



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 15, 2023 – 06:33 PM EDT

PDB ID : 8DGW  
Title : Crystal structure of HCoV-HKU1 spike stem helix peptide in complex with Fab of broadly neutralizing antibody CC95.108 isolated from a vaccinated COVID-19 convalescent  
Authors : Liu, H.; Wilson, I.A.  
Deposited on : 2022-06-24  
Resolution : 2.81 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 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

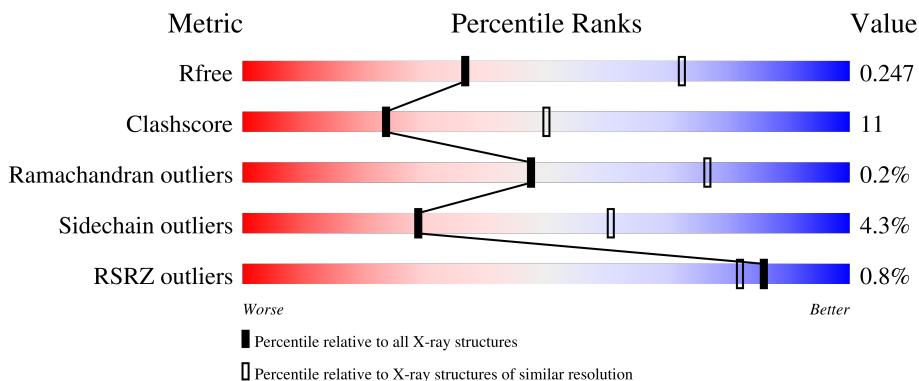
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

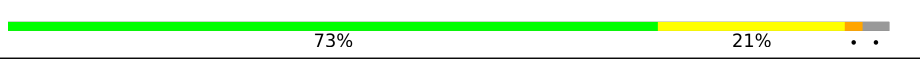
The reported resolution of this entry is 2.81 Å.

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	3617 (2.84-2.80)
Clashscore	141614	4060 (2.84-2.80)
Ramachandran outliers	138981	3978 (2.84-2.80)
Sidechain outliers	138945	3980 (2.84-2.80)
RSRZ outliers	127900	3552 (2.84-2.80)

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  $\geq 3$ , 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	220	
1	C	220	
1	E	220	
1	H	220	
2	B	216	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	D	216	 74% 24% .
2	F	216	 2% 67% 30% .
2	L	216	 75% 22% .
3	G	25	 4% 56% 24% 20%
3	I	25	 48% 32% 20%
3	J	25	 4% 64% 16% 20%
3	K	25	 4% 64% 12% . 20%

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 13052 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 Antibody CC95.108 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	H	213	Total 1587	C 996	N 275	O 309	S 7	0	0	0
1	A	218	Total 1625	C 1016	N 283	O 319	S 7	0	0	0
1	C	213	Total 1582	C 991	N 274	O 310	S 7	0	0	0
1	E	210	Total 1568	C 984	N 274	O 303	S 7	0	0	0

- Molecule 2 is a protein called Antibody CC95.108 Fab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	L	211	Total 1538	C 968	N 257	O 309	S 4	0	0	0
2	B	207	Total 1482	C 934	N 244	O 300	S 4	0	0	0
2	D	210	Total 1507	C 949	N 248	O 306	S 4	0	0	0
2	F	210	Total 1500	C 949	N 247	O 300	S 4	0	0	0

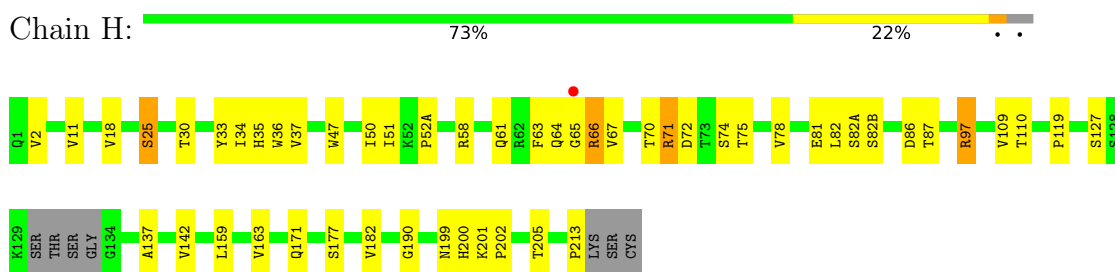
- Molecule 3 is a protein called Spike protein S2'.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	G	20	Total 162	C 105	N 26	O 31	0	0	0
3	I	20	Total 167	C 108	N 28	O 31	0	0	0
3	J	20	Total 167	C 108	N 28	O 31	0	0	0
3	K	20	Total 167	C 107	N 28	O 32	0	0	0

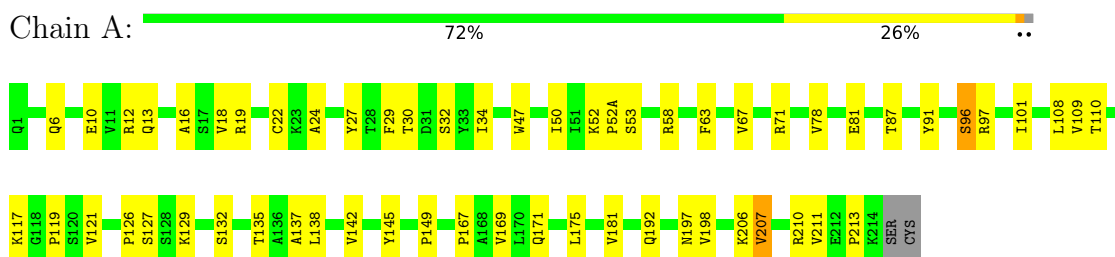
### 3 Residue-property plots

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.

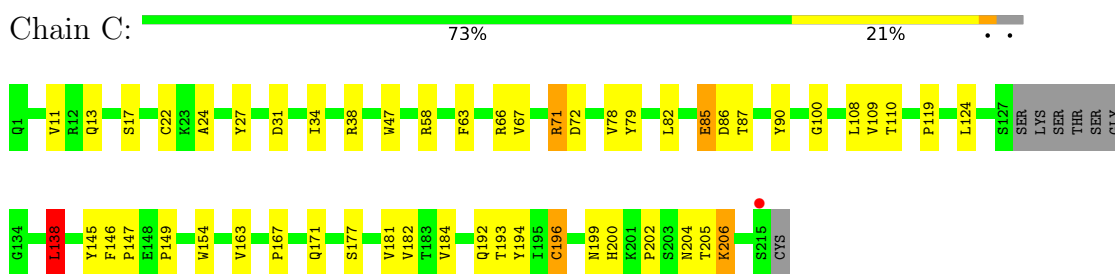
- Molecule 1: Antibody CC95.108 Fab heavy chain



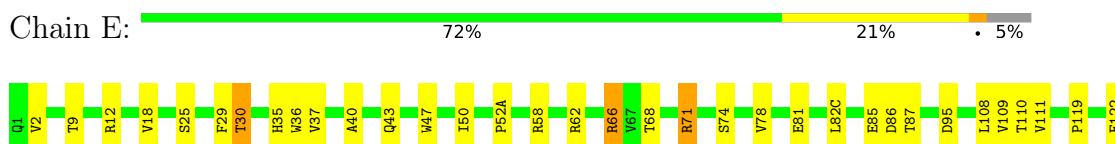
- Molecule 1: Antibody CC95.108 Fab heavy chain



- Molecule 1: Antibody CC95.108 Fab heavy chain



- Molecule 1: Antibody CC95.108 Fab heavy chain

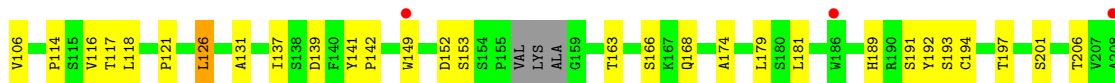
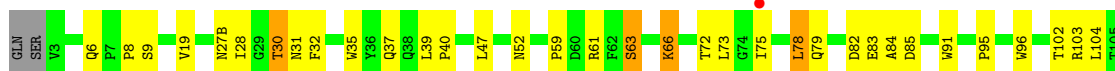




- Molecule 2: Antibody CC95.108 Fab light chain



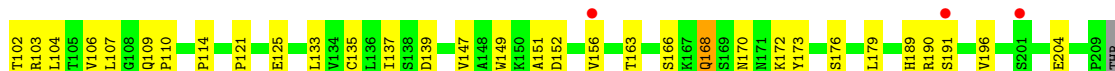
- Molecule 2: Antibody CC95.108 Fab light chain



- Molecule 2: Antibody CC95.108 Fab light chain



- Molecule 2: Antibody CC95.108 Fab light chain





## 4 Data and refinement statistics i

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	157.42Å 157.42Å 205.09Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	39.36 – 2.81 39.36 – 2.81	Depositor EDS
% Data completeness (in resolution range)	95.0 (39.36-2.81) 95.0 (39.36-2.81)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.83 (at 2.81Å)	Xtrriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, $R_{free}$	0.217 , 0.248 0.215 , 0.247	Depositor DCC
$R_{free}$ test set	2257 reflections (5.12%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	49.6	Xtrriage
Anisotropy	0.038	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 33.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.025 for h,-h-k,-l	Xtrriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	13052	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	51.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.86% 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 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	0.38	0/1663	0.64	1/2267 (0.0%)
1	C	0.31	0/1619	0.67	2/2208 (0.1%)
1	E	0.29	0/1605	0.61	0/2188
1	H	0.29	0/1624	0.58	0/2215
2	B	0.34	0/1522	0.64	3/2093 (0.1%)
2	D	0.29	0/1548	0.56	0/2131
2	F	0.32	1/1541 (0.1%)	0.54	0/2122
2	L	0.30	0/1579	0.52	0/2167
3	G	0.35	0/167	0.52	0/225
3	I	0.26	0/173	0.49	0/233
3	J	0.26	0/173	0.42	0/233
3	K	0.34	0/173	0.65	0/234
All	All	0.32	1/13387 (0.0%)	0.59	6/18316 (0.0%)

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	E	0	1
1	H	0	1
All	All	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	65	SER	CA-CB	5.19	1.60	1.52

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	138	LEU	CB-CG-CD1	-12.50	89.75	111.00

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	66	LYS	CB-CG-CD	-7.65	91.70	111.60
1	A	210	ARG	NE-CZ-NH1	7.62	124.11	120.30
1	C	85	GLU	CA-CB-CG	6.51	127.73	113.40
2	B	153	SER	C-N-CA	5.06	134.35	121.70

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	62	ARG	Sidechain
1	H	66	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	1625	0	1589	38	0
1	C	1582	0	1530	31	0
1	E	1568	0	1524	38	0
1	H	1587	0	1539	40	0
2	B	1482	0	1378	45	0
2	D	1507	0	1408	37	0
2	F	1500	0	1397	46	0
2	L	1538	0	1466	28	0
3	G	162	0	144	3	0
3	I	167	0	149	4	0
3	J	167	0	149	2	0
3	K	167	0	144	5	0
All	All	13052	0	12417	284	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 11.

The worst 5 of 284 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:H:66:ARG:NH1	1:H:86:ASP:OD2	2.04	0.90
2:B:78:LEU:HD11	2:B:104:LEU:HD21	1.58	0.85
1:A:19:ARG:NH1	1:C:79:TYR:OH	2.13	0.81
1:A:138:LEU:HB2	1:A:211:VAL:HG11	1.66	0.77
1:E:12:ARG:HG3	1:E:18:VAL:HG22	1.66	0.76

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	216/220 (98%)	208 (96%)	7 (3%)	1 (0%)	29	59
1	C	209/220 (95%)	201 (96%)	8 (4%)	0	100	100
1	E	206/220 (94%)	198 (96%)	7 (3%)	1 (0%)	29	59
1	H	209/220 (95%)	201 (96%)	7 (3%)	1 (0%)	29	59
2	B	203/216 (94%)	196 (97%)	7 (3%)	0	100	100
2	D	208/216 (96%)	201 (97%)	6 (3%)	1 (0%)	29	59
2	F	208/216 (96%)	201 (97%)	7 (3%)	0	100	100
2	L	209/216 (97%)	204 (98%)	5 (2%)	0	100	100
3	G	18/25 (72%)	18 (100%)	0	0	100	100
3	I	18/25 (72%)	18 (100%)	0	0	100	100
3	J	18/25 (72%)	18 (100%)	0	0	100	100
3	K	18/25 (72%)	18 (100%)	0	0	100	100
All	All	1740/1844 (94%)	1682 (97%)	54 (3%)	4 (0%)	47	76

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	H	127	SER
1	E	190	GLY
2	D	110	PRO
1	A	127	SER

### 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	182/187 (97%)	175 (96%)	7 (4%)	33	65
1	C	175/187 (94%)	166 (95%)	9 (5%)	24	54
1	E	173/187 (92%)	167 (96%)	6 (4%)	36	68
1	H	175/187 (94%)	167 (95%)	8 (5%)	27	58
2	B	157/181 (87%)	150 (96%)	7 (4%)	27	59
2	D	161/181 (89%)	158 (98%)	3 (2%)	57	84
2	F	156/181 (86%)	152 (97%)	4 (3%)	46	78
2	L	167/181 (92%)	158 (95%)	9 (5%)	22	51
3	G	18/24 (75%)	16 (89%)	2 (11%)	6	18
3	I	19/24 (79%)	16 (84%)	3 (16%)	2	7
3	J	19/24 (79%)	18 (95%)	1 (5%)	22	52
3	K	19/24 (79%)	17 (90%)	2 (10%)	7	20
All	All	1421/1568 (91%)	1360 (96%)	61 (4%)	29	60

5 of 61 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	78	LEU
3	I	1226	HIS
1	C	85	GLU
3	G	1236	SER
3	K	1226	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28

such sidechains are listed below:

Mol	Chain	Res	Type
1	C	199	ASN
3	K	1244	ASN
2	D	171	ASN
2	F	185	GLN
2	D	168	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, 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	218/220 (99%)	-0.15	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	29, 45, 84, 111	0
1	C	213/220 (96%)	-0.06	1 (0%) <span style="border: 1px solid blue; padding: 2px;">91</span> <span style="border: 1px solid blue; padding: 2px;">88</span>	25, 42, 82, 114	0
1	E	210/220 (95%)	-0.18	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	28, 40, 82, 120	0
1	H	213/220 (96%)	-0.29	1 (0%) <span style="border: 1px solid blue; padding: 2px;">91</span> <span style="border: 1px solid blue; padding: 2px;">88</span>	26, 38, 66, 109	0
2	B	207/216 (95%)	0.13	4 (1%) <span style="border: 1px solid gray; padding: 2px;">66</span> <span style="border: 1px solid gray; padding: 2px;">59</span>	44, 68, 92, 123	0
2	D	210/216 (97%)	-0.06	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	28, 57, 80, 119	0
2	F	210/216 (97%)	0.09	5 (2%) <span style="border: 1px solid gray; padding: 2px;">59</span> <span style="border: 1px solid gray; padding: 2px;">49</span>	37, 65, 90, 111	0
2	L	211/216 (97%)	-0.24	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	28, 42, 61, 84	0
3	G	20/25 (80%)	0.06	1 (5%) <span style="border: 1px solid red; padding: 2px;">28</span> <span style="border: 1px solid red; padding: 2px;">20</span>	37, 51, 97, 120	0
3	I	20/25 (80%)	0.31	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	48, 61, 82, 88	0
3	J	20/25 (80%)	-0.09	1 (5%) <span style="border: 1px solid red; padding: 2px;">28</span> <span style="border: 1px solid red; padding: 2px;">20</span>	33, 42, 80, 110	0
3	K	20/25 (80%)	-0.06	1 (5%) <span style="border: 1px solid red; padding: 2px;">28</span> <span style="border: 1px solid red; padding: 2px;">20</span>	39, 49, 79, 111	0
All	All	1772/1844 (96%)	-0.09	14 (0%) <span style="border: 1px solid blue; padding: 2px;">86</span> <span style="border: 1px solid blue; padding: 2px;">82</span>	25, 48, 84, 123	0

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	215	SER	4.1
3	G	1226	HIS	3.2
1	H	65	GLY	3.1
3	K	1226	HIS	2.8
2	F	201	SER	2.7

### 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.