WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 28, 2023 – 02:55 AM EDT

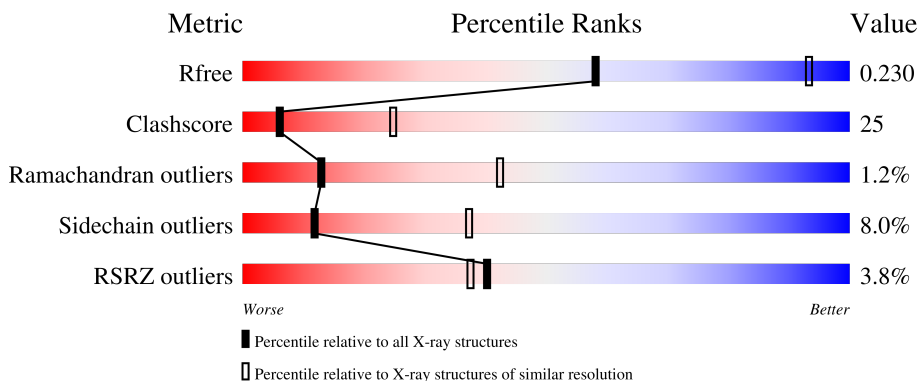|		
PDB ID	:	3JVZ
Title	:	E2 Ubiquitin-HECT
Authors	:	Souphron, J.; Kamadurai, H.B.; Schulman, B.A.
Deposited on	:	2009-09-17
Resolution	:	3.30 Å(reported)

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 3.30 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	Percentile Ranks	Value
Rfree		0.230
Clashscore		25
Ramachandran outliers		1.2%
Sidechain outliers		8.0%
RSRZ outliers		3.8%

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}	130704	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

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	146	55% 42% •
1	B	146	54% 38% 7% •
2	C	385	% 50% 44% • •
2	D	385	51% 41% 6% •
3	X	81	28% 46% 44% • 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	Y	81	

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 9814 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 Ubiquitin-conjugating enzyme E2 D2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	146	Total 1170	C 749	N 202	O 213	S 6	0	0	0
1	B	146	Total 1170	C 749	N 202	O 213	S 6	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	3	SER	LEU	engineered mutation	UNP P62837
A	85	SER	CYS	engineered mutation	UNP P62837
A	98	LYS	THR	engineered mutation	UNP P62837
B	3	SER	LEU	engineered mutation	UNP P62837
B	85	SER	CYS	engineered mutation	UNP P62837
B	98	LYS	THR	engineered mutation	UNP P62837

- Molecule 2 is a protein called E3 ubiquitin-protein ligase NEDD4-like.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	377	Total 3148	C 2037	N 520	O 574	S 17	0	0	0
2	D	374	Total 3124	C 2021	N 517	O 569	S 17	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	571	GLY	-	expression tag	UNP Q96PU5
C	572	SER	-	expression tag	UNP Q96PU5
C	573	PRO	-	expression tag	UNP Q96PU5
C	574	GLU	-	expression tag	UNP Q96PU5
C	575	PHE	-	expression tag	UNP Q96PU5
C	922	ALA	CYS	engineered mutation	UNP Q96PU5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	571	GLY	-	expression tag	UNP Q96PU5
D	572	SER	-	expression tag	UNP Q96PU5
D	573	PRO	-	expression tag	UNP Q96PU5
D	574	GLU	-	expression tag	UNP Q96PU5
D	575	PHE	-	expression tag	UNP Q96PU5
D	922	ALA	CYS	engineered mutation	UNP Q96PU5

- Molecule 3 is a protein called Ubiquitin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	X	76	Total 601	C 378	N 105	O 117	S 1	0	0	0
3	Y	76	Total 601	C 378	N 105	O 117	S 1	0	0	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
X	-4	GLY	-	expression tag	UNP P62988
X	-3	SER	-	expression tag	UNP P62988
X	-2	GLY	-	expression tag	UNP P62988
X	-1	GLY	-	expression tag	UNP P62988
X	0	SER	-	expression tag	UNP P62988
Y	-4	GLY	-	expression tag	UNP P62988
Y	-3	SER	-	expression tag	UNP P62988
Y	-2	GLY	-	expression tag	UNP P62988
Y	-1	GLY	-	expression tag	UNP P62988
Y	0	SER	-	expression tag	UNP P62988

3 Residue-property plots ⓘ

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: Ubiquitin-conjugating enzyme E2 D2

- Molecule 1: Ubiquitin-conjugating enzyme E2 D2

- Molecule 2: E3 ubiquitin-protein ligase NEDD4-like

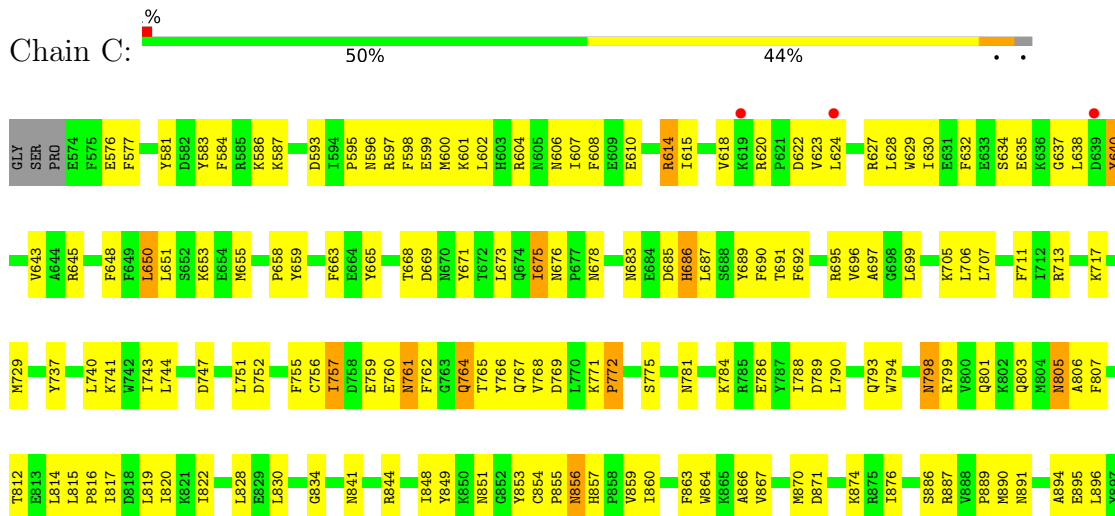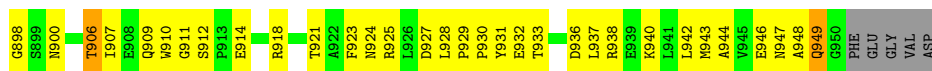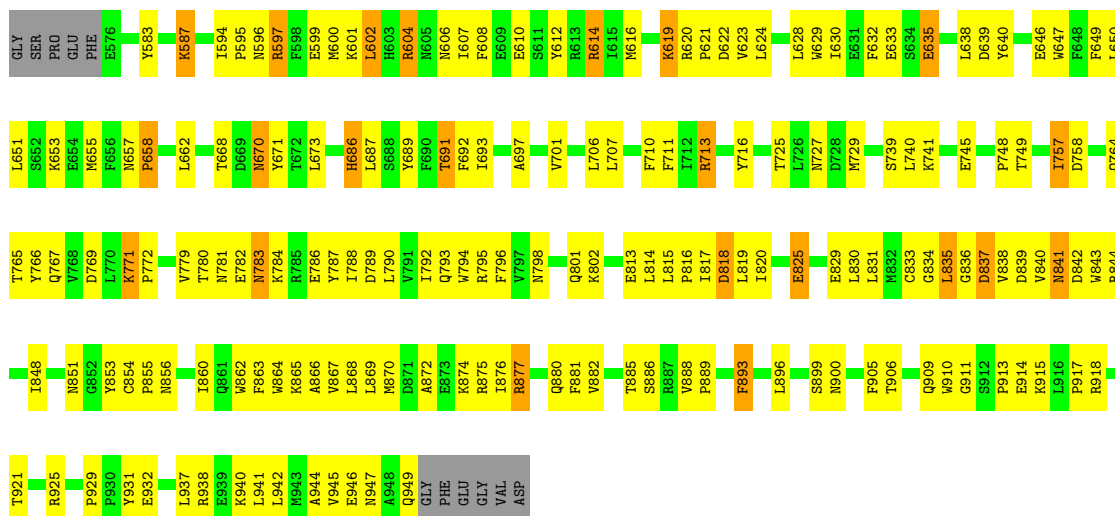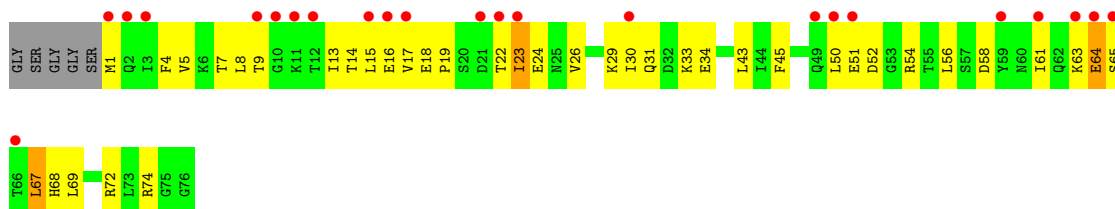G898 S899 N900 T906 I907 E908 Q909 W910 G911 S912 P913 E914 R918 T921 A922 F923 N924 R925 L926 D927 L928 P929 P930 Y931 E932 T933 D936 L937 R938 E939 K940 L941 L942 M943 A944 V945 E946 N947 A948 Q949 G950 PHE GLU GLY VAL ASP

• Molecule 2: E3 ubiquitin-protein ligase NEDD4-like

Chain D: 51%   41%   6%   •

GLY SER PRO PHE E576 Y583 K587 I594 P595 N596 R597 F598 E599 M600 K601 L602 H603 R604 N605 N606 I607 F608 E609 E610 S611 Y612 R613 R614 I615 M616 K619 R620 P621 D622 V623 L624 L628 W629 I630 E631 F632 E633 S634 E635 L638 D639 Y640 E646 W647 F648 F649 L650

L651 S652 K653 E654 M655 F656 N657 P658 L662 T668 D669 N670 Y671 T672 L673 H686 L687 S688 Y689 F690 T691 F692 I693 A697 V701 L706 L707 F710 F711 I712 R713 Y716 T725 L726 N727 D728 M729 S739 L740 K741 E745 P748 T749 I757 D758 Q764

T765 Y766 Q767 V768 D769 L770 K771 P772 V779 T780 N781 E782 N783 K784 R785 E786 Y787 I788 D789 L790 V791 I792 Q793 W794 R795 F796 V797 N798 Q801 K802 E813 L814 L815 P816 I817 D818 L819 I820 E825 E829 L830 L831 M832 C833 G834 L835 G836 D837 V838 D839 V840 N841 D842 W843 R844

I848 N851 G852 Y853 C854 P855 N856 I860 Q861 W862 F863 W864 K865 A866 V867 L868 L869 M870 D871 A872 E873 K874 R875 I876 R877 Q880 F881 V882 T885 S886 R887 V888 P889 F893 L896 S899 N900 F905 T906 Q909 W910 G911 S912 P913 E914 K915 L916 P917 R918

T921 R925 P929 P930 Y931 E932 L937 R938 E939 K940 L941 L942 M943 A944 V945 E946 N947 A948 Q949 GLY PHE GLU GLY VAL ASP

• Molecule 3: Ubiquitin

Chain X: 28%   46%   44%   •   6%

GLY SER GLY SER M1 Q2 I3 F4 V5 K6 T7 L8 T9 G10 K11 T12 I13 T14 L15 E16 V17 E18 P19 S20 D21 T22 I23 E24 N25 V26 K29 I30 Q31 D32 K33 E34 L43 I44 F45 Q49 L50 E51 D52 G53 R54 T55 L56 S57 D58 Y59 N60 I61 Q62 K63 E64 S65

T66 L67 H68 L69 R72 L73 R74 G75 G76

• Molecule 3: Ubiquitin

Chain Y: 25%   44%   46%   •   6%

GLY SER GLY SER M1 Q2 I3 F4 V5 K6 T7 L8 I13 T14 L15 E16 V17 D21 T22 I23 E24 N25 V26 K27 A28 K29 I30 Q31 D32 K33 E34 G35 I36 P37 Q40 Q41 R42 L43 I44 Q49 L50 E51 D52 G53 R54 T55 D58 Y59 N60 I61 Q62 K63 E64 S65

T66 L67 H68 L69 V70 L71 R72 L73 R74 G75 G76

4   Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	173.98Å 199.89Å 109.71Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 3.30 49.97 – 3.30	Depositor EDS
% Data completeness (in resolution range)	87.5 (50.00-3.30) 87.4 (49.97-3.30)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	2.92 (at 3.33Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.217 , 0.267 0.220 , 0.230	Depositor DCC
R_{free} test set	1288 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	74.2	Xtriage
Anisotropy	0.189	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 42.8	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	9814	wwPDB-VP
Average B, all atoms (Å^2)	80.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.85% 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 ⓘ

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	0.43	0/1206	0.67	0/1642
1	B	0.41	0/1206	0.63	0/1642
2	C	0.49	0/3232	0.66	0/4366
2	D	0.48	1/3207 (0.0%)	0.72	1/4333 (0.0%)
3	X	0.40	0/607	0.70	0/816
3	Y	0.38	0/607	0.58	0/816
All	All	0.46	1/10065 (0.0%)	0.67	1/13615 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	837	ASP	CB-CG	5.05	1.62	1.51

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	764	GLN	N-CA-C	-5.95	94.94	111.00

There are no chirality outliers.

There are no planarity outliers.

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	1170	0	1152	57	0
1	B	1170	0	1152	57	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	3148	0	3063	163	0
2	D	3124	0	3045	150	0
3	X	601	0	629	41	0
3	Y	601	0	629	40	0
All	All	9814	0	9670	480	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 25.

The worst 5 of 480 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)
2:D:597:ARG:HG3	2:D:597:ARG:HH11	1.25	1.00
2:C:817:ILE:O	2:C:820:ILE:HG22	1.61	1.00
2:C:675:ILE:HD12	2:C:675:ILE:H	1.26	0.97
2:C:848:ILE:HG13	2:C:906:THR:HG23	1.49	0.95
2:C:944:ALA:O	2:C:948:ALA:HB2	1.69	0.92

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	144/146 (99%)	133 (92%)	10 (7%)	1 (1%)	22	54
1	B	144/146 (99%)	124 (86%)	17 (12%)	3 (2%)	7	31
2	C	375/385 (97%)	321 (86%)	50 (13%)	4 (1%)	14	45
2	D	372/385 (97%)	328 (88%)	42 (11%)	2 (0%)	29	61
3	X	74/81 (91%)	64 (86%)	8 (11%)	2 (3%)	5	26
3	Y	74/81 (91%)	63 (85%)	9 (12%)	2 (3%)	5	26

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	1183/1224 (97%)	1033 (87%)	136 (12%)	14 (1%)	13 42

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	3	SER
2	D	825	GLU
1	B	129	THR
2	C	593	ASP
3	Y	2	GLN

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	130/130 (100%)	123 (95%)	7 (5%)	22 53
1	B	130/130 (100%)	118 (91%)	12 (9%)	9 31
2	C	341/347 (98%)	315 (92%)	26 (8%)	13 39
2	D	339/347 (98%)	306 (90%)	33 (10%)	8 29
3	X	68/70 (97%)	64 (94%)	4 (6%)	19 49
3	Y	68/70 (97%)	64 (94%)	4 (6%)	19 49
All	All	1076/1094 (98%)	990 (92%)	86 (8%)	12 37

5 of 86 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	670	ASN
2	D	835	LEU
2	D	691	THR
2	D	771	LYS
2	D	896	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 35 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	805	ASN
2	D	900	ASN
3	Y	2	GLN
2	C	781	ASN
2	C	764	GLN

5.3.3   RNA ⓘ

There are no RNA molecules in this entry.

5.4   Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5   Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6   Ligand geometry ⓘ

There are no ligands in this entry.

5.7   Other polymers ⓘ

There are no such residues in this entry.

5.8   Polymer linkage issues ⓘ

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, 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	146/146 (100%)	0.08	0	100	100	35, 80, 104, 114	0
1	B	146/146 (100%)	0.13	0	100	100	52, 77, 99, 114	0
2	C	377/385 (97%)	0.01	3 (0%)	86	86	45, 72, 112, 122	0
2	D	374/385 (97%)	-0.05	0	100	100	38, 67, 96, 126	0
3	X	76/81 (93%)	1.19	23 (30%)	0	0	93, 126, 127, 127	0
3	Y	76/81 (93%)	1.38	20 (26%)	0	0	95, 125, 127, 127	0
All	All	1195/1224 (97%)	0.18	46 (3%)	40	37	35, 75, 126, 127	0

The worst 5 of 46 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	X	1	MET	5.1
3	Y	15	LEU	4.1
3	Y	66	THR	4.0
3	Y	43	LEU	4.0
3	X	2	GLN	3.5

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.