



# wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 11, 2016 – 03:38 PM EDT

PDB ID : 4ZWQ  
Title : Crystal Structure of the Bacteriophage T4 recombination mediator protein  
UvsY, Lattice Type I  
Authors : Gajewski, S.; White, S.W.  
Deposited on : 2015-05-19  
Resolution : 2.35 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  
<http://wwpdb.org/validation/2016/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 : unknown  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027107  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20027107

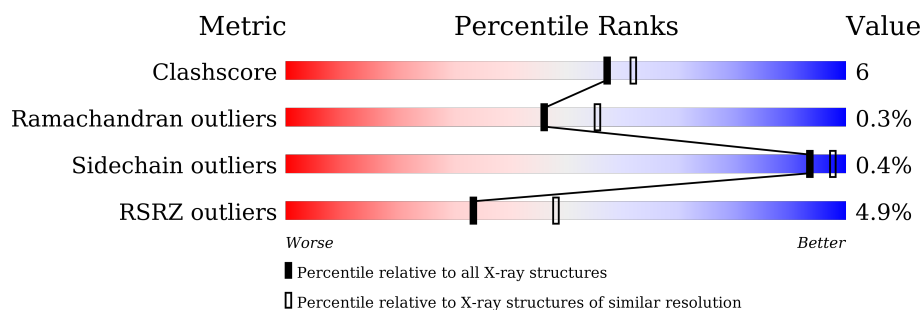
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.35 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))
Clashscore	102246	1456 (2.38-2.34)
Ramachandran outliers	100387	1435 (2.38-2.34)
Sidechain outliers	100360	1436 (2.38-2.34)
RSRZ outliers	91569	1358 (2.38-2.34)

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	157	<div> <div>7%</div> <div>71%15%13%</div> </div>
1	B	157	<div> <div>%</div> <div>68%18%14%</div> </div>
1	C	157	<div> <div>%</div> <div>70%15%15%</div> </div>
1	D	157	<div> <div>%</div> <div>74%12%14%</div> </div>
1	E	157	<div> <div>%</div> <div>73%13%14%</div> </div>
1	F	157	<div> <div>2%</div> <div>73%13%15%</div> </div>
1	G	157	<div> <div>17%</div> <div>74%10%16%</div> </div>

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 7494 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 Recombination protein uvsY.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	136	Total	C	N	O	S	0	0	0
			1084	687	184	206	7			
1	B	135	Total	C	N	O	S	0	0	0
			1083	689	185	202	7			
1	C	134	Total	C	N	O	S	0	0	0
			1072	681	181	204	6			
1	D	135	Total	C	N	O	S	0	0	0
			1084	689	183	205	7			
1	E	135	Total	C	N	O	S	0	0	0
			1077	686	183	201	7			
1	F	134	Total	C	N	O	S	0	0	0
			1068	678	180	204	6			
1	G	132	Total	C	N	O	S	0	0	0
			1026	649	174	198	5			

There are 140 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	expression tag	UNP P04537
A	-18	GLY	-	expression tag	UNP P04537
A	-17	SER	-	expression tag	UNP P04537
A	-16	SER	-	expression tag	UNP P04537
A	-15	HIS	-	expression tag	UNP P04537
A	-14	HIS	-	expression tag	UNP P04537
A	-13	HIS	-	expression tag	UNP P04537
A	-12	HIS	-	expression tag	UNP P04537
A	-11	HIS	-	expression tag	UNP P04537
A	-10	HIS	-	expression tag	UNP P04537
A	-9	SER	-	expression tag	UNP P04537
A	-8	SER	-	expression tag	UNP P04537
A	-7	GLY	-	expression tag	UNP P04537
A	-6	LEU	-	expression tag	UNP P04537
A	-5	VAL	-	expression tag	UNP P04537

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	PRO	-	expression tag	UNP P04537
A	-3	ARG	-	expression tag	UNP P04537
A	-2	GLY	-	expression tag	UNP P04537
A	-1	SER	-	expression tag	UNP P04537
A	0	HIS	-	expression tag	UNP P04537
B	-19	MET	-	expression tag	UNP P04537
B	-18	GLY	-	expression tag	UNP P04537
B	-17	SER	-	expression tag	UNP P04537
B	-16	SER	-	expression tag	UNP P04537
B	-15	HIS	-	expression tag	UNP P04537
B	-14	HIS	-	expression tag	UNP P04537
B	-13	HIS	-	expression tag	UNP P04537
B	-12	HIS	-	expression tag	UNP P04537
B	-11	HIS	-	expression tag	UNP P04537
B	-10	HIS	-	expression tag	UNP P04537
B	-9	SER	-	expression tag	UNP P04537
B	-8	SER	-	expression tag	UNP P04537
B	-7	GLY	-	expression tag	UNP P04537
B	-6	LEU	-	expression tag	UNP P04537
B	-5	VAL	-	expression tag	UNP P04537
B	-4	PRO	-	expression tag	UNP P04537
B	-3	ARG	-	expression tag	UNP P04537
B	-2	GLY	-	expression tag	UNP P04537
B	-1	SER	-	expression tag	UNP P04537
B	0	HIS	-	expression tag	UNP P04537
C	-19	MET	-	expression tag	UNP P04537
C	-18	GLY	-	expression tag	UNP P04537
C	-17	SER	-	expression tag	UNP P04537
C	-16	SER	-	expression tag	UNP P04537
C	-15	HIS	-	expression tag	UNP P04537
C	-14	HIS	-	expression tag	UNP P04537
C	-13	HIS	-	expression tag	UNP P04537
C	-12	HIS	-	expression tag	UNP P04537
C	-11	HIS	-	expression tag	UNP P04537
C	-10	HIS	-	expression tag	UNP P04537
C	-9	SER	-	expression tag	UNP P04537
C	-8	SER	-	expression tag	UNP P04537
C	-7	GLY	-	expression tag	UNP P04537
C	-6	LEU	-	expression tag	UNP P04537
C	-5	VAL	-	expression tag	UNP P04537
C	-4	PRO	-	expression tag	UNP P04537
C	-3	ARG	-	expression tag	UNP P04537

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	-2	GLY	-	expression tag	UNP P04537
C	-1	SER	-	expression tag	UNP P04537
C	0	HIS	-	expression tag	UNP P04537
D	-19	MET	-	expression tag	UNP P04537
D	-18	GLY	-	expression tag	UNP P04537
D	-17	SER	-	expression tag	UNP P04537
D	-16	SER	-	expression tag	UNP P04537
D	-15	HIS	-	expression tag	UNP P04537
D	-14	HIS	-	expression tag	UNP P04537
D	-13	HIS	-	expression tag	UNP P04537
D	-12	HIS	-	expression tag	UNP P04537
D	-11	HIS	-	expression tag	UNP P04537
D	-10	HIS	-	expression tag	UNP P04537
D	-9	SER	-	expression tag	UNP P04537
D	-8	SER	-	expression tag	UNP P04537
D	-7	GLY	-	expression tag	UNP P04537
D	-6	LEU	-	expression tag	UNP P04537
D	-5	VAL	-	expression tag	UNP P04537
D	-4	PRO	-	expression tag	UNP P04537
D	-3	ARG	-	expression tag	UNP P04537
D	-2	GLY	-	expression tag	UNP P04537
D	-1	SER	-	expression tag	UNP P04537
D	0	HIS	-	expression tag	UNP P04537
E	-19	MET	-	expression tag	UNP P04537
E	-18	GLY	-	expression tag	UNP P04537
E	-17	SER	-	expression tag	UNP P04537
E	-16	SER	-	expression tag	UNP P04537
E	-15	HIS	-	expression tag	UNP P04537
E	-14	HIS	-	expression tag	UNP P04537
E	-13	HIS	-	expression tag	UNP P04537
E	-12	HIS	-	expression tag	UNP P04537
E	-11	HIS	-	expression tag	UNP P04537
E	-10	HIS	-	expression tag	UNP P04537
E	-9	SER	-	expression tag	UNP P04537
E	-8	SER	-	expression tag	UNP P04537
E	-7	GLY	-	expression tag	UNP P04537
E	-6	LEU	-	expression tag	UNP P04537
E	-5	VAL	-	expression tag	UNP P04537
E	-4	PRO	-	expression tag	UNP P04537
E	-3	ARG	-	expression tag	UNP P04537
E	-2	GLY	-	expression tag	UNP P04537
E	-1	SER	-	expression tag	UNP P04537

*Continued on next page...*

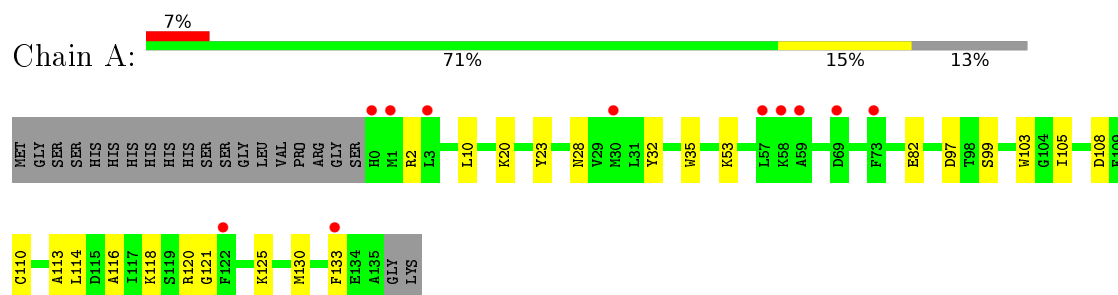
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	0	HIS	-	expression tag	UNP P04537
F	-19	MET	-	expression tag	UNP P04537
F	-18	GLY	-	expression tag	UNP P04537
F	-17	SER	-	expression tag	UNP P04537
F	-16	SER	-	expression tag	UNP P04537
F	-15	HIS	-	expression tag	UNP P04537
F	-14	HIS	-	expression tag	UNP P04537
F	-13	HIS	-	expression tag	UNP P04537
F	-12	HIS	-	expression tag	UNP P04537
F	-11	HIS	-	expression tag	UNP P04537
F	-10	HIS	-	expression tag	UNP P04537
F	-9	SER	-	expression tag	UNP P04537
F	-8	SER	-	expression tag	UNP P04537
F	-7	GLY	-	expression tag	UNP P04537
F	-6	LEU	-	expression tag	UNP P04537
F	-5	VAL	-	expression tag	UNP P04537
F	-4	PRO	-	expression tag	UNP P04537
F	-3	ARG	-	expression tag	UNP P04537
F	-2	GLY	-	expression tag	UNP P04537
F	-1	SER	-	expression tag	UNP P04537
F	0	HIS	-	expression tag	UNP P04537
G	-19	MET	-	expression tag	UNP P04537
G	-18	GLY	-	expression tag	UNP P04537
G	-17	SER	-	expression tag	UNP P04537
G	-16	SER	-	expression tag	UNP P04537
G	-15	HIS	-	expression tag	UNP P04537
G	-14	HIS	-	expression tag	UNP P04537
G	-13	HIS	-	expression tag	UNP P04537
G	-12	HIS	-	expression tag	UNP P04537
G	-11	HIS	-	expression tag	UNP P04537
G	-10	HIS	-	expression tag	UNP P04537
G	-9	SER	-	expression tag	UNP P04537
G	-8	SER	-	expression tag	UNP P04537
G	-7	GLY	-	expression tag	UNP P04537
G	-6	LEU	-	expression tag	UNP P04537
G	-5	VAL	-	expression tag	UNP P04537
G	-4	PRO	-	expression tag	UNP P04537
G	-3	ARG	-	expression tag	UNP P04537
G	-2	GLY	-	expression tag	UNP P04537
G	-1	SER	-	expression tag	UNP P04537
G	0	HIS	-	expression tag	UNP P04537

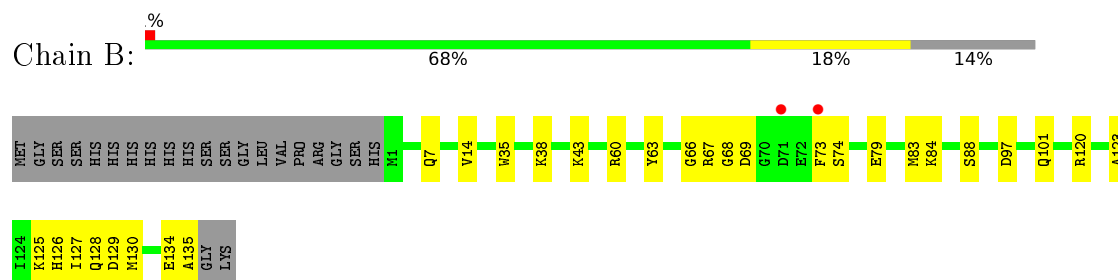
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

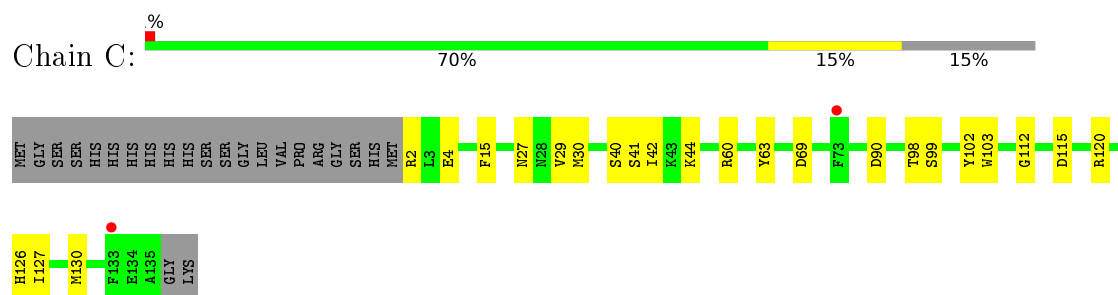
#### • Molecule 1: Recombination protein uvsY



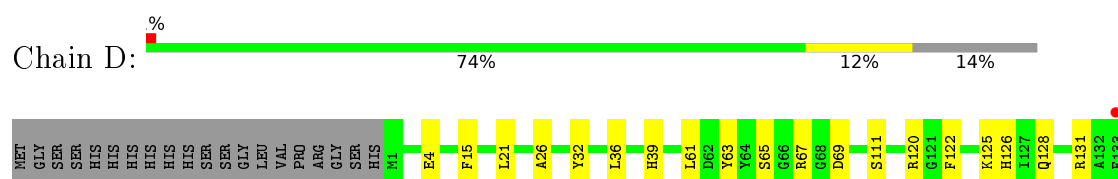
#### • Molecule 1: Recombination protein uvsY



#### • Molecule 1: Recombination protein uvsY



#### • Molecule 1: Recombination protein uvsY



E134

A135

GLY

LYS

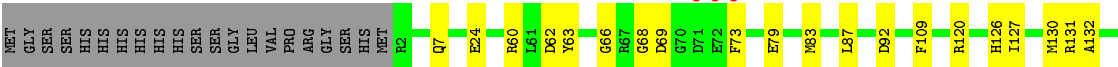
• Molecule 1: Recombination protein uvsY



GLY

LYS

• Molecule 1: Recombination protein uvsY

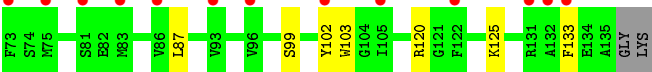
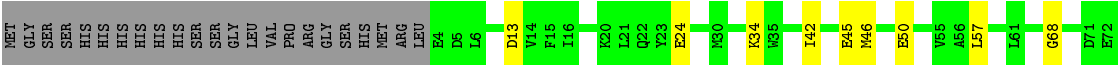
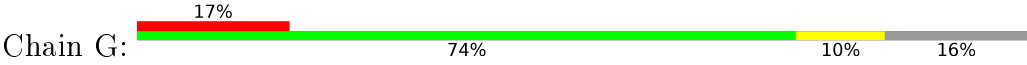


A135

GLY

LYS

• Molecule 1: Recombination protein uvsY





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	109.26Å 109.22Å 270.86Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.73 – 2.35 46.73 – 2.31	Depositor EDS
% Data completeness (in resolution range)	64.2 (46.73-2.35) 60.8 (46.73-2.31)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	9.83 (at 2.32Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.222 , 0.254 (Not available) , (Not available)	Depositor DCC
$R_{free}$ test set	NotAvailable	DCC
Wilson B-factor (Å <sup>2</sup> )	29.0	Xtriage
Anisotropy	0.088	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 38.0	EDS
Estimated twinning fraction	0.500 for -k,-h,-l 0.477 for -k,-h,-l	Xtriage
Reported twinning fraction	0.500 for -k,-h,-l	Depositor
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	10 of 43526 reflections (0.023%)	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	7494	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 40.97 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 2.5299e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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	0.35	0/1100	0.58	0/1473
1	B	0.40	0/1099	0.57	0/1468
1	C	0.44	0/1088	0.62	0/1456
1	D	0.41	0/1100	0.61	1/1470 (0.1%)
1	E	0.37	0/1093	0.58	0/1461
1	F	0.37	0/1084	0.57	0/1452
1	G	0.29	0/1042	0.51	0/1402
All	All	0.38	0/7606	0.58	1/10182 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	61	LEU	CA-CB-CG	5.51	127.98	115.30

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	1084	0	1073	15	0
1	B	1083	0	1094	20	0
1	C	1072	0	1068	18	0
1	D	1084	0	1091	15	0
1	E	1077	0	1085	12	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	F	1068	0	1057	14	0
1	G	1026	0	978	11	0
All	All	7494	0	7446	89	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 6.

The worst 5 of 89 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:63:TYR:HB2	1:F:69:ASP:HB2	1.74	0.68
1:A:105:ILE:HG12	1:B:43:LYS:HE2	1.75	0.68
1:A:130:MET:HG2	1:B:128:GLN:HG2	1.80	0.64
1:G:42:ILE:HG23	1:G:103:TRP:HE3	1.64	0.63
1:D:128:GLN:OE1	1:D:131:ARG:NH2	2.32	0.62

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	134/157 (85%)	133 (99%)	1 (1%)	0	100	100
1	B	133/157 (85%)	131 (98%)	1 (1%)	1 (1%)	24	26
1	C	132/157 (84%)	130 (98%)	2 (2%)	0	100	100
1	D	133/157 (85%)	131 (98%)	2 (2%)	0	100	100
1	E	133/157 (85%)	131 (98%)	1 (1%)	1 (1%)	24	26
1	F	132/157 (84%)	130 (98%)	2 (2%)	0	100	100
1	G	130/157 (83%)	128 (98%)	1 (1%)	1 (1%)	24	26
All	All	927/1099 (84%)	914 (99%)	10 (1%)	3 (0%)	46	55

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	68	GLY
1	B	68	GLY
1	E	68	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	113/136 (83%)	112 (99%)	1 (1%)	84	92
1	B	114/136 (84%)	114 (100%)	0	100	100
1	C	113/136 (83%)	112 (99%)	1 (1%)	84	92
1	D	115/136 (85%)	115 (100%)	0	100	100
1	E	113/136 (83%)	113 (100%)	0	100	100
1	F	112/136 (82%)	112 (100%)	0	100	100
1	G	103/136 (76%)	102 (99%)	1 (1%)	82	91
All	All	783/952 (82%)	780 (100%)	3 (0%)	93	97

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2	ARG
1	C	27	ASN
1	G	50	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	7	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

### 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 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 [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	136/157 (86%)	0.36	11 (8%) 15 23	27, 42, 64, 79	0
1	B	135/157 (85%)	-0.01	2 (1%) 76 85	19, 30, 64, 68	0
1	C	134/157 (85%)	-0.25	2 (1%) 76 85	14, 23, 42, 82	0
1	D	135/157 (85%)	-0.16	1 (0%) 89 94	12, 28, 46, 64	0
1	E	135/157 (85%)	-0.01	1 (0%) 89 94	19, 33, 56, 79	0
1	F	134/157 (85%)	0.29	3 (2%) 65 77	27, 40, 56, 67	0
1	G	132/157 (84%)	1.20	26 (19%) 1 2	45, 64, 82, 94	0
All	All	941/1099 (85%)	0.20	46 (4%) 33 48	12, 35, 68, 94	0

The worst 5 of 46 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	4	GLU	6.0
1	G	6	LEU	5.0
1	G	133	PHE	4.8
1	E	69	ASP	4.7
1	A	73	PHE	4.1

### 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

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.