wwPDB X-ray Structure Validation Summary Report (i)

Jun 13, 2024 – 09:10 AM EDT

PDB ID	:	4I2V
Title	:	X-ray structure of the unliganded uridine phosphorylase from Yersinia pseudotuberculosis at 2.12A resolution
Authors	:	Lashkov, A.A.; Balaev, V.V.; Prokofev, I.I.; Betzel, C.; Gabdoulkhakov, A.G.; Mikhailov, A.M.
Deposited on	:	2012-11-23
Resolution	:	2.12 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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:

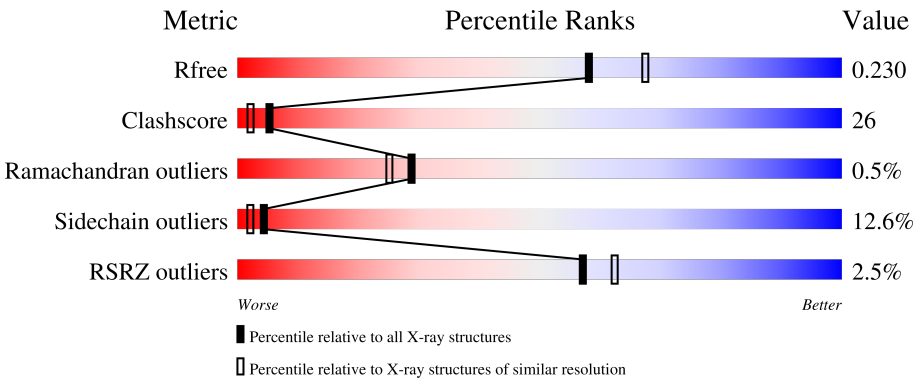
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 2.12 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	6241 (2.14-2.10)
Clashscore	141614	6778 (2.14-2.10)
Ramachandran outliers	138981	6705 (2.14-2.10)
Sidechain outliers	138945	6706 (2.14-2.10)
RSRZ outliers	127900	6112 (2.14-2.10)

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	253	2%, 47%, 46%, 6% •
1	B	253	2%, 48%, 41%, 10% •
1	C	253	2%, 43%, 47%, 8% •
1	D	253	2%, 45%, 47%, 7% •
1	E	253	4%, 48%, 43%, 8% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	253	3% 49% 45% 5% •

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 11322 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 Uridine phosphorylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	250	1875	1173	330	360	12	0	0	0
1	B	250	1875	1173	330	360	12	0	0	0
1	C	250	1875	1173	330	360	12	0	0	0
1	D	250	1875	1173	330	360	12	0	0	0
1	E	250	1875	1173	330	360	12	0	0	0
1	F	250	1875	1173	330	360	12	0	0	0

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	12	12	12	0	0
2	B	17	17	17	0	0
2	C	13	13	13	0	0
2	D	17	17	17	0	0
2	E	4	4	4	0	0
2	F	9	9	9	0	0

3 Residue-property plots (i)

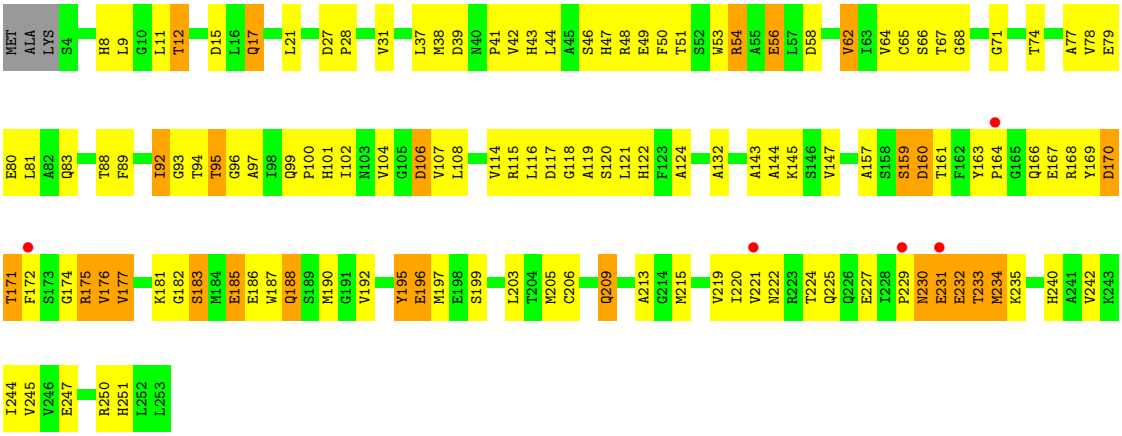
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: Uridine phosphorylase

Chain A:

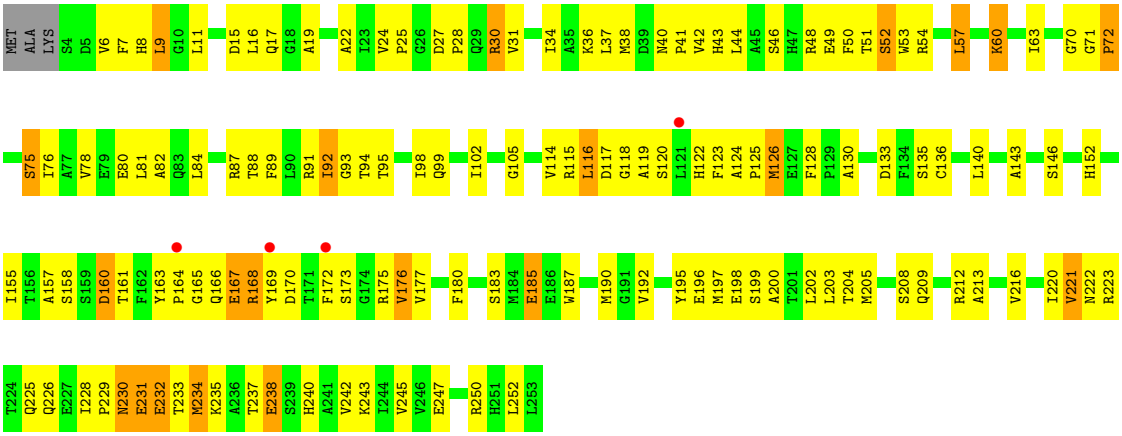
• Molecule 1: Uridine phosphorylase

Chain B:



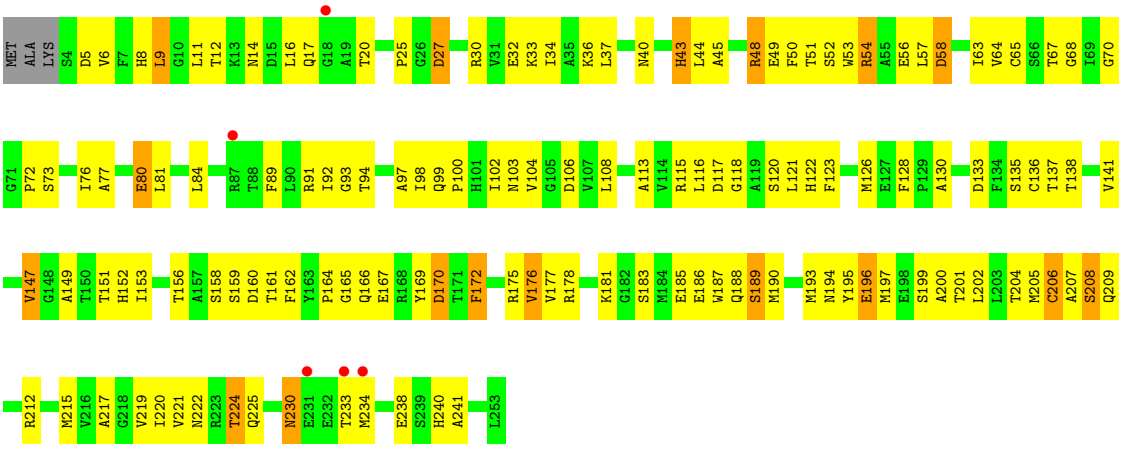
• Molecule 1: Uridine phosphorylase

Chain C:



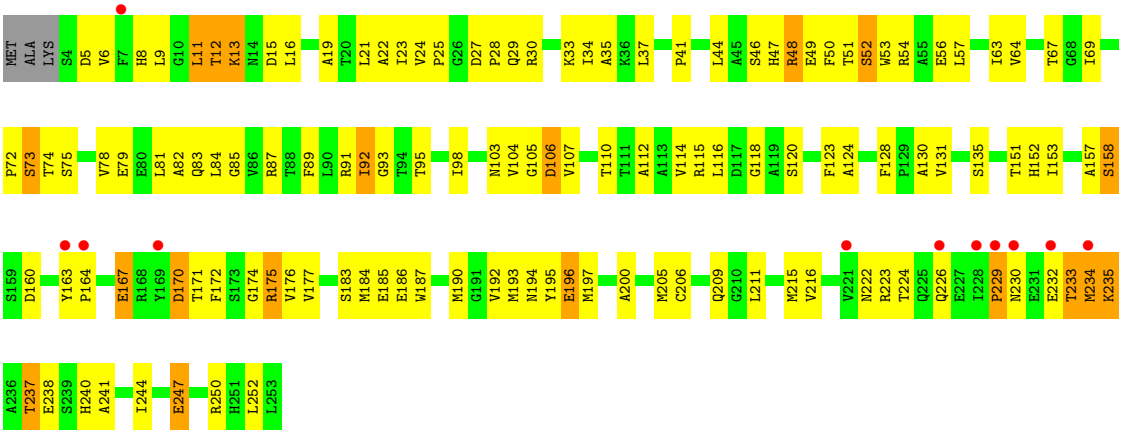
• Molecule 1: Uridine phosphorylase

Chain D:



• Molecule 1: Uridine phosphorylase

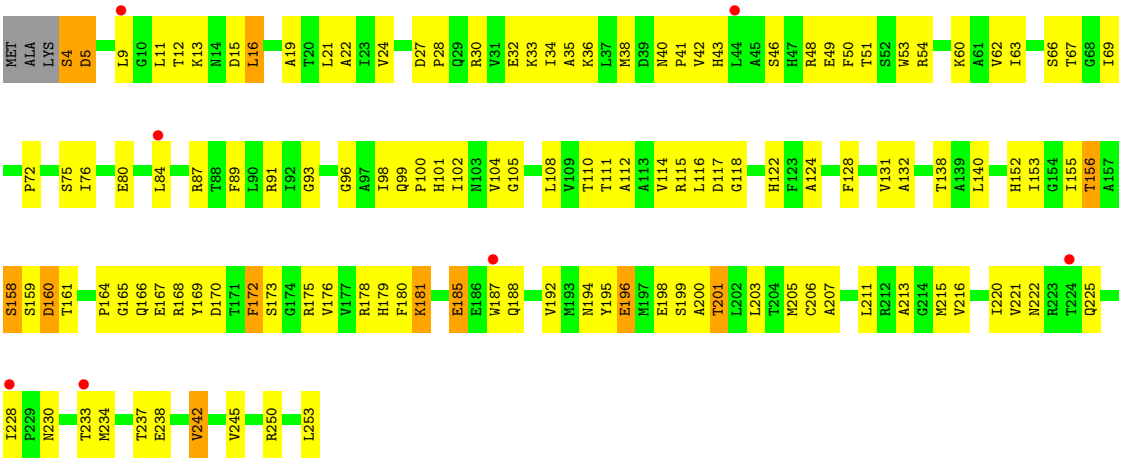
Chain E:



• Molecule 1: Uridine phosphorylase

Chain F:

3%
49%
45%
5%


MET ALA LYS S4 D5 L9 G10 L11 T12 K13 N14 D15 L16 A19 T20 L21 A22 I23 V24 D27 P28 Q29 R30 V31 E32 K33 I34 A35 K36 L37 M38 D39 N40 P41 V42 H43 L44 A45 S46 H47 R48 E49 F50 T51 S52 W53 R54 K60 A61 V62 I63 S66 T67 G68 I69
P72 S75 I76 E80 L84 R87 T88 F89 L90 R91 I92 G93 G96 A97 I98 Q99 P100 H101 I102 N103 V104 G105 L108 V109 T110 T111 A112 A113 V114 R115 L116 D117 G118 H122 F123 A124 F128 V131 A132 T138 A139 L140 H152 I153 G154 I155 T156 A157
S158 S159 D160 T161 P164 G165 Q166 E167 R168 Y169 D170 T171 F172 S173 G174 R175 V176 V177 R178 H179 F180 K181 E185 E186 W187 Q188 V192 M193 N194 Y195 E196 M197 E198 S199 A200 T201 L202 L203 T204 M205 C206 A207 L211 R212 A213 G214 M215 V216 I220 V221 N222 R223 T224 Q225
I228 P229 N230 T233 M234 T237 E238 V242 V245 R250 L253

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 3	Depositor
Cell constants a, b, c, α, β, γ	156.68Å 156.68Å 48.46Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	29.61 – 2.12 29.61 – 2.12	Depositor EDS
% Data completeness (in resolution range)	97.6 (29.61-2.12) 97.6 (29.61-2.12)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.71 (at 2.12Å)	Xtriage
Refinement program	PHENIX 1.8.1_1168	Depositor
R, R_{free}	0.200 , 0.242 0.203 , 0.230	Depositor DCC
R_{free} test set	3680 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	31.6	Xtriage
Anisotropy	0.246	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 14.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.41$, $< L^2 > = 0.23$	Xtriage
Estimated twinning fraction	0.427 for -h,-k,l 0.075 for h,-h-k,-l 0.074 for -k,-h,-l	Xtriage
Reported twinning fraction	0.460 for -h,-k,l	Depositor
Outliers	0 of 73600 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	11322	wwPDB-VP
Average B, all atoms (Å^2)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.49% 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 (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	0.46	1/1908 (0.1%)	0.67	0/2590
1	B	0.45	0/1908	0.69	0/2590
1	C	0.43	0/1908	0.69	0/2590
1	D	0.44	0/1908	0.67	0/2590
1	E	0.41	0/1908	0.69	0/2590
1	F	0.43	0/1908	0.66	0/2590
All	All	0.44	1/11448 (0.0%)	0.68	0/15540

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	B	0	1
1	E	0	1
All	All	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	136	CYS	CB-SG	-5.05	1.73	1.81

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	232	GLU	Peptide
1	E	232	GLU	Peptide

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	1875	0	1869	117	0
1	B	1875	0	1869	104	0
1	C	1875	0	1869	113	0
1	D	1875	0	1869	119	0
1	E	1875	0	1869	97	0
1	F	1875	0	1869	94	0
2	A	12	0	0	1	0
2	B	17	0	0	0	0
2	C	13	0	0	1	0
2	D	17	0	0	4	0
2	E	4	0	0	0	0
2	F	9	0	0	1	0
All	All	11322	0	11214	595	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 26.

The worst 5 of 595 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:122:HIS:HB3	1:B:164:PRO:HB2	1.43	1.00
1:F:16:LEU:HG	1:F:63:ILE:HG13	1.48	0.93
1:D:49:GLU:HG2	1:D:68:GLY:HA2	1.55	0.88
1:E:167:GLU:HG3	1:E:183:SER:HB2	1.54	0.87
1:E:164:PRO:HB2	1:F:122:HIS:HB3	1.57	0.85

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	248/253 (98%)	228 (92%)	20 (8%)	0	100	100
1	B	248/253 (98%)	236 (95%)	11 (4%)	1 (0%)	34	32
1	C	248/253 (98%)	227 (92%)	17 (7%)	4 (2%)	9	4
1	D	248/253 (98%)	229 (92%)	18 (7%)	1 (0%)	34	32
1	E	248/253 (98%)	228 (92%)	18 (7%)	2 (1%)	19	14
1	F	248/253 (98%)	233 (94%)	15 (6%)	0	100	100
All	All	1488/1518 (98%)	1381 (93%)	99 (7%)	8 (0%)	29	25

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	126	MET
1	C	221	VAL
1	C	116	LEU
1	B	176	VAL
1	C	176	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	201/203 (99%)	179 (89%)	22 (11%)	6	3
1	B	201/203 (99%)	173 (86%)	28 (14%)	3	1
1	C	201/203 (99%)	174 (87%)	27 (13%)	4	2
1	D	201/203 (99%)	174 (87%)	27 (13%)	4	2
1	E	201/203 (99%)	173 (86%)	28 (14%)	3	1
1	F	201/203 (99%)	181 (90%)	20 (10%)	7	4
All	All	1206/1218 (99%)	1054 (87%)	152 (13%)	4	2

5 of 152 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	135	SER
1	F	172	PHE
1	E	170	ASP
1	E	240	HIS
1	F	242	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	152	HIS
1	E	240	HIS
1	D	8	HIS
1	C	166	GLN
1	D	194	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	250/253 (98%)	0.17	6 (2%)	59	64	17, 31, 50, 61	0
1	B	250/253 (98%)	0.04	5 (2%)	65	69	16, 30, 44, 55	0
1	C	250/253 (98%)	0.05	4 (1%)	72	76	18, 30, 46, 53	0
1	D	250/253 (98%)	0.04	5 (2%)	65	69	16, 31, 46, 58	0
1	E	250/253 (98%)	0.19	11 (4%)	34	40	20, 32, 53, 66	0
1	F	250/253 (98%)	0.19	7 (2%)	53	59	19, 34, 50, 59	0
All	All	1500/1518 (98%)	0.11	38 (2%)	57	62	16, 32, 50, 66	0

The worst 5 of 38 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	84	LEU	5.1
1	E	169	TYR	4.4
1	B	229	PRO	4.0
1	A	4	SER	3.8
1	E	164	PRO	3.7

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.