



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 14, 2020 – 07:12 pm BST

PDB ID : 2R6C  
Title : Crystal Form BH2  
Authors : Bailey, S.; Eliason, W.K.; Steitz, T.A.  
Deposited on : 2007-09-05  
Resolution : 4.00 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

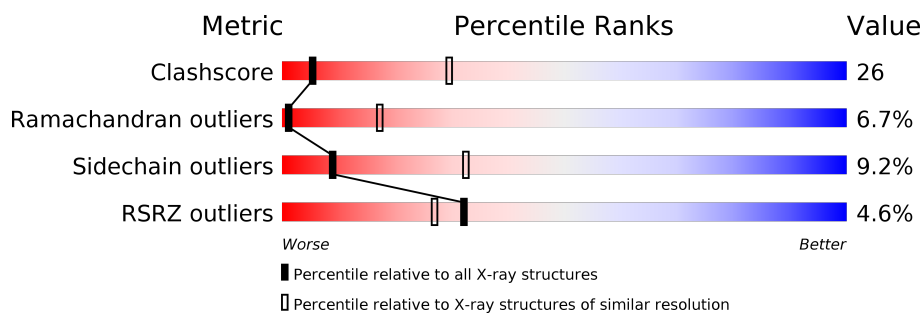
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 4.00 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))
Clashscore	141614	1148 (4.30-3.70)
Ramachandran outliers	138981	1108 (4.30-3.70)
Sidechain outliers	138945	1099 (4.30-3.70)
RSRZ outliers	127900	1028 (4.34-3.66)

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	454	<div> <div>3%</div> <div>46% 25% 7% • 21%</div> </div>
1	B	454	<div> <div>4%</div> <div>44% 27% 7% • 21%</div> </div>
1	C	454	<div> <div>2%</div> <div>45% 25% 8% • 21%</div> </div>
1	D	454	<div> <div>2%</div> <div>42% 28% 8% • 21%</div> </div>
1	E	454	<div> <div>4%</div> <div>46% 25% 8% • 21%</div> </div>
1	F	454	<div> <div>8%</div> <div>48% 24% 6% • 21%</div> </div>
2	G	143	<div> <div>•</div> <div>59% 29% 8% • •</div> </div>

*Continued on next page...*

Mol	Chain	Length	Quality of chain
2	H	143	
2	I	143	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 19920 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 Replicative helicase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	358	Total	C	N	O	S	0	0	0
			2759	1737	473	536	13			
1	B	358	Total	C	N	O	S	0	0	0
			2759	1737	473	536	13			
1	C	358	Total	C	N	O	S	0	0	0
			2759	1737	473	536	13			
1	D	358	Total	C	N	O	S	0	0	0
			2759	1737	473	536	13			
1	E	358	Total	C	N	O	S	0	0	0
			2759	1737	473	536	13			
1	F	358	Total	C	N	O	S	0	0	0
			2759	1737	473	536	13			

- Molecule 2 is a protein called DnaG Primase, Helicase Binding Domain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	G	138	Total	C	N	O	Se	0	0	0
			1122	712	199	205	6			
2	H	138	Total	C	N	O	Se	0	0	0
			1122	712	199	205	6			
2	I	138	Total	C	N	O	Se	0	0	0
			1122	712	199	205	6			

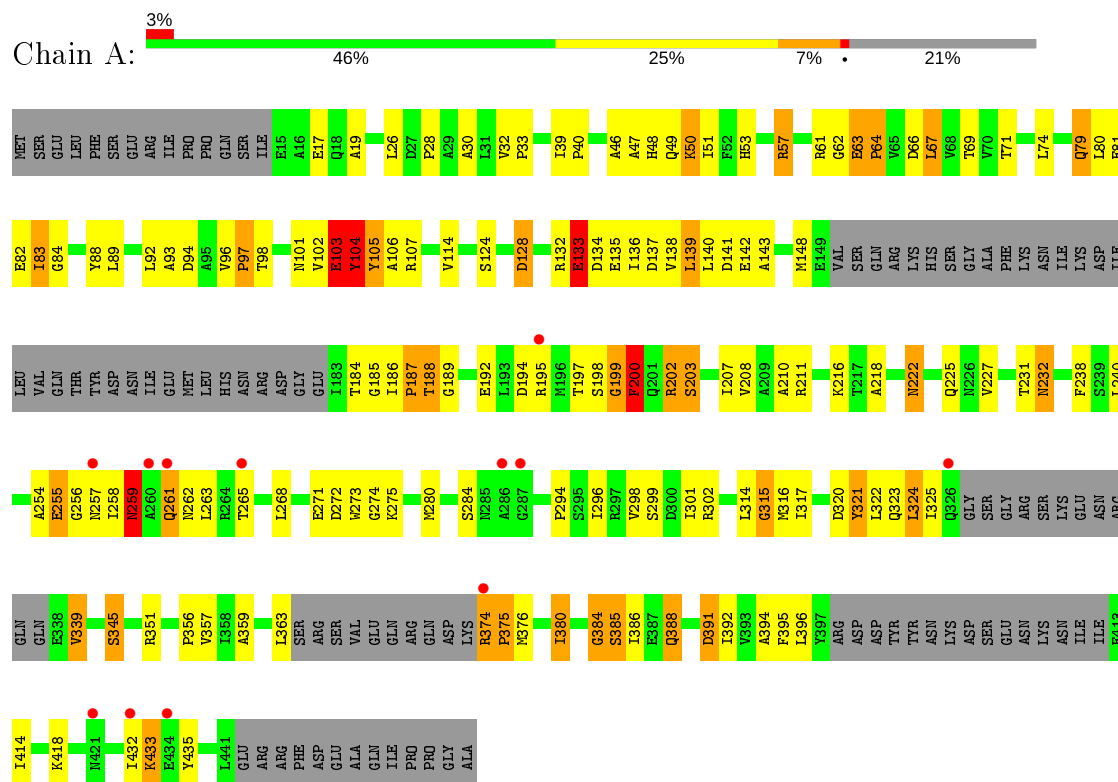
There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	530	GLU	ASP	CONFLICT	UNP Q9X4D0
G	531	LEU	VAL	CONFLICT	UNP Q9X4D0
H	530	GLU	ASP	CONFLICT	UNP Q9X4D0
H	531	LEU	VAL	CONFLICT	UNP Q9X4D0
I	530	GLU	ASP	CONFLICT	UNP Q9X4D0
I	531	LEU	VAL	CONFLICT	UNP Q9X4D0

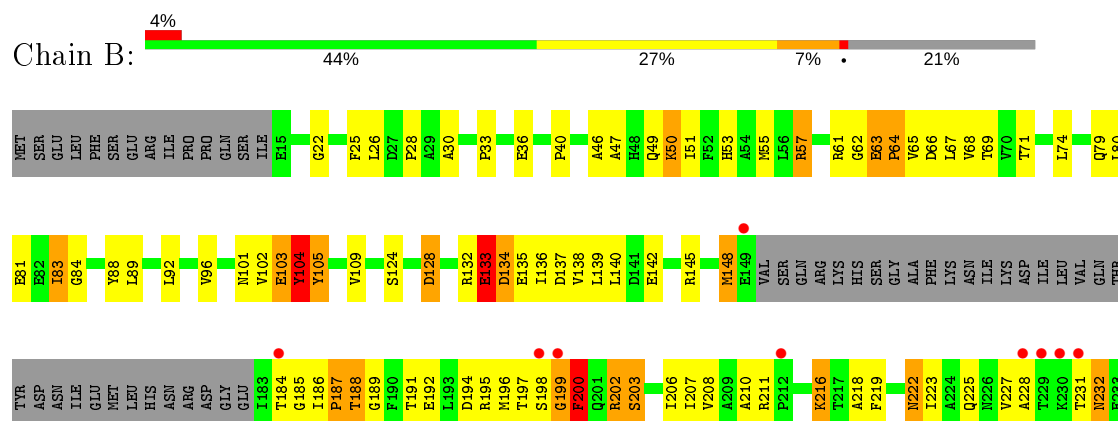
### 3 Residue-property plots

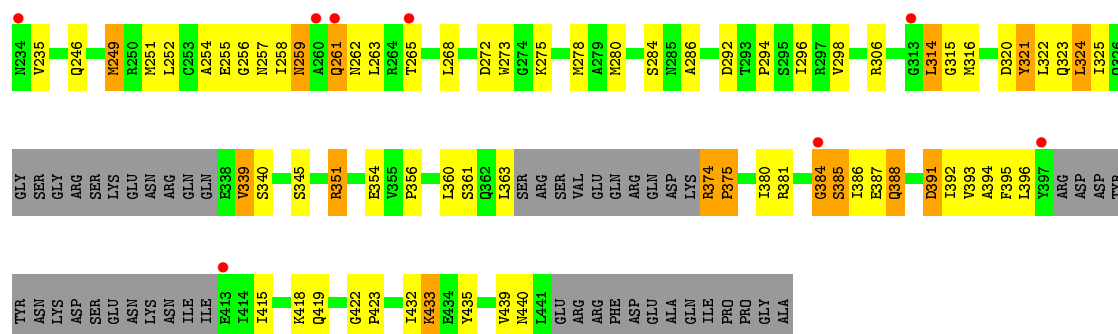
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: Replicative helicase

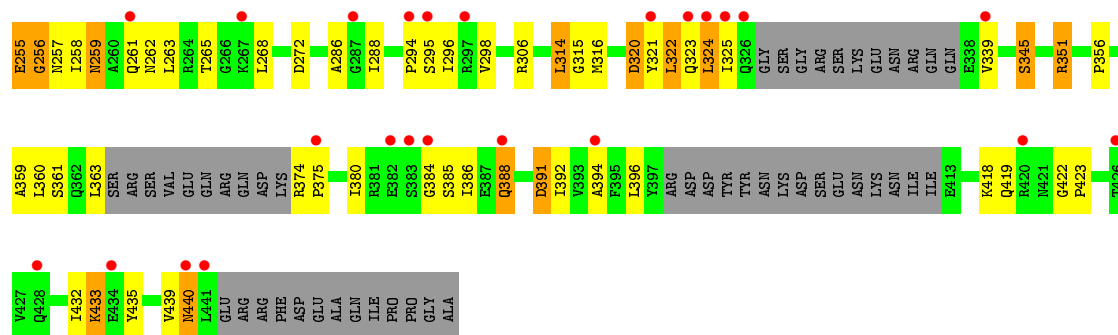


#### • Molecule 1: Replicative helicase

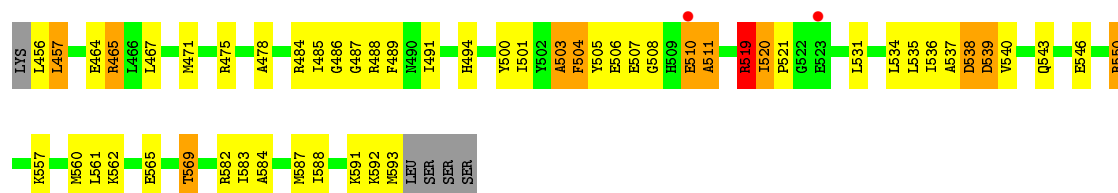




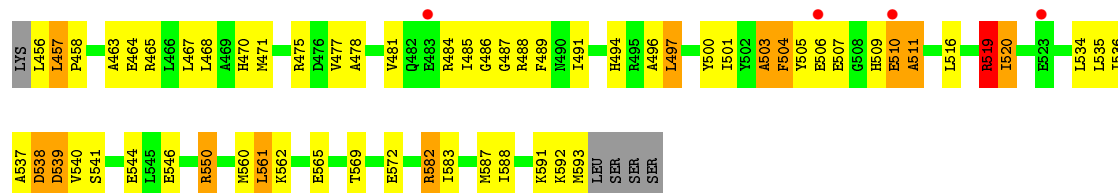




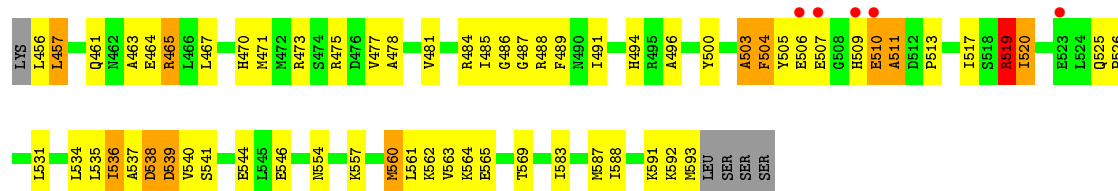
• Molecule 2: DnaG Primase, Helicase Binding Domain



• Molecule 2: DnaG Primase, Helicase Binding Domain



• Molecule 2: DnaG Primase, Helicase Binding Domain





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	228.78Å 228.78Å 192.98Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 4.00 19.98 – 4.00	Depositor EDS
% Data completeness (in resolution range)	90.8 (20.00-4.00) 90.8 (19.98-4.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.60 (at 4.07Å)	Xtriage
Refinement program	REFMAC 5.2	Depositor
R, $R_{free}$	0.320 , 0.344 0.310 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	127.3	Xtriage
Anisotropy	0.073	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.26 , 176.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	0.037 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.86	EDS
Total number of atoms	19920	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	166.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 15.51% 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

### 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	0.51	0/2789	0.79	7/3770 (0.2%)
1	B	0.47	0/2789	0.78	4/3770 (0.1%)
1	C	0.51	0/2789	0.81	7/3770 (0.2%)
1	D	0.51	0/2789	0.79	6/3770 (0.2%)
1	E	0.50	0/2789	0.80	8/3770 (0.2%)
1	F	0.47	0/2789	0.76	3/3770 (0.1%)
2	G	0.42	0/1134	0.70	1/1514 (0.1%)
2	H	0.47	0/1134	0.71	1/1514 (0.1%)
2	I	0.42	0/1134	0.69	1/1514 (0.1%)
All	All	0.49	0/20136	0.77	38/27162 (0.1%)

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	A	1	5
1	B	1	3
1	C	1	5
1	D	2	4
1	E	2	4
1	F	1	2
2	G	0	1
2	H	0	1
2	I	0	1
All	All	8	26

There are no bond length outliers.

The worst 5 of 38 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	519	ARG	N-CA-C	7.51	131.28	111.00

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	I	519	ARG	N-CA-C	7.41	131.01	111.00
2	G	519	ARG	N-CA-C	7.37	130.89	111.00
1	D	261	GLN	N-CA-C	6.80	129.35	111.00
1	A	261	GLN	N-CA-C	6.45	128.43	111.00

5 of 8 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	104	TYR	CA
1	B	104	TYR	CA
1	C	104	TYR	CA
1	D	104	TYR	CA
1	D	184	THR	CA

5 of 26 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	103	GLU	Peptide
1	A	187	PRO	Peptide
1	A	259	ASN	Peptide
1	A	320	ASP	Peptide
1	A	374	ARG	Peptide

## 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	2759	0	2806	160	0
1	B	2759	0	2806	151	0
1	C	2759	0	2805	163	0
1	D	2759	0	2807	187	0
1	E	2759	0	2806	162	0
1	F	2759	0	2807	124	0
2	G	1122	0	1144	62	0
2	H	1122	0	1144	61	0
2	I	1122	0	1144	56	0
All	All	19920	0	20269	1051	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 26.

The worst 5 of 1051 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:255:GLU:O	1:A:257:ASN:N	1.60	1.34
1:C:188:THR:OG1	1:C:194:ASP:OD1	1.52	1.27
1:C:306:ARG:CZ	1:D:32:VAL:HG21	1.64	1.26
1:D:255:GLU:O	1:D:257:ASN:N	1.70	1.23
1:A:132:ARG:O	1:A:133:GLU:HB2	1.40	1.20

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	348/454 (77%)	284 (82%)	41 (12%)	23 (7%)	1	17
1	B	348/454 (77%)	285 (82%)	40 (12%)	23 (7%)	1	17
1	C	348/454 (77%)	283 (81%)	40 (12%)	25 (7%)	1	16
1	D	348/454 (77%)	285 (82%)	42 (12%)	21 (6%)	1	19
1	E	348/454 (77%)	283 (81%)	42 (12%)	23 (7%)	1	17
1	F	348/454 (77%)	289 (83%)	38 (11%)	21 (6%)	1	19
2	G	136/143 (95%)	113 (83%)	13 (10%)	10 (7%)	1	15
2	H	136/143 (95%)	109 (80%)	17 (12%)	10 (7%)	1	15
2	I	136/143 (95%)	111 (82%)	14 (10%)	11 (8%)	1	13
All	All	2496/3153 (79%)	2042 (82%)	287 (12%)	167 (7%)	1	17

5 of 167 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	63	GLU
1	A	188	THR
1	A	200	PHE
1	A	203	SER
1	A	232	ASN

### 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	297/386 (77%)	269 (91%)	28 (9%)	8	31
1	B	297/386 (77%)	269 (91%)	28 (9%)	8	31
1	C	297/386 (77%)	269 (91%)	28 (9%)	8	31
1	D	297/386 (77%)	267 (90%)	30 (10%)	7	29
1	E	297/386 (77%)	268 (90%)	29 (10%)	8	29
1	F	297/386 (77%)	269 (91%)	28 (9%)	8	31
2	G	117/116 (101%)	108 (92%)	9 (8%)	13	40
2	H	117/116 (101%)	107 (92%)	10 (8%)	10	37
2	I	117/116 (101%)	111 (95%)	6 (5%)	24	52
All	All	2133/2664 (80%)	1937 (91%)	196 (9%)	9	32

5 of 196 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	67	LEU
1	D	339	VAL
2	H	497	LEU
1	D	81	GLU
1	D	202	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 53 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	421	ASN
1	D	323	GLN
2	H	494	HIS
1	D	53	HIS
1	D	222	ASN

### 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 ⓘ

### 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	358/454 (78%)	-0.01	12 (3%)	45	36	156, 167, 168, 170	0
1	B	358/454 (78%)	0.09	17 (4%)	31	26	163, 167, 168, 175	0
1	C	358/454 (78%)	-0.05	10 (2%)	53	42	157, 167, 168, 170	0
1	D	358/454 (78%)	-0.03	11 (3%)	49	38	163, 167, 168, 182	0
1	E	358/454 (78%)	0.06	17 (4%)	31	26	163, 167, 168, 170	0
1	F	358/454 (78%)	0.53	38 (10%)	6	6	162, 167, 168, 187	0
2	G	132/143 (92%)	-0.21	2 (1%)	73	64	163, 167, 168, 169	0
2	H	132/143 (92%)	-0.19	4 (3%)	50	39	162, 167, 168, 169	0
2	I	132/143 (92%)	0.03	5 (3%)	40	32	163, 167, 168, 169	0
All	All	2544/3153 (80%)	0.07	116 (4%)	32	27	156, 167, 168, 187	0

The worst 5 of 116 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	I	510	GLU	7.1
1	F	326	GLN	6.7
1	F	199	GLY	6.2
1	B	260	ALA	5.8
1	A	287	GLY	5.5

### 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.