



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 16, 2024 – 10:24 am BST

PDB ID : 8QJA
Title : T6SS-linked Rhs repeat protein - Advenella mimigardefordensis VgrG-Rhs core
Authors : Kielkopf, C.S.; Shneider, M.M.; Leiman, P.G.; Taylor, N.M.I.
Deposited on : 2023-09-13
Resolution : 3.36 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
Xtriage (Phenix) : 1.13
EDS : 3.0
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.002 (Gargrove)
Density-Fitness : 1.0.11
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.38.2

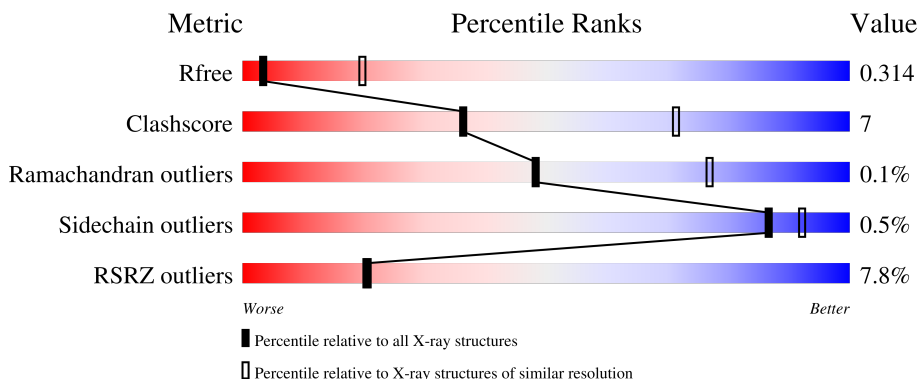
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.36 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}	164625	1012 (3.40-3.32)
Clashscore	180529	1035 (3.40-3.32)
Ramachandran outliers	177936	1037 (3.40-3.32)
Sidechain outliers	177891	1037 (3.40-3.32)
RSRZ outliers	164620	1012 (3.40-3.32)

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	1879	5% (Poor fit) 43% (0 outliers) 11% (1 outlier) 46% (Not modelled)
1	B	1879	4% (Poor fit) 44% (0 outliers) 10% (1 outlier) 45% (Not modelled)
1	C	1879	4% (Poor fit) 44% (0 outliers) 10% (1 outlier) 46% (Not modelled)
1	D	1879	5% (Poor fit) 44% (0 outliers) 10% (1 outlier) 46% (Not modelled)

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 33470 atoms, of which 0 are hydrogens and 0 are deuteriums.

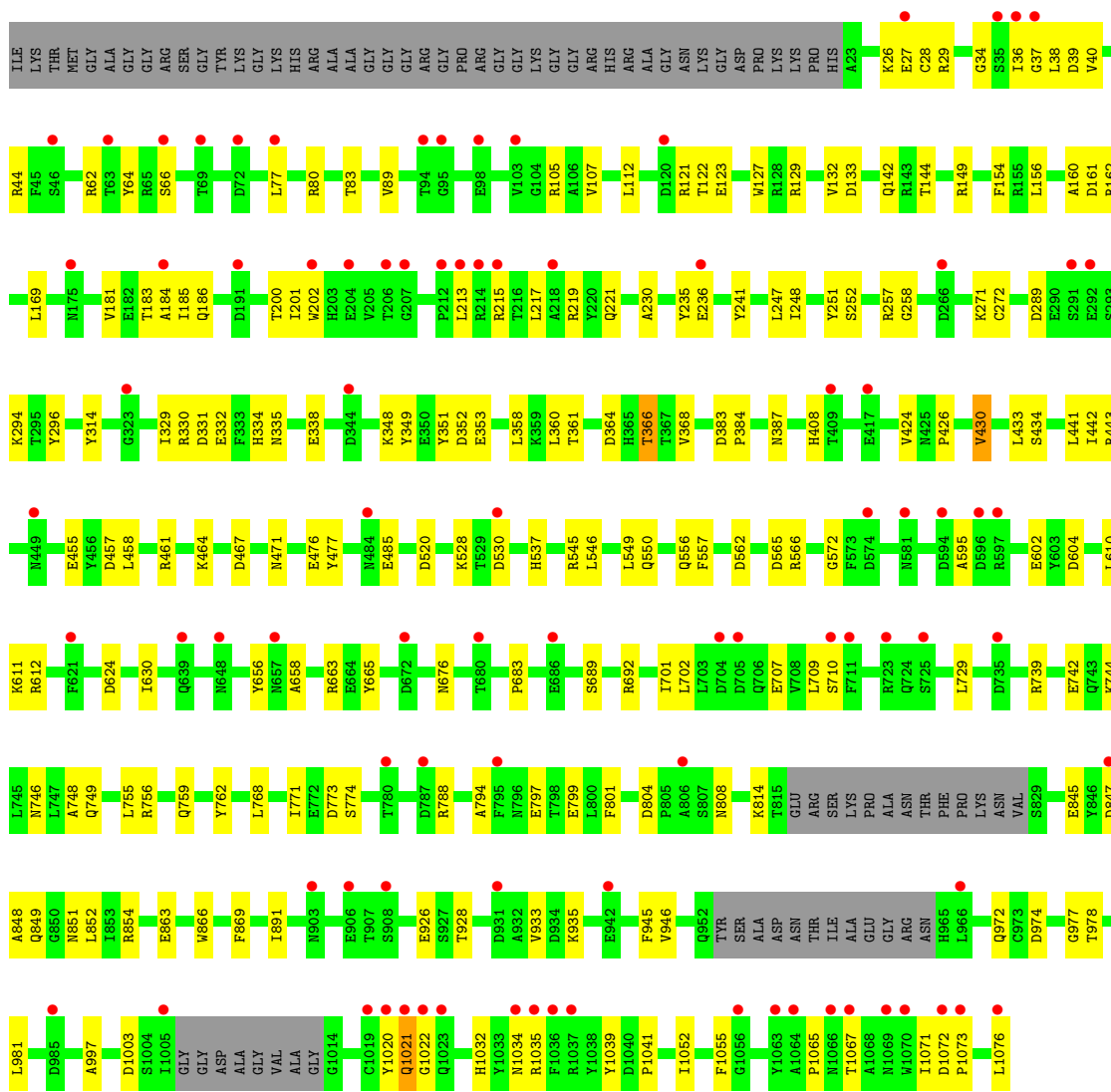
In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Putative type VI secretion system YD repeat-containing Rhs element Vgr protein.

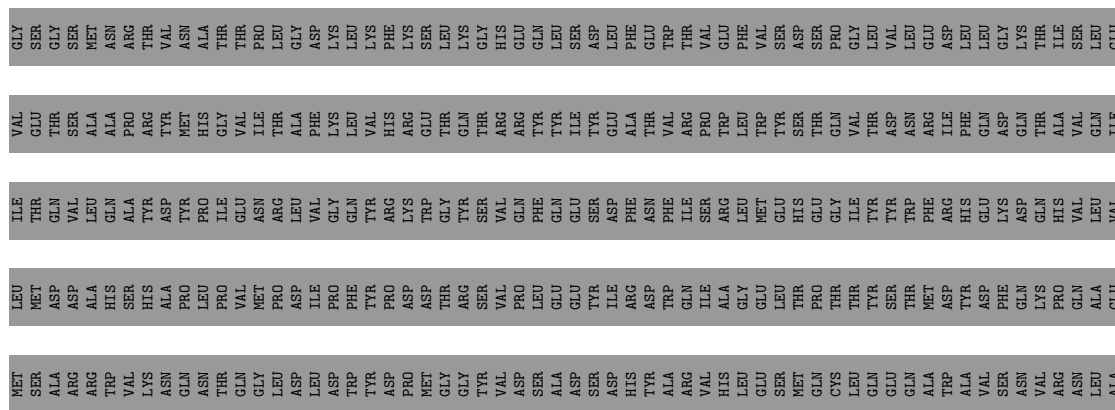
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	1021	8365	5240	1479	1631	15	0	0	0
1	B	1026	8390	5255	1484	1636	15	0	0	0
1	C	1020	8355	5233	1478	1629	15	0	0	0
1	D	1021	8360	5236	1479	1630	15	0	0	0

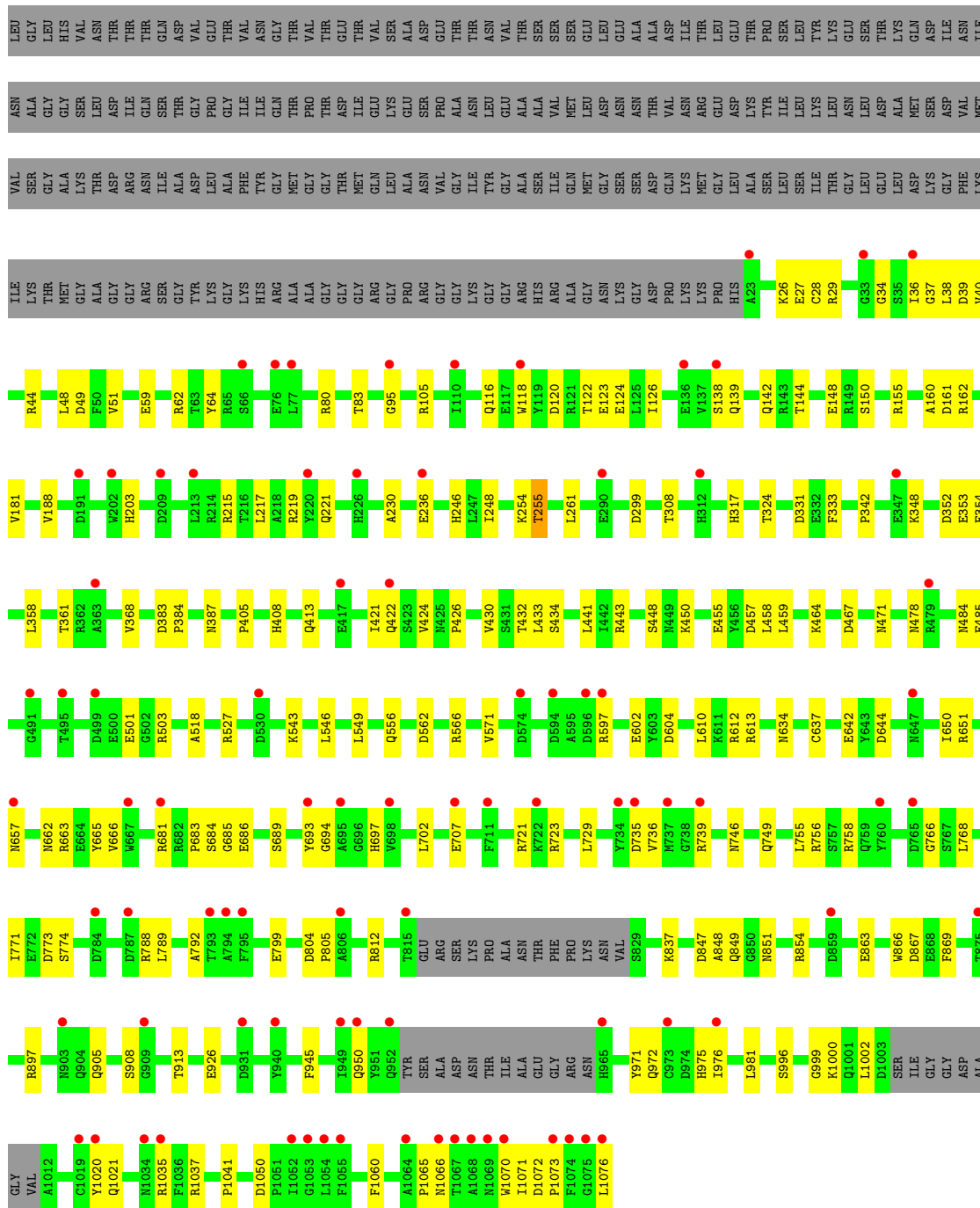
There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-802	GLY	-	expression tag	UNP W0PIA9
A	-801	SER	-	expression tag	UNP W0PIA9
A	-800	GLY	-	expression tag	UNP W0PIA9
A	-799	SER	-	expression tag	UNP W0PIA9
B	-802	GLY	-	expression tag	UNP W0PIA9
B	-801	SER	-	expression tag	UNP W0PIA9
B	-800	GLY	-	expression tag	UNP W0PIA9
B	-799	SER	-	expression tag	UNP W0PIA9
C	-802	GLY	-	expression tag	UNP W0PIA9
C	-801	SER	-	expression tag	UNP W0PIA9
C	-800	GLY	-	expression tag	UNP W0PIA9
C	-799	SER	-	expression tag	UNP W0PIA9
D	-802	GLY	-	expression tag	UNP W0PIA9
D	-801	SER	-	expression tag	UNP W0PIA9
D	-800	GLY	-	expression tag	UNP W0PIA9
D	-799	SER	-	expression tag	UNP W0PIA9



● Molecule 1: Putative type VI secretion system YD repeat-containing Rhs element Vgr protein





4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	66.24Å 97.78Å 215.01Å 85.69° 89.99° 76.89°	Depositor
Resolution (Å)	48.48 – 3.36 48.48 – 3.36	Depositor EDS
% Data completeness (in resolution range)	91.2 (48.48-3.36) 93.9 (48.48-3.36)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.24 (at 3.33Å)	Xtrriage
Refinement program	PHENIX 1.18.2_3874, PHENIX 1.20.1_4487, BUSTER	Depositor
R, R_{free}	0.257 , 0.317 0.258 , 0.314	Depositor DCC
R_{free} test set	3730 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	53.4	Xtrriage
Anisotropy	0.874	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 47.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.87	EDS
Total number of atoms	33470	wwPDB-VP
Average B, all atoms (Å ²)	65.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.44% 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

5.1 Standard geometry

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	0.25	0/8565	0.53	0/11607
1	B	0.25	0/8590	0.54	0/11641
1	C	0.25	0/8555	0.53	0/11593
1	D	0.25	0/8560	0.53	0/11600
All	All	0.25	0/34270	0.53	0/46441

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	8365	0	7857	122	0
1	B	8390	0	7882	114	0
1	C	8355	0	7844	115	0
1	D	8360	0	7849	113	0
All	All	33470	0	31432	458	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 7.

The worst 5 of 458 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:29:ARG:NH2	1:B:121:ARG:O	2.09	0.85
1:B:55:VAL:HG11	1:B:417:GLU:HB3	1.57	0.84
1:D:299:ASP:HB2	1:D:543:LYS:HE3	1.59	0.83
1:A:219:ARG:HH22	1:A:236:GLU:HG2	1.47	0.80
1:C:117:GLU:HG2	1:C:128:ARG:HG2	1.65	0.79

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	1013/1879 (54%)	969 (96%)	44 (4%)	0	100	100
1	B	1018/1879 (54%)	971 (95%)	47 (5%)	0	100	100
1	C	1012/1879 (54%)	966 (96%)	45 (4%)	1 (0%)	48	77
1	D	1013/1879 (54%)	968 (96%)	43 (4%)	2 (0%)	44	71
All	All	4056/7516 (54%)	3874 (96%)	179 (4%)	3 (0%)	48	77

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	694	GLY
1	D	694	GLY
1	D	95	GLY

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	889/1607 (55%)	883 (99%)	6 (1%)	81	89
1	B	890/1607 (55%)	884 (99%)	6 (1%)	81	89
1	C	887/1607 (55%)	882 (99%)	5 (1%)	84	91
1	D	887/1607 (55%)	886 (100%)	1 (0%)	92	97
All	All	3553/6428 (55%)	3535 (100%)	18 (0%)	86	92

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	255	THR
1	D	255	THR
1	C	1021	GLN
1	B	366	THR
1	C	216	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	186	GLN
1	C	754	ASN
1	D	1069	ASN
1	C	970	HIS
1	B	1048	GLN

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](#)

There are no ligands in this entry.

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	1021/1879 (54%)	0.74	89 (8%) 17 18	30, 56, 98, 228	0
1	B	1026/1879 (54%)	0.68	75 (7%) 22 22	30, 56, 99, 175	0
1	C	1020/1879 (54%)	0.78	69 (6%) 25 24	39, 68, 108, 156	0
1	D	1021/1879 (54%)	0.81	86 (8%) 18 19	38, 69, 111, 201	0
All	All	4088/7516 (54%)	0.75	319 (7%) 20 20	30, 62, 107, 228	0

The worst 5 of 319 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1035	ARG	6.6
1	D	1020	TYR	5.6
1	D	1067	THR	5.5
1	A	1073	PRO	5.3
1	A	1020	TYR	4.9

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 monosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.