Full wwPDB X-ray Structure Validation Report ⓘ

Jan 24, 2021 – 01:12 PM EST

PDB ID	:	3BIL
Title	:	Crystal structure of a probable LacI family transcriptional regulator from Corynebacterium glutamicum
Authors	:	Bonanno, J.B.; Freeman, J.; Bain, K.T.; Mendoza, M.; Ozyurt, S.; Smith, D.; Wasserman, S.; Sauder, J.M.; Burley, S.K.; Almo, S.C.; New York SGX Research Center for Structural Genomics (NYSGXRC)
Deposited on	:	2007-11-30
Resolution	:	2.50 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	:	2.16
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)	:	2.16

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 2.50 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	Percentile Ranks	Value
Rfree		0.272
Clashscore		9
Ramachandran outliers		0.6%
Sidechain outliers		10.9%
RSRZ outliers		7.2%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
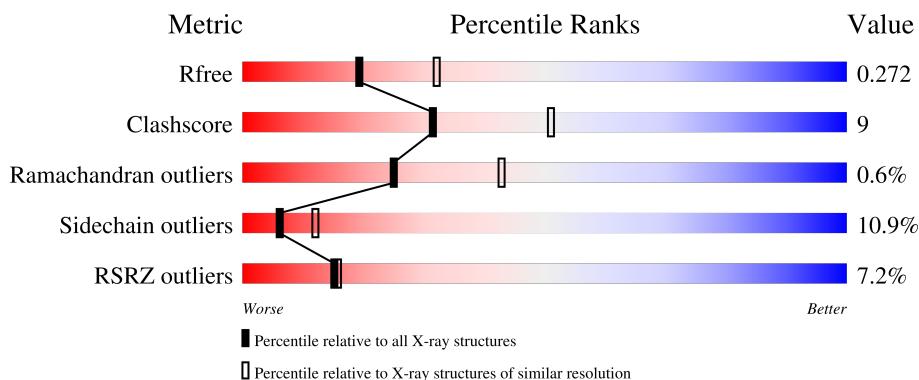
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	348	4% — 54% / 17% • 26%
1	B	348	7% — 59% / 13% • 26%
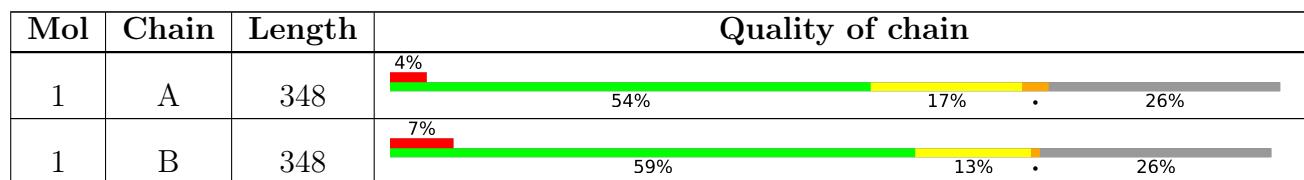
2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 3842 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 Probable LacI-family transcriptional regulator.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	257	Total	C	N	O	S	0	1	0
			1911	1202	319	380	10			
1	B	257	Total	C	N	O	S	0	1	0
			1911	1202	319	380	10			

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	expression tag	UNP Q8NQQ9
A	1	SER	-	expression tag	UNP Q8NQQ9
A	2	LEU	-	expression tag	UNP Q8NQQ9
A	340	GLU	-	expression tag	UNP Q8NQQ9
A	341	GLY	-	expression tag	UNP Q8NQQ9
A	342	HIS	-	expression tag	UNP Q8NQQ9
A	343	HIS	-	expression tag	UNP Q8NQQ9
A	344	HIS	-	expression tag	UNP Q8NQQ9
A	345	HIS	-	expression tag	UNP Q8NQQ9
A	346	HIS	-	expression tag	UNP Q8NQQ9
A	347	HIS	-	expression tag	UNP Q8NQQ9
B	0	MET	-	expression tag	UNP Q8NQQ9
B	1	SER	-	expression tag	UNP Q8NQQ9
B	2	LEU	-	expression tag	UNP Q8NQQ9
B	340	GLU	-	expression tag	UNP Q8NQQ9
B	341	GLY	-	expression tag	UNP Q8NQQ9
B	342	HIS	-	expression tag	UNP Q8NQQ9
B	343	HIS	-	expression tag	UNP Q8NQQ9
B	344	HIS	-	expression tag	UNP Q8NQQ9
B	345	HIS	-	expression tag	UNP Q8NQQ9
B	346	HIS	-	expression tag	UNP Q8NQQ9
B	347	HIS	-	expression tag	UNP Q8NQQ9

- Molecule 2 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	11	Total 11	O 11	0	0
2	B	9	Total 9	O 9	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: Probable LacI-family transcriptional regulator
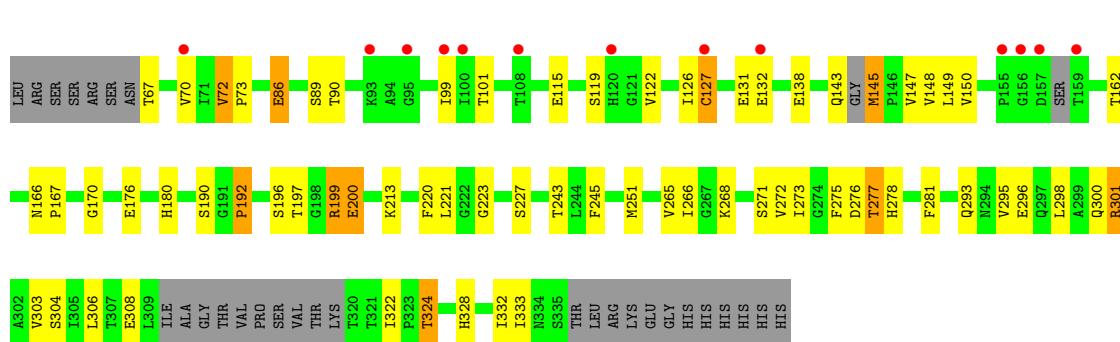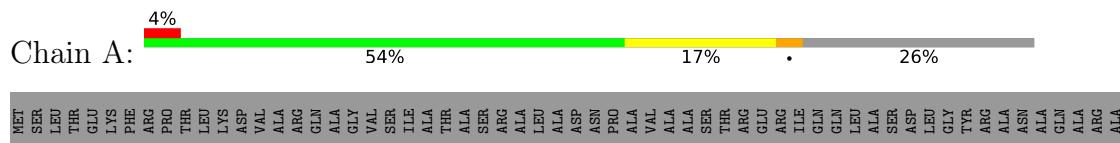
- Molecule 1: Probable LacI-family transcriptional regulator
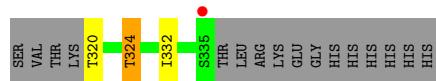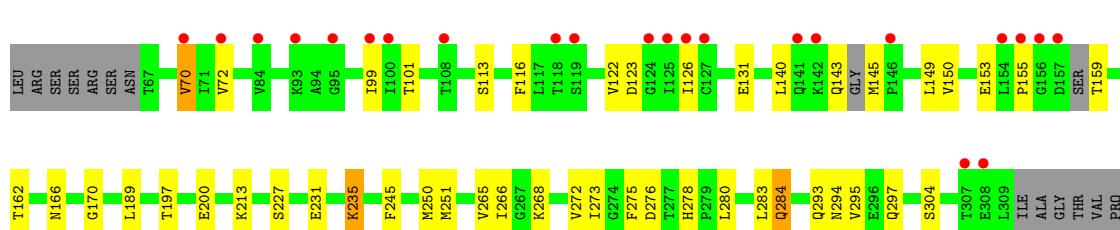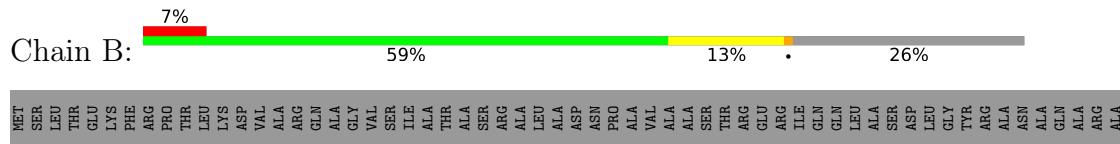
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	75.21Å 54.34Å 78.65Å 90.00° 94.19° 90.00°	Depositor
Resolution (Å)	20.00 – 2.50 29.92 – 2.50	Depositor EDS
% Data completeness (in resolution range)	100.0 (20.00-2.50) 100.0 (29.92-2.50)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	0.13	Depositor
$< I/\sigma(I) >$ [1]	2.70 (at 2.51Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.218 , 0.278 0.215 , 0.272	Depositor DCC
R_{free} test set	1117 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	45.8	Xtriage
Anisotropy	0.650	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 50.9	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.026 for l,-k,h	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	3842	wwPDB-VP
Average B, all atoms (Å^2)	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.96% 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	A	0.78	0/1942	0.81	0/2641
1	B	0.77	0/1942	0.79	1/2641 (0.0%)
All	All	0.77	0/3884	0.80	1/5282 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	250	MET	CG-SD-CE	5.05	108.28	100.20

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	1911	0	1914	46	0
1	B	1911	0	1914	29	0
2	A	11	0	0	2	0
2	B	9	0	0	0	0
All	All	3842	0	3828	69	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 9.

All (69) 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:149:LEU:HB2	1:B:162:THR:HG22	1.50	0.93
1:A:170:GLY:H	1:A:324:THR:HG22	1.43	0.82
1:B:170:GLY:H	1:B:324:THR:HG22	1.46	0.80
1:A:99:ILE:CG2	1:B:99:ILE:HG23	2.14	0.77
1:A:99:ILE:CG2	1:B:99:ILE:CG2	2.62	0.77
1:A:90:THR:CG2	1:A:303:VAL:HG11	2.17	0.75
1:A:149:LEU:HB2	1:A:162:THR:HG22	1.69	0.73
1:B:170:GLY:H	1:B:324:THR:CG2	2.01	0.73
1:A:192:PRO:HD3	1:A:223:GLY:O	1.95	0.65
1:A:99:ILE:HG23	1:B:99:ILE:CG2	2.27	0.65
1:A:122:VAL:O	1:A:145:MET:HE1	1.97	0.65
1:B:278:HIS:HD1	1:B:280:LEU:H	1.46	0.64
1:A:298:LEU:HG	1:A:322:ILE:HG13	1.80	0.63
1:B:276:ASP:OD1	1:B:293:GLN:NE2	2.31	0.61
1:A:266:ILE:HD13	1:A:272:VAL:HG21	1.83	0.61
1:A:90:THR:CG2	1:A:303:VAL:CG1	2.79	0.60
1:A:328[B]:HIS:CE1	2:A:351:HOH:O	2.54	0.60
1:B:170:GLY:HA3	1:B:324:THR:HG23	1.82	0.60
1:B:143:GLN:O	1:B:145:MET:N	2.35	0.59
1:A:166:ASN:O	1:A:324:THR:HB	2.02	0.59
1:A:72:VAL:HG22	1:A:73:PRO:HD2	1.85	0.59
1:A:190:SER:HB3	1:A:221:LEU:HA	1.85	0.58
1:A:328[B]:HIS:NE2	2:A:351:HOH:O	2.32	0.58
1:A:90:THR:HG21	1:A:303:VAL:CG1	2.34	0.57
1:A:273:ILE:CD1	1:A:332:ILE:HD11	2.36	0.56
1:B:166:ASN:O	1:B:324:THR:HB	2.05	0.56
1:A:99:ILE:HG22	1:B:99:ILE:HG23	1.86	0.55
1:A:196:SER:O	1:A:200:GLU:HB2	2.07	0.55
1:B:273:ILE:CD1	1:B:332:ILE:HD11	2.38	0.53
1:A:276:ASP:OD1	1:A:293:GLN:NE2	2.41	0.53
1:A:243:THR:OG1	1:A:271:SER:HB2	2.09	0.53
1:B:294:ASN:OD1	1:B:297:GLN:HB2	2.09	0.53
1:A:90:THR:HG22	1:A:303:VAL:HG11	1.89	0.52
1:A:170:GLY:H	1:A:324:THR:CG2	2.20	0.52
1:A:148:VAL:HG23	1:A:306:LEU:HD13	1.92	0.52
1:A:273:ILE:HD11	1:A:332:ILE:HD11	1.92	0.51
1:B:131:GLU:HG2	1:B:155:PRO:HD3	1.92	0.51
1:A:67:THR:HG21	1:B:116:PHE:CE1	2.46	0.51
1:A:170:GLY:N	1:A:324:THR:HG22	2.20	0.51
1:A:176:GLU:OE1	1:A:180:HIS:CE1	2.64	0.51
1:B:170:GLY:N	1:B:324:THR:CG2	2.73	0.50
1:A:126:ILE:HG13	1:A:306:LEU:HD22	1.94	0.50

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:265:VAL:HB	1:A:268:LYS:HB3	1.94	0.49
1:A:277:THR:O	1:A:278:HIS:C	2.52	0.49
1:B:265:VAL:HB	1:B:268:LYS:HB3	1.95	0.48
1:B:266:ILE:HD13	1:B:272:VAL:HG21	1.95	0.48
1:A:143:GLN:O	1:A:145:MET:N	2.47	0.47
1:A:251:MET:HA	1:A:251:MET:HE2	1.97	0.47
1:A:301:ARG:HG3	1:A:322:ILE:HD11	1.96	0.46
1:B:170:GLY:HA3	1:B:324:THR:CG2	2.46	0.45
1:B:273:ILE:HD13	1:B:332:ILE:HD11	1.98	0.45
1:A:271:SER:OG	1:A:333:ILE:HD12	2.18	0.44
1:B:140:LEU:HD23	1:B:140:LEU:HA	1.90	0.44
1:B:273:ILE:HD11	1:B:332:ILE:HD11	1.99	0.44
1:A:190:SER:HB3	1:A:220:PHE:O	2.18	0.44
1:B:231:GLU:HG3	1:B:235:LYS:HE2	2.00	0.44
1:A:131:GLU:OE1	1:A:199:ARG:NH2	2.37	0.43
1:A:273:ILE:HD13	1:A:332:ILE:HD11	2.01	0.43
1:A:127:CYS:O	1:A:149:LEU:HA	2.18	0.43
1:A:170:GLY:HA3	1:A:324:THR:CG2	2.49	0.43
1:A:170:GLY:HA3	1:A:324:THR:HG23	2.00	0.43
1:B:283:LEU:O	1:B:284:GLN:O	2.38	0.42
1:B:122:VAL:O	1:B:145:MET:CE	2.68	0.42
1:A:281:PHE:HE2	1:B:280:LEU:HD11	1.85	0.41
1:B:70:VAL:HA	1:B:126:ILE:O	2.21	0.41
1:A:86:GLU:HG2	1:A:296:GLU:HG3	2.02	0.41
1:A:90:THR:HG22	1:A:303:VAL:CG1	2.50	0.40
1:A:166:ASN:HA	1:A:167:PRO:HD3	1.94	0.40
1:B:251:MET:HA	1:B:251:MET:HE2	2.04	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	250/348 (72%)	236 (94%)	12 (5%)	2 (1%)	19	35
1	B	250/348 (72%)	236 (94%)	13 (5%)	1 (0%)	34	54
All	All	500/696 (72%)	472 (94%)	25 (5%)	3 (1%)	25	43

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	284	GLN
1	A	308	GLU
1	A	132	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	
1	A	212/282 (75%)	186 (88%)	26 (12%)	4	9
1	B	212/282 (75%)	192 (91%)	20 (9%)	8	17
All	All	424/564 (75%)	378 (89%)	46 (11%)	6	12

All (46) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	70	VAL
1	A	72	VAL
1	A	86	GLU
1	A	89	SER
1	A	101	THR
1	A	115	GLU
1	A	119	SER
1	A	127	CYS
1	A	138	GLU
1	A	145	MET
1	A	147	VAL
1	A	150	VAL
1	A	192	PRO

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	197	THR
1	A	199	ARG
1	A	200	GLU
1	A	213	LYS
1	A	227	SER
1	A	245	PHE
1	A	275	PHE
1	A	277	THR
1	A	295	VAL
1	A	300	GLN
1	A	301	ARG
1	A	304	SER
1	A	324	THR
1	B	70	VAL
1	B	72	VAL
1	B	101	THR
1	B	113	SER
1	B	123	ASP
1	B	150	VAL
1	B	153	GLU
1	B	159	THR
1	B	189	LEU
1	B	197	THR
1	B	200	GLU
1	B	213	LYS
1	B	227	SER
1	B	235	LYS
1	B	245	PHE
1	B	275	PHE
1	B	295	VAL
1	B	304	SER
1	B	320	THR
1	B	324	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	120	HIS
1	A	181	ASN
1	A	211	ASN
1	A	334	ASN
1	B	211	ASN

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	A	257/348 (73%)	0.49	13 (5%)	28	29	16, 45, 78, 88	0
1	B	257/348 (73%)	0.58	24 (9%)	8	8	16, 45, 81, 95	0
All	All	514/696 (73%)	0.53	37 (7%)	15	16	16, 45, 78, 95	0

All (37) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	157	ASP	6.2
1	B	100	ILE	5.5
1	A	100	ILE	5.3
1	B	70	VAL	5.2
1	B	156	GLY	5.0
1	A	70	VAL	4.8
1	A	157	ASP	4.4
1	B	155	PRO	4.2
1	A	156	GLY	3.8
1	B	126	ILE	3.7
1	B	146	PRO	3.7
1	B	127	CYS	3.5
1	A	95	GLY	3.5
1	B	141	GLN	3.3
1	B	142	LYS	3.0
1	B	335	SER	2.9
1	A	155	PRO	2.9
1	A	159	THR	2.8
1	B	154	LEU	2.7
1	B	108	THR	2.6
1	B	118	THR	2.6
1	B	72	VAL	2.5
1	B	84	VAL	2.5
1	B	307	THR	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	132	GLU	2.5
1	B	124	GLY	2.4
1	A	127	CYS	2.3
1	B	119	SER	2.3
1	B	95	GLY	2.3
1	A	120	HIS	2.3
1	B	99	ILE	2.3
1	B	308	GLU	2.2
1	A	99	ILE	2.2
1	B	125	ILE	2.2
1	B	93	LYS	2.1
1	A	108	THR	2.1
1	A	93	LYS	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 ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.