



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 23, 2020 – 07:18 pm BST

PDB ID : 1E0K  
Title : gp4d helicase from phage T7  
Authors : Singleton, M.R.; Sawaya, M.R.; Ellenberger, T.; Wigley, D.B.  
Deposited on : 2000-03-30  
Resolution : 3.30 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  
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

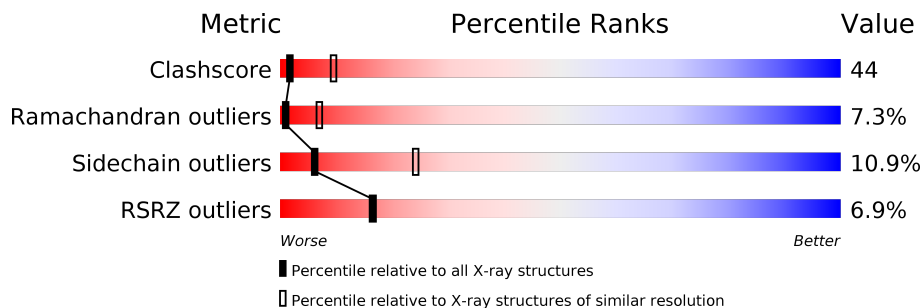
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.30 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	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

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	289	 7% 36% 50% 13%
1	B	289	 4% 37% 50% 12%
1	C	289	 4% 38% 51% 11%
1	D	289	 10% 35% 49% 15%
1	E	289	 9% 36% 50% 13%
1	F	289	 7% 35% 54% 10%

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 13284 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.

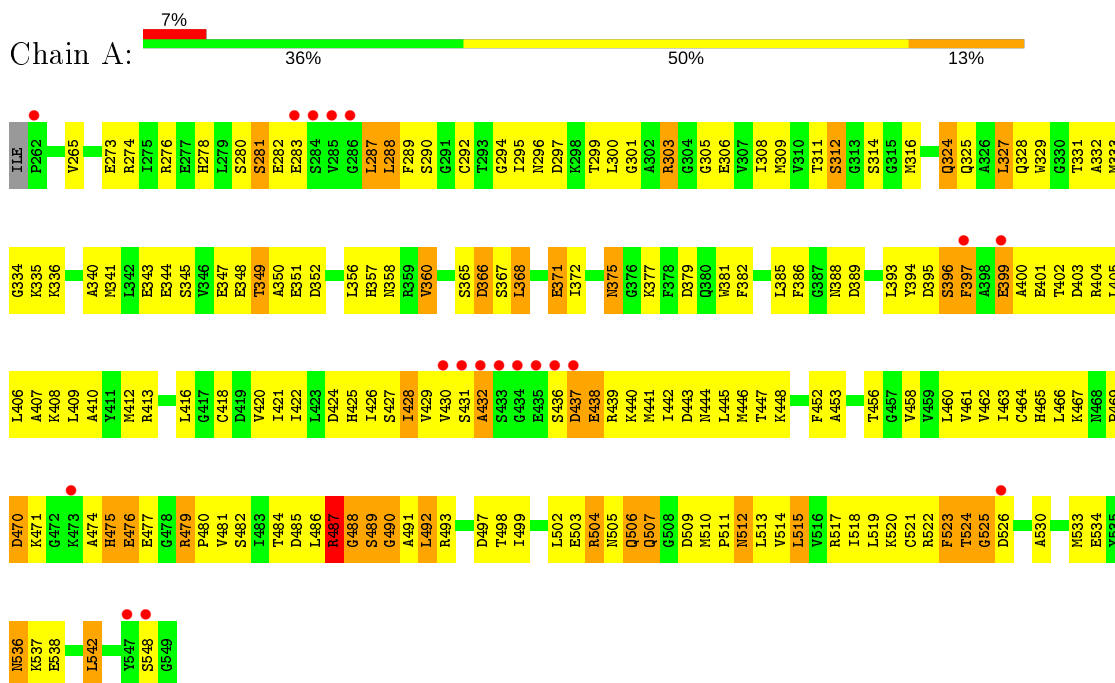
- Molecule 1 is a protein called DNA HELICASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	288	2214	1378	390	433	13	0	0	0
1	B	288	2214	1378	390	433	13	0	0	0
1	C	288	2214	1378	390	433	13	0	0	0
1	D	288	2214	1378	390	433	13	0	0	0
1	E	288	2214	1378	390	433	13	0	0	0
1	F	288	2214	1378	390	433	13	0	0	0

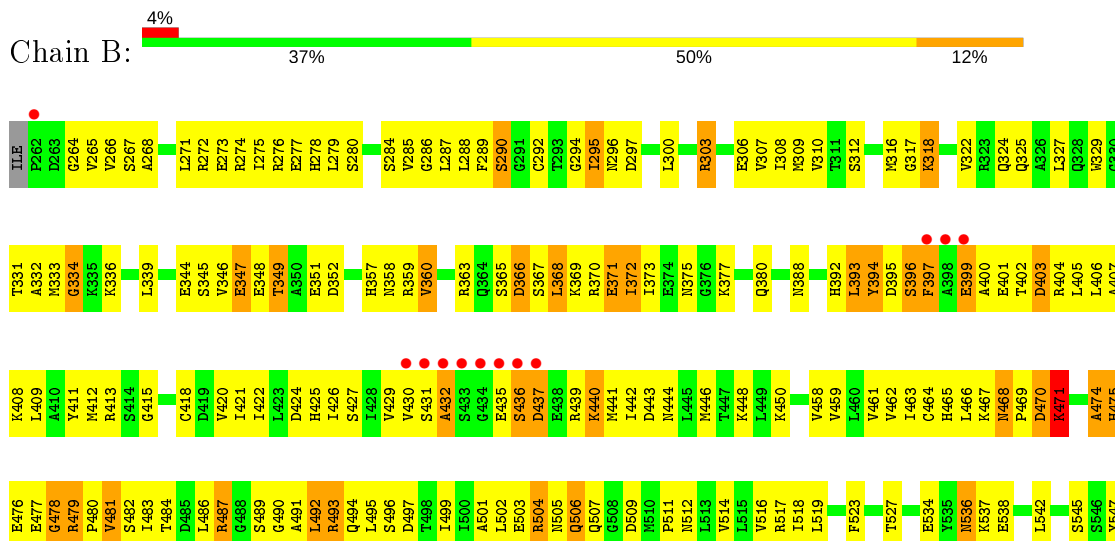
### 3 Residue-property plots [i](#)

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: DNA HELICASE

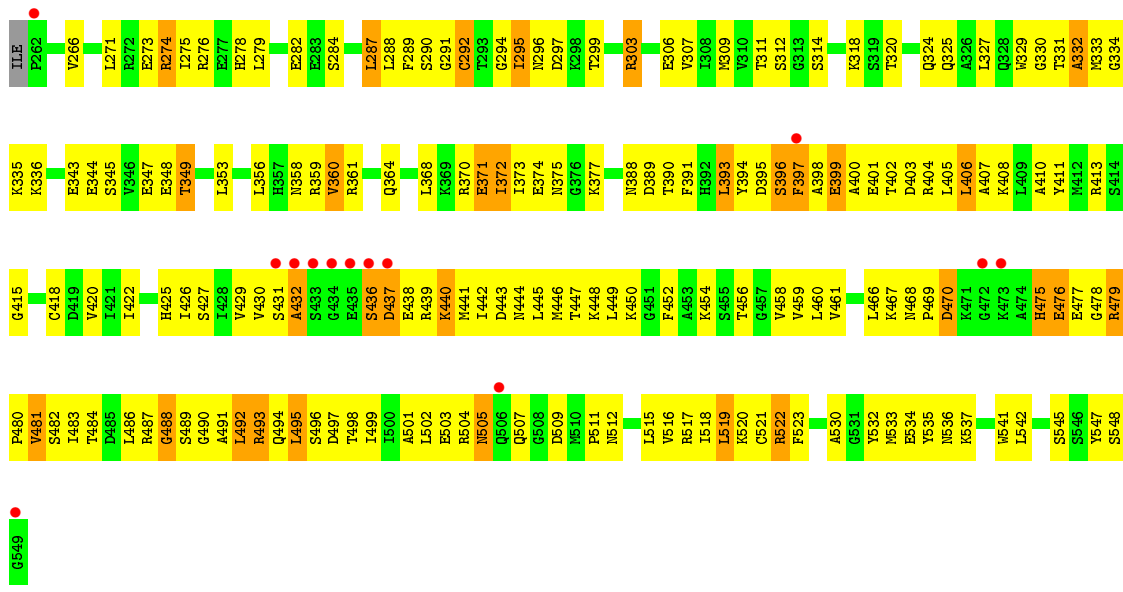


#### • Molecule 1: DNA HELICASE

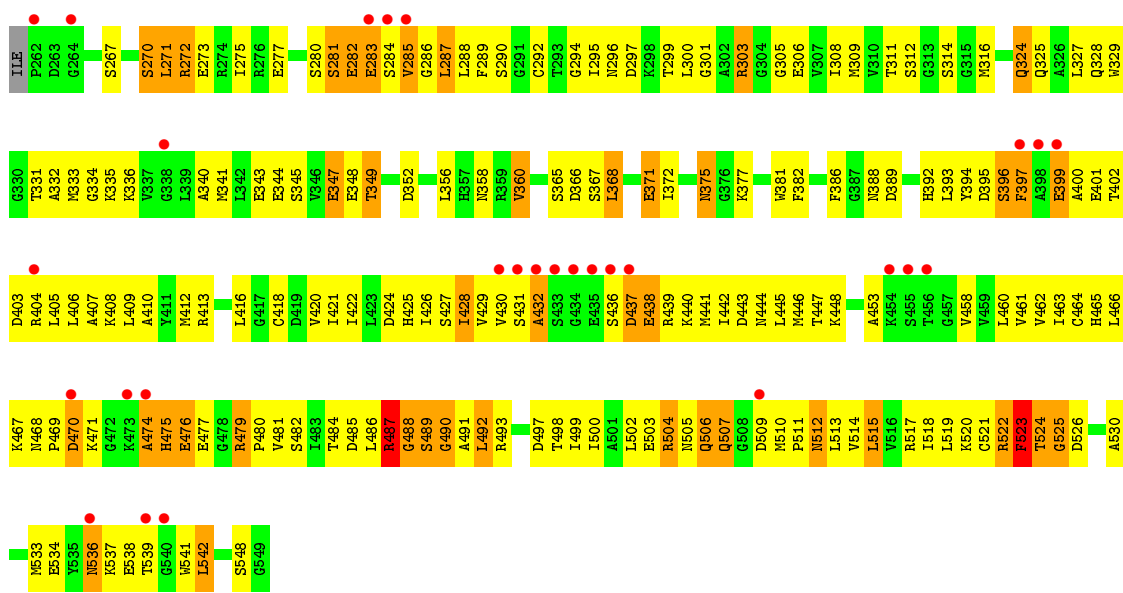




• Molecule 1: DNA HELICASE

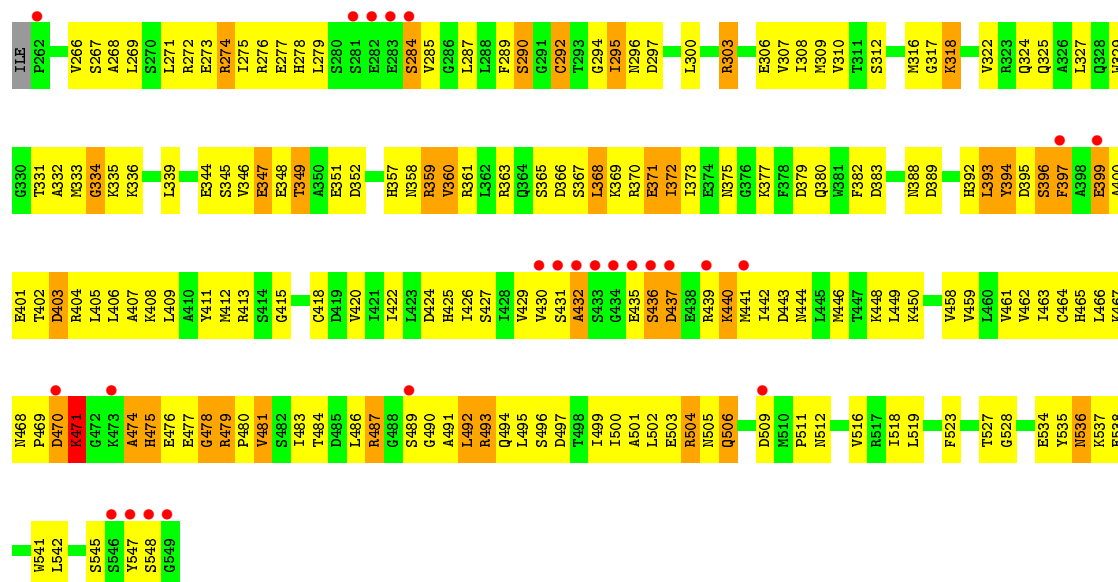


• Molecule 1: DNA HELICASE

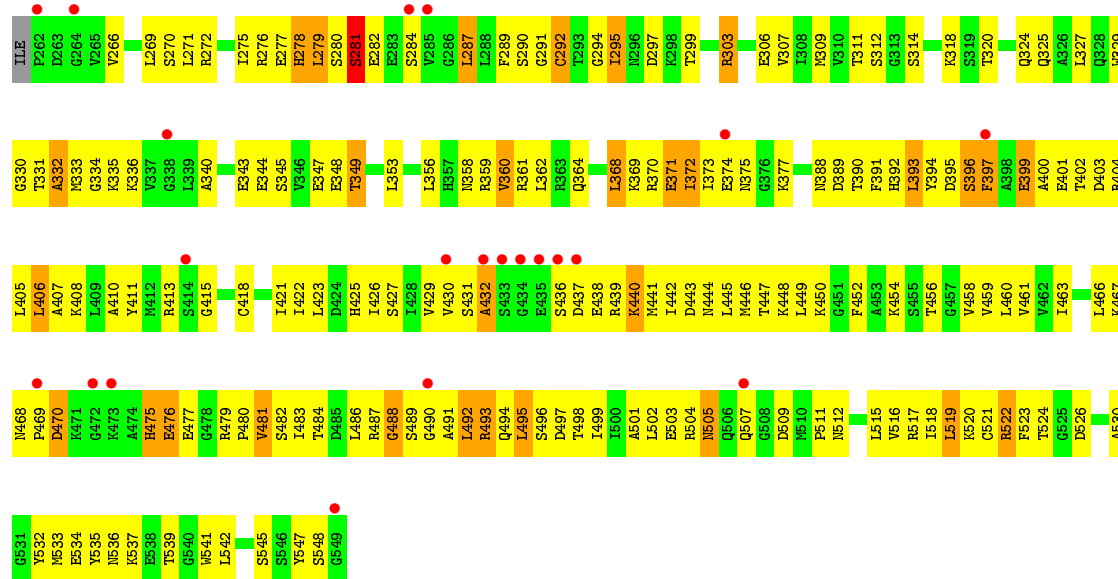


• Molecule 1: DNA HELICASE





• Molecule 1: DNA HELICASE



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	120.65Å 120.65Å 284.50Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.30 19.83 – 3.30	Depositor EDS
% Data completeness (in resolution range)	94.6 (20.00-3.30) 94.7 (19.83-3.30)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.24 (at 3.29Å)	Xtrriage
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.254 , 0.308 0.239 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	96.0	Xtrriage
Anisotropy	0.221	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 90.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	13284	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	99.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.75% 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](#)

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	0.48	0/2243	0.73	1/3011 (0.0%)
1	B	0.45	0/2244	0.77	3/3014 (0.1%)
1	C	0.47	0/2244	0.75	1/3014 (0.0%)
1	D	1.02	8/2243 (0.4%)	0.78	3/3011 (0.1%)
1	E	0.45	0/2244	0.77	3/3014 (0.1%)
1	F	0.43	0/2244	0.75	1/3014 (0.0%)
All	All	0.59	8/13462 (0.1%)	0.76	12/18078 (0.1%)

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	523	PHE	CE1-CZ	19.75	1.74	1.37
1	D	523	PHE	CD1-CE1	19.15	1.77	1.39
1	D	523	PHE	CE2-CZ	19.09	1.73	1.37
1	D	523	PHE	CD2-CE2	15.77	1.70	1.39
1	D	283	GLU	N-CA	14.31	1.75	1.46

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	282	GLU	C-N-CA	10.08	146.91	121.70
1	D	283	GLU	N-CA-C	6.61	128.84	111.00
1	B	295	ILE	N-CA-C	-5.85	95.19	111.00
1	E	474	ALA	N-CA-C	5.80	126.67	111.00
1	B	474	ALA	N-CA-C	5.72	126.46	111.00

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	2214	0	2216	184	0
1	B	2214	0	2217	206	0
1	C	2214	0	2217	189	0
1	D	2214	0	2215	239	0
1	E	2214	0	2217	202	0
1	F	2214	0	2217	184	0
All	All	13284	0	13299	1161	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 44.

The worst 5 of 1161 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:D:523:PHE:CE1	1:D:523:PHE:CD1	1.77	1.66
1:D:523:PHE:CE1	1:D:523:PHE:CZ	1.74	1.65
1:D:283:GLU:N	1:D:523:PHE:CE1	1.70	1.51
1:D:283:GLU:HA	1:D:523:PHE:CE1	1.42	1.50
1:D:283:GLU:CA	1:D:283:GLU:N	1.75	1.50

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	284/289 (98%)	225 (79%)	34 (12%)	25 (9%)	<b>1</b> <b>5</b>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	286/289 (99%)	226 (79%)	42 (15%)	18 (6%)	1	9
1	C	286/289 (99%)	225 (79%)	45 (16%)	16 (6%)	2	11
1	D	284/289 (98%)	223 (78%)	34 (12%)	27 (10%)	0	4
1	E	286/289 (99%)	231 (81%)	37 (13%)	18 (6%)	1	9
1	F	286/289 (99%)	225 (79%)	40 (14%)	21 (7%)	1	7
All	All	1712/1734 (99%)	1355 (79%)	232 (14%)	125 (7%)	1	7

5 of 125 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	396	SER
1	A	436	SER
1	A	437	ASP
1	A	475	HIS
1	A	489	SER

### 5.3.2 Protein sidechains [i](#)

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	241/242 (100%)	211 (88%)	30 (12%)	4	19
1	B	241/242 (100%)	217 (90%)	24 (10%)	7	27
1	C	241/242 (100%)	216 (90%)	25 (10%)	7	25
1	D	241/242 (100%)	212 (88%)	29 (12%)	5	20
1	E	241/242 (100%)	216 (90%)	25 (10%)	7	25
1	F	241/242 (100%)	217 (90%)	24 (10%)	7	27
All	All	1446/1452 (100%)	1289 (89%)	157 (11%)	6	24

5 of 157 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	492	LEU

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	360	VAL
1	F	470	ASP
1	C	495	LEU
1	D	270	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 62 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	494	GLN
1	D	325	GLN
1	F	392	HIS
1	D	278	HIS
1	D	375	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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1
1	D	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	D	283:GLU	C	284:SER	N	7.79
1	A	283:GLU	C	284:SER	N	6.86

## 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	288/289 (99%)	0.13	19 (6%) 18 18	40, 88, 175, 200	0
1	B	288/289 (99%)	-0.12	13 (4%) 33 32	23, 71, 167, 200	0
1	C	288/289 (99%)	-0.09	13 (4%) 33 32	30, 73, 168, 192	0
1	D	288/289 (99%)	0.41	28 (9%) 7 8	65, 117, 178, 200	0
1	E	288/289 (99%)	0.12	25 (8%) 10 10	47, 103, 176, 200	0
1	F	288/289 (99%)	0.04	21 (7%) 15 15	43, 94, 172, 200	0
All	All	1728/1734 (99%)	0.08	119 (6%) 16 16	23, 92, 175, 200	0

The worst 5 of 119 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	284	SER	18.5
1	D	284	SER	12.2
1	E	431	SER	10.7
1	A	435	GLU	10.4
1	D	434	GLY	10.3

### 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 [i](#)

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.