Preliminary Full wwPDB X-ray Structure Validation Report (i)

Feb 22, 2021 – 10:21 am GMT

Deposition ID : D_1292113439

This is a Preliminary Full wwPDB X-ray Structure Validation Report.

This report is produced by the wwPDB Deposition System during initial deposition but before annotation of the structure.

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:

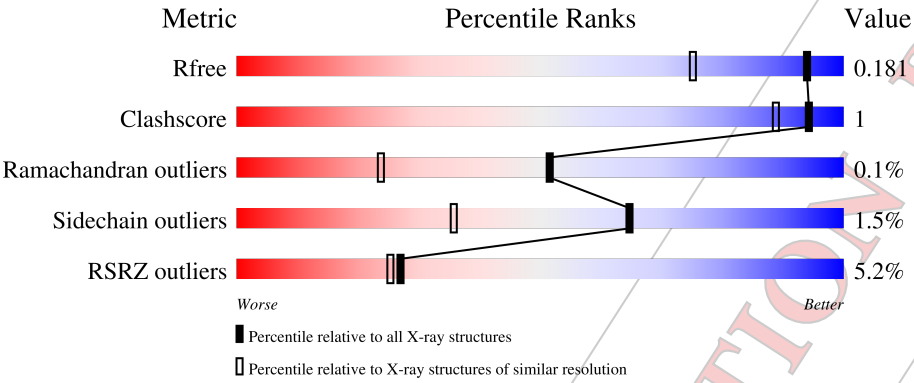
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.17.1.dev1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.17.1.dev1

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 1.42 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	2579 (1.44-1.40)
Clashscore	141614	2696 (1.44-1.40)
Ramachandran outliers	138981	2632 (1.44-1.40)
Sidechain outliers	138945	2631 (1.44-1.40)
RSRZ outliers	127900	2528 (1.44-1.40)

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	751	4% 93%
1	B	751	6% 92%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 13069 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 Alginate lyase, family PL17.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	727	Total	C	N	O	S	0	5	0
			5837	3721	984	1118	14			
1	B	727	Total	C	N	O	S	0	9	0
			5863	3736	988	1125	14			

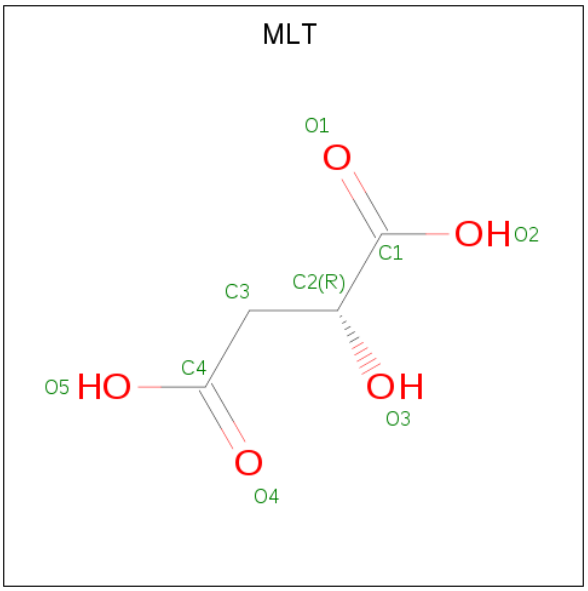
- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

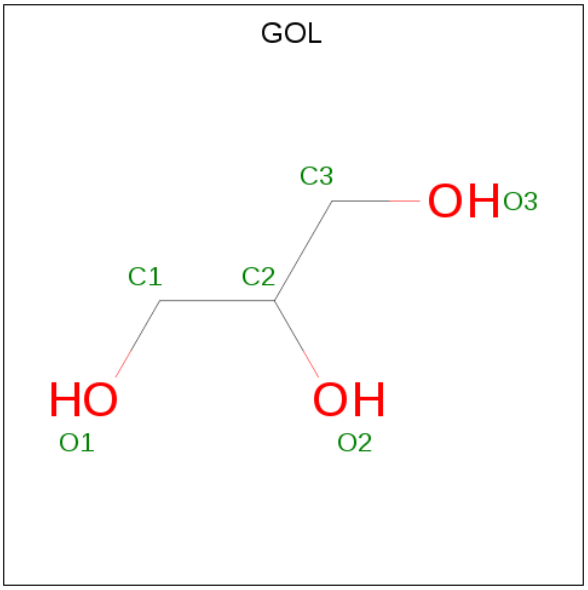
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Mg	0	0
			1	1		
3	A	1	Total	Mg	0	0
			1	1		
3	B	1	Total	Mg	0	0
			1	1		
3	B	1	Total	Mg	0	0
			1	1		

- Molecule 4 is D-MALATE (three-letter code: MLT) (formula: $C_4H_6O_5$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 9 4 5	0	0
4	B	1	Total C O 9 4 5	0	0

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
5	B	1	6	3	3	0	0

- Molecule 6 is water.

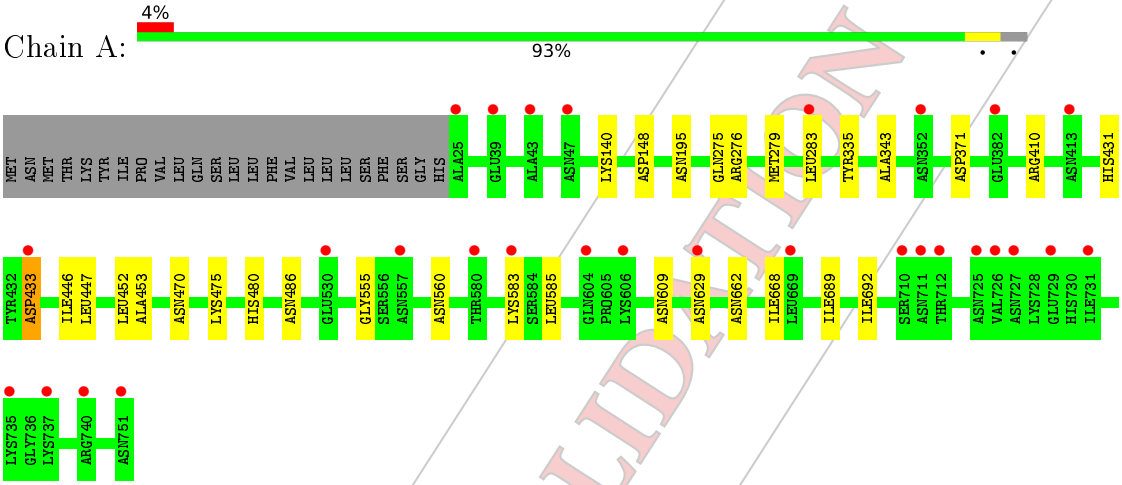
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
6	Z	1326	1326	1326	0	0

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.

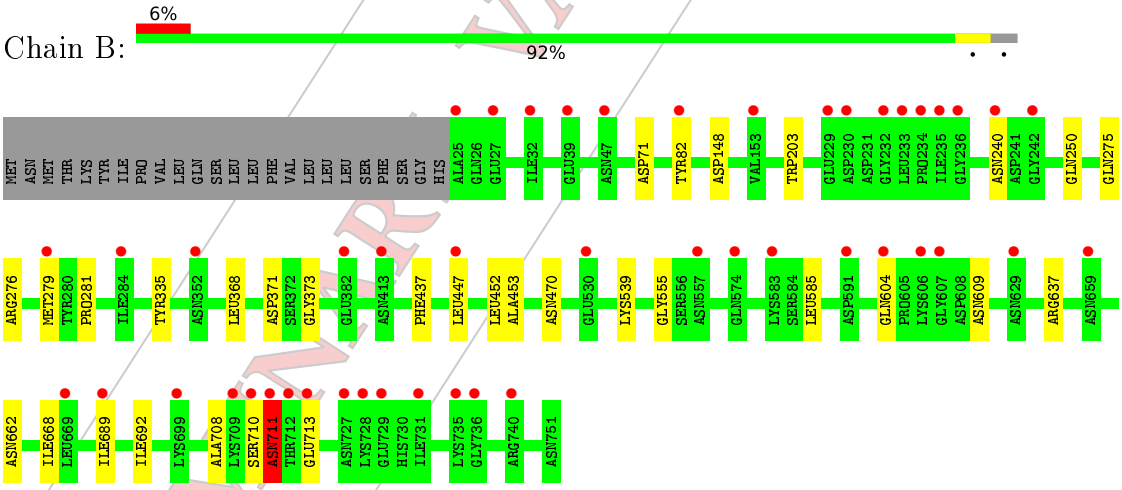
- Molecule 1: Alginate lyase, family PL17

Chain A:



- Molecule 1: Alginate lyase, family PL17

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, α, β, γ	163.39Å 163.39Å 166.73Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	43.70 – 1.42 43.70 – 1.42	Depositor EDS
% Data completeness (in resolution range)	99.8 (43.70-1.42) 99.8 (43.70-1.42)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.10	Depositor
$< I/\sigma(I) >$ [1]	2.04 (at 1.42Å)	Xtriage
Refinement program	REFMAC 5.8.0131	Depositor
R, R_{free}	0.161 , 0.170 0.171 , 0.181	Depositor DCC
R_{free} test set	23668 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	14.9	Xtriage
Anisotropy	0.009	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.40 , 41.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.015 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	13069	wwPDB-VP
Average B, all atoms (Å^2)	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.59% 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: MG, GOL, CA, MLT

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	0.35	0/5987	0.61	1/8099 (0.0%)
1	B	0.35	0/6032	0.61	2/8160 (0.0%)
All	All	0.35	0/12019	0.61	3/16259 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	637	ARG	NE-CZ-NH2	-5.41	117.59	120.30
1	B	637	ARG	NE-CZ-NH1	5.23	122.91	120.30
1	A	410	ARG	NE-CZ-NH1	5.12	122.86	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	5837	0	5667	15	0
1	B	5863	0	5697	16	0
2	A	1	0	0	0	0
2	B	2	0	0	0	0
3	A	2	0	0	0	0
3	B	2	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	9	0	4	0	0
4	B	9	0	4	0	0
5	A	12	0	16	0	0
5	B	6	0	8	0	0
6	Z	1326	0	0	2	0
All	All	13069	0	11396	30	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 1.

All (30) 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:585:LEU:HD12	1:B:585:LEU:HD12	1.53	0.91
1:B:453:ALA:H	1:B:470:ASN:HD22	1.27	0.83
1:A:453:ALA:H	1:A:470:ASN:HD22	1.24	0.80
1:B:240:ASN:ND2	6:Z:1323:HOH:O	2.18	0.72
1:B:609:ASN:HD21	1:B:692:ILE:H	1.40	0.70
1:A:609:ASN:HD21	1:A:692:ILE:H	1.41	0.67
1:B:275:GLN:O	1:B:279:MET:HG2	2.05	0.56
1:A:275:GLN:O	1:A:279:MET:HG2	2.06	0.55
1:A:431:HIS:HB3	1:A:433:ASP:OD2	2.06	0.54
1:B:668:ILE:HG21	1:B:689:ILE:HD11	1.90	0.54
1:A:433:ASP:OD2	1:A:480:HIS:CE1	2.62	0.52
1:B:708:ALA:HB3	1:B:713:GLU:HB2	1.91	0.52
1:B:710:SER:O	1:B:711:ASN:O	2.28	0.51
1:B:71:ASP:OD2	1:B:82:TYR:OH	2.16	0.50
1:A:486:ASN:HD21	1:A:560:ASN:HD22	1.61	0.48
1:A:668:ILE:HG21	1:A:689:ILE:HD11	1.96	0.48
1:B:710:SER:O	1:B:713:GLU:HG2	2.14	0.48
1:A:283[A]:LEU:HD13	1:A:343:ALA:HA	1.95	0.47
1:A:555:GLY:HA2	1:A:662:ASN:HD22	1.78	0.47
1:A:433:ASP:OD2	1:A:480:HIS:NE2	2.49	0.46
1:B:437:PHE:CZ	1:B:447:LEU:HD23	2.53	0.43
1:A:453:ALA:H	1:A:470:ASN:ND2	2.05	0.43
1:A:283[B]:LEU:HD22	1:A:343:ALA:HA	2.01	0.43
1:A:446:ILE:O	1:A:447:LEU:HD12	2.19	0.42
1:B:668:ILE:CG2	1:B:689:ILE:HD11	2.50	0.42
1:B:555:GLY:HA2	1:B:662:ASN:HD22	1.85	0.42
1:B:203:TRP:CZ2	1:B:281:PRO:HD3	2.55	0.41
1:A:431:HIS:CD2	1:A:475:LYS:HG2	2.55	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:539[B]:LYS:NZ	6:Z:811:HOH:O	2.36	0.41
1:B:368:LEU:O	1:B:373:GLY:HA3	2.20	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	729/751 (97%)	706 (97%)	23 (3%)	0	100	100
1	B	735/751 (98%)	710 (97%)	24 (3%)	1 (0%)	51	24
All	All	1464/1502 (98%)	1416 (97%)	47 (3%)	1 (0%)	51	24

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	711	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	622/643 (97%)	612 (98%)	10 (2%)	62	32
1	B	629/643 (98%)	621 (99%)	8 (1%)	69	41
All	All	1251/1286 (97%)	1233 (99%)	18 (1%)	65	38

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	140	LYS
1	A	148	ASP
1	A	195	ASN
1	A	276	ARG
1	A	335	TYR
1	A	371	ASP
1	A	433	ASP
1	A	452	LEU
1	A	583	LYS
1	A	629	ASN
1	B	148	ASP
1	B	250	GLN
1	B	276	ARG
1	B	335	TYR
1	B	371	ASP
1	B	452	LEU
1	B	604	GLN
1	B	711	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (29) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	28	HIS
1	A	147	ASN
1	A	150	ASN
1	A	177	ASN
1	A	199	ASN
1	A	205	ASN
1	A	301	HIS
1	A	470	ASN
1	A	476	GLN
1	A	560	ASN
1	A	574	GLN
1	A	609	ASN
1	A	634	HIS
1	A	641	ASN
1	A	662	ASN
1	A	686	ASN
1	B	147	ASN
1	B	150	ASN
1	B	205	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	275	GLN
1	B	301	HIS
1	B	353	HIS
1	B	470	ASN
1	B	476	GLN
1	B	560	ASN
1	B	609	ASN
1	B	634	HIS
1	B	662	ASN
1	B	686	ASN

5.3.3 RNA

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

Of 12 ligands modelled in this entry, 7 are monoatomic - leaving 5 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	$\#\|Z\| > 2$
4	MLT	B	801	-	2,8,8	0.91	0	3,10,10	1.72	1 (33%)
5	GOL	A	809	-	5,5,5	0.33	0	5,5,5	0.30	0
5	GOL	B	803	-	5,5,5	0.64	0	5,5,5	0.51	0
4	MLT	A	808	-	2,8,8	0.37	0	3,10,10	0.83	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
5	GOL	A	900	-	5,5,5	0.32	0	5,5,5	0.32	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	MLT	B	801	-	-	0/2/8/8	-
5	GOL	A	809	-	-	0/4/4/4	-
5	GOL	B	803	-	-	0/4/4/4	-
4	MLT	A	808	-	-	0/2/8/8	-
5	GOL	A	900	-	-	0/4/4/4	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	B	801	MLT	C3-C2-C1	2.42	114.18	111.10

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	727/751 (96%)	0.37	29 (3%)	38	38	12, 16, 27, 37	0
1	B	727/751 (96%)	0.48	47 (6%)	18	16	12, 17, 29, 40	13 (1%)
All	All	1454/1502 (96%)	0.43	76 (5%)	27	25	12, 17, 28, 40	13 (0%)

All (76) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	235	ILE	10.2
1	B	711	ASN	9.7
1	A	711	ASN	9.4
1	B	712	THR	7.9
1	B	25	ALA	5.9
1	A	710	ALA	5.9
1	A	25	ALA	5.7
1	A	606	LYS	5.0
1	A	629	ASN	4.9
1	B	557	ASN	4.8
1	B	229	GLU	4.8
1	A	712	THR	4.8
1	A	583	LYS	4.8
1	B	230	ASP	4.7
1	B	606	LYS	4.6
1	B	629	ASN	4.6
1	B	82	TYR	4.6
1	B	413	ASN	4.5
1	B	242	GLY	4.4
1	A	604	GLN	4.4
1	A	557	ASN	4.3
1	B	731	ILE	4.2
1	B	736	GLY	3.9
1	A	413	ASN	3.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	47	ASN	3.8
1	B	604	GLN	3.8
1	B	710	SER	3.8
1	A	47	ASN	3.7
1	B	735	LYS	3.6
1	A	740	ARG	3.6
1	B	740	ARG	3.5
1	A	731	ILE	3.5
1	A	727	ASN	3.3
1	B	713	GLU	3.2
1	B	352	ASN	3.1
1	A	726	VAL	2.9
1	B	240	ASN	2.9
1	B	729	GLU	2.8
1	B	727	ASN	2.8
1	B	574	GLN	2.7
1	A	433	ASP	2.7
1	B	234	PRO	2.7
1	A	751[A]	ASN	2.7
1	B	583	LYS	2.7
1	B	236	GLY	2.6
1	B	382	GLU	2.5
1	A	352	ASN	2.5
1	A	729	GLU	2.5
1	A	725	ASN	2.5
1	B	709	LYS	2.5
1	A	737	LYS	2.5
1	B	39	GLU	2.4
1	B	279	MET	2.4
1	A	382	GLU	2.3
1	B	591	ASP	2.3
1	B	530	GLU	2.2
1	A	580	THR	2.2
1	B	153	VAL	2.2
1	A	669	LEU	2.2
1	B	232	GLY	2.2
1	B	699	LYS	2.2
1	B	659	ASN	2.1
1	B	233	LEU	2.1
1	B	447	LEU	2.1
1	B	669	LEU	2.1
1	A	39	GLU	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	530	GLU	2.1
1	A	735	LYS	2.1
1	A	283[A]	LEU	2.1
1	A	43	ALA	2.1
1	B	607	GLY	2.1
1	B	728	LYS	2.1
1	B	32	ILE	2.0
1	B	284	ILE	2.0
1	B	689	ILE	2.0
1	B	27	GLU	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (i)

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 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
5	GOL	B	803	6/?	0.88	0.18	20,20,20,20	0
3	MG	A	807	1/?	0.92	0.13	34,34,34,34	0
5	GOL	A	900	6/?	0.93	0.10	19,21,22,22	0
4	MLT	B	801	9/?	0.93	0.12	18,19,21,22	0
4	MLT	A	808	9/?	0.96	0.09	16,16,18,18	0
5	GOL	A	809	6/?	0.97	0.08	16,16,17,17	0
3	MG	B	805	1/?	0.99	0.15	11,11,11,11	0
3	MG	A	805	1/?	0.99	0.07	22,22,22,22	0
2	CA	B	807	1/?	0.99	0.05	17,17,17,17	0
3	MG	B	806	1/?	1.00	0.15	12,12,12,12	0
2	CA	A	802	1/?	1.00	0.17	8,8,8,8	0
2	CA	B	802	1/?	1.00	0.17	8,8,8,8	0

6.5 Other polymers (i)

There are no such residues in this entry.