WORLDWIDE PDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 29, 2024 – 07:13 AM EDT

PDB ID	:	3R1X
Title	:	Crystal structure of 2-oxo-3-deoxygalactonate kinase from Klebsiella pneumoniae
Authors	:	Michalska, K.; Cuff, M.E.; Tesar, C.; Feldmann, B.; Joachimiak, A.; Midwest Center for Structural Genomics (MCSG)
Deposited on	:	2011-03-11
Resolution	:	2.09 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:

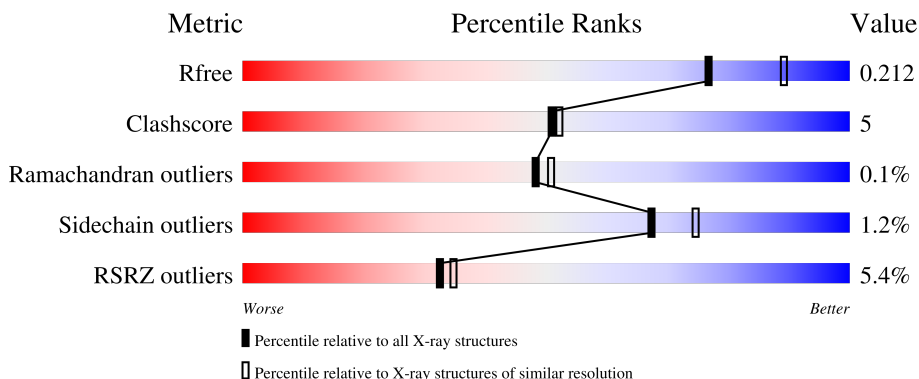|		
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 2.09 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	Percentile Ranks	Value
Rfree		0.212
Clashscore		5
Ramachandran outliers		0.1%
Sidechain outliers		1.2%
RSRZ outliers		5.4%

Worse *Better*

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	6234 (2.10-2.10)
Clashscore	180529	6893 (2.10-2.10)
Ramachandran outliers	177936	6839 (2.10-2.10)
Sidechain outliers	177891	6840 (2.10-2.10)
RSRZ outliers	164620	6234 (2.10-2.10)

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 <=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	295	4% / 92% / 7% ·
1	B	295	4% / 89% / 9% ··
1	C	295	4% / 90% / 9% ·
1	D	295	9% / 89% / 9% ·

2　Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 9789 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 2-oxo-3-deoxygalactonate kinase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	295	Total	C	N	O	S	Se	0	9	0
			2298	1440	426	422	4	6			
1	B	292	Total	C	N	O	S	Se	0	20	0
			2363	1484	436	435	3	5			
1	C	292	Total	C	N	O	S	Se	0	7	0
			2249	1411	404	425	4	5			
1	D	292	Total	C	N	O	S	Se	0	4	0
			2225	1399	401	416	4	5			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	expression tag	UNP A6TFZ6
A	-1	ASN	-	expression tag	UNP A6TFZ6
A	0	ALA	-	expression tag	UNP A6TFZ6
B	-2	SER	-	expression tag	UNP A6TFZ6
B	-1	ASN	-	expression tag	UNP A6TFZ6
B	0	ALA	-	expression tag	UNP A6TFZ6
C	-2	SER	-	expression tag	UNP A6TFZ6
C	-1	ASN	-	expression tag	UNP A6TFZ6
C	0	ALA	-	expression tag	UNP A6TFZ6
D	-2	SER	-	expression tag	UNP A6TFZ6
D	-1	ASN	-	expression tag	UNP A6TFZ6
D	0	ALA	-	expression tag	UNP A6TFZ6

- Molecule 2 is FORMIC ACID (three-letter code: FMT) (formula: CH_2O_2).

FMT

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

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).

GOL

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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	169	Total 169	O 169	0	1
4	B	177	Total 177	O 177	0	2
4	C	103	Total 103	O 103	0	0

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	D	91	Total O 91 91		0	0

3 Residue-property plots ⓘ

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.
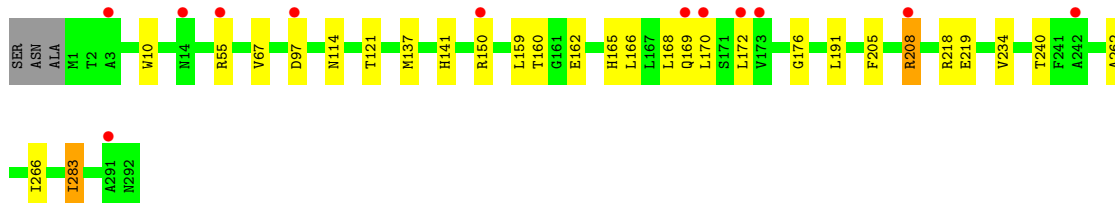
- Molecule 1: 2-oxo-3-deoxygalactonate kinase

Chain A:
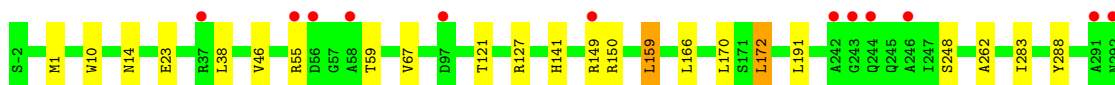
- Molecule 1: 2-oxo-3-deoxygalactonate kinase

Chain B:

- Molecule 1: 2-oxo-3-deoxygalactonate kinase
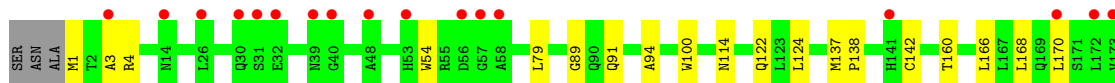
Chain C:

- Molecule 1: 2-oxo-3-deoxygalactonate kinase

Chain D:

G174 A175 G176 I195 T240 F241 A242 G243 Q244 Q245 A246 I247 V250 L255 A265 R268 E269 I286 A287 Y288 A291 N292

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	85.51Å 134.03Å 139.91Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	35.20 – 2.09 35.20 – 2.09	Depositor EDS
% Data completeness (in resolution range)	99.2 (35.20-2.09) 99.1 (35.20-2.09)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.60 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.174 , 0.199 0.186 , 0.212	Depositor DCC
R_{free} test set	1909 reflections (2.01%)	wwPDB-VP
Wilson B-factor (Å^2)	31.2	Xtriage
Anisotropy	0.225	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 50.5	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.011 for -h,l,k	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	9789	wwPDB-VP
Average B, all atoms (Å^2)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.20% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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: FMT, 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	0.76	0/2344	0.78	1/3178 (0.0%)
1	B	0.77	1/2408 (0.0%)	0.75	0/3269
1	C	0.68	0/2290	0.70	1/3113 (0.0%)
1	D	0.67	0/2269	0.72	0/3084
All	All	0.72	1/9311 (0.0%)	0.74	2/12644 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	137	MSE	SE-CE	-6.00	1.60	1.95

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	127	ARG	NE-CZ-NH1	5.16	122.88	120.30
1	C	208	ARG	NE-CZ-NH2	-5.06	117.77	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	2298	0	2299	27	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	2363	0	2365	46	0
1	C	2249	0	2219	18	0
1	D	2225	0	2211	18	0
2	A	6	0	2	0	0
2	D	6	0	2	0	0
3	A	12	0	16	0	0
3	B	48	0	64	7	0
3	C	6	0	8	0	0
3	D	36	0	48	1	0
4	A	169	0	0	2	0
4	B	177	0	0	1	0
4	C	103	0	0	1	0
4	D	91	0	0	0	0
All	All	9789	0	9234	95	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 95 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:166:LEU:HD21	1:B:170[B]:LEU:CD2	1.50	1.38
1:B:166:LEU:CD2	1:B:170[A]:LEU:HD22	1.81	1.09
1:A:166:LEU:CD2	1:B:170[B]:LEU:HD21	1.91	1.00
1:A:166:LEU:CD2	1:B:170[B]:LEU:CD2	2.41	0.98
1:A:166:LEU:HD21	1:B:170[B]:LEU:HD21	0.96	0.95

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	302/295 (102%)	296 (98%)	6 (2%)	0	100	100
1	B	311/295 (105%)	305 (98%)	6 (2%)	0	100	100
1	C	297/295 (101%)	292 (98%)	5 (2%)	0	100	100
1	D	294/295 (100%)	288 (98%)	5 (2%)	1 (0%)	37	37
All	All	1204/1180 (102%)	1181 (98%)	22 (2%)	1 (0%)	48	51

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	247	ILE

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	236/222 (106%)	233 (99%)	3 (1%)	65	72
1	B	242/222 (109%)	236 (98%)	6 (2%)	42	47
1	C	232/222 (104%)	231 (100%)	1 (0%)	89	93
1	D	229/222 (103%)	226 (99%)	3 (1%)	65	72
All	All	939/888 (106%)	926 (99%)	13 (1%)	67	70

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	219	GLU
1	B	283	ILE
1	D	269	GLU
1	D	114	ASN
1	D	244	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	141	HIS
1	D	91	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 ⓘ

21 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	FMT	A	294	-	2,2,2	0.65	0	1,1,1	0.39	0
3	GOL	D	298	-	5,5,5	0.44	0	5,5,5	0.50	0
3	GOL	A	296	-	5,5,5	0.23	0	5,5,5	0.63	0
3	GOL	B	295	-	5,5,5	0.44	0	5,5,5	0.95	0
3	GOL	B	296	-	5,5,5	0.42	0	5,5,5	0.62	0
3	GOL	B	299	-	5,5,5	0.62	0	5,5,5	0.53	0
3	GOL	B	294	-	5,5,5	0.34	0	5,5,5	0.54	0
2	FMT	A	293	-	2,2,2	0.66	0	1,1,1	0.33	0
3	GOL	B	297	-	5,5,5	0.46	0	5,5,5	0.66	0
3	GOL	B	298	-	5,5,5	0.35	0	5,5,5	0.46	0
3	GOL	B	300	-	5,5,5	0.39	0	5,5,5	0.49	0
3	GOL	A	295	-	5,5,5	0.38	0	5,5,5	0.27	0
3	GOL	D	296	-	5,5,5	0.27	0	5,5,5	0.82	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$\|Z\|>2$	Counts	RMSZ	#$\|Z\|>2$
3	GOL	B	293	-	5,5,5	0.29	0	5,5,5	1.31	1 (20%)
3	GOL	D	294	-	5,5,5	0.38	0	5,5,5	0.55	0
2	FMT	D	300	-	2,2,2	0.68	0	1,1,1	0.41	0
3	GOL	C	293	-	5,5,5	0.40	0	5,5,5	0.31	0
3	GOL	D	295	-	5,5,5	0.33	0	5,5,5	0.46	0
2	FMT	D	293	-	2,2,2	0.67	0	1,1,1	0.37	0
3	GOL	D	299	-	5,5,5	0.42	0	5,5,5	0.70	0
3	GOL	D	297	-	5,5,5	0.42	0	5,5,5	0.58	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
3	GOL	B	295	-	-	2/4/4/4	-
3	GOL	B	297	-	-	3/4/4/4	-
3	GOL	B	296	-	-	0/4/4/4	-
3	GOL	C	293	-	-	4/4/4/4	-
3	GOL	D	295	-	-	4/4/4/4	-
3	GOL	B	298	-	-	1/4/4/4	-
3	GOL	B	299	-	-	0/4/4/4	-
3	GOL	A	295	-	-	0/4/4/4	-
3	GOL	B	300	-	-	2/4/4/4	-
3	GOL	D	298	-	-	0/4/4/4	-
3	GOL	D	296	-	-	2/4/4/4	-
3	GOL	A	296	-	-	2/4/4/4	-
3	GOL	D	299	-	-	0/4/4/4	-
3	GOL	B	293	-	-	2/4/4/4	-
3	GOL	B	294	-	-	2/4/4/4	-
3	GOL	D	294	-	-	3/4/4/4	-
3	GOL	D	297	-	-	0/4/4/4	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	293	GOL	O2-C2-C1	2.18	118.20	109.18

There are no chirality outliers.

5 of 27 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	293	GOL	C1-C2-C3-O3
3	B	297	GOL	O1-C1-C2-C3
3	B	295	GOL	O1-C1-C2-C3
3	C	293	GOL	O1-C1-C2-C3
3	C	293	GOL	C1-C2-C3-O3

There are no ring outliers.

4 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	295	GOL	3	0
3	B	296	GOL	1	0
3	D	296	GOL	1	0
3	B	293	GOL	3	0

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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	290/295 (98%)	0.11	12 (4%)	42	44	23, 42, 61, 87	8 (2%)
1	B	287/295 (97%)	0.10	12 (4%)	41	43	12, 41, 59, 83	20 (6%)
1	C	287/295 (97%)	0.45	11 (3%)	44	47	21, 53, 82, 110	7 (2%)
1	D	287/295 (97%)	0.67	27 (9%)	15	16	22, 56, 95, 120	4 (1%)
All	All	1151/1180 (97%)	0.33	62 (5%)	32	35	12, 47, 80, 120	39 (3%)

The worst 5 of 62 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	172	LEU	5.8
1	B	172[A]	LEU	5.3
1	D	246	ALA	5.2
1	B	170[A]	LEU	4.8
1	C	243	GLY	4.2

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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($Å^2$)	Q<0.9
3	GOL	B	296	6/6	0.83	0.16	63,65,67,69	0
3	GOL	B	297	6/6	0.83	0.20	81,81,82,83	0
2	FMT	A	293	3/3	0.84	0.21	77,77,77,77	0
3	GOL	B	299	6/6	0.85	0.17	57,66,67,69	0
3	GOL	B	298	6/6	0.86	0.24	80,82,83,83	0
2	FMT	D	300	3/3	0.86	0.15	82,82,83,83	0
3	GOL	B	300	6/6	0.86	0.21	76,78,79,80	0
3	GOL	D	294	6/6	0.87	0.15	74,76,77,78	0
3	GOL	D	295	6/6	0.87	0.17	73,75,76,76	0
3	GOL	D	297	6/6	0.87	0.15	59,62,67,68	0
3	GOL	D	298	6/6	0.87	0.21	94,94,95,95	0
3	GOL	A	296	6/6	0.88	0.17	60,67,69,69	0
3	GOL	C	293	6/6	0.88	0.16	80,81,81,81	0
3	GOL	B	293	6/6	0.89	0.14	56,56,59,60	0
3	GOL	D	296	6/6	0.89	0.17	67,69,70,72	0
3	GOL	B	294	6/6	0.90	0.15	63,64,65,67	0
3	GOL	A	295	6/6	0.90	0.16	74,76,77,79	0
3	GOL	D	299	6/6	0.90	0.13	66,72,72,74	0
2	FMT	D	293	3/3	0.91	0.15	65,65,66,66	0
2	FMT	A	294	3/3	0.91	0.14	67,67,69,70	0
3	GOL	B	295	6/6	0.92	0.12	48,57,60,61	0

6.5 Other polymers ⓘ

There are no such residues in this entry.