Full wwPDB X-ray Structure Validation Report (i)

Aug 20, 2023 – 12:58 AM EDT

PDB ID	:	2FJY
Title	:	Crystal Structure of B-form Bombyx mori Pheromone Binding Protein
Authors	:	Lautenschlager, C.; Leal, W.S.; Clardy, J.
Deposited on	:	2006-01-03
Resolution	:	2.30 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

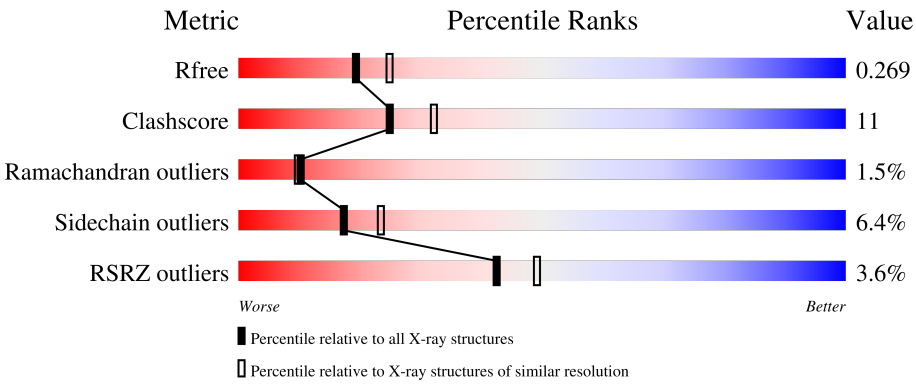
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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}	130704	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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	142	% 68% 23% 5% •
1	B	142	6% 58% 27% 12% •

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2261 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 Pheromone-binding protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	136	1059	664	176	207	12	0	0	0
1	B	142	1107	693	184	217	13	0	0	0

- Molecule 2 is water.

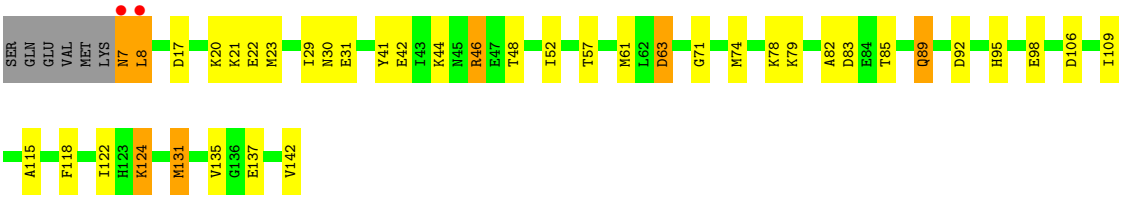
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	48	48	48	0	0
2	B	47	47	47	0	0

3 Residue-property plots (i)

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.

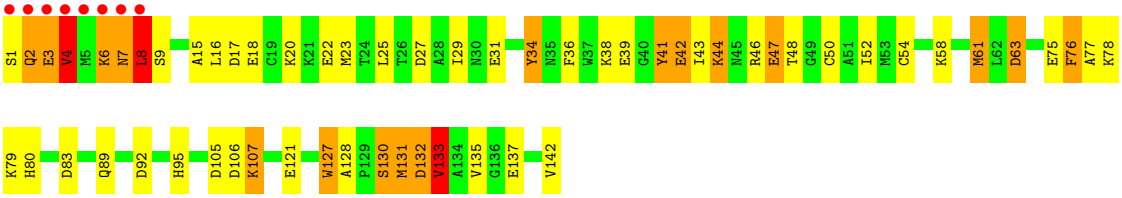
• Molecule 1: Pheromone-binding protein

Chain A:



• Molecule 1: Pheromone-binding protein

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	54.24Å 70.79Å 75.49Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.00 – 2.30 44.05 – 2.30	Depositor EDS
% Data completeness (in resolution range)	89.9 (43.00-2.30) 89.9 (44.05-2.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.75 (at 2.29Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.220 , 0.270 0.216 , 0.269	Depositor DCC
R_{free} test set	605 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	20.8	Xtriage
Anisotropy	0.157	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 27.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	2261	wwPDB-VP
Average B, all atoms (Å^2)	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 46.69 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.1045e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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 (i)

5.1 Standard geometry (i)

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	2.03	27/1079 (2.5%)	1.34	13/1454 (0.9%)
1	B	2.18	36/1127 (3.2%)	1.41	17/1517 (1.1%)
All	All	2.11	63/2206 (2.9%)	1.38	30/2971 (1.0%)

All (63) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	63	ASP	CB-CG	14.58	1.82	1.51
1	B	63	ASP	CB-CG	13.42	1.79	1.51
1	B	107	LYS	CE-NZ	10.62	1.75	1.49
1	A	89	GLN	CG-CD	10.16	1.74	1.51
1	A	31	GLU	CD-OE1	9.64	1.36	1.25
1	B	39	GLU	CG-CD	9.62	1.66	1.51
1	B	107	LYS	CD-CE	9.22	1.74	1.51
1	B	89	GLN	CG-CD	9.08	1.72	1.51
1	B	18	GLU	CD-OE1	-8.47	1.16	1.25
1	B	22	GLU	CD-OE2	8.45	1.34	1.25
1	B	61	MET	SD-CE	-8.42	1.30	1.77
1	B	77	ALA	CA-CB	8.14	1.69	1.52
1	B	89	GLN	CD-NE2	7.99	1.52	1.32
1	A	23	MET	CG-SD	-7.88	1.60	1.81
1	B	142	VAL	CB-CG1	-7.56	1.36	1.52
1	A	46	ARG	CG-CD	7.54	1.70	1.51
1	B	79	LYS	CD-CE	7.50	1.70	1.51
1	A	124	LYS	CE-NZ	7.40	1.67	1.49
1	A	44	LYS	CB-CG	7.21	1.72	1.52
1	B	42	GLU	CG-CD	7.15	1.62	1.51
1	A	41	TYR	CE1-CZ	-6.87	1.29	1.38
1	B	41	TYR	CZ-OH	6.68	1.49	1.37
1	A	142	VAL	CB-CG1	-6.57	1.39	1.52
1	B	78	LYS	CD-CE	6.53	1.67	1.51
1	B	39	GLU	CB-CG	6.49	1.64	1.52
1	B	31	GLU	CD-OE1	6.48	1.32	1.25

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	4	VAL	CA-CB	6.35	1.68	1.54
1	B	44	LYS	CB-CG	6.22	1.69	1.52
1	B	2	GLN	CB-CG	6.20	1.69	1.52
1	A	46	ARG	CZ-NH2	6.09	1.41	1.33
1	B	128	ALA	CA-CB	6.08	1.65	1.52
1	A	82	ALA	CA-CB	-5.96	1.40	1.52
1	A	31	GLU	CD-OE2	5.96	1.32	1.25
1	B	41	TYR	CG-CD1	-5.95	1.31	1.39
1	B	22	GLU	CG-CD	5.93	1.60	1.51
1	A	7	ASN	CB-CG	5.85	1.64	1.51
1	A	98	GLU	CD-OE1	5.80	1.32	1.25
1	B	41	TYR	CD1-CE1	5.74	1.48	1.39
1	B	34	TYR	CG-CD1	-5.64	1.31	1.39
1	B	75	GLU	CD-OE1	5.63	1.31	1.25
1	A	83	ASP	CB-CG	5.60	1.63	1.51
1	B	137	GLU	CD-OE2	-5.60	1.19	1.25
1	A	41	TYR	CD1-CE1	5.53	1.47	1.39
1	A	137	GLU	CG-CD	5.52	1.60	1.51
1	A	95	HIS	C-O	-5.49	1.12	1.23
1	B	106	ASP	CB-CG	-5.49	1.40	1.51
1	A	79	LYS	CD-CE	5.46	1.64	1.51
1	B	76	PHE	CG-CD2	-5.44	1.30	1.38
1	B	76	PHE	CE1-CZ	-5.42	1.27	1.37
1	A	115	ALA	CA-CB	-5.41	1.41	1.52
1	A	61	MET	CG-SD	5.32	1.95	1.81
1	A	22	GLU	CG-CD	5.29	1.59	1.51
1	B	130	SER	CA-CB	-5.28	1.45	1.52
1	B	131	MET	CG-SD	5.25	1.94	1.81
1	B	2	GLN	CG-CD	5.23	1.63	1.51
1	A	89	GLN	CD-NE2	5.22	1.46	1.32
1	A	71	GLY	C-O	-5.21	1.15	1.23
1	A	78	LYS	CD-CE	5.20	1.64	1.51
1	B	127	TRP	CB-CG	-5.15	1.41	1.50
1	A	44	LYS	CE-NZ	5.13	1.61	1.49
1	A	74	MET	CG-SD	5.10	1.94	1.81
1	B	95	HIS	N-CA	5.10	1.56	1.46
1	B	121	GLU	CD-OE1	-5.05	1.20	1.25

All (30) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	46	ARG	NE-CZ-NH2	9.85	125.22	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	92	ASP	CB-CG-OD1	9.63	126.96	118.30
1	A	46	ARG	NE-CZ-NH2	9.43	125.02	120.30
1	B	92	ASP	CB-CG-OD2	9.23	126.61	118.30
1	B	39	GLU	OE1-CD-OE2	-8.99	112.51	123.30
1	A	31	GLU	OE1-CD-OE2	8.45	133.44	123.30
1	A	106	ASP	CB-CG-OD2	8.38	125.84	118.30
1	A	83	ASP	CB-CG-OD2	7.39	124.95	118.30
1	B	133	VAL	CB-CA-C	-6.93	98.24	111.40
1	B	17	ASP	CB-CG-OD2	6.74	124.37	118.30
1	B	132	ASP	CB-CG-OD2	6.73	124.36	118.30
1	B	135	VAL	CG1-CB-CG2	-6.70	100.19	110.90
1	B	131	MET	CG-SD-CE	6.64	110.83	100.20
1	B	106	ASP	CB-CG-OD1	6.50	124.15	118.30
1	A	8	LEU	CA-CB-CG	6.40	130.03	115.30
1	B	63	ASP	CB-CG-OD2	6.18	123.86	118.30
1	A	131	MET	CG-SD-CE	6.17	110.07	100.20
1	A	63	ASP	CB-CG-OD1	5.99	123.69	118.30
1	B	18	GLU	OE1-CD-OE2	-5.95	116.16	123.30
1	B	46	ARG	NE-CZ-NH1	-5.92	117.34	120.30
1	A	44	LYS	CD-CE-NZ	5.91	125.28	111.70
1	B	47	GLU	OE1-CD-OE2	5.89	130.36	123.30
1	A	63	ASP	CB-CG-OD2	5.63	123.37	118.30
1	B	8	LEU	CA-CB-CG	5.58	128.13	115.30
1	A	63	ASP	OD1-CG-OD2	-5.48	112.89	123.30
1	A	21	LYS	CD-CE-NZ	-5.43	99.22	111.70
1	B	92	ASP	CB-CG-OD1	-5.29	113.54	118.30
1	B	25	LEU	CA-CB-CG	5.19	127.23	115.30
1	B	39	GLU	CG-CD-OE1	5.10	128.51	118.30
1	A	92	ASP	CB-CG-OD2	-5.00	113.80	118.30

There are no chirality outliers.

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	A	1059	0	1018	15	0
1	B	1107	0	1071	31	0
2	A	48	0	0	1	0
2	B	47	0	0	3	0
All	All	2261	0	2089	46	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 (46) 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:124:LYS:CE	1:A:124:LYS:NZ	1.67	1.57
1:A:63:ASP:CB	1:A:63:ASP:CG	1.82	1.48
1:B:107:LYS:CE	1:B:107:LYS:NZ	1.75	1.47
1:B:63:ASP:CB	1:B:63:ASP:CG	1.79	1.46
1:B:131:MET:SD	1:B:131:MET:CE	2.05	1.43
1:B:130:SER:OG	1:B:133:VAL:HG23	1.64	0.96
1:B:7:ASN:O	1:B:8:LEU:HB2	1.79	0.80
1:B:58:LYS:HD3	1:B:61:MET:HE2	1.69	0.75
1:B:7:ASN:O	1:B:8:LEU:CB	2.48	0.61
1:A:29:ILE:C	1:A:29:ILE:HD12	2.21	0.60
1:B:127:TRP:HB2	2:B:150:HOH:O	2.03	0.58
1:B:23:MET:HE1	1:B:50:CYS:O	2.04	0.58
1:B:3:GLU:HG2	1:B:4:VAL:H	1.70	0.57
1:B:20:LYS:HE2	1:B:27:ASP:HA	1.87	0.57
1:B:38:LYS:HG3	2:B:175:HOH:O	2.06	0.55
1:B:80:HIS:HB3	1:B:133:VAL:HG22	1.89	0.54
1:A:46:ARG:NH1	2:A:174:HOH:O	2.40	0.54
1:A:57:THR:HG22	1:A:57:THR:O	2.07	0.53
1:B:130:SER:OG	1:B:133:VAL:CG2	2.47	0.53
1:A:57:THR:O	1:A:57:THR:CG2	2.56	0.53
1:B:7:ASN:OD1	1:B:131:MET:HE3	2.09	0.52
1:B:20:LYS:HG3	1:B:29:ILE:HD11	1.93	0.51
1:A:118:PHE:O	1:A:122:ILE:HG23	2.11	0.50
1:B:58:LYS:HD3	1:B:61:MET:CE	2.39	0.49
1:B:3:GLU:HG2	1:B:4:VAL:N	2.27	0.49
1:B:36:PHE:HA	1:B:43:ILE:HD12	1.93	0.49
1:A:85:THR:O	1:A:89:GLN:HG2	2.14	0.47
1:A:20:LYS:HE2	1:A:30:ASN:HD21	1.80	0.47
1:B:8:LEU:HD11	1:B:34:TYR:HE2	1.80	0.46
1:A:20:LYS:HG3	1:A:29:ILE:HD11	1.96	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:15:ALA:HA	1:B:61:MET:HE1	1.97	0.46
1:B:9:SER:HB3	1:B:132:ASP:OD1	2.16	0.46
1:B:83:ASP:OD1	1:B:83:ASP:C	2.55	0.45
1:A:46:ARG:HG2	1:A:109:ILE:HG12	1.98	0.44
1:B:38:LYS:HD2	1:B:41:TYR:CZ	2.52	0.44
1:A:131:MET:O	1:A:135:VAL:HG23	2.17	0.44
1:B:48:THR:O	1:B:52:ILE:HG13	2.18	0.44
1:B:23:MET:HE3	1:B:50:CYS:HB3	2.00	0.43
1:A:29:ILE:HD12	1:A:30:ASN:N	2.34	0.43
1:B:7:ASN:HD21	1:B:131:MET:HE1	1.85	0.42
1:B:8:LEU:HD11	2:B:170:HOH:O	2.20	0.42
1:A:48:THR:O	1:A:52:ILE:HG12	2.19	0.42
1:B:7:ASN:OD1	1:B:131:MET:CE	2.68	0.41
1:A:7:ASN:OD1	1:A:131:MET:CE	2.69	0.41
1:B:23:MET:CE	1:B:54:CYS:HB2	2.51	0.40
1:B:16:LEU:HD22	1:B:29:ILE:HD12	2.03	0.40

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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	134/142 (94%)	129 (96%)	5 (4%)	0	100 100
1	B	140/142 (99%)	128 (91%)	8 (6%)	4 (3%)	4 3
All	All	274/284 (96%)	257 (94%)	13 (5%)	4 (2%)	10 10

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	6	LYS
1	B	8	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	7	ASN
1	B	4	VAL

5.3.2 Protein sidechains (i)

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	114/120 (95%)	111 (97%)	3 (3%)	46 63
1	B	120/120 (100%)	108 (90%)	12 (10%)	7 9
All	All	234/240 (98%)	219 (94%)	15 (6%)	17 23

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	8	LEU
1	A	17	ASP
1	A	42	GLU
1	B	1	SER
1	B	2	GLN
1	B	3	GLU
1	B	4	VAL
1	B	6	LYS
1	B	8	LEU
1	B	42	GLU
1	B	44	LYS
1	B	47	GLU
1	B	76	PHE
1	B	105	ASP
1	B	133	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	67	ASN

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 monosaccharides 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 (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	136/142 (95%)	-0.38	2 (1%)	73	79	14, 23, 42, 108	0
1	B	142/142 (100%)	-0.03	8 (5%)	24	30	14, 25, 67, 109	0
All	All	278/284 (97%)	-0.20	10 (3%)	42	49	14, 24, 43, 109	0

All (10) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	2	GLN	13.2
1	B	1	SER	12.2
1	A	7	ASN	5.4
1	B	7	ASN	5.2
1	A	8	LEU	4.7
1	B	6	LYS	4.6
1	B	8	LEU	4.5
1	B	3	GLU	4.5
1	B	5	MET	4.0
1	B	4	VAL	2.8

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 monosaccharides 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.