



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 4, 2020 – 09:35 AM EST

PDB ID : 6NQ3  
Title : Crystal Structure of a SUZ12-RBBP4-PHF19-JARID2 Heterotetrameric Complex  
Authors : Chen, S.; Jiao, L.; Liu, X.  
Deposited on : 2019-01-19  
Resolution : 2.89 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
Xtriage (Phenix)	:	1.13
EDS	:	2.4
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac	:	5.8.0158
CCP4	:	7.0 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.4

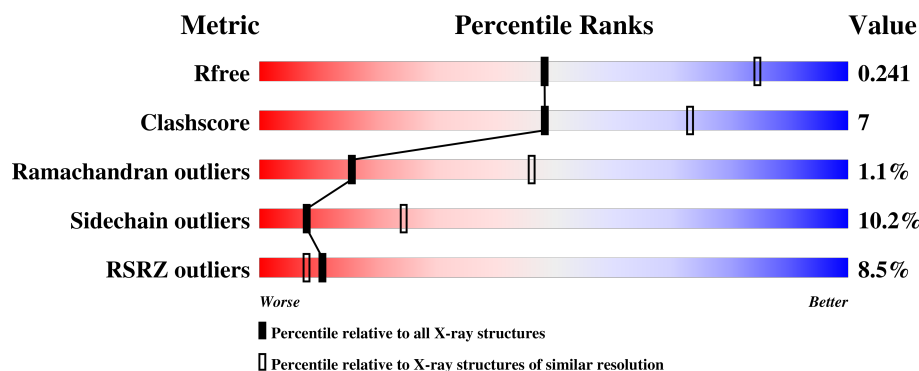
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.89 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))
$R_{free}$	111664	1716 (2.90-2.90)
Clashscore	122126	1924 (2.90-2.90)
Ramachandran outliers	120053	1884 (2.90-2.90)
Sidechain outliers	120020	1886 (2.90-2.90)
RSRZ outliers	108989	1669 (2.90-2.90)

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. 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	439	<div> <div>67%</div> <div>18%</div> <div>•</div> <div>13%</div> </div>
1	E	439	<div> <div>68%</div> <div>18%</div> <div>•</div> <div>13%</div> </div>
2	B	478	<div> <div>9%</div> <div>42%</div> <div>17%</div> <div>•</div> <div>39%</div> </div>
2	F	478	<div> <div>11%</div> <div>44%</div> <div>13%</div> <div>•</div> <div>42%</div> </div>
3	C	84	<div> <div>7%</div> <div>46%</div> <div>14%</div> <div>•</div> <div>37%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	84	<p>19% 38% 19% 39%</p>
4	D	19	<p>53% 37% 11%</p>
4	H	19	<p>11% 74% 16% 11%</p>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 11928 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 Histone-binding protein RBBP4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	381	Total	C	N	O	S	0	0	0
			3044	1920	518	596	10			
1	E	382	Total	C	N	O	S	0	0	0
			3045	1920	519	596	10			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-13	MET	-	initiating methionine	UNP Q09028
A	-12	SER	-	expression tag	UNP Q09028
A	-11	HIS	-	expression tag	UNP Q09028
A	-10	HIS	-	expression tag	UNP Q09028
A	-9	HIS	-	expression tag	UNP Q09028
A	-8	HIS	-	expression tag	UNP Q09028
A	-7	HIS	-	expression tag	UNP Q09028
A	-6	HIS	-	expression tag	UNP Q09028
A	-5	LEU	-	expression tag	UNP Q09028
A	-4	VAL	-	expression tag	UNP Q09028
A	-3	PRO	-	expression tag	UNP Q09028
A	-2	ARG	-	expression tag	UNP Q09028
A	-1	GLY	-	expression tag	UNP Q09028
A	0	SER	-	expression tag	UNP Q09028
E	-13	MET	-	initiating methionine	UNP Q09028
E	-12	SER	-	expression tag	UNP Q09028
E	-11	HIS	-	expression tag	UNP Q09028
E	-10	HIS	-	expression tag	UNP Q09028
E	-9	HIS	-	expression tag	UNP Q09028
E	-8	HIS	-	expression tag	UNP Q09028
E	-7	HIS	-	expression tag	UNP Q09028
E	-6	HIS	-	expression tag	UNP Q09028
E	-5	LEU	-	expression tag	UNP Q09028
E	-4	VAL	-	expression tag	UNP Q09028
E	-3	PRO	-	expression tag	UNP Q09028

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	-2	ARG	-	expression tag	UNP Q09028
E	-1	GLY	-	expression tag	UNP Q09028
E	0	SER	-	expression tag	UNP Q09028

- Molecule 2 is a protein called Polycomb protein SUZ12.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	291	Total	C	N	O	S	0	0	0
			2418	1551	442	411	14			
2	F	278	Total	C	N	O	S	0	0	0
			2317	1487	422	394	14			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	546	TRP	-	expression tag	UNP Q15022
B	547	SER	-	expression tag	UNP Q15022
B	548	HIS	-	expression tag	UNP Q15022
B	549	PRO	-	expression tag	UNP Q15022
B	550	GLN	-	expression tag	UNP Q15022
B	551	PHE	-	expression tag	UNP Q15022
B	552	GLU	-	expression tag	UNP Q15022
B	553	LYS	-	expression tag	UNP Q15022
F	546	TRP	-	expression tag	UNP Q15022
F	547	SER	-	expression tag	UNP Q15022
F	548	HIS	-	expression tag	UNP Q15022
F	549	PRO	-	expression tag	UNP Q15022
F	550	GLN	-	expression tag	UNP Q15022
F	551	PHE	-	expression tag	UNP Q15022
F	552	GLU	-	expression tag	UNP Q15022
F	553	LYS	-	expression tag	UNP Q15022

- Molecule 3 is a protein called PHD finger protein 19.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	53	Total	C	N	O	S	0	0	0
			416	262	71	82	1			
3	G	51	Total	C	N	O	S	0	0	0
			400	254	69	76	1			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	497	SER	-	expression tag	UNP Q5T6S3
C	498	ASN	-	expression tag	UNP Q5T6S3
C	499	ALA	-	expression tag	UNP Q5T6S3
G	497	SER	-	expression tag	UNP Q5T6S3
G	498	ASN	-	expression tag	UNP Q5T6S3
G	499	ALA	-	expression tag	UNP Q5T6S3

- Molecule 4 is a protein called Protein Jumonji.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	17	Total	C	N	O	S	0	0	0
			143	93	26	23	1			
4	H	17	Total	C	N	O	S	0	0	0
			143	93	26	23	1			

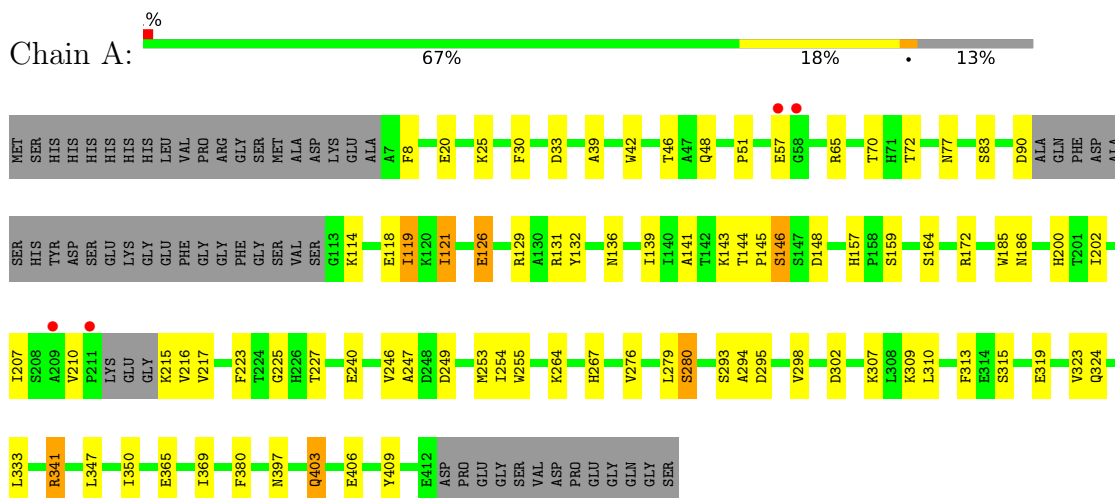
- Molecule 5 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total	Zn	0	0
			1	1		
5	F	1	Total	Zn	0	0
			1	1		

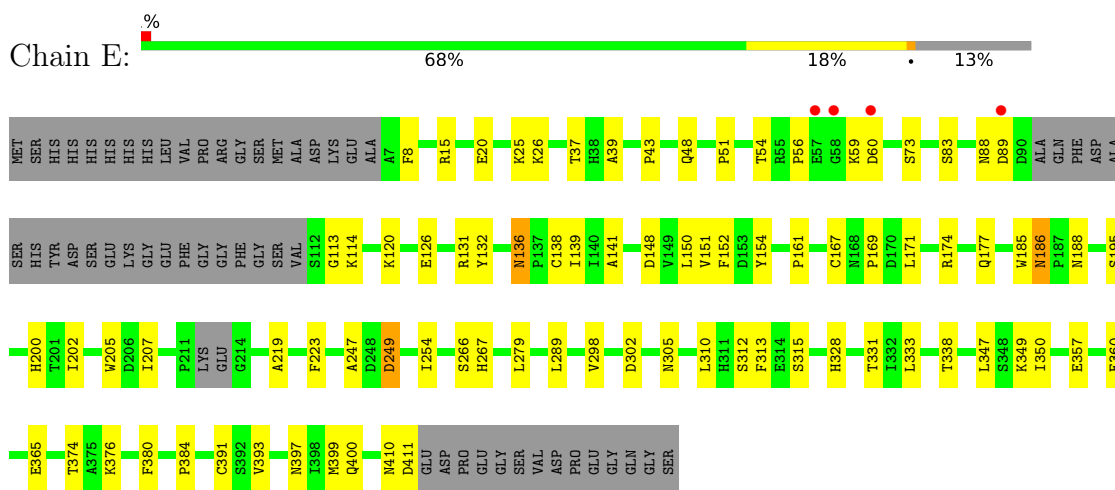
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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.

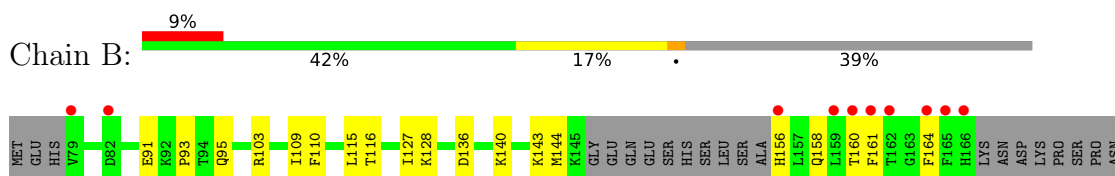
#### • Molecule 1: Histone-binding protein RBBP4



#### • Molecule 1: Histone-binding protein RBBP4

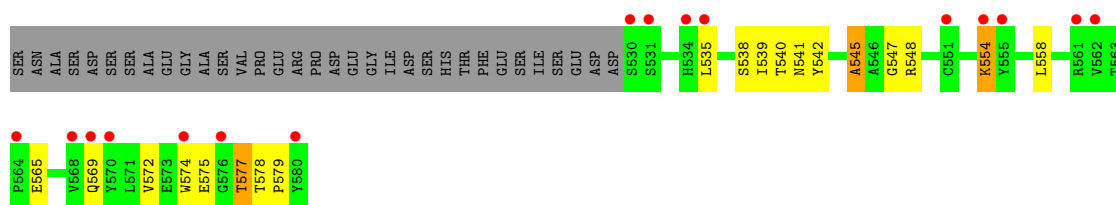


#### • Molecule 2: Polycomb protein SUZ12





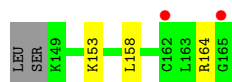
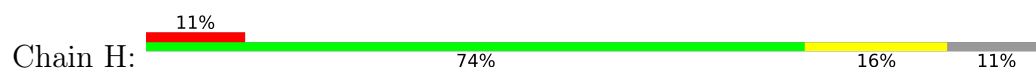




- Molecule 4: Protein Jumonji



- Molecule 4: Protein Jumonji



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	127.71Å 139.60Å 268.08Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.60 – 2.89 46.60 – 2.89	Depositor EDS
% Data completeness (in resolution range)	82.0 (46.60-2.89) 82.0 (46.60-2.89)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.06 (at 2.91Å)	Xtriage
Refinement program	BUSTER 2.10.3	Depositor
R, $R_{free}$	0.170 , 0.230 0.183 , 0.241	Depositor DCC
$R_{free}$ test set	2180 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	50.3	Xtriage
Anisotropy	0.198	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 80.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	11928	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	67.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.87% 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:  
ZN

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	0.53	0/3127	0.76	0/4264
1	E	0.49	0/3128	0.74	0/4265
2	B	0.55	0/2469	0.73	0/3317
2	F	0.51	0/2365	0.71	0/3172
3	C	0.48	0/425	0.71	0/574
3	G	0.46	0/409	0.71	0/552
4	D	0.53	0/145	0.65	0/191
4	H	0.49	0/145	0.62	0/191
All	All	0.52	0/12213	0.73	0/16526

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 [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	3044	0	2890	41	0
1	E	3045	0	2892	39	0
2	B	2418	0	2461	47	0
2	F	2317	0	2351	31	0
3	C	416	0	402	7	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	G	400	0	394	7	0
4	D	143	0	154	4	0
4	H	143	0	154	3	0
5	B	1	0	0	0	0
5	F	1	0	0	0	0
All	All	11928	0	11698	154	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 7.

The worst 5 of 154 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)
2:B:308:LEU:HD12	2:B:308:LEU:H	1.42	0.82
1:A:48:GLN:HE22	1:A:131:ARG:HA	1.45	0.81
1:A:121:ILE:HD13	1:A:157:HIS:CD2	2.21	0.76
1:A:48:GLN:NE2	1:A:131:ARG:HA	2.03	0.73
2:B:241:ASN:N	2:B:241:ASN:OD1	2.22	0.72

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	375/439 (85%)	354 (94%)	20 (5%)	1 (0%)	43	75
1	E	376/439 (86%)	357 (95%)	15 (4%)	4 (1%)	16	46
2	B	271/478 (57%)	253 (93%)	16 (6%)	2 (1%)	24	58
2	F	258/478 (54%)	225 (87%)	29 (11%)	4 (2%)	11	35
3	C	51/84 (61%)	48 (94%)	2 (4%)	1 (2%)	8	30

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	G	49/84 (58%)	41 (84%)	4 (8%)	4 (8%)	1	2
4	D	15/19 (79%)	14 (93%)	1 (7%)	0	100	100
4	H	15/19 (79%)	15 (100%)	0	0	100	100
All	All	1410/2040 (69%)	1307 (93%)	87 (6%)	16 (1%)	16	46

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	89	ASP
2	F	435	ASN
1	A	315	SER
3	C	529	ASP
1	E	113	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	343/388 (88%)	317 (92%)	26 (8%)	14	39
1	E	343/388 (88%)	317 (92%)	26 (8%)	14	39
2	B	272/442 (62%)	237 (87%)	35 (13%)	5	14
2	F	259/442 (59%)	227 (88%)	32 (12%)	5	15
3	C	44/70 (63%)	37 (84%)	7 (16%)	3	8
3	G	42/70 (60%)	34 (81%)	8 (19%)	1	5
4	D	16/18 (89%)	15 (94%)	1 (6%)	20	50
4	H	16/18 (89%)	15 (94%)	1 (6%)	20	50
All	All	1335/1836 (73%)	1199 (90%)	136 (10%)	8	24

5 of 136 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	528	ASP

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
1	E	88	ASN
2	F	525	ILE
3	C	549	LEU
1	E	8	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 31 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	431	GLN
3	C	534	HIS
2	F	456	ASN
2	B	435	ASN
1	E	38	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

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	381/439 (86%)	-0.25	4 (1%) 82 81	18, 44, 77, 110	0
1	E	382/439 (87%)	-0.12	4 (1%) 82 81	22, 56, 106, 124	0
2	B	291/478 (60%)	0.42	42 (14%) 2 2	24, 63, 110, 131	0
2	F	278/478 (58%)	0.75	51 (18%) 1 1	40, 84, 139, 159	0
3	C	53/84 (63%)	0.55	6 (11%) 5 4	43, 86, 116, 130	0
3	G	51/84 (60%)	1.62	16 (31%) 0 0	88, 129, 155, 167	0
4	D	17/19 (89%)	0.14	0 100 100	33, 50, 100, 101	0
4	H	17/19 (89%)	1.16	2 (11%) 4 3	71, 86, 114, 135	0
All	All	1470/2040 (72%)	0.22	125 (8%) 11 8	18, 59, 125, 167	0

The worst 5 of 125 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	161	PHE	9.7
2	B	230	LEU	7.7
2	F	354	THR	7.1
2	F	164	PHE	6.6
2	F	159	LEU	6.6

### 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](#)

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
5	ZN	F	1001	1/1	1.00	0.14	61,61,61,61	0
5	ZN	B	1001	1/1	1.00	0.14	40,40,40,40	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.