wwPDB X-ray Structure Validation Summary Report (i)

Dec 16, 2024 – 08:23 PM JST

PDB ID	:	8XIL
Title	:	Cellodextrin phosphorylase from Clostridium thermocellum mutant - all cysteine residues were substituted with serines
Authors	:	Kuga, T.; Sunagawa, N.; Igarashi, K.
Deposited on	:	2023-12-19
Resolution	:	1.21 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The 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:

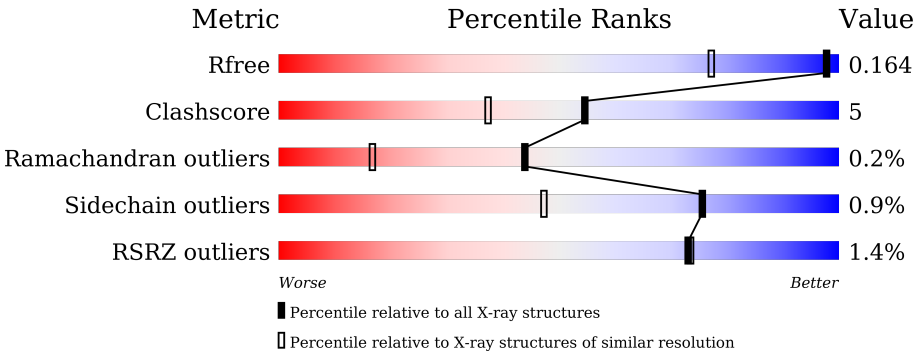
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.21
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

1 Overall quality at a glance (i)

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

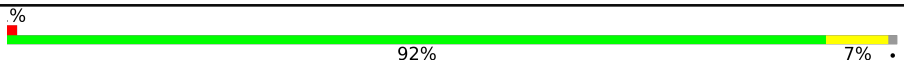
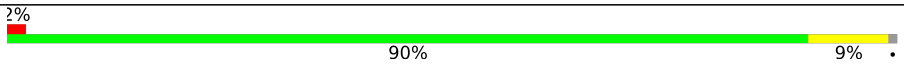
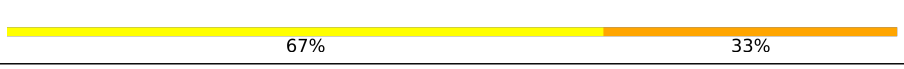
The reported resolution of this entry is 1.21 Å.

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}	164625	1745 (1.24-1.20)
Clashscore	180529	1895 (1.24-1.20)
Ramachandran outliers	177936	1845 (1.24-1.20)
Sidechain outliers	177891	1844 (1.24-1.20)
RSRZ outliers	164620	1744 (1.24-1.20)

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	992	
1	B	992	
2	C	3	
2	D	3	

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
3	ACT	A	1001	-	-	X	-
3	ACT	B	1001	-	-	X	-
4	TRS	B	1003	-	-	X	-
5	SO4	A	1006	-	-	X	-
6	PEG	A	1007	-	-	X	-
7	GOL	B	1011	-	-	X	-
7	GOL	B	1017	-	-	X	-
7	GOL	B	1018	-	-	X	-

2 Entry composition (i)

There are 9 unique types of molecules in this entry. The entry contains 34642 atoms, of which 15432 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 Cellodextrin phosphorylase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	984	Total	C	H	N	O	S	177	55	0
			15858	5312	7532	1393	1596	25			
1	B	984	Total	C	H	N	O	S	188	71	0
			16079	5396	7606	1425	1628	24			

There are 40 discrepancies between the modelled and reference sequences:

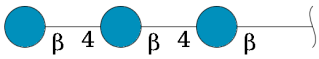
Chain	Residue	Modelled	Actual	Comment	Reference
A	64	SER	CYS	engineered mutation	UNP Q93HT8
A	80	SER	CYS	engineered mutation	UNP Q93HT8
A	225	SER	CYS	engineered mutation	UNP Q93HT8
A	229	SER	CYS	engineered mutation	UNP Q93HT8
A	240	SER	CYS	engineered mutation	UNP Q93HT8
A	353	SER	CYS	engineered mutation	UNP Q93HT8
A	372	SER	CYS	engineered mutation	UNP Q93HT8
A	606	SER	CYS	engineered mutation	UNP Q93HT8
A	625	SER	CYS	engineered mutation	UNP Q93HT8
A	629	ASP	ALA	engineered mutation	UNP Q93HT8
A	872	SER	CYS	engineered mutation	UNP Q93HT8
A	934	SER	CYS	engineered mutation	UNP Q93HT8
A	985	LEU	-	expression tag	UNP Q93HT8
A	986	GLU	-	expression tag	UNP Q93HT8
A	987	HIS	-	expression tag	UNP Q93HT8
A	988	HIS	-	expression tag	UNP Q93HT8
A	989	HIS	-	expression tag	UNP Q93HT8
A	990	HIS	-	expression tag	UNP Q93HT8
A	991	HIS	-	expression tag	UNP Q93HT8
A	992	HIS	-	expression tag	UNP Q93HT8
B	64	SER	CYS	engineered mutation	UNP Q93HT8
B	80	SER	CYS	engineered mutation	UNP Q93HT8
B	225	SER	CYS	engineered mutation	UNP Q93HT8
B	229	SER	CYS	engineered mutation	UNP Q93HT8
B	240	SER	CYS	engineered mutation	UNP Q93HT8

Continued on next page...

Continued from previous page...

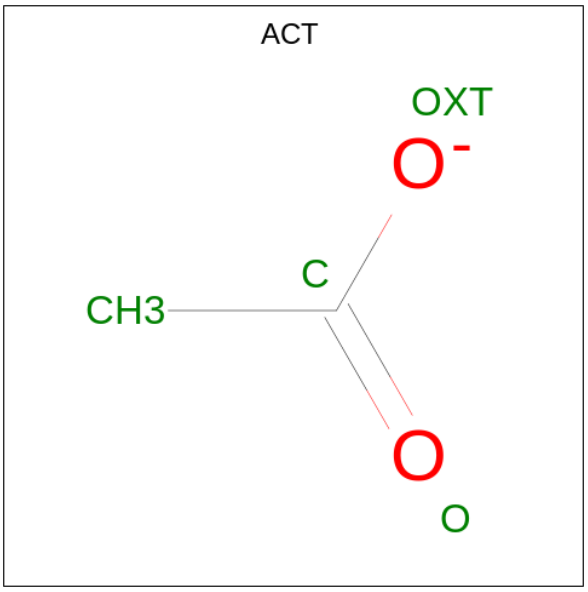
Chain	Residue	Modelled	Actual	Comment	Reference
B	353	SER	CYS	engineered mutation	UNP Q93HT8
B	372	SER	CYS	engineered mutation	UNP Q93HT8
B	606	SER	CYS	engineered mutation	UNP Q93HT8
B	625	SER	CYS	engineered mutation	UNP Q93HT8
B	629	ASP	ALA	engineered mutation	UNP Q93HT8
B	872	SER	CYS	engineered mutation	UNP Q93HT8
B	934	SER	CYS	engineered mutation	UNP Q93HT8
B	985	LEU	-	expression tag	UNP Q93HT8
B	986	GLU	-	expression tag	UNP Q93HT8
B	987	HIS	-	expression tag	UNP Q93HT8
B	988	HIS	-	expression tag	UNP Q93HT8
B	989	HIS	-	expression tag	UNP Q93HT8
B	990	HIS	-	expression tag	UNP Q93HT8
B	991	HIS	-	expression tag	UNP Q93HT8
B	992	HIS	-	expression tag	UNP Q93HT8

- Molecule 2 is an oligosaccharide called beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose.



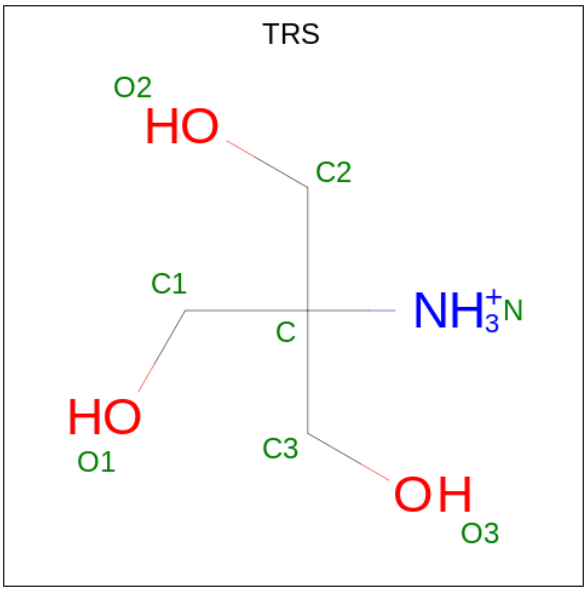
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
2	C	3	Total	C	O	0	0	0
			34	18	16			
2	D	3	Total	C	O	0	0	0
			34	18	16			

- Molecule 3 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).



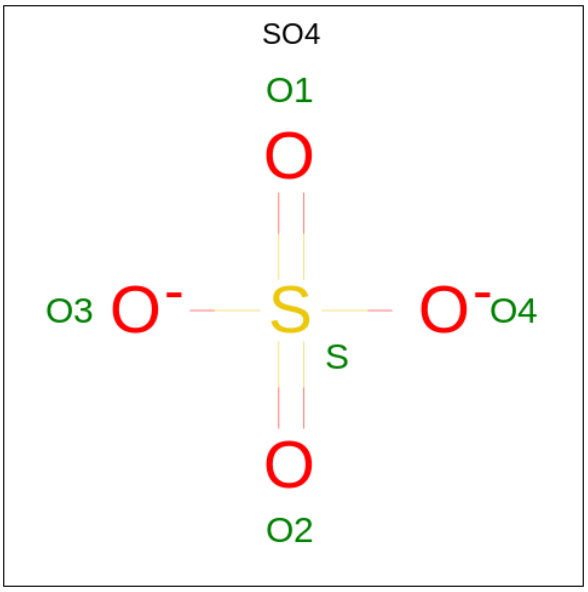
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C H O 7 2 3 2	0	0
3	A	1	Total C H O 7 2 3 2	0	0
3	A	1	Total C H O 7 2 3 2	0	0
3	A	1	Total C H O 7 2 3 2	0	0
3	A	1	Total C H O 7 2 3 2	0	0
3	B	1	Total C H O 7 2 3 2	0	0
3	B	1	Total C H O 7 2 3 2	0	0
3	B	1	Total C H O 7 2 3 2	0	0

- Molecule 4 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: TRS) (formula: $C_4H_{12}NO_3$).



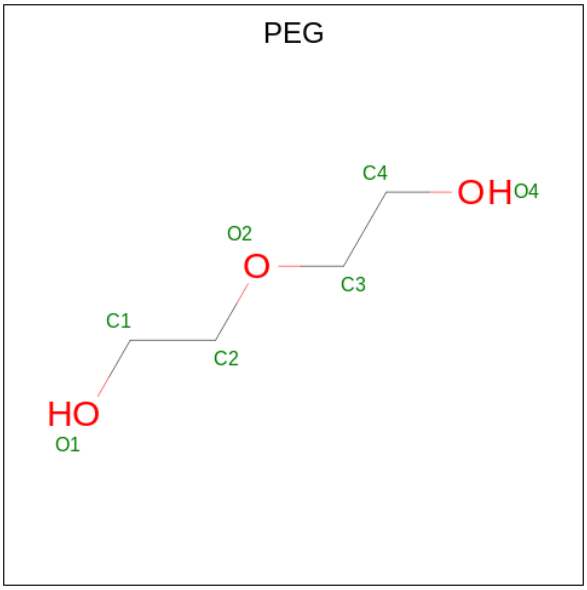
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	H	N	O	0	0
			20	4	12	1	3		
4	B	1	Total	C	H	N	O	0	0
			20	4	12	1	3		
4	B	1	Total	C	H	N	O	0	0
			20	4	12	1	3		

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O_4S) (labeled as "Ligand of Interest" by depositor).



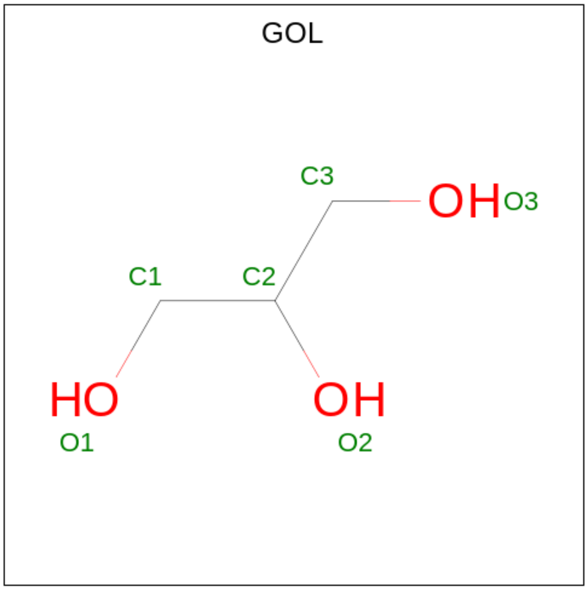
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total O S 5 4 1	0	0
5	B	1	Total O S 5 4 1	0	0

- Molecule 6 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C H O 17 4 10 3	0	0
6	A	1	Total C H O 17 4 10 3	0	0
6	A	1	Total C H O 17 4 10 3	0	0
6	B	1	Total C H O 17 4 10 3	0	0
6	B	1	Total C H O 17 4 10 3	0	0

- Molecule 7 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

Continued on next page...

Continued from previous page...

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

- Molecule 8 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	2	Total Cl 2 2	0	0
8	B	3	Total Cl 3 3	0	0

- Molecule 9 is water.

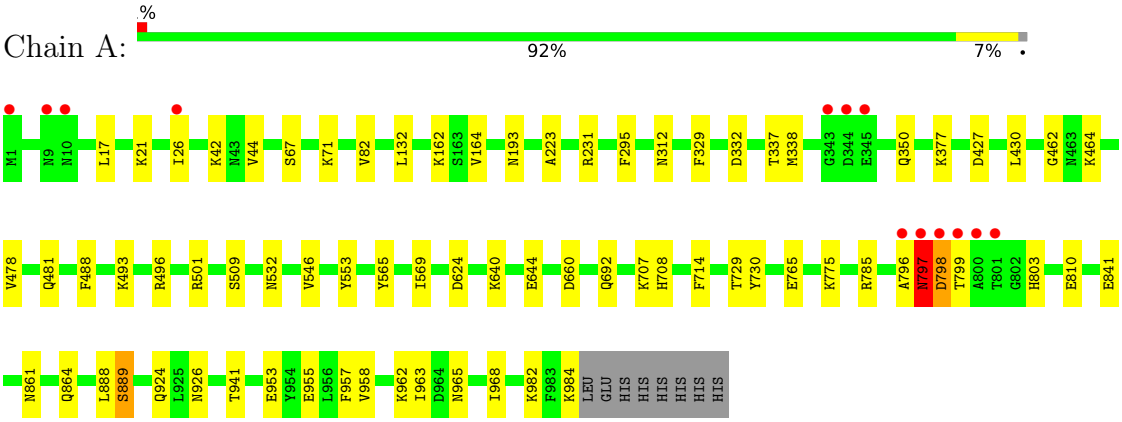
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	1052	Total O 1052 1052	0	0
9	B	1047	Total O 1047 1047	0	0

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.

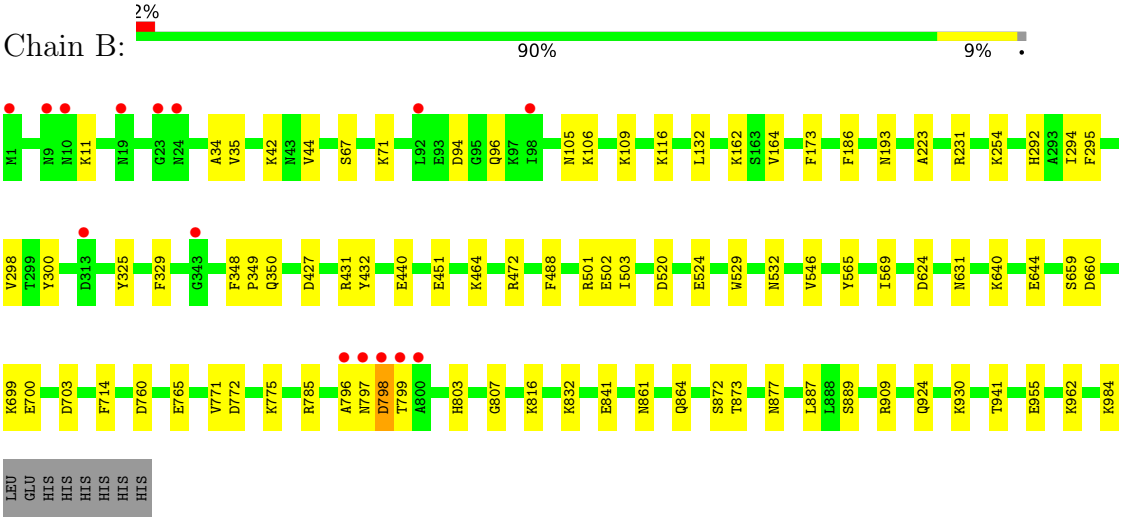
• Molecule 1: Cellodextrin phosphorylase

Chain A:



• Molecule 1: Cellodextrin phosphorylase

Chain B:



• Molecule 2: beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose

Chain C:

• Molecule 2: beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose

Chain D:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	83.21Å 88.78Å 88.76Å 98.58° 110.55° 110.56°	Depositor
Resolution (Å)	43.82 – 1.21 43.82 – 1.21	Depositor EDS
% Data completeness (in resolution range)	94.3 (43.82-1.21) 94.3 (43.82-1.21)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.40 (at 1.21Å)	Xtriage
Refinement program	PHENIX 1.21rc-5127	Depositor
R, R_{free}	0.143 , 0.164 0.143 , 0.164	Depositor DCC
R_{free} test set	30525 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	14.5	Xtriage
Anisotropy	0.065	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 36.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.447 for -h,-l,-k	Xtriage
F_o,F_c correlation	0.98	EDS
Total number of atoms	34642	wwPDB-VP
Average B, all atoms (Å^2)	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.12% 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: BGC, GOL, ACT, CL, TRS, SO4, PEG

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	0.39	0/8504	0.65	0/11497
1	B	0.39	0/8652	0.65	0/11691
All	All	0.39	0/17156	0.65	0/23188

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	B	0	2
All	All	0	3

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	797	ASN	Peptide
1	B	502	GLU	Mainchain

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	8326	7532	8159	72	0
1	B	8473	7606	8279	88	0
2	C	34	0	30	3	0
2	D	34	0	30	5	0
3	A	20	15	15	3	0
3	B	12	9	9	9	0
4	A	8	12	12	2	0
4	B	16	24	24	6	0
5	A	5	0	0	2	0
5	B	5	0	0	0	0
6	A	21	30	30	7	0
6	B	14	20	20	0	0
7	A	66	88	88	4	0
7	B	72	96	96	17	0
8	A	2	0	0	0	0
8	B	3	0	0	1	0
9	A	1052	0	0	11	0
9	B	1047	0	0	9	0
All	All	19210	15432	16792	161	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 5.

The worst 5 of 161 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:B:295[B]:PHE:O	3:B:1001:ACT:H3	1.71	0.90
1:B:771:VAL:HG12	1:B:775[A]:LYS:HE2	1.55	0.88
1:B:631:ASN:HD21	7:B:1018:GOL:H11	1.41	0.86
1:A:955:GLU:OE2	1:A:962:LYS:NZ	2.09	0.84
1:B:295[A]:PHE:O	3:B:1001:ACT:H3	1.80	0.81

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	1039/992 (105%)	1011 (97%)	26 (2%)	2 (0%)	44	17
1	B	1054/992 (106%)	1024 (97%)	29 (3%)	1 (0%)	48	17
All	All	2093/1984 (106%)	2035 (97%)	55 (3%)	3 (0%)	44	17

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	797	ASN
1	A	798	ASP
1	A	797	ASN

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	906/859 (106%)	899 (99%)	7 (1%)	79	53
1	B	922/859 (107%)	911 (99%)	11 (1%)	67	33
All	All	1828/1718 (106%)	1810 (99%)	18 (1%)	75	43

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	798	ASP
1	B	930	LYS
1	B	889[B]	SER
1	B	162	LYS
1	B	765[B]	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	219	ASN
1	A	926	ASN
1	B	219	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)

6 monosaccharides 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	BGC	C	1	2	12,12,12	0.53	0	17,17,17	1.31	2 (11%)
2	BGC	C	2	2	11,11,12	0.69	0	15,15,17	1.62	3 (20%)
2	BGC	C	3	2	11,11,12	0.68	0	15,15,17	1.20	2 (13%)
2	BGC	D	1	2	12,12,12	0.55	0	17,17,17	1.55	3 (17%)
2	BGC	D	2	2	11,11,12	0.68	0	15,15,17	1.27	2 (13%)
2	BGC	D	3	2	11,11,12	0.82	0	15,15,17	1.11	1 (6%)

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
2	BGC	C	1	2	-	1/2/22/22	0/1/1/1
2	BGC	C	2	2	-	1/2/19/22	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	BGC	C	3	2	-	2/2/19/22	0/1/1/1
2	BGC	D	1	2	-	2/2/22/22	0/1/1/1
2	BGC	D	2	2	-	2/2/19/22	0/1/1/1
2	BGC	D	3	2	-	2/2/19/22	0/1/1/1

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	2	BGC	C1-O5-C5	3.90	117.48	112.19
2	D	2	BGC	C1-O5-C5	3.50	116.93	112.19
2	D	1	BGC	O5-C5-C4	-3.47	103.39	109.69
2	D	1	BGC	O4-C4-C3	3.22	117.79	110.35
2	C	2	BGC	C3-C4-C5	-3.12	104.67	110.24

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

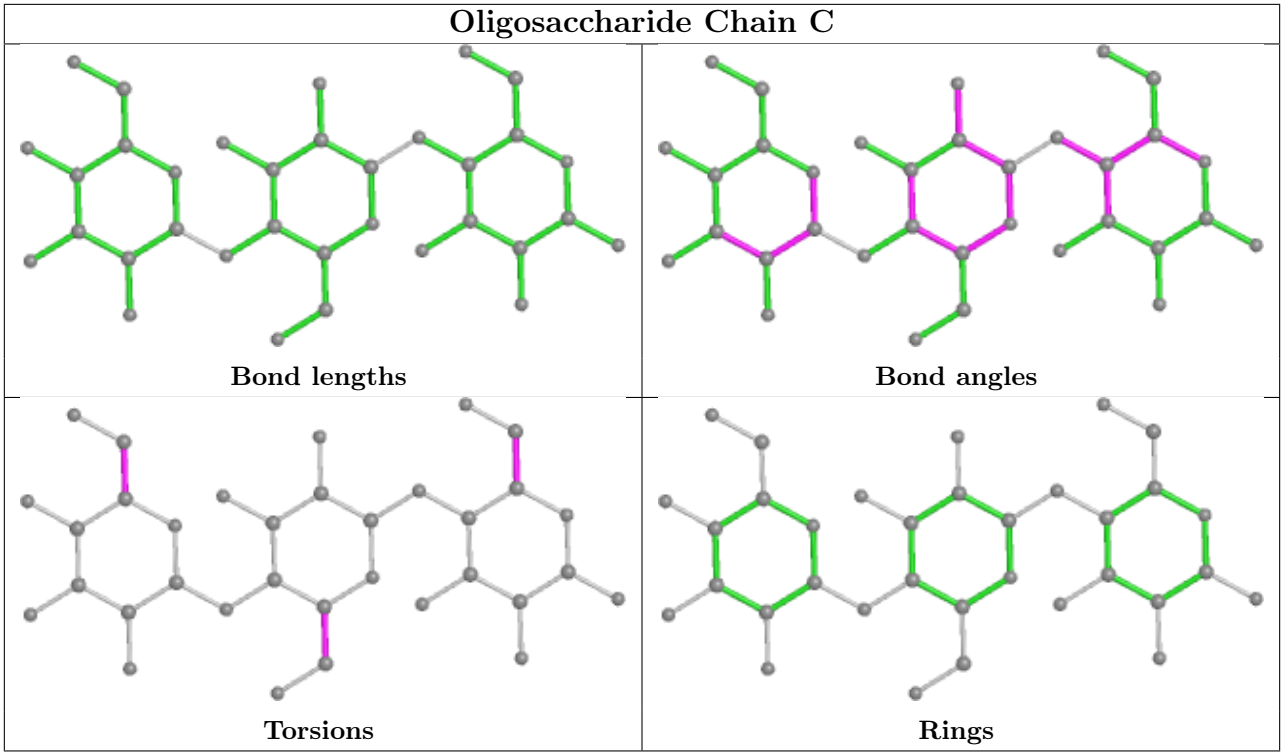
Mol	Chain	Res	Type	Atoms
2	D	3	BGC	O5-C5-C6-O6
2	D	1	BGC	O5-C5-C6-O6
2	D	1	BGC	C4-C5-C6-O6
2	D	3	BGC	C4-C5-C6-O6
2	C	3	BGC	O5-C5-C6-O6

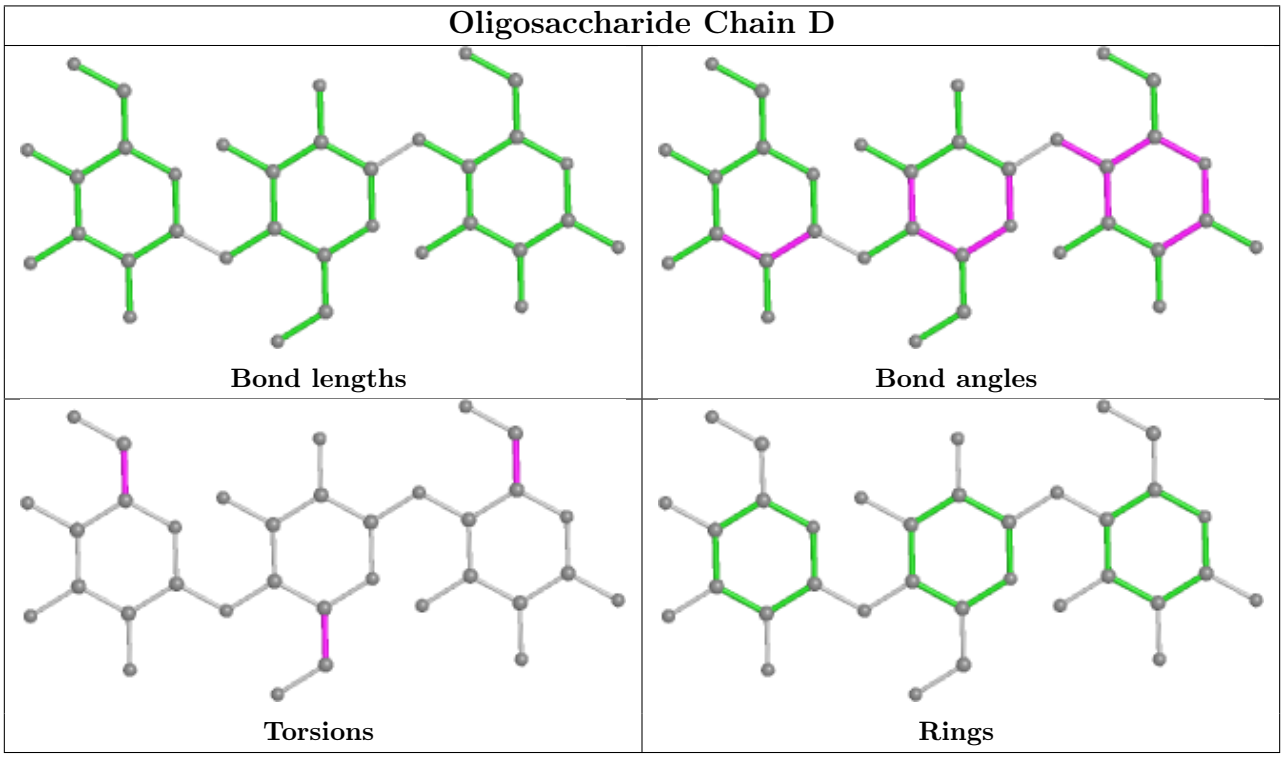
There are no ring outliers.

3 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	1	BGC	2	0
2	C	2	BGC	1	0
2	D	1	BGC	5	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.





5.6 Ligand geometry (i)

Of 46 ligands modelled in this entry, 5 are monoatomic - leaving 41 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	$\#\|Z\| > 2$
7	GOL	B	1012	-	5,5,5	0.35	0	5,5,5	0.75	0
5	SO4	A	1006	-	4,4,4	0.53	0	6,6,6	0.26	0
7	GOL	A	1016	-	5,5,5	0.38	0	5,5,5	0.46	0
7	GOL	B	1009	-	5,5,5	0.35	0	5,5,5	0.41	0
7	GOL	A	1012	-	5,5,5	0.36	0	5,5,5	0.40	0
7	GOL	A	1018	-	5,5,5	0.52	0	5,5,5	0.82	0
4	TRS	A	1002	-	7,7,7	0.33	0	9,9,9	0.82	0
6	PEG	A	1007	-	6,6,6	0.32	0	5,5,5	0.47	0
7	GOL	B	1017	-	5,5,5	0.40	0	5,5,5	0.55	0
7	GOL	B	1015	-	5,5,5	0.53	0	5,5,5	0.67	0
7	GOL	A	1019	-	5,5,5	0.35	0	5,5,5	0.40	0
3	ACT	A	1005	-	3,3,3	1.26	0	3,3,3	1.16	0
3	ACT	B	1001	-	3,3,3	1.22	0	3,3,3	0.18	0
7	GOL	B	1014	-	5,5,5	0.38	0	5,5,5	0.51	0
7	GOL	A	1014	-	5,5,5	0.31	0	5,5,5	0.59	0
7	GOL	B	1008	-	5,5,5	0.47	0	5,5,5	1.05	0
6	PEG	A	1009	-	6,6,6	0.25	0	5,5,5	0.14	0
7	GOL	A	1010	-	5,5,5	0.44	0	5,5,5	0.82	0
3	ACT	B	1002	-	3,3,3	1.26	0	3,3,3	1.28	0
3	ACT	A	1003	-	3,3,3	1.25	0	3,3,3	1.09	0
7	GOL	A	1013	-	5,5,5	0.31	0	5,5,5	0.96	0
6	PEG	B	1006	-	6,6,6	0.24	0	5,5,5	0.20	0
3	ACT	B	1004	-	3,3,3	1.27	0	3,3,3	1.20	0
7	GOL	A	1020	-	5,5,5	0.34	0	5,5,5	0.45	0
7	GOL	B	1016	-	5,5,5	0.39	0	5,5,5	0.59	0
3	ACT	A	1004	-	3,3,3	1.31	0	3,3,3	1.34	0
4	TRS	B	1003	-	7,7,7	0.60	0	9,9,9	0.90	0
7	GOL	B	1011	-	5,5,5	0.47	0	5,5,5	1.32	0
6	PEG	B	1007	-	6,6,6	0.23	0	5,5,5	0.30	0
7	GOL	B	1019	-	5,5,5	0.32	0	5,5,5	0.47	0
5	SO4	B	1005	-	4,4,4	0.55	0	6,6,6	0.04	0
7	GOL	A	1017	-	5,5,5	0.35	0	5,5,5	0.39	0
6	PEG	A	1008	-	6,6,6	0.22	0	5,5,5	0.23	0
4	TRS	B	1000	-	7,7,7	0.35	0	9,9,9	0.51	0
7	GOL	A	1011	-	5,5,5	0.35	0	5,5,5	0.42	0
7	GOL	B	1010	-	5,5,5	0.46	0	5,5,5	0.83	0
3	ACT	A	1001	-	3,3,3	1.36	0	3,3,3	1.10	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
7	GOL	B	1013	-	5,5,5	0.38	0	5,5,5	0.44	0
7	GOL	B	1018	-	5,5,5	0.38	0	5,5,5	0.46	0
7	GOL	A	1015	-	5,5,5	0.35	0	5,5,5	0.50	0
3	ACT	A	1000	-	3,3,3	1.27	0	3,3,3	1.30	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
7	GOL	B	1012	-	-	0/4/4/4	-
7	GOL	A	1016	-	-	4/4/4/4	-
7	GOL	B	1009	-	-	1/4/4/4	-
7	GOL	A	1012	-	-	4/4/4/4	-
7	GOL	A	1018	-	-	2/4/4/4	-
4	TRS	A	1002	-	-	4/9/9/9	-
6	PEG	A	1007	-	-	3/4/4/4	-
7	GOL	B	1017	-	-	4/4/4/4	-
7	GOL	B	1015	-	-	3/4/4/4	-
7	GOL	A	1019	-	-	4/4/4/4	-
7	GOL	B	1014	-	-	0/4/4/4	-
7	GOL	A	1014	-	-	0/4/4/4	-
7	GOL	B	1008	-	-	2/4/4/4	-
6	PEG	A	1009	-	-	2/4/4/4	-
7	GOL	A	1010	-	-	2/4/4/4	-
7	GOL	A	1013	-	-	2/4/4/4	-
6	PEG	B	1006	-	-	2/4/4/4	-
7	GOL	A	1020	-	-	3/4/4/4	-
7	GOL	B	1016	-	-	2/4/4/4	-
4	TRS	B	1003	-	-	3/9/9/9	-
7	GOL	B	1011	-	-	3/4/4/4	-
6	PEG	B	1007	-	-	3/4/4/4	-
7	GOL	B	1019	-	-	2/4/4/4	-
7	GOL	A	1017	-	-	2/4/4/4	-
6	PEG	A	1008	-	-	3/4/4/4	-
4	TRS	B	1000	-	-	9/9/9/9	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
7	GOL	A	1011	-	-	0/4/4/4	-
7	GOL	B	1010	-	-	2/4/4/4	-
7	GOL	B	1013	-	-	2/4/4/4	-
7	GOL	B	1018	-	-	2/4/4/4	-
7	GOL	A	1015	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 77 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	1002	TRS	N-C-C3-O3
4	B	1003	TRS	C2-C-C1-O1
4	B	1003	TRS	C3-C-C1-O1
7	A	1010	GOL	O1-C1-C2-C3
7	A	1016	GOL	C1-C2-C3-O3

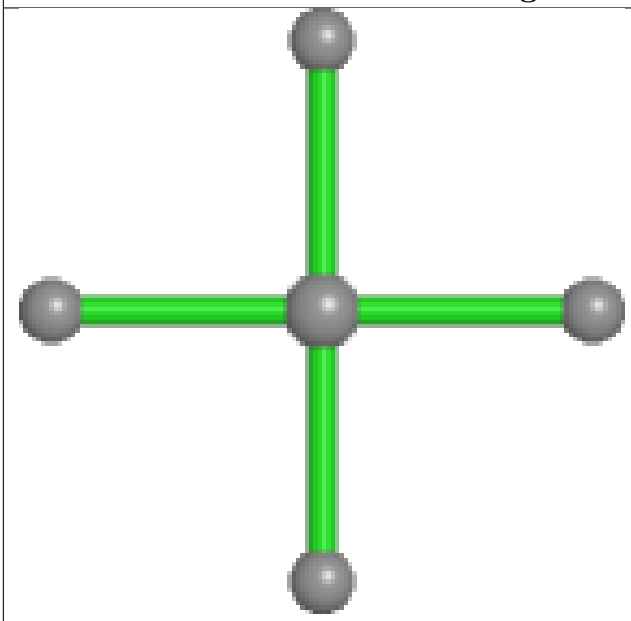
There are no ring outliers.

17 monomers are involved in 50 short contacts:

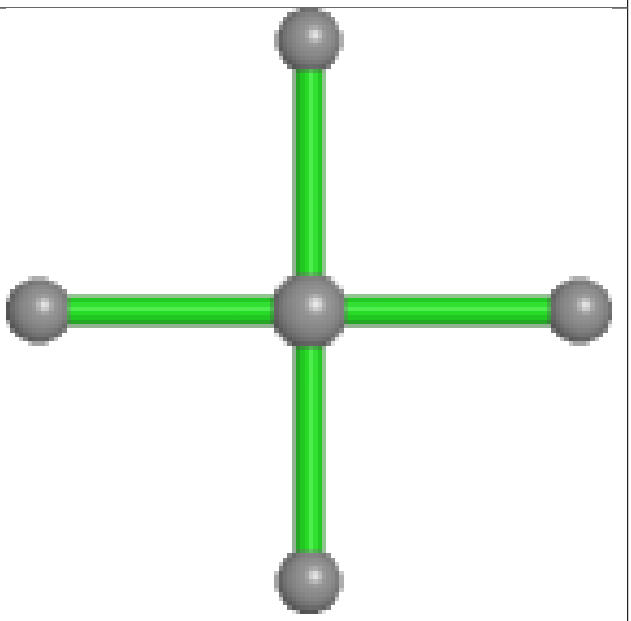
Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	B	1012	GOL	1	0
5	A	1006	SO4	2	0
7	A	1016	GOL	1	0
7	A	1018	GOL	1	0
4	A	1002	TRS	2	0
6	A	1007	PEG	6	0
7	B	1017	GOL	4	0
3	B	1001	ACT	9	0
7	B	1008	GOL	2	0
7	A	1010	GOL	2	0
7	B	1016	GOL	2	0
4	B	1003	TRS	6	0
7	B	1011	GOL	4	0
6	A	1008	PEG	1	0
3	A	1001	ACT	2	0
7	B	1018	GOL	4	0
3	A	1000	ACT	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

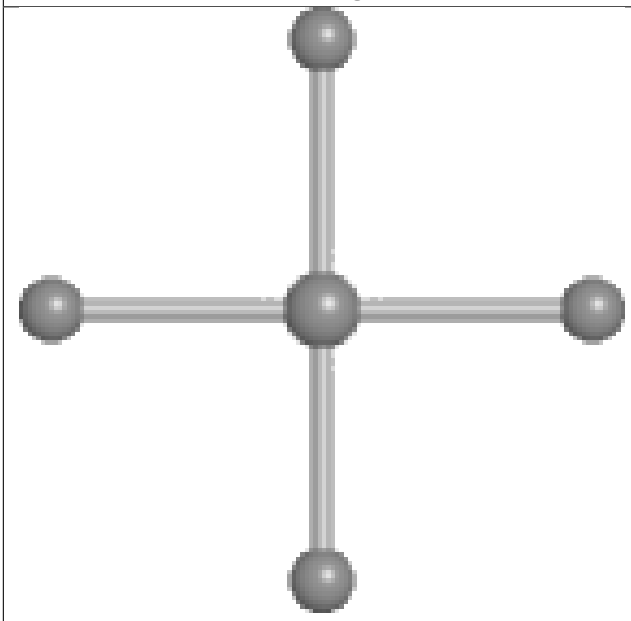
Ligand SO4 A 1006



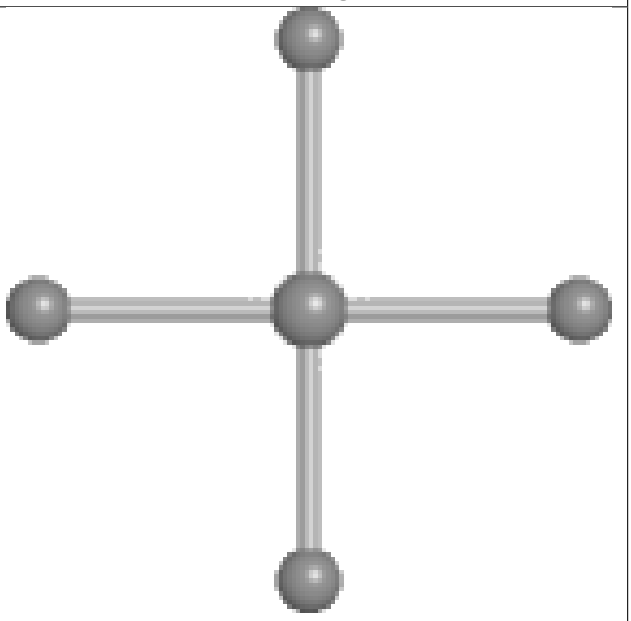
Bond lengths



Bond angles

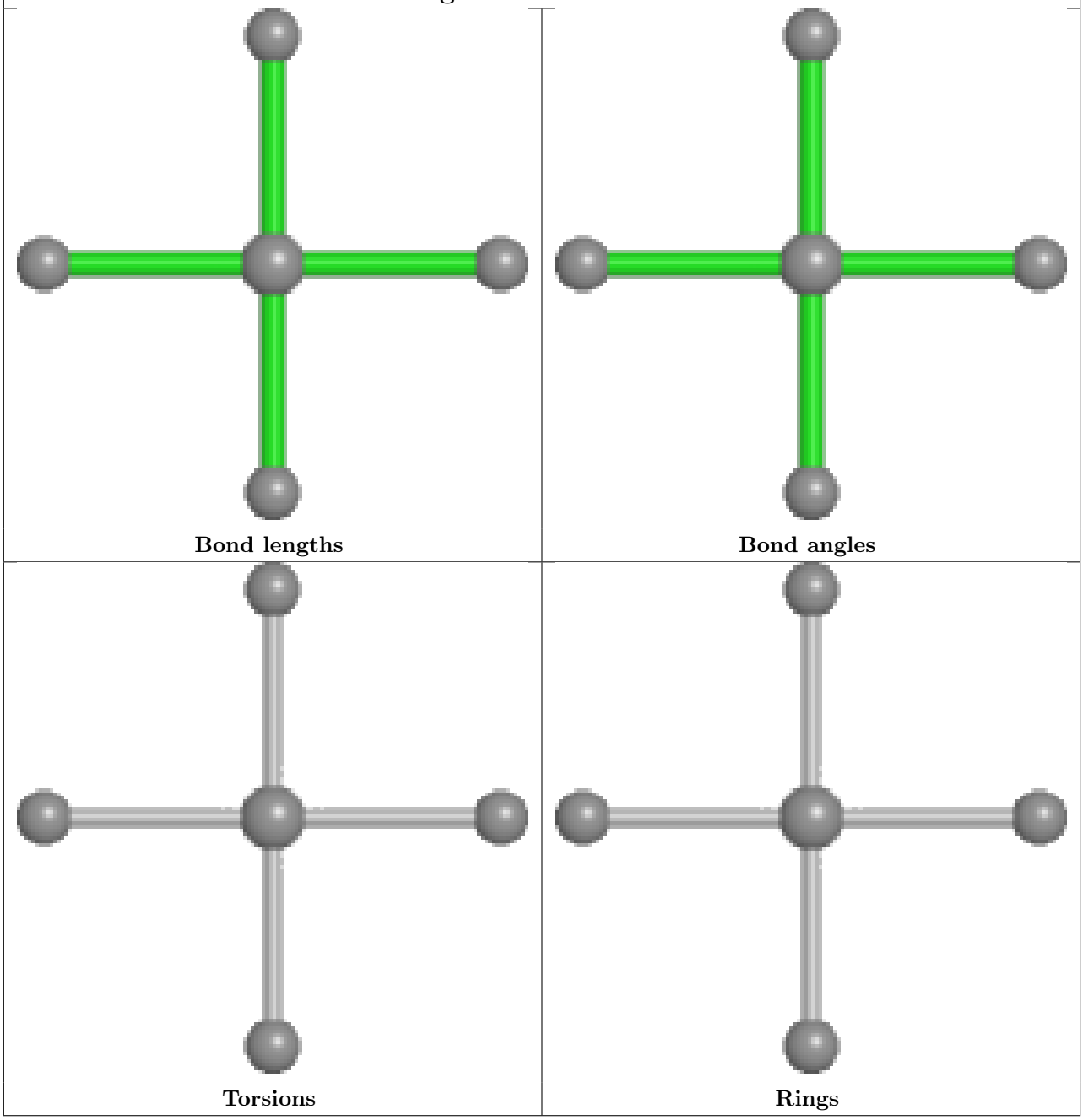


Torsions



Rings

Ligand SO4 B 1005



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	984/992 (99%)	-0.49	13 (1%)	74	74	5, 17, 38, 114	58 (5%)
1	B	984/992 (99%)	-0.44	15 (1%)	71	72	5, 17, 37, 116	73 (7%)
All	All	1968/1984 (99%)	-0.46	28 (1%)	73	73	5, 17, 38, 116	131 (6%)

The worst 5 of 28 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	800	ALA	4.3
1	B	98	ILE	4.2
1	A	10	ASN	3.4
1	B	10	ASN	3.4
1	A	799	THR	3.2

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

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
2	BGC	C	3	11/12	0.90	0.15	34,44,49,51	11
2	BGC	D	3	11/12	0.90	0.17	46,55,57,58	11
2	BGC	D	2	11/12	0.95	0.15	60,68,76,83	0
2	BGC	D	1	12/12	0.95	0.17	20,64,102,112	12
2	BGC	C	1	12/12	0.96	0.10	31,43,45,45	12

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	BGC	C	2	11/12	0.96	0.09	34,39,41,41	6

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	ACT	B	1002	4/4	0.91	0.17	58,69,97,99	0
3	ACT	A	1003	4/4	0.92	0.15	73,85,102,102	0
3	ACT	B	1004	4/4	0.92	0.16	51,61,78,80	0
4	TRS	A	1002	8/8	0.92	0.16	44,72,119,138	0
6	PEG	A	1009	7/7	0.92	0.12	64,77,91,93	0
7	GOL	B	1019	6/6	0.92	0.15	33,81,104,122	0
5	SO4	B	1005	5/5	0.93	0.17	42,50,112,128	0
6	PEG	B	1007	7/7	0.93	0.13	50,88,107,107	0
7	GOL	A	1019	6/6	0.93	0.15	60,78,87,105	0
6	PEG	A	1007	7/7	0.93	0.16	18,63,110,132	0
6	PEG	B	1006	7/7	0.94	0.10	43,57,75,75	0
3	ACT	A	1000	4/4	0.94	0.16	41,50,68,103	0
6	PEG	A	1008	7/7	0.94	0.12	51,83,104,104	0
7	GOL	A	1020	6/6	0.94	0.14	28,63,86,103	0
7	GOL	B	1010	6/6	0.94	0.13	33,54,101,101	0
3	ACT	A	1005	4/4	0.94	0.15	51,62,77,89	0
7	GOL	A	1017	6/6	0.95	0.11	67,81,93,111	0
4	TRS	B	1003	8/8	0.95	0.14	18,52,100,116	0
5	SO4	A	1006	5/5	0.95	0.16	38,41,42,44	5
7	GOL	A	1012	6/6	0.95	0.12	38,60,95,103	0
7	GOL	A	1014	6/6	0.95	0.12	26,66,83,85	0
4	TRS	B	1000	8/8	0.96	0.09	42,53,59,59	0
7	GOL	A	1016	6/6	0.96	0.13	21,30,95,95	0
7	GOL	B	1008	6/6	0.96	0.12	24,62,109,109	0
3	ACT	A	1004	4/4	0.96	0.08	31,37,63,81	0
7	GOL	B	1011	6/6	0.96	0.12	19,68,89,106	0
7	GOL	B	1012	6/6	0.96	0.12	38,59,92,110	0
7	GOL	B	1016	6/6	0.96	0.12	26,63,97,97	0
7	GOL	B	1017	6/6	0.96	0.09	35,60,89,103	0
7	GOL	B	1018	6/6	0.96	0.09	31,52,74,89	0
7	GOL	A	1018	6/6	0.96	0.09	29,54,66,71	0
7	GOL	B	1014	6/6	0.97	0.06	25,33,39,42	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
7	GOL	A	1011	6/6	0.97	0.09	21,28,83,83	0
3	ACT	B	1001	4/4	0.97	0.14	17,21,22,121	0
7	GOL	A	1010	6/6	0.97	0.11	31,68,125,125	0
7	GOL	B	1013	6/6	0.97	0.09	27,51,86,103	0
7	GOL	B	1015	6/6	0.98	0.09	21,60,120,120	0
3	ACT	A	1001	4/4	0.98	0.06	29,34,53,55	0
7	GOL	A	1015	6/6	0.98	0.07	25,44,61,61	0
7	GOL	B	1009	6/6	0.98	0.07	25,41,72,72	0
7	GOL	A	1013	6/6	0.98	0.09	17,35,71,71	0
8	CL	B	1020	1/1	0.99	0.14	46,46,46,46	0
8	CL	A	1022	1/1	1.00	0.02	12,12,12,12	0
8	CL	A	1021	1/1	1.00	0.03	13,13,13,13	0
8	CL	B	1021	1/1	1.00	0.03	13,13,13,13	0
8	CL	B	1022	1/1	1.00	0.02	12,12,12,12	0

6.5 Other polymers (i)

There are no such residues in this entry.