WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Sep 19, 2023 – 04:43 AM EDT

PDB ID	:	5C7E
Title	:	Crystal structure of the rice Topless related protein 2 (TPR2) N-terminal domain (1-209) in complex with Arabidopsis IAA10 peptide
Authors	:	Ke, J.; Ma, H.; Gu, X.; Brunzelle, J.S.; Xu, H.E.; Melcher, K.
Deposited on	:	2015-06-24
Resolution	:	3.10 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

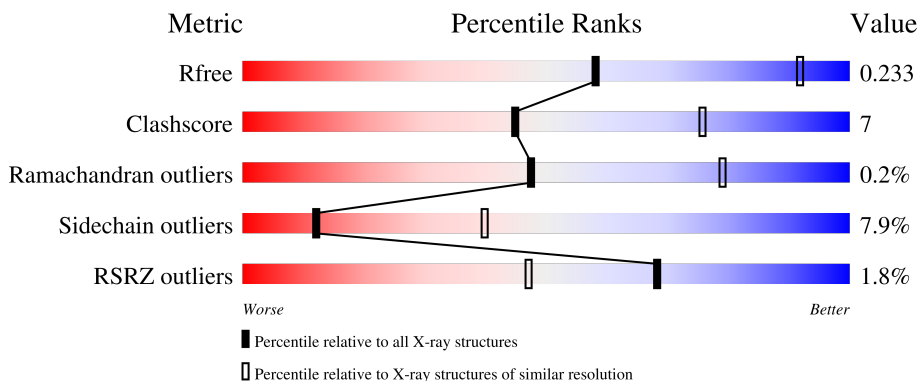|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 3.10 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		0.233
Clashscore		7
Ramachandran outliers		0.2%
Sidechain outliers		7.9%
RSRZ outliers		1.8%

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	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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	209	76% / 20%
1	B	209	82% / 14%
1	C	209	80% / 16%
1	D	209	78% / 18%
1	E	209	72% / 24%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	209	76% ... 21%
2	G	11	27% 64% 18% 9% 9%
2	H	11	9% 27% 27% 45%
2	I	11	55% 18% 9% 18%
2	J	11	9% 55% 27% 18%
2	K	11	45% 9% 9% 36%
2	L	11	27% 36% 9% 27%

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 10658 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 ASPR2 protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	205	Total	C	N	O	S	0	0	0
			1710	1103	287	314	6			
1	B	205	Total	C	N	O	S	0	0	0
			1704	1101	281	316	6			
1	C	206	Total	C	N	O	S	0	0	0
			1725	1113	288	318	6			
1	D	206	Total	C	N	O	S	0	0	0
			1720	1108	287	319	6			
1	E	205	Total	C	N	O	S	0	0	0
			1714	1105	286	317	6			
1	F	205	Total	C	N	O	S	0	0	0
			1718	1108	287	317	6			

- Molecule 2 is a protein called Auxin-responsive protein IAA10.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	G	10	Total	C	N	O	0	0	0
			65	41	10	14			
2	H	6	Total	C	N	O	0	0	0
			43	29	6	8			
2	I	9	Total	C	N	O	0	0	0
			68	42	9	17			
2	J	9	Total	C	N	O	0	0	0
			57	37	9	11			
2	K	7	Total	C	N	O	0	0	0
			46	31	7	8			
2	L	8	Total	C	N	O	0	0	0
			59	38	8	13			

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total 1	Zn 1	0	0
3	B	1	Total 1	Zn 1	0	0
3	C	1	Total 1	Zn 1	0	0
3	D	1	Total 1	Zn 1	0	0
3	E	1	Total 1	Zn 1	0	0
3	F	1	Total 1	Zn 1	0	0

- Molecule 4 is water.

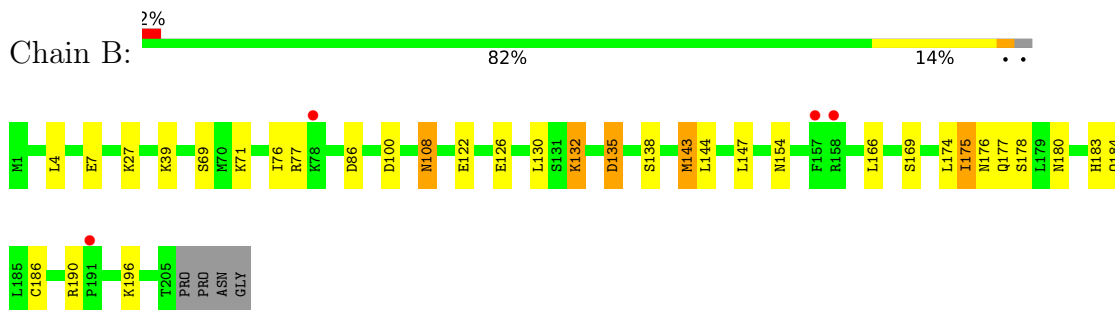| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	4	Total 4	O 4	0	0
4	B	1	Total 1	O 1	0	0
4	C	7	Total 7	O 7	0	0
4	D	4	Total 4	O 4	0	0
4	E	5	Total 5	O 5	0	0
4	F	1	Total 1	O 1	0	0
4	L	1	Total 1	O 1	0	0

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.
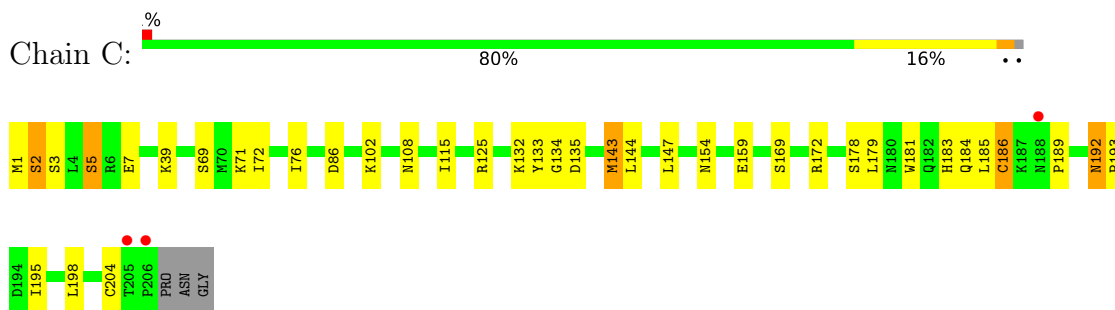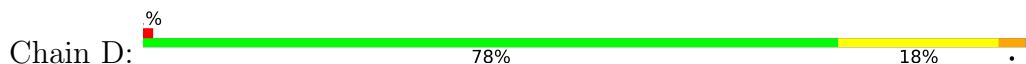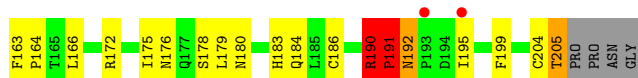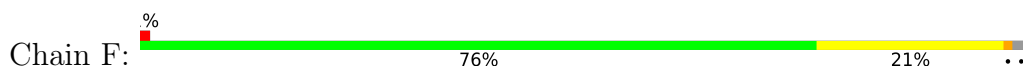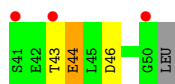
- Molecule 1: ASPR2 protein
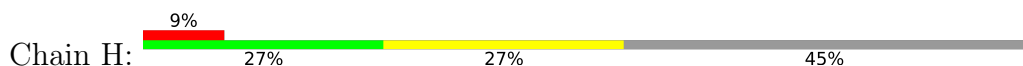

- Molecule 1: ASPR2 protein

- Molecule 1: ASPR2 protein

- Molecule 1: ASPR2 protein

M1 S2 S3 L4 S5 R6 M38 K39 E42 E49 V53 R67 Y68 S69 M70 K71 I76 R77 D86 D100 E122 R125 E126 K132 Y133 G134 D135 M143 L144 L147 K148 N154 L166 R172 S178 L179 N180 W181 Q182 H183 Q184 L185 C186

K187 N188 P189 R190 P191 N192 P193 D194 I195 F199 P206 PRO ASN GLY

• Molecule 1: ASPR2 protein

Chain E:

M1 S2 S5 R6 E7 K27 M38 K39 E42 D51 K55 R67 Y68 S69 M70 F74 E75 I76 R77 K78 L85 D86 D100 N108 E109 E110 L111 L119 E122 R125 E126 Y133 G134 S138 I142 M143 L144 L147 K148 N154

F163 P164 T165 L166 R172 I175 N176 Q177 S178 L179 N180 H183 Q184 L185 C186 R190 P191 N192 P193 D194 I195 F199 C204 T205 PRO PRO ASN GLY

• Molecule 1: ASPR2 protein

Chain F:

M1 S2 S3 L4 S5 R6 E7 K27 K39 S69 M70 K71 F74 E75 I76 R77 R90 A91 K92 D100 L101 K102 V103 F104 S105 T106 F107 N108 L111 E122 R125 E126 L130 S131 K132 Y133 G134 M143 L144 L147 N154 F157 N176 Q177 S178

L179 N180 W181 Q182 H183 Q184 L185 C186 R190 P191 N192 P193 D194 I195 F199 T205 PRO PRO ASN GLY

• Molecule 2: Auxin-responsive protein IAA10

Chain G:

S41 E42 T43 E44 L45 D46 G50 LEU

• Molecule 2: Auxin-responsive protein IAA10

Chain H:

SER GLU THR E44 L45 L49 GLY LEU

• Molecule 2: Auxin-responsive protein IAA10

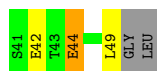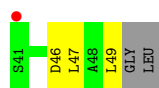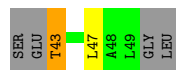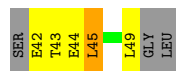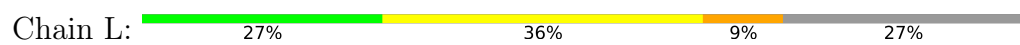Chain I:

S41 E42 T43 E44 L49 GLY LEU

• Molecule 2: Auxin-responsive protein IAA10

Chain J:

• Molecule 2: Auxin-responsive protein IAA10

Chain K:

• Molecule 2: Auxin-responsive protein IAA10

Chain L:

4   Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	162.67Å 162.67Å 157.25Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	44.99 – 3.10 44.99 – 3.10	Depositor EDS
% Data completeness (in resolution range)	99.9 (44.99-3.10) 99.9 (44.99-3.10)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.28 (at 3.12Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.196 , 0.229 0.202 , 0.233	Depositor DCC
R_{free} test set	2237 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å^2)	98.5	Xtriage
Anisotropy	0.426	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 63.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.026 for -h,-k,l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	10658	wwPDB-VP
Average B, all atoms (Å^2)	102.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.42% 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: ZN

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	0.40	0/1745	0.65	0/2344
1	B	0.36	0/1739	0.64	1/2338 (0.0%)
1	C	0.44	0/1761	0.69	1/2367 (0.0%)
1	D	0.39	0/1755	0.68	3/2359 (0.1%)
1	E	0.47	1/1749 (0.1%)	0.72	3/2351 (0.1%)
1	F	0.39	0/1753	0.63	0/2355
2	G	0.38	0/64	0.68	0/86
2	H	0.36	0/42	0.79	0/56
2	I	0.65	0/67	1.21	0/90
2	J	0.51	0/56	1.06	0/76
2	K	0.32	0/45	0.79	0/61
2	L	0.36	0/58	0.91	0/78
All	All	0.41	1/10834 (0.0%)	0.68	8/14561 (0.1%)

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	C	0	1
1	D	0	3
1	F	0	1
2	K	0	1
All	All	0	6

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	191	PRO	N-CD	5.87	1.56	1.47

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	190	ARG	N-CA-C	7.60	131.53	111.00
1	E	192	ASN	N-CA-C	-6.06	94.64	111.00
1	D	190	ARG	N-CA-C	-5.71	95.59	111.00
1	E	190	ARG	N-CA-CB	-5.34	100.98	110.60
1	D	135	ASP	N-CA-C	5.22	125.10	111.00

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	192	ASN	Peptide
1	D	132	LYS	Peptide
1	D	134	GLY	Peptide
1	D	188	ASN	Peptide
1	F	193	PRO	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	1710	0	1707	27	0
1	B	1704	0	1696	18	0
1	C	1725	0	1730	25	0
1	D	1720	0	1727	24	0
1	E	1714	0	1712	33	1
1	F	1718	0	1723	26	1
2	G	65	0	62	3	0
2	H	43	0	45	3	0
2	I	68	0	65	3	0
2	J	57	0	55	1	0
2	K	46	0	48	0	0
2	L	59	0	58	3	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	E	1	0	0	0	0
3	F	1	0	0	0	0
4	A	4	0	0	0	0
4	B	1	0	0	0	0
4	C	7	0	0	1	0
4	D	4	0	0	0	0
4	E	5	0	0	0	0
4	F	1	0	0	0	0
4	L	1	0	0	0	0
All	All	10658	0	10628	139	1

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 139 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:C:189:PRO:HB3	1:C:204:CYS:SG	2.18	0.83
1:A:71:LYS:HE2	2:G:44:GLU:HG2	1.62	0.81
1:F:190:ARG:HB3	1:F:191:PRO:HD2	1.63	0.79
1:F:102:LYS:O	1:F:105:SER:HB2	1.84	0.77
1:E:190:ARG:HB3	1:E:191:PRO:HD3	1.64	0.77

All (1) 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)
1:E:67:ARG:NH1	1:F:193:PRO:O[4_555]	2.05	0.15

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	203/209 (97%)	199 (98%)	4 (2%)	0	100	100
1	B	203/209 (97%)	202 (100%)	1 (0%)	0	100	100
1	C	204/209 (98%)	202 (99%)	2 (1%)	0	100	100
1	D	204/209 (98%)	202 (99%)	2 (1%)	0	100	100
1	E	203/209 (97%)	198 (98%)	3 (2%)	2 (1%)	15	49
1	F	203/209 (97%)	199 (98%)	3 (2%)	1 (0%)	29	64
2	G	8/11 (73%)	8 (100%)	0	0	100	100
2	H	4/11 (36%)	4 (100%)	0	0	100	100
2	I	7/11 (64%)	7 (100%)	0	0	100	100
2	J	7/11 (64%)	6 (86%)	1 (14%)	0	100	100
2	K	5/11 (46%)	4 (80%)	1 (20%)	0	100	100
2	L	6/11 (54%)	6 (100%)	0	0	100	100
All	All	1257/1320 (95%)	1237 (98%)	17 (1%)	3 (0%)	47	79

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	191	PRO
1	E	190	ARG
1	F	193	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	188/197 (95%)	176 (94%)	12 (6%)	17	48
1	B	188/197 (95%)	174 (93%)	14 (7%)	13	42
1	C	192/197 (98%)	177 (92%)	15 (8%)	12	40
1	D	192/197 (98%)	179 (93%)	13 (7%)	16	45
1	E	190/197 (96%)	175 (92%)	15 (8%)	12	40
1	F	191/197 (97%)	177 (93%)	14 (7%)	14	43

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	G	6/9 (67%)	4 (67%)	2 (33%)	0	0
2	H	4/9 (44%)	4 (100%)	0	100	100
2	I	8/9 (89%)	7 (88%)	1 (12%)	4	18
2	J	5/9 (56%)	3 (60%)	2 (40%)	0	0
2	K	4/9 (44%)	2 (50%)	2 (50%)	0	0
2	L	6/9 (67%)	3 (50%)	3 (50%)	0	0
All	All	1174/1236 (95%)	1081 (92%)	93 (8%)	12	40

5 of 93 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	70	MET
1	F	39	LYS
1	E	108	ASN
1	E	176	ASN
1	F	133	TYR

Sometimes 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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 6 ligands modelled in this entry, 6 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 ⓘ

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	205/209 (98%)	-0.08	2 (0%)	82	67	69, 86, 163, 195	0
1	B	205/209 (98%)	0.11	4 (1%)	65	44	81, 113, 149, 171	0
1	C	206/209 (98%)	-0.02	3 (1%)	73	54	63, 81, 123, 155	0
1	D	206/209 (98%)	0.10	3 (1%)	73	54	67, 97, 137, 180	0
1	E	205/209 (98%)	0.04	4 (1%)	65	44	78, 98, 155, 194	0
1	F	205/209 (98%)	0.07	2 (0%)	82	67	83, 105, 158, 198	0
2	G	10/11 (90%)	1.14	3 (30%)	0	0	86, 101, 157, 162	0
2	H	6/11 (54%)	0.71	1 (16%)	1	1	127, 141, 150, 155	0
2	I	9/11 (81%)	0.27	0	100	100	81, 110, 148, 150	0
2	J	9/11 (81%)	0.52	1 (11%)	5	2	107, 121, 156, 158	0
2	K	7/11 (63%)	-0.29	0	100	100	96, 111, 138, 146	0
2	L	8/11 (72%)	0.60	0	100	100	119, 137, 163, 180	0
All	All	1281/1320 (97%)	0.06	23 (1%)	68	47	63, 98, 154, 198	0

The worst 5 of 23 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	191	PRO	6.1
1	C	206	PRO	3.8
1	D	206	PRO	3.2
1	B	157	PHE	3.0
2	G	43	THR	3.0

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	ZN	E	400	1/1	0.79	0.07	149,149,149,149	1
3	ZN	F	400	1/1	0.87	0.05	162,162,162,162	1
3	ZN	A	400	1/1	0.90	0.07	161,161,161,161	1
3	ZN	D	400	1/1	0.92	0.09	110,110,110,110	1
3	ZN	B	400	1/1	0.97	0.07	119,119,119,119	1
3	ZN	C	400	1/1	0.98	0.15	84,84,84,84	1

6.5 Other polymers ⓘ

There are no such residues in this entry.