wwPDB X-ray Structure Validation Summary Report (i)

Sep 27, 2022 – 11:00 am BST

PDB ID	:	7ZMV
Title	:	Crystal structure of human RECQL5 helicase APO form in complex with engineered nanobody (Gluebody) G5-006
Authors	:	Ye, M.; Makola, M.; Newman, J.A.; Fairhead, M.; MacLean, E.; Krojer, T.; Aitkenhead, H.; Bountra, C.; Gileadi, O.; von Delft, F.
Deposited on	:	2022-04-19
Resolution	:	2.00 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

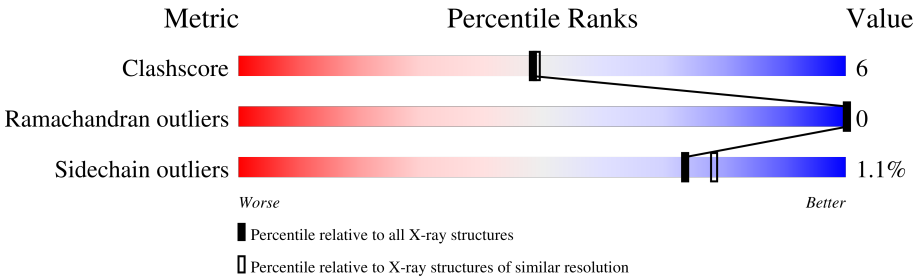
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	FAILED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)

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 of the lower bar indicate the fraction of residues that contain outliers for >=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 <=5%

Note EDS failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	445	84% 14%
1	B	445	86% 12%
1	C	445	87% 11%
1	D	445	86% 13%
2	E	127	83% 14%
2	F	127	82% 16%
2	G	127	81% 17%
2	K	127	83% 14%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 18647 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	440	Total	C	N	O	S	0	5	0
			3455	2182	623	627	23			
1	B	440	Total	C	N	O	S	0	4	0
			3445	2177	618	627	23			
1	C	440	Total	C	N	O	S	0	5	0
			3452	2181	621	627	23			
1	D	440	Total	C	N	O	S	0	5	0
			3459	2185	624	627	23			

There are 8 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
B	9	SER	-	expression tag	UNP O94762
B	10	MET	-	expression tag	UNP O94762
C	9	SER	-	expression tag	UNP O94762
C	10	MET	-	expression tag	UNP O94762
D	9	SER	-	expression tag	UNP O94762
D	10	MET	-	expression tag	UNP O94762

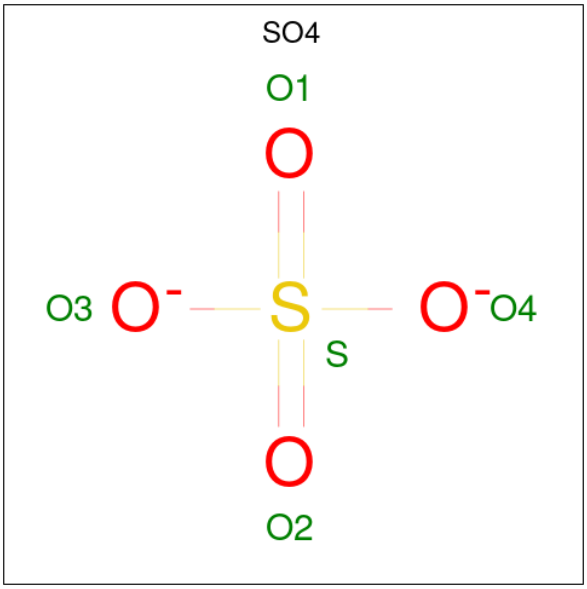
- Molecule 2 is a protein called Gluebody G5-006.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	K	124	Total	C	N	O	S	0	1	0
			946	586	161	193	6			
2	E	124	Total	C	N	O	S	0	1	0
			946	586	161	193	6			
2	F	124	Total	C	N	O	S	0	1	0
			946	586	161	193	6			
2	G	124	Total	C	N	O	S	0	1	0
			946	586	161	193	6			

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Zn 1 1	0	0
3	B	1	Total Zn 1 1	0	0
3	C	1	Total Zn 1 1	0	0
3	D	1	Total Zn 1 1	0	0

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total O S 5 4 1	0	0
4	B	1	Total O S 5 4 1	0	0
4	C	1	Total O S 5 4 1	0	0
4	D	1	Total O S 5 4 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	187	Total O 187 187	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	186	Total O 186 186	0	0
5	C	193	Total O 193 193	0	0
5	D	186	Total O 186 186	0	0
5	K	78	Total O 78 78	0	0
5	E	60	Total O 60 60	0	0
5	F	72	Total O 72 72	0	0
5	G	66	Total O 66 66	0	0

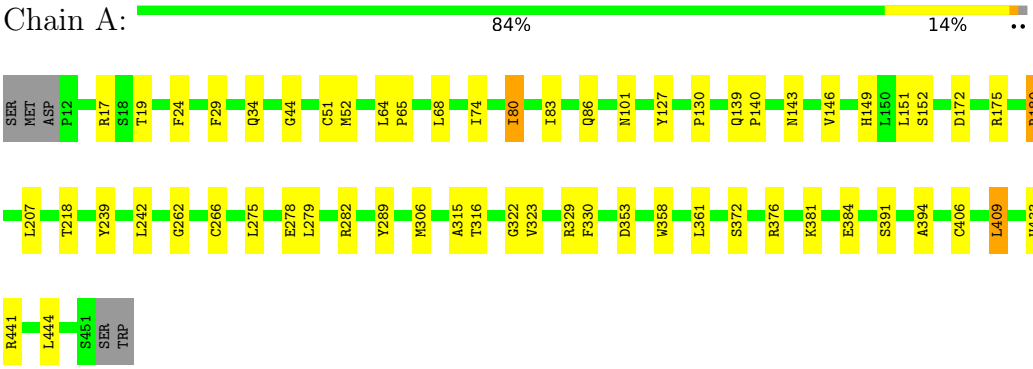
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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. 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. 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.

Note EDS failed to run properly.

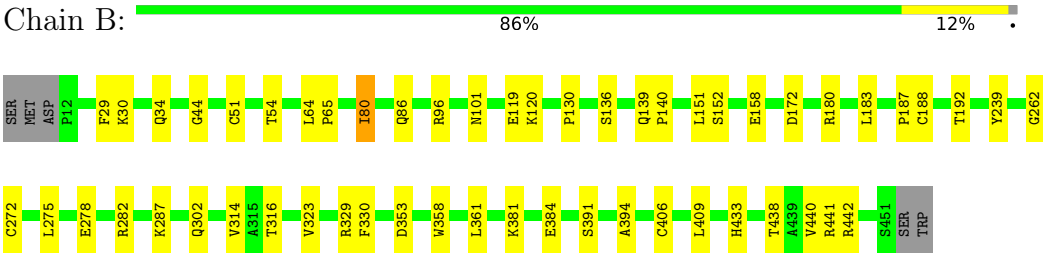
• Molecule 1: ATP-dependent DNA helicase Q5

Chain A:



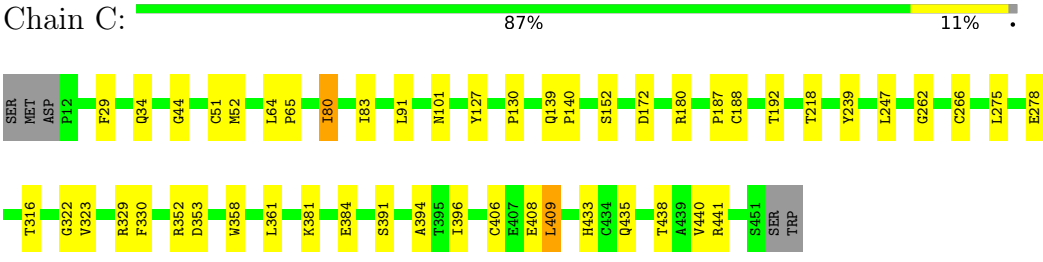
• Molecule 1: ATP-dependent DNA helicase Q5

Chain B:



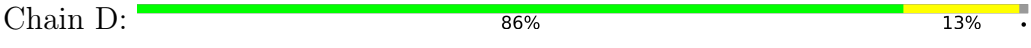
• Molecule 1: ATP-dependent DNA helicase Q5

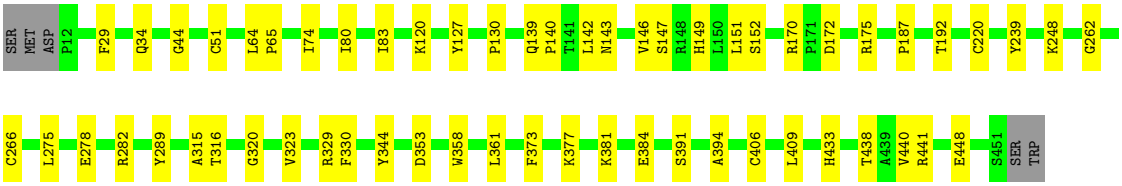
Chain C:



• Molecule 1: ATP-dependent DNA helicase Q5

Chain D:



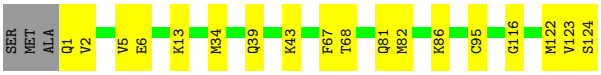


• Molecule 2: Gluebody G5-006

Chain K:

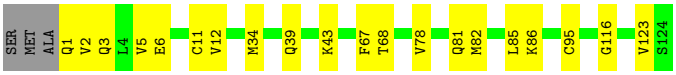
• Molecule 2: Gluebody G5-006

Chain E:



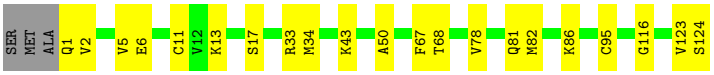
• Molecule 2: Gluebody G5-006

Chain F:



• Molecule 2: Gluebody G5-006

Chain G:



4 Data and refinement statistics (i)

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	232.59Å 89.97Å 164.13Å 90.00° 110.06° 90.00°	Depositor
Resolution (Å)	77.69 – 2.00	Depositor
% Data completeness (in resolution range)	97.6 (77.69-2.00)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.45 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.243 , 0.275	Depositor
Wilson B-factor (Å^2)	39.8	Xtriage
Anisotropy	0.067	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	18647	wwPDB-VP
Average B, all atoms (Å^2)	51.0	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 61.04 % 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 1.3908e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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, SO4

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	0.69	0/3531	0.79	1/4770 (0.0%)
1	B	0.71	1/3518 (0.0%)	0.80	0/4753
1	C	0.69	0/3535	0.80	0/4774
1	D	0.70	0/3535	0.79	0/4774
2	E	0.66	0/968	0.85	0/1314
2	F	0.69	0/968	0.86	0/1314
2	G	0.71	0/968	0.85	0/1314
2	K	0.70	0/968	0.87	0/1314
All	All	0.69	1/17991 (0.0%)	0.81	1/24327 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	158	GLU	CD-OE2	-5.01	1.20	1.25

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	180	ARG	CG-CD-NE	5.48	123.32	111.80

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	3455	0	3486	48	0
1	B	3445	0	3473	36	0
1	C	3452	0	3479	40	0
1	D	3459	0	3497	38	0
2	E	946	0	887	9	0
2	F	946	0	887	12	0
2	G	946	0	887	14	0
2	K	946	0	887	13	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	5	0	0	0	0
4	B	5	0	0	0	0
4	C	5	0	0	0	0
4	D	5	0	0	0	0
5	A	187	0	0	8	0
5	B	186	0	0	6	0
5	C	193	0	0	9	0
5	D	186	0	0	6	0
5	E	60	0	0	1	0
5	F	72	0	0	2	0
5	G	66	0	0	2	0
5	K	78	0	0	5	0
All	All	18647	0	17483	209	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 209 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:143:ASN:O	1:D:146:VAL:HG22	1.60	1.00
1:D:248:LYS:HG2	5:D:647:HOH:O	1.62	0.97
1:C:352:ARG:HD3	5:C:660:HOH:O	1.71	0.90
1:B:442:ARG:HG3	5:B:730:HOH:O	1.71	0.89
1:A:306:MET:HE1	1:A:323:VAL:HG21	1.62	0.82

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	443/445 (100%)	432 (98%)	11 (2%)	0	100	100
1	B	442/445 (99%)	433 (98%)	9 (2%)	0	100	100
1	C	443/445 (100%)	430 (97%)	13 (3%)	0	100	100
1	D	443/445 (100%)	430 (97%)	13 (3%)	0	100	100
2	E	123/127 (97%)	119 (97%)	4 (3%)	0	100	100
2	F	123/127 (97%)	120 (98%)	3 (2%)	0	100	100
2	G	123/127 (97%)	118 (96%)	5 (4%)	0	100	100
2	K	123/127 (97%)	119 (97%)	4 (3%)	0	100	100
All	All	2263/2288 (99%)	2201 (97%)	62 (3%)	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	369/373 (99%)	366 (99%)	3 (1%)	81	86
1	B	368/373 (99%)	364 (99%)	4 (1%)	73	78
1	C	370/373 (99%)	368 (100%)	2 (0%)	88	92
1	D	370/373 (99%)	367 (99%)	3 (1%)	81	86
2	E	98/100 (98%)	95 (97%)	3 (3%)	40	40
2	F	98/100 (98%)	96 (98%)	2 (2%)	55	58

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
2	G	98/100 (98%)	96 (98%)	2 (2%)	55 58
2	K	98/100 (98%)	97 (99%)	1 (1%)	76 81
All	All	1869/1892 (99%)	1849 (99%)	20 (1%)	73 78

5 of 20 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	E	68	THR
2	F	95	CYS
2	G	95	CYS
2	G	68	THR
1	B	409	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	302	GLN
2	K	120	GLN
2	G	120	GLN
2	K	81	GLN
2	E	3	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 8 ligands modelled in this entry, 4 are monoatomic - leaving 4 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
4	SO4	C	502	-	4,4,4	0.30	0	6,6,6	0.22	0
4	SO4	D	502	-	4,4,4	0.12	0	6,6,6	0.27	0
4	SO4	A	502	-	4,4,4	0.29	0	6,6,6	0.12	0
4	SO4	B	502	-	4,4,4	0.31	0	6,6,6	0.23	0

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

6.1 Protein, DNA and RNA chains

EDS failed to run properly - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

EDS failed to run properly - this section is therefore empty.

6.3 Carbohydrates

EDS failed to run properly - this section is therefore empty.

6.4 Ligands

EDS failed to run properly - this section is therefore empty.

6.5 Other polymers

EDS failed to run properly - this section is therefore empty.