Full wwPDB X-ray Structure Validation Report (i)

May 13, 2020 – 05:49 am BST

PDB ID	:	6GXS
Title	:	Crystal structure of CV39L lectin from Chromobacterium violaceum at 1.8 A resolution
Authors	:	Sykorova, P.; Novotna, J.; Demo, G.; Pompidor, G.; Dubska, E.; Komarek, J.; Fujdiarova, E.; Haronikova, L.; Varrot, A.; Imberty, A.; Shilova, N.; Bovin, N.; Pokorna, M.; Wimmerova, M.
Deposited on	:	2018-06-27
Resolution	:	1.80 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 following versions of software and data (see references (1)) were used in the production of this report:

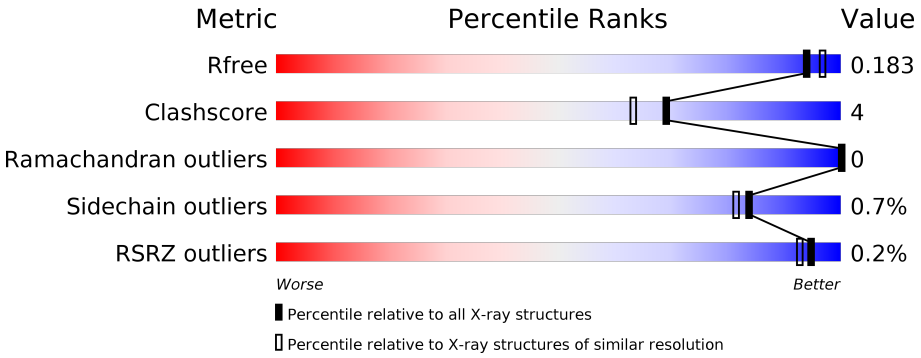
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA,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

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 1.80 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	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	376	88% 7%
1	B	376	88% 8%
1	C	376	87% 9%
1	D	376	88% 7%

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	SO4	C	403	-	-	X	-

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 12196 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 CV39L lectin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	362	Total	C	N	O	S	0	2	0
			2707	1722	465	517	3			
1	B	363	Total	C	N	O	S	0	0	0
			2704	1716	466	520	2			
1	C	361	Total	C	N	O	S	0	4	0
			2712	1724	467	518	3			
1	D	362	Total	C	N	O	S	0	0	0
			2698	1713	465	518	2			

There are 56 discrepancies between the modelled and reference sequences:

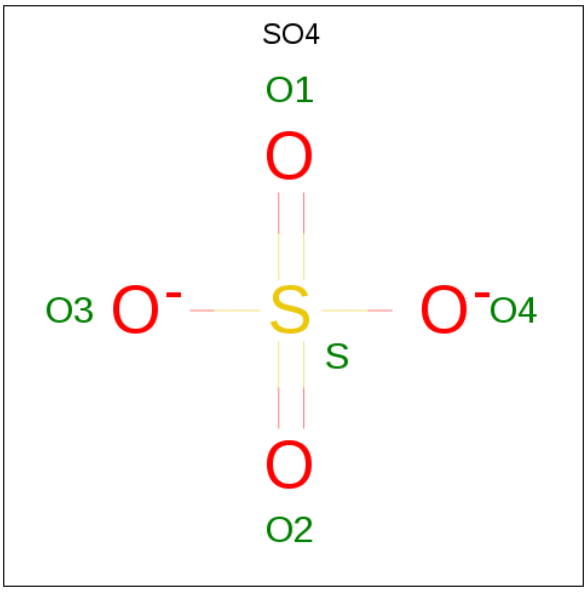
Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP Q7NZ70
A	1	HIS	-	expression tag	UNP Q7NZ70
A	2	HIS	-	expression tag	UNP Q7NZ70
A	3	HIS	-	expression tag	UNP Q7NZ70
A	4	HIS	-	expression tag	UNP Q7NZ70
A	5	HIS	-	expression tag	UNP Q7NZ70
A	6	HIS	-	expression tag	UNP Q7NZ70
A	7	GLU	-	expression tag	UNP Q7NZ70
A	8	ASN	-	expression tag	UNP Q7NZ70
A	9	LEU	-	expression tag	UNP Q7NZ70
A	10	TYR	-	expression tag	UNP Q7NZ70
A	11	PHE	-	expression tag	UNP Q7NZ70
A	12	GLN	-	expression tag	UNP Q7NZ70
A	13	SER	-	expression tag	UNP Q7NZ70
B	0	MET	-	initiating methionine	UNP Q7NZ70
B	1	HIS	-	expression tag	UNP Q7NZ70
B	2	HIS	-	expression tag	UNP Q7NZ70
B	3	HIS	-	expression tag	UNP Q7NZ70
B	4	HIS	-	expression tag	UNP Q7NZ70
B	5	HIS	-	expression tag	UNP Q7NZ70
B	6	HIS	-	expression tag	UNP Q7NZ70

Continued on next page...

Continued from previous page...

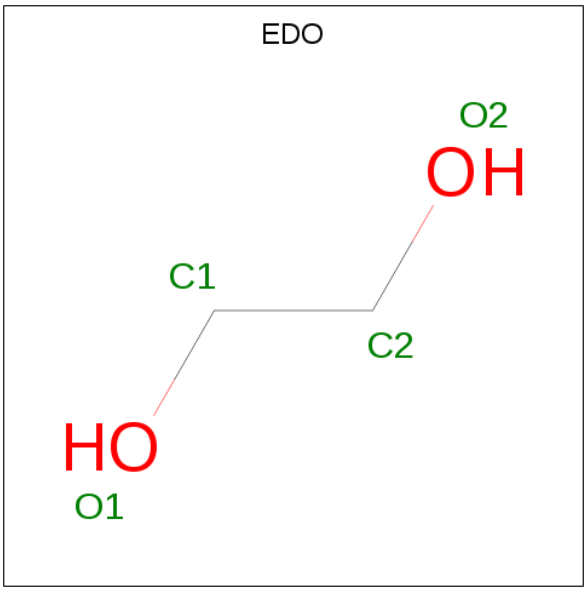
Chain	Residue	Modelled	Actual	Comment	Reference
B	7	GLU	-	expression tag	UNP Q7NZ70
B	8	ASN	-	expression tag	UNP Q7NZ70
B	9	LEU	-	expression tag	UNP Q7NZ70
B	10	TYR	-	expression tag	UNP Q7NZ70
B	11	PHE	-	expression tag	UNP Q7NZ70
B	12	GLN	-	expression tag	UNP Q7NZ70
B	13	SER	-	expression tag	UNP Q7NZ70
C	0	MET	-	initiating methionine	UNP Q7NZ70
C	1	HIS	-	expression tag	UNP Q7NZ70
C	2	HIS	-	expression tag	UNP Q7NZ70
C	3	HIS	-	expression tag	UNP Q7NZ70
C	4	HIS	-	expression tag	UNP Q7NZ70
C	5	HIS	-	expression tag	UNP Q7NZ70
C	6	HIS	-	expression tag	UNP Q7NZ70
C	7	GLU	-	expression tag	UNP Q7NZ70
C	8	ASN	-	expression tag	UNP Q7NZ70
C	9	LEU	-	expression tag	UNP Q7NZ70
C	10	TYR	-	expression tag	UNP Q7NZ70
C	11	PHE	-	expression tag	UNP Q7NZ70
C	12	GLN	-	expression tag	UNP Q7NZ70
C	13	SER	-	expression tag	UNP Q7NZ70
D	0	MET	-	initiating methionine	UNP Q7NZ70
D	1	HIS	-	expression tag	UNP Q7NZ70
D	2	HIS	-	expression tag	UNP Q7NZ70
D	3	HIS	-	expression tag	UNP Q7NZ70
D	4	HIS	-	expression tag	UNP Q7NZ70
D	5	HIS	-	expression tag	UNP Q7NZ70
D	6	HIS	-	expression tag	UNP Q7NZ70
D	7	GLU	-	expression tag	UNP Q7NZ70
D	8	ASN	-	expression tag	UNP Q7NZ70
D	9	LEU	-	expression tag	UNP Q7NZ70
D	10	TYR	-	expression tag	UNP Q7NZ70
D	11	PHE	-	expression tag	UNP Q7NZ70
D	12	GLN	-	expression tag	UNP Q7NZ70
D	13	SER	-	expression tag	UNP Q7NZ70

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



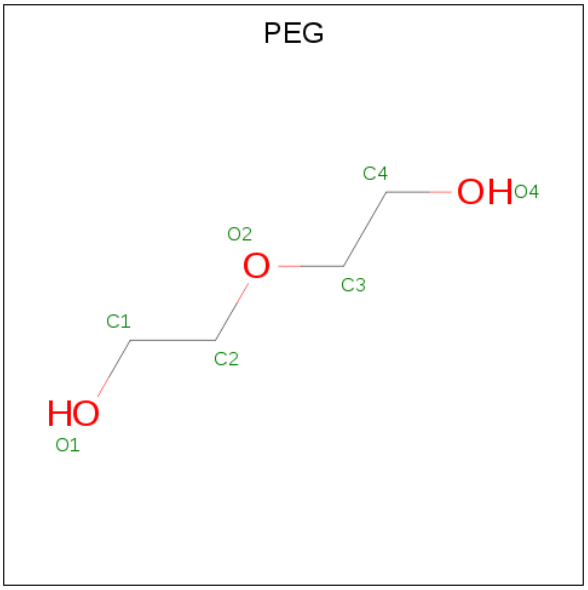
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	C	1	Total O S 5 4 1	0	0
2	C	1	Total O S 5 4 1	0	0
2	C	1	Total O S 5 4 1	0	0
2	C	1	Total O S 5 4 1	0	0
2	D	1	Total O S 5 4 1	0	0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

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



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

- Molecule 4 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



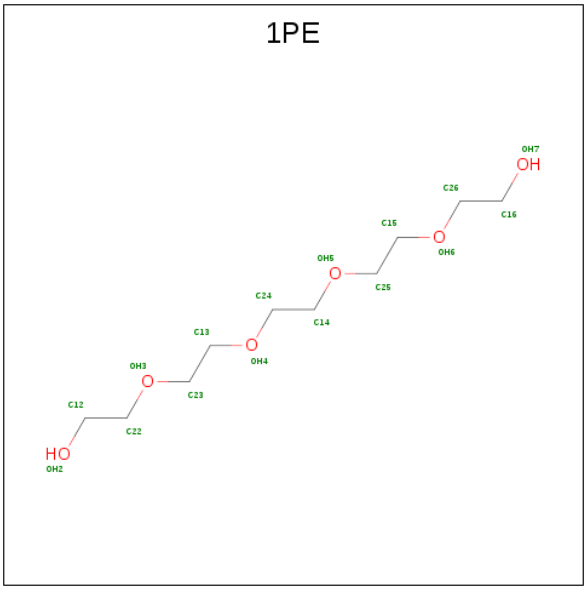
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total	C	O	0	0
			7	4	3		
4	B	1	Total	C	O	0	0
			7	4	3		

Continued on next page...

Continued from previous page...

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

- Molecule 5 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula: $C_{10}H_{22}O_6$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total C O 16 10 6	0	0

- Molecule 6 is water.

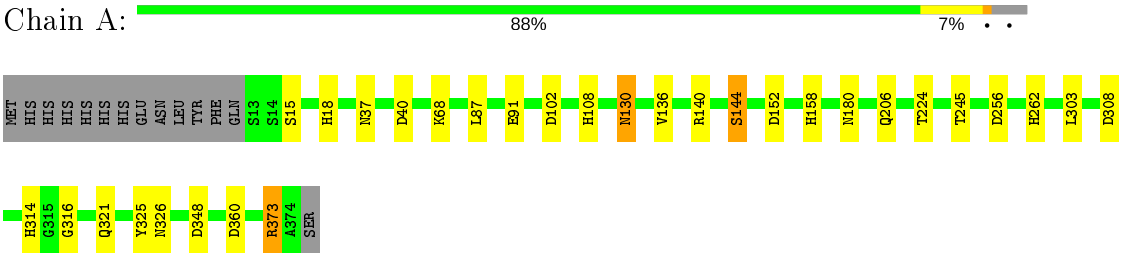
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	334	Total O 334 334	0	0
6	B	343	Total O 343 343	0	0
6	C	312	Total O 312 312	0	0
6	D	286	Total O 286 286	0	0

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 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.

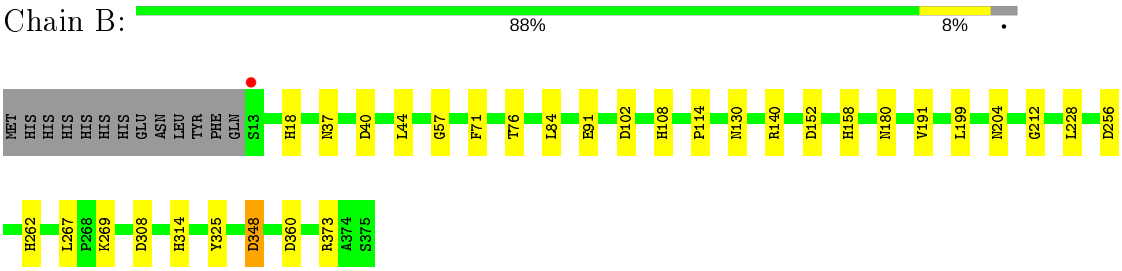
- Molecule 1: CV39L lectin

Chain A:



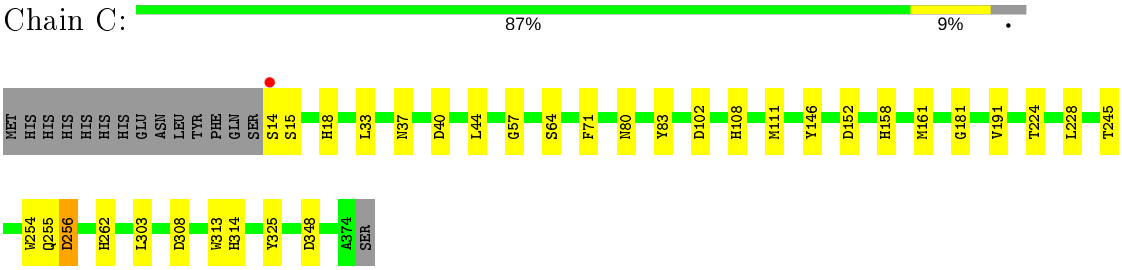
- Molecule 1: CV39L lectin

Chain B:



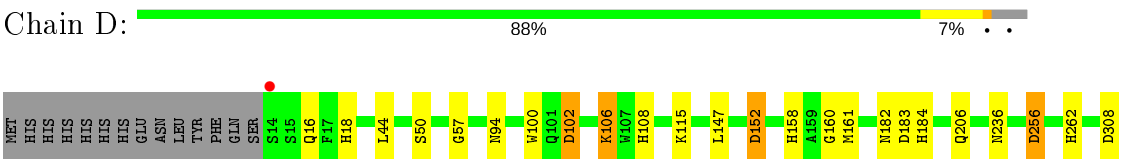
- Molecule 1: CV39L lectin

Chain C:



- Molecule 1: CV39L lectin

Chain D:




H314
Y325
N338
D348
N352
L355
E362
G368
S375

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	65.71Å 123.78Å 180.94Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	73.04 – 1.80 73.04 – 1.80	Depositor EDS
% Data completeness (in resolution range)	100.0 (73.04-1.80) 100.0 (73.04-1.80)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	6.28 (at 1.80Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.153 , 0.189 0.153 , 0.183	Depositor DCC
R_{free} test set	6893 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	11.9	Xtriage
Anisotropy	0.166	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 46.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	12196	wwPDB-VP
Average B, all atoms (Å^2)	12.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.07% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: PEG, 1PE, EDO, SO4

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	1.19	1/2803 (0.0%)	1.02	3/3835 (0.1%)
1	B	1.16	0/2794	0.99	5/3821 (0.1%)
1	C	1.19	5/2812 (0.2%)	1.03	5/3847 (0.1%)
1	D	1.17	2/2788 (0.1%)	1.03	7/3813 (0.2%)
All	All	1.18	8/11197 (0.1%)	1.02	20/15316 (0.1%)

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	100	TRP	CG-CD1	7.12	1.46	1.36
1	C	64	SER	CB-OG	-5.97	1.34	1.42
1	C	146	TYR	CE1-CZ	-5.92	1.30	1.38
1	C	83	TYR	CB-CG	5.89	1.60	1.51
1	A	144	SER	CB-OG	5.65	1.49	1.42
1	C	313	TRP	CG-CD1	-5.60	1.28	1.36
1	D	50	SER	CB-OG	5.25	1.49	1.42
1	C	254	TRP	CG-CD1	5.07	1.43	1.36

All (20) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	102	ASP	CB-CG-OD1	7.80	125.32	118.30
1	C	256	ASP	CB-CG-OD1	7.75	125.28	118.30
1	D	161	MET	CG-SD-CE	-7.46	88.27	100.20
1	A	140	ARG	NE-CZ-NH2	-7.41	116.59	120.30
1	D	183	ASP	CB-CG-OD1	6.58	124.22	118.30
1	D	102	ASP	CB-CG-OD2	-6.54	112.41	118.30
1	D	183	ASP	CB-CG-OD2	-6.37	112.57	118.30
1	B	140	ARG	NE-CZ-NH1	5.85	123.23	120.30
1	D	256	ASP	CB-CG-OD1	5.82	123.54	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	68	LYS	CD-CE-NZ	-5.66	98.69	111.70
1	C	256	ASP	CB-CG-OD2	-5.65	113.22	118.30
1	B	360	ASP	CB-CG-OD2	-5.65	113.22	118.30
1	A	360	ASP	CB-CG-OD1	5.61	123.35	118.30
1	B	267	LEU	CB-CG-CD1	-5.57	101.54	111.00
1	C	33	LEU	CA-CB-CG	-5.51	102.63	115.30
1	B	348	ASP	CB-CG-OD1	5.50	123.25	118.30
1	C	111	MET	CG-SD-CE	-5.39	91.57	100.20
1	C	303	LEU	N-CA-C	-5.11	97.22	111.00
1	B	373	ARG	NE-CZ-NH2	-5.10	117.75	120.30
1	D	152	ASP	CB-CG-OD1	5.01	122.81	118.30

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	2707	0	2499	19	0
1	B	2704	0	2484	22	0
1	C	2712	0	2499	21	0
1	D	2698	0	2479	14	0
2	A	10	0	0	0	0
2	B	15	0	0	1	0
2	C	20	0	0	2	0
2	D	10	0	0	0	0
3	A	4	0	6	0	0
3	B	4	0	6	0	0
4	B	14	0	20	2	0
4	C	7	0	10	0	0
5	B	16	0	22	2	0
6	A	334	0	0	5	0
6	B	343	0	0	10	0
6	C	312	0	0	2	0
6	D	286	0	0	2	0
All	All	12196	0	10025	78	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 4.

All (78) 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:C:158:HIS:HD2	2:C:403:SO4:O3	1.60	0.83
1:A:321:GLN:NE2	1:A:326:ASN:OD1	2.23	0.72
1:C:256:ASP:OD2	1:C:262:HIS:HE1	1.73	0.72
1:B:256:ASP:OD2	1:B:262:HIS:HE1	1.76	0.69
1:C:158:HIS:CD2	2:C:403:SO4:O3	2.48	0.63
1:A:87[B]:LEU:HD11	1:A:136:VAL:HG22	1.81	0.61
1:C:18:HIS:HD2	1:C:37:ASN:OD1	1.84	0.61
1:A:256:ASP:OD2	1:A:262:HIS:HE1	1.84	0.61
1:B:152:ASP:OD2	1:B:158:HIS:HE1	1.85	0.60
1:B:114:PRO:HD2	6:B:784:HOH:O	2.00	0.60
6:A:777:HOH:O	1:B:158:HIS:HD2	1.85	0.60
1:D:256:ASP:OD2	1:D:262:HIS:HE1	1.85	0.60
1:C:152:ASP:OD2	1:C:158:HIS:HE1	1.84	0.59
1:C:191:VAL:HG22	1:C:228:LEU:HG	1.84	0.59
1:B:204:ASN:HD22	4:B:406:PEG:H32	1.69	0.58
1:C:308:ASP:OD2	1:C:314:HIS:HE1	1.87	0.57
1:B:102:ASP:OD2	1:B:108:HIS:HE1	1.87	0.57
1:C:181:GLY:HA3	1:C:255:GLN:HE22	1.70	0.56
1:A:15:SER:HB3	1:A:373:ARG:HG2	1.86	0.55
1:B:18:HIS:HE1	6:B:657:HOH:O	1.90	0.54
1:A:262:HIS:HD2	6:A:809:HOH:O	1.90	0.54
5:B:407:1PE:H232	6:B:505:HOH:O	2.08	0.54
1:C:14:SER:O	1:C:15:SER:OG	2.13	0.53
1:A:102:ASP:OD2	1:A:108:HIS:HE1	1.90	0.53
1:D:18:HIS:HE1	6:D:696:HOH:O	1.91	0.53
1:B:114:PRO:CD	6:B:784:HOH:O	2.54	0.53
1:B:204:ASN:HD22	4:B:406:PEG:C3	2.21	0.53
1:C:224:THR:OG1	1:C:245:THR:HG21	2.08	0.52
1:C:18:HIS:HE1	6:C:697:HOH:O	1.92	0.52
1:B:18:HIS:HD2	1:B:37:ASN:OD1	1.93	0.52
1:A:87[B]:LEU:HD13	1:A:136:VAL:HG13	1.91	0.52
1:D:308:ASP:OD2	1:D:314:HIS:HE1	1.92	0.51
1:B:308:ASP:OD2	1:B:314:HIS:HE1	1.92	0.51
1:C:14:SER:C	1:C:15:SER:HG	2.09	0.51
1:A:18:HIS:HE1	6:A:751:HOH:O	1.92	0.51
6:B:506:HOH:O	1:C:245:THR:HG22	2.09	0.51
1:A:308:ASP:OD2	1:A:314:HIS:HE1	1.94	0.51

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:152:ASP:OD2	1:D:158:HIS:HE1	1.95	0.50
1:C:191:VAL:CG2	1:C:228:LEU:HG	2.41	0.50
1:D:325:TYR:CZ	1:D:348:ASP:HB2	2.47	0.50
6:B:506:HOH:O	1:C:245:THR:CG2	2.60	0.49
1:C:325:TYR:CZ	1:C:348:ASP:HB2	2.48	0.49
1:C:102:ASP:OD2	1:C:108:HIS:HE1	1.96	0.49
1:D:182:ASN:HD21	1:D:236:ASN:HB2	1.78	0.48
1:B:199:LEU:O	1:B:212:GLY:HA3	2.14	0.47
1:C:40:ASP:HB3	1:C:71:PHE:O	2.14	0.46
1:D:44:LEU:O	1:D:57:GLY:HA3	2.16	0.46
1:D:102:ASP:OD2	1:D:108:HIS:HE1	1.99	0.46
1:C:44:LEU:O	1:C:57:GLY:HA3	2.16	0.45
1:A:224:THR:OG1	1:A:245:THR:HG21	2.16	0.45
1:A:325:TYR:CZ	1:A:348:ASP:HB2	2.51	0.45
1:B:108:HIS:HD2	2:B:403:SO4:O1	1.98	0.45
1:B:325:TYR:CZ	1:B:348:ASP:HB2	2.51	0.45
1:B:269:LYS:HE2	6:B:533:HOH:O	2.16	0.45
1:C:80:ASN:HD22	1:C:80:ASN:H	1.65	0.44
1:A:18:HIS:HD2	1:A:37:ASN:OD1	2.01	0.43
1:A:152:ASP:OD2	1:A:158:HIS:HE1	2.01	0.43
1:B:191:VAL:CG2	1:B:228:LEU:HG	2.49	0.43
1:B:76:THR:HB	1:B:84:LEU:HD11	2.00	0.43
1:C:161[A]:MET:HE3	6:C:609:HOH:O	2.19	0.43
1:A:15:SER:CB	1:A:373:ARG:HG2	2.48	0.43
1:B:262:HIS:HD2	6:B:806:HOH:O	2.01	0.43
1:B:40:ASP:HB3	1:B:71:PHE:O	2.20	0.42
1:D:16:GLN:HE21	1:D:352:ASN:HD21	1.66	0.42
1:A:130:ASN:H	1:A:130:ASN:HD22	1.68	0.42
5:B:407:1PE:H261	6:B:505:HOH:O	2.19	0.42
1:B:44:LEU:O	1:B:57:GLY:HA3	2.19	0.42
1:B:180:ASN:ND2	6:B:515:HOH:O	2.51	0.42
1:A:40:ASP:OD2	1:A:91:GLU:OE2	2.38	0.42
1:D:184:HIS:CE1	6:D:512:HOH:O	2.71	0.41
1:D:94:ASN:HD22	1:D:115:LYS:NZ	2.19	0.41
1:A:303:LEU:O	1:A:316:GLY:HA3	2.21	0.41
1:B:40:ASP:OD2	1:B:91:GLU:OE2	2.38	0.41
1:D:355:LEU:O	1:D:368:GLY:HA3	2.21	0.41
1:A:206:GLN:HB3	6:A:791:HOH:O	2.21	0.40
1:D:147:LEU:O	1:D:160:GLY:HA3	2.21	0.40
1:D:102:ASP:OD2	1:D:106:LYS:HG3	2.20	0.40
1:A:180:ASN:ND2	6:A:504:HOH:O	2.47	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	362/376 (96%)	356 (98%)	6 (2%)	0	100	100
1	B	361/376 (96%)	356 (99%)	5 (1%)	0	100	100
1	C	363/376 (96%)	358 (99%)	5 (1%)	0	100	100
1	D	360/376 (96%)	353 (98%)	7 (2%)	0	100	100
All	All	1446/1504 (96%)	1423 (98%)	23 (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	273/285 (96%)	270 (99%)	3 (1%)	73	68
1	B	272/285 (95%)	271 (100%)	1 (0%)	91	89
1	C	274/285 (96%)	274 (100%)	0	100	100
1	D	271/285 (95%)	267 (98%)	4 (2%)	65	56
All	All	1090/1140 (96%)	1082 (99%)	8 (1%)	84	81

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	130	ASN
1	A	144	SER
1	A	373	ARG
1	B	130	ASN
1	D	106	LYS
1	D	206	GLN
1	D	338	ASN
1	D	362	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (63) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	18	HIS
1	A	24	ASN
1	A	94	ASN
1	A	108	HIS
1	A	130	ASN
1	A	153	ASN
1	A	158	HIS
1	A	174	GLN
1	A	180	ASN
1	A	193	ASN
1	A	196	ASN
1	A	236	ASN
1	A	262	HIS
1	A	314	HIS
1	A	321	GLN
1	A	336	ASN
1	A	352	ASN
1	B	18	HIS
1	B	94	ASN
1	B	108	HIS
1	B	130	ASN
1	B	158	HIS
1	B	180	ASN
1	B	193	ASN
1	B	196	ASN
1	B	257	ASN
1	B	262	HIS
1	B	314	HIS
1	B	336	ASN
1	B	352	ASN
1	C	18	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	24	ASN
1	C	80	ASN
1	C	94	ASN
1	C	108	HIS
1	C	130	ASN
1	C	158	HIS
1	C	180	ASN
1	C	193	ASN
1	C	255	GLN
1	C	257	ASN
1	C	262	HIS
1	C	314	HIS
1	C	336	ASN
1	C	352	ASN
1	D	18	HIS
1	D	94	ASN
1	D	108	HIS
1	D	153	ASN
1	D	158	HIS
1	D	174	GLN
1	D	180	ASN
1	D	182	ASN
1	D	193	ASN
1	D	196	ASN
1	D	236	ASN
1	D	255	GLN
1	D	257	ASN
1	D	262	HIS
1	D	314	HIS
1	D	321	GLN
1	D	336	ASN
1	D	352	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)

17 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	SO4	B	402	-	4,4,4	0.52	0	6,6,6	1.00	1 (16%)
2	SO4	C	404	-	4,4,4	0.77	0	6,6,6	1.03	0
4	PEG	C	405	-	6,6,6	0.46	0	5,5,5	0.45	0
4	PEG	B	406	-	6,6,6	0.98	0	5,5,5	1.02	0
2	SO4	B	401	-	4,4,4	0.79	0	6,6,6	1.26	0
2	SO4	A	401	-	4,4,4	1.16	0	6,6,6	0.72	0
2	SO4	D	401	-	4,4,4	0.65	0	6,6,6	0.61	0
3	EDO	A	403	-	3,3,3	0.94	0	2,2,2	0.48	0
2	SO4	C	402	-	4,4,4	0.46	0	6,6,6	1.07	0
2	SO4	A	402	-	4,4,4	0.52	0	6,6,6	0.67	0
2	SO4	C	401	-	4,4,4	1.42	1 (25%)	6,6,6	1.10	0
5	1PE	B	407	-	15,15,15	0.56	0	14,14,14	0.93	1 (7%)
4	PEG	B	405	-	6,6,6	1.03	0	5,5,5	1.22	0
2	SO4	B	403	-	4,4,4	0.45	0	6,6,6	1.04	0
3	EDO	B	404	-	3,3,3	0.39	0	2,2,2	0.51	0
2	SO4	C	403	-	4,4,4	0.38	0	6,6,6	1.39	1 (16%)
2	SO4	D	402	-	4,4,4	0.65	0	6,6,6	0.69	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
4	PEG	C	405	-	-	0/4/4/4	-
4	PEG	B	406	-	-	3/4/4/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	EDO	A	403	-	-	1/1/1/1	-
4	PEG	B	405	-	-	2/4/4/4	-
5	1PE	B	407	-	-	7/13/13/13	-
3	EDO	B	404	-	-	1/1/1/1	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	401	SO4	O1-S	2.14	1.57	1.46

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	403	SO4	O4-S-O3	-2.37	98.96	109.06
5	B	407	1PE	OH2-C12-C22	-2.32	98.35	111.81
2	B	402	SO4	O4-S-O2	2.04	119.98	109.31

There are no chirality outliers.

All (14) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	407	1PE	OH4-C13-C23-OH3
4	B	405	PEG	O1-C1-C2-O2
4	B	406	PEG	O2-C3-C4-O4
5	B	407	1PE	OH7-C16-C26-OH6
5	B	407	1PE	OH5-C14-C24-OH4
4	B	406	PEG	O1-C1-C2-O2
3	A	403	EDO	O1-C1-C2-O2
3	B	404	EDO	O1-C1-C2-O2
5	B	407	1PE	C16-C26-OH6-C15
4	B	405	PEG	C1-C2-O2-C3
5	B	407	1PE	OH2-C12-C22-OH3
5	B	407	1PE	OH6-C15-C25-OH5
5	B	407	1PE	C15-C25-OH5-C14
4	B	406	PEG	C1-C2-O2-C3

There are no ring outliers.

4 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	406	PEG	2	0
5	B	407	1PE	2	0
2	B	403	SO4	1	0
2	C	403	SO4	2	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	362/376 (96%)	-0.60	0	100	100	6, 10, 19, 32	0
1	B	363/376 (96%)	-0.63	1 (0%)	94	92	7, 10, 20, 37	0
1	C	361/376 (96%)	-0.64	1 (0%)	94	92	6, 10, 20, 35	0
1	D	362/376 (96%)	-0.53	1 (0%)	94	92	7, 12, 24, 40	0
All	All	1448/1504 (96%)	-0.60	3 (0%)	95	93	6, 10, 21, 40	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	14	SER	4.5
1	B	13	SER	3.5
1	C	14	SER	2.6

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	EDO	A	403	4/4	0.82	0.13	28,30,30,32	0
4	PEG	B	405	7/7	0.84	0.12	31,32,38,40	0
5	1PE	B	407	16/16	0.85	0.15	31,38,43,44	0
4	PEG	B	406	7/7	0.85	0.20	25,27,29,30	0
2	SO4	B	403	5/5	0.88	0.20	41,42,54,56	0
4	PEG	C	405	7/7	0.89	0.12	37,43,44,45	0
3	EDO	B	404	4/4	0.91	0.16	31,32,33,40	0
2	SO4	C	404	5/5	0.92	0.17	26,32,36,36	5
2	SO4	C	401	5/5	0.93	0.16	16,20,28,30	5
2	SO4	B	402	5/5	0.94	0.18	35,44,45,56	0
2	SO4	A	402	5/5	0.95	0.15	36,44,46,48	0
2	SO4	C	403	5/5	0.95	0.17	32,37,42,45	0
2	SO4	A	401	5/5	0.96	0.12	20,23,26,26	0
2	SO4	D	402	5/5	0.96	0.07	32,35,37,39	0
2	SO4	D	401	5/5	0.97	0.09	20,22,28,28	5
2	SO4	B	401	5/5	0.98	0.09	17,18,21,23	0
2	SO4	C	402	5/5	0.98	0.11	16,21,23,28	0

6.5 Other polymers (i)

There are no such residues in this entry.