



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 9, 2026 – 06:38 AM UTC

PDB ID : 8GQI / pdb_00008gqi
Title : Structure of Thiolase from Pseudomonas aeruginosa PAO1
Authors : Hong, J.; Son, H.F.; Kim, K.J.
Deposited on : 2022-08-30
Resolution : 2.21 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 ⓘ 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 : 4-5-2 with Phenix2.0
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

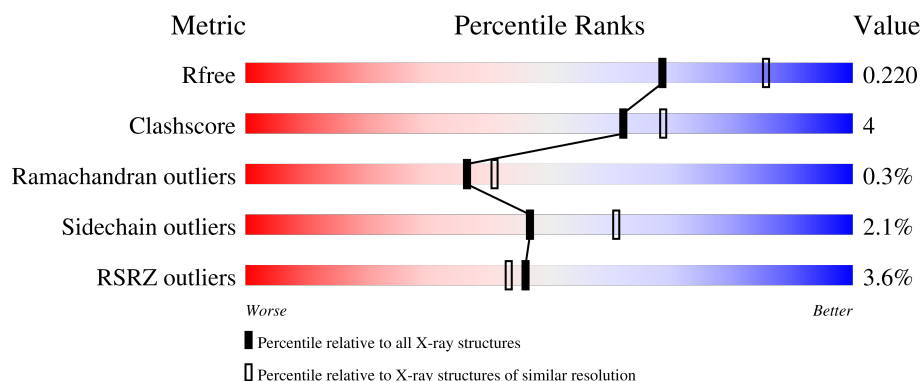
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.21 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}	180053	7682 (2.24-2.20)
Clashscore	190562	8402 (2.24-2.20)
Ramachandran outliers	187476	8303 (2.24-2.20)
Sidechain outliers	187428	8304 (2.24-2.20)
RSRZ outliers	180081	7683 (2.24-2.20)

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 $\leq 5\%$. 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	409	2% 89% 9%
1	B	409	5% 87% 11%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 6160 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
1	A	401	Total	C	N	O	S	0	0	0
			2958	1838	554	554	12			
1	B	400	Total	C	N	O	S	0	0	0
			2947	1832	550	553	12			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	402	LEU	-	expression tag	UNP Q9HY37
A	403	GLU	-	expression tag	UNP Q9HY37
A	404	HIS	-	expression tag	UNP Q9HY37
A	405	HIS	-	expression tag	UNP Q9HY37
A	406	HIS	-	expression tag	UNP Q9HY37
A	407	HIS	-	expression tag	UNP Q9HY37
A	408	HIS	-	expression tag	UNP Q9HY37
A	409	HIS	-	expression tag	UNP Q9HY37
B	402	LEU	-	expression tag	UNP Q9HY37
B	403	GLU	-	expression tag	UNP Q9HY37
B	404	HIS	-	expression tag	UNP Q9HY37
B	405	HIS	-	expression tag	UNP Q9HY37
B	406	HIS	-	expression tag	UNP Q9HY37
B	407	HIS	-	expression tag	UNP Q9HY37
B	408	HIS	-	expression tag	UNP Q9HY37
B	409	HIS	-	expression tag	UNP Q9HY37

- Molecule 2 is GLYCEROL (CCD ID: GOL) (formula: C₃H₈O₃).



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

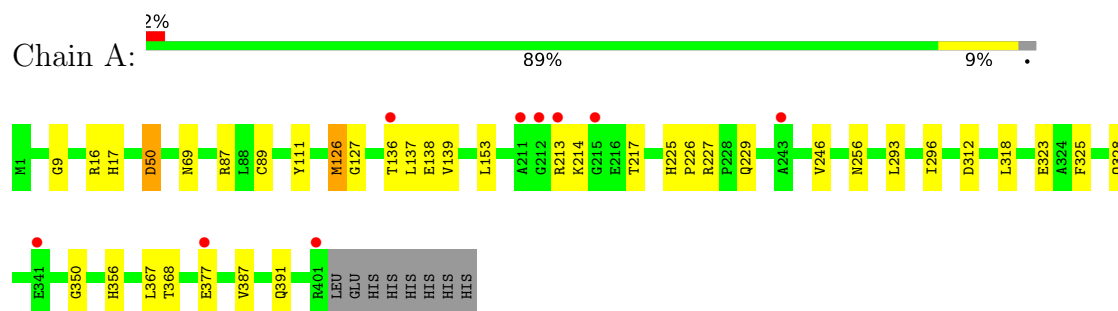
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	126	Total	O	0	0
			126	126		
3	B	111	Total	O	0	0
			111	111		

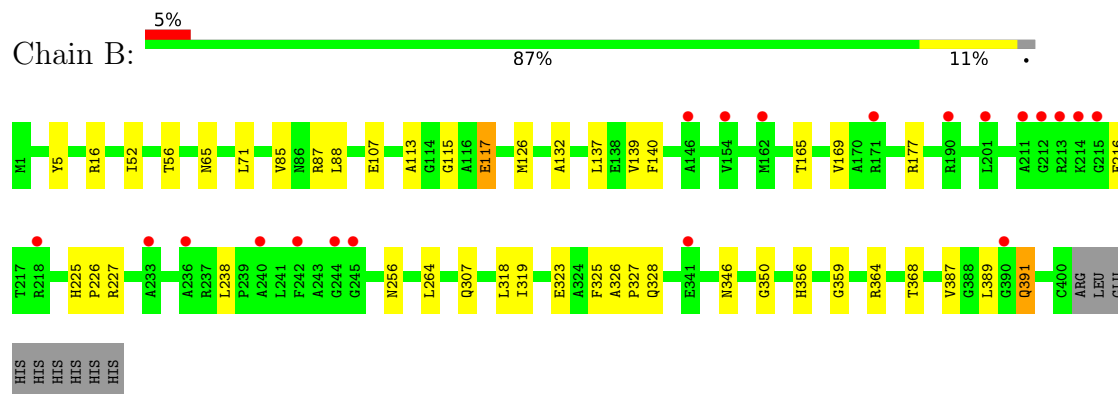
3 Residue-property plots [i](#)

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: Thiolase



• Molecule 1: Thiolase



4 Data and refinement statistics

Property	Value	Source
Space group	I 41	Depositor
Cell constants a, b, c, α , β , γ	177.37Å 177.37Å 70.59Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.39 – 2.21 40.39 – 2.21	Depositor EDS
% Data completeness (in resolution range)	99.2 (40.39-2.21) 99.4 (40.39-2.21)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	5.18 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.174 , 0.214 0.184 , 0.220	Depositor DCC
R_{free} test set	2757 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å ²)	30.3	Xtriage
Anisotropy	0.141	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 34.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.033 for -k,-h,-l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	6160	wwPDB-VP
Average B, all atoms (Å ²)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.43% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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

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	1.15	2/2998 (0.1%)	1.42	7/4057 (0.2%)
1	B	1.14	2/2987 (0.1%)	1.48	8/4043 (0.2%)
All	All	1.15	4/5985 (0.1%)	1.45	15/8100 (0.2%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	115	GLY	C-O	5.53	1.29	1.24
1	A	16	ARG	C-O	5.32	1.31	1.23
1	B	140	PHE	C-O	5.30	1.30	1.23
1	A	127	GLY	C-O	5.19	1.29	1.23

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	391	GLN	CA-C-N	7.42	126.69	121.65
1	B	391	GLN	C-N-CA	7.42	126.69	121.65
1	A	136	THR	CB-CA-C	7.04	121.31	109.84
1	B	117	GLU	CB-CA-C	-6.87	96.07	109.68
1	A	126	MET	CA-C-N	6.72	126.37	119.92
1	A	126	MET	C-N-CA	6.72	126.37	119.92
1	A	69	ASN	N-CA-C	-5.71	105.14	111.36
1	B	87	ARG	N-CA-C	-5.58	103.75	110.44
1	B	117	GLU	CB-CG-CD	5.43	121.83	112.60
1	B	346	ASN	CB-CA-C	5.33	115.72	110.71
1	A	138	GLU	CB-CA-C	5.25	118.80	110.19
1	A	17	HIS	CA-CB-CG	-5.12	108.68	113.80
1	A	50	ASP	CA-CB-CG	5.06	117.66	112.60
1	B	364	ARG	CA-C-O	-5.04	115.53	120.82
1	B	107	GLU	CB-CG-CD	5.01	121.11	112.60

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	2958	0	2996	21	0
1	B	2947	0	2983	29	0
2	A	6	0	8	2	0
2	B	12	0	16	1	0
3	A	126	0	0	0	0
3	B	111	0	0	3	0
All	All	6160	0	6003	49	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 (49) 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:126:MET:HG2	1:B:139:VAL:HG22	1.61	0.83
1:B:126:MET:CG	1:B:139:VAL:HG22	2.23	0.69
1:B:325:PHE:H	1:B:328:GLN:HE21	1.40	0.67
1:B:126:MET:HE3	1:B:137:LEU:HD11	1.79	0.63
1:B:323:GLU:CD	1:B:350:GLY:HA3	2.25	0.61
1:A:325:PHE:H	1:A:328:GLN:HE21	1.46	0.61
1:A:126:MET:HG2	1:A:139:VAL:HG22	1.84	0.59
1:B:132:ALA:O	3:B:601:HOH:O	2.17	0.58
1:A:356:HIS:HB3	2:A:501:GOL:H11	1.87	0.56
1:B:117:GLU:HG2	1:B:359:GLY:H	1.72	0.55
1:A:387:VAL:HB	1:A:391:GLN:HB2	1.89	0.54
1:B:177:ARG:NH2	1:B:238:LEU:O	2.39	0.53
1:A:9:GLY:HA3	1:A:367:LEU:HD13	1.91	0.52
1:B:387:VAL:HB	1:B:391:GLN:HB2	1.92	0.51
1:A:126:MET:HE1	1:B:126:MET:HG3	1.94	0.49
1:B:117:GLU:CG	1:B:359:GLY:H	2.24	0.49
1:A:256:ASN:ND2	1:A:356:HIS:H	2.12	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:87:ARG:HH11	1:A:391:GLN:HE22	1.61	0.48
1:B:16:ARG:NH2	3:B:605:HOH:O	2.38	0.47
1:A:296:ILE:C	1:A:296:ILE:HD12	2.38	0.47
1:B:225:HIS:N	1:B:226:PRO:CD	2.78	0.47
1:B:5:TYR:O	1:B:264:LEU:HA	2.14	0.47
1:A:256:ASN:HD22	1:A:356:HIS:H	1.62	0.46
1:A:153:LEU:C	1:A:153:LEU:HD23	2.42	0.45
1:B:226:PRO:O	1:B:227:ARG:NH1	2.46	0.45
1:A:225:HIS:N	1:A:226:PRO:CD	2.79	0.45
1:B:165:THR:O	1:B:169:VAL:HG23	2.16	0.45
1:B:56:THR:HG22	1:B:88:LEU:HA	1.99	0.44
1:A:356:HIS:CD2	2:A:501:GOL:H12	2.53	0.44
1:A:227:ARG:HD2	1:A:229:GLN:NE2	2.32	0.44
1:B:52:ILE:O	1:B:113:ALA:HA	2.17	0.44
1:A:137:LEU:HG	1:A:139:VAL:HG23	2.00	0.43
1:B:256:ASN:HD22	1:B:356:HIS:H	1.66	0.43
1:B:326:ALA:N	1:B:327:PRO:HD2	2.34	0.43
1:A:50:ASP:HB3	1:A:111:TYR:CE2	2.54	0.42
1:B:126:MET:HE2	1:B:126:MET:HB3	1.82	0.42
1:A:325:PHE:O	1:A:328:GLN:HG3	2.19	0.42
1:B:325:PHE:O	1:B:328:GLN:HG3	2.20	0.42
1:A:318:LEU:C	1:A:318:LEU:HD23	2.45	0.42
1:B:356:HIS:HD2	2:B:501:GOL:H11	1.85	0.42
1:B:227:ARG:HD3	1:B:227:ARG:HA	1.92	0.41
1:A:137:LEU:HD12	1:A:137:LEU:HA	1.85	0.41
1:B:225:HIS:HE1	3:B:610:HOH:O	2.03	0.41
1:B:71:LEU:C	1:B:71:LEU:HD12	2.46	0.41
1:B:256:ASN:ND2	1:B:356:HIS:H	2.18	0.41
1:B:318:LEU:HD23	1:B:319:ILE:N	2.35	0.41
1:A:293:LEU:O	1:A:296:ILE:HG13	2.20	0.40
1:B:264:LEU:N	1:B:264:LEU:HD12	2.36	0.40
1:A:323:GLU:CD	1:A:350:GLY:HA3	2.46	0.40

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	399/409 (98%)	383 (96%)	15 (4%)	1 (0%)	36	41
1	B	398/409 (97%)	383 (96%)	14 (4%)	1 (0%)	36	41
All	All	797/818 (97%)	766 (96%)	29 (4%)	2 (0%)	36	41

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	65	ASN
1	A	213	ARG

5.3.2 Protein sidechains ⓘ

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	291/299 (97%)	284 (98%)	7 (2%)	43	56
1	B	290/299 (97%)	285 (98%)	5 (2%)	53	68
All	All	581/598 (97%)	569 (98%)	12 (2%)	47	61

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	89	CYS
1	A	214	LYS
1	A	217	THR
1	A	246	VAL
1	A	312	ASP
1	A	368	THR
1	A	377	GLU
1	B	85	VAL
1	B	216	GLU
1	B	307	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	368	THR
1	B	389	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (15) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	225	HIS
1	A	256	ASN
1	A	307	GLN
1	A	328	GLN
1	A	348	HIS
1	A	391	GLN
1	B	58	GLN
1	B	83	GLN
1	B	192	GLN
1	B	225	HIS
1	B	256	ASN
1	B	300	GLN
1	B	328	GLN
1	B	348	HIS
1	B	391	GLN

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 oligosaccharides 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	B	501	-	5,5,5	0.24	0	5,5,5	0.67	0
2	GOL	B	502	-	5,5,5	0.09	0	5,5,5	0.26	0
2	GOL	A	501	-	5,5,5	0.31	0	5,5,5	0.77	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	B	501	-	-	2/4/4/4	-
2	GOL	B	502	-	-	2/4/4/4	-
2	GOL	A	501	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	501	GOL	O1-C1-C2-C3
2	B	501	GOL	O1-C1-C2-C3
2	A	501	GOL	O1-C1-C2-O2
2	B	502	GOL	O1-C1-C2-C3
2	B	501	GOL	O1-C1-C2-O2
2	B	502	GOL	O1-C1-C2-O2

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	501	GOL	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	501	GOL	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, 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	401/409 (98%)	0.10	9 (2%) 62 60	20, 32, 55, 109	0
1	B	400/409 (97%)	0.29	20 (5%) 34 31	20, 35, 63, 103	0
All	All	801/818 (97%)	0.20	29 (3%) 46 43	20, 33, 60, 109	0

All (29) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	212	GLY	5.5
1	B	211	ALA	3.8
1	B	212	GLY	3.8
1	A	401	ARG	3.5
1	A	211	ALA	3.3
1	B	240	ALA	2.9
1	B	214	LYS	2.8
1	A	213	ARG	2.8
1	B	213	ARG	2.7
1	B	242	PHE	2.7
1	A	243	ALA	2.6
1	B	215	GLY	2.6
1	B	390	GLY	2.6
1	B	244	GLY	2.3
1	A	341	GLU	2.3
1	B	171	ARG	2.3
1	B	162	MET	2.2
1	B	245	GLY	2.2
1	B	190	ARG	2.2
1	A	215	GLY	2.2
1	B	154	VAL	2.1
1	B	233	ALA	2.1
1	A	377	GLU	2.1
1	A	136	THR	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	201	LEU	2.1
1	B	236	ALA	2.0
1	B	341	GLU	2.0
1	B	146	ALA	2.0
1	B	218	ARG	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 oligosaccharides 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($\text{\AA}^2$)	Q<0.9
2	GOL	A	501	6/6	0.83	0.17	37,51,54,56	0
2	GOL	B	501	6/6	0.90	0.12	44,48,49,55	0
2	GOL	B	502	6/6	0.90	0.16	54,57,58,59	0

6.5 Other polymers [i](#)

There are no such residues in this entry.