



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2026 – 04:09 AM UTC

PDB ID : 5ELS / pdb\_00005els  
Title : Structure of the KH domain of T-STAR in complex with AAAUAA RNA  
Authors : Dominguez, C.; Feracci, M.  
Deposited on : 2015-11-05  
Resolution : 2.87 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 : 4-5-2 with Phenix2.0  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 2.0  
EDS : 3.0  
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)  
CCP4 : 9.0.010 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.49

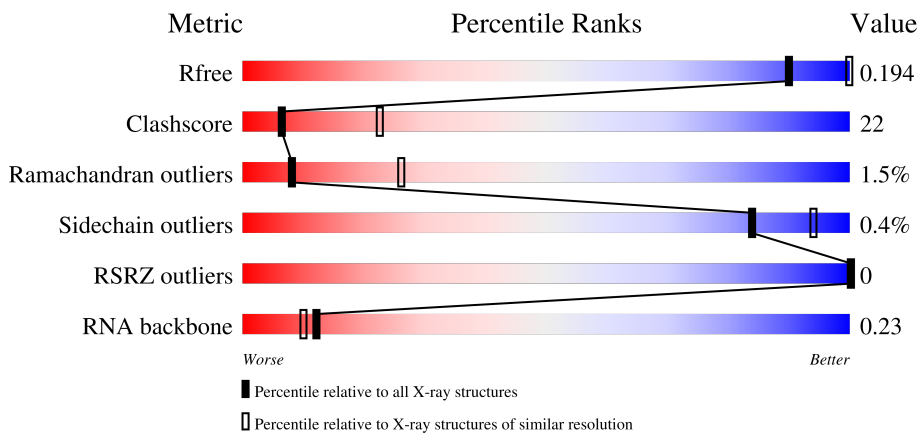
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.87 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.




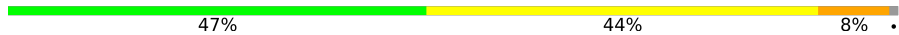


Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	180053	3557 (2.90-2.86)
Clashscore	190562	3801 (2.90-2.86)
Ramachandran outliers	187476	3699 (2.90-2.86)
Sidechain outliers	187428	3702 (2.90-2.86)
RSRZ outliers	180081	3558 (2.90-2.86)
RNA backbone	3983	1174 (3.10-2.66)

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	113	70% 30%
1	B	113	66% 30% .
1	C	113	62% 37% .
1	D	113	58% 37% . . .

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	E	113	 <p>52% 45%</p>
1	F	113	 <p>47% 44% 8%</p>
2	H	6	 <p>33% 17% 50%</p>
2	I	6	 <p>50% 33% 17%</p>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 5631 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 KH domain-containing, RNA-binding, signal transduction-associated protein 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	113	902	581	158	159	4	0	0	0
1	B	113	902	581	158	159	4	0	0	0
1	C	113	902	581	158	159	4	0	0	0
1	D	111	882	568	155	155	4	0	0	0
1	E	110	874	564	154	152	4	0	0	0
1	F	112	894	577	156	157	4	0	0	0

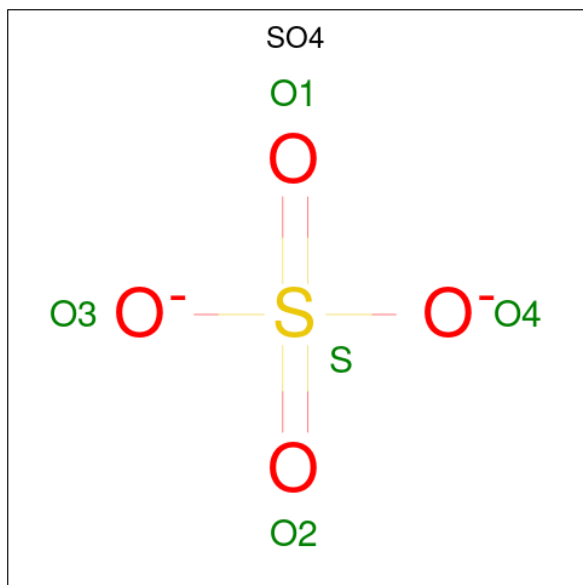
There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	48	GLY	-	expression tag	UNP O75525
A	49	ALA	-	expression tag	UNP O75525
B	48	GLY	-	expression tag	UNP O75525
B	49	ALA	-	expression tag	UNP O75525
C	48	GLY	-	expression tag	UNP O75525
C	49	ALA	-	expression tag	UNP O75525
D	48	GLY	-	expression tag	UNP O75525
D	49	ALA	-	expression tag	UNP O75525
E	48	GLY	-	expression tag	UNP O75525
E	49	ALA	-	expression tag	UNP O75525
F	48	GLY	-	expression tag	UNP O75525
F	49	ALA	-	expression tag	UNP O75525

- Molecule 2 is a RNA chain called RNA (5'-R(P\*AP\*AP\*AP\*UP\*AP\*A)-3').

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	H	6	130	59	27	38	6	0	0
2	I	6	130	59	27	38	6	0	0

- Molecule 3 is SULFATE ION (CCD ID: SO4) (formula: O<sub>4</sub>S).

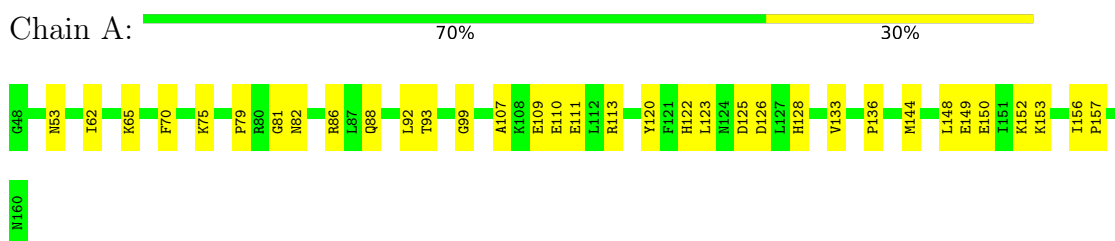


Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O S		
3	A	1	5	4 1	0	0
3	B	1	5	4 1	0	0
3	B	1	5	4 1	0	0

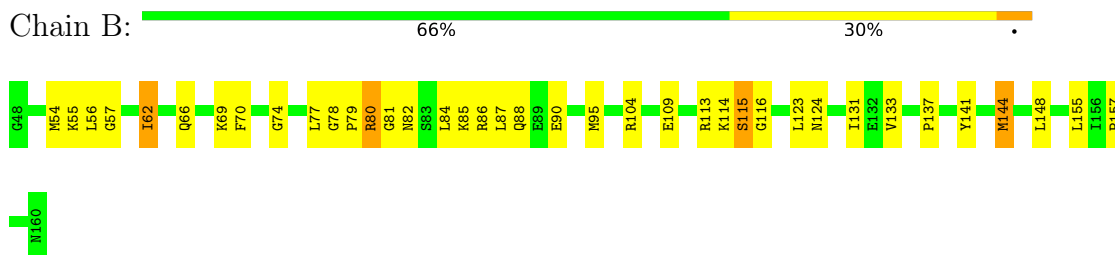
### 3 Residue-property plots

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.

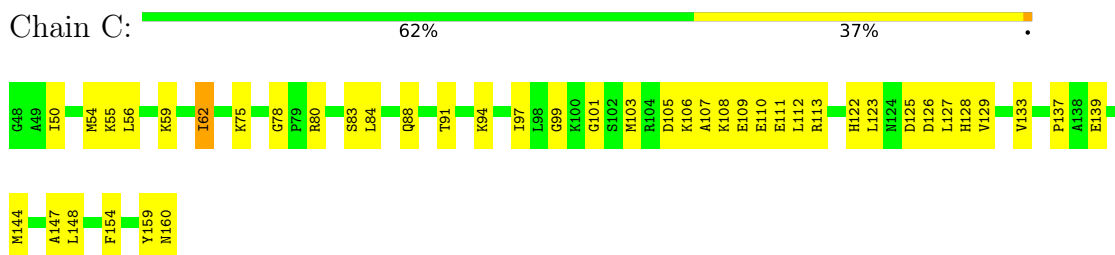
- Molecule 1: KH domain-containing, RNA-binding, signal transduction-associated protein 3



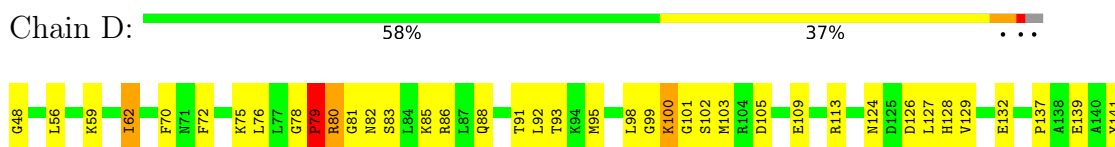
- Molecule 1: KH domain-containing, RNA-binding, signal transduction-associated protein 3



- Molecule 1: KH domain-containing, RNA-binding, signal transduction-associated protein 3

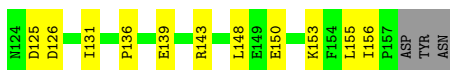


- Molecule 1: KH domain-containing, RNA-binding, signal transduction-associated protein 3

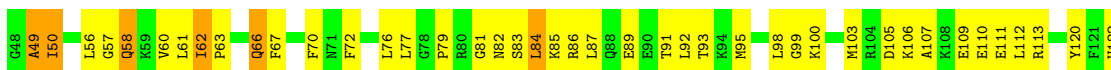




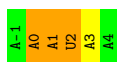
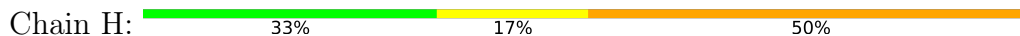
- Molecule 1: KH domain-containing, RNA-binding, signal transduction-associated protein 3



- Molecule 1: KH domain-containing, RNA-binding, signal transduction-associated protein 3



- Molecule 2: RNA (5'-R(P\*AP\*AP\*AP\*UP\*AP\*A)-3')



- Molecule 2: RNA (5'-R(P\*AP\*AP\*AP\*UP\*AP\*A)-3')



## 4 Data and refinement statistics i

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	93.73Å 162.22Å 113.04Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.87 – 2.87 46.87 – 2.87	Depositor EDS
% Data completeness (in resolution range)	99.8 (46.87-2.87) 97.8 (46.87-2.87)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.17	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.18 (at 2.86Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.188 , 0.236 0.176 , 0.194	Depositor DCC
$R_{free}$ test set	1020 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	43.4	Xtriage
Anisotropy	0.467	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 13.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.39$ , $\langle L^2 \rangle = 0.22$	Xtriage
Estimated twinning fraction	0.218 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.216 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
Reported twinning fraction	0.250 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Depositor
Outliers	0 of 19995 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	5631	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.59% 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	0.60	1/919 (0.1%)	0.76	0/1228
1	B	0.57	1/919 (0.1%)	0.84	0/1228
1	C	0.64	1/919 (0.1%)	0.77	0/1228
1	D	0.55	1/898 (0.1%)	0.83	0/1199
1	E	0.68	2/890 (0.2%)	0.79	0/1188
1	F	0.68	3/911 (0.3%)	1.00	5/1217 (0.4%)
2	H	0.25	0/146	0.37	0/225
2	I	0.20	0/146	0.34	0/225
All	All	0.61	9/5748 (0.2%)	0.82	5/7738 (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
1	D	0	1
1	F	0	2
All	All	0	3

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	62	ILE	C-N	9.86	1.56	1.33
1	E	113	ARG	NE-CZ	-9.81	1.22	1.33
1	A	62	ILE	C-N	8.10	1.52	1.33
1	E	113	ARG	CZ-NH1	-7.87	1.21	1.32
1	D	62	ILE	C-N	6.44	1.49	1.33

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	141	TYR	CA-CB-CG	-7.88	99.72	113.90
1	F	58	GLN	CA-CB-CG	-5.87	102.37	114.10
1	F	49	ALA	CA-C-N	5.52	131.91	121.97
1	F	49	ALA	C-N-CA	5.52	131.91	121.97
1	F	133	VAL	CG1-CB-CG2	-5.40	98.91	110.80

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	79	PRO	Peptide
1	F	127	LEU	Peptide
1	F	84	LEU	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	902	0	941	22	0
1	B	902	0	941	36	0
1	C	902	0	941	37	0
1	D	882	0	926	48	0
1	E	874	0	922	48	0
1	F	894	0	935	65	0
2	H	130	0	58	4	0
2	I	130	0	57	3	0
3	A	5	0	0	0	0
3	B	10	0	0	1	0
All	All	5631	0	5721	249	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 22.

The worst 5 of 249 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:F:77:LEU:HD13	1:F:84:LEU:HD23	1.20	1.16

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:77:LEU:CD1	1:F:84:LEU:HD23	1.90	1.01
1:A:150:GLU:O	1:A:153:LYS:NZ	1.97	0.95
1:E:84:LEU:HD22	1:E:88:GLN:HE21	1.30	0.94
1:B:86:ARG:NH1	3:B:202:SO4:O4	2.06	0.89

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	111/113 (98%)	108 (97%)	3 (3%)	0	100	100
1	B	111/113 (98%)	107 (96%)	2 (2%)	2 (2%)	6	22
1	C	111/113 (98%)	108 (97%)	3 (3%)	0	100	100
1	D	109/113 (96%)	102 (94%)	1 (1%)	6 (6%)	1	4
1	E	108/113 (96%)	104 (96%)	4 (4%)	0	100	100
1	F	110/113 (97%)	98 (89%)	10 (9%)	2 (2%)	6	22
All	All	660/678 (97%)	627 (95%)	23 (4%)	10 (2%)	8	26

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	115	SER
1	D	79	PRO
1	D	80	ARG
1	D	101	GLY
1	D	157	PRO

### 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	96/96 (100%)	96 (100%)	0	100	100
1	B	96/96 (100%)	95 (99%)	1 (1%)	68	87
1	C	96/96 (100%)	96 (100%)	0	100	100
1	D	94/96 (98%)	94 (100%)	0	100	100
1	E	93/96 (97%)	92 (99%)	1 (1%)	65	86
1	F	95/96 (99%)	95 (100%)	0	100	100
All	All	570/576 (99%)	568 (100%)	2 (0%)	84	94

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	144	MET
1	E	96	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	58	GLN
1	D	71	ASN
1	E	88	GLN
1	D	82	ASN
1	B	82	ASN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	H	5/6 (83%)	3 (60%)	0
2	I	5/6 (83%)	2 (40%)	0
All	All	10/12 (83%)	5 (50%)	0

All (5) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	H	0[A]	A
2	H	1[A]	A
2	H	2[A]	U
2	I	1[B]	A
2	I	2[B]	A

There are no RNA pucker outliers to report.

## 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 oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

3 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	$\# Z  > 2$
3	SO4	B	201	-	4,4,4	0.24	0	6,6,6	0.39	0
3	SO4	B	202	-	4,4,4	0.30	0	6,6,6	0.50	0
3	SO4	A	201	-	4,4,4	0.24	0	6,6,6	0.24	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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	202	SO4	1	0

## 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	113/113 (100%)	-1.61	0 100 100	17, 28, 46, 53	0
1	B	113/113 (100%)	-1.46	0 100 100	23, 39, 80, 95	0
1	C	113/113 (100%)	-1.48	0 100 100	14, 27, 68, 82	0
1	D	111/113 (98%)	-1.28	0 100 100	25, 51, 88, 113	0
1	E	110/113 (97%)	-1.13	0 100 100	35, 57, 104, 133	0
1	F	112/113 (99%)	-0.62	0 100 100	47, 82, 132, 189	0
2	H	6/6 (100%)	-0.40	0 100 100	15, 19, 24, 24	6 (100%)
2	I	6/6 (100%)	-0.37	0 100 100	15, 19, 22, 24	6 (100%)
All	All	684/690 (99%)	-1.25	0 100 100	14, 46, 106, 189	12 (1%)

There are no RSRZ outliers to report.

### 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 oligosaccharides 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
3	SO4	A	201	5/5	0.99	0.04	59,59,67,75	0
3	SO4	B	201	5/5	0.99	0.04	54,56,61,73	0
3	SO4	B	202	5/5	0.99	0.03	44,45,48,60	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.