



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 13, 2017 – 04:42 am GMT

PDB ID : 4Z9J  
Title : Apo-Tar from E. coli  
Authors : Mise, T.  
Deposited on : 2015-04-10  
Resolution : 1.78 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc28949

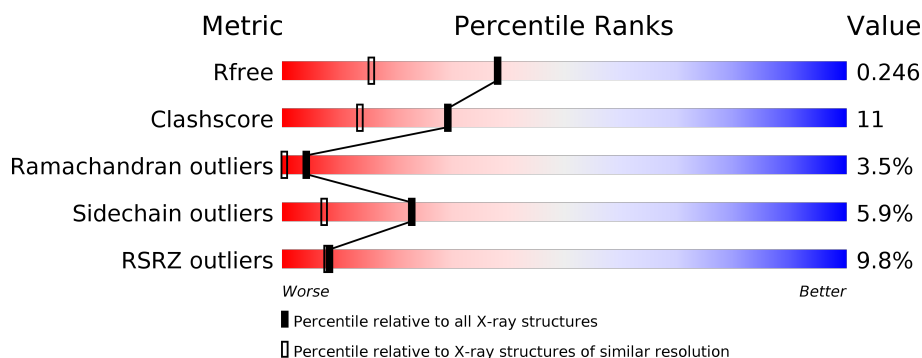
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.78 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))
$R_{free}$	100719	7172 (1.80-1.76)
Clashscore	112137	8247 (1.80-1.76)
Ramachandran outliers	110173	8154 (1.80-1.76)
Sidechain outliers	110143	8153 (1.80-1.76)
RSRZ outliers	101464	7262 (1.80-1.76)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	196	<div> <div>7%</div> <div>64%</div> <div>12%</div> <div>• •</div> <div>19%</div> </div>
1	B	196	<div> <div>9%</div> <div>61%</div> <div>14%</div> <div>5%</div> <div>•</div> <div>19%</div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2752 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 Methyl-accepting chemotaxis protein II.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	158	Total	C	N	O	S	0	0	0
			1269	791	218	252	8			
1	B	158	Total	C	N	O	S	0	0	0
			1269	791	218	252	8			

There are 56 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	MET	-	expression tag	UNP P07017
A	-1	GLY	-	expression tag	UNP P07017
A	0	SER	-	expression tag	UNP P07017
A	1	SER	-	expression tag	UNP P07017
A	2	HIS	-	expression tag	UNP P07017
A	3	HIS	-	expression tag	UNP P07017
A	4	HIS	-	expression tag	UNP P07017
A	5	HIS	-	expression tag	UNP P07017
A	6	HIS	-	expression tag	UNP P07017
A	7	HIS	-	expression tag	UNP P07017
A	8	SER	-	expression tag	UNP P07017
A	9	SER	-	expression tag	UNP P07017
A	10	GLY	-	expression tag	UNP P07017
A	11	LEU	-	expression tag	UNP P07017
A	12	VAL	-	expression tag	UNP P07017
A	13	PRO	-	expression tag	UNP P07017
A	14	ARG	-	expression tag	UNP P07017
A	15	GLY	-	expression tag	UNP P07017
A	16	SER	-	expression tag	UNP P07017
A	17	HIS	-	expression tag	UNP P07017
A	18	MET	-	expression tag	UNP P07017
A	19	ALA	-	expression tag	UNP P07017
A	20	SER	-	expression tag	UNP P07017
A	21	ASP	-	expression tag	UNP P07017
A	22	ASP	-	expression tag	UNP P07017

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	23	ASP	-	expression tag	UNP P07017
A	24	ASP	-	expression tag	UNP P07017
A	25	LYS	-	expression tag	UNP P07017
B	-2	MET	-	expression tag	UNP P07017
B	-1	GLY	-	expression tag	UNP P07017
B	0	SER	-	expression tag	UNP P07017
B	1	SER	-	expression tag	UNP P07017
B	2	HIS	-	expression tag	UNP P07017
B	3	HIS	-	expression tag	UNP P07017
B	4	HIS	-	expression tag	UNP P07017
B	5	HIS	-	expression tag	UNP P07017
B	6	HIS	-	expression tag	UNP P07017
B	7	HIS	-	expression tag	UNP P07017
B	8	SER	-	expression tag	UNP P07017
B	9	SER	-	expression tag	UNP P07017
B	10	GLY	-	expression tag	UNP P07017
B	11	LEU	-	expression tag	UNP P07017
B	12	VAL	-	expression tag	UNP P07017
B	13	PRO	-	expression tag	UNP P07017
B	14	ARG	-	expression tag	UNP P07017
B	15	GLY	-	expression tag	UNP P07017
B	16	SER	-	expression tag	UNP P07017
B	17	HIS	-	expression tag	UNP P07017
B	18	MET	-	expression tag	UNP P07017
B	19	ALA	-	expression tag	UNP P07017
B	20	SER	-	expression tag	UNP P07017
B	21	ASP	-	expression tag	UNP P07017
B	22	ASP	-	expression tag	UNP P07017
B	23	ASP	-	expression tag	UNP P07017
B	24	ASP	-	expression tag	UNP P07017
B	25	LYS	-	expression tag	UNP P07017

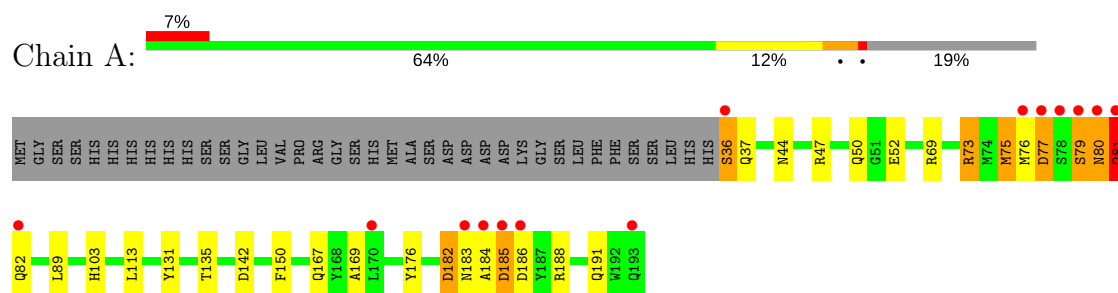
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	115	Total O 115 115	0	0
2	B	99	Total O 99 99	0	0

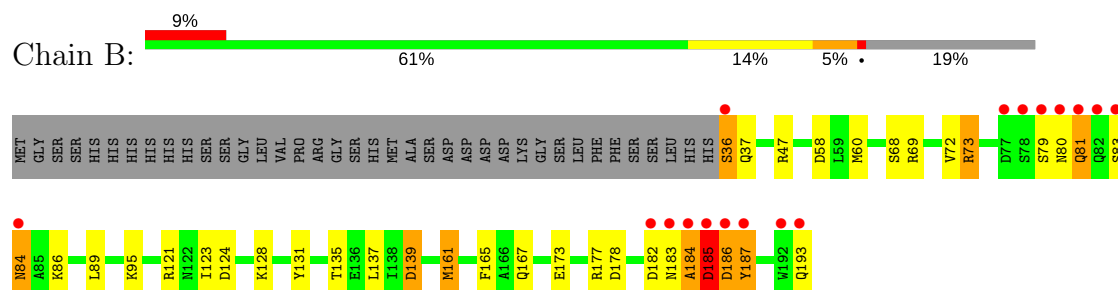
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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: Methyl-accepting chemotaxis protein II



#### • Molecule 1: Methyl-accepting chemotaxis protein II



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	60.41Å 73.51Å 79.05Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	53.83 – 1.78 48.00 – 1.78	Depositor EDS
% Data completeness (in resolution range)	99.8 (53.83-1.78) 99.8 (48.00-1.78)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.99 (at 1.78Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, $R_{free}$	0.182 , 0.234 0.194 , 0.246	Depositor DCC
$R_{free}$ test set	1665 reflections (5.09%)	DCC
Wilson B-factor (Å <sup>2</sup> )	22.7	Xtriage
Anisotropy	0.455	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 47.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	2752	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.45% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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	1.03	1/1291 (0.1%)	1.03	4/1742 (0.2%)
1	B	1.07	2/1291 (0.2%)	1.11	7/1742 (0.4%)
All	All	1.05	3/2582 (0.1%)	1.07	11/3484 (0.3%)

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	A	0	1
1	B	0	1
All	All	0	2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	139	ASP	CB-CG	-5.54	1.40	1.51
1	B	47	ARG	CZ-NH2	5.49	1.40	1.33
1	A	176	TYR	CB-CG	-5.06	1.44	1.51

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	161	MET	CG-SD-CE	-13.76	78.18	100.20
1	B	47	ARG	NE-CZ-NH2	-12.38	114.11	120.30
1	A	69	ARG	NE-CZ-NH1	9.94	125.27	120.30
1	A	69	ARG	NE-CZ-NH2	-7.08	116.76	120.30
1	B	69	ARG	NE-CZ-NH2	-6.77	116.92	120.30
1	B	95	LYS	CD-CE-NZ	6.01	125.52	111.70
1	B	69	ARG	NE-CZ-NH1	5.91	123.25	120.30
1	A	73	ARG	NE-CZ-NH1	5.51	123.05	120.30
1	B	137	LEU	CB-CG-CD2	5.39	120.17	111.00

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	73	ARG	NE-CZ-NH2	-5.28	117.66	120.30
1	B	73	ARG	CA-CB-CG	-5.23	101.89	113.40

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	185	ASP	Peptide
1	B	36	SER	Peptide

## 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	1269	0	1228	34	0
1	B	1269	0	1228	22	0
2	A	115	0	0	4	0
2	B	99	0	0	3	0
All	All	2752	0	2456	54	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 11.

All (54) 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:167:GLN:OE1	2:A:201:HOH:O	1.94	0.84
1:B:58:ASP:HB2	2:B:258:HOH:O	1.82	0.78
1:A:182:ASP:CB	1:A:183:ASN:HA	2.18	0.74
1:A:182:ASP:HB3	1:A:183:ASN:HA	1.71	0.73
1:A:36:SER:CB	1:A:37:GLN:HA	2.20	0.72
1:B:185:ASP:O	1:B:186:ASP:HB3	1.92	0.69
1:A:185:ASP:HB2	1:A:186:ASP:HB2	1.78	0.66
1:B:124:ASP:OD2	1:B:128:LYS:HE2	1.96	0.65
1:A:76:MET:SD	1:A:82:GLN:HB3	2.39	0.62

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:36:SER:HB3	1:A:37:GLN:HA	1.81	0.61
1:A:52:GLU:OE1	1:A:103:HIS:HD2	1.83	0.61
1:B:185:ASP:HB3	1:B:187:TYR:HB3	1.82	0.60
1:A:80:ASN:O	1:A:81:GLN:HB2	2.03	0.57
1:B:173:GLU:O	1:B:177:ARG:HG2	2.04	0.57
1:A:47:ARG:NH1	1:B:173:GLU:OE1	2.37	0.57
1:A:131:TYR:O	1:A:135:THR:HG23	2.04	0.57
1:B:187:TYR:N	1:B:187:TYR:CD1	2.73	0.56
1:A:79:SER:OG	1:A:80:ASN:N	2.39	0.56
1:B:186:ASP:HB3	1:B:187:TYR:HA	1.87	0.56
1:B:193:GLN:HA	1:B:193:GLN:NE2	2.21	0.55
1:A:185:ASP:HB3	1:A:186:ASP:HA	1.87	0.55
1:B:73:ARG:HD2	1:B:89:LEU:CD1	2.35	0.55
1:A:73:ARG:HD2	1:A:89:LEU:HD12	1.89	0.55
1:A:36:SER:O	2:A:202:HOH:O	2.19	0.53
1:A:75:MET:HE1	1:A:150:PHE:HZ	1.74	0.53
1:A:47:ARG:HH12	1:B:173:GLU:CD	2.13	0.52
1:A:75:MET:CE	1:A:150:PHE:HZ	2.23	0.50
1:A:80:ASN:O	1:A:81:GLN:CB	2.60	0.49
1:B:185:ASP:O	1:B:186:ASP:CB	2.61	0.48
1:A:36:SER:HB3	1:A:37:GLN:CA	2.44	0.47
1:B:123:ILE:HD11	1:B:165:PHE:HA	1.96	0.47
1:A:185:ASP:HB2	1:A:186:ASP:CB	2.45	0.47
1:B:68:SER:O	1:B:72:VAL:HG13	2.15	0.47
1:B:131:TYR:O	1:B:135:THR:HG23	2.15	0.46
1:A:44:ASN:OD1	1:A:47:ARG:NH2	2.47	0.46
1:A:185:ASP:HB3	1:A:186:ASP:CA	2.45	0.46
1:A:185:ASP:CB	1:A:186:ASP:CA	2.94	0.45
1:B:186:ASP:CB	1:B:187:TYR:HA	2.43	0.45
1:B:121:ARG:NE	2:B:205:HOH:O	2.50	0.45
1:B:121:ARG:CD	2:B:205:HOH:O	2.65	0.45
1:A:76:MET:O	1:A:77:ASP:O	2.34	0.45
1:B:183:ASN:O	1:B:184:ALA:HB2	2.16	0.45
1:A:76:MET:SD	1:A:82:GLN:CB	3.04	0.44
1:A:50:GLN:HE22	1:A:169:ALA:HA	1.83	0.43
1:B:60:MET:SD	1:B:161:MET:CE	3.07	0.43
1:A:103:HIS:HE1	2:A:310:HOH:O	2.01	0.43
1:A:142:ASP:OD1	2:A:203:HOH:O	2.21	0.43
1:A:185:ASP:CB	1:A:186:ASP:CB	2.97	0.42
1:A:113:LEU:HD12	1:A:113:LEU:N	2.33	0.42
1:B:178:ASP:HA	1:B:183:ASN:HB3	2.02	0.42

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:36:SER:HB3	1:A:37:GLN:NE2	2.35	0.41
1:A:184:ALA:HA	1:A:185:ASP:C	2.40	0.41
1:A:76:MET:CE	1:A:82:GLN:HB3	2.52	0.40
1:B:81:GLN:HG2	1:B:81:GLN:O	2.21	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	156/196 (80%)	148 (95%)	4 (3%)	4 (3%)	6	1
1	B	156/196 (80%)	145 (93%)	4 (3%)	7 (4%)	3	0
All	All	312/392 (80%)	293 (94%)	8 (3%)	11 (4%)	4	0

All (11) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	77	ASP
1	A	79	SER
1	A	81	GLN
1	B	37	GLN
1	B	79	SER
1	B	84	ASN
1	B	184	ALA
1	B	186	ASP
1	A	80	ASN
1	B	80	ASN
1	B	185	ASP

### 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	136/169 (80%)	130 (96%)	6 (4%)	33	14
1	B	136/169 (80%)	126 (93%)	10 (7%)	16	4
All	All	272/338 (80%)	256 (94%)	16 (6%)	23	7

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	36	SER
1	A	75	MET
1	A	81	GLN
1	A	182	ASP
1	A	188	ARG
1	A	191	GLN
1	B	36	SER
1	B	81	GLN
1	B	83	SER
1	B	84	ASN
1	B	86	LYS
1	B	139	ASP
1	B	167	GLN
1	B	182	ASP
1	B	185	ASP
1	B	187	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	50	GLN
1	A	80	ASN
1	A	81	GLN
1	A	82	GLN
1	A	84	ASN
1	A	99	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	103	HIS
1	A	152	GLN
1	A	191	GLN
1	B	80	ASN
1	B	152	GLN
1	B	191	GLN
1	B	193	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 carbohydrates 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

### 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	158/196 (80%)	0.57	14 (8%) 10 10	14, 22, 78, 112	0
1	B	158/196 (80%)	0.80	17 (10%) 6 6	15, 27, 86, 114	0
All	All	316/392 (80%)	0.68	31 (9%) 8 8	14, 25, 86, 114	0

All (31) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	78	SER	12.8
1	B	186	ASP	12.1
1	A	183	ASN	11.0
1	A	79	SER	8.2
1	B	36	SER	8.1
1	A	185	ASP	7.5
1	B	80	ASN	6.3
1	B	184	ALA	6.3
1	A	36	SER	5.8
1	B	187	TYR	5.7
1	B	81	GLN	5.5
1	B	79	SER	5.3
1	A	77	ASP	5.1
1	B	83	SER	4.9
1	B	185	ASP	4.3
1	B	82	GLN	4.2
1	A	81	GLN	4.0
1	B	77	ASP	4.0
1	B	183	ASN	4.0
1	A	80	ASN	3.9
1	A	186	ASP	3.6
1	A	82	GLN	3.6
1	B	84	ASN	3.1
1	B	193	GLN	2.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	78	SER	2.8
1	A	193	GLN	2.5
1	B	182	ASP	2.3
1	A	184	ALA	2.3
1	B	192	TRP	2.2
1	A	76	MET	2.1
1	A	170	LEU	2.1

## 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 carbohydrates 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.