



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2018 – 10:19 pm GMT

PDB ID : 5T5L
Title : LECTIN FROM BAUHINIA FORFICATA IN COMPLEX WITH TN-PEPTIDE
Authors : Lubkowski, J.; Wlodawer, A.
Deposited on : 2016-08-31
Resolution : 1.17 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.7.3 (157068), CSD as539be (2018)
Xtriage (Phenix) : 1.13
EDS : trunk30967
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk30967

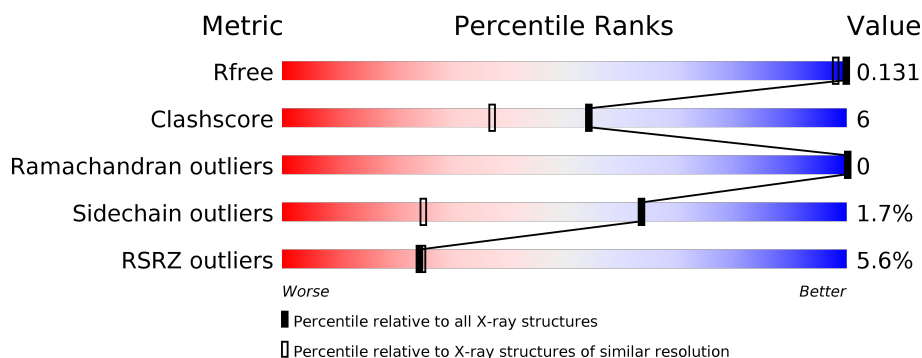
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.17 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}	111664	1413 (1.22-1.14)
Clashscore	122126	1482 (1.22-1.14)
Ramachandran outliers	120053	1429 (1.22-1.14)
Sidechain outliers	120020	1428 (1.22-1.14)
RSRZ outliers	108989	1388 (1.22-1.14)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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	242	<div> <div>5%</div> <div> <div></div> <div>84%</div> <div>9% • 5%</div> </div> </div>
1	B	242	<div> <div>4%</div> <div> <div></div> <div>88%</div> <div>6% • 5%</div> </div> </div>
2	a	5	<div> <div>60%</div> <div> <div></div> <div>80%</div> <div>20%</div> </div> </div>
2	b	5	<div> <div>40%</div> <div> <div></div> <div>100%</div> </div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	EDO	A	306	-	-	X	-

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 4337 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 Lectin.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	A	229	Total	C	N	O	0	16	0
			1878	1192	321	365			
1	B	229	Total	C	N	O	0	13	0
			1846	1175	309	362			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	234	GLY	-	expression tag	UNP P86993
A	235	ALA	-	expression tag	UNP P86993
A	236	ARG	-	expression tag	UNP P86993
A	237	HIS	-	expression tag	UNP P86993
A	238	HIS	-	expression tag	UNP P86993
A	239	HIS	-	expression tag	UNP P86993
A	240	HIS	-	expression tag	UNP P86993
A	241	HIS	-	expression tag	UNP P86993
A	242	HIS	-	expression tag	UNP P86993
B	234	GLY	-	expression tag	UNP P86993
B	235	ALA	-	expression tag	UNP P86993
B	236	ARG	-	expression tag	UNP P86993
B	237	HIS	-	expression tag	UNP P86993
B	238	HIS	-	expression tag	UNP P86993
B	239	HIS	-	expression tag	UNP P86993
B	240	HIS	-	expression tag	UNP P86993
B	241	HIS	-	expression tag	UNP P86993
B	242	HIS	-	expression tag	UNP P86993

- Molecule 2 is a protein called TN ANTIGEN ACE-SER-SER-VAL-GLY.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	a	4	Total	C	N	O	0	0	0
			22	13	3	6			

Continued on next page...

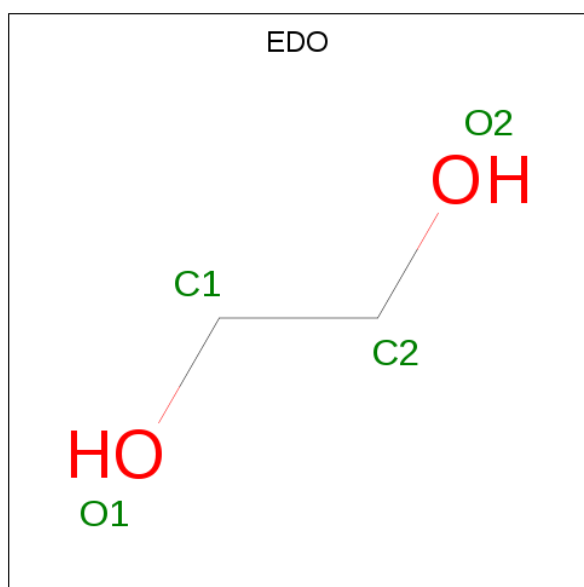
Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	b	5	Total	C	N	O	0	0	0
			26	15	4	7			

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		

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

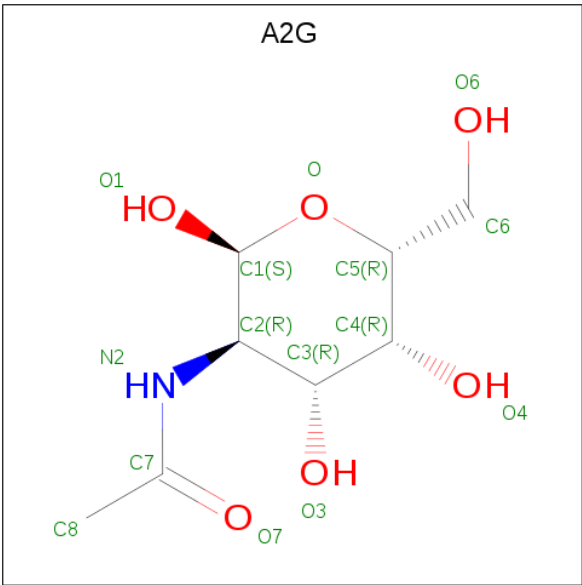


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	1	Total	Cl	0	0
			1	1		

- Molecule 7 is N-ACETYL-2-DEOXY-2-AMINO-GALACTOSE (three-letter code: A2G) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	a	1	Total	C	N	O	0	0
			14	8	1	5		
7	b	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	234	Total	O	0	2
			236	236		
8	B	232	Total	O	0	0
			232	232		
8	a	7	Total	O	0	0
			7	7		
8	b	11	Total	O	0	0
			11	11		

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	44.87Å 88.50Å 110.78Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.30 – 1.17 32.44 – 1.17	Depositor EDS
% Data completeness (in resolution range)	97.2 (30.30-1.17) 97.2 (32.44-1.17)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.28 (at 1.17Å)	Xtriage
Refinement program	REFMAC 5.8.0107	Depositor
R, R_{free}	0.107 , 0.130 0.108 , 0.131	Depositor DCC
R_{free} test set	4270 reflections (2.97%)	wwPDB-VP
Wilson B-factor (Å ²)	15.0	Xtriage
Anisotropy	0.276	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 57.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.98	EDS
Total number of atoms	4337	wwPDB-VP
Average B, all atoms (Å ²)	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.03% 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, ACE, CL, CA, EDO, A2G

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	0.91	9/2002 (0.4%)	0.98	10/2732 (0.4%)
1	B	0.97	5/1960 (0.3%)	0.94	2/2679 (0.1%)
2	a	1.40	0/19	0.78	0/25
2	b	1.32	0/23	1.04	0/30
All	All	0.95	14/4004 (0.3%)	0.96	12/5466 (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

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	14	GLU	CD-OE1	13.31	1.40	1.25
1	B	14	GLU	CD-OE2	8.29	1.34	1.25
1	A	27	GLU	CD-OE2	6.50	1.32	1.25
1	B	19[A]	GLN	CD-OE1	-5.65	1.11	1.24
1	B	19[B]	GLN	CD-OE1	-5.65	1.11	1.24

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	203	ARG	NE-CZ-NH1	9.36	124.98	120.30
1	A	188[A]	LEU	CB-CG-CD1	7.24	123.30	111.00
1	A	188[B]	LEU	CB-CG-CD1	7.24	123.30	111.00
1	A	14[A]	GLU	N-CA-CB	-6.19	99.45	110.60

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	14[B]	GLU	N-CA-CB	-6.19	99.45	110.60

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	131	ARG	Mainchain

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	1878	0	1773	20	0
1	B	1846	0	1755	22	0
2	a	22	0	21	0	0
2	b	26	0	24	0	0
3	A	2	0	0	0	0
3	B	2	0	0	0	0
4	A	20	0	30	9	0
4	B	20	0	28	2	0
5	B	6	0	8	0	0
6	B	1	0	0	0	0
7	a	14	0	13	0	0
7	b	14	0	13	0	0
8	A	236	0	0	6	2
8	B	232	0	0	14	1
8	a	7	0	0	0	0
8	b	11	0	0	0	0
All	All	4337	0	3665	43	2

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 6.

The worst 5 of 43 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:42[B]:ARG:NH1	8:A:403:HOH:O	1.86	1.08
1:B:104[B]:ASN:OD1	8:B:401:HOH:O	1.77	1.01
1:B:70:ARG:NH1	8:B:403:HOH:O	1.94	0.99
1:B:104[B]:ASN:ND2	8:B:402:HOH:O	1.94	0.96
1:B:104[B]:ASN:CG	8:B:401:HOH:O	2.02	0.95

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
8:A:553:HOH:O	8:B:573:HOH:O[3_555]	1.94	0.26
8:A:584:HOH:O	8:A:616:HOH:O[4_456]	2.09	0.11

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	243/242 (100%)	237 (98%)	6 (2%)	0	100	100
1	B	240/242 (99%)	234 (98%)	6 (2%)	0	100	100
2	a	2/5 (40%)	2 (100%)	0	0	100	100
2	b	3/5 (60%)	3 (100%)	0	0	100	100
All	All	488/494 (99%)	476 (98%)	12 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	214/209 (102%)	208 (97%)	6 (3%)	47	8
1	B	211/209 (101%)	207 (98%)	4 (2%)	60	20
2	a	3/3 (100%)	3 (100%)	0	100	100
2	b	3/3 (100%)	3 (100%)	0	100	100
All	All	431/424 (102%)	421 (98%)	10 (2%)	63	14

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	188[A]	LEU
1	A	188[B]	LEU
1	B	70	ARG
1	A	130	PRO
1	B	25	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 18 ligands modelled in this entry, 5 are monoatomic - leaving 13 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	EDO	A	303	-	3,3,3	0.50	0	2,2,2	1.07	0
4	EDO	A	304	-	3,3,3	0.96	0	2,2,2	3.42	2 (100%)
4	EDO	A	305	-	3,3,3	0.63	0	2,2,2	0.44	0
4	EDO	A	306	-	3,3,3	0.75	0	2,2,2	0.96	0
4	EDO	A	307	-	3,3,3	0.41	0	2,2,2	0.17	0
4	EDO	B	303	-	3,3,3	0.32	0	2,2,2	0.63	0
4	EDO	B	304	-	3,3,3	0.41	0	2,2,2	0.32	0
4	EDO	B	305	-	3,3,3	0.79	0	2,2,2	0.23	0
5	GOL	B	306	-	5,5,5	1.27	1 (20%)	5,5,5	2.11	3 (60%)
4	EDO	B	308	-	3,3,3	3.24	2 (66%)	2,2,2	0.18	0
4	EDO	B	309	-	3,3,3	0.44	0	2,2,2	0.03	0
7	A2G	a	201	2	14,14,15	0.87	1 (7%)	17,19,21	0.82	0
7	A2G	b	201	2	14,14,15	0.71	0	17,19,21	0.90	1 (5%)

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	EDO	A	303	-	-	0/1/1/1	0/0/0/0
4	EDO	A	304	-	-	0/1/1/1	0/0/0/0
4	EDO	A	305	-	-	0/1/1/1	0/0/0/0
4	EDO	A	306	-	-	0/1/1/1	0/0/0/0
4	EDO	A	307	-	-	0/1/1/1	0/0/0/0
4	EDO	B	303	-	-	0/1/1/1	0/0/0/0
4	EDO	B	304	-	-	0/1/1/1	0/0/0/0
4	EDO	B	305	-	-	0/1/1/1	0/0/0/0
5	GOL	B	306	-	-	0/4/4/4	0/0/0/0
4	EDO	B	308	-	-	0/1/1/1	0/0/0/0
4	EDO	B	309	-	-	0/1/1/1	0/0/0/0
7	A2G	a	201	2	-	0/6/23/26	0/1/1/1
7	A2G	b	201	2	-	0/6/23/26	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	308	EDO	O2-C2	-4.79	1.17	1.42
4	B	308	EDO	O1-C1	-2.93	1.26	1.42
7	a	201	A2G	O-C5	2.28	1.48	1.43
5	B	306	GOL	O2-C2	2.29	1.50	1.43

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	306	GOL	C3-C2-C1	-2.99	100.12	111.63
4	A	304	EDO	O2-C2-C1	-2.80	92.32	112.09
7	b	201	A2G	C1-O-C5	2.06	115.03	112.19
5	B	306	GOL	O2-C2-C3	2.28	119.40	109.00
5	B	306	GOL	O2-C2-C1	2.50	120.42	109.00

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	304	EDO	2	0
4	A	306	EDO	7	0
4	B	308	EDO	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 [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	229/242 (94%)	0.38	11 (4%) 30 31	12, 15, 30, 64	0
1	B	229/242 (94%)	0.28	10 (4%) 34 35	11, 16, 34, 71	0
2	a	3/5 (60%)	3.65	3 (100%) 0 0	25, 25, 26, 38	0
2	b	4/5 (80%)	2.61	2 (50%) 0 0	22, 24, 30, 49	0
All	All	465/494 (94%)	0.37	26 (5%) 24 25	11, 16, 32, 71	0

The worst 5 of 26 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	229	ASN	6.9
1	B	229	ASN	6.2
2	a	103	VAL	5.6
2	b	104	GLY	5.1
1	A	131	ARG	3.9

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 carbohydrates 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
5	GOL	B	306	6/6	0.84	0.34	19,26,30,32	6
4	EDO	A	306	4/4	0.86	0.24	25,32,40,40	0
4	EDO	A	303	4/4	0.87	0.10	25,31,34,34	0
4	EDO	B	303	4/4	0.89	0.10	23,30,31,33	0
4	EDO	A	305	4/4	0.91	0.13	36,41,42,49	0
4	EDO	B	305	4/4	0.92	0.16	28,43,55,55	0
4	EDO	B	304	4/4	0.94	0.07	25,25,26,28	0
4	EDO	A	304	4/4	0.94	0.13	20,28,30,52	0
4	EDO	B	308	4/4	0.95	0.13	18,19,23,26	0
4	EDO	A	307	4/4	0.97	0.10	19,22,25,26	0
4	EDO	B	309	4/4	0.97	0.10	20,23,24,26	0
7	A2G	b	201	14/15	0.97	0.06	13,14,16,18	0
7	A2G	a	201	14/15	0.97	0.07	13,14,17,20	0
3	CA	B	301	1/1	1.00	0.09	12,12,12,12	0
3	CA	A	302	1/1	1.00	0.07	12,12,12,12	0
3	CA	B	302	1/1	1.00	0.06	12,12,12,12	0
6	CL	B	307	1/1	1.00	0.04	17,17,17,17	0
3	CA	A	301	1/1	1.00	0.07	12,12,12,12	0

6.5 Other polymers ⓘ

There are no such residues in this entry.