



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 1, 2021 – 12:11 PM EST

PDB ID : 7KJ0  
Title : hyperoxidized human peroxiredoxin 2  
Authors : Kean, K.M.; Karplus, P.A.  
Deposited on : 2020-10-25  
Resolution : 2.29 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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.17.1  
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) : 2.17.1

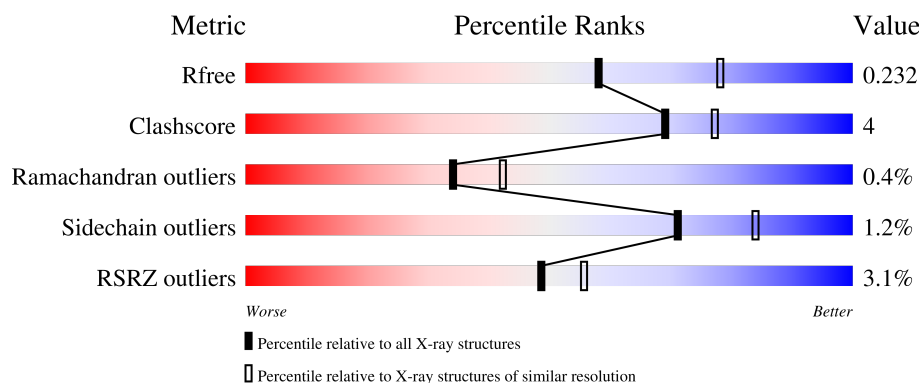
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.29 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 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	197	<div> <div>2%</div> <div>91%</div> <div>9%</div> </div>
1	B	197	<div> <div>4%</div> <div>92%</div> <div>8%</div> </div>
1	C	197	<div> <div>2%</div> <div>89%</div> <div>11%</div> </div>
1	D	197	<div> <div>3%</div> <div>91%</div> <div>8%</div> </div>
1	E	197	<div> <div>0%</div> <div>90%</div> <div>9%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	197	<div><div>%</div><div><div></div><div>92%</div><div>8%</div></div></div>
1	G	197	<div><div>2%</div><div><div></div><div>94%</div><div>5%</div></div></div>
1	H	197	<div><div>5%</div><div><div></div><div>89%</div><div>10%</div></div></div>
1	I	197	<div><div>9%</div><div><div></div><div>92%</div><div>8%</div></div></div>
1	J	197	<div><div>3%</div><div><div></div><div>93%</div><div>7%</div></div></div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 33464 atoms, of which 15795 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 Peroxiredoxin-2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	197	Total	C	H	N	O	S	0	11	0
			3183	1020	1588	272	300	3			
1	B	197	Total	C	H	N	O	S	0	15	0
			3198	1026	1590	274	305	3			
1	C	197	Total	C	H	N	O	S	0	9	0
			3167	1015	1579	271	299	3			
1	D	197	Total	C	H	N	O	S	0	10	0
			3142	1010	1558	270	301	3			
1	E	197	Total	C	H	N	O	S	0	11	0
			3189	1021	1590	273	302	3			
1	F	197	Total	C	H	N	O	S	0	9	0
			3151	1011	1567	270	300	3			
1	G	197	Total	C	H	N	O	S	0	13	0
			3176	1017	1579	273	304	3			
1	H	197	Total	C	H	N	O	S	0	8	0
			3150	1011	1568	269	299	3			
1	I	197	Total	C	H	N	O	S	0	13	0
			3192	1021	1590	275	303	3			
1	J	197	Total	C	H	N	O	S	0	10	0
			3178	1017	1586	272	300	3			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	192	Total	O	0	9
			192	192		
2	B	196	Total	O	0	10
			196	196		
2	C	169	Total	O	0	10
			169	169		
2	D	167	Total	O	0	9
			167	167		

*Continued on next page...*

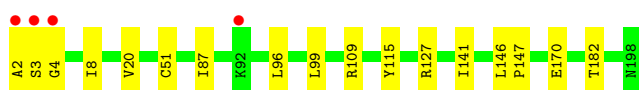
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	E	199	Total 199	O 199	0	8
2	F	209	Total 209	O 209	0	9
2	G	130	Total 130	O 130	0	8
2	H	144	Total 144	O 144	0	10
2	I	169	Total 169	O 169	0	8
2	J	163	Total 163	O 163	0	11

### 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.

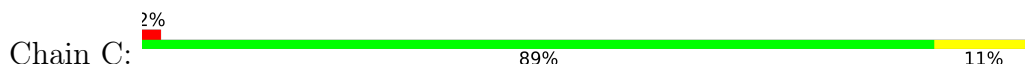
- Molecule 1: Peroxiredoxin-2



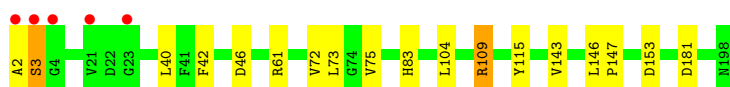
- Molecule 1: Peroxiredoxin-2



- Molecule 1: Peroxiredoxin-2



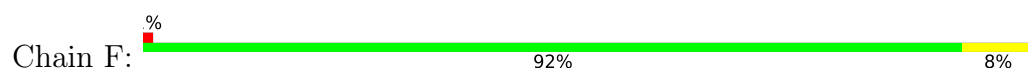
- Molecule 1: Peroxiredoxin-2



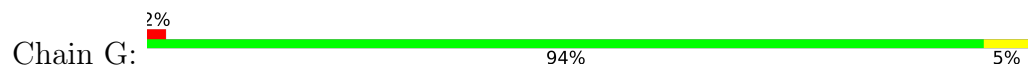
- Molecule 1: Peroxiredoxin-2



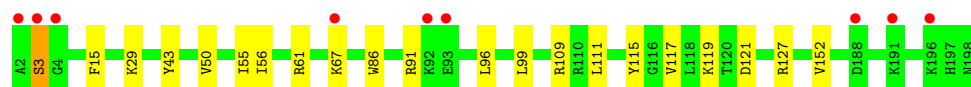
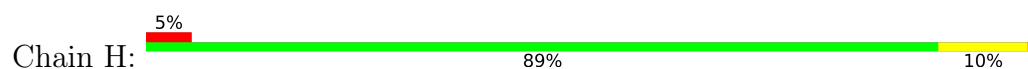
- Molecule 1: Peroxiredoxin-2



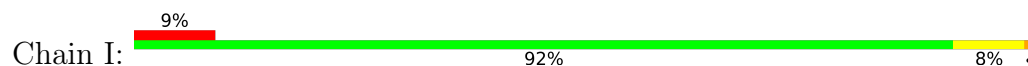
• Molecule 1: Peroxiredoxin-2



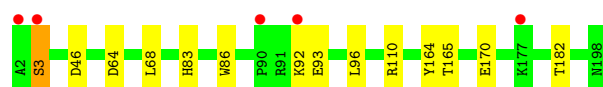
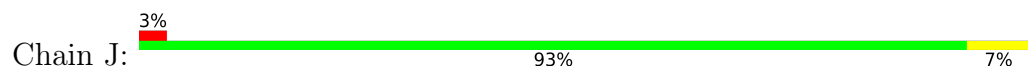
• Molecule 1: Peroxiredoxin-2



• Molecule 1: Peroxiredoxin-2



• Molecule 1: Peroxiredoxin-2



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	231.00Å 88.00Å 125.60Å 90.00° 100.20° 90.00°	Depositor
Resolution (Å)	49.60 – 2.29 49.60 – 2.29	Depositor EDS
% Data completeness (in resolution range)	97.9 (49.60-2.29) 97.9 (49.60-2.29)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.40 (at 2.29Å)	Xtriage
Refinement program	PHENIX 1.17.1 _3660	Depositor
R, $R_{free}$	0.182 , 0.233 0.182 , 0.232	Depositor DCC
$R_{free}$ test set	5449 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	45.9	Xtriage
Anisotropy	0.190	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 44.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	33464	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.12% 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

Bond lengths and bond angles in the following residue types are not validated in this section: CSD

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	0.25	0/1658	0.44	0/2243
1	B	0.25	0/1676	0.44	0/2270
1	C	0.25	0/1633	0.45	0/2212
1	D	0.25	0/1639	0.44	0/2222
1	E	0.26	0/1652	0.45	0/2237
1	F	0.25	0/1634	0.44	0/2213
1	G	0.25	0/1662	0.44	0/2250
1	H	0.25	0/1624	0.44	0/2200
1	I	0.25	0/1668	0.45	0/2259
1	J	0.25	0/1645	0.44	0/2226
All	All	0.25	0/16491	0.44	0/22332

There are no bond length outliers.

There are no bond angle outliers.

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	1595	1588	1557	14	0
1	B	1608	1590	1551	13	0
1	C	1588	1579	1572	15	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	1584	1558	1525	12	0
1	E	1599	1590	1577	16	0
1	F	1584	1567	1546	13	0
1	G	1597	1579	1528	7	0
1	H	1582	1568	1562	12	0
1	I	1602	1590	1552	11	0
1	J	1592	1586	1565	8	0
2	A	192	0	0	5	0
2	B	196	0	0	7	0
2	C	169	0	0	5	1
2	D	167	0	0	3	0
2	E	199	0	0	6	0
2	F	209	0	0	5	0
2	G	130	0	0	3	0
2	H	144	0	0	1	0
2	I	169	0	0	4	0
2	J	163	0	0	0	1
All	All	17669	15795	15535	118	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 4.

The worst 5 of 118 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:D:75[A]:VAL:HG12	1:D:104:LEU:HB2	1.59	0.84
1:G:2[A]:ALA:N	2:G:201[A]:HOH:O	2.12	0.81
1:C:2[A]:ALA:N	2:C:203[A]:HOH:O	2.16	0.79
1:I:2[A]:ALA:N	2:I:202[A]:HOH:O	2.15	0.78
1:I:100[B]:ASN:OD1	2:I:201:HOH:O	2.02	0.78

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:C:326:HOH:O	2:J:345:HOH:O[3_455]	2.10	0.10

## 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	204/197 (104%)	193 (95%)	9 (4%)	2 (1%)	15	17
1	B	208/197 (106%)	204 (98%)	4 (2%)	0	100	100
1	C	202/197 (102%)	192 (95%)	9 (4%)	1 (0%)	29	35
1	D	203/197 (103%)	197 (97%)	6 (3%)	0	100	100
1	E	204/197 (104%)	196 (96%)	5 (2%)	3 (2%)	10	10
1	F	202/197 (102%)	196 (97%)	5 (2%)	1 (0%)	29	35
1	G	206/197 (105%)	195 (95%)	8 (4%)	3 (2%)	10	10
1	H	201/197 (102%)	195 (97%)	6 (3%)	0	100	100
1	I	206/197 (105%)	195 (95%)	11 (5%)	0	100	100
1	J	203/197 (103%)	195 (96%)	8 (4%)	0	100	100
All	All	2039/1970 (104%)	1958 (96%)	71 (4%)	10 (0%)	34	35

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	4[A]	GLY
1	E	4[B]	GLY
1	E	44	PRO
1	G	4[A]	GLY
1	G	4[B]	GLY

### 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	172/164 (105%)	172 (100%)	0	100	100
1	B	173/164 (106%)	171 (99%)	2 (1%)	71	84
1	C	170/164 (104%)	170 (100%)	0	100	100
1	D	170/164 (104%)	165 (97%)	5 (3%)	42	58
1	E	172/164 (105%)	169 (98%)	3 (2%)	60	76
1	F	170/164 (104%)	168 (99%)	2 (1%)	71	84
1	G	173/164 (106%)	171 (99%)	2 (1%)	71	84
1	H	169/164 (103%)	163 (96%)	6 (4%)	35	49
1	I	174/164 (106%)	171 (98%)	3 (2%)	60	76
1	J	171/164 (104%)	165 (96%)	6 (4%)	36	50
All	All	1714/1640 (104%)	1685 (98%)	29 (2%)	71	76

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	3[A]	SER
1	J	110[A]	ARG
1	H	67	LYS
1	J	3[B]	SER
1	H	29	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

10 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
1	CSD	F	51	1	3,7,8	1.09	0	1,8,10	0.94	0
1	CSD	J	51	1	3,7,8	1.08	0	1,8,10	1.28	0
1	CSD	I	51	1	3,7,8	1.07	0	1,8,10	1.32	0
1	CSD	E	51	1	3,7,8	1.10	0	1,8,10	1.60	0
1	CSD	C	51	1	3,7,8	1.06	0	1,8,10	0.62	0
1	CSD	H	51	1	3,7,8	1.08	0	1,8,10	1.45	0
1	CSD	D	51	1	3,7,8	1.07	0	1,8,10	0.98	0
1	CSD	B	51	1	3,7,8	1.08	0	1,8,10	1.33	0
1	CSD	A	51	1	3,7,8	1.10	0	1,8,10	1.37	0
1	CSD	G	51	1	3,7,8	1.10	0	1,8,10	0.92	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	CSD	F	51	1	-	0/2/6/8	-
1	CSD	J	51	1	-	0/2/6/8	-
1	CSD	I	51	1	-	0/2/6/8	-
1	CSD	E	51	1	-	0/2/6/8	-
1	CSD	C	51	1	-	0/2/6/8	-
1	CSD	H	51	1	-	0/2/6/8	-
1	CSD	D	51	1	-	0/2/6/8	-
1	CSD	B	51	1	-	0/2/6/8	-
1	CSD	A	51	1	-	0/2/6/8	-
1	CSD	G	51	1	-	0/2/6/8	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	I	51	CSD	1	0
1	E	51	CSD	2	0
1	C	51	CSD	1	0
1	A	51	CSD	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	G	51	CSD	1	0

## 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	196/197 (99%)	-0.09	4 (2%) 65 71	35, 44, 58, 78	0
1	B	196/197 (99%)	-0.02	7 (3%) 42 49	32, 43, 61, 67	0
1	C	196/197 (99%)	-0.03	4 (2%) 65 71	37, 49, 67, 77	0
1	D	196/197 (99%)	0.13	5 (2%) 56 63	38, 49, 63, 71	0
1	E	196/197 (99%)	0.06	2 (1%) 82 86	34, 47, 64, 75	0
1	F	196/197 (99%)	0.10	2 (1%) 82 86	35, 44, 59, 65	0
1	G	196/197 (99%)	0.13	4 (2%) 65 71	42, 53, 68, 79	0
1	H	196/197 (99%)	0.13	9 (4%) 32 39	37, 46, 69, 79	0
1	I	196/197 (99%)	0.47	18 (9%) 9 12	36, 48, 73, 82	0
1	J	196/197 (99%)	0.18	5 (2%) 56 63	40, 54, 67, 76	0
All	All	1960/1970 (99%)	0.11	60 (3%) 49 56	32, 48, 66, 82	0

The worst 5 of 60 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	3[A]	SER	6.1
1	D	2[A]	ALA	6.0
1	D	3[A]	SER	5.9
1	I	2[A]	ALA	5.9
1	E	3[A]	SER	5.3

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q < 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
1	CSD	G	51	8/9	0.95	0.10	48,53,66,66	0
1	CSD	E	51	8/9	0.97	0.12	37,42,49,50	0
1	CSD	A	51	8/9	0.97	0.12	39,46,56,63	0
1	CSD	H	51	8/9	0.97	0.13	42,45,55,60	0
1	CSD	I	51	8/9	0.97	0.13	35,42,52,63	0
1	CSD	J	51	8/9	0.97	0.15	49,54,65,66	0
1	CSD	D	51	8/9	0.98	0.15	41,44,52,53	0
1	CSD	C	51	8/9	0.99	0.11	34,43,53,63	0
1	CSD	F	51	8/9	0.99	0.11	39,43,50,52	0
1	CSD	B	51	8/9	0.99	0.11	32,37,45,45	0

### 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.