



# wwPDB X-ray Structure Validation Summary Report

May 7, 2014 – 12:29 AM EDT

PDB ID : 4JHF  
Title : Crystal Structure of Translin  
Authors : Dvir, H.; Eliahoo, E.; Alian, A.  
Deposited on : 2013-03-05  
Resolution : 3.00 Å(reported)

This is a wwPDB 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 <http://wwpdb.org/ValidationPDFNotes.html>

---

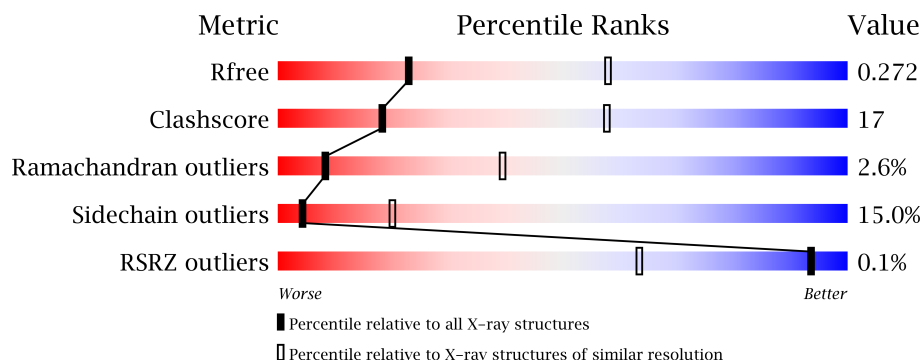
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.16 November 2013  
Xtriage (Phenix) : dev-1439  
EDS : stable22978  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22978

# 1 Overall quality at a glance

The reported resolution of this entry is 3.00 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))
$R_{free}$	66092	1216 (3.00-3.00)
Clashscore	79885	1594 (3.00-3.00)
Ramachandran outliers	78287	1537 (3.00-3.00)
Sidechain outliers	78261	1540 (3.00-3.00)
RSRZ outliers	66119	1217 (3.00-3.00)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	240	
1	B	240	
1	C	240	
1	D	240	
1	E	240	
1	F	240	
1	G	240	
1	H	240	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 13933 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Translin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	213	Total	C	N	O	S	0	0	0
			1741	1116	300	323	2			
1	B	212	Total	C	N	O	S	0	0	0
			1733	1111	298	322	2			
1	C	214	Total	C	N	O	S	0	0	0
			1741	1113	301	325	2			
1	D	216	Total	C	N	O	S	0	0	0
			1767	1133	304	328	2			
1	E	213	Total	C	N	O	S	0	0	0
			1743	1117	300	324	2			
1	F	212	Total	C	N	O	S	0	0	0
			1734	1112	298	322	2			
1	G	212	Total	C	N	O	S	0	0	0
			1737	1114	299	322	2			
1	H	210	Total	C	N	O	S	0	0	0
			1717	1100	296	319	2			

There are 96 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-11	MET	-	EXPRESSION TAG	UNP Q15631
A	-10	ARG	-	EXPRESSION TAG	UNP Q15631
A	-9	GLY	-	EXPRESSION TAG	UNP Q15631
A	-8	SER	-	EXPRESSION TAG	UNP Q15631
A	-7	HIS	-	EXPRESSION TAG	UNP Q15631
A	-6	HIS	-	EXPRESSION TAG	UNP Q15631
A	-5	HIS	-	EXPRESSION TAG	UNP Q15631
A	-4	HIS	-	EXPRESSION TAG	UNP Q15631
A	-3	HIS	-	EXPRESSION TAG	UNP Q15631
A	-2	HIS	-	EXPRESSION TAG	UNP Q15631
A	-1	GLY	-	EXPRESSION TAG	UNP Q15631
A	0	SER	-	EXPRESSION TAG	UNP Q15631
B	-11	MET	-	EXPRESSION TAG	UNP Q15631

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	-10	ARG	-	EXPRESSION TAG	UNP Q15631
B	-9	GLY	-	EXPRESSION TAG	UNP Q15631
B	-8	SER	-	EXPRESSION TAG	UNP Q15631
B	-7	HIS	-	EXPRESSION TAG	UNP Q15631
B	-6	HIS	-	EXPRESSION TAG	UNP Q15631
B	-5	HIS	-	EXPRESSION TAG	UNP Q15631
B	-4	HIS	-	EXPRESSION TAG	UNP Q15631
B	-3	HIS	-	EXPRESSION TAG	UNP Q15631
B	-2	HIS	-	EXPRESSION TAG	UNP Q15631
B	-1	GLY	-	EXPRESSION TAG	UNP Q15631
B	0	SER	-	EXPRESSION TAG	UNP Q15631
C	-11	MET	-	EXPRESSION TAG	UNP Q15631
C	-10	ARG	-	EXPRESSION TAG	UNP Q15631
C	-9	GLY	-	EXPRESSION TAG	UNP Q15631
C	-8	SER	-	EXPRESSION TAG	UNP Q15631
C	-7	HIS	-	EXPRESSION TAG	UNP Q15631
C	-6	HIS	-	EXPRESSION TAG	UNP Q15631
C	-5	HIS	-	EXPRESSION TAG	UNP Q15631
C	-4	HIS	-	EXPRESSION TAG	UNP Q15631
C	-3	HIS	-	EXPRESSION TAG	UNP Q15631
C	-2	HIS	-	EXPRESSION TAG	UNP Q15631
C	-1	GLY	-	EXPRESSION TAG	UNP Q15631
C	0	SER	-	EXPRESSION TAG	UNP Q15631
D	-11	MET	-	EXPRESSION TAG	UNP Q15631
D	-10	ARG	-	EXPRESSION TAG	UNP Q15631
D	-9	GLY	-	EXPRESSION TAG	UNP Q15631
D	-8	SER	-	EXPRESSION TAG	UNP Q15631
D	-7	HIS	-	EXPRESSION TAG	UNP Q15631
D	-6	HIS	-	EXPRESSION TAG	UNP Q15631
D	-5	HIS	-	EXPRESSION TAG	UNP Q15631
D	-4	HIS	-	EXPRESSION TAG	UNP Q15631
D	-3	HIS	-	EXPRESSION TAG	UNP Q15631
D	-2	HIS	-	EXPRESSION TAG	UNP Q15631
D	-1	GLY	-	EXPRESSION TAG	UNP Q15631
D	0	SER	-	EXPRESSION TAG	UNP Q15631
E	-11	MET	-	EXPRESSION TAG	UNP Q15631
E	-10	ARG	-	EXPRESSION TAG	UNP Q15631
E	-9	GLY	-	EXPRESSION TAG	UNP Q15631
E	-8	SER	-	EXPRESSION TAG	UNP Q15631
E	-7	HIS	-	EXPRESSION TAG	UNP Q15631
E	-6	HIS	-	EXPRESSION TAG	UNP Q15631
E	-5	HIS	-	EXPRESSION TAG	UNP Q15631

*Continued on next page...*

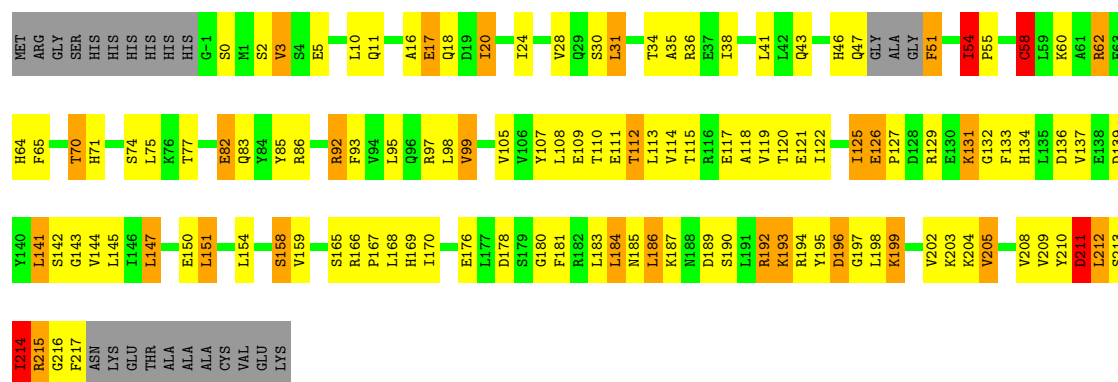
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	-4	HIS	-	EXPRESSION TAG	UNP Q15631
E	-3	HIS	-	EXPRESSION TAG	UNP Q15631
E	-2	HIS	-	EXPRESSION TAG	UNP Q15631
E	-1	GLY	-	EXPRESSION TAG	UNP Q15631
E	0	SER	-	EXPRESSION TAG	UNP Q15631
F	-11	MET	-	EXPRESSION TAG	UNP Q15631
F	-10	ARG	-	EXPRESSION TAG	UNP Q15631
F	-9	GLY	-	EXPRESSION TAG	UNP Q15631
F	-8	SER	-	EXPRESSION TAG	UNP Q15631
F	-7	HIS	-	EXPRESSION TAG	UNP Q15631
F	-6	HIS	-	EXPRESSION TAG	UNP Q15631
F	-5	HIS	-	EXPRESSION TAG	UNP Q15631
F	-4	HIS	-	EXPRESSION TAG	UNP Q15631
F	-3	HIS	-	EXPRESSION TAG	UNP Q15631
F	-2	HIS	-	EXPRESSION TAG	UNP Q15631
F	-1	GLY	-	EXPRESSION TAG	UNP Q15631
F	0	SER	-	EXPRESSION TAG	UNP Q15631
G	-11	MET	-	EXPRESSION TAG	UNP Q15631
G	-10	ARG	-	EXPRESSION TAG	UNP Q15631
G	-9	GLY	-	EXPRESSION TAG	UNP Q15631
G	-8	SER	-	EXPRESSION TAG	UNP Q15631
G	-7	HIS	-	EXPRESSION TAG	UNP Q15631
G	-6	HIS	-	EXPRESSION TAG	UNP Q15631
G	-5	HIS	-	EXPRESSION TAG	UNP Q15631
G	-4	HIS	-	EXPRESSION TAG	UNP Q15631
G	-3	HIS	-	EXPRESSION TAG	UNP Q15631
G	-2	HIS	-	EXPRESSION TAG	UNP Q15631
G	-1	GLY	-	EXPRESSION TAG	UNP Q15631
G	0	SER	-	EXPRESSION TAG	UNP Q15631
H	-11	MET	-	EXPRESSION TAG	UNP Q15631
H	-10	ARG	-	EXPRESSION TAG	UNP Q15631
H	-9	GLY	-	EXPRESSION TAG	UNP Q15631
H	-8	SER	-	EXPRESSION TAG	UNP Q15631
H	-7	HIS	-	EXPRESSION TAG	UNP Q15631
H	-6	HIS	-	EXPRESSION TAG	UNP Q15631
H	-5	HIS	-	EXPRESSION TAG	UNP Q15631
H	-4	HIS	-	EXPRESSION TAG	UNP Q15631
H	-3	HIS	-	EXPRESSION TAG	UNP Q15631
H	-2	HIS	-	EXPRESSION TAG	UNP Q15631
H	-1	GLY	-	EXPRESSION TAG	UNP Q15631
H	0	SER	-	EXPRESSION TAG	UNP Q15631

- Molecule 2 is water.

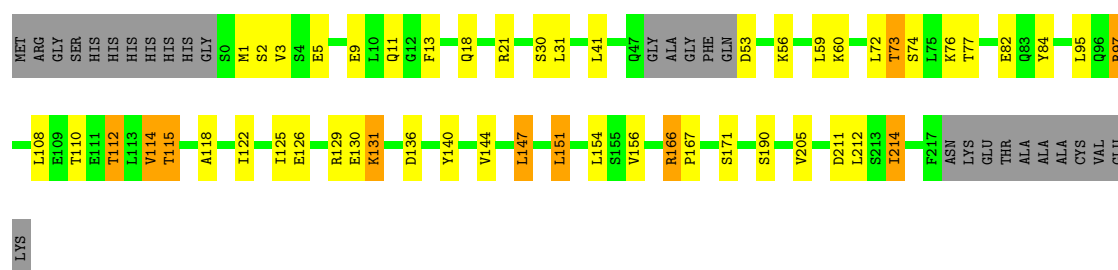
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	6	Total 6	O 6	0	0
2	B	3	Total 3	O 3	0	0
2	C	2	Total 2	O 2	0	0
2	D	1	Total 1	O 1	0	0
2	E	2	Total 2	O 2	0	0
2	G	3	Total 3	O 3	0	0
2	H	3	Total 3	O 3	0	0





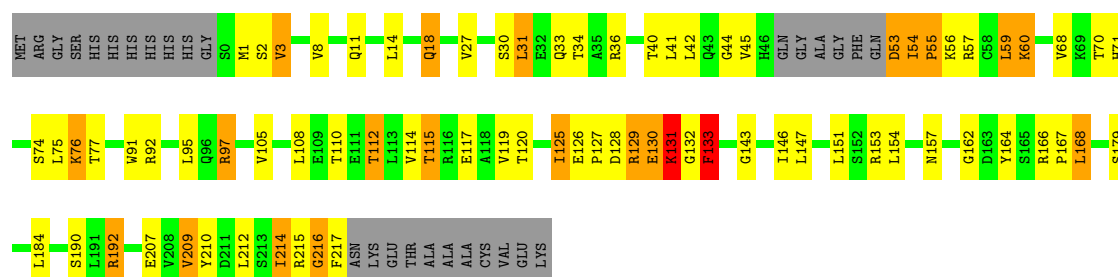
• Molecule 1: Translin

Chain E:



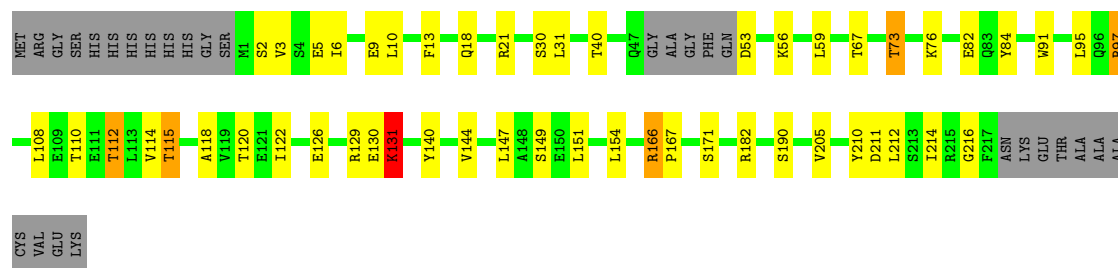
• Molecule 1: Translin

Chain F:



• Molecule 1: Translin

Chain G:



• Molecule 1: Translin

Chain H:





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	82.31Å 82.31Å 635.16Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	58.20 – 3.00 57.96 – 3.00	Depositor EDS
% Data completeness (in resolution range)	98.5 (58.20-3.00) 98.6 (57.96-3.00)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.93 (at 3.01Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.205 , 0.271 0.207 , 0.272	Depositor DCC
$R_{free}$ test set	2272 reflections (5.31%)	DCC
Wilson B-factor (Å <sup>2</sup> )	75.0	Xtriage
Anisotropy	0.493	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 35.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.42$ , $\langle L^2 \rangle = 0.24$	Xtriage
Outliers	1 of 45031 reflections (0.002%)	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	13933	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	79.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.33% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

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	0.69	1/1772 (0.1%)	0.77	0/2390
1	B	0.72	1/1764 (0.1%)	0.79	0/2381
1	C	0.73	0/1771	0.87	1/2389 (0.0%)
1	D	0.76	0/1799	0.90	1/2426 (0.0%)
1	E	0.61	0/1774	0.72	0/2393
1	F	0.73	1/1765 (0.1%)	0.82	0/2381
1	G	0.60	1/1768 (0.1%)	0.73	0/2385
1	H	0.69	0/1747	0.77	0/2357
All	All	0.70	4/14160 (0.0%)	0.80	2/19102 (0.0%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	91	TRP	CD2-CE2	7.88	1.50	1.41
1	G	91	TRP	CD2-CE2	5.68	1.48	1.41
1	B	91	TRP	CD2-CE2	5.64	1.48	1.41
1	A	91	TRP	CD2-CE2	5.18	1.47	1.41

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	97	ARG	NE-CZ-NH1	5.28	122.94	120.30
1	D	141	LEU	CA-CB-CG	-5.05	103.68	115.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit,

and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1741	0	1761	42	0
1	B	1733	0	1747	41	0
1	C	1741	0	1762	82	0
1	D	1767	0	1783	127	0
1	E	1743	0	1763	30	0
1	F	1734	0	1755	73	0
1	G	1737	0	1758	31	0
1	H	1717	0	1741	81	0
2	A	6	0	0	0	0
2	B	3	0	0	0	0
2	C	2	0	0	0	0
2	D	1	0	0	0	0
2	E	2	0	0	0	0
2	G	3	0	0	0	0
2	H	3	0	0	0	0
All	All	13933	0	14070	472	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 17.

The worst 5 of 472 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:F:125:ILE:CG2	1:F:126:GLU:N	1.98	1.23
1:F:125:ILE:CG2	1:F:126:GLU:H	1.53	1.19
1:D:215:ARG:HH11	1:D:215:ARG:HG2	1.01	1.12
1:F:54:ILE:HB	1:F:55:PRO:HD3	1.32	1.11
1:F:125:ILE:HG23	1:F:126:GLU:H	1.11	1.11

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	209/240 (87%)	197 (94%)	8 (4%)	4 (2%)	12	51
1	B	208/240 (87%)	195 (94%)	10 (5%)	3 (1%)	16	60
1	C	210/240 (88%)	183 (87%)	20 (10%)	7 (3%)	6	32
1	D	212/240 (88%)	172 (81%)	29 (14%)	11 (5%)	3	18
1	E	209/240 (87%)	192 (92%)	13 (6%)	4 (2%)	12	51
1	F	208/240 (87%)	183 (88%)	20 (10%)	5 (2%)	9	42
1	G	208/240 (87%)	192 (92%)	12 (6%)	4 (2%)	12	51
1	H	206/240 (86%)	188 (91%)	13 (6%)	5 (2%)	9	42
All	All	1670/1920 (87%)	1502 (90%)	125 (8%)	43 (3%)	8	39

5 of 43 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	136	ASP
1	D	54	ILE
1	D	214	ILE
1	F	55	PRO
1	F	216	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	192/212 (91%)	165 (86%)	27 (14%)	5	23
1	B	191/212 (90%)	168 (88%)	23 (12%)	7	30
1	C	192/212 (91%)	165 (86%)	27 (14%)	5	23
1	D	195/212 (92%)	154 (79%)	41 (21%)	1	8
1	E	193/212 (91%)	171 (89%)	22 (11%)	8	33
1	F	192/212 (91%)	155 (81%)	37 (19%)	2	12
1	G	192/212 (91%)	166 (86%)	26 (14%)	6	24
1	H	190/212 (90%)	162 (85%)	28 (15%)	4	21
All	All	1537/1696 (91%)	1306 (85%)	231 (15%)	4	20

5 of 231 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	185	ASN
1	E	112	THR
1	H	95	LEU
1	D	192	ARG
1	D	215	ARG

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	D	134	HIS
1	D	169	HIS
1	F	185	ASN
1	D	83	GLN
1	H	83	GLN

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

There are no ligands in this entry.

## 5.7 Other polymers ⓘ

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 ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

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	213/240 (88%)	-0.37	0 100 100	32, 68, 107, 152	0
1	B	212/240 (88%)	-0.35	0 100 100	31, 70, 112, 155	0
1	C	214/240 (89%)	-0.33	1 (0%) 88 36	34, 72, 120, 194	0
1	D	216/240 (90%)	-0.37	0 100 100	40, 73, 110, 146	0
1	E	213/240 (88%)	-0.33	0 100 100	53, 82, 119, 163	0
1	F	212/240 (88%)	-0.27	0 100 100	49, 81, 113, 159	0
1	G	212/240 (88%)	-0.24	0 100 100	54, 86, 130, 163	0
1	H	210/240 (87%)	-0.33	1 (0%) 88 36	40, 77, 122, 177	0
All	All	1702/1920 (88%)	-0.32	2 (0%) 93 63	31, 76, 120, 194	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	129	ARG	3.7
1	H	129	ARG	3.2

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

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

### 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

### 6.4 Ligands ⓘ

There are no ligands in this entry.



## 6.5 Other polymers ⓘ

There are no such residues in this entry.