



# wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 17, 2019 – 03:38 PM EDT

PDB ID : 5N4W  
Title : Crystal structure of the Cul2-Rbx1-EloBC-VHL ubiquitin ligase complex  
Authors : Cardote, T.A.F.; Gadd, M.S.; Ciulli, A.  
Deposited on : 2017-02-11  
Resolution : 3.90 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	:	rb-20031633
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac	:	5.8.0158
CCP4	:	7.0 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20031633

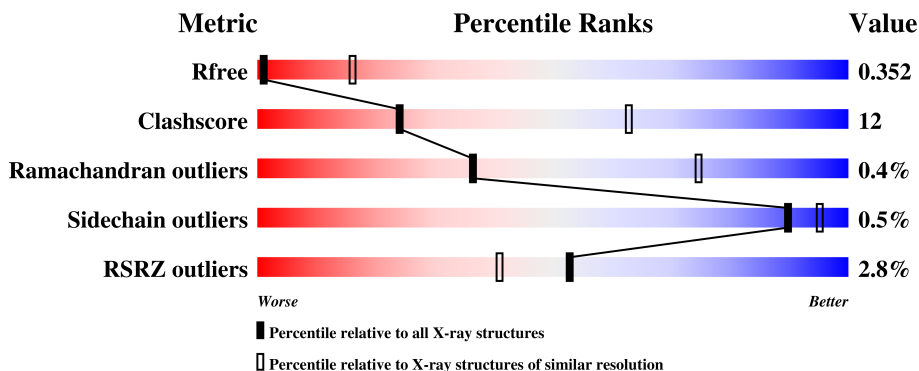
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.90 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))
$R_{free}$	111664	1145 (4.20-3.60)
Clashscore	122126	1225 (4.20-3.60)
Ramachandran outliers	120053	1184 (4.20-3.60)
Sidechain outliers	120020	1175 (4.20-3.60)
RSRZ outliers	108989	1046 (4.20-3.60)

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	748	<div> <div> <div></div> <div>65%</div> <div>24%</div> <div>10%</div> </div> </div>
2	V	160	<div> <div> <div></div> <div>18%</div> <div>6%</div> <div>77%</div> </div> </div>
3	R	102	<div> <div> <div></div> <div>51%</div> <div>33%</div> <div>16%</div> </div> </div>
4	B	104	<div> <div> <div></div> <div>62%</div> <div>17%</div> <div>21%</div> </div> </div>
5	C	97	<div> <div> <div></div> <div>65%</div> <div>29%</div> <div>6%</div> </div> </div>

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 7722 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 Cullin-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	671	Total	C	N	O	S	0	0	0
			5384	3402	921	1019	42			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	expression tag	UNP Q13617
A	-1	GLY	-	expression tag	UNP Q13617
A	0	SER	-	expression tag	UNP Q13617

- Molecule 2 is a protein called Von Hippel-Lindau disease tumor suppressor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	V	37	Total	C	N	O	S	0	0	0
			315	196	59	59	1			

- Molecule 3 is a protein called E3 ubiquitin-protein ligase RBX1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	R	86	Total	C	N	O	S	0	0	0
			690	433	128	120	9			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
R	18	LYS	GLY	conflict	UNP P62877

- Molecule 4 is a protein called Elongin-B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	B	82	Total	C	N	O	S	0	1	0
			629	399	105	122	3			

- Molecule 5 is a protein called Elongin-C.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	C	91	Total	C	N	O	S	0	1	0
			701	450	113	132	6			

There is a discrepancy between the modelled and reference sequences:

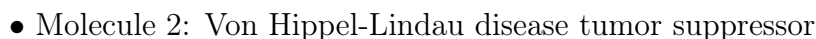
Chain	Residue	Modelled	Actual	Comment	Reference
C	16	MET	-	initiating methionine	UNP Q15369

- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn).

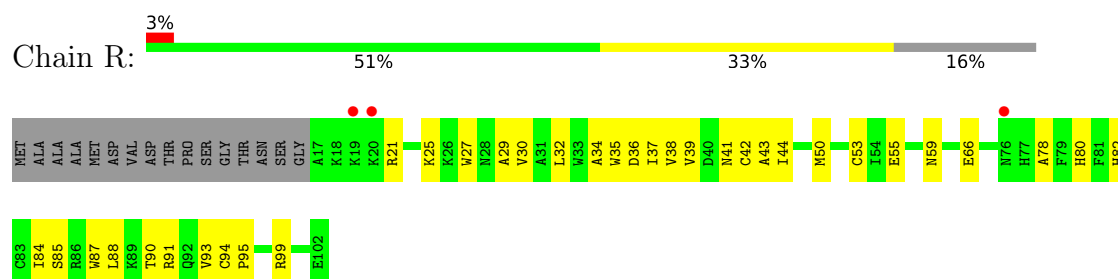
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	R	3	Total	Zn	0	0
			3	3		



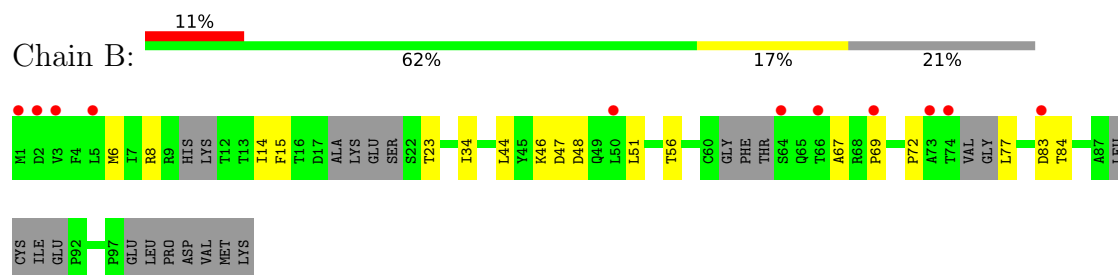
- Molecule 1: Cullin-2



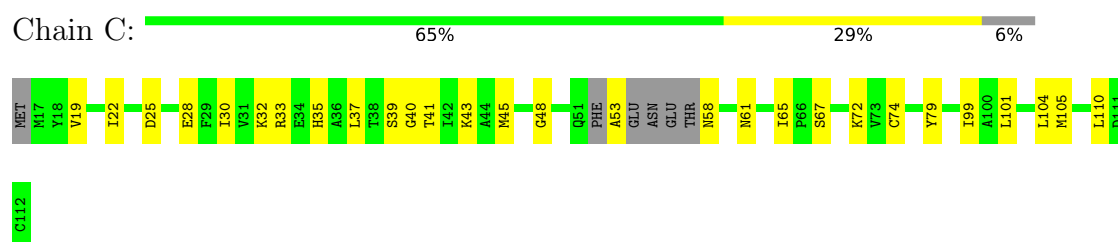
- Molecule 3: E3 ubiquitin-protein ligase RBX1



- Molecule 4: Elongin-B



- Molecule 5: Elongin-C



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	86.04Å 190.96Å 238.88Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	95.48 – 3.90 95.48 – 3.90	Depositor EDS
% Data completeness (in resolution range)	99.3 (95.48-3.90) 90.2 (95.48-3.90)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.21 (at 3.89Å)	Xtriage
Refinement program	PHENIX (1.11_2558: ???)	Depositor
R, $R_{free}$	0.302 , 0.346 0.306 , 0.352	Depositor DCC
$R_{free}$ test set	956 reflections (5.23%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	119.2	Xtriage
Anisotropy	0.624	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.23 , 117.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	7722	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	190.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.17% 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](#)

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	0.25	0/5468	0.46	1/7354 (0.0%)
2	V	0.20	0/319	0.40	0/430
3	R	0.24	0/707	0.51	0/958
4	B	0.23	0/639	0.45	0/858
5	C	0.25	0/717	0.40	0/965
All	All	0.25	0/7850	0.45	1/10565 (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	A	530	PRO	N-CA-CB	6.02	110.53	103.30

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	5384	0	5262	128	0
2	V	315	0	320	10	0
3	R	690	0	654	29	0
4	B	629	0	611	11	0
5	C	701	0	690	22	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	R	3	0	0	0	0
All	All	7722	0	7537	181	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 12.

The worst 5 of 181 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)
3:R:55:GLU:O	3:R:59:ASN:HB2	1.69	0.92
1:A:472:LYS:O	1:A:476:MET:HB2	1.74	0.88
3:R:84:ILE:O	3:R:88:LEU:HB2	1.74	0.87
1:A:535:LYS:O	1:A:536:SER:OG	1.94	0.85
1:A:475:ARG:NE	3:R:66:GLU:OE2	2.09	0.84

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	638/748 (85%)	574 (90%)	62 (10%)	2 (0%)	43	79
2	V	35/160 (22%)	34 (97%)	1 (3%)	0	100	100
3	R	84/102 (82%)	69 (82%)	14 (17%)	1 (1%)	14	55
4	B	71/104 (68%)	64 (90%)	6 (8%)	1 (1%)	12	51
5	C	87/97 (90%)	83 (95%)	4 (5%)	0	100	100
All	All	915/1211 (76%)	824 (90%)	87 (10%)	4 (0%)	36	74

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	534	GLU
3	R	43	ALA
4	B	47	ASP
1	A	536	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	589/682 (86%)	585 (99%)	4 (1%)	85	93
2	V	37/147 (25%)	37 (100%)	0	100	100
3	R	73/86 (85%)	73 (100%)	0	100	100
4	B	67/92 (73%)	67 (100%)	0	100	100
5	C	77/86 (90%)	77 (100%)	0	100	100
All	All	843/1093 (77%)	839 (100%)	4 (0%)	90	95

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	116	ASN
1	A	497	GLN
1	A	697	ASN
1	A	708	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	116	ASN
1	A	697	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 3 ligands modelled in this entry, 3 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	671/748 (89%)	-0.22	11 (1%) 72 61	25, 148, 245, 265	0
2	V	37/160 (23%)	0.18	2 (5%) 26 20	270, 300, 314, 316	0
3	R	86/102 (84%)	-0.10	3 (3%) 44 33	110, 219, 286, 315	0
4	B	82/104 (78%)	0.74	11 (13%) 3 3	221, 268, 305, 321	0
5	C	91/97 (93%)	-0.42	0 100 100	225, 261, 303, 315	0
All	All	967/1211 (79%)	-0.13	27 (2%) 53 41	25, 206, 298, 321	0

The worst 5 of 27 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	B	64	SER	6.5
4	B	1	MET	6.1
4	B	2	ASP	4.3
3	R	19	LYS	4.0
1	A	734	SER	3.8

### 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. 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	B-factors(Å <sup>2</sup> )	Q<0.9
6	ZN	R	203	1/1	0.75	0.09	271,271,271,271	0
6	ZN	R	201	1/1	0.87	0.17	220,220,220,220	0
6	ZN	R	202	1/1	0.93	0.13	213,213,213,213	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.