO4	D	1146	5/5	0.96	0.28	42,42,44,44	0
2	SO4	A	1144	5/5	0.97	0.18	33,35,37,37	0
2	SO4	C	1145	5/5	0.97	0.09	19,20,22,22	5
2	SO4	D	1145	5/5	0.98	0.13	34,35,37,38	0
2	SO4	E	1145	5/5	0.99	0.14	32,33,35,36	0
2	SO4	F	1145	5/5	0.99	0.10	28,28,29,30	0
4	GD	F	1150	1/1	0.99	0.04	27,27,27,27	0
4	GD	B	1152	1/1	1.00	0.04	23,23,23,23	0
4	GD	C	1149	1/1	1.00	0.05	20,20,20,20	0
4	GD	D	1153	1/1	1.00	0.05	19,19,19,19	0
4	GD	E	1150	1/1	1.00	0.03	26,26,26,26	0
4	GD	A	1151	1/1	1.00	0.04	23,23,23,23	0

6.5 Other polymers (i)

There are no such residues in this entry.