



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 24, 2020 – 01:16 pm BST

PDB ID : 6ELU  
Title : Structure of Serum Resistance Associated protein from T. b. rhodesiense  
Authors : Zoll, S.; Higgins, M.K.  
Deposited on : 2017-09-29  
Resolution : 2.30 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](mailto: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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

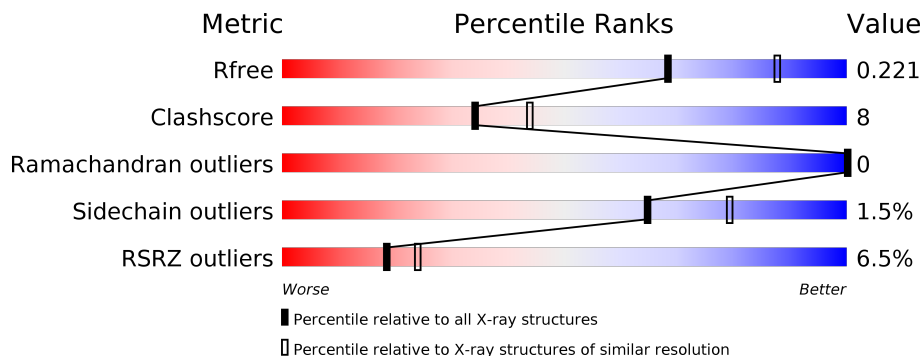
MolProbity : 4.02b-467  
Xtrriage (Phenix) : 1.13  
EDS : 2.11  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

# 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.





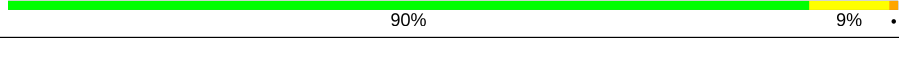
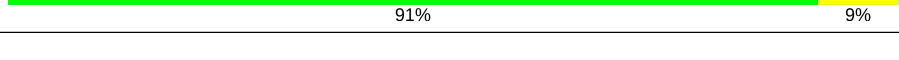
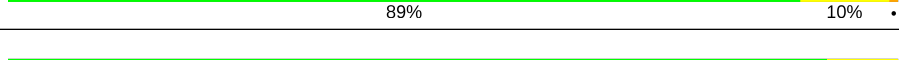
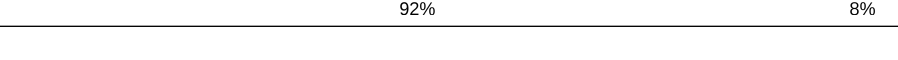
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	233	
1	D	233	
1	G	233	
1	J	233	
2	B	229	
2	E	229	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	H	229	 <p>% 83% 14%</p>
2	K	229	 <p>% 82% 14%</p>
3	C	217	 <p>90% 9%</p>
3	F	217	 <p>91% 9%</p>
3	I	217	 <p>89% 10%</p>
3	L	217	 <p>92% 8%</p>

## 2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 19153 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 Serum resistance associated; VSG protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	193	1413	883	247	280	3	0	0	0
1	D	161	1188	745	208	232	3	0	0	0
1	G	168	1236	775	216	242	3	0	0	0
1	J	172	1264	793	221	247	3	0	0	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	29	GLY	-	expression tag	UNP Q8T309
A	30	SER	-	expression tag	UNP Q8T309
A	31	HIS	-	expression tag	UNP Q8T309
A	32	MET	-	expression tag	UNP Q8T309
D	29	GLY	-	expression tag	UNP Q8T309
D	30	SER	-	expression tag	UNP Q8T309
D	31	HIS	-	expression tag	UNP Q8T309
D	32	MET	-	expression tag	UNP Q8T309
G	29	GLY	-	expression tag	UNP Q8T309
G	30	SER	-	expression tag	UNP Q8T309
G	31	HIS	-	expression tag	UNP Q8T309
G	32	MET	-	expression tag	UNP Q8T309
J	29	GLY	-	expression tag	UNP Q8T309
J	30	SER	-	expression tag	UNP Q8T309
J	31	HIS	-	expression tag	UNP Q8T309
J	32	MET	-	expression tag	UNP Q8T309

- Molecule 2 is a protein called G10\_3 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	222	Total	C	N	O	S	0	0	0
			1686	1064	280	334	8			
2	E	222	Total	C	N	O	S	0	0	0
			1686	1064	280	334	8			
2	H	222	Total	C	N	O	S	0	0	0
			1686	1064	280	334	8			
2	K	222	Total	C	N	O	S	0	0	0
			1686	1064	280	334	8			

- Molecule 3 is a protein called G10\_3 Light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	217	Total	C	N	O	S	0	0	0
			1682	1046	282	347	7			
3	F	217	Total	C	N	O	S	0	0	0
			1682	1046	282	347	7			
3	I	217	Total	C	N	O	S	0	0	0
			1682	1046	282	347	7			
3	L	217	Total	C	N	O	S	0	0	0
			1682	1046	282	347	7			

- Molecule 4 is water.

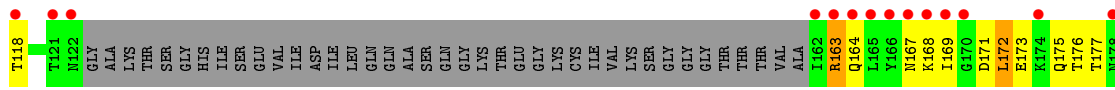
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	O	0	0
			1	1		
4	B	49	Total	O	0	0
			49	49		
4	C	83	Total	O	0	0
			83	83		
4	E	50	Total	O	0	0
			50	50		
4	F	77	Total	O	0	0
			77	77		
4	G	4	Total	O	0	0
			4	4		
4	H	86	Total	O	0	0
			86	86		
4	I	91	Total	O	0	0
			91	91		
4	J	3	Total	O	0	0
			3	3		
4	K	74	Total	O	0	0
			74	74		

*Continued on next page...*

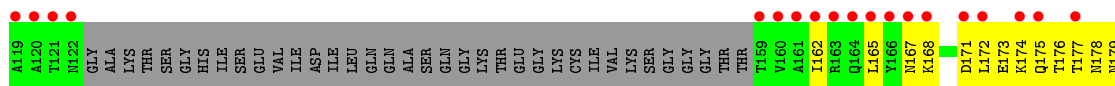
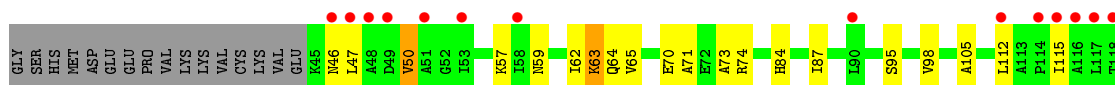
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
4	L	62	Total	O	0	0
			62	62		

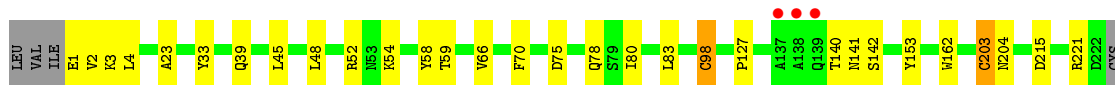
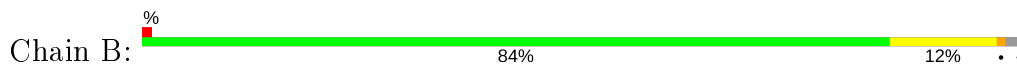




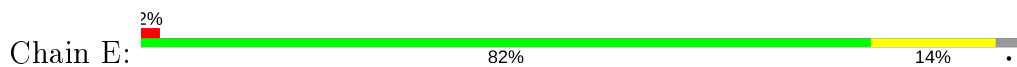
• Molecule 1: Serum resistance associated; VSG protein



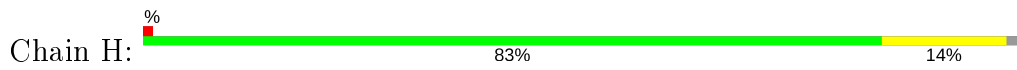
• Molecule 2: G10\_3 heavy chain



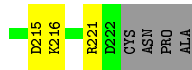
• Molecule 2: G10\_3 heavy chain



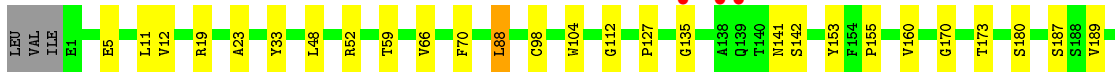
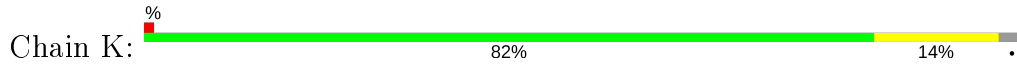
• Molecule 2: G10\_3 heavy chain



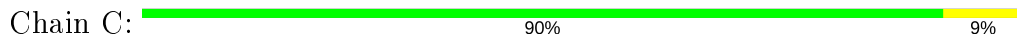




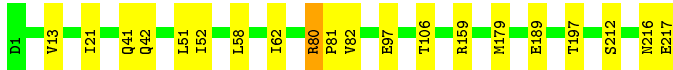
- Molecule 2: G10\_3 heavy chain



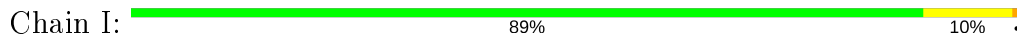
- Molecule 3: G10\_3 Light chain



- Molecule 3: G10\_3 Light chain



- Molecule 3: G10\_3 Light chain



- Molecule 3: G10\_3 Light chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	127.60Å 103.55Å 150.64Å 90.00° 114.58° 90.00°	Depositor
Resolution (Å)	48.43 – 2.30 48.43 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.9 (48.43-2.30) 99.9 (48.43-2.30)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.74 (at 2.29Å)	Xtrriage
Refinement program	PHENIX (1.11.1_2575: ???)	Depositor
R, $R_{free}$	0.191 , 0.220 0.191 , 0.221	Depositor DCC
$R_{free}$ test set	7914 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	46.9	Xtrriage
Anisotropy	0.209	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 55.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtrriage
Estimated twinning fraction	0.026 for h,-k,-h-l	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	19153	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	71.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.11% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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

### 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	0.36	0/1422	0.66	2/1932 (0.1%)
1	D	0.33	0/1196	0.58	2/1625 (0.1%)
1	G	0.35	0/1244	0.64	2/1692 (0.1%)
1	J	0.41	1/1272 (0.1%)	0.64	1/1730 (0.1%)
2	B	0.40	0/1731	0.65	1/2364 (0.0%)
2	E	0.36	1/1731 (0.1%)	0.60	1/2364 (0.0%)
2	H	0.30	0/1731	0.58	0/2364
2	K	0.33	0/1731	0.61	0/2364
3	C	0.32	0/1722	0.62	2/2342 (0.1%)
3	F	0.41	2/1722 (0.1%)	0.64	4/2342 (0.2%)
3	I	0.32	0/1722	0.56	1/2342 (0.0%)
3	L	0.29	0/1722	0.53	0/2342
All	All	0.35	4/18946 (0.0%)	0.61	16/25803 (0.1%)

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
2	B	0	1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	F	80	ARG	CG-CD	-6.22	1.36	1.51
3	F	97	GLU	CG-CD	6.13	1.61	1.51
1	J	50	VAL	CB-CG1	5.63	1.64	1.52
2	E	91	GLU	CG-CD	5.37	1.59	1.51

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	80	ARG	CG-CD-NE	-8.52	93.91	111.80
1	A	77	LEU	CB-CG-CD2	-8.48	96.58	111.00
1	D	191	ILE	CG1-CB-CG2	-8.33	93.08	111.40
1	D	197	LEU	CB-CG-CD1	-8.10	97.23	111.00
3	F	80	ARG	CA-CB-CG	-8.02	95.75	113.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	140	THR	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	1413	0	1493	50	0
1	D	1188	0	1263	49	0
1	G	1236	0	1313	38	0
1	J	1264	0	1347	41	0
2	B	1686	0	1629	20	0
2	E	1686	0	1629	25	0
2	H	1686	0	1629	22	0
2	K	1686	0	1629	24	0
3	C	1682	0	1593	12	0
3	F	1682	0	1593	16	0
3	I	1682	0	1593	14	0
3	L	1682	0	1593	9	0
4	A	1	0	0	0	0
4	B	49	0	0	0	0
4	C	83	0	0	0	0
4	E	50	0	0	1	0
4	F	77	0	0	0	0
4	G	4	0	0	1	0
4	H	86	0	0	1	0
4	I	91	0	0	1	0
4	J	3	0	0	1	0
4	K	74	0	0	2	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	L	62	0	0	0	0
All	All	19153	0	18304	310	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 8.

The worst 5 of 310 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:161:ALA:O	1:A:164:GLN:CG	2.05	1.04
1:A:161:ALA:HA	1:A:164:GLN:HG2	1.42	1.00
3:L:193:HIS:O	3:L:215:ARG:NH2	1.94	1.00
1:A:161:ALA:O	1:A:164:GLN:HG2	1.64	0.97
1:J:192:LEU:HB3	1:J:243:ILE:HD12	1.49	0.94

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	189/233 (81%)	185 (98%)	4 (2%)	0	100	100
1	D	157/233 (67%)	153 (98%)	4 (2%)	0	100	100
1	G	164/233 (70%)	160 (98%)	4 (2%)	0	100	100
1	J	168/233 (72%)	163 (97%)	5 (3%)	0	100	100
2	B	220/229 (96%)	215 (98%)	5 (2%)	0	100	100
2	E	220/229 (96%)	215 (98%)	5 (2%)	0	100	100
2	H	220/229 (96%)	216 (98%)	4 (2%)	0	100	100
2	K	220/229 (96%)	215 (98%)	5 (2%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	C	215/217 (99%)	209 (97%)	6 (3%)	0	100	100
3	F	215/217 (99%)	208 (97%)	7 (3%)	0	100	100
3	I	215/217 (99%)	209 (97%)	6 (3%)	0	100	100
3	L	215/217 (99%)	209 (97%)	6 (3%)	0	100	100
All	All	2418/2716 (89%)	2357 (98%)	61 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	154/188 (82%)	154 (100%)	0	100	100
1	D	130/188 (69%)	123 (95%)	7 (5%)	22	30
1	G	135/188 (72%)	132 (98%)	3 (2%)	52	69
1	J	138/188 (73%)	135 (98%)	3 (2%)	52	69
2	B	188/194 (97%)	185 (98%)	3 (2%)	62	78
2	E	188/194 (97%)	185 (98%)	3 (2%)	62	78
2	H	188/194 (97%)	184 (98%)	4 (2%)	53	70
2	K	188/194 (97%)	186 (99%)	2 (1%)	73	86
3	C	190/190 (100%)	187 (98%)	3 (2%)	62	78
3	F	190/190 (100%)	189 (100%)	1 (0%)	88	95
3	I	190/190 (100%)	188 (99%)	2 (1%)	73	86
3	L	190/190 (100%)	189 (100%)	1 (0%)	88	95
All	All	2069/2288 (90%)	2037 (98%)	32 (2%)	65	79

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	E	98	CYS

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
1	G	163	ARG
2	K	88	LEU
3	F	179	MET
1	G	173	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
2	E	39	GLN
3	L	216	ASN
2	E	139	GLN
2	E	13	GLN
3	F	149	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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	193/233 (82%)	0.83	33 (17%) 1 2	65, 98, 165, 185	0
1	D	161/233 (69%)	1.19	43 (26%) 0 0	74, 113, 188, 213	0
1	G	168/233 (72%)	0.76	32 (19%) 1 1	53, 97, 177, 196	0
1	J	172/233 (73%)	0.95	38 (22%) 0 1	49, 94, 183, 210	0
2	B	222/229 (96%)	-0.11	3 (1%) 75 80	32, 58, 101, 148	0
2	E	222/229 (96%)	-0.11	4 (1%) 68 74	30, 54, 88, 141	1 (0%)
2	H	222/229 (96%)	-0.23	3 (1%) 75 80	32, 51, 90, 170	1 (0%)
2	K	222/229 (96%)	-0.23	3 (1%) 75 80	35, 52, 88, 165	0
3	C	217/217 (100%)	-0.26	0 100 100	32, 52, 84, 122	0
3	F	217/217 (100%)	-0.28	0 100 100	31, 51, 86, 114	0
3	I	217/217 (100%)	-0.30	0 100 100	30, 47, 79, 145	0
3	L	217/217 (100%)	-0.19	0 100 100	36, 56, 103, 152	0
All	All	2450/2716 (90%)	0.11	159 (6%) 18 24	30, 60, 156, 213	2 (0%)

The worst 5 of 159 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	114	PRO	10.9
1	D	48	ALA	9.1
1	J	209	LYS	8.0
1	G	166	TYR	7.6
1	D	210	GLY	7.1

### 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 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.