



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 24, 2024 – 07:02 AM EDT

PDB ID : 7CSO  
Title : Structure of Ephexin4 DH-PH-SH3  
Authors : Zhang, M.; Lin, L.; Wang, C.; Zhu, J.  
Deposited on : 2020-08-15  
Resolution : 2.39 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  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 1.20.1  
EDS : 2.37.1  
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) : 2.37.1

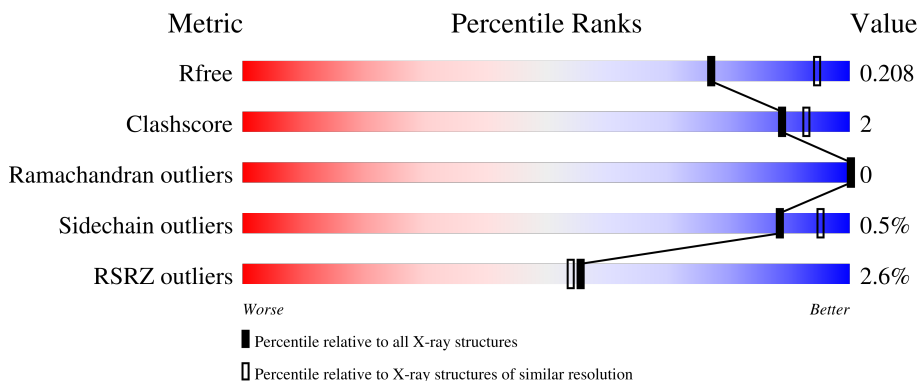
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.39 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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	459	 2% 90% 5% 6%
1	B	459	 3% 89% 5% 6%
1	C	459	 3% 90% 5% 5%
1	D	459	 2% 90% 5% 5%

## 2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 14658 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 Rho guanine nucleotide exchange factor 16.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	432	3468	2198	603	654	13	0	0	0
1	B	433	3501	2218	611	659	13	0	0	0
1	C	437	3501	2216	609	663	13	0	0	0
1	D	434	3510	2222	615	660	13	0	0	0

There are 24 discrepancies between the modelled and reference sequences:

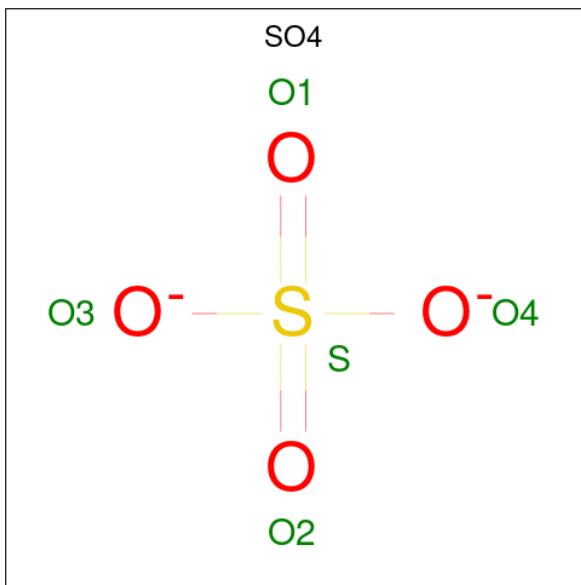
Chain	Residue	Modelled	Actual	Comment	Reference
A	255	GLY	-	expression tag	UNP Q3U5C8
A	256	PRO	-	expression tag	UNP Q3U5C8
A	257	GLY	-	expression tag	UNP Q3U5C8
A	258	SER	-	expression tag	UNP Q3U5C8
A	259	GLY	-	expression tag	UNP Q3U5C8
A	260	PRO	-	expression tag	UNP Q3U5C8
B	255	GLY	-	expression tag	UNP Q3U5C8
B	256	PRO	-	expression tag	UNP Q3U5C8
B	257	GLY	-	expression tag	UNP Q3U5C8
B	258	SER	-	expression tag	UNP Q3U5C8
B	259	GLY	-	expression tag	UNP Q3U5C8
B	260	PRO	-	expression tag	UNP Q3U5C8
C	255	GLY	-	expression tag	UNP Q3U5C8
C	256	PRO	-	expression tag	UNP Q3U5C8
C	257	GLY	-	expression tag	UNP Q3U5C8
C	258	SER	-	expression tag	UNP Q3U5C8
C	259	GLY	-	expression tag	UNP Q3U5C8
C	260	PRO	-	expression tag	UNP Q3U5C8
D	255	GLY	-	expression tag	UNP Q3U5C8
D	256	PRO	-	expression tag	UNP Q3U5C8
D	257	GLY	-	expression tag	UNP Q3U5C8

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	258	SER	-	expression tag	UNP Q3U5C8
D	259	GLY	-	expression tag	UNP Q3U5C8
D	260	PRO	-	expression tag	UNP Q3U5C8

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	C	1	Total O S 5 4 1	0	0
2	C	1	Total O S 5 4 1	0	0
2	C	1	Total O S 5 4 1	0	0
2	D	1	Total O S 5 4 1	0	0

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	D	1	Total	O	S	0	0
			5	4	1		
2	D	1	Total	O	S	0	0
			5	4	1		
2	D	1	Total	O	S	0	0
			5	4	1		

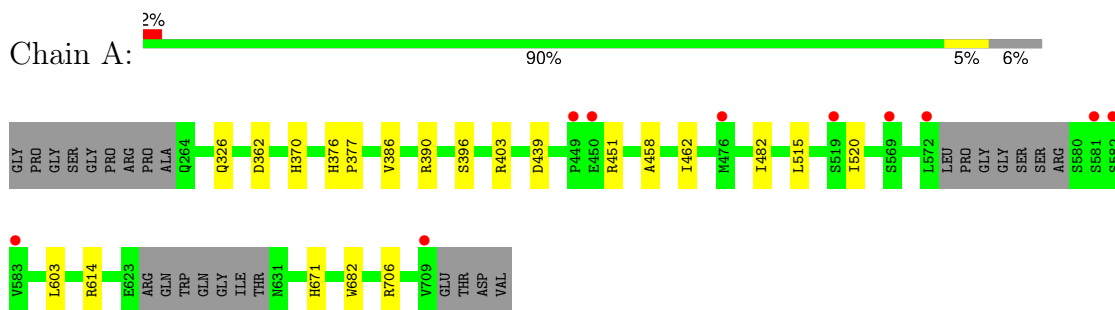
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	152	Total	O	0	0
			152	152		
3	B	132	Total	O	0	0
			132	132		
3	C	139	Total	O	0	0
			139	139		
3	D	190	Total	O	0	0
			190	190		

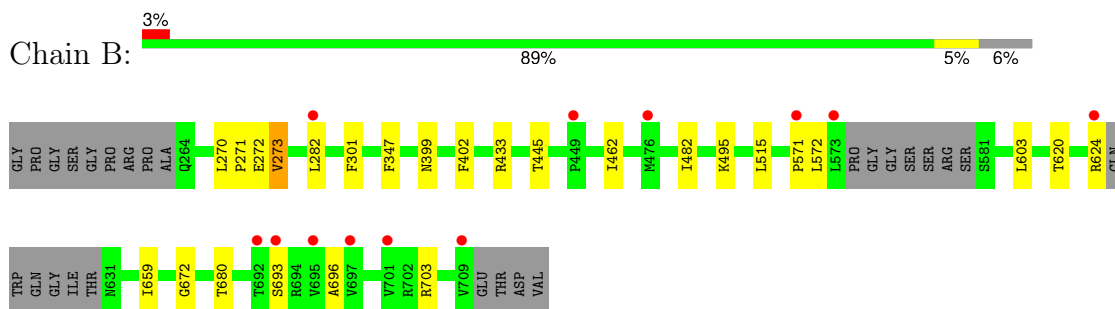
### 3 Residue-property plots [i](#)

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.

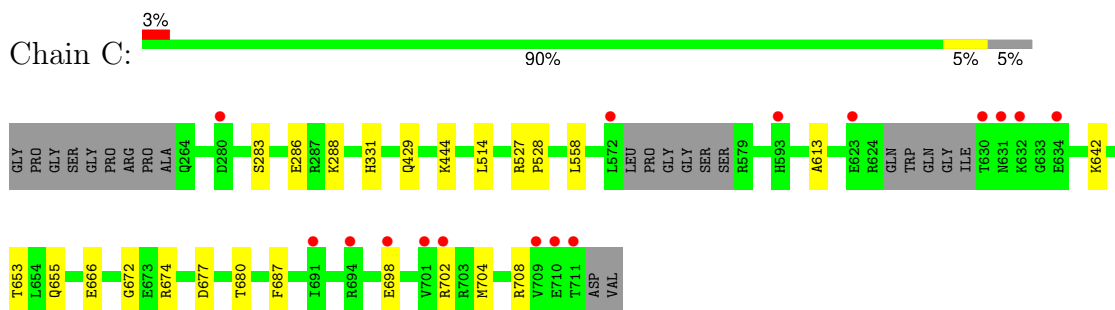
- Molecule 1: Rho guanine nucleotide exchange factor 16



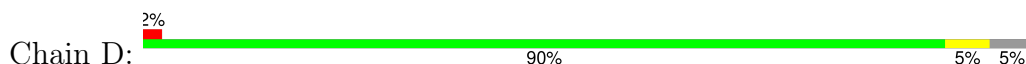
- Molecule 1: Rho guanine nucleotide exchange factor 16

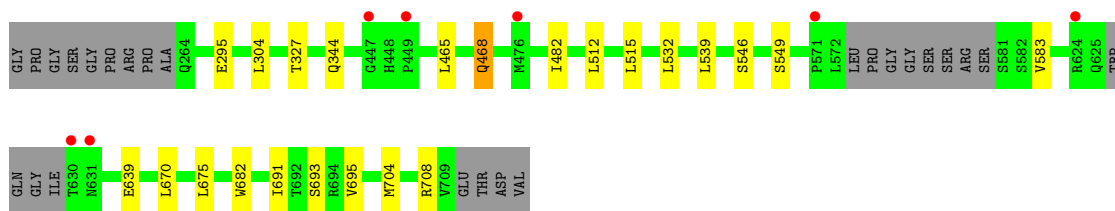


- Molecule 1: Rho guanine nucleotide exchange factor 16



- Molecule 1: Rho guanine nucleotide exchange factor 16





## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 31 1 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	144.44Å 144.44Å 290.52Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.42 – 2.39 48.42 – 2.39	Depositor EDS
% Data completeness (in resolution range)	99.8 (48.42-2.39) 99.8 (48.42-2.39)	Depositor EDS
$R_{merge}$	0.18	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.00 (at 2.39Å)	Xtrriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, $R_{free}$	0.176 , 0.204 0.184 , 0.208	Depositor DCC
$R_{free}$ test set	6640 reflections (4.87%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	43.5	Xtrriage
Anisotropy	0.165	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 42.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.025 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	14658	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	53.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.09% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: SO4

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	0.46	0/3535	0.65	0/4781
1	B	0.45	0/3567	0.64	0/4819
1	C	0.43	0/3567	0.59	0/4825
1	D	0.46	0/3577	0.65	1/4833 (0.0%)
All	All	0.45	0/14246	0.63	1/19258 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	468	GLN	CB-CA-C	5.62	121.65	110.40

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	3468	0	3442	11	0
1	B	3501	0	3501	16	0
1	C	3501	0	3470	14	0
1	D	3510	0	3506	12	0
2	A	15	0	0	0	0
2	B	15	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	15	0	0	0	0
2	D	20	0	0	0	0
3	A	152	0	0	2	0
3	B	132	0	0	1	0
3	C	139	0	0	1	0
3	D	190	0	0	0	0
All	All	14658	0	13919	52	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 2.

The worst 5 of 52 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:675:LEU:HA	1:D:704:MET:HE2	1.76	0.67
1:B:515:LEU:HB2	1:B:603:LEU:HB3	1.82	0.60
1:C:653:THR:O	1:C:674:ARG:NH2	2.36	0.57
1:C:331:HIS:CE1	1:D:327:THR:HA	2.41	0.56
1:C:674:ARG:HD2	1:C:677:ASP:OD1	2.06	0.55

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	426/459 (93%)	417 (98%)	9 (2%)	0	100	100
1	B	427/459 (93%)	419 (98%)	8 (2%)	0	100	100
1	C	431/459 (94%)	423 (98%)	8 (2%)	0	100	100
1	D	428/459 (93%)	422 (99%)	6 (1%)	0	100	100
All	All	1712/1836 (93%)	1681 (98%)	31 (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	383/411 (93%)	381 (100%)	2 (0%)	88	95
1	B	389/411 (95%)	387 (100%)	2 (0%)	88	95
1	C	386/411 (94%)	386 (100%)	0	100	100
1	D	390/411 (95%)	387 (99%)	3 (1%)	81	91
All	All	1548/1644 (94%)	1541 (100%)	7 (0%)	88	95

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	273	VAL
1	D	295	GLU
1	D	583	VAL
1	D	515	LEU
1	B	272	GLU

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

### 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](#)

13 ligands 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	$\# Z  > 2$
2	SO4	A	801	-	4,4,4	0.30	0	6,6,6	0.28	0
2	SO4	B	802	-	4,4,4	0.34	0	6,6,6	0.44	0
2	SO4	C	803	-	4,4,4	0.51	0	6,6,6	0.08	0
2	SO4	D	802	-	4,4,4	0.52	0	6,6,6	0.10	0
2	SO4	D	803	-	4,4,4	0.26	0	6,6,6	0.35	0
2	SO4	D	804	-	4,4,4	0.23	0	6,6,6	0.31	0
2	SO4	A	802	-	4,4,4	0.31	0	6,6,6	0.24	0
2	SO4	C	801	-	4,4,4	0.25	0	6,6,6	0.20	0
2	SO4	A	803	-	4,4,4	0.51	0	6,6,6	0.08	0
2	SO4	D	801	-	4,4,4	0.51	0	6,6,6	0.09	0
2	SO4	C	802	-	4,4,4	0.29	0	6,6,6	0.52	0
2	SO4	B	803	-	4,4,4	0.49	0	6,6,6	0.08	0
2	SO4	B	801	-	4,4,4	0.29	0	6,6,6	0.18	0

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.

No monomer is involved in short contacts.

## 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	432/459 (94%)	0.29	10 (2%) 60 58	32, 50, 83, 107	0
1	B	433/459 (94%)	0.35	12 (2%) 53 51	33, 52, 88, 122	0
1	C	437/459 (95%)	0.32	16 (3%) 41 41	33, 55, 90, 126	0
1	D	434/459 (94%)	0.24	7 (1%) 72 70	30, 47, 78, 121	0
All	All	1736/1836 (94%)	0.30	45 (2%) 56 54	30, 50, 87, 126	0

The worst 5 of 45 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	709	VAL	5.4
1	D	630	THR	5.2
1	C	572	LEU	4.7
1	A	582	SER	4.4
1	C	711	THR	4.3

### 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](#)

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
2	SO4	B	802	5/5	0.91	0.16	58,58,95,100	0
2	SO4	C	803	5/5	0.92	0.15	76,80,92,95	0
2	SO4	D	803	5/5	0.93	0.23	64,69,97,103	0
2	SO4	A	802	5/5	0.94	0.11	62,70,85,93	0
2	SO4	B	803	5/5	0.94	0.13	67,82,91,96	0
2	SO4	C	802	5/5	0.95	0.13	50,51,71,79	0
2	SO4	A	801	5/5	0.95	0.15	54,63,75,91	0
2	SO4	A	803	5/5	0.95	0.14	71,72,80,92	0
2	SO4	B	801	5/5	0.96	0.15	55,64,80,91	0
2	SO4	D	801	5/5	0.97	0.13	64,69,79,80	0
2	SO4	D	804	5/5	0.97	0.14	56,61,73,79	0
2	SO4	C	801	5/5	0.99	0.19	46,46,47,49	0
2	SO4	D	802	5/5	0.99	0.12	51,71,75,83	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.