



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 4, 2024 – 10:59 PM EST

PDB ID : 8GQF
Title : Crystal structure of Thiolase
Authors : Hong, J.; Son, H.F.; Kim, K.J.
Deposited on : 2022-08-30
Resolution : 1.80 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 ⓘ 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 ⓘ](#)) 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 : 2.36
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) : 2.36

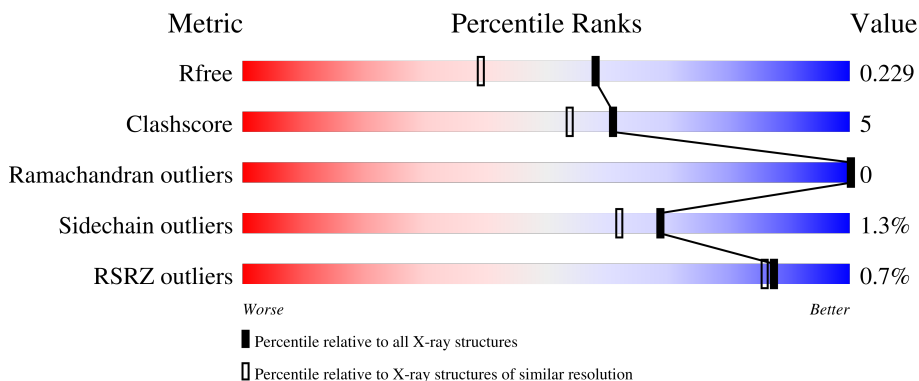
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 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 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 $\leq 5\%$. 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	407	
1	B	407	
1	C	407	
1	D	407	

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 11805 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 Thiolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	379	2743	1713	492	519	19	0	0	0
1	B	379	2743	1713	492	519	19	0	0	0
1	C	377	2729	1705	490	516	18	0	0	0
1	D	378	2735	1708	491	518	18	0	0	0

There are 112 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	initiating methionine	UNP A0A0F7QQ39
A	-18	GLY	-	expression tag	UNP A0A0F7QQ39
A	-17	SER	-	expression tag	UNP A0A0F7QQ39
A	-16	SER	-	expression tag	UNP A0A0F7QQ39
A	-15	HIS	-	expression tag	UNP A0A0F7QQ39
A	-14	HIS	-	expression tag	UNP A0A0F7QQ39
A	-13	HIS	-	expression tag	UNP A0A0F7QQ39
A	-12	HIS	-	expression tag	UNP A0A0F7QQ39
A	-11	HIS	-	expression tag	UNP A0A0F7QQ39
A	-10	HIS	-	expression tag	UNP A0A0F7QQ39
A	-9	SER	-	expression tag	UNP A0A0F7QQ39
A	-8	SER	-	expression tag	UNP A0A0F7QQ39
A	-7	GLY	-	expression tag	UNP A0A0F7QQ39
A	-6	LEU	-	expression tag	UNP A0A0F7QQ39
A	-5	VAL	-	expression tag	UNP A0A0F7QQ39
A	-4	PRO	-	expression tag	UNP A0A0F7QQ39
A	-3	ARG	-	expression tag	UNP A0A0F7QQ39
A	-2	GLY	-	expression tag	UNP A0A0F7QQ39
A	-1	SER	-	expression tag	UNP A0A0F7QQ39
A	0	HIS	-	expression tag	UNP A0A0F7QQ39
A	380	LEU	-	expression tag	UNP A0A0F7QQ39

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	381	GLU	-	expression tag	UNP A0A0F7QQ39
A	382	HIS	-	expression tag	UNP A0A0F7QQ39
A	383	HIS	-	expression tag	UNP A0A0F7QQ39
A	384	HIS	-	expression tag	UNP A0A0F7QQ39
A	385	HIS	-	expression tag	UNP A0A0F7QQ39
A	386	HIS	-	expression tag	UNP A0A0F7QQ39
A	387	HIS	-	expression tag	UNP A0A0F7QQ39
B	-19	MET	-	initiating methionine	UNP A0A0F7QQ39
B	-18	GLY	-	expression tag	UNP A0A0F7QQ39
B	-17	SER	-	expression tag	UNP A0A0F7QQ39
B	-16	SER	-	expression tag	UNP A0A0F7QQ39
B	-15	HIS	-	expression tag	UNP A0A0F7QQ39
B	-14	HIS	-	expression tag	UNP A0A0F7QQ39
B	-13	HIS	-	expression tag	UNP A0A0F7QQ39
B	-12	HIS	-	expression tag	UNP A0A0F7QQ39
B	-11	HIS	-	expression tag	UNP A0A0F7QQ39
B	-10	HIS	-	expression tag	UNP A0A0F7QQ39
B	-9	SER	-	expression tag	UNP A0A0F7QQ39
B	-8	SER	-	expression tag	UNP A0A0F7QQ39
B	-7	GLY	-	expression tag	UNP A0A0F7QQ39
B	-6	LEU	-	expression tag	UNP A0A0F7QQ39
B	-5	VAL	-	expression tag	UNP A0A0F7QQ39
B	-4	PRO	-	expression tag	UNP A0A0F7QQ39
B	-3	ARG	-	expression tag	UNP A0A0F7QQ39
B	-2	GLY	-	expression tag	UNP A0A0F7QQ39
B	-1	SER	-	expression tag	UNP A0A0F7QQ39
B	0	HIS	-	expression tag	UNP A0A0F7QQ39
B	380	LEU	-	expression tag	UNP A0A0F7QQ39
B	381	GLU	-	expression tag	UNP A0A0F7QQ39
B	382	HIS	-	expression tag	UNP A0A0F7QQ39
B	383	HIS	-	expression tag	UNP A0A0F7QQ39
B	384	HIS	-	expression tag	UNP A0A0F7QQ39
B	385	HIS	-	expression tag	UNP A0A0F7QQ39
B	386	HIS	-	expression tag	UNP A0A0F7QQ39
B	387	HIS	-	expression tag	UNP A0A0F7QQ39
C	-19	MET	-	initiating methionine	UNP A0A0F7QQ39
C	-18	GLY	-	expression tag	UNP A0A0F7QQ39
C	-17	SER	-	expression tag	UNP A0A0F7QQ39
C	-16	SER	-	expression tag	UNP A0A0F7QQ39
C	-15	HIS	-	expression tag	UNP A0A0F7QQ39
C	-14	HIS	-	expression tag	UNP A0A0F7QQ39
C	-13	HIS	-	expression tag	UNP A0A0F7QQ39

Continued on next page...

Continued from previous page...

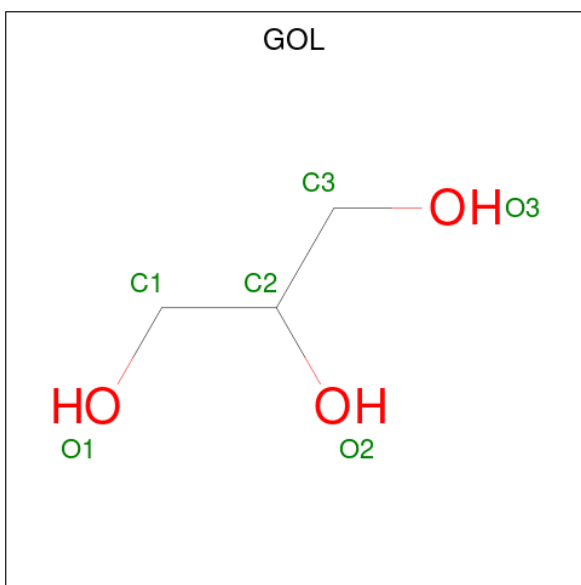
Chain	Residue	Modelled	Actual	Comment	Reference
C	-12	HIS	-	expression tag	UNP A0A0F7QQ39
C	-11	HIS	-	expression tag	UNP A0A0F7QQ39
C	-10	HIS	-	expression tag	UNP A0A0F7QQ39
C	-9	SER	-	expression tag	UNP A0A0F7QQ39
C	-8	SER	-	expression tag	UNP A0A0F7QQ39
C	-7	GLY	-	expression tag	UNP A0A0F7QQ39
C	-6	LEU	-	expression tag	UNP A0A0F7QQ39
C	-5	VAL	-	expression tag	UNP A0A0F7QQ39
C	-4	PRO	-	expression tag	UNP A0A0F7QQ39
C	-3	ARG	-	expression tag	UNP A0A0F7QQ39
C	-2	GLY	-	expression tag	UNP A0A0F7QQ39
C	-1	SER	-	expression tag	UNP A0A0F7QQ39
C	0	HIS	-	expression tag	UNP A0A0F7QQ39
C	380	LEU	-	expression tag	UNP A0A0F7QQ39
C	381	GLU	-	expression tag	UNP A0A0F7QQ39
C	382	HIS	-	expression tag	UNP A0A0F7QQ39
C	383	HIS	-	expression tag	UNP A0A0F7QQ39
C	384	HIS	-	expression tag	UNP A0A0F7QQ39
C	385	HIS	-	expression tag	UNP A0A0F7QQ39
C	386	HIS	-	expression tag	UNP A0A0F7QQ39
C	387	HIS	-	expression tag	UNP A0A0F7QQ39
D	-19	MET	-	initiating methionine	UNP A0A0F7QQ39
D	-18	GLY	-	expression tag	UNP A0A0F7QQ39
D	-17	SER	-	expression tag	UNP A0A0F7QQ39
D	-16	SER	-	expression tag	UNP A0A0F7QQ39
D	-15	HIS	-	expression tag	UNP A0A0F7QQ39
D	-14	HIS	-	expression tag	UNP A0A0F7QQ39
D	-13	HIS	-	expression tag	UNP A0A0F7QQ39
D	-12	HIS	-	expression tag	UNP A0A0F7QQ39
D	-11	HIS	-	expression tag	UNP A0A0F7QQ39
D	-10	HIS	-	expression tag	UNP A0A0F7QQ39
D	-9	SER	-	expression tag	UNP A0A0F7QQ39
D	-8	SER	-	expression tag	UNP A0A0F7QQ39
D	-7	GLY	-	expression tag	UNP A0A0F7QQ39
D	-6	LEU	-	expression tag	UNP A0A0F7QQ39
D	-5	VAL	-	expression tag	UNP A0A0F7QQ39
D	-4	PRO	-	expression tag	UNP A0A0F7QQ39
D	-3	ARG	-	expression tag	UNP A0A0F7QQ39
D	-2	GLY	-	expression tag	UNP A0A0F7QQ39
D	-1	SER	-	expression tag	UNP A0A0F7QQ39
D	0	HIS	-	expression tag	UNP A0A0F7QQ39
D	380	LEU	-	expression tag	UNP A0A0F7QQ39

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	381	GLU	-	expression tag	UNP A0A0F7QQ39
D	382	HIS	-	expression tag	UNP A0A0F7QQ39
D	383	HIS	-	expression tag	UNP A0A0F7QQ39
D	384	HIS	-	expression tag	UNP A0A0F7QQ39
D	385	HIS	-	expression tag	UNP A0A0F7QQ39
D	386	HIS	-	expression tag	UNP A0A0F7QQ39
D	387	HIS	-	expression tag	UNP A0A0F7QQ39

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



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

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	229	Total O 229 229	0	0
3	B	236	Total O 236 236	0	0
3	C	183	Total O 183 183	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	D	189	Total 189	O 189	0	0



4 Data and refinement statistics i

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	50.56Å 82.27Å 97.74Å 90.01° 75.00° 89.95°	Depositor
Resolution (Å)	31.49 – 1.80 31.47 – 1.80	Depositor EDS
% Data completeness (in resolution range)	95.8 (31.49-1.80) 95.8 (31.47-1.80)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.29 (at 1.80Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.181 , 0.220 0.194 , 0.229	Depositor DCC
R_{free} test set	6597 reflections (4.89%)	wwPDB-VP
Wilson B-factor (Å ²)	19.2	Xtriage
Anisotropy	0.062	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 33.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.021 for h,-k,h-l 0.438 for -h,k,-l 0.019 for -h,-k,-h+l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	11805	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 5.82% 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL

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	0.80	0/2781	0.93	1/3767 (0.0%)
1	B	0.80	0/2781	0.90	0/3767
1	C	0.76	0/2767	0.89	1/3748 (0.0%)
1	D	0.77	0/2773	0.88	1/3757 (0.0%)
All	All	0.78	0/11102	0.90	3/15039 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	76	ARG	NE-CZ-NH2	-5.47	117.56	120.30
1	A	76	ARG	NE-CZ-NH2	-5.42	117.59	120.30
1	C	188	ARG	NE-CZ-NH1	5.18	122.89	120.30

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	2743	0	2788	23	0
1	B	2743	0	2788	26	0
1	C	2729	0	2772	31	0
1	D	2735	0	2776	38	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	12	0	16	0	0
2	B	6	0	8	0	0
3	A	229	0	0	1	0
3	B	236	0	0	2	0
3	C	183	0	0	2	0
3	D	189	0	0	2	0
All	All	11805	0	11148	100	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 100 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:88:LEU:HD23	1:C:88:LEU:HD13	1.39	1.05
1:A:88:LEU:HD13	1:D:88:LEU:HD13	1.47	0.93
1:A:175:HIS:HD2	1:A:208:ARG:H	1.30	0.79
1:D:149:MET:SD	1:D:274:MET:CE	2.72	0.78
1:C:173:ARG:NH2	1:C:321:ASP:OD1	2.17	0.77

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	377/407 (93%)	366 (97%)	11 (3%)	0	100	100
1	B	377/407 (93%)	368 (98%)	9 (2%)	0	100	100
1	C	375/407 (92%)	360 (96%)	15 (4%)	0	100	100
1	D	376/407 (92%)	363 (96%)	13 (4%)	0	100	100
All	All	1505/1628 (92%)	1457 (97%)	48 (3%)	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	273/298 (92%)	270 (99%)	3 (1%)	73	68
1	B	273/298 (92%)	269 (98%)	4 (2%)	65	56
1	C	271/298 (91%)	268 (99%)	3 (1%)	73	68
1	D	272/298 (91%)	268 (98%)	4 (2%)	65	56
All	All	1089/1192 (91%)	1075 (99%)	14 (1%)	69	62

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	142	TYR
1	C	215	SER
1	D	293	ILE
1	D	142	TYR
1	D	274	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	135	ASN
1	D	175	HIS
1	D	369	GLN
1	B	175	HIS
1	B	135	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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

3 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	GOL	A	401	-	5,5,5	0.09	0	5,5,5	0.27	0
2	GOL	A	402	-	5,5,5	0.14	0	5,5,5	0.35	0
2	GOL	B	401	-	5,5,5	0.06	0	5,5,5	0.42	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
2	GOL	A	401	-	-	0/4/4/4	-
2	GOL	A	402	-	-	0/4/4/4	-
2	GOL	B	401	-	-	3/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	401	GOL	C1-C2-C3-O3
2	B	401	GOL	O1-C1-C2-C3
2	B	401	GOL	O2-C2-C3-O3

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 [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	379/407 (93%)	-0.23	2 (0%) 91 89	11, 18, 33, 56	0
1	B	379/407 (93%)	-0.21	3 (0%) 86 84	12, 18, 32, 55	0
1	C	377/407 (92%)	-0.09	3 (0%) 86 84	13, 22, 39, 56	0
1	D	378/407 (92%)	-0.09	3 (0%) 86 84	13, 22, 39, 69	0
All	All	1513/1628 (92%)	-0.15	11 (0%) 87 86	11, 20, 37, 69	0

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	2	SER	4.3
1	C	3	PRO	3.2
1	B	140	GLN	3.0
1	B	4	ILE	2.9
1	C	293	ILE	2.8

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, 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(\AA^2)	Q<0.9
2	GOL	A	401	6/6	0.81	0.12	30,31,34,35	0
2	GOL	B	401	6/6	0.86	0.12	28,32,34,35	0
2	GOL	A	402	6/6	0.88	0.16	28,33,39,40	0

6.5 Other polymers [i](#)

There are no such residues in this entry.