



Full wwPDB X-ray Structure Validation Report ⓘ

Dec 9, 2023 – 03:16 pm GMT

PDB ID : 2VW9
Title : Single stranded DNA binding protein complex from Helicobacter pylori
Authors : Chan, K.-W.; Wang, C.-H.; Lee, Y.-J.; Sun, Y.-J.
Deposited on : 2008-06-18
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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Xtriage (Phenix) : 1.13
EDS : 2.36
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) : 2.36

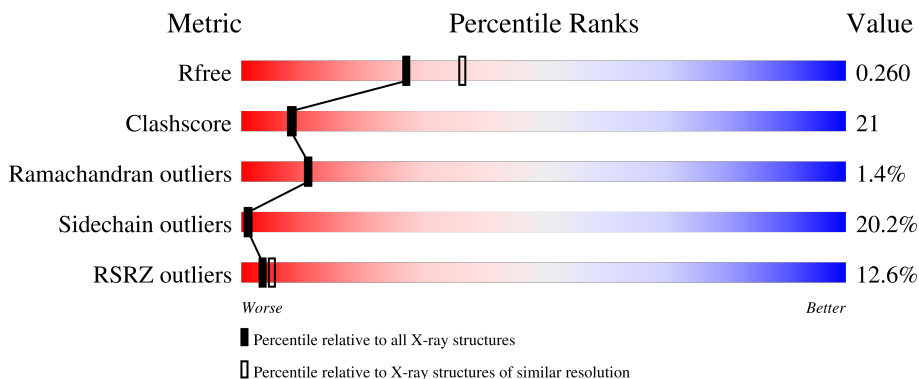
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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 $\leq 5\%$. 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	134	
1	B	134	
2	C	35	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2272 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 SINGLE-STRANDED DNA BINDING PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	109	Total 836	C 522	N 150	O 159	S 5	0	0	1
1	B	106	Total 811	C 508	N 146	O 152	S 5	0	0	1

- Molecule 2 is a DNA chain called POLY-DT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	C	24	Total 465	C 240	N 48	O 158	P 19	0	0	0

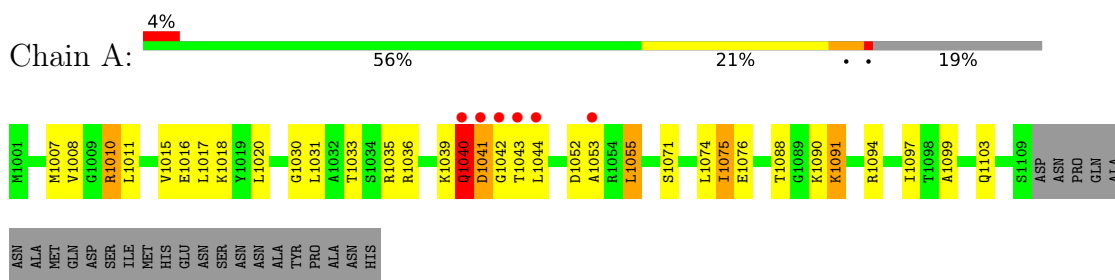
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	85	Total 85	O 85	0	0
3	B	47	Total 47	O 47	0	0
3	C	28	Total 28	O 28	0	0

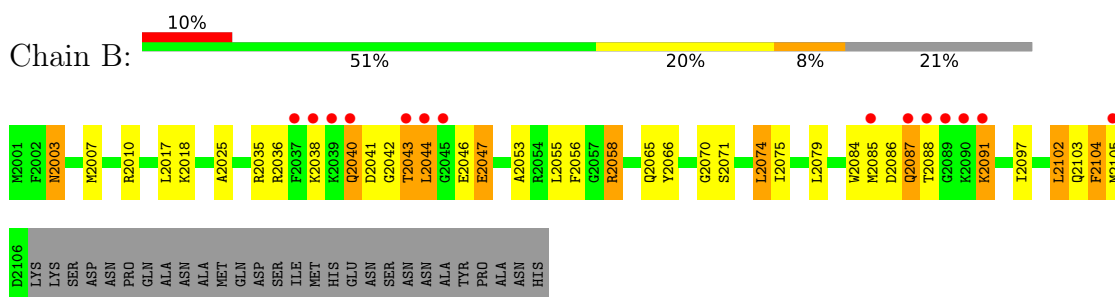
3 Residue-property plots [i](#)

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.

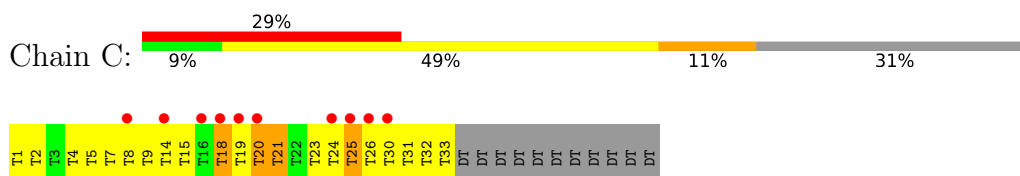
- Molecule 1: SINGLE-STRANDED DNA BINDING PROTEIN



- Molecule 1: SINGLE-STRANDED DNA BINDING PROTEIN



- Molecule 2: POLY-DT



4 Data and refinement statistics i

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α , β , γ	75.56Å 75.56Å 117.69Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	25.13 – 2.30 25.13 – 2.30	Depositor EDS
% Data completeness (in resolution range)	89.5 (25.13-2.30) 93.9 (25.13-2.30)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	6.74 (at 2.31Å)	Xtrriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.238 , 0.260 0.240 , 0.260	Depositor DCC
R_{free} test set	843 reflections (4.83%)	wwPDB-VP
Wilson B-factor (Å ²)	40.8	Xtrriage
Anisotropy	0.201	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 55.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.033 for -h,-k,l	Xtrriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	2272	wwPDB-VP
Average B, all atoms (Å ²)	49.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.27% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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 [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	0.57	0/847	0.78	0/1138
1	B	0.56	0/822	0.79	0/1105
2	C	0.91	0/508	1.41	7/776 (0.9%)
All	All	0.66	0/2177	0.99	7/3019 (0.2%)

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

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	20	DT	O4'-C1'-N1	8.16	113.71	108.00
2	C	25	DT	O4'-C1'-C2'	6.40	111.02	105.90
2	C	25	DT	O4'-C4'-C3'	5.59	109.35	106.00
2	C	32	DT	O4'-C1'-C2'	5.56	110.35	105.90
2	C	18	DT	O4'-C1'-N1	5.50	111.85	108.00
2	C	25	DT	N1-C1'-C2'	5.20	122.47	112.60
2	C	21	DT	O4'-C1'-C2'	5.16	110.03	105.90

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1035	ARG	Sidechain

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	836	0	832	20	0
1	B	811	0	811	34	0
2	C	465	0	298	44	0
3	A	85	0	0	2	0
3	B	47	0	0	1	0
3	C	28	0	0	0	0
All	All	2272	0	1941	86	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 21.

All (86) 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:1091:LYS:H	1:A:1091:LYS:HE2	1.23	0.98
1:B:2087:GLN:H	1:B:2087:GLN:HE21	1.21	0.86
2:C:19:DT:O2	2:C:20:DT:H73	1.78	0.83
1:B:2087:GLN:HE21	1:B:2087:GLN:N	1.82	0.77
1:A:1074:LEU:HB3	1:A:1103:GLN:HB3	1.71	0.72
2:C:30:DT:H2''	2:C:31:DT:OP1	1.89	0.70
1:A:1075:ILE:HD11	1:A:1099:ALA:HB2	1.72	0.69
1:B:2084:TRP:CZ2	2:C:23:DT:H72	2.29	0.67
2:C:19:DT:O2	2:C:20:DT:C7	2.42	0.67
1:B:2075:ILE:HG22	1:B:2102:LEU:HD22	1.75	0.65
2:C:14:DT:H6	2:C:14:DT:H3'	1.60	0.65
1:A:1041:ASP:OD2	1:A:1041:ASP:N	2.27	0.65
2:C:7:DT:H4'	2:C:8:DT:OP1	1.96	0.64
1:B:2056:PHE:CE2	2:C:20:DT:H71	2.32	0.64
1:B:2084:TRP:CZ2	2:C:23:DT:C7	2.80	0.64
1:B:2035:ARG:HD3	1:B:2047:GLU:OE2	1.98	0.63
1:A:1075:ILE:HD11	1:A:1099:ALA:CB	2.29	0.62
1:B:2084:TRP:HZ2	2:C:23:DT:C7	2.12	0.62
2:C:14:DT:H4'	2:C:15:DT:OP1	2.00	0.62
1:B:2018:LYS:NZ	2:C:20:DT:O2	2.27	0.61
2:C:25:DT:C2'	2:C:26:DT:H71	2.29	0.61

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:1:DT:H4'	2:C:2:DT:O5'	2.02	0.60
2:C:18:DT:H2''	2:C:19:DT:H5''	1.84	0.60
1:B:2056:PHE:CZ	2:C:20:DT:H71	2.37	0.59
1:A:1052:ASP:OD1	1:A:1094:ARG:NH1	2.36	0.59
1:B:2038:LYS:HA	1:B:2044:LEU:HA	1.85	0.59
1:A:1091:LYS:HE2	1:A:1091:LYS:N	2.07	0.58
1:A:1055:LEU:HD12	1:A:1099:ALA:HB3	1.83	0.58
1:B:2056:PHE:CE2	2:C:20:DT:C7	2.88	0.57
1:B:2084:TRP:CH2	2:C:23:DT:H72	2.40	0.56
1:A:1091:LYS:H	1:A:1091:LYS:CE	2.10	0.55
1:B:2017:LEU:HD11	1:B:2025:ALA:HB1	1.90	0.54
2:C:14:DT:H3'	2:C:14:DT:C6	2.43	0.54
2:C:25:DT:H2'	2:C:26:DT:H71	1.90	0.54
3:B:3017:HOH:O	2:C:33:DT:H73	2.07	0.53
2:C:18:DT:H1'	2:C:19:DT:H5''	1.90	0.53
2:C:23:DT:O2	2:C:23:DT:H2'	2.08	0.53
1:B:2040:GLN:O	1:B:2042:GLY:N	2.42	0.53
1:B:2104:PHE:HE1	2:C:14:DT:H2'	1.74	0.52
1:A:1040:GLN:O	1:A:1040:GLN:HG2	2.10	0.52
1:B:2043:THR:HG23	1:B:2044:LEU:N	2.25	0.52
2:C:18:DT:C2'	2:C:19:DT:H5''	2.39	0.51
1:B:2040:GLN:HG3	1:B:2043:THR:O	2.11	0.50
2:C:19:DT:C2	2:C:20:DT:C7	2.94	0.49
1:B:2056:PHE:HE2	2:C:20:DT:C7	2.24	0.49
1:A:1075:ILE:HD13	1:A:1076:GLU:N	2.27	0.49
1:B:2074:LEU:HD13	1:B:2075:ILE:N	2.28	0.48
3:A:2006:HOH:O	2:C:33:DT:H4'	2.13	0.48
1:B:2065:GLN:HG3	1:B:2066:TYR:CD1	2.48	0.48
2:C:24:DT:H2''	2:C:25:DT:OP2	2.13	0.48
1:A:1039:LYS:O	1:A:1041:ASP:N	2.47	0.48
2:C:7:DT:H2''	2:C:8:DT:H72	1.96	0.48
1:A:1010:ARG:HA	1:A:1071:SER:O	2.15	0.47
2:C:14:DT:C6	2:C:14:DT:C3'	2.97	0.47
1:B:2074:LEU:HD13	1:B:2074:LEU:C	2.35	0.47
2:C:20:DT:H2''	2:C:21:DT:OP2	2.13	0.46
1:A:1007:MET:HE3	1:A:1007:MET:HB3	1.75	0.46
1:B:2085:MET:SD	1:B:2091:LYS:HB3	2.56	0.46
2:C:19:DT:H2'	2:C:20:DT:C7	2.45	0.45
2:C:7:DT:H2''	2:C:8:DT:C7	2.46	0.45
1:B:2010:ARG:HA	1:B:2071:SER:O	2.17	0.45
2:C:18:DT:H2''	2:C:19:DT:C5'	2.45	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:24:DT:O2	2:C:24:DT:O4'	2.34	0.45
1:B:2091:LYS:HE2	1:B:2091:LYS:HB2	1.74	0.45
1:A:1053:ALA:HA	1:A:1097:ILE:O	2.17	0.44
1:B:2053:ALA:HA	1:B:2097:ILE:O	2.16	0.44
2:C:8:DT:C6	2:C:9:DT:H72	2.52	0.44
2:C:7:DT:C4'	2:C:8:DT:OP1	2.65	0.44
2:C:19:DT:C2	2:C:20:DT:H72	2.53	0.44
1:B:2074:LEU:C	1:B:2074:LEU:CD1	2.87	0.43
1:A:1040:GLN:NE2	3:A:2033:HOH:O	2.51	0.43
1:A:1011:LEU:HA	1:A:1030:GLY:O	2.18	0.43
1:B:2056:PHE:CE2	2:C:20:DT:O4	2.72	0.43
2:C:30:DT:O2	2:C:30:DT:O4'	2.35	0.43
1:B:2003:ASN:HD21	1:B:2079:LEU:H	1.67	0.43
2:C:1:DT:H4'	2:C:2:DT:O4'	2.20	0.42
2:C:4:DT:H1'	2:C:5:DT:H71	2.01	0.42
1:B:2035:ARG:CD	1:B:2047:GLU:OE2	2.66	0.42
1:B:2058:ARG:O	1:B:2058:ARG:HG3	2.19	0.42
1:B:2044:LEU:O	1:B:2044:LEU:HG	2.20	0.42
1:B:2087:GLN:H	1:B:2087:GLN:NE2	2.02	0.41
1:A:1007:MET:HE3	1:A:1033:THR:OG1	2.20	0.41
1:B:2070:GLY:HA2	2:C:26:DT:H2''	2.02	0.41
1:A:1016:GLU:H	1:A:1016:GLU:HG2	1.73	0.41
1:A:1008:VAL:O	1:A:1033:THR:HG23	2.21	0.40
2:C:20:DT:C2'	2:C:21:DT:OP2	2.70	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	107/134 (80%)	103 (96%)	2 (2%)	2 (2%)	8 7

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	104/134 (78%)	100 (96%)	3 (3%)	1 (1%)	15	17
All	All	211/268 (79%)	203 (96%)	5 (2%)	3 (1%)	11	11

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1040	GLN
1	B	2041	ASP
1	A	1042	GLY

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	88/113 (78%)	72 (82%)	16 (18%)	1	1
1	B	85/113 (75%)	66 (78%)	19 (22%)	1	0
All	All	173/226 (76%)	138 (80%)	35 (20%)	1	1

All (35) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1010	ARG
1	A	1015	VAL
1	A	1017	LEU
1	A	1018	LYS
1	A	1020	LEU
1	A	1031	LEU
1	A	1036	ARG
1	A	1040	GLN
1	A	1041	ASP
1	A	1043	THR
1	A	1044	LEU
1	A	1055	LEU
1	A	1075	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	1088	THR
1	A	1090	LYS
1	A	1091	LYS
1	B	2003	ASN
1	B	2007	MET
1	B	2036	ARG
1	B	2040	GLN
1	B	2043	THR
1	B	2044	LEU
1	B	2046	GLU
1	B	2047	GLU
1	B	2055	LEU
1	B	2058	ARG
1	B	2074	LEU
1	B	2086	ASP
1	B	2087	GLN
1	B	2088	THR
1	B	2091	LYS
1	B	2102	LEU
1	B	2103	GLN
1	B	2104	PHE
1	B	2105	MET

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	1065	GLN
1	A	1103	GLN
1	B	2003	ASN
1	B	2040	GLN
1	B	2087	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 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	109/134 (81%)	0.38	6 (5%) 25 31	21, 36, 61, 68	0
1	B	106/134 (79%)	0.76	14 (13%) 3 4	23, 45, 67, 72	0
2	C	24/35 (68%)	1.66	10 (41%) 0 0	46, 80, 91, 96	0
All	All	239/303 (78%)	0.68	30 (12%) 3 5	21, 43, 80, 96	0

All (30) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1044	LEU	5.8
1	B	2044	LEU	5.4
1	B	2038	LYS	5.3
2	C	26	DT	4.9
1	A	1040	GLN	4.6
2	C	18	DT	4.5
1	B	2040	GLN	4.4
1	B	2085	MET	4.1
2	C	25	DT	3.8
2	C	24	DT	3.8
1	A	1041	ASP	3.7
1	B	2105	MET	3.7
1	B	2087	GLN	3.7
1	B	2088	THR	3.7
2	C	14	DT	3.5
2	C	30	DT	3.3
1	B	2039	LYS	3.2
2	C	20	DT	3.2
1	B	2091	LYS	3.2
1	A	1042	GLY	3.0
1	B	2037	PHE	2.8
1	B	2045	GLY	2.8
1	A	1043	THR	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	C	16	DT	2.3
1	A	1053	ALA	2.3
2	C	19	DT	2.3
2	C	8	DT	2.2
1	B	2090	LYS	2.2
1	B	2043	THR	2.2
1	B	2089	GLY	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 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.