



# wwPDB X-ray Structure Validation Summary Report

Feb 27, 2014 – 08:56 PM GMT

PDB ID : 2Q4U  
Title : Ensemble refinement of the crystal structure of an EF-hand protein from Danio rerio Dr.36843  
Authors : Levin, E.J.; Kondrashov, D.A.; Wesenberg, G.E.; Phillips Jr., G.N.; Center for Eukaryotic Structural Genomics (CESG)  
Deposited on : 2007-05-31  
Resolution : 2.10 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>

---

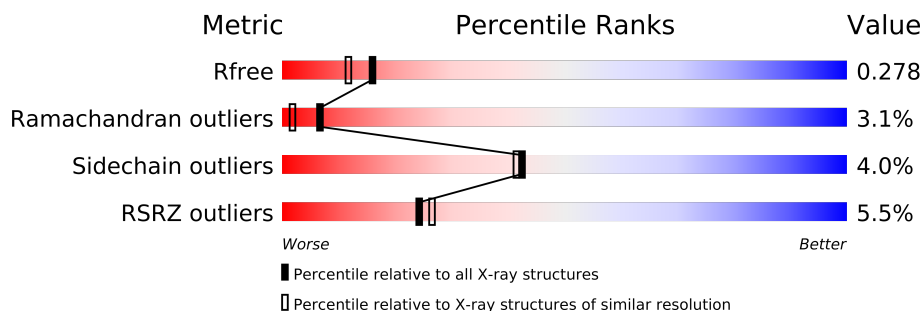
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
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) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.10 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	3012 (2.10-2.10)
Ramachandran outliers	78287	3610 (2.10-2.10)
Sidechain outliers	78261	3611 (2.10-2.10)
RSRZ outliers	66119	3013 (2.10-2.10)

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	1-A	272	
1	10-A	272	
1	11-A	272	
1	12-A	272	
1	13-A	272	
1	14-A	272	
1	15-A	272	
1	16-A	272	
1	2-A	272	
1	3-A	272	
1	4-A	272	
1	5-A	272	
1	6-A	272	
1	7-A	272	
1	8-A	272	
1	9-A	272	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 39088 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 Protein Zgc:100843.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	1-A	270	Total	C	N	O	S	Se	0	0	0
			2203	1402	374	414	3	10			
1	2-A	270	Total	C	N	O	S	Se	0	0	0
			2203	1402	374	414	3	10			
1	3-A	270	Total	C	N	O	S	Se	0	0	0
			2203	1402	374	414	3	10			
1	4-A	270	Total	C	N	O	S	Se	0	0	0
			2203	1402	374	414	3	10			
1	5-A	270	Total	C	N	O	S	Se	0	0	0
			2203	1402	374	414	3	10			
1	6-A	270	Total	C	N	O	S	Se	0	0	0
			2203	1402	374	414	3	10			
1	7-A	270	Total	C	N	O	S	Se	0	0	0
			2203	1402	374	414	3	10			
1	8-A	270	Total	C	N	O	S	Se	0	0	0
			2203	1402	374	414	3	10			
1	9-A	270	Total	C	N	O	S	Se	0	0	0
			2203	1402	374	414	3	10			
1	10-A	270	Total	C	N	O	S	Se	0	0	0
			2203	1402	374	414	3	10			
1	11-A	270	Total	C	N	O	S	Se	0	0	0
			2203	1402	374	414	3	10			
1	12-A	270	Total	C	N	O	S	Se	0	0	0
			2203	1402	374	414	3	10			
1	13-A	270	Total	C	N	O	S	Se	0	0	0
			2203	1402	374	414	3	10			
1	14-A	270	Total	C	N	O	S	Se	0	0	0
			2203	1402	374	414	3	10			
1	15-A	270	Total	C	N	O	S	Se	0	0	0
			2203	1402	374	414	3	10			
1	16-A	270	Total	C	N	O	S	Se	0	0	0
			2203	1402	374	414	3	10			

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	SER	-	EXPRESSION TAG	UNP Q5XJX1
A	40	MSE	MET	MODIFIED RESIDUE	UNP Q5XJX1
A	63	MSE	MET	MODIFIED RESIDUE	UNP Q5XJX1
A	82	MSE	MET	MODIFIED RESIDUE	UNP Q5XJX1
A	107	MSE	MET	MODIFIED RESIDUE	UNP Q5XJX1
A	153	MSE	MET	MODIFIED RESIDUE	UNP Q5XJX1
A	154	MSE	MET	MODIFIED RESIDUE	UNP Q5XJX1
A	186	MSE	MET	MODIFIED RESIDUE	UNP Q5XJX1
A	225	MSE	MET	MODIFIED RESIDUE	UNP Q5XJX1
A	226	MSE	MET	MODIFIED RESIDUE	UNP Q5XJX1
A	251	MSE	MET	MODIFIED RESIDUE	UNP Q5XJX1

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	1-A	240	Total O 240 240	0	0
2	2-A	240	Total O 240 240	0	0
2	3-A	240	Total O 240 240	0	0
2	4-A	240	Total O 240 240	0	0
2	5-A	240	Total O 240 240	0	0
2	6-A	240	Total O 240 240	0	0
2	7-A	240	Total O 240 240	0	0
2	8-A	240	Total O 240 240	0	0
2	9-A	240	Total O 240 240	0	0
2	10-A	240	Total O 240 240	0	0
2	11-A	240	Total O 240 240	0	0
2	12-A	240	Total O 240 240	0	0
2	13-A	240	Total O 240 240	0	0

*Continued on next page...*

*Continued from previous page...*

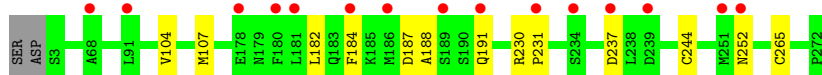
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	14-A	240	Total 240	O 240	0	0
2	15-A	240	Total 240	O 240	0	0
2	16-A	240	Total 240	O 240	0	0

### 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 errors 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.

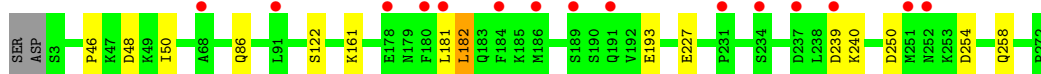
- Molecule 1: Protein Zgc:100843

Chain 1-A: 



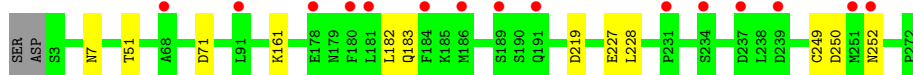
- Molecule 1: Protein Zgc:100843

Chain 2-A: 



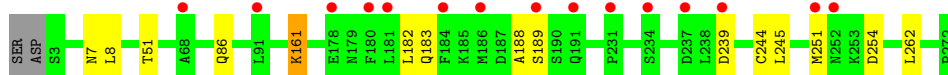
- Molecule 1: Protein Zgc:100843

Chain 3-A: 



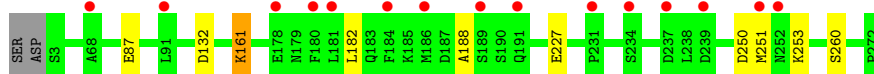
- Molecule 1: Protein Zgc:100843

Chain 4-A: 



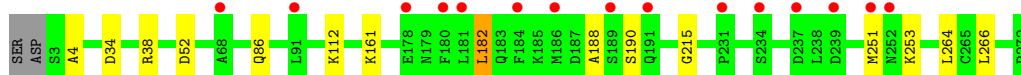
- Molecule 1: Protein Zgc:100843

Chain 5-A: 

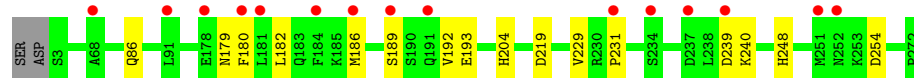


- Molecule 1: Protein Zgc:100843

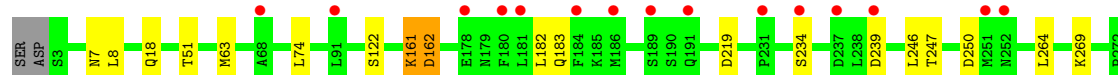
Chain 6-A: 



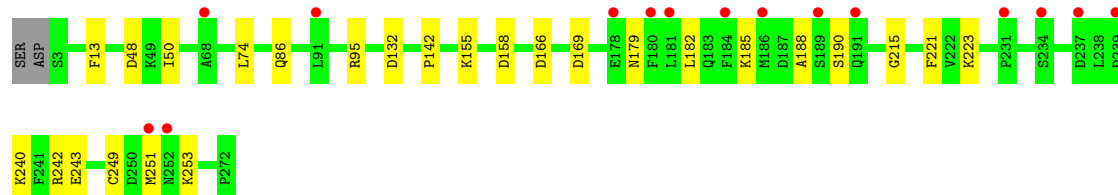
## • Molecule 1: Protein Zgc:100843

Chain 7-A: 

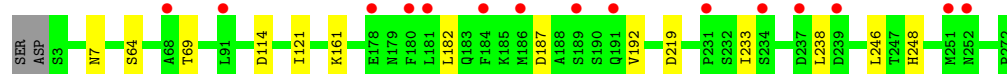
## • Molecule 1: Protein Zgc:100843

Chain 8-A: 

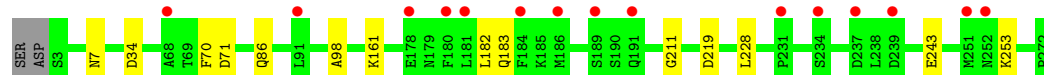
## • Molecule 1: Protein Zgc:100843

Chain 9-A: 

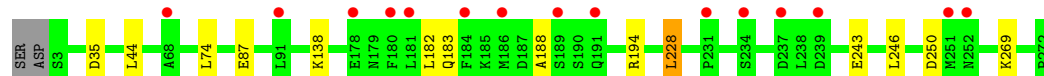
## • Molecule 1: Protein Zgc:100843

Chain 10-A: 

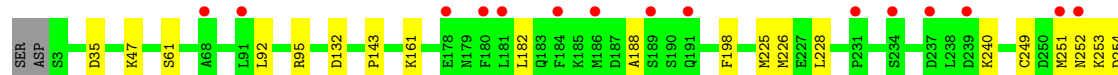
## • Molecule 1: Protein Zgc:100843

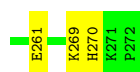
Chain 11-A: 

## • Molecule 1: Protein Zgc:100843

Chain 12-A: 

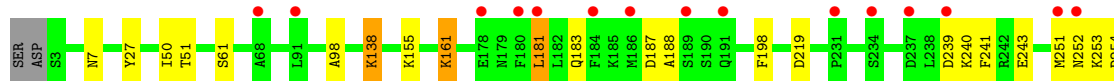
## • Molecule 1: Protein Zgc:100843

Chain 13-A: 



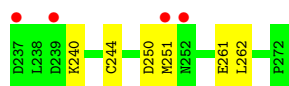
• Molecule 1: Protein Zgc:100843

Chain 14-A:



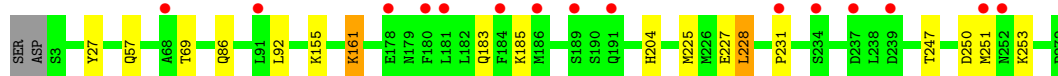
• Molecule 1: Protein Zgc:100843

Chain 15-A:



• Molecule 1: Protein Zgc:100843

Chain 16-A:





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	47.88Å 52.75Å 114.36Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.17 – 2.10 44.17 – 2.10	Depositor EDS
% Data completeness (in resolution range)	98.7 (44.17-2.10) 98.8 (44.17-2.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.35 (at 2.10Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.154 , 0.242 0.189 , 0.278	Depositor DCC
$R_{free}$ test set	885 reflections (5.38%)	DCC
Wilson B-factor (Å <sup>2</sup> )	35.8	Xtriage
Anisotropy	0.172	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 172.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 17336 reflections	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	39088	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.07% 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	1-A	0.50	0/2235	0.64	0/2976
1	2-A	0.50	0/2235	0.62	0/2976
1	3-A	0.50	0/2235	0.64	0/2976
1	4-A	0.50	0/2235	0.63	0/2976
1	5-A	0.50	0/2235	0.62	0/2976
1	6-A	0.50	0/2235	0.65	0/2976
1	7-A	0.50	0/2235	0.66	0/2976
1	8-A	0.51	0/2235	0.65	0/2976
1	9-A	0.51	0/2235	0.65	0/2976
1	10-A	0.48	0/2235	0.63	0/2976
1	11-A	0.51	0/2235	0.65	1/2976 (0.0%)
1	12-A	0.51	0/2235	0.65	1/2976 (0.0%)
1	13-A	0.55	0/2235	0.73	1/2976 (0.0%)
1	14-A	0.57	0/2235	0.75	0/2976
1	15-A	0.58	0/2235	0.73	1/2976 (0.0%)
1	16-A	0.56	0/2235	0.74	0/2976
All	All	0.52	0/35760	0.67	4/47616 (0.0%)

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	16-A	0	1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	15-A	228	LEU	CA-CB-CG	5.61	128.21	115.30
1	12-A	228	LEU	CA-CB-CG	5.32	127.53	115.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	13-A	228	LEU	CA-CB-CG	5.31	127.51	115.30
1	11-A	228	LEU	CA-CB-CG	5.18	127.22	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	16-A	27	TYR	Sidechain

## 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	1-A	2203	0	2186	0	0
1	2-A	2203	0	2186	0	0
1	3-A	2203	0	2186	0	0
1	4-A	2203	0	2186	0	0
1	5-A	2203	0	2186	0	0
1	6-A	2203	0	2186	0	0
1	7-A	2203	0	2186	0	0
1	8-A	2203	0	2186	0	0
1	9-A	2203	0	2186	0	0
1	10-A	2203	0	2186	0	0
1	11-A	2203	0	2186	0	0
1	12-A	2203	0	2186	0	0
1	13-A	2203	0	2186	0	0
1	14-A	2203	0	2186	0	0
1	15-A	2203	0	2186	0	0
1	16-A	2203	0	2186	0	0
2	1-A	240	0	0	0	0
2	2-A	240	0	0	0	0
2	3-A	240	0	0	0	0
2	4-A	240	0	0	0	0
2	5-A	240	0	0	0	0
2	6-A	240	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	7-A	240	0	0	0	0
2	8-A	240	0	0	0	0
2	9-A	240	0	0	0	0
2	10-A	240	0	0	0	0
2	11-A	240	0	0	0	0
2	12-A	240	0	0	0	0
2	13-A	240	0	0	0	0
2	14-A	240	0	0	0	0
2	15-A	240	0	0	0	0
2	16-A	240	0	0	0	0
All	All	39088	0	34976	0	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. Clashscore could not be calculated for this entry.

There are no clashes within the asymmetric unit.

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	1-A	268/272 (98%)	234 (87%)	30 (11%)	4 (2%)	15	8
1	2-A	268/272 (98%)	243 (91%)	16 (6%)	9 (3%)	6	1
1	3-A	268/272 (98%)	239 (89%)	24 (9%)	5 (2%)	12	5
1	4-A	268/272 (98%)	245 (91%)	17 (6%)	6 (2%)	10	4
1	5-A	268/272 (98%)	244 (91%)	18 (7%)	6 (2%)	10	4
1	6-A	268/272 (98%)	236 (88%)	22 (8%)	10 (4%)	5	1
1	7-A	268/272 (98%)	225 (84%)	35 (13%)	8 (3%)	7	2
1	8-A	268/272 (98%)	240 (90%)	20 (8%)	8 (3%)	7	2
1	9-A	268/272 (98%)	227 (85%)	28 (10%)	13 (5%)	3	1
1	10-A	268/272 (98%)	228 (85%)	32 (12%)	8 (3%)	7	2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	11-A	268/272 (98%)	238 (89%)	22 (8%)	8 (3%)	7	2
1	12-A	268/272 (98%)	238 (89%)	25 (9%)	5 (2%)	12	5
1	13-A	268/272 (98%)	235 (88%)	23 (9%)	10 (4%)	5	1
1	14-A	268/272 (98%)	211 (79%)	44 (16%)	13 (5%)	3	1
1	15-A	268/272 (98%)	225 (84%)	33 (12%)	10 (4%)	5	1
1	16-A	268/272 (98%)	235 (88%)	22 (8%)	11 (4%)	4	1
All	All	4288/4352 (98%)	3743 (87%)	411 (10%)	134 (3%)	7	2

5 of 134 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	1-A	182	LEU
1	1-A	188	ALA
1	2-A	50	ILE
1	2-A	227	GLU
1	2-A	250	ASP

### 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	1-A	238/230 (104%)	229 (96%)	9 (4%)	44	44
1	2-A	238/230 (104%)	231 (97%)	7 (3%)	55	57
1	3-A	238/230 (104%)	231 (97%)	7 (3%)	55	57
1	4-A	238/230 (104%)	228 (96%)	10 (4%)	40	38
1	5-A	238/230 (104%)	233 (98%)	5 (2%)	66	70
1	6-A	238/230 (104%)	232 (98%)	6 (2%)	60	63
1	7-A	238/230 (104%)	230 (97%)	8 (3%)	49	49
1	8-A	238/230 (104%)	225 (94%)	13 (6%)	30	26
1	9-A	238/230 (104%)	225 (94%)	13 (6%)	30	26
1	10-A	238/230 (104%)	232 (98%)	6 (2%)	60	63
1	11-A	238/230 (104%)	233 (98%)	5 (2%)	66	70

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	12-A	238/230 (104%)	229 (96%)	9 (4%)	44	44
1	13-A	238/230 (104%)	226 (95%)	12 (5%)	34	30
1	14-A	238/230 (104%)	223 (94%)	15 (6%)	25	21
1	15-A	238/230 (104%)	218 (92%)	20 (8%)	16	10
1	16-A	238/230 (104%)	230 (97%)	8 (3%)	49	49
All	All	3808/3680 (104%)	3655 (96%)	153 (4%)	42	41

5 of 153 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	9-A	132	ASP
1	11-A	34	ASP
1	15-A	244	CYS
1	9-A	166	ASP
1	9-A	251	MSE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 81 such sidechains are listed below:

Mol	Chain	Res	Type
1	8-A	252	ASN
1	10-A	56	GLN
1	15-A	252	ASN
1	8-A	270	HIS
1	9-A	191	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	1-A	270/272 (99%)	0.10	15 (5%)	24	26	13, 29, 53, 69	270 (100%)
1	2-A	270/272 (99%)	0.10	15 (5%)	24	26	13, 29, 53, 69	270 (100%)
1	3-A	270/272 (99%)	0.10	15 (5%)	24	26	13, 29, 53, 69	270 (100%)
1	4-A	270/272 (99%)	0.10	15 (5%)	24	26	13, 29, 53, 69	270 (100%)
1	5-A	270/272 (99%)	0.10	15 (5%)	24	26	13, 29, 53, 69	270 (100%)
1	6-A	270/272 (99%)	0.10	15 (5%)	24	26	13, 29, 53, 69	270 (100%)
1	7-A	270/272 (99%)	0.10	15 (5%)	24	26	13, 29, 53, 69	270 (100%)
1	8-A	270/272 (99%)	0.10	15 (5%)	24	26	13, 29, 53, 69	270 (100%)
1	9-A	270/272 (99%)	0.10	15 (5%)	24	26	13, 29, 53, 69	270 (100%)
1	10-A	270/272 (99%)	0.10	15 (5%)	24	26	13, 29, 53, 69	270 (100%)
1	11-A	270/272 (99%)	0.10	15 (5%)	24	26	13, 29, 53, 69	270 (100%)
1	12-A	270/272 (99%)	0.10	15 (5%)	24	26	13, 29, 53, 69	270 (100%)
1	13-A	270/272 (99%)	0.10	15 (5%)	24	26	13, 29, 53, 69	270 (100%)
1	14-A	270/272 (99%)	0.10	15 (5%)	24	26	13, 29, 53, 69	270 (100%)
1	15-A	270/272 (99%)	0.10	15 (5%)	24	26	13, 29, 53, 69	270 (100%)
1	16-A	270/272 (99%)	0.10	15 (5%)	24	26	13, 29, 53, 69	270 (100%)
All	All	4320/4352 (99%)	0.10	240 (5%)	24	26	13, 29, 53, 69	4320 (100%)

The worst 5 of 240 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	1-A	186	MSE	5.8
1	2-A	186	MSE	5.8
1	3-A	186	MSE	5.8
1	4-A	186	MSE	5.8
1	5-A	186	MSE	5.8



## 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.