Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 06:06 PM GMT

PDB ID	:	4KI0
Title	:	Crystal structure of the maltose-binding protein/maltose transporter complex in an outward-facing conformation bound to maltohexaose
Authors	:	Oldham, M.L.; Chen, S.; Chen, J.
Deposited on	:	2013-05-01
Resolution	:	2.38 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

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

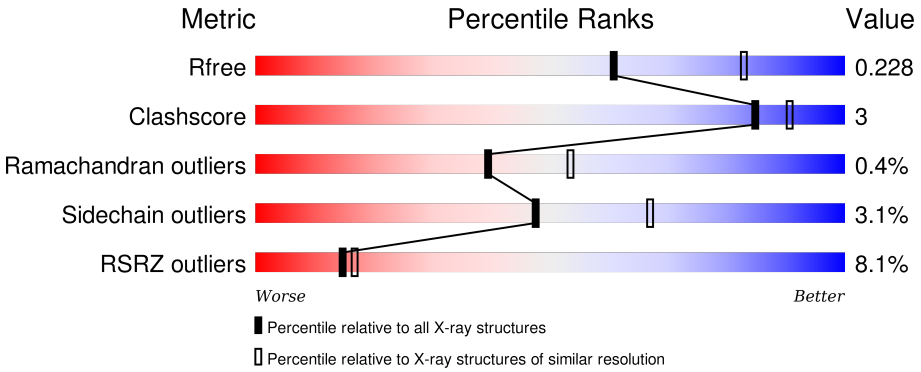
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance

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

The reported resolution of this entry is 2.38 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	4019 (2.40-2.36)
Clashscore	102246	4595 (2.40-2.36)
Ramachandran outliers	100387	4520 (2.40-2.36)
Sidechain outliers	100360	4522 (2.40-2.36)
RSRZ outliers	91569	4034 (2.40-2.36)

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. 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	381	6% 87% 10% •
1	B	381	11% 86% 11% •
2	E	380	7% 94% • •
3	F	514	9% 85% 9% • 5%
4	G	296	6% 88% 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
7	UMQ	F	607	-	-	-	X
7	UMQ	F	608	-	-	-	X
7	UMQ	F	609	-	-	-	X
7	UMQ	G	303	-	-	-	X
7	UMQ	G	306	-	-	-	X
7	UMQ	G	307	-	-	-	X
8	GLC	F	602	X	-	-	-

2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 15618 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 ABC transporter related protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	371	Total	C	N	O	S	0	1	0
			2884	1824	518	529	13			
1	B	371	Total	C	N	O	S	0	0	0
			2876	1819	515	529	13			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	372	ALA	-	EXPRESSION TAG	UNP C9QV42
A	373	SER	-	EXPRESSION TAG	UNP C9QV42
A	374	ALA	-	EXPRESSION TAG	UNP C9QV42
A	375	SER	-	EXPRESSION TAG	UNP C9QV42
A	376	HIS	-	EXPRESSION TAG	UNP C9QV42
A	377	HIS	-	EXPRESSION TAG	UNP C9QV42
A	378	HIS	-	EXPRESSION TAG	UNP C9QV42
A	379	HIS	-	EXPRESSION TAG	UNP C9QV42
A	380	HIS	-	EXPRESSION TAG	UNP C9QV42
A	381	HIS	-	EXPRESSION TAG	UNP C9QV42
B	372	ALA	-	EXPRESSION TAG	UNP C9QV42
B	373	SER	-	EXPRESSION TAG	UNP C9QV42
B	374	ALA	-	EXPRESSION TAG	UNP C9QV42
B	375	SER	-	EXPRESSION TAG	UNP C9QV42
B	376	HIS	-	EXPRESSION TAG	UNP C9QV42
B	377	HIS	-	EXPRESSION TAG	UNP C9QV42
B	378	HIS	-	EXPRESSION TAG	UNP C9QV42
B	379	HIS	-	EXPRESSION TAG	UNP C9QV42
B	380	HIS	-	EXPRESSION TAG	UNP C9QV42
B	381	HIS	-	EXPRESSION TAG	UNP C9QV42

- Molecule 2 is a protein called Maltose-binding periplasmic protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	370	Total	C	N	O	S	0	0	0
			2877	1853	469	549	6			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	371	ALA	-	EXPRESSION TAG	UNP P0AEX9
E	372	SER	-	EXPRESSION TAG	UNP P0AEX9
E	373	ALA	-	EXPRESSION TAG	UNP P0AEX9
E	374	SER	-	EXPRESSION TAG	UNP P0AEX9
E	375	HIS	-	EXPRESSION TAG	UNP P0AEX9
E	376	HIS	-	EXPRESSION TAG	UNP P0AEX9
E	377	HIS	-	EXPRESSION TAG	UNP P0AEX9
E	378	HIS	-	EXPRESSION TAG	UNP P0AEX9
E	379	HIS	-	EXPRESSION TAG	UNP P0AEX9
E	380	HIS	-	EXPRESSION TAG	UNP P0AEX9

- Molecule 3 is a protein called Maltose transport system permease protein MalF.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	F	490	Total	C	N	O	S	0	1	0
			3819	2512	607	683	17			

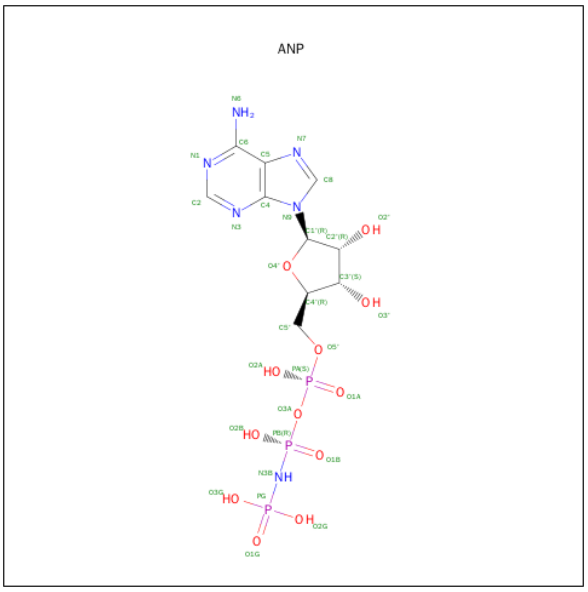
- Molecule 4 is a protein called Binding-protein-dependent transport systems inner membrane component.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	G	289	Total	C	N	O	S	0	1	0
			2238	1500	357	373	8			

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

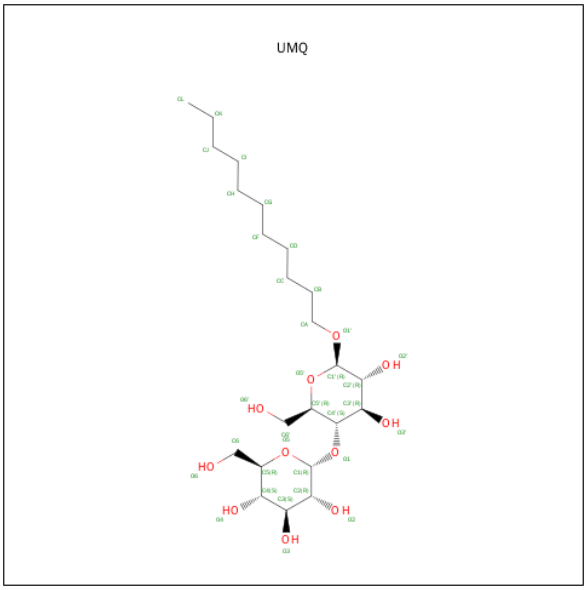
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total	Mg	0	0
			1	1		
5	A	1	Total	Mg	0	0
			1	1		

- Molecule 6 is PHOSPHOAMINOPHOSPHONIC ACID-ADENYLATE ESTER (three-letter code: ANP) (formula: $C_{10}H_{17}N_6O_{12}P_3$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	A	1	Total	C	N	O	P	0	0
			31	10	6	12	3		
6	B	1	Total	C	N	O	P	0	0
			31	10	6	12	3		

- Molecule 7 is UNDECYL-MALTOSIDE (three-letter code: UMQ) (formula: $C_{23}H_{44}O_{11}$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	E	1	Total	C	O	0	0
			34	23	11		
7	F	1	Total	C	O	0	0
			13	12	1		

Continued on next page...

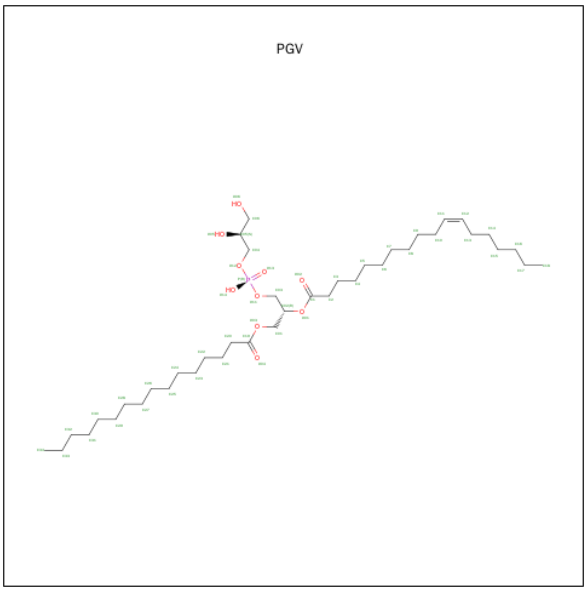
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	F	1	Total C O 13 12 1	0	0
7	F	1	Total C O 13 12 1	0	0
7	G	1	Total C O 34 23 11	0	0
7	G	1	Total C O 13 12 1	0	0
7	G	1	Total C O 13 12 1	0	0
7	G	1	Total C O 13 12 1	0	0
7	G	1	Total C O 23 17 6	0	0
7	G	1	Total C O 13 12 1	0	0
7	G	1	Total C O 34 23 11	0	0

- Molecule 8 is a polymer of unknown type called SUGAR (4-MER).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	F	4	Total C O 34 18 16	0	0

- Molecule 9 is (1R)-2-{[{[(2S)-2,3-DIHYDROXYPROPYL]OXY}(HYDROXY)PHOSPHORYL]OXY}-1-[(PALMITOYLOXY)METHYL]ETHYL (11E)-OCTADEC-11-ENOATE (three-letter code: PGV) (formula: $C_{40}H_{77}O_{10}P$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
9	F	1	Total	C	O	P	0	0
			51	40	10	1		
9	F	1	Total	C	O	P	0	0
			51	40	10	1		

- Molecule 10 is water.

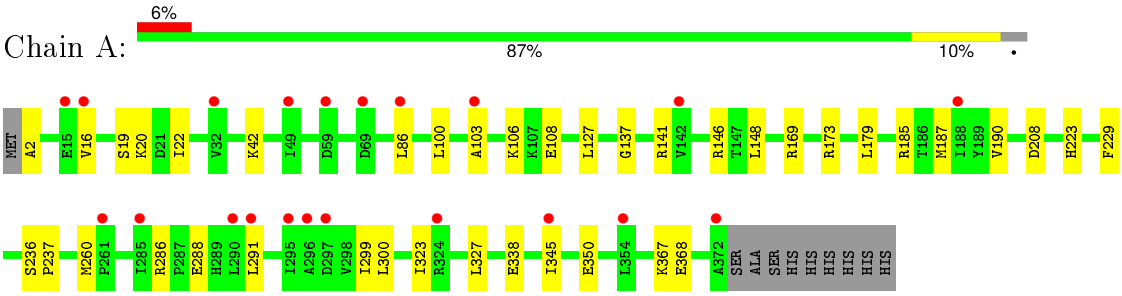
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	112	Total	O	0	0
			112	112		
10	B	123	Total	O	0	0
			123	123		
10	E	105	Total	O	0	0
			105	105		
10	F	87	Total	O	0	0
			87	87		
10	G	81	Total	O	0	0
			81	81		

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

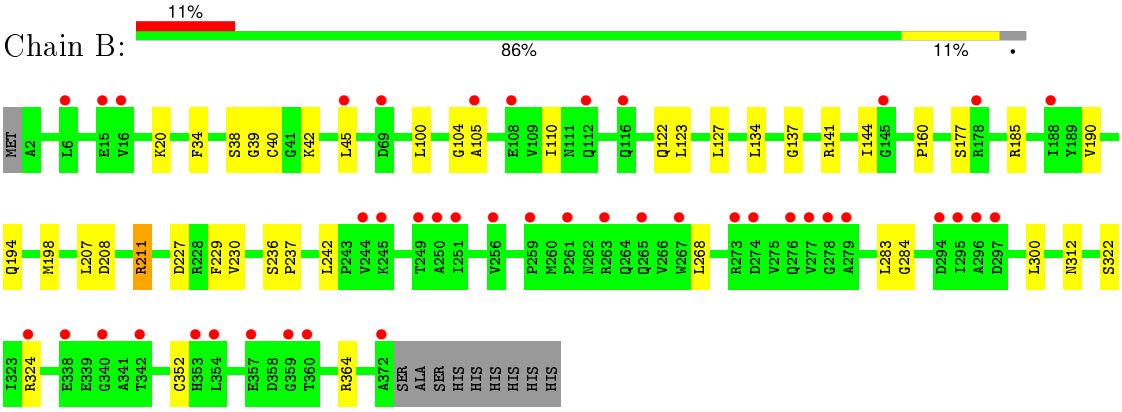
- Molecule 1: ABC transporter related protein

Chain A:



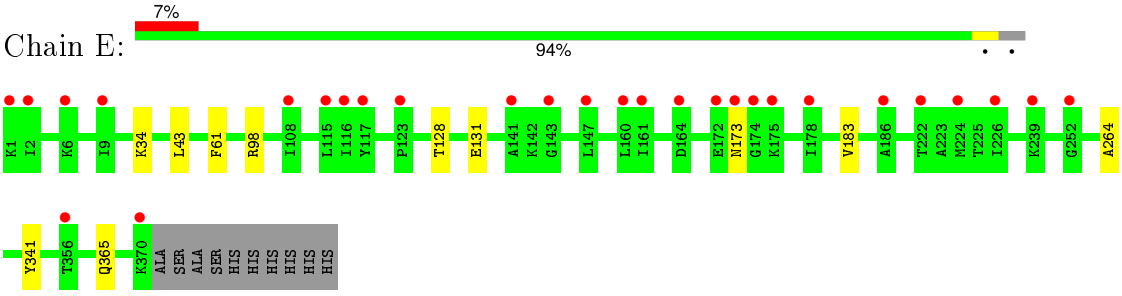
- Molecule 1: ABC transporter related protein

Chain B:



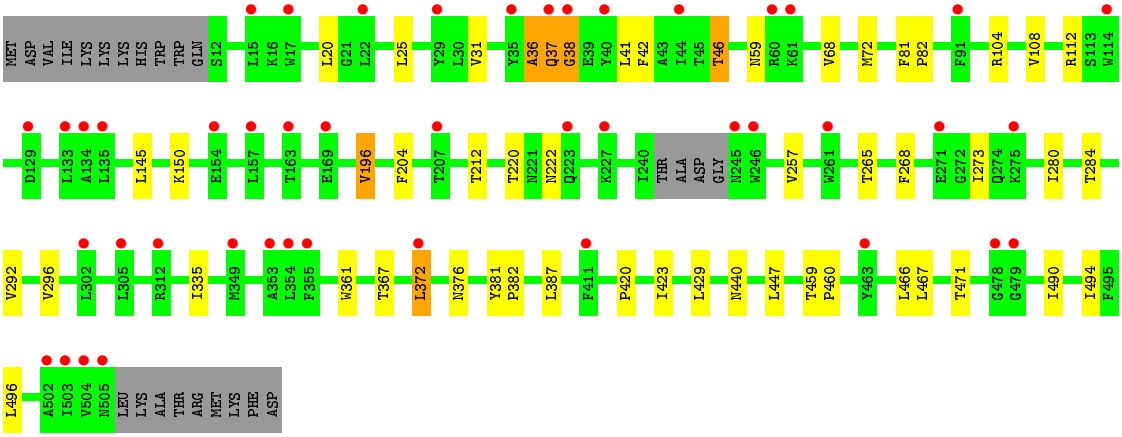
- Molecule 2: Maltose-binding periplasmic protein

Chain E:



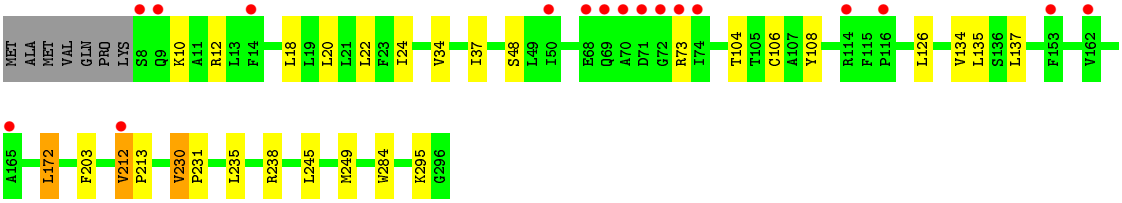
- Molecule 3: Maltose transport system permease protein MalF

Chain F:



• Molecule 4: Binding-protein-dependent transport systems inner membrane component

Chain G:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	81.59Å 96.63Å 112.32Å 85.30° 79.11° 73.00°	Depositor
Resolution (Å)	19.82 – 2.38 19.81 – 2.38	Depositor EDS
% Data completeness (in resolution range)	87.9 (19.82-2.38) 87.6 (19.81-2.38)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.31 (at 2.38Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.192 , 0.228 0.196 , 0.228	Depositor DCC
R_{free} test set	6156 reflections (5.70%)	DCC
Wilson B-factor (Å^2)	58.8	Xtriage
Anisotropy	0.812	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 64.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.35$	Xtriage
Outliers	0 of 113613 reflections	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	15618	wwPDB-VP
Average B, all atoms (Å^2)	86.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.73% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: PGV, ANP, MG, GLC, UMQ

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	0.31	0/2937	0.55	0/3982
1	B	0.31	0/2926	0.54	0/3968
2	E	0.32	0/2946	0.48	0/3998
3	F	0.32	0/3917	0.51	0/5333
4	G	0.33	0/2304	0.52	0/3150
All	All	0.32	0/15030	0.52	0/20431

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
8	F	1	0

There are no bond length outliers.

There are no bond angle outliers.

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
8	F	602	GLC	C1

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen

atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2884	0	2954	22	0
1	B	2876	0	2942	19	0
2	E	2877	0	2859	6	0
3	F	3819	0	3853	29	0
4	G	2238	0	2323	20	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
6	A	31	0	13	0	0
6	B	31	0	13	1	0
7	E	34	0	44	0	0
7	F	39	0	69	0	0
7	G	143	0	213	1	0
8	F	34	0	28	0	0
9	F	102	0	152	0	0
10	A	112	0	0	6	0
10	B	123	0	0	3	0
10	E	105	0	0	1	0
10	F	87	0	0	1	0
10	G	81	0	0	2	0
All	All	15618	0	15463	90	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 3.

All (90) 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)
3:F:429:LEU:HD23	4:G:172:LEU:HD12	1.53	0.88
1:A:173:ARG:CZ	10:A:1690:HOH:O	2.36	0.73
4:G:295:LYS:NZ	10:G:410:HOH:O	2.21	0.73
1:A:223:HIS:CE1	1:A:368:GLU:HG2	2.27	0.68
4:G:212:VAL:HG22	4:G:284[A]:TRP:CD2	2.30	0.67
3:F:284:THR:HG22	3:F:466:LEU:HA	1.77	0.66
3:F:280:ILE:O	3:F:284:THR:HG23	1.97	0.64
1:A:100:LEU:O	1:A:103:ALA:O	2.14	0.63
1:B:45:LEU:HD12	1:B:207:LEU:HD11	1.82	0.62
4:G:238:ARG:NH1	10:G:474:HOH:O	2.34	0.60
3:F:212:THR:HG23	3:F:222:ASN:HD21	1.68	0.58
1:B:364:ARG:NH1	10:B:563:HOH:O	2.35	0.57

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:F:387:LEU:HD13	3:F:429:LEU:HD13	1.85	0.57
1:A:260:MET:HE2	1:A:300:LEU:HD22	1.87	0.57
3:F:372:LEU:HD13	3:F:447:LEU:HD23	1.86	0.57
3:F:36:ALA:O	3:F:38:GLY:N	2.38	0.56
1:A:260:MET:CE	1:A:300:LEU:HD22	2.36	0.55
3:F:196:VAL:CG1	3:F:204:PHE:HB3	2.37	0.55
1:B:122:GLN:NE2	10:B:576:HOH:O	2.39	0.55
1:B:194:GLN:NE2	10:B:598:HOH:O	2.40	0.54
4:G:20:LEU:O	4:G:24:ILE:HG12	2.07	0.54
3:F:335:ILE:CD1	4:G:34:VAL:HG22	2.39	0.53
3:F:284:THR:HG21	3:F:467:LEU:H	1.73	0.53
2:E:43:LEU:HD12	2:E:43:LEU:C	2.30	0.52
1:B:236:SER:HA	1:B:237:PRO:C	2.30	0.52
1:A:288:GLU:HG2	1:B:312:ASN:HB2	1.93	0.51
1:A:16:VAL:HG11	10:A:1702:HOH:O	2.09	0.51
3:F:273:ILE:HG22	3:F:459:THR:HG21	1.93	0.51
4:G:212:VAL:HG22	4:G:284[A]:TRP:CE3	2.45	0.51
3:F:335:ILE:HD12	4:G:34:VAL:HG22	1.92	0.51
3:F:265:THR:O	3:F:268:PHE:N	2.44	0.50
1:A:42:LYS:HD2	1:A:190:VAL:HG13	1.93	0.50
1:A:327:LEU:CD2	1:A:345:ILE:HD11	2.41	0.50
4:G:134:VAL:HG23	4:G:135:LEU:CD1	2.42	0.50
3:F:196:VAL:HG13	3:F:204:PHE:HB3	1.93	0.50
2:E:341:TYR:CE2	3:F:460:PRO:HB2	2.47	0.50
3:F:471:THR:CG2	3:F:490:ILE:HG21	2.42	0.50
1:A:19:SER:HB3	1:A:22:ILE:HD12	1.94	0.49
1:A:291:LEU:HD12	10:A:1679:HOH:O	2.12	0.49
1:A:148:LEU:CD2	1:A:179:LEU:HD22	2.42	0.49
2:E:61:PHE:CE2	2:E:264:ALA:HB2	2.49	0.48
4:G:134:VAL:HG23	4:G:135:LEU:HD13	1.95	0.47
1:A:286:ARG:HB3	1:A:288:GLU:OE1	2.14	0.47
1:A:187:MET:HB3	10:A:1678:HOH:O	2.13	0.47
3:F:471:THR:HG21	3:F:490:ILE:HG21	1.97	0.47
1:A:2:ALA:N	10:A:1672:HOH:O	2.48	0.47
1:B:268:LEU:HD22	1:B:352:CYS:HB2	1.96	0.47
1:A:208:ASP:HB2	1:A:229:PHE:CE2	2.49	0.47
4:G:104:THR:HG22	4:G:203:PHE:HZ	1.80	0.47
3:F:68:VAL:HG12	3:F:72:MET:HG2	1.97	0.47
4:G:212:VAL:HG12	4:G:213:PRO:HD3	1.97	0.46
1:B:194:GLN:O	1:B:198:MET:HG2	2.16	0.46
3:F:372:LEU:CD1	3:F:447:LEU:HD23	2.45	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:F:361:TRP:NE1	10:F:758:HOH:O	2.36	0.46
4:G:230:VAL:HB	4:G:231:PRO:HD3	1.97	0.46
1:A:146:ARG:HD2	10:A:1610:HOH:O	2.14	0.46
1:A:260:MET:HE3	1:A:323:ILE:HD11	1.98	0.46
1:A:137:GLY:O	1:A:141:ARG:HD3	2.17	0.45
4:G:245:LEU:HG	4:G:249:MET:CE	2.46	0.45
3:F:81:PHE:HB3	3:F:82:PRO:HD3	1.98	0.45
3:F:31:VAL:HG22	3:F:46:THR:CG2	2.46	0.45
1:B:208:ASP:HB2	1:B:229:PHE:CE2	2.52	0.45
1:B:34:PHE:HB2	1:B:190:VAL:HG22	1.98	0.44
2:E:98:ARG:NH2	10:E:597:HOH:O	2.51	0.44
2:E:183:VAL:HG22	2:E:365:GLN:OE1	2.18	0.44
1:B:283:LEU:HD23	1:B:284:GLY:N	2.33	0.44
3:F:361:TRP:CE3	3:F:367:THR:HG22	2.53	0.44
3:F:335:ILE:HD11	4:G:37:ILE:HD12	2.00	0.43
1:B:137:GLY:O	1:B:141:ARG:HD3	2.19	0.43
3:F:381:TYR:N	3:F:382:PRO:CD	2.82	0.43
4:G:18:LEU:HD23	4:G:22:LEU:HD23	2.01	0.42
3:F:42:PHE:O	3:F:46:THR:HB	2.19	0.42
3:F:104:ARG:O	3:F:108:VAL:HG23	2.20	0.42
1:A:86:LEU:HA	1:A:146:ARG:NH2	2.35	0.42
1:A:236:SER:HA	1:A:237:PRO:C	2.40	0.42
2:E:128:THR:OG1	2:E:131:GLU:HB2	2.19	0.41
1:B:40:CYS:SG	1:B:42:LYS:HG3	2.60	0.41
4:G:108:TYR:CZ	4:G:203:PHE:HB2	2.55	0.41
4:G:212:VAL:HG13	4:G:284[B]:TRP:HB3	2.02	0.41
1:B:144:ILE:HD11	1:B:160:PRO:O	2.21	0.41
3:F:420:PRO:HA	3:F:423:ILE:HG12	2.02	0.41
1:B:208:ASP:OD2	1:B:211:ARG:NH1	2.52	0.41
1:B:105:ALA:HA	1:B:110:ILE:HD11	2.01	0.41
4:G:104:THR:HG22	4:G:203:PHE:CZ	2.56	0.41
3:F:292:VAL:O	3:F:296:VAL:HG23	2.21	0.40
4:G:212:VAL:HG21	7:G:307:UMQ:O2'	2.21	0.40
1:B:227:ASP:OD1	1:B:230:VAL:HG23	2.20	0.40
1:B:300:LEU:N	1:B:300:LEU:HD12	2.37	0.40
1:A:299:ILE:O	1:A:300:LEU:HD23	2.21	0.40
1:B:39:GLY:N	6:B:402:ANP:HNB1	2.18	0.40

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	370/381 (97%)	363 (98%)	7 (2%)	0	100	100
1	B	369/381 (97%)	349 (95%)	18 (5%)	2 (0%)	34	46
2	E	368/380 (97%)	360 (98%)	7 (2%)	1 (0%)	46	61
3	F	487/514 (95%)	469 (96%)	15 (3%)	3 (1%)	30	40
4	G	288/296 (97%)	283 (98%)	4 (1%)	1 (0%)	46	61
All	All	1882/1952 (96%)	1824 (97%)	51 (3%)	7 (0%)	39	53

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	F	37	GLN
2	E	173	ASN
3	F	38	GLY
1	B	38	SER
3	F	36	ALA
4	G	230	VAL
1	B	104	GLY

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	315/323 (98%)	305 (97%)	10 (3%)	46	66
1	B	314/323 (97%)	303 (96%)	11 (4%)	43	62

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
2	E	297/305 (97%)	296 (100%)	1 (0%)	94 98
3	F	402/424 (95%)	385 (96%)	17 (4%)	36 53
4	G	232/237 (98%)	222 (96%)	10 (4%)	35 52
All	All	1560/1612 (97%)	1511 (97%)	49 (3%)	47 67

All (49) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	20	LYS
1	A	106	LYS
1	A	108	GLU
1	A	127	LEU
1	A	169[A]	ARG
1	A	169[B]	ARG
1	A	185	ARG
1	A	338	GLU
1	A	350	GLU
1	A	367	LYS
1	B	20	LYS
1	B	100	LEU
1	B	123	LEU
1	B	127	LEU
1	B	134	LEU
1	B	177	SER
1	B	185	ARG
1	B	211	ARG
1	B	242	LEU
1	B	322	SER
1	B	324	ARG
2	E	34	LYS
3	F	20	LEU
3	F	25	LEU
3	F	37	GLN
3	F	41	LEU
3	F	46	THR
3	F	59	ASN
3	F	112	ARG
3	F	145	LEU
3	F	150	LYS
3	F	196	VAL
3	F	220	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	F	257	VAL
3	F	372	LEU
3	F	376	ASN
3	F	440	ASN
3	F	494	ILE
3	F	496	LEU
4	G	10	LYS
4	G	12	ARG
4	G	48	SER
4	G	73	ARG
4	G	106	CYS
4	G	126	LEU
4	G	137	LEU
4	G	172	LEU
4	G	212	VAL
4	G	235	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (11) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	7	GLN
1	B	8	ASN
1	B	125	HIS
2	E	218	ASN
3	F	37	GLN
3	F	106	GLN
3	F	232	ASN
3	F	376	ASN
3	F	437	ASN
3	F	440	ASN
4	G	241	ASN

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)

Of 4 carbohydrates modelled in this entry, 1 is modelled with single atom - 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	$\#\|Z\| > 2$
8	GLC	F	602	8	11,11,12	0.64	0	14,15,17	1.67	5 (35%)
8	GLC	F	603	8	11,11,12	0.64	0	14,15,17	1.09	1 (7%)
8	GLC	F	604	8	11,11,12	0.61	0	14,15,17	0.61	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
8	GLC	F	602	8	1/1/4/5	0/2/19/22	0/1/1/1
8	GLC	F	603	8	-	0/2/19/22	0/1/1/1
8	GLC	F	604	8	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
8	F	602	GLC	O4-C4-C3	-2.77	104.11	110.34
8	F	602	GLC	C2-C3-C4	2.06	114.54	111.04
8	F	602	GLC	C1-C2-C3	2.25	112.21	109.54
8	F	602	GLC	C3-C4-C5	2.49	114.54	110.20
8	F	603	GLC	C1-O5-C5	2.98	116.03	112.25
8	F	602	GLC	O5-C1-C2	3.40	116.36	110.86

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
8	F	602	GLC	C1

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.6 Ligand geometry (i)

Of 17 ligands modelled in this entry, 2 are monoatomic - leaving 15 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	$\#\|Z\| > 2$
6	ANP	A	1502	5	27,33,33	1.99	6 (22%)	30,52,52	2.15	7 (23%)
6	ANP	B	402	5	27,33,33	2.00	7 (25%)	30,52,52	2.27	6 (20%)
7	UMQ	E	401	-	35,35,35	0.41	0	46,46,46	0.71	0
9	PGV	F	605	-	50,50,50	1.07	3 (6%)	51,56,56	1.05	4 (7%)
9	PGV	F	606	-	50,50,50	1.07	3 (6%)	51,56,56	0.97	3 (5%)
7	UMQ	F	607	-	12,12,35	0.36	0	11,11,46	0.40	0
7	UMQ	F	608	-	12,12,35	0.38	0	11,11,46	0.37	0
7	UMQ	F	609	-	12,12,35	0.35	0	11,11,46	0.42	0
7	UMQ	G	301	-	35,35,35	0.52	1 (2%)	46,46,46	1.05	2 (4%)
7	UMQ	G	302	-	12,12,35	0.38	0	11,11,46	0.43	0
7	UMQ	G	303	-	12,12,35	0.35	0	11,11,46	0.42	0
7	UMQ	G	304	-	12,12,35	0.38	0	11,11,46	0.38	0
7	UMQ	G	305	-	23,23,35	0.48	0	28,28,46	0.68	0
7	UMQ	G	306	-	12,12,35	0.36	0	11,11,46	0.42	0
7	UMQ	G	307	-	35,35,35	0.52	1 (2%)	46,46,46	1.05	4 (8%)

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
6	ANP	A	1502	5	-	1/12/38/38	0/3/3/3
6	ANP	B	402	5	-	1/12/38/38	0/3/3/3

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
7	UMQ	E	401	-	-	0/20/60/60	0/2/2/2
9	PGV	F	605	-	-	0/55/55/55	0/0/0/0
9	PGV	F	606	-	-	0/55/55/55	0/0/0/0
7	UMQ	F	607	-	-	0/10/10/60	0/0/0/2
7	UMQ	F	608	-	-	0/10/10/60	0/0/0/2
7	UMQ	F	609	-	-	0/10/10/60	0/0/0/2
7	UMQ	G	301	-	-	0/20/60/60	0/2/2/2
7	UMQ	G	302	-	-	0/10/10/60	0/0/0/2
7	UMQ	G	303	-	-	0/10/10/60	0/0/0/2
7	UMQ	G	304	-	-	0/10/10/60	0/0/0/2
7	UMQ	G	305	-	-	0/14/34/60	0/1/1/2
7	UMQ	G	306	-	-	0/10/10/60	0/0/0/2
7	UMQ	G	307	-	-	1/20/60/60	0/2/2/2

All (21) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	B	402	ANP	PG-O3G	-2.08	1.50	1.56
6	B	402	ANP	PG-O2G	-2.06	1.50	1.56
6	A	1502	ANP	PG-O3G	-2.01	1.51	1.56
7	G	307	UMQ	O1'-C1'	2.24	1.44	1.40
7	G	301	UMQ	O1'-C1'	2.36	1.44	1.40
6	A	1502	ANP	C5-C4	3.09	1.47	1.40
6	B	402	ANP	C5-C4	3.20	1.47	1.40
6	A	1502	ANP	PB-N3B	3.53	1.72	1.63
6	B	402	ANP	PG-N3B	3.57	1.72	1.63
9	F	606	PGV	C12-C11	3.63	1.52	1.31
6	A	1502	ANP	PG-N3B	3.65	1.73	1.63
9	F	605	PGV	C12-C11	3.80	1.53	1.31
6	B	402	ANP	PB-N3B	3.84	1.73	1.63
9	F	605	PGV	O03-C19	4.00	1.45	1.33
6	A	1502	ANP	PG-O1G	4.13	1.50	1.46
9	F	605	PGV	O01-C1	4.18	1.46	1.34
9	F	606	PGV	O01-C1	4.20	1.46	1.34
9	F	606	PGV	O03-C19	4.26	1.46	1.33
6	B	402	ANP	PB-O1B	4.30	1.51	1.46
6	B	402	ANP	PG-O1G	4.88	1.51	1.46
6	A	1502	ANP	PB-O1B	5.33	1.52	1.46

All (26) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	B	402	ANP	N3-C2-N1	-8.25	122.58	128.89
6	A	1502	ANP	N3-C2-N1	-7.86	122.88	128.89
6	B	402	ANP	O1G-PG-N3B	-5.18	103.95	111.90
6	A	1502	ANP	O1G-PG-N3B	-3.97	105.81	111.90
6	B	402	ANP	O1B-PB-N3B	-2.78	107.64	111.90
6	A	1502	ANP	C4-C5-N7	-2.57	107.12	109.48
6	A	1502	ANP	PA-O3A-PB	-2.53	124.17	132.67
9	F	605	PGV	O03-C19-O04	-2.42	117.25	123.49
6	B	402	ANP	C4-C5-N7	-2.36	107.31	109.48
9	F	605	PGV	C03-C02-C01	-2.02	107.33	112.07
9	F	606	PGV	C13-C12-C11	-2.01	111.36	125.34
6	A	1502	ANP	C1'-N9-C4	-2.01	123.92	126.94
6	A	1502	ANP	O3G-PG-O2G	2.06	113.68	107.58
6	B	402	ANP	O3G-PG-O2G	2.15	113.96	107.58
7	G	307	UMQ	O5-C5-C4	2.23	113.88	109.68
7	G	307	UMQ	O1'-C1'-C2'	2.33	110.98	108.04
7	G	307	UMQ	C1'-C2'-C3'	2.33	114.56	109.97
7	G	301	UMQ	CA-O1'-C1'	2.46	118.24	113.94
9	F	606	PGV	O03-C19-C20	2.88	120.68	111.90
7	G	307	UMQ	C3-C4-C5	3.02	115.46	110.20
9	F	605	PGV	O03-C19-C20	3.37	122.16	111.90
9	F	606	PGV	O01-C1-C2	3.46	119.05	111.53
9	F	605	PGV	O01-C1-C2	3.81	119.80	111.53
6	B	402	ANP	O2B-PB-O1B	4.25	118.88	110.00
7	G	301	UMQ	O1'-C1'-C2'	4.67	113.93	108.04
6	A	1502	ANP	O2B-PB-O1B	4.84	120.11	110.00

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	B	402	ANP	O1B-PB-N3B-PG
6	A	1502	ANP	O1B-PB-N3B-PG
7	G	307	UMQ	CA-O1'-C1'-O5'

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	B	402	ANP	1	0
7	G	307	UMQ	1	0

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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	371/381 (97%)	0.30	21 (5%) 27 31	54, 79, 108, 126	0
1	B	371/381 (97%)	0.57	43 (11%) 6 7	54, 92, 147, 170	0
2	E	370/380 (97%)	0.26	28 (7%) 17 19	59, 83, 111, 132	0
3	F	490/514 (95%)	0.50	45 (9%) 11 13	53, 85, 134, 153	0
4	G	289/296 (97%)	0.30	17 (5%) 26 30	51, 71, 115, 148	0
All	All	1891/1952 (96%)	0.40	154 (8%) 15 17	51, 82, 130, 170	0

All (154) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	372	ALA	9.6
1	B	251	ILE	8.2
1	B	295	ILE	7.1
3	F	37	GLN	6.9
4	G	70	ALA	6.7
4	G	71	ASP	6.4
3	F	40	TYR	5.0
1	B	276	GLN	5.0
2	E	1	LYS	5.0
3	F	503	ILE	4.8
1	B	105	ALA	4.6
2	E	174	GLY	4.5
1	A	372	ALA	4.5
1	B	261	PRO	4.5
4	G	69	GLN	4.4
3	F	245	ASN	4.4
3	F	312	ARG	4.3
1	B	16	VAL	4.3
1	A	295	ILE	4.2
3	F	353	ALA	4.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	45	LEU	4.1
2	E	173	ASN	4.1
3	F	354	LEU	4.1
2	E	115	LEU	4.0
1	A	324	ARG	4.0
4	G	73	ARG	3.9
1	B	277	VAL	3.9
1	B	263	ARG	3.9
3	F	349	MET	3.8
4	G	114	ARG	3.8
3	F	134	ALA	3.7
1	B	359	GLY	3.6
3	F	29	TYR	3.6
3	F	133	LEU	3.6
3	F	504	VAL	3.6
3	F	355	PHE	3.6
1	B	340	GLY	3.6
1	B	265	GLN	3.5
1	A	69	ASP	3.5
2	E	370	LYS	3.5
1	B	357	GLU	3.5
1	B	249	THR	3.4
1	B	353	HIS	3.4
3	F	246	TRP	3.4
2	E	224	MET	3.4
1	A	297	ASP	3.4
3	F	505	ASN	3.3
1	B	278	GLY	3.3
1	A	188	ILE	3.3
4	G	74	ILE	3.2
2	E	172	GLU	3.2
1	B	354	LEU	3.1
1	B	274	ASP	3.1
1	B	245	LYS	3.1
2	E	252	GLY	3.1
4	G	68	GLU	3.1
2	E	226	ILE	3.1
3	F	478	GLY	3.1
3	F	114	TRP	3.1
1	B	267	TRP	3.0
1	B	296	ALA	3.0
1	A	285	ILE	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	294	ASP	3.0
1	A	49	ILE	3.0
2	E	161	ILE	3.0
1	B	279	ALA	2.9
3	F	61	LYS	2.9
1	B	244	VAL	2.9
3	F	271	GLU	2.8
3	F	154	GLU	2.8
4	G	72	GLY	2.8
2	E	147	LEU	2.8
3	F	17	TRP	2.8
2	E	239	LYS	2.8
1	B	116	GLN	2.7
1	A	345	ILE	2.7
1	B	112	GLN	2.7
4	G	116	PRO	2.7
3	F	22	LEU	2.7
4	G	153	PHE	2.7
3	F	275	LYS	2.7
1	B	15	GLU	2.7
4	G	162	VAL	2.7
1	A	16	VAL	2.7
1	B	324	ARG	2.6
1	B	273	ARG	2.6
2	E	178	ILE	2.6
1	A	261	PRO	2.6
2	E	186	ALA	2.6
3	F	372	LEU	2.6
1	A	15	GLU	2.6
1	B	360	THR	2.6
2	E	2	ILE	2.6
4	G	14	PHE	2.6
1	B	69	ASP	2.6
3	F	305	LEU	2.6
3	F	38	GLY	2.5
3	F	35	TYR	2.5
3	F	463	TYR	2.5
1	B	108	GLU	2.5
2	E	175	LYS	2.5
3	F	479	GLY	2.5
1	B	259	PRO	2.5
1	A	142	VAL	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	F	227	LYS	2.4
2	E	356	THR	2.4
3	F	302	LEU	2.4
3	F	411	PHE	2.4
1	B	297	ASP	2.4
3	F	502	ALA	2.4
3	F	135	LEU	2.4
1	B	256	VAL	2.4
3	F	129	ASP	2.4
1	B	342	THR	2.4
1	B	250	ALA	2.4
2	E	160	LEU	2.4
3	F	223	GLN	2.4
1	A	290	LEU	2.3
2	E	116	ILE	2.3
3	F	163	THR	2.3
4	G	212	VAL	2.3
1	A	103	ALA	2.3
3	F	60	ARG	2.3
2	E	222	THR	2.2
3	F	44	ILE	2.2
3	F	91	PHE	2.2
2	E	141	ALA	2.2
2	E	117	TYR	2.2
1	B	338	GLU	2.2
1	A	291	LEU	2.2
1	B	6	LEU	2.2
3	F	15	LEU	2.2
1	A	296	ALA	2.2
4	G	50	ILE	2.2
4	G	8	SER	2.2
2	E	9	ILE	2.1
3	F	169	GLU	2.1
1	B	145	GLY	2.1
1	A	354	LEU	2.1
4	G	9	GLN	2.1
1	B	188	ILE	2.1
1	A	59	ASP	2.1
3	F	157	LEU	2.1
2	E	143	GLY	2.1
2	E	164	ASP	2.1
1	B	178	ARG	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	F	261	TRP	2.1
3	F	207	THR	2.0
4	G	165	ALA	2.0
2	E	6	LYS	2.0
2	E	108	ILE	2.0
1	A	32	VAL	2.0
1	A	86	LEU	2.0
2	E	123	PRO	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)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95^{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	LLDF	B-factors(Å^2)	Q<0.9
8	GLC	F	602	11/12	0.95	0.11	-0.88	73,80,87,95	0
8	GLC	F	604	11/12	0.99	0.07	-1.63	61,62,66,67	0
8	GLC	F	603	11/12	0.98	0.09	-1.65	61,64,68,68	0
8	GLC	F	601	1/12	0.82	0.26	-	91,91,91,91	0

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95^{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	LLDF	B-factors(Å^2)	Q<0.9
7	UMQ	F	609	13/34	0.81	0.34	6.48	95,101,111,111	0
7	UMQ	F	607	13/34	0.81	0.26	5.49	87,90,92,94	0
7	UMQ	G	306	13/34	0.67	0.29	5.00	90,99,118,119	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
7	UMQ	F	608	13/34	0.70	0.29	4.89	97,100,115,115	0
7	UMQ	G	307	34/34	0.74	0.35	3.45	99,135,156,157	0
7	UMQ	G	303	13/34	0.86	0.19	2.52	78,86,99,101	0
7	UMQ	G	301	34/34	0.68	0.33	1.70	91,132,157,158	0
9	PGV	F	605	51/51	0.88	0.23	1.51	89,98,109,111	0
9	PGV	F	606	51/51	0.78	0.25	1.38	86,106,129,134	0
7	UMQ	G	305	23/34	0.78	0.22	1.20	94,106,127,136	0
7	UMQ	E	401	34/34	0.93	0.19	0.98	80,87,96,97	0
6	ANP	A	1502	31/31	0.98	0.11	-0.20	55,62,75,81	0
6	ANP	B	402	31/31	0.96	0.11	-0.46	58,70,84,89	0
5	MG	B	401	1/1	0.96	0.10	-0.79	59,59,59,59	0
5	MG	A	1501	1/1	0.99	0.04	-2.36	57,57,57,57	0
7	UMQ	G	304	13/34	0.81	0.28	-	88,100,110,111	0
7	UMQ	G	302	13/34	0.80	0.21	-	86,93,96,99	0

6.5 Other polymers (i)

There are no such residues in this entry.