



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 23, 2017 – 01:08 PM EST

PDB ID : 5LB8  
Title : Crystal structure of human RECQL5 helicase APO form.  
Authors : Newman, J.A.; Aitkenhead, H.; Savitsky, P.; Krojer, T.; Von Delft, F.; Arrow-smith, C.H.; Edwards, A.M.; Bountra, C.; Gileadi, O.; Structural Genomics Consortium (SGC)  
Deposited on : 2016-06-15  
Resolution : 3.40 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-20028442
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-20028442

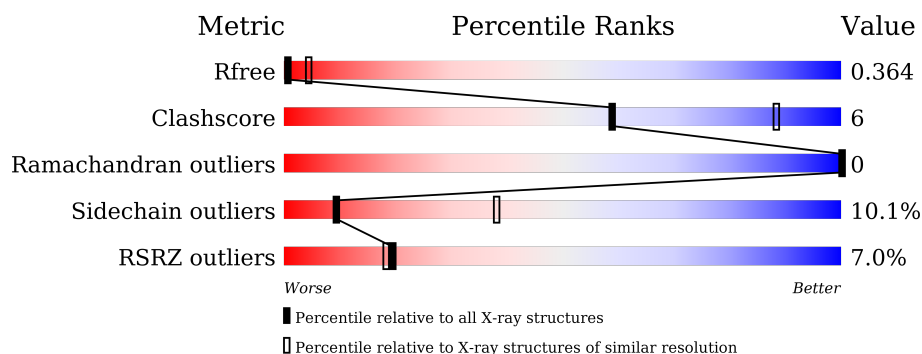
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.40 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))
$R_{free}$	91344	1476 (3.50-3.30)
Clashscore	102246	1611 (3.50-3.30)
Ramachandran outliers	100387	1571 (3.50-3.30)
Sidechain outliers	100360	1571 (3.50-3.30)
RSRZ outliers	91569	1485 (3.50-3.30)

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	518	<div> <div>6%</div> <div>63%</div> <div>18%</div> <div>•</div> <div>18%</div> </div>
1	D	518	<div> <div>6%</div> <div>65%</div> <div>17%</div> <div>•</div> <div>18%</div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6566 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 ATP-dependent DNA helicase Q5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	427	Total	C	N	O	S	0	0	0
			3266	2070	572	602	22			
1	D	426	Total	C	N	O	S	0	0	0
			3298	2087	584	605	22			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	9	SER	-	expression tag	UNP O94762
A	10	MET	-	expression tag	UNP O94762
D	9	SER	-	expression tag	UNP O94762
D	10	MET	-	expression tag	UNP O94762

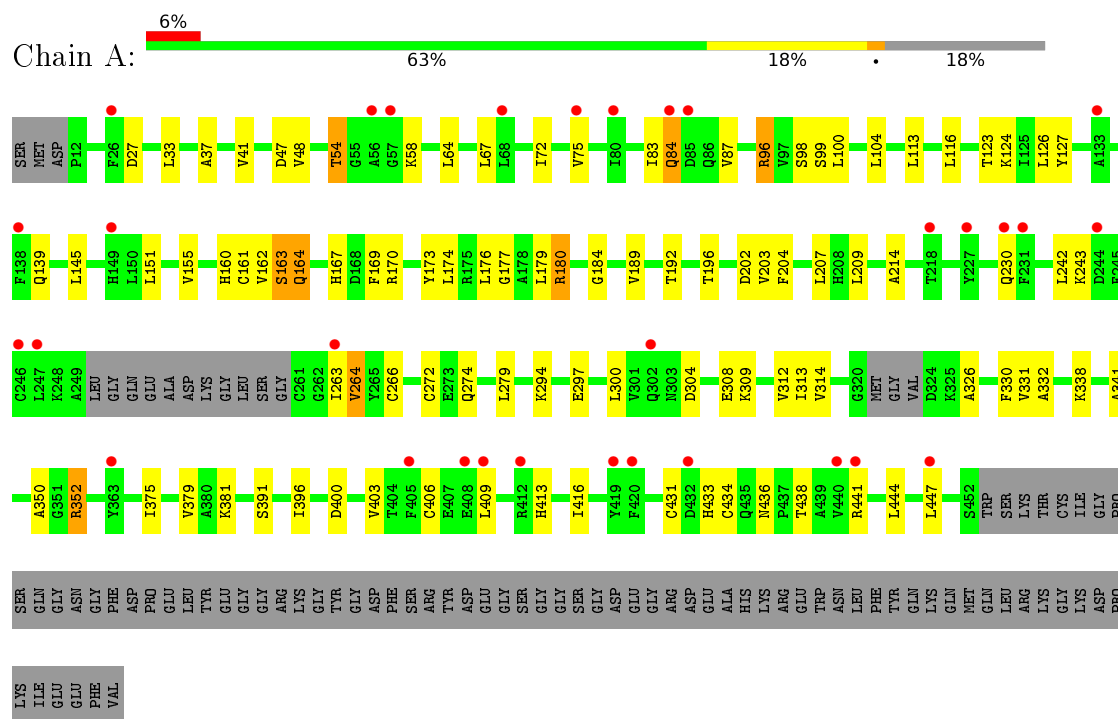
- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Zn	0	0
			1	1		
2	D	1	Total	Zn	0	0
			1	1		

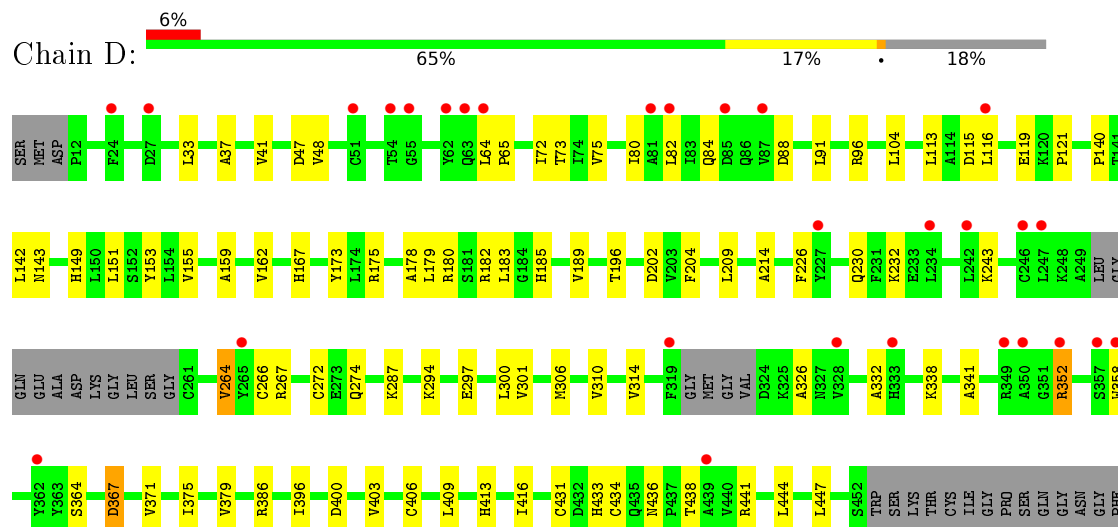
### 3 Residue-property plots [i](#)

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: ATP-dependent DNA helicase Q5



#### • Molecule 1: ATP-dependent DNA helicase Q5



ASP	PRO	GLU	LEU	TYR	GLU	GLY	GLY	ARG	LYS	GLY	TYR	GLY	ASP	PHE	SER	ARG	TYR	ASP	GLU	GLY	SER	GLY	GLY	SER	GLY	ASP	GLU	GLY	ARG	ASP	GLU	ALA	HIS	LYS	ARG	GLU	TRP	ASN	LEU	PHE	TYR	GLN	LYS	GLN	MET	GLN	LEU	ARG	LYS	GLY	LYS	ASP	PRO	LYS	ILE	GLU	GLU	PHE	VAL
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	68.74Å 68.74Å 396.31Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.61 – 3.40 48.61 – 3.40	Depositor EDS
% Data completeness (in resolution range)	99.8 (48.61-3.40) 99.8 (48.61-3.40)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.04 (at 3.40Å)	Xtriage
Refinement program	BUSTER 2.10.2	Depositor
R, $R_{free}$	0.301 , 0.322 0.317 , 0.364	Depositor DCC
$R_{free}$ test set	684 reflections (5.12%)	DCC
Wilson B-factor (Å <sup>2</sup> )	139.0	Xtriage
Anisotropy	0.065	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 118.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	6566	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	145.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.46% 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

Bond lengths and bond angles in the following residue types are not validated in this section: ZN

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	0.45	0/3330	0.63	0/4513
1	D	0.40	0/3362	0.61	0/4550
All	All	0.43	0/6692	0.62	0/9063

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	3266	0	3229	43	0
1	D	3298	0	3290	34	0
2	A	1	0	0	0	0
2	D	1	0	0	0	0
All	All	6566	0	6519	76	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 76 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:162:VAL:CG1	1:A:173:TYR:HB3	2.09	0.81
1:A:162:VAL:HG13	1:A:173:TYR:HB3	1.62	0.81
1:A:75:VAL:HA	1:A:155:VAL:HB	1.78	0.66
1:A:54:THR:HA	1:A:58:LYS:HD3	1.80	0.63
1:A:163:SER:O	1:A:169:PHE:HD1	1.80	0.63

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	421/518 (81%)	394 (94%)	27 (6%)	0	100	100
1	D	420/518 (81%)	390 (93%)	30 (7%)	0	100	100
All	All	841/1036 (81%)	784 (93%)	57 (7%)	0	100	100

There are no Ramachandran outliers to report.

### 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	344/432 (80%)	306 (89%)	38 (11%)	8	34
1	D	351/432 (81%)	319 (91%)	32 (9%)	12	44
All	All	695/864 (80%)	625 (90%)	70 (10%)	9	38

5 of 70 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	381	LYS
1	D	47	ASP
1	D	406	CYS
1	A	406	CYS
1	A	436	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	143	ASN
1	A	164	GLN
1	D	143	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](#)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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	427/518 (82%)	0.45	31 (7%)	18 17	83, 141, 176, 189	0
1	D	426/518 (82%)	0.43	29 (6%)	20 19	115, 149, 179, 199	0
All	All	853/1036 (82%)	0.44	60 (7%)	19 18	83, 146, 178, 199	0

The worst 5 of 60 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	350	ALA	5.8
1	A	246	CYS	5.6
1	D	85	ASP	5.3
1	D	357	SER	5.2
1	D	82	LEU	4.9

### 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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	ZN	A	601	1/1	0.88	0.17	-0.90	166,166,166,166	0
2	ZN	D	601	1/1	0.90	0.11	-1.28	132,132,132,132	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.