W O R L D W I D E
wwPDB
P R O T E I N D A T A B A N K

Full wwPDB X-ray Structure Validation Report ⓘ

Oct 31, 2023 – 06:58 PM JST

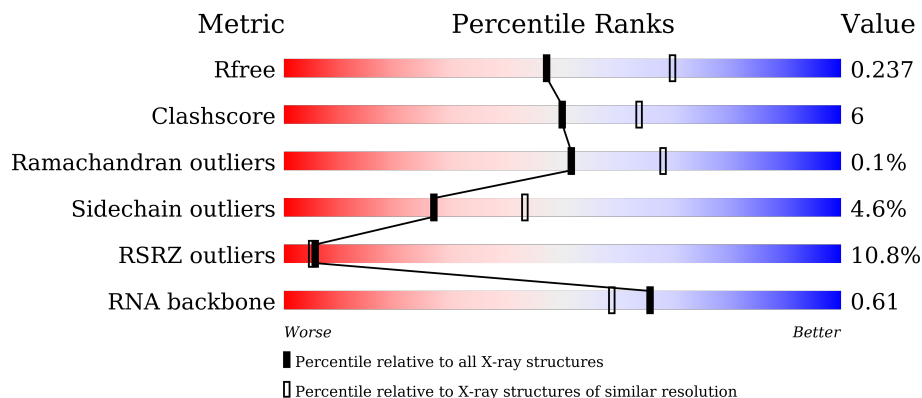|		
PDB ID	:	5GUH
Title	:	Crystal structure of silkworm PIWI-clade Argonaute Siwi bound to piRNA
Authors	:	Matsumoto, N.; Nishimasu, H.; Ishitani, R.; Nureki, O.
Deposited on	:	2016-08-29
Resolution	:	2.40 Å(reported)

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 2.40 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		0.237
Clashscore		6
Ramachandran outliers		0.1%
Sidechain outliers		4.6%
RSRZ outliers		10.8%
RNA backbone		0.61

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$_{free}$	130704	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)
RNA backbone	3102	1174 (2.80-2.00)

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	899	9% · 71% · 12% · 16%
2	B	28	4% · 21% · · · 71%

WORLDWIDE
wwPDB
PROTEIN DATA BANK

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 6078 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 PIWI.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	759	Total 5828	C 3733	N 993	O 1074	S 28	0	0	0

- Molecule 2 is a RNA chain called RNA (28-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	8	Total 164	C 74	N 19	O 63	P 8	0	0	0

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	2	Total 2	Mg 2	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	83	Total 83	O 83	0	0
4	B	1	Total 1	O 1	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: PIWI

Chain A:
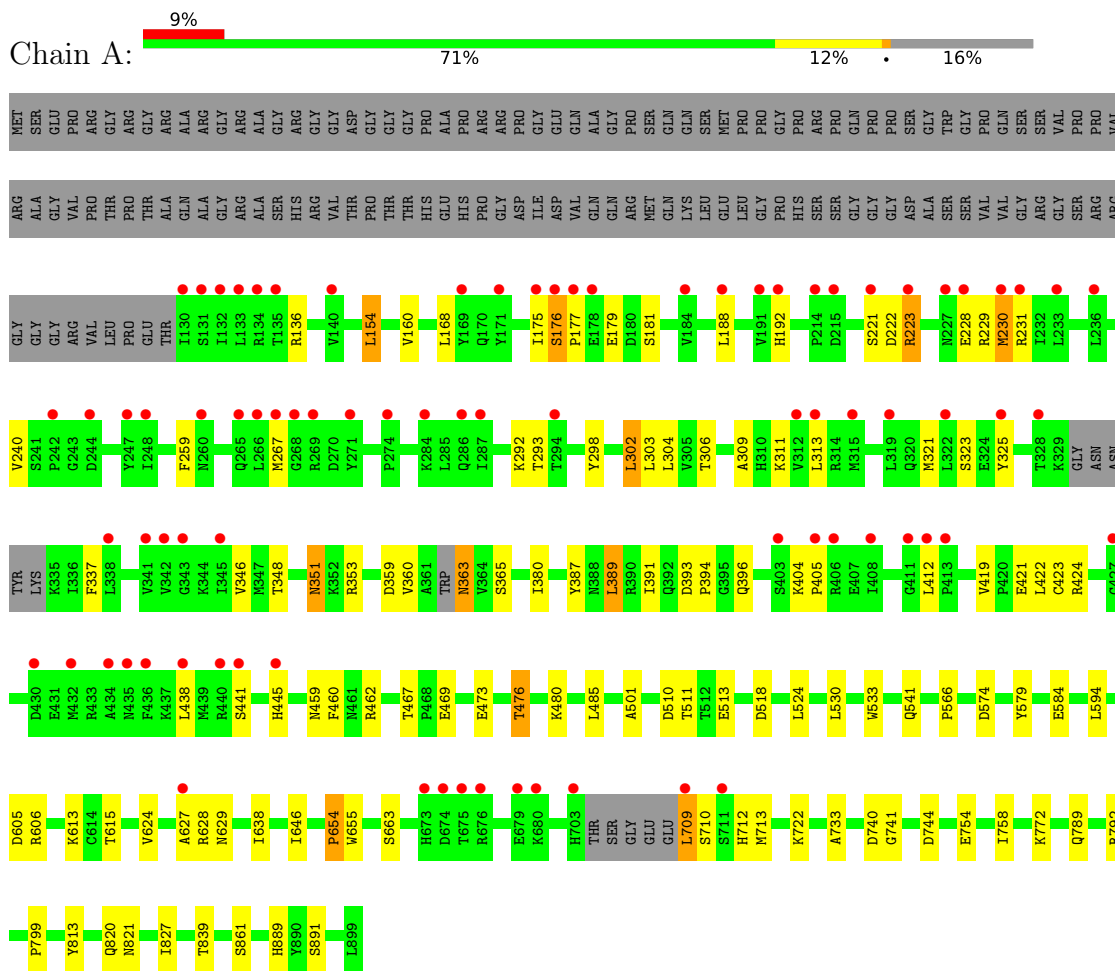
• Molecule 2: RNA (28-MER)

Chain B:
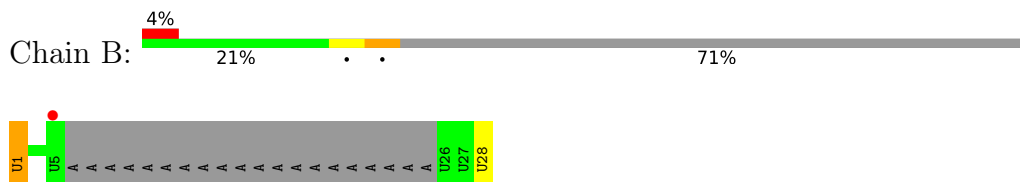
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	61.90Å 114.62Å 136.72Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	87.83 – 2.40 87.84 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.5 (87.83-2.40) 99.6 (87.84-2.40)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.68 (at 2.40Å)	Xtriage
Refinement program	PHENIX (1.10_2155: ???)	Depositor
R, R_{free}	0.209 , 0.235 0.210 , 0.237	Depositor DCC
R_{free} test set	1999 reflections (5.17%)	wwPDB-VP
Wilson B-factor (Å^2)	46.9	Xtriage
Anisotropy	0.319	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 59.3	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	6078	wwPDB-VP
Average B, all atoms (Å^2)	67.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.03% 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: MG, OMU

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	0/5952	0.43	0/8104
2	B	0.87	1/156 (0.6%)	0.64	0/235
All	All	0.28	1/6108 (0.0%)	0.44	0/8339

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1	U	OP3-P	-10.83	1.48	1.61

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	5828	0	5696	68	0
2	B	164	0	85	1	0
3	A	2	0	0	0	0
4	A	83	0	0	1	0
4	B	1	0	0	0	0
All	All	6078	0	5781	68	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 6.

All (68) 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:179:GLU:HG2	1:A:223:ARG:HE	1.47	0.80
1:A:136:ARG:NH1	4:A:1001:HOH:O	2.20	0.74
1:A:709:LEU:HB2	1:A:712:HIS:HB2	1.75	0.68
1:A:511:THR:HG22	1:A:513:GLU:H	1.59	0.67
1:A:889:HIS:HD2	1:A:891:SER:H	1.45	0.65
1:A:230:MET:SD	1:A:230:MET:N	2.70	0.65
1:A:380:ILE:HA	1:A:391:ILE:HD13	1.78	0.64
1:A:175:ILE:HG22	1:A:177:PRO:HD2	1.80	0.64
1:A:223:ARG:HB2	1:A:230:MET:HE3	1.80	0.63
1:A:293:THR:HG23	1:A:304:LEU:HD11	1.82	0.61
1:A:168:LEU:HB3	1:A:240:VAL:HB	1.83	0.60
1:A:510:ASP:HB2	1:A:799:PRO:HD2	1.85	0.58
1:A:348:THR:OG1	1:A:351:ASN:ND2	2.26	0.56
1:A:441:SER:O	1:A:445:HIS:ND1	2.31	0.56
1:A:710:SER:HA	1:A:754:GLU:HG2	1.89	0.55
1:A:223:ARG:H	1:A:230:MET:HE3	1.72	0.54
1:A:387:TYR:HB2	1:A:389:LEU:HD11	1.89	0.53
1:A:605:ASP:OD1	1:A:606:ARG:N	2.43	0.52
1:A:188:LEU:HD21	1:A:223:ARG:HG3	1.91	0.52
1:A:363:ASN:N	1:A:363:ASN:OD1	2.43	0.52
1:A:298:TYR:CD1	1:A:303:LEU:HB2	2.46	0.51
1:A:404:LYS:N	1:A:405:PRO:HD2	2.26	0.50
1:A:533:TRP:CE2	1:A:566:PRO:HB3	2.47	0.50
1:A:574:ASP:O	1:A:606:ARG:NH2	2.45	0.50
1:A:176:SER:HB2	1:A:231:ARG:HB3	1.93	0.50
1:A:223:ARG:HD3	1:A:230:MET:HE3	1.92	0.50
1:A:389:LEU:HD13	1:A:422:LEU:HD21	1.93	0.50
1:A:292:LYS:HG3	1:A:309:ALA:HB2	1.94	0.49
1:A:192:HIS:HE1	1:A:221:SER:HB2	1.78	0.49
1:A:359:ASP:OD1	1:A:360:VAL:N	2.44	0.49
1:A:459:ASN:OD1	1:A:462:ARG:NH1	2.42	0.49
1:A:321:MET:HG2	1:A:337:PHE:HE1	1.78	0.49
1:A:740:ASP:OD1	1:A:741:GLY:N	2.45	0.49
1:A:160:VAL:HB	1:A:302:LEU:HB3	1.94	0.48
1:A:579:TYR:CD2	1:A:606:ARG:HG2	2.50	0.47
1:A:365:SER:OG	1:A:394:PRO:O	2.33	0.46
1:A:629:ASN:HB3	1:A:638:ILE:HD13	1.98	0.46
1:A:306:THR:HG21	1:A:460:PHE:CE2	2.51	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:346:VAL:HG21	1:A:423:CYS:HB3	1.97	0.46
1:A:624:VAL:O	2:B:1:U:H4'	2.16	0.46
1:A:594:LEU:HD13	1:A:646:ILE:HG23	1.98	0.45
1:A:713:MET:HG3	1:A:758:ILE:HD11	1.98	0.45
1:A:821:ASN:HA	1:A:827:ILE:HD11	1.97	0.45
1:A:654:PRO:HB2	1:A:655:TRP:CE3	2.52	0.45
1:A:168:LEU:HD12	1:A:168:LEU:HA	1.83	0.45
1:A:467:THR:HG22	1:A:469:GLU:HG2	1.99	0.45
1:A:154:LEU:HD13	1:A:789:GLN:HB2	1.99	0.44
1:A:223:ARG:CB	1:A:230:MET:HE3	2.46	0.44
1:A:393:ASP:OD1	1:A:396:GLN:N	2.49	0.44
1:A:709:LEU:O	1:A:713:MET:N	2.44	0.44
1:A:663:SER:HB3	1:A:733:ALA:H	1.82	0.44
1:A:222:ASP:HA	1:A:230:MET:HG2	1.98	0.44
1:A:223:ARG:H	1:A:230:MET:CE	2.30	0.43
1:A:584:GLU:HB2	1:A:613:LYS:HD2	2.00	0.43
1:A:501:ALA:HB2	1:A:524:LEU:O	2.19	0.43
1:A:419:VAL:HG23	1:A:422:LEU:HB2	2.01	0.42
1:A:351:ASN:ND2	1:A:353:ARG:H	2.18	0.41
1:A:223:ARG:HD3	1:A:230:MET:CE	2.51	0.41
1:A:228:GLU:CD	1:A:229:ARG:H	2.23	0.41
1:A:541:GLN:HG2	1:A:627:ALA:HB1	2.03	0.41
1:A:473:GLU:O	1:A:476:THR:HG22	2.21	0.41
1:A:259:PHE:CE1	1:A:293:THR:HG21	2.55	0.41
1:A:311:LYS:HD2	1:A:311:LYS:HA	1.90	0.40
1:A:396:GLN:HE22	1:A:421:GLU:HB2	1.85	0.40
1:A:709:LEU:HD13	1:A:712:HIS:HB2	2.02	0.40
1:A:772:LYS:HB3	1:A:839:THR:HA	2.03	0.40
1:A:530:LEU:HD11	1:A:533:TRP:HB3	2.03	0.40
1:A:480:LYS:HB2	1:A:480:LYS:HE3	1.89	0.40

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	751/899 (84%)	718 (96%)	32 (4%)	1 (0%)	51	68

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	654	PRO

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	613/783 (78%)	585 (95%)	28 (5%)	27	43

All (28) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	154	LEU
1	A	176	SER
1	A	181	SER
1	A	223	ARG
1	A	230	MET
1	A	267	MET
1	A	302	LEU
1	A	313	LEU
1	A	323	SER
1	A	325	TYR
1	A	351	ASN
1	A	363	ASN
1	A	389	LEU
1	A	412	LEU
1	A	424	ARG
1	A	438	LEU
1	A	476	THR
1	A	485	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	518	ASP
1	A	615	THR
1	A	628	ARG
1	A	709	LEU
1	A	722	LYS
1	A	744	ASP
1	A	792	ARG
1	A	813	TYR
1	A	820	GLN
1	A	861	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (14) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	192	HIS
1	A	310	HIS
1	A	320	GLN
1	A	351	ASN
1	A	363	ASN
1	A	425	GLN
1	A	692	GLN
1	A	695	GLN
1	A	701	ASN
1	A	703	HIS
1	A	716	ASN
1	A	783	ASN
1	A	789	GLN
1	A	889	HIS

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	6/28 (21%)	0	0

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

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

1 non-standard protein/DNA/RNA residue 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	OMU	B	28	2	19,22,23	1.19	2 (10%)	26,31,34	1.73	5 (19%)

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	OMU	B	28	2	-	0/9/27/28	0/2/2/2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	28	OMU	C4-N3	-2.41	1.34	1.38
2	B	28	OMU	C2-N3	-2.09	1.34	1.38

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	28	OMU	C4-N3-C2	-4.55	120.57	126.58
2	B	28	OMU	N3-C2-N1	4.27	120.56	114.89
2	B	28	OMU	C5-C4-N3	3.50	120.07	114.84
2	B	28	OMU	O4-C4-C5	-2.93	120.01	125.16
2	B	28	OMU	O2-C2-N1	-2.49	119.47	122.79

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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 (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	759/899 (84%)	0.61	82 (10%)	5	5	30, 63, 116, 149	0
2	B	7/28 (25%)	0.74	1 (14%)	2	2	63, 66, 114, 175	0
All	All	766/927 (82%)	0.61	83 (10%)	5	5	30, 63, 116, 175	0

All (83) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	132	ILE	6.0
1	A	130	ILE	6.0
1	A	673	HIS	5.5
1	A	131	SER	5.3
1	A	223	ARG	5.3
1	A	430	ASP	5.0
1	A	188	LEU	5.0
1	A	411	GLY	4.1
1	A	268	GLY	4.1
1	A	438	LEU	4.0
1	A	412	LEU	3.9
1	A	434	ALA	3.9
1	A	405	PRO	3.9
1	A	134	ARG	3.9
1	A	342	VAL	3.7
1	A	703	HIS	3.7
1	A	175	ILE	3.7
1	A	680	LYS	3.6
1	A	177	PRO	3.6
1	A	266	LEU	3.6
1	A	230	MET	3.5
1	A	133	LEU	3.4
1	A	233	LEU	3.4
1	A	436	PHE	3.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	341	VAL	3.4
1	A	176	SER	3.3
1	A	269	ARG	3.3
1	A	322	LEU	3.3
1	A	231	ARG	3.2
1	A	328	THR	3.2
1	A	319	LEU	3.2
1	A	191	VAL	3.2
1	A	171	TYR	3.2
1	A	441	SER	3.2
1	A	184	VAL	3.0
1	A	432	MET	3.0
1	A	274	PRO	2.9
1	A	140	VAL	2.8
1	A	228	GLU	2.8
1	A	345	ILE	2.8
1	A	313	LEU	2.8
1	A	413	PRO	2.8
1	A	214	PRO	2.8
1	A	287	ILE	2.7
1	A	267	MET	2.7
1	A	260	ASN	2.7
1	A	403	SER	2.7
1	A	343	GLY	2.7
1	A	435	ASN	2.7
1	A	215	ASP	2.7
1	A	709	LEU	2.7
1	A	445	HIS	2.6
1	A	271	TYR	2.6
1	A	427	GLY	2.6
1	A	627	ALA	2.6
1	A	679	GLU	2.6
1	A	674	ASP	2.5
1	A	178	GLU	2.5
1	A	676	ARG	2.4
1	A	408	ILE	2.3
1	A	315	MET	2.3
1	A	236	LEU	2.3
1	A	440	ARG	2.3
1	A	294	THR	2.3
1	A	247	TYR	2.3
1	A	675	THR	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	312	VAL	2.2
1	A	192	HIS	2.2
1	A	244	ASP	2.1
1	A	338	LEU	2.1
1	A	325	TYR	2.1
1	A	284	LYS	2.1
1	A	227	ASN	2.1
1	A	286	GLN	2.1
1	A	221	SER	2.1
1	A	265	GLN	2.1
1	A	169	TYR	2.1
1	A	248	ILE	2.1
1	A	406	ARG	2.1
1	A	711	SER	2.0
2	B	5	U	2.0
1	A	242	PRO	2.0
1	A	135	THR	2.0

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

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
2	OMU	B	28	21/22	0.94	0.17	68,82,92,97	0

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	MG	A	902	1/1	0.94	0.25	60,60,60,60	0
3	MG	A	901	1/1	0.99	0.07	30,30,30,30	0

6.5 Other polymers ⓘ

There are no such residues in this entry.