wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 12, 2021 – 01:55 PM EDT

PDB ID	:	1Z9C
Title	:	Crystal structure of OhrR bound to the ohrA promoter: Structure of MarR family protein with operator DNA
Authors	:	Hong, M.; Fuangthong, M.; Helmann, J.D.; Brennan, R.G.
Deposited on	:	2005-04-01
Resolution	:	2.64 Å(reported)

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

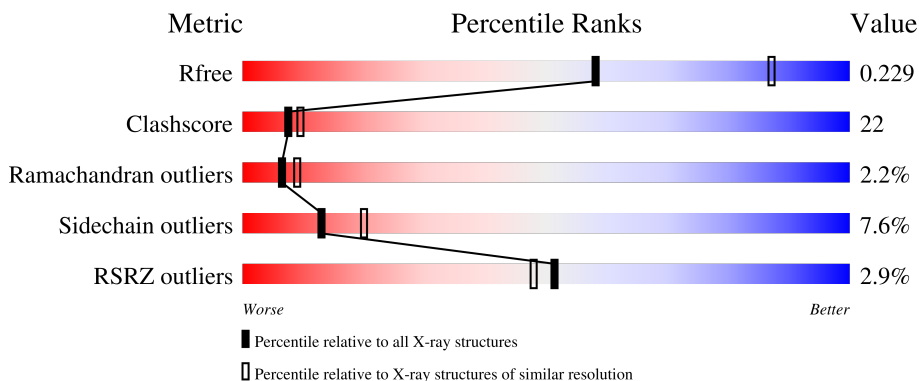|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 2.64 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		0.229
Clashscore		22
Ramachandran outliers		2.2%
Sidechain outliers		7.6%
RSRZ outliers		2.9%

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	1426 (2.66-2.62)
Clashscore	141614	1472 (2.66-2.62)
Ramachandran outliers	138981	1446 (2.66-2.62)
Sidechain outliers	138945	1446 (2.66-2.62)
RSRZ outliers	127900	1408 (2.66-2.62)

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	G	29	7% 48% 52%
1	I	29	7% 41% 52% · ·
1	K	29	7% 72% 24% ·
2	H	29	3% 52% 48%
2	J	29	7% 38% 55% · ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	L	29	38% · 55% · · ·
3	A	147	% · 52% · 36% · 5% · 7%
3	B	147	% · 52% · 36% · 5% · 7%
3	C	147	3% · 54% · 36% · · · 6%
3	D	147	5% · 42% · 47% · · · 7%
3	E	147	% · 61% · 29% · · · 7%
3	F	147	% · 46% · 44% · · 6%

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 10073 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 DNA chain called DNA (29-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	G	29	Total	C	N	O	P	0	0	0
			590	288	99	175	28			
1	I	28	Total	C	N	O	P	0	0	0
			570	278	97	168	27			
1	K	29	Total	C	N	O	P	0	0	0
			590	288	99	175	28			

- Molecule 2 is a DNA chain called DNA (29-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	29	Total	C	N	O	P	0	0	0
			592	288	105	171	28			
2	J	28	Total	C	N	O	P	0	0	0
			572	278	103	164	27			
2	L	28	Total	C	N	O	P	0	0	0
			572	278	103	164	27			

- Molecule 3 is a protein called Organic hydroperoxide resistance transcriptional regulator.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	137	Total	C	N	O	S	0	0	0
			1075	688	178	205	4			
3	B	137	Total	C	N	O	S	0	0	0
			1068	684	172	207	5			
3	C	138	Total	C	N	O	S	0	0	0
			1057	677	173	202	5			
3	D	137	Total	C	N	O	S	0	0	0
			1051	670	173	203	5			
3	E	137	Total	C	N	O	S	0	0	0
			1085	693	177	210	5			
3	F	138	Total	C	N	O	S	0	0	0
			1077	690	175	207	5			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	15	SER	CYS	engineered mutation	UNP O34777
B	15	SER	CYS	engineered mutation	UNP O34777
C	15	SER	CYS	engineered mutation	UNP O34777
D	15	SER	CYS	engineered mutation	UNP O34777
E	15	SER	CYS	engineered mutation	UNP O34777
F	15	SER	CYS	engineered mutation	UNP O34777

- Molecule 4 is water.

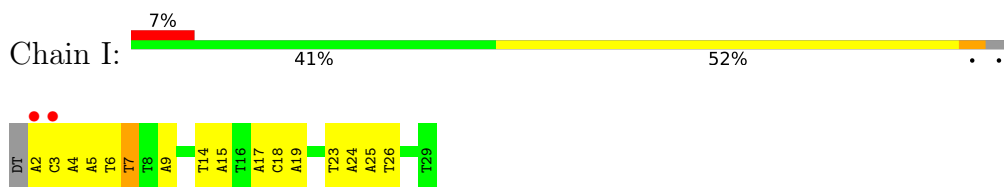| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	G	13	Total 13	O 13	0	0
4	H	11	Total 11	O 11	0	0
4	I	7	Total 7	O 7	0	0
4	J	8	Total 8	O 8	0	0
4	K	11	Total 11	O 11	0	0
4	L	10	Total 10	O 10	0	0
4	A	25	Total 25	O 25	0	0
4	B	16	Total 16	O 16	0	0
4	C	16	Total 16	O 16	0	0
4	D	14	Total 14	O 14	0	0
4	E	21	Total 21	O 21	0	0
4	F	22	Total 22	O 22	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.
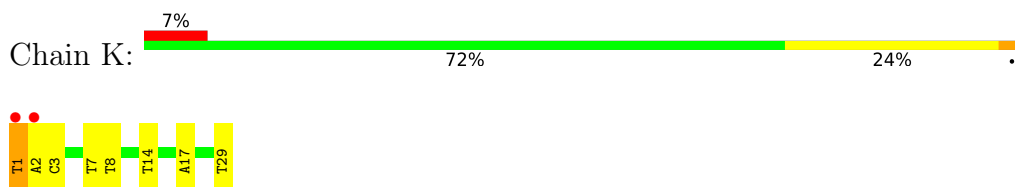
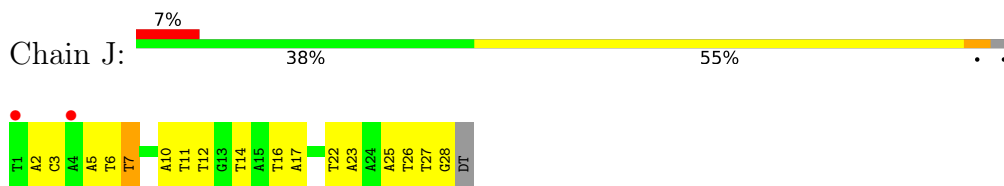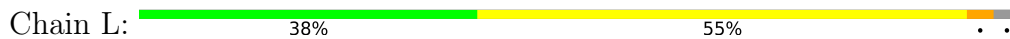- Molecule 1: DNA (29-MER)
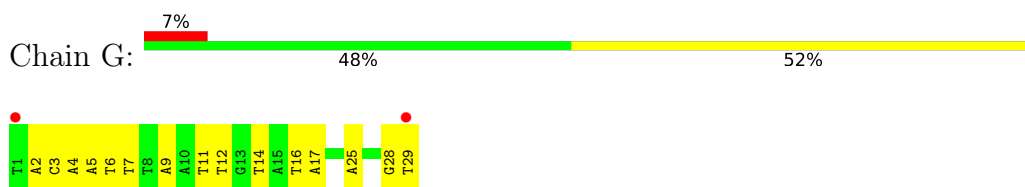
Chain G:

- Molecule 1: DNA (29-MER)

Chain I:

- Molecule 1: DNA (29-MER)

Chain K:

- Molecule 2: DNA (29-MER)

Chain H:

- Molecule 2: DNA (29-MER)

Chain J:
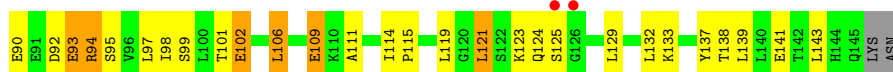

- Molecule 2: DNA (29-MER)

Chain L:

• Molecule 3: Organic hydroperoxide resistance transcriptional regulator

Chain A:

• Molecule 3: Organic hydroperoxide resistance transcriptional regulator

Chain B:

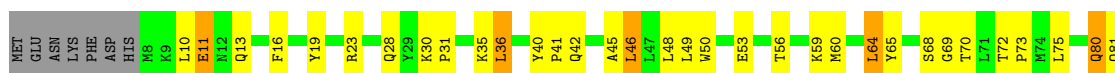• Molecule 3: Organic hydroperoxide resistance transcriptional regulator
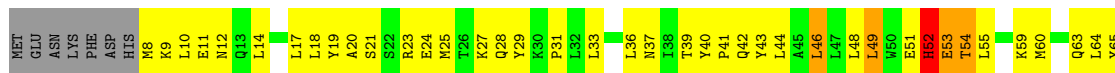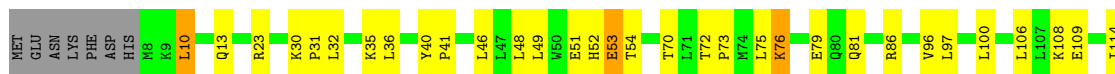
Chain C:

• Molecule 3: Organic hydroperoxide resistance transcriptional regulator

Chain D:

• Molecule 3: Organic hydroperoxide resistance transcriptional regulator

Chain E:

%

61%   29%   7%

MET GLU ASN LYS PHE ASP HIS M8 K9 L10 Q13 R23 K30 P31 L32 K35 L36 Y40 P41 L46 L47 L48 L49 W50 E51 H52 E53 T54 T70 L71 T72 P73 M74 L75 K76 E79 Q80 Q81 R86 V96 L97 L100 L106 L107 K108 E109 I114

P115 G116 T117 I118 L119 G120 L121 S122 K123 G126 E127 D128 L129 K130 Q131 Y137 T138 L139 L143 H144 GLN LYS ASN

• Molecule 3: Organic hydroperoxide resistance transcriptional regulator

Chain F:

%

46%   44%   6%

MET GLU ASN LYS PHE ASP H7 M8 K9 L10 E11 F16 L17 L18 Y19 A20 S21 S22 R23 Q28 Y29 K30 P31 L32 L33 D34 K35 L36 Y40 P41 Q42 Y43 L44 A45 L46 L47 L48 L49 W50 E51 H52 E53 T54 L55 T56 V57 K58 K59 E62 Q63 L64 Y65 L66 D67 S68 G69

T70 L71 T72 P73 M74 L75 K76 R77 M78 L83 R86 D92 E93 R94 S95 I98 S99 L100 T101 E102 K108 E109 K110 A111 V112 D113 I114 P115 G116 L121 S122 K123 Q124 S125 G126 L129 S134 Y137 E141 T142 L143 H144 GLN LYS ASN
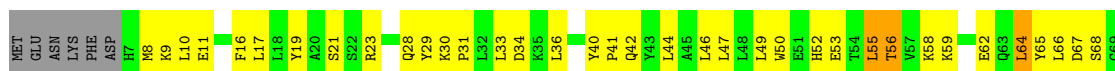
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	81.90Å 80.96Å 109.34Å 90.00° 100.20° 90.00°	Depositor
Resolution (Å)	59.65 – 2.64 57.12 – 2.64	Depositor EDS
% Data completeness (in resolution range)	96.0 (59.65-2.64) 91.6 (57.12-2.64)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	0.05	Depositor
$< I/\sigma(I) >$ [1]	2.82 (at 2.65Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.216 , 0.277 0.223 , 0.229	Depositor DCC
R_{free} test set	1914 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	42.7	Xtriage
Anisotropy	0.376	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 63.2	EDS
L-test for twinning[2]	$< |L| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	10073	wwPDB-VP
Average B, all atoms (Å^2)	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.07% 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 ⓘ

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	G	1.56	2/660 (0.3%)	0.83	0/1017
1	I	1.31	2/638 (0.3%)	0.78	1/983 (0.1%)
1	K	1.53	2/660 (0.3%)	0.92	3/1017 (0.3%)
2	H	1.17	1/664 (0.2%)	0.75	1/1023 (0.1%)
2	J	1.16	2/642 (0.3%)	0.77	1/989 (0.1%)
2	L	1.36	2/642 (0.3%)	0.78	0/989
3	A	0.39	0/1089	0.63	0/1471
3	B	0.40	0/1082	0.67	2/1463 (0.1%)
3	C	0.37	0/1072	0.60	0/1454
3	D	0.34	0/1065	0.62	0/1443
3	E	0.37	0/1099	0.65	0/1483
3	F	0.38	0/1090	0.69	0/1471
All	All	0.88	11/10403 (0.1%)	0.71	8/14803 (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	K	1	0

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	G	14	DT	C5-C7	31.41	1.68	1.50
1	K	14	DT	C5-C7	29.12	1.67	1.50
2	H	14	DT	C5-C7	27.39	1.66	1.50
2	L	14	DT	C5-C7	23.66	1.64	1.50
1	I	7	DT	C5-C7	23.09	1.64	1.50

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	K	1	DT	C1'-O4'-C4'	-8.90	101.20	110.10
1	K	1	DT	C4'-C3'-C2'	-6.95	96.85	103.10
1	K	1	DT	C2'-C3'-O3'	-6.12	92.42	112.60
1	I	14	DT	C6-C5-C7	-5.54	119.58	122.90
2	H	14	DT	C6-C5-C7	-5.50	119.60	122.90

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	K	1	DT	C3'

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	G	590	0	335	14	0
1	I	570	0	323	17	0
1	K	590	0	335	7	0
2	H	592	0	333	23	0
2	J	572	0	321	20	0
2	L	572	0	321	22	0
3	A	1075	0	1108	51	0
3	B	1068	0	1089	54	0
3	C	1057	0	1060	58	0
3	D	1051	0	1050	76	0
3	E	1085	0	1119	50	0
3	F	1077	0	1111	68	0
4	A	25	0	0	2	0
4	B	16	0	0	1	0
4	C	16	0	0	2	0
4	D	14	0	0	1	0
4	E	21	0	0	0	0
4	F	22	0	0	4	0
4	G	13	0	0	0	0
4	H	11	0	0	0	0
4	I	7	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	J	8	0	0	0	0
4	K	11	0	0	0	0
4	L	10	0	0	2	0
All	All	10073	0	8505	403	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 22.

The worst 5 of 403 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)
2:L:19:DA:H5"	3:F:30:LYS:HE3	1.31	1.12
3:E:130:LYS:H	3:E:130:LYS:HD2	1.17	1.10
1:I:17:DA:H2"	1:I:18:DC:H5'	1.37	1.05
2:H:26:DT:H2"	2:H:27:DT:H5'	1.40	1.02
3:C:123:LYS:HB2	3:C:129:LEU:HD22	1.44	0.99

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	
3	A	135/147 (92%)	120 (89%)	13 (10%)	2 (2%)	10	14
3	B	135/147 (92%)	122 (90%)	8 (6%)	5 (4%)	3	4
3	C	136/147 (92%)	119 (88%)	14 (10%)	3 (2%)	6	9
3	D	135/147 (92%)	119 (88%)	14 (10%)	2 (2%)	10	14
3	E	135/147 (92%)	124 (92%)	8 (6%)	3 (2%)	6	9
3	F	136/147 (92%)	123 (90%)	10 (7%)	3 (2%)	6	9
All	All	812/882 (92%)	727 (90%)	67 (8%)	18 (2%)	6	9

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	11	GLU
3	D	53	GLU
3	E	127	GLU
3	F	125	SER
3	A	93	GLU

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	
3	A	118/135 (87%)	109 (92%)	9 (8%)	13	20
3	B	117/135 (87%)	109 (93%)	8 (7%)	16	24
3	C	113/135 (84%)	105 (93%)	8 (7%)	14	22
3	D	112/135 (83%)	104 (93%)	8 (7%)	14	22
3	E	121/135 (90%)	112 (93%)	9 (7%)	13	21
3	F	118/135 (87%)	107 (91%)	11 (9%)	9	12
All	All	699/810 (86%)	646 (92%)	53 (8%)	13	20

5 of 53 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	52	HIS
3	E	46	LEU
3	F	113	ASP
3	D	54	THR
3	D	131	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
3	D	12	ASN
3	D	37	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	E	52	HIS
3	E	13	GLN
3	E	42	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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

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	G	29/29 (100%)	0.13	2 (6%)	16	13	18, 45, 114, 148	0
1	I	28/29 (96%)	0.27	2 (7%)	16	12	26, 63, 124, 141	0
1	K	29/29 (100%)	0.07	2 (6%)	16	13	21, 40, 80, 140	0
2	H	29/29 (100%)	0.04	1 (3%)	45	41	17, 45, 109, 150	0
2	J	28/29 (96%)	0.41	2 (7%)	16	12	25, 57, 120, 129	0
2	L	28/29 (96%)	-0.01	0	100	100	19, 42, 73, 85	0
3	A	137/147 (93%)	-0.09	2 (1%)	73	71	18, 36, 87, 104	0
3	B	137/147 (93%)	-0.14	2 (1%)	73	71	14, 38, 83, 112	0
3	C	138/147 (93%)	-0.02	5 (3%)	42	39	21, 49, 89, 121	0
3	D	137/147 (93%)	0.23	7 (5%)	28	24	29, 53, 100, 129	0
3	E	137/147 (93%)	-0.22	2 (1%)	73	71	14, 36, 97, 115	0
3	F	138/147 (93%)	-0.16	2 (1%)	75	73	17, 35, 78, 102	0
All	All	995/1056 (94%)	-0.03	29 (2%)	51	48	14, 42, 98, 150	0

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	D	127	GLU	6.1
1	G	29	DT	4.5
3	D	125	SER	4.5
3	F	125	SER	4.3
3	C	125	SER	4.3

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

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

WORLDWIDE
wwPDB
PROTEIN DATA BANK

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.