Full wwPDB X-ray Structure Validation Report (i)

May 23, 2023 – 01:28 pm BST

PDB ID	:	8OPZ
Title	:	Crystal structure of a tailspike depolymerase (APK16_gp47) from Acinetobacter phage APK16
Authors	:	Matyuta, I.O.; Boyko, K.M.; Nikolaeva, A.Y.; Shneider, M.M.; Timoshina, O.Y.; Miroshnikov, K.A.; Popov, V.O.
Deposited on	:	2023-04-10
Resolution	:	1.50 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 (i) 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 (i)) were used in the production of this report:

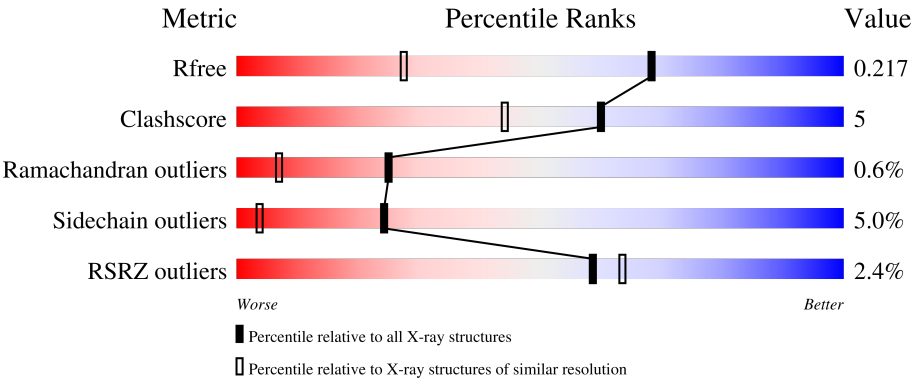
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.33
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.33

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.50 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}	130704	2936 (1.50-1.50)
Clashscore	141614	3144 (1.50-1.50)
Ramachandran outliers	138981	3066 (1.50-1.50)
Sidechain outliers	138945	3064 (1.50-1.50)
RSRZ outliers	127900	2884 (1.50-1.50)

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	641	2% 77% 19%

2 Entry composition

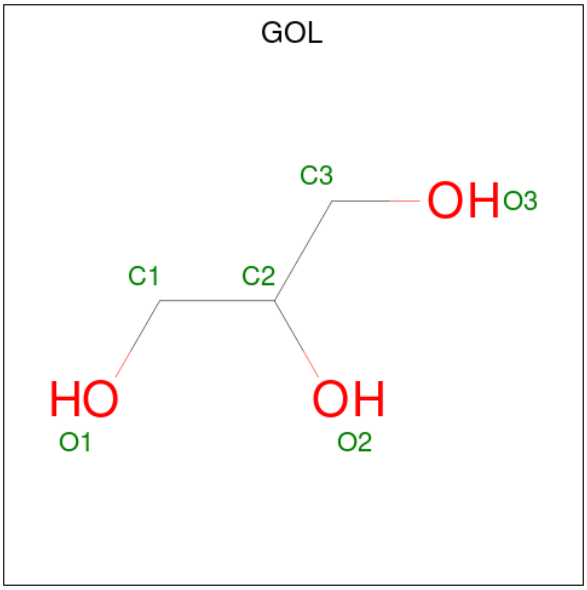
There are 3 unique types of molecules in this entry. The entry contains 4817 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Tailspike depolymerase (APK16_gp47) from Acinetobacter phage APK16.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	632	4733	2955	817	942	19	0	4	0

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

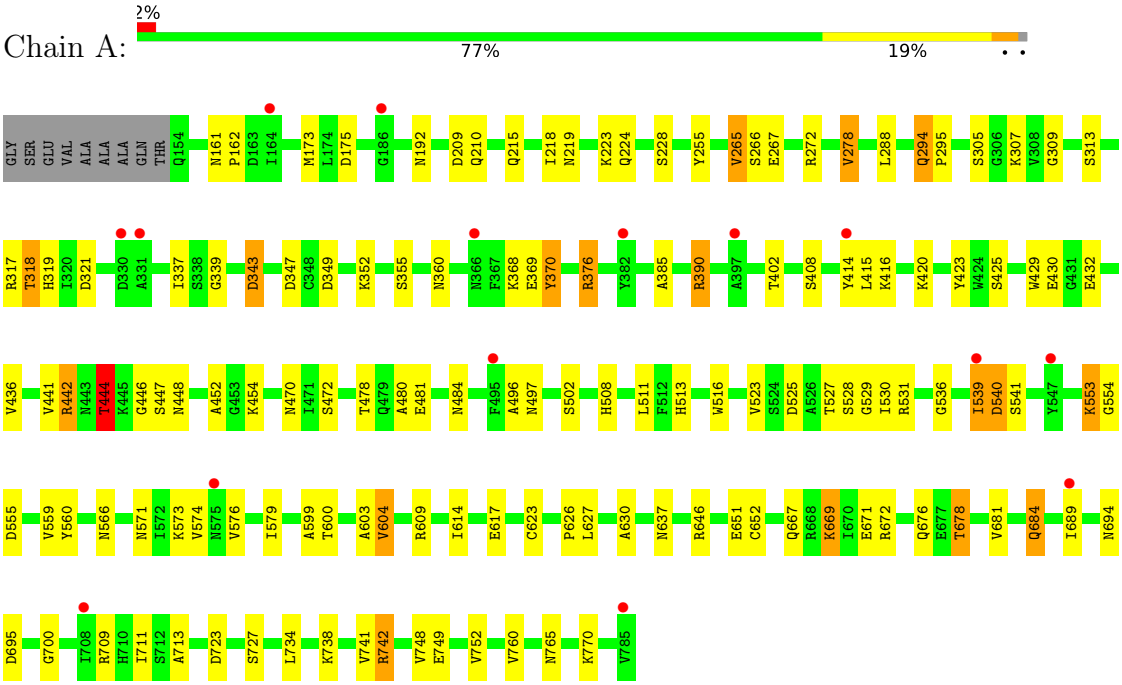
- Molecule 3 is water.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O			
3	A	78	78	78		0	0

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: Tailspike depolymerase (APK16_gp47) from Acinetobacter phage APK16



4 Data and refinement statistics (i)

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α, β, γ	88.05Å 88.05Å 254.97Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	84.99 – 1.50 84.99 – 1.50	Depositor EDS
% Data completeness (in resolution range)	99.8 (84.99-1.50) 99.8 (84.99-1.50)	Depositor EDS
R_{merge}	0.25	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.98 (at 1.50Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.175 , 0.215 0.178 , 0.217	Depositor DCC
R_{free} test set	5830 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å^2)	23.3	Xtriage
Anisotropy	0.242	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 16.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.38$, $< L^2 > = 0.20$	Xtriage
Estimated twinning fraction	0.417 for -h-k,k,-l	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	4817	wwPDB-VP
Average B, all atoms (Å^2)	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.97% 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

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	0.86	1/4836 (0.0%)	1.69	64/6576 (1.0%)

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	9

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	651	GLU	CD-OE2	-5.49	1.19	1.25

All (64) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	390	ARG	NE-CZ-NH2	-10.76	114.92	120.30
1	A	390	ARG	NE-CZ-NH1	10.18	125.39	120.30
1	A	390	ARG	CD-NE-CZ	9.62	137.07	123.60
1	A	531	ARG	NE-CZ-NH2	-9.60	115.50	120.30
1	A	531	ARG	NE-CZ-NH1	7.79	124.19	120.30
1	A	749	GLU	CB-CA-C	7.76	125.92	110.40
1	A	709	ARG	NE-CZ-NH1	7.44	124.02	120.30
1	A	390	ARG	CG-CD-NE	7.39	127.31	111.80
1	A	511	LEU	O-C-N	7.36	134.48	122.70
1	A	637	ASN	CB-CA-C	7.00	124.40	110.40
1	A	441	VAL	CG1-CB-CG2	-6.95	99.79	110.90
1	A	423	TYR	CB-CG-CD2	-6.64	117.01	121.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	343	ASP	CB-CA-C	6.64	123.68	110.40
1	A	278	VAL	CA-CB-CG1	6.49	120.63	110.90
1	A	667	GLN	CB-CA-C	6.39	123.18	110.40
1	A	671	GLU	CB-CA-C	6.38	123.16	110.40
1	A	265	VAL	CA-CB-CG1	6.34	120.42	110.90
1	A	370	TYR	CB-CG-CD2	-6.33	117.20	121.00
1	A	609	ARG	NE-CZ-NH1	6.30	123.45	120.30
1	A	349	ASP	CB-CA-C	6.23	122.86	110.40
1	A	752	VAL	CA-CB-CG2	6.13	120.09	110.90
1	A	531	ARG	CG-CD-NE	6.11	124.62	111.80
1	A	559	VAL	CA-CB-CG2	6.09	120.03	110.90
1	A	442	ARG	NE-CZ-NH1	6.03	123.31	120.30
1	A	224	GLN	CB-CA-C	5.99	122.38	110.40
1	A	272[A]	ARG	CB-CA-C	5.97	122.34	110.40
1	A	272[B]	ARG	CB-CA-C	5.97	122.34	110.40
1	A	604	VAL	CA-CB-CG2	5.90	119.74	110.90
1	A	376	ARG	NE-CZ-NH2	-5.89	117.36	120.30
1	A	525	ASP	CB-CG-OD2	-5.86	113.03	118.30
1	A	748	VAL	CA-CB-CG2	5.82	119.63	110.90
1	A	600	THR	C-N-CA	5.75	134.38	122.30
1	A	672	ARG	NE-CZ-NH1	5.75	123.17	120.30
1	A	672	ARG	NE-CZ-NH2	-5.71	117.44	120.30
1	A	529	GLY	O-C-N	5.69	131.80	122.70
1	A	560	TYR	CB-CG-CD2	-5.67	117.60	121.00
1	A	741	VAL	CA-CB-CG1	5.64	119.36	110.90
1	A	368	LYS	CB-CA-C	5.62	121.64	110.40
1	A	695	ASP	CB-CA-C	5.61	121.61	110.40
1	A	352	LYS	N-CA-CB	5.45	120.42	110.60
1	A	555	ASP	CB-CA-C	-5.44	99.52	110.40
1	A	436	VAL	CA-CB-CG2	5.39	118.99	110.90
1	A	752	VAL	CA-CB-CG1	5.39	118.99	110.90
1	A	630	ALA	CB-CA-C	5.37	118.16	110.10
1	A	321	ASP	CB-CA-C	5.36	121.11	110.40
1	A	603	ALA	CB-CA-C	5.35	118.13	110.10
1	A	723	ASP	CB-CA-C	5.32	121.04	110.40
1	A	609	ARG	CG-CD-NE	5.30	122.94	111.80
1	A	319	HIS	CB-CA-C	5.25	120.90	110.40
1	A	444	THR	N-CA-CB	-5.24	100.34	110.30
1	A	480	ALA	CB-CA-C	5.23	117.95	110.10
1	A	540	ASP	N-CA-CB	5.22	120.00	110.60
1	A	444	THR	CA-CB-OG1	5.22	119.96	109.00
1	A	402	THR	CA-CB-CG2	5.21	119.70	112.40

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	317	ARG	CD-NE-CZ	5.20	130.88	123.60
1	A	376	ARG	NE-CZ-NH1	5.20	122.90	120.30
1	A	470	ASN	CB-CA-C	5.17	120.75	110.40
1	A	614	ILE	CA-C-N	-5.15	105.86	117.20
1	A	337	ILE	N-CA-CB	5.14	122.62	110.80
1	A	553	LYS	CB-CA-C	5.13	120.67	110.40
1	A	385	ALA	CB-CA-C	5.09	117.73	110.10
1	A	684	GLN	CB-CA-C	5.04	120.47	110.40
1	A	760	VAL	CA-CB-CG2	5.02	118.43	110.90
1	A	742	ARG	CB-CG-CD	5.01	124.63	111.60

There are no chirality outliers.

All (9) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	173	MET	Peptide
1	A	294	GLN	Peptide
1	A	539	ILE	Peptide
1	A	579	ILE	Peptide
1	A	669	LYS	Peptide
1	A	676	GLN	Peptide
1	A	700	GLY	Peptide
1	A	713	ALA	Peptide
1	A	738	LYS	Peptide

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	4733	0	4575	43	0
2	A	6	0	8	0	0
3	A	78	0	0	0	0
All	All	4817	0	4583	43	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.

All (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:360:ASN:HD21	1:A:369:GLU:HA	1.54	0.73
1:A:484:ASN:HD22	1:A:508:HIS:H	1.39	0.70
1:A:318:THR:HG22	1:A:355:SER:HB3	1.79	0.64
1:A:209:ASP:HB2	1:A:218:ILE:HD11	1.79	0.64
1:A:318:THR:HG21	1:A:339:GLY:CA	2.34	0.57
1:A:444:THR:HG21	1:A:448:ASN:HA	1.87	0.56
1:A:536:GLY:HA2	1:A:571:ASN:O	2.07	0.55
1:A:228:SER:HA	1:A:255:TYR:O	2.07	0.54
1:A:496:ALA:HA	1:A:528:SER:O	2.07	0.53
1:A:318:THR:HG21	1:A:339:GLY:HA3	1.90	0.52
1:A:309:GLY:HA2	1:A:347:ASP:O	2.10	0.52
1:A:523:VAL:HG22	1:A:553:LYS:HG3	1.90	0.52
1:A:623:CYS:O	1:A:652:CYS:HA	2.10	0.52
1:A:604:VAL:O	1:A:627:LEU:HA	2.10	0.52
1:A:278:VAL:HG22	1:A:313:SER:HB3	1.92	0.51
1:A:541:SER:HB3	1:A:576:VAL:HG22	1.91	0.51
1:A:414:TYR:HE1	1:A:416:LYS:HE3	1.76	0.50
1:A:617:GLU:HA	1:A:646:ARG:O	2.12	0.50
1:A:215:GLN:HE21	1:A:219:ASN:HD21	1.59	0.50
1:A:420:LYS:HA	1:A:442:ARG:O	2.14	0.48
1:A:175:ASP:H	1:A:192:ASN:HD21	1.62	0.47
1:A:484:ASN:ND2	1:A:508:HIS:H	2.09	0.47
1:A:454:LYS:HA	1:A:481:GLU:O	2.14	0.47
1:A:530:ILE:HG21	1:A:539:ILE:HG12	1.97	0.47
1:A:305:SER:HA	1:A:343:ASP:O	2.16	0.46
1:A:430:GLU:HG2	1:A:452:ALA:HB3	1.97	0.46
1:A:574:VAL:HG13	1:A:576:VAL:HG23	1.99	0.45
1:A:209:ASP:HB2	1:A:218:ILE:CD1	2.45	0.45
1:A:527:THR:O	1:A:554:GLY:HA3	2.18	0.44
1:A:681:VAL:H	1:A:684:GLN:HE21	1.65	0.43
1:A:727:SER:HB2	1:A:734:LEU:HG	2.01	0.43
1:A:360:ASN:ND2	1:A:370:TYR:H	2.16	0.43
1:A:425:SER:HA	1:A:447:SER:O	2.19	0.42
1:A:599:ALA:O	1:A:623:CYS:HA	2.20	0.42
1:A:681:VAL:H	1:A:684:GLN:NE2	2.18	0.42
1:A:497:ASN:ND2	1:A:516:TRP:HB3	2.35	0.41
1:A:513:HIS:HE1	1:A:541:SER:OG	2.03	0.41
1:A:415:LEU:HB2	1:A:429:TRP:CZ2	2.55	0.41
1:A:343:ASP:HA	1:A:376:ARG:O	2.20	0.41
1:A:161:ASN:HA	1:A:162:PRO:HD3	1.95	0.40
1:A:415:LEU:HB2	1:A:429:TRP:CH2	2.57	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:444:THR:HG22	1:A:446:GLY:O	2.21	0.40
1:A:432:GLU:HA	1:A:454:LYS:O	2.21	0.40

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	634/641 (99%)	596 (94%)	34 (5%)	4 (1%)	25 7

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	295	PRO
1	A	678	THR
1	A	288	LEU
1	A	626	PRO

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	501/511 (98%)	476 (95%)	25 (5%)	24 4

All (25) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	210	GLN
1	A	223	LYS
1	A	265	VAL
1	A	266	SER
1	A	267	GLU
1	A	294	GLN
1	A	307	LYS
1	A	318	THR
1	A	390	ARG
1	A	408	SER
1	A	444	THR
1	A	472	SER
1	A	478	THR
1	A	502	SER
1	A	540	ASP
1	A	566	ASN
1	A	573	LYS
1	A	669	LYS
1	A	678	THR
1	A	689	ILE
1	A	694	ASN
1	A	711	ILE
1	A	742	ARG
1	A	765	ASN
1	A	770	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	192	ASN
1	A	219	ASN
1	A	360	ASN
1	A	448	ASN
1	A	470	ASN
1	A	484	ASN
1	A	497	ASN
1	A	501	HIS
1	A	508	HIS
1	A	513	HIS
1	A	684	GLN
1	A	765	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

1 ligand is 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	A	801	-	5,5,5	0.44	0	5,5,5	0.81	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	A	801	-	-	4/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	801	GOL	O1-C1-C2-C3
2	A	801	GOL	C1-C2-C3-O3
2	A	801	GOL	O2-C2-C3-O3
2	A	801	GOL	O1-C1-C2-O2

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

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, 95^{th} 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	632/641 (98%)	0.39	15 (2%) 59 63	19, 28, 42, 54	0

All (15) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	397	ALA	4.4
1	A	785	VAL	3.6
1	A	331	ALA	3.3
1	A	414	TYR	3.2
1	A	366	ASN	3.0
1	A	382	TYR	2.8
1	A	689	ILE	2.5
1	A	164	ILE	2.5
1	A	708	ILE	2.4
1	A	495	PHE	2.4
1	A	575	ASN	2.2
1	A	547	TYR	2.2
1	A	186	GLY	2.2
1	A	330	ASP	2.1
1	A	539	ILE	2.1

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, 95^{th} 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
2	GOL	A	801	6/6	0.94	0.07	25,27,29,30	0

6.5 Other polymers

There are no such residues in this entry.