



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 6, 2026 – 04:33 AM UTC

PDB ID : 4RZS / pdb_00004rzs
Title : Lac repressor engineered to bind sucralose, unliganded tetramer
Authors : Arbing, M.A.; Cascio, D.; Kosuri, S.; Church, G.M.
Deposited on : 2014-12-24
Resolution : 2.71 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	:	4-5-2 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

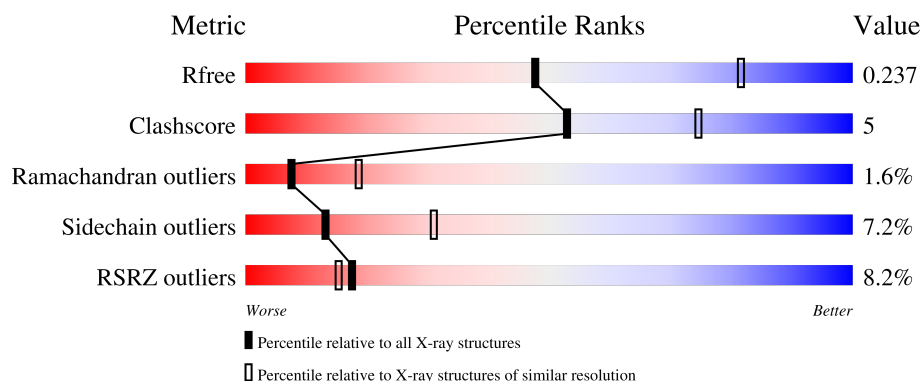
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.71 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	4348 (2.74-2.70)
Clashscore	190562	4665 (2.74-2.70)
Ramachandran outliers	187476	4584 (2.74-2.70)
Sidechain outliers	187428	4585 (2.74-2.70)
RSRZ outliers	180081	4348 (2.74-2.70)

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	381	7% 78% 10% • 9%
1	B	381	14% 79% 9% • 7%
1	C	381	3% 69% 8% • 22%
1	D	381	3% 68% 7% • 22%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 9530 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 Lac repressor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	347	Total	C	N	O	S	0	1	0
			2547	1587	451	497	12			
1	B	355	Total	C	N	O	S	0	7	0
			2595	1619	459	505	12			
1	C	298	Total	C	N	O	S	0	0	0
			2157	1343	384	419	11			
1	D	296	Total	C	N	O	S	0	1	0
			2184	1363	388	422	11			

There are 100 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-20	MET	-	expression tag	UNP C9QQT3
A	-19	GLY	-	expression tag	UNP C9QQT3
A	-18	SER	-	expression tag	UNP C9QQT3
A	-17	SER	-	expression tag	UNP C9QQT3
A	-16	HIS	-	expression tag	UNP C9QQT3
A	-15	HIS	-	expression tag	UNP C9QQT3
A	-14	HIS	-	expression tag	UNP C9QQT3
A	-13	HIS	-	expression tag	UNP C9QQT3
A	-12	HIS	-	expression tag	UNP C9QQT3
A	-11	HIS	-	expression tag	UNP C9QQT3
A	-10	SER	-	expression tag	UNP C9QQT3
A	-9	SER	-	expression tag	UNP C9QQT3
A	-8	GLY	-	expression tag	UNP C9QQT3
A	-7	LEU	-	expression tag	UNP C9QQT3
A	-6	VAL	-	expression tag	UNP C9QQT3
A	-5	PRO	-	expression tag	UNP C9QQT3
A	-4	ARG	-	expression tag	UNP C9QQT3
A	-3	GLY	-	expression tag	UNP C9QQT3
A	-2	SER	-	expression tag	UNP C9QQT3
A	-1	HIS	-	expression tag	UNP C9QQT3
A	0	MET	-	expression tag	UNP C9QQT3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	149	THR	ASP	engineered mutation	UNP C9QQT3
A	150	ALA	VAL	engineered mutation	UNP C9QQT3
A	156	LEU	ILE	engineered mutation	UNP C9QQT3
A	193	ASP	SER	engineered mutation	UNP C9QQT3
B	-20	MET	-	expression tag	UNP C9QQT3
B	-19	GLY	-	expression tag	UNP C9QQT3
B	-18	SER	-	expression tag	UNP C9QQT3
B	-17	SER	-	expression tag	UNP C9QQT3
B	-16	HIS	-	expression tag	UNP C9QQT3
B	-15	HIS	-	expression tag	UNP C9QQT3
B	-14	HIS	-	expression tag	UNP C9QQT3
B	-13	HIS	-	expression tag	UNP C9QQT3
B	-12	HIS	-	expression tag	UNP C9QQT3
B	-11	HIS	-	expression tag	UNP C9QQT3
B	-10	SER	-	expression tag	UNP C9QQT3
B	-9	SER	-	expression tag	UNP C9QQT3
B	-8	GLY	-	expression tag	UNP C9QQT3
B	-7	LEU	-	expression tag	UNP C9QQT3
B	-6	VAL	-	expression tag	UNP C9QQT3
B	-5	PRO	-	expression tag	UNP C9QQT3
B	-4	ARG	-	expression tag	UNP C9QQT3
B	-3	GLY	-	expression tag	UNP C9QQT3
B	-2	SER	-	expression tag	UNP C9QQT3
B	-1	HIS	-	expression tag	UNP C9QQT3
B	0	MET	-	expression tag	UNP C9QQT3
B	149	THR	ASP	engineered mutation	UNP C9QQT3
B	150	ALA	VAL	engineered mutation	UNP C9QQT3
B	156	LEU	ILE	engineered mutation	UNP C9QQT3
B	193	ASP	SER	engineered mutation	UNP C9QQT3
C	-20	MET	-	expression tag	UNP C9QQT3
C	-19	GLY	-	expression tag	UNP C9QQT3
C	-18	SER	-	expression tag	UNP C9QQT3
C	-17	SER	-	expression tag	UNP C9QQT3
C	-16	HIS	-	expression tag	UNP C9QQT3
C	-15	HIS	-	expression tag	UNP C9QQT3
C	-14	HIS	-	expression tag	UNP C9QQT3
C	-13	HIS	-	expression tag	UNP C9QQT3
C	-12	HIS	-	expression tag	UNP C9QQT3
C	-11	HIS	-	expression tag	UNP C9QQT3
C	-10	SER	-	expression tag	UNP C9QQT3
C	-9	SER	-	expression tag	UNP C9QQT3
C	-8	GLY	-	expression tag	UNP C9QQT3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-7	LEU	-	expression tag	UNP C9QQT3
C	-6	VAL	-	expression tag	UNP C9QQT3
C	-5	PRO	-	expression tag	UNP C9QQT3
C	-4	ARG	-	expression tag	UNP C9QQT3
C	-3	GLY	-	expression tag	UNP C9QQT3
C	-2	SER	-	expression tag	UNP C9QQT3
C	-1	HIS	-	expression tag	UNP C9QQT3
C	0	MET	-	expression tag	UNP C9QQT3
C	149	THR	ASP	engineered mutation	UNP C9QQT3
C	150	ALA	VAL	engineered mutation	UNP C9QQT3
C	156	LEU	ILE	engineered mutation	UNP C9QQT3
C	193	ASP	SER	engineered mutation	UNP C9QQT3
D	-20	MET	-	expression tag	UNP C9QQT3
D	-19	GLY	-	expression tag	UNP C9QQT3
D	-18	SER	-	expression tag	UNP C9QQT3
D	-17	SER	-	expression tag	UNP C9QQT3
D	-16	HIS	-	expression tag	UNP C9QQT3
D	-15	HIS	-	expression tag	UNP C9QQT3
D	-14	HIS	-	expression tag	UNP C9QQT3
D	-13	HIS	-	expression tag	UNP C9QQT3
D	-12	HIS	-	expression tag	UNP C9QQT3
D	-11	HIS	-	expression tag	UNP C9QQT3
D	-10	SER	-	expression tag	UNP C9QQT3
D	-9	SER	-	expression tag	UNP C9QQT3
D	-8	GLY	-	expression tag	UNP C9QQT3
D	-7	LEU	-	expression tag	UNP C9QQT3
D	-6	VAL	-	expression tag	UNP C9QQT3
D	-5	PRO	-	expression tag	UNP C9QQT3
D	-4	ARG	-	expression tag	UNP C9QQT3
D	-3	GLY	-	expression tag	UNP C9QQT3
D	-2	SER	-	expression tag	UNP C9QQT3
D	-1	HIS	-	expression tag	UNP C9QQT3
D	0	MET	-	expression tag	UNP C9QQT3
D	149	THR	ASP	engineered mutation	UNP C9QQT3
D	150	ALA	VAL	engineered mutation	UNP C9QQT3
D	156	LEU	ILE	engineered mutation	UNP C9QQT3
D	193	ASP	SER	engineered mutation	UNP C9QQT3

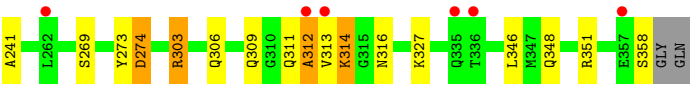
- Molecule 2 is GLYCEROL (CCD ID: GOL) (formula: $C_3H_8O_3$).



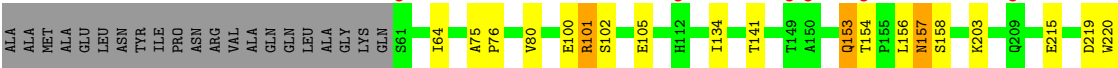
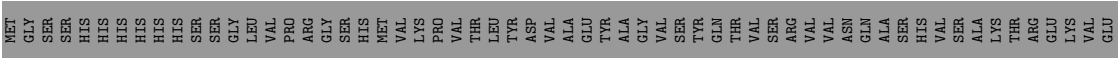
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			6	3	3		
2	B	1	Total	C	O	0	0
			6	3	3		
2	C	1	Total	C	O	0	0
			6	3	3		
2	D	1	Total	C	O	0	0
			6	3	3		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	12	Total	O	0	0
			12	12		
3	B	4	Total	O	0	0
			4	4		
3	C	2	Total	O	0	0
			2	2		
3	D	5	Total	O	0	0
			5	5		



● Molecule 1: Lac repressor



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	91.09Å 111.60Å 189.33Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	94.49 – 2.71 94.49 – 2.71	Depositor EDS
% Data completeness (in resolution range)	68.1 (94.49-2.71) 68.4 (94.49-2.71)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.03 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.199 , 0.239 0.200 , 0.237	Depositor DCC
R_{free} test set	3638 reflections (6.87%)	wwPDB-VP
Wilson B-factor (Å ²)	73.3	Xtriage
Anisotropy	0.074	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 46.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	9530	wwPDB-VP
Average B, all atoms (Å ²)	79.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.17% 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	1.11	0/2586	1.06	6/3527 (0.2%)
1	B	1.13	2/2635 (0.1%)	1.06	6/3600 (0.2%)
1	C	1.00	0/2186	1.02	3/2982 (0.1%)
1	D	1.05	0/2214	1.05	5/3015 (0.2%)
All	All	1.08	2/9621 (0.0%)	1.05	20/13124 (0.2%)

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	5
1	D	0	1
All	All	0	7

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	48	ILE	CA-C	5.87	1.59	1.52
1	B	149	THR	CA-C	-5.26	1.47	1.53

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	50	ASN	N-CA-C	10.34	125.79	110.48
1	B	315	GLY	N-CA-C	7.90	122.84	112.25
1	D	335	GLN	N-CA-C	7.78	122.12	109.59
1	D	315	GLY	N-CA-C	7.65	122.51	112.25

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	100	GLU	N-CA-C	-7.42	97.96	109.76

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	315	GLY	Peptide
1	B	149	THR	Peptide
1	B	151[B]	SER	Peptide
1	B	49	PRO	Peptide
1	B	61	SER	Peptide

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	2547	0	2515	22	0
1	B	2595	0	2541	50	0
1	C	2157	0	2131	14	0
1	D	2184	0	2188	18	0
2	A	6	0	8	0	0
2	B	6	0	8	0	0
2	C	6	0	8	0	0
2	D	6	0	8	0	0
3	A	12	0	0	0	0
3	B	4	0	0	0	0
3	C	2	0	0	0	0
3	D	5	0	0	0	0
All	All	9530	0	9407	94	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 94 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:152[B]:ASP:HB3	1:B:153[B]:GLN:NE2	1.34	1.41
1:B:154[B]:THR:CG2	1:B:156[B]:LEU:HG	1.79	1.11
1:B:154[B]:THR:HG22	1:B:156[B]:LEU:HG	1.15	1.09
1:B:152[B]:ASP:CB	1:B:153[B]:GLN:NE2	2.16	1.08
1:B:130:ASP:CB	1:B:152[B]:ASP:HB2	1.97	0.94

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	344/381 (90%)	320 (93%)	20 (6%)	4 (1%)	10	25
1	B	360/381 (94%)	330 (92%)	17 (5%)	13 (4%)	2	5
1	C	296/381 (78%)	280 (95%)	12 (4%)	4 (1%)	9	22
1	D	295/381 (77%)	282 (96%)	10 (3%)	3 (1%)	12	30
All	All	1295/1524 (85%)	1212 (94%)	59 (5%)	24 (2%)	7	15

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	152[A]	ASP
1	B	152[B]	ASP
1	B	156[A]	LEU
1	B	156[B]	LEU
1	C	101	ARG

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	267/311 (86%)	246 (92%)	21 (8%)	11	27
1	B	265/311 (85%)	245 (92%)	20 (8%)	12	29
1	C	224/311 (72%)	209 (93%)	15 (7%)	15	34
1	D	231/311 (74%)	213 (92%)	18 (8%)	11	28
All	All	987/1244 (79%)	913 (92%)	74 (8%)	13	29

5 of 74 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	102	SER
1	D	342	LEU
1	D	134	ILE
1	D	303	ARG
1	B	151[A]	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 39 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	311	GLN
1	D	234	ASN
1	D	74	HIS
1	D	157	ASN
1	D	311	GLN

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 oligosaccharides in this entry.

5.6 Ligand geometry

4 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	C	501	-	5,5,5	0.55	0	5,5,5	0.65	0
2	GOL	B	501	-	5,5,5	0.77	0	5,5,5	0.56	0
2	GOL	D	501	-	5,5,5	0.59	0	5,5,5	0.65	0
2	GOL	A	501	-	5,5,5	0.28	0	5,5,5	0.68	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	C	501	-	-	2/4/4/4	-
2	GOL	B	501	-	-	1/4/4/4	-
2	GOL	D	501	-	-	0/4/4/4	-
2	GOL	A	501	-	-	0/4/4/4	-

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	C	501	GOL	O1-C1-C2-C3
2	C	501	GOL	O1-C1-C2-O2
2	B	501	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	347/381 (91%)	0.27	28 (8%) 18 16	33, 63, 136, 162	1 (0%)
1	B	355/381 (93%)	0.48	53 (14%) 5 5	34, 67, 182, 208	7 (1%)
1	C	298/381 (78%)	0.49	12 (4%) 42 39	55, 91, 123, 144	0
1	D	296/381 (77%)	0.24	13 (4%) 39 36	44, 71, 101, 120	1 (0%)
All	All	1296/1524 (85%)	0.37	106 (8%) 17 15	33, 74, 149, 208	9 (0%)

The worst 5 of 106 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	49	PRO	6.2
1	B	151[A]	SER	6.1
1	B	27	ALA	6.1
1	B	17	TYR	5.6
1	B	18	GLN	5.5

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	C	501	6/6	0.80	0.15	59,64,70,71	0
2	GOL	A	501	6/6	0.86	0.12	48,53,67,68	0
2	GOL	B	501	6/6	0.90	0.12	51,60,70,70	0
2	GOL	D	501	6/6	0.90	0.14	56,64,75,78	0

6.5 Other polymers [i](#)

There are no such residues in this entry.