



Full wwPDB X-ray Structure Validation Report ⓘ

Aug 1, 2016 – 12:37 PM EDT

PDB ID : 5CVY
Title : The Structure of Bacillus pumilus GH48 in complex with cellobiose and cel-
lohexaose
Authors : Alahuhta, P.M.; Lunin, V.V.
Deposited on : 2015-07-27
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

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20027939
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20027939

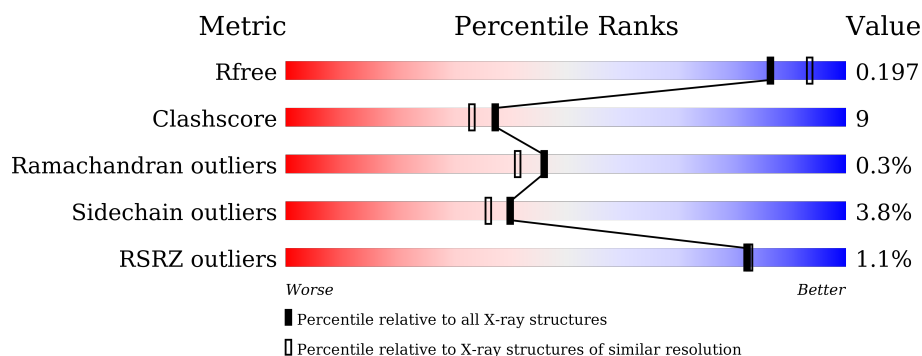
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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}	91344	6249 (2.00-2.00)
Clashscore	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (2.00-2.00)
RSRZ outliers	91569	6262 (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. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	709	<div> <div></div> <div>88%</div> <div>10% ..</div> </div>

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
5	MLI	A	804	-	-	X	X
5	MLI	A	805	-	-	X	X

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	MLI	A	806	-	-	-	X
6	EDO	A	808	-	-	-	X
6	EDO	A	809	-	-	-	X
6	EDO	A	810	-	-	X	X
6	EDO	A	812[A]	-	-	-	X
6	EDO	A	812[B]	-	-	-	X
6	EDO	A	815[A]	-	-	-	X
6	EDO	A	815[B]	-	-	-	X
6	EDO	A	817	-	-	-	X
6	EDO	A	818	-	-	X	X
6	EDO	A	820	-	-	X	X
7	GOL	A	821	-	-	X	X
7	GOL	A	822	-	-	X	X
7	GOL	A	823	-	-	X	X
7	GOL	A	824	-	-	-	X
8	NA	A	830	-	-	-	X
8	NA	A	834	-	-	-	X
8	NA	A	836	-	-	-	X
8	NA	A	839	-	-	-	X

2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 6866 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 Glycoside hydrolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	701	Total	C	N	O	S	0	28	0
			5816	3689	974	1135	18			

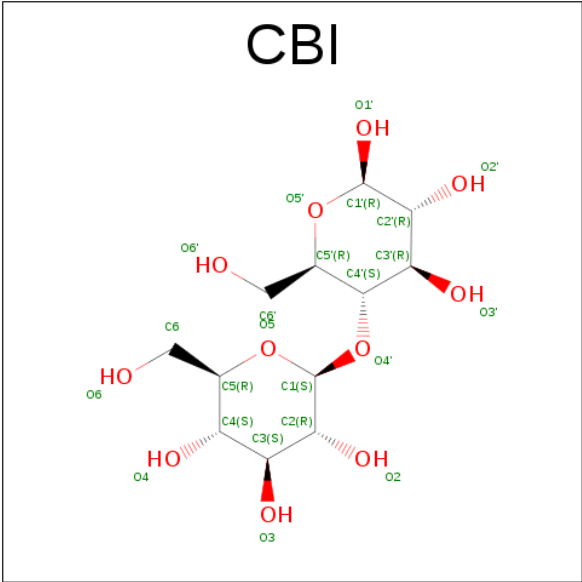
There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	702	LEU	-	expression tag	UNP A8FDC4
A	703	GLU	-	expression tag	UNP A8FDC4
A	704	HIS	-	expression tag	UNP A8FDC4
A	705	HIS	-	expression tag	UNP A8FDC4
A	706	HIS	-	expression tag	UNP A8FDC4
A	707	HIS	-	expression tag	UNP A8FDC4
A	708	HIS	-	expression tag	UNP A8FDC4
A	709	HIS	-	expression tag	UNP A8FDC4

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

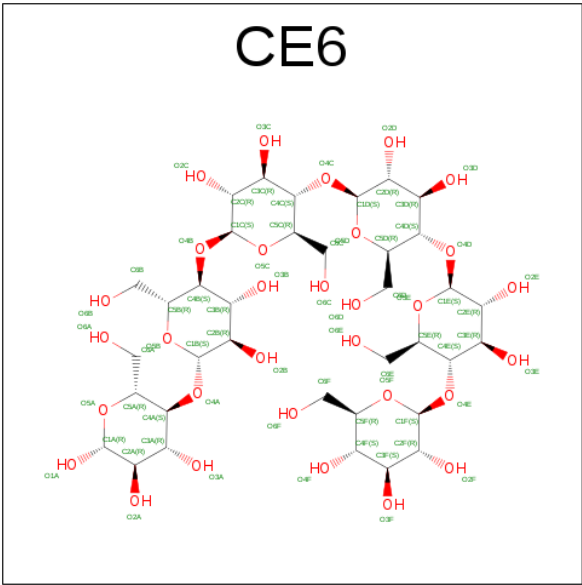
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Ca	0	0
			1	1		

- Molecule 3 is CELLOBIOSE (three-letter code: CBI) (formula: C₁₂H₂₂O₁₁).



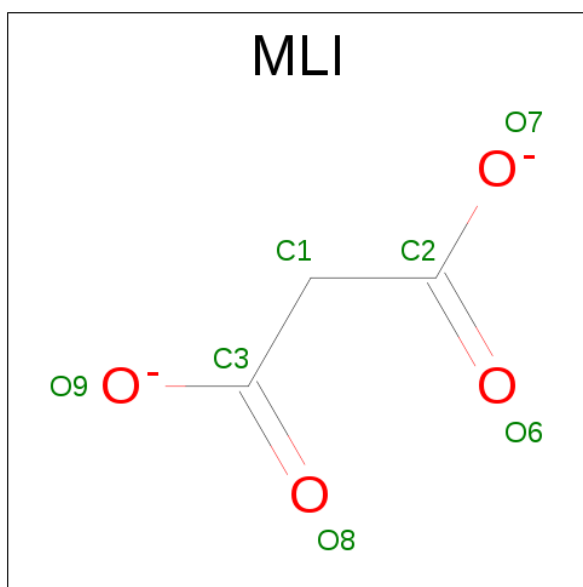
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			23	12	11		

- Molecule 4 is CELLOHEXAOSE (three-letter code: CE6) (formula: $C_{36}H_{62}O_{31}$).



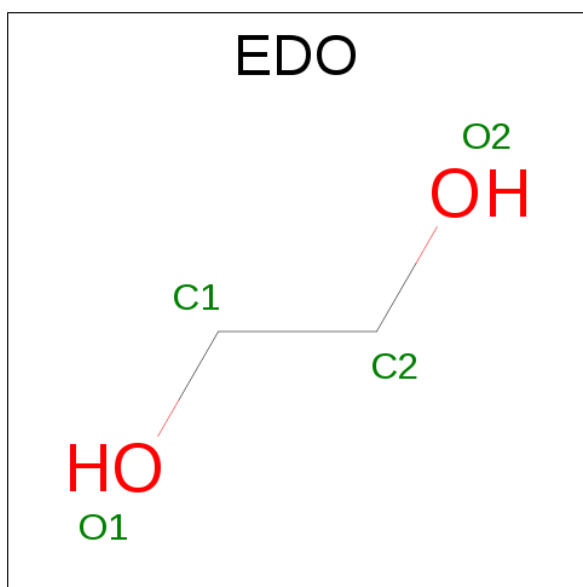
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			67	36	31		

- Molecule 5 is MALONATE ION (three-letter code: MLI) (formula: $C_3H_2O_4$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			7	3	4		
5	A	1	Total	C	O	0	0
			7	3	4		
5	A	1	Total	C	O	0	0
			7	3	4		

- Molecule 6 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



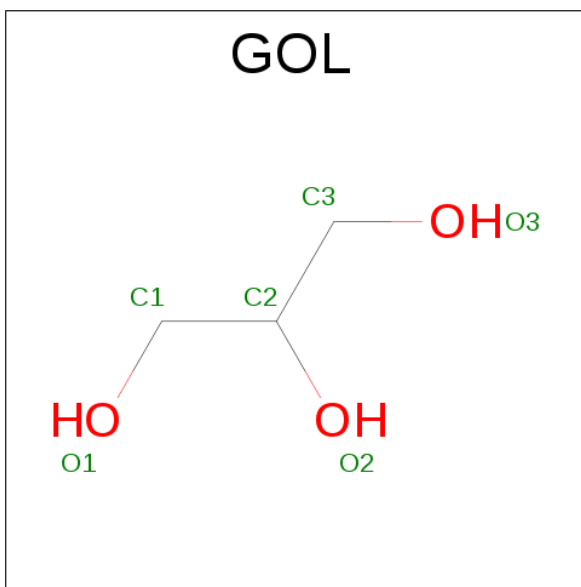
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			4	2	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total 4	C 2	O 2	0	0
6	A	1	Total 4	C 2	O 2	0	0
6	A	1	Total 4	C 2	O 2	0	0
6	A	1	Total 4	C 2	O 2	0	0
6	A	1	Total 8	C 4	O 4	0	1
6	A	1	Total 4	C 2	O 2	0	0
6	A	1	Total 4	C 2	O 2	0	0
6	A	1	Total 8	C 4	O 4	0	1
6	A	1	Total 4	C 2	O 2	0	0
6	A	1	Total 4	C 2	O 2	0	0
6	A	1	Total 4	C 2	O 2	0	0
6	A	1	Total 8	C 4	O 4	0	1
6	A	1	Total 4	C 2	O 2	0	0

- Molecule 7 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	C	O	0	0
			6	3	3		
7	A	1	Total	C	O	0	0
			6	3	3		
7	A	1	Total	C	O	0	0
			6	3	3		
7	A	1	Total	C	O	0	0
			6	3	3		

- Molecule 8 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	15	Total	Na	0	0
			15	15		

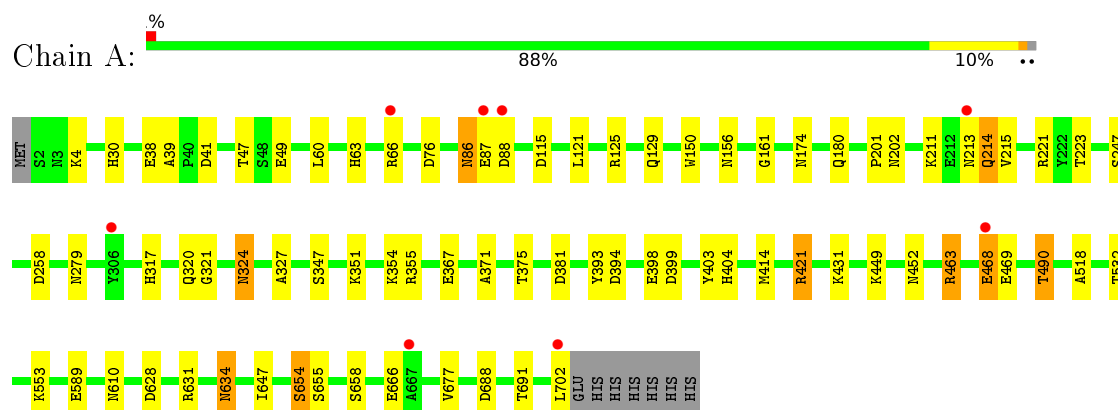
- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	794	Total	O	0	37
			831	831		

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 ($\text{RSRZ} > 2$). 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: Glycoside hydrolase



4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α , β , γ	98.04Å 98.04Å 218.96Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.00 50.26 – 2.00	Depositor EDS
% Data completeness (in resolution range)	97.3 (50.00-2.00) 97.3 (50.26-2.00)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.22 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, R_{free}	0.146 , 0.189 0.157 , 0.197	Depositor DCC
R_{free} test set	3584 reflections (5.32%)	DCC
Wilson B-factor (Å ²)	14.3	Xtriage
Anisotropy	0.522	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 60.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	6866	wwPDB-VP
Average B, all atoms (Å ²)	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.58% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, MLI, NA, CA, CBI, EDO, CE6

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	0.98	1/6001 (0.0%)	0.96	15/8172 (0.2%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	654	SER	CB-OG	-6.75	1.33	1.42

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	125	ARG	NE-CZ-NH2	-11.40	114.60	120.30
1	A	628	ASP	CB-CG-OD1	9.00	126.40	118.30
1	A	125	ARG	NE-CZ-NH1	8.94	124.77	120.30
1	A	355	ARG	NE-CZ-NH1	8.66	124.63	120.30
1	A	631	ARG	NE-CZ-NH1	8.04	124.32	120.30
1	A	628	ASP	CB-CG-OD2	-7.22	111.80	118.30
1	A	355	ARG	NE-CZ-NH2	-7.00	116.80	120.30
1	A	421[A]	ARG	NE-CZ-NH2	-6.75	116.93	120.30
1	A	421[B]	ARG	NE-CZ-NH2	-6.75	116.93	120.30
1	A	399	ASP	CB-CG-OD1	6.43	124.09	118.30
1	A	76	ASP	CB-CG-OD1	5.92	123.63	118.30
1	A	463	ARG	NE-CZ-NH1	5.61	123.10	120.30
1	A	381	ASP	CB-CG-OD1	5.59	123.33	118.30
1	A	41	ASP	CB-CG-OD1	5.51	123.26	118.30
1	A	631	ARG	NE-CZ-NH2	-5.24	117.68	120.30

There are no chirality outliers.

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	5816	0	5345	73	0
2	A	1	0	0	0	0
3	A	23	0	22	0	0
4	A	67	0	62	2	0
5	A	21	0	5	14	0
6	A	68	0	100	25	0
7	A	24	0	24	24	0
8	A	15	0	0	0	0
9	A	831	0	0	57	1
All	All	6866	0	5558	108	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 9.

All (108) 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:449[C]:LYS:HE3	9:A:945:HOH:O	1.14	1.32
6:A:809:EDO:H22	9:A:1246:HOH:O	1.26	1.26
1:A:174[A]:ASN:ND2	9:A:919:HOH:O	1.63	1.26
6:A:818:EDO:H11	7:A:823:GOL:C1	1.66	1.24
1:A:449[B]:LYS:HE2	9:A:945:HOH:O	1.32	1.21
1:A:532[B]:THR:OG1	9:A:920:HOH:O	1.65	1.15
1:A:431[A]:LYS:HE2	6:A:820:EDO:H21	1.17	1.13
1:A:174[A]:ASN:ND2	9:A:921:HOH:O	1.67	1.12
7:A:821:GOL:H11	9:A:926:HOH:O	1.47	1.11
7:A:821:GOL:C1	9:A:926:HOH:O	1.97	1.10
7:A:823:GOL:H32	9:A:934:HOH:O	1.53	1.09
1:A:610[B]:ASN:OD1	9:A:923:HOH:O	1.73	1.05
6:A:818:EDO:H11	7:A:823:GOL:H11	1.08	1.05
5:A:805:MLI:H11	7:A:822:GOL:H2	1.39	1.05
6:A:814:EDO:O2	9:A:922:HOH:O	1.73	1.04
1:A:63[A]:HIS:HD2	9:A:1167:HOH:O	1.41	1.02
1:A:174[A]:ASN:OD1	9:A:919:HOH:O	1.81	0.93
7:A:821:GOL:O1	9:A:926:HOH:O	1.85	0.92

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
6:A:814:EDO:O1	9:A:925:HOH:O	1.80	0.90
1:A:666[B]:GLU:OE2	9:A:929:HOH:O	1.91	0.88
5:A:805:MLI:C1	7:A:822:GOL:H2	2.04	0.87
6:A:810:EDO:O2	9:A:930:HOH:O	1.94	0.85
6:A:818:EDO:C1	7:A:823:GOL:H11	2.01	0.85
5:A:805:MLI:H11	7:A:822:GOL:C2	2.06	0.84
1:A:449[B]:LYS:HE3	9:A:908:HOH:O	1.77	0.84
1:A:174[A]:ASN:CG	9:A:919:HOH:O	1.92	0.83
6:A:818:EDO:C1	7:A:823:GOL:C1	2.54	0.82
7:A:821:GOL:O2	9:A:928:HOH:O	1.87	0.81
1:A:449[B]:LYS:CE	9:A:945:HOH:O	2.05	0.80
1:A:431[B]:LYS:NZ	6:A:820:EDO:H11	1.97	0.80
1:A:49[B]:GLU:OE1	9:A:932:HOH:O	2.00	0.79
1:A:63[A]:HIS:CD2	9:A:1167:HOH:O	2.20	0.77
1:A:449[A]:LYS:NZ	9:A:934:HOH:O	2.10	0.76
1:A:431[A]:LYS:HE2	6:A:820:EDO:C2	2.09	0.76
1:A:553[A]:LYS:NZ	9:A:936:HOH:O	2.13	0.76
1:A:431[A]:LYS:CE	6:A:820:EDO:H21	2.09	0.75
1:A:49[B]:GLU:HG3	6:A:811:EDO:H11	1.70	0.74
1:A:449[B]:LYS:CE	9:A:908:HOH:O	2.34	0.73
5:A:805:MLI:O6	7:A:822:GOL:O2	2.06	0.73
6:A:818:EDO:H11	7:A:823:GOL:H12	1.69	0.73
5:A:805:MLI:O6	9:A:933:HOH:O	2.07	0.72
1:A:393:TYR:CZ	7:A:821:GOL:H2	2.25	0.71
1:A:49[B]:GLU:CD	9:A:932:HOH:O	2.27	0.71
1:A:431[B]:LYS:HZ2	6:A:820:EDO:H11	1.55	0.71
6:A:816:EDO:O1	9:A:903:HOH:O	0.68	0.68
1:A:610[A]:ASN:OD1	9:A:935[A]:HOH:O	2.12	0.68
1:A:213[B]:ASN:CG	9:A:1370[B]:HOH:O	2.33	0.67
1:A:156:ASN:HD21	1:A:161:GLY:H	1.44	0.65
1:A:180[B]:GLN:OE1	1:A:404:HIS:CE1	2.49	0.64
6:A:810:EDO:C1	9:A:1527:HOH:O	2.45	0.64
1:A:449[B]:LYS:NZ	9:A:908:HOH:O	2.30	0.63
6:A:809:EDO:C2	9:A:1246:HOH:O	2.04	0.63
5:A:805:MLI:O9	5:A:805:MLI:O7	2.07	0.63
1:A:490[B]:THR:HG23	9:A:1419[B]:HOH:O	2.00	0.61
1:A:30:HIS:HD2	9:A:1066:HOH:O	1.85	0.60
6:A:815[A]:EDO:C1	9:A:1308:HOH:O	2.51	0.59
1:A:666[B]:GLU:N	1:A:666[B]:GLU:OE1	2.28	0.57
7:A:823:GOL:C3	9:A:934:HOH:O	2.27	0.57
7:A:824:GOL:O1	9:A:913:HOH:O	2.10	0.57

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
6:A:815[A]:EDO:H11	9:A:1308:HOH:O	2.05	0.56
5:A:805:MLI:C1	7:A:822:GOL:C2	2.75	0.56
1:A:452:ASN:OD1	9:A:937:HOH:O	2.18	0.56
1:A:317:HIS:HD2	1:A:375:THR:OG1	1.89	0.56
5:A:804:MLI:O9	5:A:805:MLI:H11	2.06	0.55
5:A:805:MLI:C1	7:A:822:GOL:O2	2.56	0.54
1:A:247[A]:SER:OG	6:A:810:EDO:C2	2.56	0.54
1:A:86:ASN:HD22	1:A:86:ASN:C	2.11	0.53
1:A:398:GLU:OE2	5:A:804:MLI:O9	2.27	0.53
1:A:468:GLU:HG2	1:A:469:GLU:OE2	2.08	0.53
1:A:431[B]:LYS:HZ3	6:A:820:EDO:H11	1.71	0.53
1:A:431[A]:LYS:HE3	9:A:1525:HOH:O	2.08	0.52
1:A:213[B]:ASN:ND2	9:A:1370[B]:HOH:O	2.43	0.52
1:A:634[A]:ASN:HA	9:A:978:HOH:O	2.10	0.51
5:A:805:MLI:C2	7:A:822:GOL:O2	2.58	0.50
1:A:247[A]:SER:OG	6:A:810:EDO:H21	2.11	0.50
1:A:449[C]:LYS:CE	9:A:945:HOH:O	1.99	0.50
5:A:805:MLI:H11	7:A:822:GOL:O2	2.11	0.50
5:A:805:MLI:O8	9:A:901:HOH:O	0.53	0.48
1:A:666[B]:GLU:H	1:A:666[B]:GLU:CD	2.15	0.48
1:A:4:LYS:NZ	9:A:962:HOH:O	2.45	0.48
1:A:129[B]:GLN:NE2	9:A:938:HOH:O	2.19	0.48
1:A:213[B]:ASN:OD1	9:A:1370[B]:HOH:O	2.20	0.48
1:A:518:ALA:HB1	6:A:818:EDO:H22	1.95	0.48
1:A:702:LEU:O	9:A:939:HOH:O	2.20	0.47
1:A:393:TYR:CE2	7:A:821:GOL:H2	2.50	0.46
5:A:805:MLI:H12	7:A:822:GOL:H2	1.94	0.46
1:A:320:GLN:HG2	1:A:371:ALA:O	2.15	0.46
1:A:223:THR:HG21	4:A:803:CE6:H3B	1.99	0.45
1:A:354:LYS:HE2	9:A:1638:HOH:O	2.17	0.45
1:A:47:THR:HA	1:A:150:TRP:O	2.17	0.45
1:A:247[A]:SER:OG	6:A:810:EDO:H22	2.18	0.44
1:A:324:ASN:C	1:A:324:ASN:HD22	2.21	0.44
1:A:324:ASN:ND2	1:A:327:ALA:H	2.15	0.43
1:A:86:ASN:C	1:A:86:ASN:ND2	2.71	0.43
1:A:421[A]:ARG:HD2	9:A:1247:HOH:O	2.18	0.43
1:A:279:ASN:HB2	9:A:968:HOH:O	2.18	0.43
1:A:60:LEU:HD12	1:A:691:THR:HG22	2.01	0.42
7:A:823:GOL:O2	9:A:1310:HOH:O	0.48	0.42
1:A:213[B]:ASN:OD1	1:A:214:GLN:NE2	2.52	0.42
1:A:63[B]:HIS:HD2	9:A:924:HOH:O	0.34	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:63[B]:HIS:CD2	9:A:924:HOH:O	0.81	0.42
1:A:201:PRO:O	1:A:202[A]:ASN:HB2	2.19	0.42
1:A:321:GLY:HA3	1:A:414[B]:MET:SD	2.59	0.41
1:A:403:TYR:CE2	4:A:803:CE6:O3A	2.73	0.41
1:A:589:GLU:CG	1:A:677:VAL:HG13	2.51	0.41
1:A:351:LYS:HE3	9:A:1608:HOH:O	2.21	0.41
1:A:30:HIS:HE1	1:A:38:GLU:OE1	2.04	0.41
1:A:647[A]:ILE:HA	1:A:647[A]:ILE:HD12	1.96	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)
9:A:1193:HOH:O	9:A:1451:HOH:O[3_454]	1.93	0.27

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	728/709 (103%)	707 (97%)	19 (3%)	2 (0%)	46 41

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	88	ASP
1	A	39	ALA

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	601/580 (104%)	576 (96%)	25 (4%)	36	31

All (25) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	66	ARG
1	A	86	ASN
1	A	87	GLU
1	A	115	ASP
1	A	121	LEU
1	A	211	LYS
1	A	214	GLN
1	A	215	VAL
1	A	221	ARG
1	A	258	ASP
1	A	324	ASN
1	A	347[A]	SER
1	A	347[B]	SER
1	A	367	GLU
1	A	394	ASP
1	A	463	ARG
1	A	468	GLU
1	A	490[A]	THR
1	A	490[B]	THR
1	A	634[A]	ASN
1	A	634[B]	ASN
1	A	654	SER
1	A	655	SER
1	A	658	SER
1	A	688	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	30	HIS
1	A	86	ASN
1	A	120	GLN
1	A	156	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	177	GLN
1	A	234	GLN
1	A	317	HIS
1	A	324	ASN
1	A	410	ASN
1	A	439	ASN
1	A	452	ASN
1	A	516	GLN
1	A	607	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](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 42 ligands modelled in this entry, 16 are monoatomic - leaving 26 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	CBI	A	802	8	24,24,24	1.02	1 (4%)	35,35,35	1.51	5 (14%)
4	CE6	A	803	8	72,72,72	0.99	2 (2%)	107,107,107	1.29	11 (10%)
5	MLI	A	804	7	0,6,6	0.00	-	0,7,7	0.00	-
5	MLI	A	805	-	0,6,6	0.00	-	0,7,7	0.00	-

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
5	MLI	A	806	-	0,6,6	0.00	-	0,7,7	0.00	-
6	EDO	A	807	-	3,3,3	0.44	0	2,2,2	0.25	0
6	EDO	A	808	-	3,3,3	0.24	0	2,2,2	1.16	0
6	EDO	A	809	-	3,3,3	0.54	0	2,2,2	0.53	0
6	EDO	A	810	-	3,3,3	0.43	0	2,2,2	0.39	0
6	EDO	A	811	8	3,3,3	0.84	0	2,2,2	0.87	0
6	EDO	A	812[A]	-	3,3,3	0.92	0	2,2,2	0.04	0
6	EDO	A	812[B]	-	3,3,3	0.71	0	2,2,2	0.27	0
6	EDO	A	813	-	3,3,3	0.40	0	2,2,2	0.78	0
6	EDO	A	814	-	3,3,3	0.44	0	2,2,2	0.56	0
6	EDO	A	815[A]	-	3,3,3	0.53	0	2,2,2	0.33	0
6	EDO	A	815[B]	-	3,3,3	0.44	0	2,2,2	0.10	0
6	EDO	A	816	-	3,3,3	0.34	0	2,2,2	0.88	0
6	EDO	A	817	8	3,3,3	0.37	0	2,2,2	0.99	0
6	EDO	A	818	8	3,3,3	0.82	0	2,2,2	0.43	0
6	EDO	A	819[A]	-	3,3,3	0.52	0	2,2,2	0.38	0
6	EDO	A	819[B]	-	3,3,3	0.48	0	2,2,2	0.30	0
6	EDO	A	820	-	3,3,3	0.43	0	2,2,2	0.50	0
7	GOL	A	821	-	5,5,5	0.83	0	5,5,5	1.41	1 (20%)
7	GOL	A	822	5	5,5,5	0.62	0	5,5,5	1.18	0
7	GOL	A	823	8	5,5,5	0.47	0	5,5,5	1.35	1 (20%)
7	GOL	A	824	-	5,5,5	0.40	0	5,5,5	0.93	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	CBI	A	802	8	-	0/8/48/48	0/2/2/2
4	CE6	A	803	8	-	0/32/152/152	0/6/6/6
5	MLI	A	804	7	-	0/0/4/4	0/0/0/0
5	MLI	A	805	-	-	0/0/4/4	0/0/0/0
5	MLI	A	806	-	-	0/0/4/4	0/0/0/0
6	EDO	A	807	-	-	0/1/1/1	0/0/0/0
6	EDO	A	808	-	-	0/1/1/1	0/0/0/0
6	EDO	A	809	-	-	0/1/1/1	0/0/0/0
6	EDO	A	810	-	-	0/1/1/1	0/0/0/0
6	EDO	A	811	8	-	0/1/1/1	0/0/0/0
6	EDO	A	812[A]	-	-	0/1/1/1	0/0/0/0
6	EDO	A	812[B]	-	-	0/1/1/1	0/0/0/0
6	EDO	A	813	-	-	0/1/1/1	0/0/0/0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	EDO	A	814	-	-	0/1/1/1	0/0/0/0
6	EDO	A	815[A]	-	-	0/1/1/1	0/0/0/0
6	EDO	A	815[B]	-	-	0/1/1/1	0/0/0/0
6	EDO	A	816	-	-	0/1/1/1	0/0/0/0
6	EDO	A	817	8	-	0/1/1/1	0/0/0/0
6	EDO	A	818	8	-	0/1/1/1	0/0/0/0
6	EDO	A	819[A]	-	-	0/1/1/1	0/0/0/0
6	EDO	A	819[B]	-	-	0/1/1/1	0/0/0/0
6	EDO	A	820	-	-	0/1/1/1	0/0/0/0
7	GOL	A	821	-	-	0/4/4/4	0/0/0/0
7	GOL	A	822	5	-	0/4/4/4	0/0/0/0
7	GOL	A	823	8	-	0/4/4/4	0/0/0/0
7	GOL	A	824	-	-	0/4/4/4	0/0/0/0

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	802	CBI	O4-C4	2.04	1.47	1.43
4	A	803	CE6	O3A-C3A	2.46	1.48	1.43
4	A	803	CE6	O5C-C1C	3.12	1.49	1.41

All (18) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	802	CBI	C1'-C2'-C3'	-3.94	104.23	110.68
4	A	803	CE6	O1A-C1A-O5A	-3.93	99.36	110.33
3	A	802	CBI	C1'-O5'-C5'	-3.71	106.45	113.54
4	A	803	CE6	O4A-C4A-C5A	-2.78	101.93	109.33
4	A	803	CE6	O4B-C4B-C5B	-2.33	103.12	109.33
7	A	823	GOL	C3-C2-C1	-2.20	101.78	111.06
4	A	803	CE6	O2C-C2C-C3C	-2.13	105.55	110.36
3	A	802	CBI	C1-O5-C5	-2.04	109.74	113.74
4	A	803	CE6	O4D-C4D-C3D	2.03	112.47	107.18
7	A	821	GOL	C3-C2-C1	2.10	119.89	111.06
4	A	803	CE6	O5D-C1D-C2D	2.12	114.69	110.28
4	A	803	CE6	O5A-C1A-C2A	2.28	114.00	110.00
4	A	803	CE6	O3A-C3A-C2A	2.30	115.56	110.36
4	A	803	CE6	O2C-C2C-C1C	2.32	115.14	110.01
3	A	802	CBI	C3'-C4'-C5'	2.34	116.19	110.85
4	A	803	CE6	O1A-C1A-C2A	2.38	115.63	109.05
3	A	802	CBI	O2'-C2'-C1'	2.91	116.12	109.74
4	A	803	CE6	O3B-C3B-C4B	3.18	117.43	109.89

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

15 monomers are involved in 51 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	803	CE6	2	0
5	A	804	MLI	2	0
5	A	805	MLI	13	0
6	A	809	EDO	2	0
6	A	810	EDO	5	0
6	A	811	EDO	1	0
6	A	814	EDO	2	0
6	A	815[A]	EDO	2	0
6	A	816	EDO	1	0
6	A	818	EDO	6	0
6	A	820	EDO	6	0
7	A	821	GOL	6	0
7	A	822	GOL	9	0
7	A	823	GOL	8	0
7	A	824	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	701/709 (98%)	-0.57	8 (1%) 82 83	7, 13, 31, 85	0

All (8) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	306	TYR	4.7
1	A	702	LEU	4.2
1	A	88	ASP	3.4
1	A	87	GLU	3.2
1	A	213[A]	ASN	2.8
1	A	66	ARG	2.2
1	A	468	GLU	2.1
1	A	667	ALA	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 carbohydrates in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 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	LLDF	B-factors(\AA^2)	Q<0.9
8	NA	A	839	1/1	0.75	0.37	12.68	28,28,28,28	1
7	GOL	A	824	6/6	0.89	0.24	12.05	18,19,20,22	6
5	MLI	A	805	7/7	0.88	0.19	11.57	16,17,19,19	7
6	EDO	A	812[A]	4/4	0.88	0.21	10.69	10,13,14,16	4
6	EDO	A	812[B]	4/4	0.88	0.21	10.69	29,33,33,33	4
6	EDO	A	818	4/4	0.79	0.41	9.84	14,14,15,16	4
5	MLI	A	804	7/7	0.93	0.16	9.21	18,18,20,22	7
7	GOL	A	821	6/6	0.90	0.23	7.65	13,14,15,16	6
8	NA	A	834	1/1	0.96	0.14	7.38	32,32,32,32	0
7	GOL	A	822	6/6	0.94	0.14	7.26	26,28,29,29	6
8	NA	A	830	1/1	0.86	0.20	6.83	39,39,39,39	1
6	EDO	A	809	4/4	0.95	0.13	3.92	9,9,9,10	4
6	EDO	A	817	4/4	0.91	0.17	3.41	25,26,26,26	4
7	GOL	A	823	6/6	0.88	0.22	3.15	22,25,26,27	6
6	EDO	A	815[B]	4/4	0.92	0.15	2.30	18,18,19,20	4
6	EDO	A	820	4/4	0.95	0.15	2.26	28,29,31,31	4
8	NA	A	836	1/1	0.91	0.14	2.22	27,27,27,27	1
6	EDO	A	810	4/4	0.90	0.14	2.20	23,24,24,25	4
5	MLI	A	806	7/7	0.82	0.36	2.18	40,41,45,45	7
6	EDO	A	808	4/4	0.87	0.13	2.15	24,24,24,24	4
6	EDO	A	815[A]	4/4	0.92	0.15	2.06	26,29,30,30	4
4	CE6	A	803	67/67	0.90	0.14	1.38	14,24,32,35	0
6	EDO	A	811	4/4	0.91	0.16	1.38	24,27,29,32	0
8	NA	A	828	1/1	0.82	0.15	0.55	41,41,41,41	0
6	EDO	A	807	4/4	0.97	0.08	0.51	18,18,18,19	0
6	EDO	A	816	4/4	0.84	0.19	0.45	38,39,40,41	4
6	EDO	A	813	4/4	0.90	0.11	0.41	31,37,37,38	0
8	NA	A	837	1/1	0.96	0.13	0.15	38,38,38,38	0
8	NA	A	832	1/1	0.95	0.08	-0.20	34,34,34,34	0
3	CBI	A	802	23/23	0.93	0.09	-0.30	18,24,31,45	0
2	CA	A	801	1/1	0.99	0.06	-0.99	25,25,25,25	0
6	EDO	A	814	4/4	0.97	0.12	-	14,15,15,17	4
8	NA	A	826	1/1	0.97	0.09	-	32,32,32,32	0
8	NA	A	831	1/1	0.98	0.08	-	28,28,28,28	0
8	NA	A	827	1/1	0.93	0.28	-	36,36,36,36	0
8	NA	A	833	1/1	0.94	0.16	-	42,42,42,42	0
8	NA	A	835	1/1	0.82	0.10	-	41,41,41,41	0
6	EDO	A	819[A]	4/4	0.74	0.23	-	53,55,57,58	4
8	NA	A	838	1/1	0.97	0.07	-	31,31,31,31	0
6	EDO	A	819[B]	4/4	0.74	0.23	-	51,53,53,55	4
8	NA	A	829	1/1	0.92	0.21	-	27,27,27,27	1
8	NA	A	825	1/1	0.97	0.06	-	31,31,31,31	0

6.5 Other polymers [i](#)

There are no such residues in this entry.