Full wwPDB X-ray Structure Validation Report (i)

Jan 31, 2016 – 09:02 PM GMT

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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

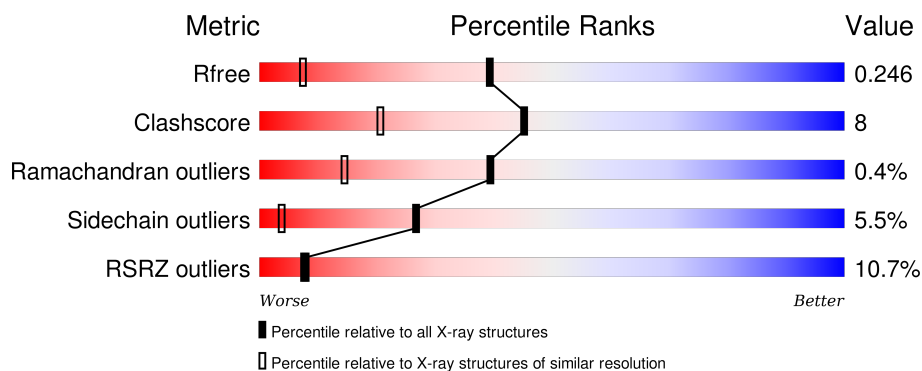|		
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 1.55 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		0.246
Clashscore		8
Ramachandran outliers		0.4%
Sidechain outliers		5.5%
RSRZ outliers		10.7%

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$_{free}$	91344	1665 (1.58-1.54)
Clashscore	102246	1014 (1.56-1.56)
Ramachandran outliers	100387	1704 (1.58-1.54)
Sidechain outliers	100360	1702 (1.58-1.54)
RSRZ outliers	91569	1668 (1.58-1.54)

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 >=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 <=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	721	11% / 83% / 14% / •

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 6441 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 HYALURONIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	721	Total 5771	C 3626	N 967	O 1156	S 22	0	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	292	ALA	TRP	ENGINEERED	UNP Q54873
A	343	VAL	PHE	ENGINEERED	UNP Q54873
A	731	VAL	GLY	SEE REMARK 999	UNP Q54873

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	670	Total 670	O 670	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: HYALURONIDASE

Chain A:
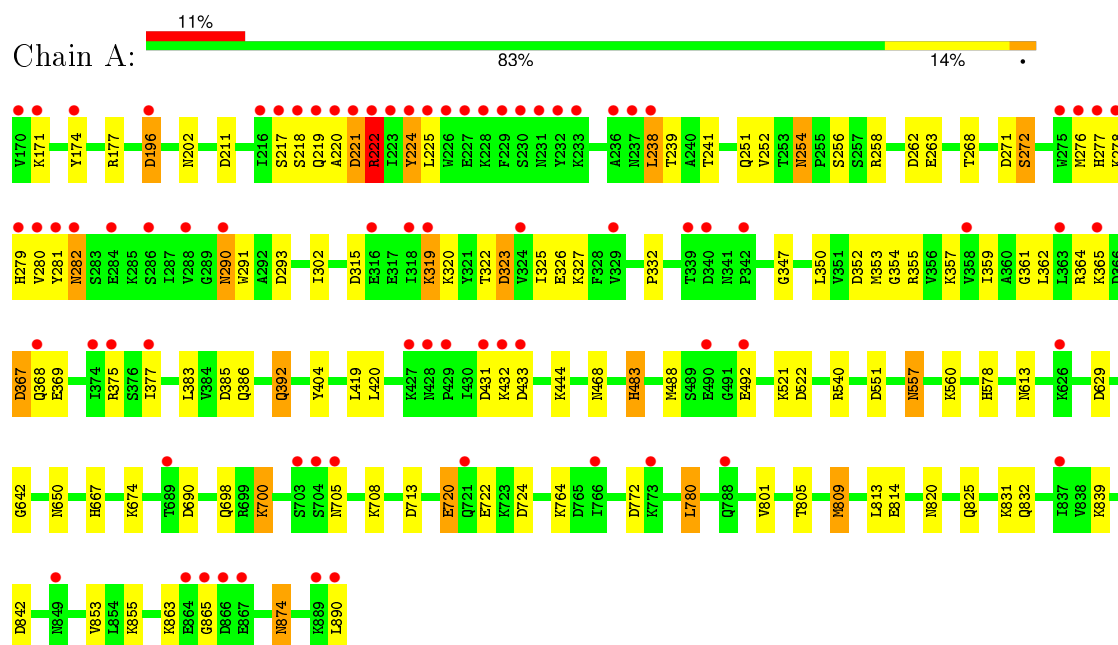
4   Data and refinement statistics   (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	83.98Å 103.72Å 100.52Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 1.55 19.91 – 1.55	Depositor EDS
% Data completeness (in resolution range)	99.5 (20.00-1.55) 99.4 (19.91-1.55)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) >$ [1]	1.51 (at 1.55Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.190 , 0.234 0.207 , 0.246	Depositor DCC
R_{free} test set	1275 reflections (1.02%)	DCC
Wilson B-factor (Å^2)	19.5	Xtriage
Anisotropy	0.137	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.40 , 54.7	EDS
Estimated twinning fraction	0.015 for -h,l,k	Xtriage
L-test for twinning [2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Outliers	0 of 127182 reflections	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	6441	wwPDB-VP
Average B, all atoms (Å^2)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.07% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.62	1/5887 (0.0%)	0.84	20/7951 (0.3%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	809	MET	SD-CE	-5.12	1.49	1.77

All (20) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	271	ASP	CB-CG-OD2	8.66	126.09	118.30
1	A	323	ASP	CB-CG-OD2	7.97	125.48	118.30
1	A	724	ASP	CB-CG-OD2	7.39	124.95	118.30
1	A	842	ASP	CB-CG-OD2	7.33	124.90	118.30
1	A	315	ASP	CB-CG-OD2	6.97	124.57	118.30
1	A	690	ASP	CB-CG-OD2	6.56	124.20	118.30
1	A	211	ASP	CB-CG-OD2	6.54	124.19	118.30
1	A	262	ASP	CB-CG-OD2	6.01	123.71	118.30
1	A	367	ASP	CB-CG-OD2	5.86	123.57	118.30
1	A	713	ASP	CB-CG-OD2	5.58	123.32	118.30
1	A	431	ASP	CB-CG-OD2	5.43	123.19	118.30
1	A	433	ASP	CB-CG-OD2	5.42	123.18	118.30
1	A	196	ASP	CB-CG-OD2	5.41	123.17	118.30
1	A	293	ASP	CB-CG-OD2	5.41	123.17	118.30
1	A	352	ASP	CB-CG-OD2	5.34	123.11	118.30
1	A	780	LEU	CA-CB-CG	5.32	127.52	115.30
1	A	780	LEU	CB-CG-CD1	5.29	120.00	111.00
1	A	551	ASP	CB-CG-OD2	5.24	123.02	118.30
1	A	522	ASP	CB-CG-OD2	5.22	123.00	118.30
1	A	540	ARG	NE-CZ-NH1	5.04	122.82	120.30

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	5771	0	5590	88	0
2	A	670	0	0	56	1
All	All	6441	0	5590	88	1

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 8.

All (88) 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:890:LEU:C	2:A:1541:HOH:O	1.85	1.12
1:A:613:ASN:H	1:A:698:GLN:HE22	1.14	0.93
1:A:863:LYS:NZ	1:A:865:GLY:O	2.04	0.91
1:A:355:ARG:C	2:A:1490:HOH:O	2.18	0.80
1:A:650:ASN:HD21	1:A:832:GLN:HE22	1.32	0.78
1:A:392:GLN:NE2	2:A:1480:HOH:O	2.17	0.75
1:A:357:LYS:HD3	2:A:1420:HOH:O	1.89	0.71
1:A:177:ARG:HG3	2:A:1495:HOH:O	1.90	0.71
1:A:323:ASP:HA	2:A:1500:HOH:O	1.91	0.70
1:A:319:LYS:HG2	2:A:1506:HOH:O	1.92	0.69
1:A:354:GLY:HA3	1:A:377:ILE:HD11	1.75	0.68
1:A:369:GLU:HG2	2:A:1516:HOH:O	1.92	0.67
1:A:483:HIS:HE1	2:A:1215:HOH:O	1.77	0.67
1:A:277:HIS:NE2	2:A:1386:HOH:O	2.27	0.67
1:A:238:LEU:HD12	2:A:1436:HOH:O	1.94	0.66
1:A:357:LYS:HB3	2:A:1516:HOH:O	1.95	0.66
1:A:801:VAL:HG11	1:A:809:MET:HE1	1.78	0.66
1:A:362:LEU:CD2	2:A:1495:HOH:O	2.44	0.66
1:A:362:LEU:HA	2:A:1481:HOH:O	1.95	0.65
1:A:874:ASN:HD22	1:A:874:ASN:C	1.99	0.65
1:A:320:LYS:HG3	2:A:1386:HOH:O	1.96	0.64
1:A:805:THR:HG22	1:A:809:MET:HE2	1.79	0.63

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:557:ASN:HD22	1:A:557:ASN:C	2.04	0.60
1:A:277:HIS:CE1	2:A:1386:HOH:O	2.55	0.60
1:A:353:MET:HG3	2:A:1420:HOH:O	2.01	0.60
1:A:705:ASN:ND2	2:A:1402:HOH:O	2.36	0.59
1:A:302:ILE:HG21	2:A:1511:HOH:O	2.03	0.58
1:A:674:LYS:NZ	2:A:1169:HOH:O	2.37	0.58
1:A:805:THR:HG22	1:A:809:MET:CE	2.33	0.58
1:A:350:LEU:HB3	2:A:1389:HOH:O	2.03	0.58
1:A:322:THR:HG21	2:A:1549:HOH:O	2.04	0.57
1:A:361:GLY:O	2:A:1481:HOH:O	2.17	0.57
1:A:364:ARG:CZ	2:A:1506:HOH:O	2.53	0.56
1:A:362:LEU:HD22	2:A:1495:HOH:O	2.07	0.55
1:A:217:SER:HB3	1:A:222:ARG:HG2	1.88	0.55
1:A:325:ILE:HA	2:A:1357:HOH:O	2.06	0.55
1:A:241:THR:HG21	1:A:276:MET:HE3	1.87	0.55
1:A:814:GLU:HG2	2:A:1364:HOH:O	2.07	0.54
1:A:359:ILE:HG12	2:A:1490:HOH:O	2.08	0.54
1:A:557:ASN:ND2	1:A:560:LYS:H	2.06	0.54
1:A:720:GLU:CG	2:A:1318:HOH:O	2.56	0.54
1:A:801:VAL:HG21	1:A:809:MET:HE3	1.90	0.54
1:A:332:PRO:HB3	2:A:1420:HOH:O	2.08	0.53
1:A:831:LYS:NZ	2:A:1383:HOH:O	2.29	0.53
1:A:839:LYS:HD2	1:A:853:VAL:HG23	1.91	0.52
1:A:831:LYS:CE	2:A:1383:HOH:O	2.57	0.52
1:A:667:HIS:HE1	2:A:1228:HOH:O	1.92	0.52
1:A:347:GLY:O	2:A:1389:HOH:O	2.19	0.51
1:A:254:ASN:HD22	1:A:254:ASN:C	2.12	0.51
1:A:225:LEU:HD12	1:A:276:MET:HE3	1.92	0.51
1:A:764:LYS:HD2	1:A:772:ASP:HB3	1.92	0.51
1:A:386:GLN:NE2	2:A:1056:HOH:O	2.44	0.51
1:A:820:ASN:HD22	1:A:825:GLN:HG2	1.74	0.51
1:A:420:LEU:HB3	1:A:488:MET:HE1	1.93	0.50
1:A:578:HIS:CD2	1:A:578:HIS:C	2.85	0.50
1:A:281:TYR:OH	2:A:1357:HOH:O	2.18	0.50
1:A:279:HIS:CE1	2:A:1380:HOH:O	2.66	0.49
1:A:280:VAL:HG13	2:A:1487:HOH:O	2.12	0.49
1:A:720:GLU:OE2	2:A:1318:HOH:O	2.20	0.49
1:A:290:ASN:HD22	1:A:291:TRP:N	2.10	0.49
1:A:700:LYS:HD2	2:A:961:HOH:O	2.12	0.48
1:A:217:SER:CB	1:A:222:ARG:HG2	2.43	0.48
1:A:326:GLU:HB3	2:A:1484:HOH:O	2.12	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:282:ASN:HB2	2:A:1292:HOH:O	2.12	0.48
1:A:174:TYR:HA	2:A:1495:HOH:O	2.14	0.47
1:A:224:TYR:O	1:A:224:TYR:HD2	1.98	0.47
1:A:280:VAL:HG22	2:A:1487:HOH:O	2.14	0.47
1:A:354:GLY:CA	1:A:377:ILE:HD11	2.44	0.47
1:A:222:ARG:HD3	2:A:1423:HOH:O	2.15	0.46
1:A:708:LYS:NZ	2:A:1543:HOH:O	2.47	0.46
1:A:225:LEU:HG	1:A:276:MET:HE2	1.97	0.46
1:A:578:HIS:CE1	1:A:629:ASP:O	2.69	0.45
1:A:367:ASP:N	2:A:1415:HOH:O	2.49	0.44
1:A:225:LEU:N	2:A:1423:HOH:O	2.50	0.44
1:A:254:ASN:ND2	1:A:256:SER:H	2.16	0.44
1:A:171:LYS:HB2	2:A:1542:HOH:O	2.17	0.44
1:A:353:MET:CG	2:A:1420:HOH:O	2.64	0.44
1:A:202:ASN:ND2	1:A:251:GLN:HE22	2.15	0.44
1:A:241:THR:HG21	1:A:276:MET:CE	2.47	0.44
1:A:722:GLU:HG3	2:A:1536:HOH:O	2.17	0.44
1:A:642:GLY:HA2	2:A:1366:HOH:O	2.18	0.43
1:A:302:ILE:HD13	2:A:1511:HOH:O	2.19	0.43
1:A:220:ALA:O	1:A:222:ARG:N	2.51	0.43
1:A:364:ARG:HG2	2:A:1506:HOH:O	2.19	0.42
1:A:432:LYS:HD2	2:A:1534:HOH:O	2.20	0.42
1:A:855:LYS:HE3	1:A:855:LYS:HB3	1.82	0.41
1:A:365:LYS:HA	2:A:1481:HOH:O	2.18	0.41
1:A:268:THR:O	1:A:272:SER:HB2	2.21	0.41

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:1379:HOH:O	2:A:1540:HOH:O[4_446]	2.12	0.08

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	719/721 (100%)	692 (96%)	24 (3%)	3 (0%)	39	14

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	219	GLN
1	A	221	ASP
1	A	222	ARG

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	637/639 (100%)	602 (94%)	35 (6%)	27	4

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

Mol	Chain	Res	Type
1	A	196	ASP
1	A	218	SER
1	A	221	ASP
1	A	222	ARG
1	A	224	TYR
1	A	238	LEU
1	A	239	THR
1	A	252	VAL
1	A	254	ASN
1	A	258	ARG
1	A	263	GLU
1	A	272	SER
1	A	278	LYS
1	A	282	ASN
1	A	290	ASN
1	A	319	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	327	LYS
1	A	368	GLN
1	A	375	ARG
1	A	383	LEU
1	A	385	ASP
1	A	392	GLN
1	A	404	TYR
1	A	419	LEU
1	A	444	LYS
1	A	468	ASN
1	A	483	HIS
1	A	492	GLU
1	A	521	LYS
1	A	557	ASN
1	A	700	LYS
1	A	720	GLU
1	A	780	LEU
1	A	813	LEU
1	A	874	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (20) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	202	ASN
1	A	237	ASN
1	A	254	ASN
1	A	261	GLN
1	A	282	ASN
1	A	290	ASN
1	A	349	ASN
1	A	386	GLN
1	A	436	GLN
1	A	468	ASN
1	A	483	HIS
1	A	557	ASN
1	A	578	HIS
1	A	667	HIS
1	A	698	GLN
1	A	759	GLN
1	A	820	ASN
1	A	825	GLN
1	A	832	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	874	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 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 (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(Å²)	Q<0.9
1	A	721/721 (100%)	0.50	77 (10%)	8	7	12, 22, 46, 66	0

All (77) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	170	VAL	19.8
1	A	223	ILE	9.5
1	A	224	TYR	9.1
1	A	232	TYR	8.3
1	A	219	GLN	7.0
1	A	225	LEU	6.4
1	A	221	ASP	6.4
1	A	427	LYS	6.1
1	A	233	LYS	5.6
1	A	174	TYR	5.5
1	A	866	ASP	5.0
1	A	340	ASP	4.9
1	A	319	LYS	4.7
1	A	171	LYS	4.5
1	A	433	ASP	4.4
1	A	284	GLU	4.1
1	A	279	HIS	4.0
1	A	229	PHE	3.8
1	A	238	LEU	3.7
1	A	281	TYR	3.5
1	A	375	ARG	3.4
1	A	342	PRO	3.3
1	A	849	ASN	3.3
1	A	766	ILE	3.3
1	A	431	ASP	3.2
1	A	218	SER	3.2
1	A	277	HIS	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	704	SER	3.1
1	A	368	GLN	3.1
1	A	288	VAL	3.0
1	A	275	TRP	3.0
1	A	280	VAL	3.0
1	A	689	THR	3.0
1	A	290	ASN	3.0
1	A	220	ALA	3.0
1	A	837	ILE	3.0
1	A	276	MET	2.9
1	A	865	GLY	2.9
1	A	363	LEU	2.8
1	A	227	GLU	2.8
1	A	316	GLU	2.8
1	A	226	TRP	2.8
1	A	864	GLU	2.8
1	A	222	ARG	2.8
1	A	432	LYS	2.7
1	A	721	GLN	2.7
1	A	490	GLU	2.5
1	A	237	ASN	2.5
1	A	492	GLU	2.5
1	A	703	SER	2.5
1	A	230	SER	2.5
1	A	374	ILE	2.5
1	A	228	LYS	2.5
1	A	236	ALA	2.5
1	A	286	SER	2.4
1	A	890	LEU	2.4
1	A	282	ASN	2.4
1	A	318	ILE	2.4
1	A	867	GLU	2.4
1	A	329	VAL	2.3
1	A	216	ILE	2.3
1	A	377	ILE	2.2
1	A	358	VAL	2.2
1	A	217	SER	2.2
1	A	196	ASP	2.1
1	A	705	ASN	2.1
1	A	278	LYS	2.1
1	A	889	LYS	2.1
1	A	339	THR	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	365	LYS	2.1
1	A	626	LYS	2.1
1	A	428	ASN	2.1
1	A	324	VAL	2.1
1	A	231	ASN	2.0
1	A	429	PRO	2.0
1	A	788	GLN	2.0
1	A	773	LYS	2.0

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.