



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 29, 2026 – 10:12 PM UTC

PDB ID : 2ZUT / pdb_00002zut
Title : Crystal structure of Galacto-N-biose/Lacto-N-biose I phosphorylase in complex with GalNAc
Authors : Hidaka, M.; Nishimoto, M.; Kitaoka, M.; Wakagi, T.; Shoun, H.; Fushinobu, S.
Deposited on : 2008-10-28
Resolution : 1.90 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

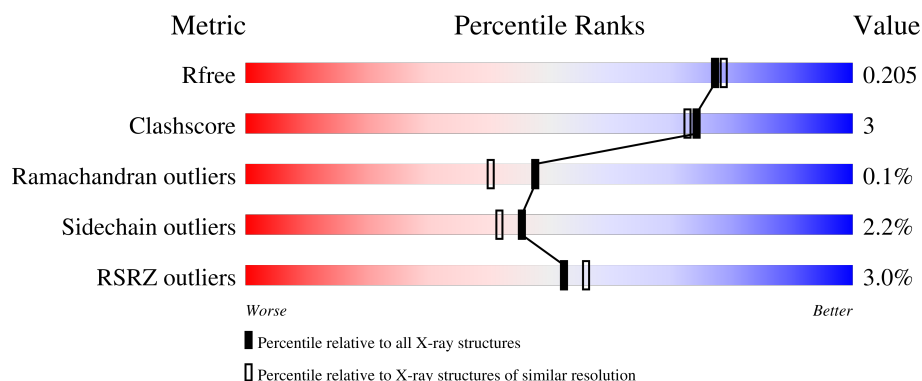
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.90 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	7789 (1.90-1.90)
Clashscore	190562	8410 (1.90-1.90)
Ramachandran outliers	187476	8333 (1.90-1.90)
Sidechain outliers	187428	8333 (1.90-1.90)
RSRZ outliers	180081	7790 (1.90-1.90)

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	759	% 91% 6% ..
1	B	759	8% 87% 8% ..
1	C	759	% 93% 5% ..
1	D	759	2% 91% 6% .

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 26811 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 Lacto-N-biose phosphorylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	744	Total	C	N	O	S	0	0	0
			5920	3779	996	1129	16			
1	B	735	Total	C	N	O	S	0	0	0
			5851	3734	985	1116	16			
1	C	744	Total	C	N	O	S	0	0	0
			5920	3779	996	1129	16			
1	D	739	Total	C	N	O	S	0	0	0
			5882	3756	989	1121	16			

There are 32 discrepancies between the modelled and reference sequences:

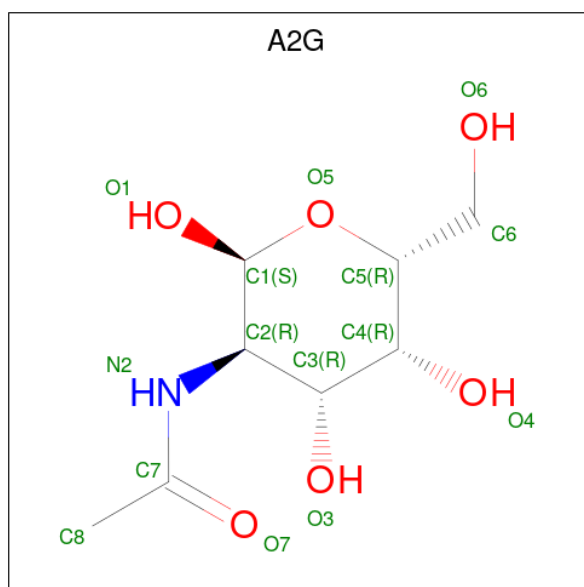
Chain	Residue	Modelled	Actual	Comment	Reference
A	752	LEU	-	expression tag	UNP Q5NU17
A	753	GLU	-	expression tag	UNP Q5NU17
A	754	HIS	-	expression tag	UNP Q5NU17
A	755	HIS	-	expression tag	UNP Q5NU17
A	756	HIS	-	expression tag	UNP Q5NU17
A	757	HIS	-	expression tag	UNP Q5NU17
A	758	HIS	-	expression tag	UNP Q5NU17
A	759	HIS	-	expression tag	UNP Q5NU17
B	752	LEU	-	expression tag	UNP Q5NU17
B	753	GLU	-	expression tag	UNP Q5NU17
B	754	HIS	-	expression tag	UNP Q5NU17
B	755	HIS	-	expression tag	UNP Q5NU17
B	756	HIS	-	expression tag	UNP Q5NU17
B	757	HIS	-	expression tag	UNP Q5NU17
B	758	HIS	-	expression tag	UNP Q5NU17
B	759	HIS	-	expression tag	UNP Q5NU17
C	752	LEU	-	expression tag	UNP Q5NU17
C	753	GLU	-	expression tag	UNP Q5NU17
C	754	HIS	-	expression tag	UNP Q5NU17
C	755	HIS	-	expression tag	UNP Q5NU17
C	756	HIS	-	expression tag	UNP Q5NU17

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	757	HIS	-	expression tag	UNP Q5NU17
C	758	HIS	-	expression tag	UNP Q5NU17
C	759	HIS	-	expression tag	UNP Q5NU17
D	752	LEU	-	expression tag	UNP Q5NU17
D	753	GLU	-	expression tag	UNP Q5NU17
D	754	HIS	-	expression tag	UNP Q5NU17
D	755	HIS	-	expression tag	UNP Q5NU17
D	756	HIS	-	expression tag	UNP Q5NU17
D	757	HIS	-	expression tag	UNP Q5NU17
D	758	HIS	-	expression tag	UNP Q5NU17
D	759	HIS	-	expression tag	UNP Q5NU17

- Molecule 2 is 2-acetamido-2-deoxy- α -D-galactopyranose (CCD ID: A2G) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			15	8	1	6		
2	B	1	Total	C	N	O	0	0
			15	8	1	6		
2	C	1	Total	C	N	O	0	0
			15	8	1	6		
2	D	1	Total	C	N	O	0	0
			15	8	1	6		

- Molecule 3 is GLYCEROL (CCD ID: GOL) (formula: $C_3H_8O_3$).



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

- Molecule 4 is NITRATE ION (CCD ID: NO3) (formula: NO₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	N	O	0	0
			4	1	3		

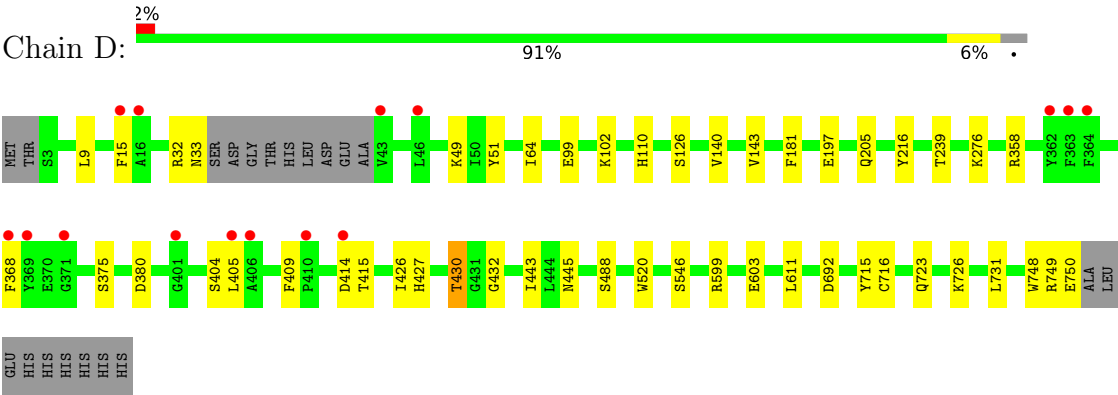
- Molecule 5 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	Mg	0	0
			1	1		
5	B	2	Total	Mg	0	0
			2	2		
5	C	1	Total	Mg	0	0
			1	1		
5	D	2	Total	Mg	0	0
			2	2		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	825	Total	O	0	0
			825	825		
6	B	687	Total	O	0	0
			687	687		
6	C	874	Total	O	0	0
			874	874		
6	D	728	Total	O	1	0
			728	728		

● Molecule 1: Lacto-N-biose phosphorylase



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	67.86Å 111.66Å 118.66Å 105.20° 90.48° 107.27°	Depositor
Resolution (Å)	44.81 – 1.90 44.81 – 1.90	Depositor EDS
% Data completeness (in resolution range)	96.7 (44.81-1.90) 96.6 (44.81-1.90)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.99 (at 1.89Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.161 , 0.204 0.161 , 0.205	Depositor DCC
R_{free} test set	12254 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å ²)	23.7	Xtriage
Anisotropy	0.049	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 43.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.039 for h,-h-k,-l	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	26811	wwPDB-VP
Average B, all atoms (Å ²)	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.32% 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: NO3, A2G, MG, 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	0.83	0/6088	0.87	2/8292 (0.0%)
1	B	0.79	1/6018 (0.0%)	0.88	5/8196 (0.1%)
1	C	0.81	0/6088	0.84	0/8292
1	D	0.78	0/6049	0.85	2/8239 (0.0%)
All	All	0.80	1/24243 (0.0%)	0.86	9/33019 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	346	ILE	CA-CB	5.10	1.61	1.54

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	364	PHE	CB-CA-C	7.57	120.75	109.63
1	D	102	LYS	CA-C-N	-6.86	112.70	119.76
1	D	102	LYS	C-N-CA	-6.86	112.70	119.76
1	B	71	THR	CA-C-N	-5.86	114.23	120.03
1	B	71	THR	C-N-CA	-5.86	114.23	120.03

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	5920	0	5612	29	0
1	B	5851	0	5541	56	0
1	C	5920	0	5612	34	0
1	D	5882	0	5578	23	0
2	A	15	0	12	0	0
2	B	15	0	12	0	0
2	C	15	0	12	0	0
2	D	15	0	12	0	0
3	A	18	0	24	1	0
3	B	12	0	16	0	0
3	C	12	0	16	0	0
3	D	12	0	16	0	0
4	A	4	0	0	0	0
5	A	1	0	0	0	0
5	B	2	0	0	0	0
5	C	1	0	0	0	0
5	D	2	0	0	0	0
6	A	825	0	0	3	0
6	B	687	0	0	12	0
6	C	874	0	0	10	0
6	D	728	0	0	6	0
All	All	26811	0	22463	142	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 3.

The worst 5 of 142 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:99:GLU:HG3	6:B:787:HOH:O	1.65	0.97
1:B:364:PHE:HD1	1:B:365:PRO:HD2	1.40	0.85
1:A:110:HIS:HD2	1:A:126:SER:HB2	1.40	0.85
1:A:33:ASN:HD22	1:A:34:SER:H	1.22	0.83
1:B:411:LYS:HD3	6:B:2332:HOH:O	1.78	0.83

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	740/759 (98%)	716 (97%)	22 (3%)	2 (0%)	36	29
1	B	731/759 (96%)	703 (96%)	27 (4%)	1 (0%)	48	40
1	C	740/759 (98%)	718 (97%)	22 (3%)	0	100	100
1	D	735/759 (97%)	709 (96%)	26 (4%)	0	100	100
All	All	2946/3036 (97%)	2846 (97%)	97 (3%)	3 (0%)	48	40

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	173	GLY
1	A	743	ASP
1	B	713	GLY

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	621/634 (98%)	609 (98%)	12 (2%)	50	47
1	B	614/634 (97%)	595 (97%)	19 (3%)	35	29
1	C	621/634 (98%)	613 (99%)	8 (1%)	61	61
1	D	617/634 (97%)	602 (98%)	15 (2%)	43	38
All	All	2473/2536 (98%)	2419 (98%)	54 (2%)	45	42

5 of 54 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	717	VAL
1	C	370	GLU
1	D	611	LEU
1	B	731	LEU
1	C	353	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 34 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	166	ASN
1	D	170	ASN
1	D	690	ASN
1	B	166	ASN
1	B	52	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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 20 ligands modelled in this entry, 6 are monoatomic - leaving 14 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
2	A2G	B	4002	-	15,15,15	0.71	0	21,21,21	0.84	1 (4%)
3	GOL	C	4010	-	5,5,5	0.23	0	5,5,5	0.88	0
4	NO3	A	4014	-	1,3,3	3.63	1 (100%)	0,3,3	-	-
2	A2G	A	4001	-	15,15,15	0.61	0	21,21,21	0.77	0
3	GOL	D	4012	-	5,5,5	0.26	0	5,5,5	0.65	0
3	GOL	B	4009	-	5,5,5	0.30	0	5,5,5	0.79	0
3	GOL	B	4008	-	5,5,5	0.34	0	5,5,5	0.87	0
3	GOL	A	4007	-	5,5,5	0.29	0	5,5,5	0.37	0
2	A2G	D	4004	-	15,15,15	0.76	0	21,21,21	0.81	1 (4%)
3	GOL	C	4011	-	5,5,5	0.31	0	5,5,5	0.68	0
3	GOL	A	4006	-	5,5,5	0.34	0	5,5,5	0.22	0
2	A2G	C	4003	-	15,15,15	0.49	0	21,21,21	0.85	1 (4%)
3	GOL	D	4013	-	5,5,5	0.37	0	5,5,5	0.69	0
3	GOL	A	4005	-	5,5,5	0.32	0	5,5,5	0.82	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	A2G	B	4002	-	-	0/6/26/26	0/1/1/1
3	GOL	C	4010	-	-	0/4/4/4	-
2	A2G	A	4001	-	-	0/6/26/26	0/1/1/1
3	GOL	D	4012	-	-	0/4/4/4	-
3	GOL	B	4009	-	-	2/4/4/4	-
3	GOL	B	4008	-	-	0/4/4/4	-
3	GOL	A	4007	-	-	3/4/4/4	-
2	A2G	D	4004	-	-	0/6/26/26	0/1/1/1
3	GOL	C	4011	-	-	0/4/4/4	-
3	GOL	A	4006	-	-	3/4/4/4	-
2	A2G	C	4003	-	-	0/6/26/26	0/1/1/1
3	GOL	D	4013	-	-	0/4/4/4	-
3	GOL	A	4005	-	-	3/4/4/4	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	4014	NO3	O1-N	3.63	1.42	1.24

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	4002	A2G	O5-C1-C2	2.34	111.87	109.52
2	D	4004	A2G	O5-C1-C2	2.24	111.76	109.52
2	C	4003	A2G	O5-C1-C2	2.05	111.58	109.52

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	4006	GOL	C1-C2-C3-O3
3	A	4007	GOL	O1-C1-C2-C3
3	B	4009	GOL	O1-C1-C2-C3
3	A	4005	GOL	O1-C1-C2-C3
3	A	4006	GOL	O2-C2-C3-O3

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	4005	GOL	1	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 [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	744/759 (98%)	-0.39	7 (0%) 81 83	15, 22, 39, 54	0
1	B	735/759 (96%)	0.07	62 (8%) 17 18	16, 25, 57, 73	0
1	C	744/759 (98%)	-0.38	4 (0%) 87 89	15, 23, 38, 52	0
1	D	739/759 (97%)	-0.12	15 (2%) 65 69	16, 25, 48, 58	0
All	All	2962/3036 (97%)	-0.21	88 (2%) 52 56	15, 24, 48, 73	0

The worst 5 of 88 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	364	PHE	5.6
1	B	368	PHE	4.5
1	C	43	VAL	4.4
1	B	10	PRO	4.1
1	B	406	ALA	3.8

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 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
4	NO3	A	4014	4/4	0.86	0.12	33,33,35,35	0
3	GOL	B	4009	6/6	0.89	0.11	35,43,43,45	0
3	GOL	D	4013	6/6	0.90	0.11	29,41,42,44	0
3	GOL	B	4008	6/6	0.90	0.09	26,30,32,34	0
5	MG	B	4015	1/1	0.90	0.07	47,47,47,47	0
3	GOL	A	4007	6/6	0.91	0.10	53,53,53,54	0
3	GOL	C	4011	6/6	0.92	0.09	32,38,40,40	0
3	GOL	C	4010	6/6	0.92	0.09	24,24,28,29	0
3	GOL	D	4012	6/6	0.94	0.08	26,28,29,32	0
3	GOL	A	4006	6/6	0.94	0.08	30,39,41,45	0
3	GOL	A	4005	6/6	0.95	0.07	23,26,28,31	0
2	A2G	B	4002	15/15	0.96	0.06	21,23,25,27	0
5	MG	A	4016	1/1	0.96	0.05	48,48,48,48	0
2	A2G	A	4001	15/15	0.96	0.06	20,21,23,24	0
2	A2G	C	4003	15/15	0.97	0.05	19,20,23,23	0
2	A2G	D	4004	15/15	0.97	0.05	21,22,24,25	0
5	MG	D	4020	1/1	0.97	0.08	43,43,43,43	0
5	MG	C	4018	1/1	0.98	0.06	42,42,42,42	0
5	MG	D	4017	1/1	0.99	0.02	25,25,25,25	0
5	MG	B	4019	1/1	1.00	0.06	31,31,31,31	0

6.5 Other polymers [i](#)

There are no such residues in this entry.