



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 26, 2020 – 09:17 am BST

PDB ID : 6M9X  
Title : X-ray Structure of Branchiostoma floridae fluorescent protein lanFP10A  
Authors : Muslinkina, L.; Pletneva, N.; Pletnev, V.; Pletnev, S.  
Deposited on : 2018-08-24  
Resolution : 1.81 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 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 : 2.11  
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) : 2.11

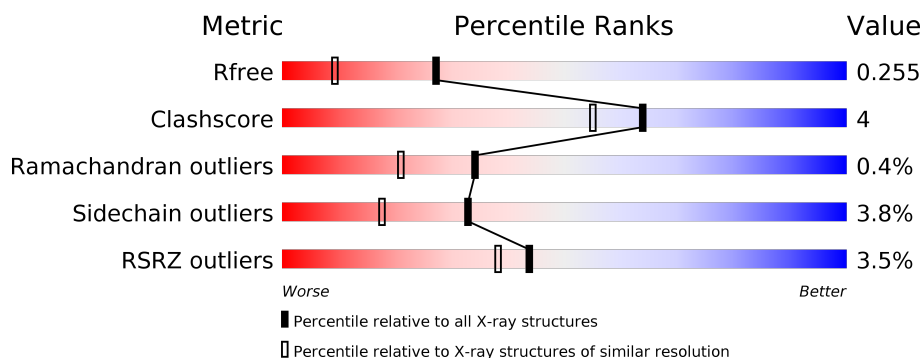
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.81 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	7484 (1.84-1.80)
Clashscore	141614	8401 (1.84-1.80)
Ramachandran outliers	138981	8290 (1.84-1.80)
Sidechain outliers	138945	8290 (1.84-1.80)
RSRZ outliers	127900	7371 (1.84-1.80)

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 on 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	227	<div> <div>2%</div> <div> <div></div> <div>85%</div> <div>9%</div> <div>5%</div> </div> </div>
1	B	227	<div> <div>5%</div> <div> <div></div> <div>82%</div> <div>12%</div> <div>5%</div> </div> </div>
1	C	227	<div> <div>3%</div> <div> <div></div> <div>84%</div> <div>10%</div> <div>5%</div> </div> </div>
1	D	227	<div> <div>3%</div> <div> <div></div> <div>83%</div> <div>11%</div> <div>5%</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 7444 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 Fluorescent protein lanFP10A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	216	Total	C	N	O	S	0	3	0
			1765	1132	295	327	11			
1	B	216	Total	C	N	O	S	0	0	0
			1739	1117	288	323	11			
1	C	216	Total	C	N	O	S	0	3	0
			1762	1129	293	329	11			
1	D	216	Total	C	N	O	S	0	3	0
			1763	1130	292	330	11			

There are 56 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	expression tag	UNP C3YRA2
A	58	JBYP	GLY	chromophore	UNP C3YRA2
A	58	JBYP	TYR	chromophore	UNP C3YRA2
A	58	JBYP	ALA	chromophore	UNP C3YRA2
A	220	SER	-	expression tag	UNP C3YRA2
A	221	GLY	-	expression tag	UNP C3YRA2
A	222	GLY	-	expression tag	UNP C3YRA2
A	223	SER	-	expression tag	UNP C3YRA2
A	224	HIS	-	expression tag	UNP C3YRA2
A	225	HIS	-	expression tag	UNP C3YRA2
A	226	HIS	-	expression tag	UNP C3YRA2
A	227	HIS	-	expression tag	UNP C3YRA2
A	228	HIS	-	expression tag	UNP C3YRA2
A	229	HIS	-	expression tag	UNP C3YRA2
B	1	MET	-	expression tag	UNP C3YRA2
B	58	JBYP	GLY	chromophore	UNP C3YRA2
B	58	JBYP	TYR	chromophore	UNP C3YRA2
B	58	JBYP	ALA	chromophore	UNP C3YRA2
B	220	SER	-	expression tag	UNP C3YRA2
B	221	GLY	-	expression tag	UNP C3YRA2
B	222	GLY	-	expression tag	UNP C3YRA2

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	223	SER	-	expression tag	UNP C3YRA2
B	224	HIS	-	expression tag	UNP C3YRA2
B	225	HIS	-	expression tag	UNP C3YRA2
B	226	HIS	-	expression tag	UNP C3YRA2
B	227	HIS	-	expression tag	UNP C3YRA2
B	228	HIS	-	expression tag	UNP C3YRA2
B	229	HIS	-	expression tag	UNP C3YRA2
C	1	MET	-	expression tag	UNP C3YRA2
C	58	JBYP	GLY	chromophore	UNP C3YRA2
C	58	JBYP	TYR	chromophore	UNP C3YRA2
C	58	JBYP	ALA	chromophore	UNP C3YRA2
C	220	SER	-	expression tag	UNP C3YRA2
C	221	GLY	-	expression tag	UNP C3YRA2
C	222	GLY	-	expression tag	UNP C3YRA2
C	223	SER	-	expression tag	UNP C3YRA2
C	224	HIS	-	expression tag	UNP C3YRA2
C	225	HIS	-	expression tag	UNP C3YRA2
C	226	HIS	-	expression tag	UNP C3YRA2
C	227	HIS	-	expression tag	UNP C3YRA2
C	228	HIS	-	expression tag	UNP C3YRA2
C	229	HIS	-	expression tag	UNP C3YRA2
D	1	MET	-	expression tag	UNP C3YRA2
D	58	JBYP	GLY	chromophore	UNP C3YRA2
D	58	JBYP	TYR	chromophore	UNP C3YRA2
D	58	JBYP	ALA	chromophore	UNP C3YRA2
D	220	SER	-	expression tag	UNP C3YRA2
D	221	GLY	-	expression tag	UNP C3YRA2
D	222	GLY	-	expression tag	UNP C3YRA2
D	223	SER	-	expression tag	UNP C3YRA2
D	224	HIS	-	expression tag	UNP C3YRA2
D	225	HIS	-	expression tag	UNP C3YRA2
D	226	HIS	-	expression tag	UNP C3YRA2
D	227	HIS	-	expression tag	UNP C3YRA2
D	228	HIS	-	expression tag	UNP C3YRA2
D	229	HIS	-	expression tag	UNP C3YRA2

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	125	Total O 125 125	0	0
2	B	96	Total O 96 96	0	0

*Continued on next page...*

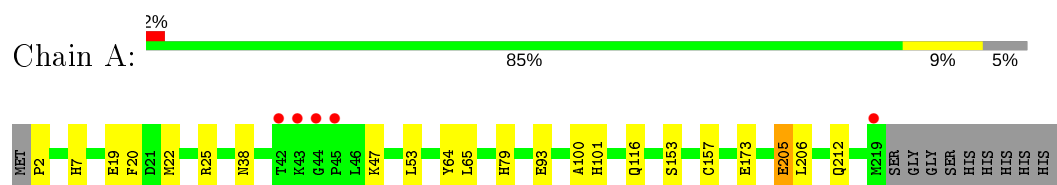
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	C	94	Total 94	O 94	0	0
2	D	100	Total 100	O 100	0	0

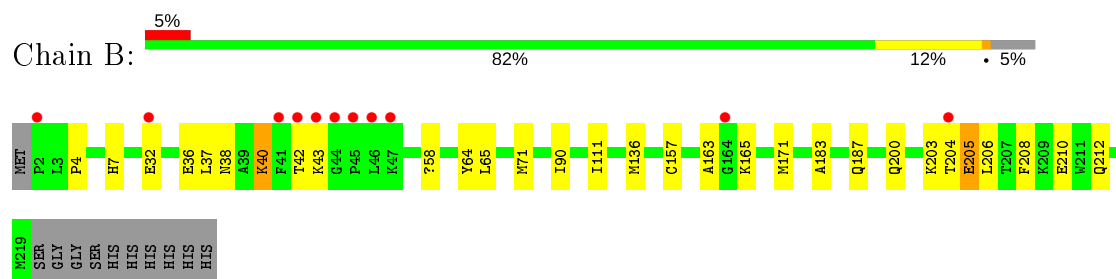
### 3 Residue-property plots [i](#)

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.

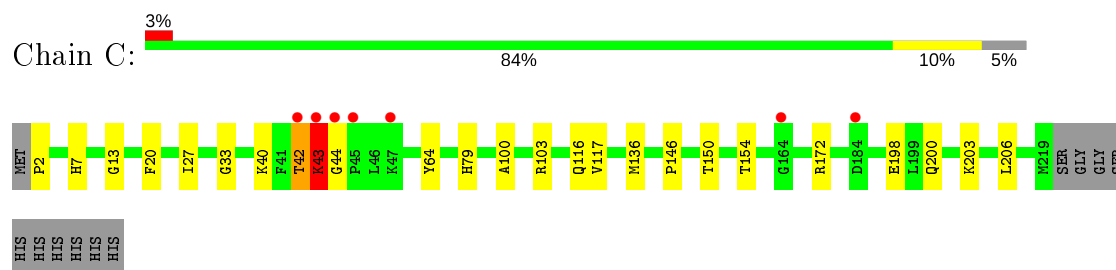
- Molecule 1: Fluorescent protein lanFP10A



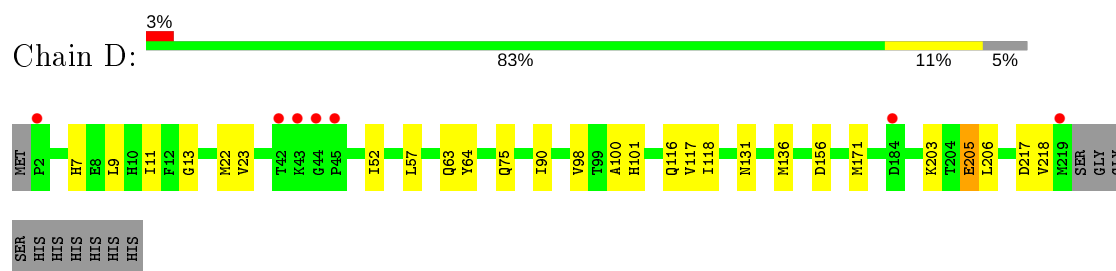
- Molecule 1: Fluorescent protein lanFP10A



- Molecule 1: Fluorescent protein lanFP10A



- Molecule 1: Fluorescent protein lanFP10A



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	58.01Å 115.84Å 128.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.65 – 1.81 29.64 – 1.81	Depositor EDS
% Data completeness (in resolution range)	98.8 (29.65-1.81) 98.9 (29.64-1.81)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.03 (at 1.82Å)	Xtriage
Refinement program	REFMAC 5.8.0222	Depositor
R, $R_{free}$	0.201 , 0.254 0.207 , 0.255	Depositor DCC
$R_{free}$ test set	1571 reflections (1.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	26.6	Xtriage
Anisotropy	0.414	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 36.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	7444	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.48% 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: JBY

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	0.63	1/1796 (0.1%)	0.77	0/2426
1	B	0.58	0/1770	0.74	0/2393
1	C	0.59	0/1793	0.76	0/2424
1	D	0.58	0/1794	0.76	0/2425
All	All	0.60	1/7153 (0.0%)	0.76	0/9668

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	C	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	93	GLU	CD-OE1	-7.87	1.17	1.25

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	103	ARG	Sidechain



## 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	1765	0	1687	15	0
1	B	1739	0	1659	15	0
1	C	1762	0	1675	12	0
1	D	1763	0	1675	16	0
2	A	125	0	0	4	0
2	B	96	0	0	4	0
2	C	94	0	0	1	0
2	D	100	0	0	2	0
All	All	7444	0	6696	58	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 4.

The worst 5 of 58 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:A:19:GLU:OE2	2:A:301:HOH:O	1.96	0.82
1:A:2:PRO:N	1:A:79:HIS:HD1	1.79	0.79
1:B:38:ASN:HD21	1:B:205:GLU:HG3	1.50	0.74
1:C:2:PRO:N	1:C:79:HIS:HD1	1.91	0.69
1:C:7:HIS:HD2	1:C:64:TYR:OH	1.79	0.65

There are no symmetry-related clashes.

## 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	216/227 (95%)	212 (98%)	4 (2%)	0	100	100
1	B	213/227 (94%)	211 (99%)	1 (0%)	1 (0%)	29	15
1	C	216/227 (95%)	211 (98%)	3 (1%)	2 (1%)	17	6
1	D	216/227 (95%)	214 (99%)	2 (1%)	0	100	100
All	All	861/908 (95%)	848 (98%)	10 (1%)	3 (0%)	34	27

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	44	GLY
1	C	43	LYS
1	B	163	ALA

### 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	189/195 (97%)	185 (98%)	4 (2%)	53	41
1	B	186/195 (95%)	175 (94%)	11 (6%)	19	7
1	C	189/195 (97%)	180 (95%)	9 (5%)	25	11
1	D	189/195 (97%)	185 (98%)	4 (2%)	53	41
All	All	753/780 (96%)	725 (96%)	28 (4%)	33	19

5 of 28 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	203	LYS
1	C	20	PHE
1	D	136	MET
1	B	204	THR
1	B	205	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	131	ASN
1	B	200	GLN
1	C	174	ASN
1	B	101	HIS
1	C	201	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

4 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	JBY	A	58	1	19,21,22	3.60	6 (31%)	24,29,31	3.69	8 (33%)
1	JBY	D	58	1	19,21,22	3.81	5 (26%)	24,29,31	3.46	7 (29%)
1	JBY	B	58	1	19,21,22	3.83	5 (26%)	24,29,31	4.18	7 (29%)
1	JBY	C	58	1	19,21,22	3.65	5 (26%)	24,29,31	3.42	4 (16%)

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	JBY	A	58	1	-	0/8/28/30	0/2/2/2
1	JBY	D	58	1	-	0/8/28/30	0/2/2/2
1	JBY	B	58	1	-	0/8/28/30	0/2/2/2
1	JBY	C	58	1	-	1/8/28/30	0/2/2/2

The worst 5 of 21 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	58	JB	CB2-CA2	14.43	1.47	1.35
1	D	58	JB	CB2-CA2	14.34	1.47	1.35
1	C	58	JB	CB2-CA2	13.37	1.46	1.35
1	A	58	JB	CB2-CA2	12.75	1.45	1.35
1	C	58	JB	CA2-C2	-6.00	1.42	1.48

The worst 5 of 26 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	58	JB	CA2-C2-N3	14.73	110.18	103.39
1	B	58	JB	O2-C2-CA2	-12.06	124.19	130.96
1	C	58	JB	CA2-C2-N3	11.87	108.86	103.39
1	D	58	JB	CA2-C2-N3	11.66	108.76	103.39
1	A	58	JB	O2-C2-CA2	-11.59	124.45	130.96

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	C	58	JB	CB3-CA3-N3-C1

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	B	58	JB	1	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates 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 ⓘ

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	215/227 (94%)	-0.06	5 (2%) 60 56	18, 26, 50, 103	0
1	B	215/227 (94%)	0.20	11 (5%) 28 22	20, 33, 62, 89	0
1	C	215/227 (94%)	0.13	7 (3%) 46 40	20, 30, 54, 95	0
1	D	215/227 (94%)	0.09	7 (3%) 46 40	19, 31, 53, 89	0
All	All	860/908 (94%)	0.09	30 (3%) 44 38	18, 30, 54, 103	0

The worst 5 of 30 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	42	THR	12.3
1	B	42	THR	6.3
1	D	43	LYS	5.2
1	D	42	THR	5.2
1	D	45	PRO	4.7

### 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(Å <sup>2</sup> )	Q<0.9
1	JBV	B	58	20/21	0.92	0.13	25,28,32,32	0
1	JBV	D	58	20/21	0.95	0.12	23,27,33,33	0
1	JBV	A	58	20/21	0.96	0.12	19,22,25,27	0
1	JBV	C	58	20/21	0.96	0.13	22,25,29,30	0

### 6.3 Carbohydrates [i](#)

There are no carbohydrates 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.