

# wwPDB X-ray Structure Validation Summary Report (i)

#### Feb 18, 2024 – 06:57 AM EST

PDB ID : 4F52

Title : Structure of a Glomulin-RBX1-CUL1 complex Authors : Duda, D.M.; Olszewski, J.L.; Schulman, B.A.

Deposited on : 2012-05-11

Resolution : 3.00 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 (1)) were used in the production of this report:

 $\begin{array}{ccc} & Mol Probity & : & 4.02b\text{-}467 \\ & Xtriage \text{ (Phenix)} & : & 1.13 \end{array}$ 

EDS : 2.36

Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)

Refmac : 5.8.0158

CCP4 : 7.0.044 (Gargrove) roteins) : Engh & Huber (2001)

Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)

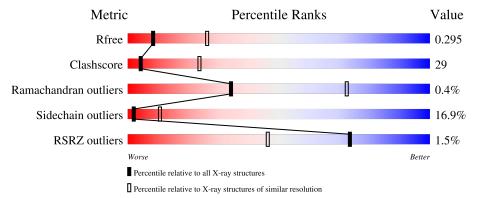
Validation Pipeline (wwPDB-VP) : 2.36

### 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 3.00 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 $(\# \mathrm{Entries})$	$\begin{array}{c} {\rm Similar\ resolution} \\ (\#{\rm Entries},{\rm resolution\ range}(\mathring{\rm A})) \end{array}$
$R_{free}$	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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% 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	282	47%	35%		9% 9%			
1	С	282	48%	35%		7% 10%			
2	В	106	36%	40%	•	21%			
2	D	106	37%	33%	10%	20%			
3	Е	596	40%	39%	_	8% 12%			

Continued on next page...



Continued from previous page...

Mol	Chain	Length		Quality of chain		
3	F	596	33%	44%	9%	14%



## 2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 13897 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-1.

Mol	Chain	Residues	${f Atoms}$				ZeroOcc	AltConf	Trace	
1	A	258	Total 2103	C 1341	N 351	O 400	S 11	0	0	0
1	С	254	Total 2073	C 1326	N 345	O 392	S 10	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	409	GLY	-	expression tag	UNP Q13616
A	410	SER	-	expression tag	UNP Q13616
A	421	GLU	LEU	engineered mutation	UNP Q13616
A	451	GLU	VAL	engineered mutation	UNP Q13616
A	452	LYS	VAL	engineered mutation	UNP Q13616
A	455	LYS	TYR	engineered mutation	UNP Q13616
С	409	GLY	-	expression tag	UNP Q13616
С	410	SER	-	expression tag	UNP Q13616
С	421	GLU	LEU	engineered mutation	UNP Q13616
С	451	GLU	VAL	engineered mutation	UNP Q13616
С	452	LYS	VAL	engineered mutation	UNP Q13616
С	455	LYS	TYR	engineered mutation	UNP Q13616

• Molecule 2 is a protein called E3 ubiquitin-protein ligase RBX1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	D	84	Total C N O S	0	0	0				
	2   B	04	692	438	127	118	9	U	0	
9	D	85	Total	С	N	О	S	0	0	0
	ש	00	701	443	128	121	9			U

There are 4 discrepancies between the modelled and reference sequences:



Chain	Residue	Modelled	Actual	Comment	Reference
В	3	GLY	-	expression tag	UNP P62877
В	4	SER	-	expression tag	UNP P62877
D	3	GLY	-	expression tag	UNP P62877
D	4	SER	-	expression tag	UNP P62877

• Molecule 3 is a protein called Glomulin.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
9	E	523	Total C N O S	0	0	0				
3	E	923	4202	2720	685	772	25	U	0	
9	E	512	Total	С	N	О	S	0	0	0
3	Г	312	4120	2676	666	753	25			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
Е	-1	GLY	-	expression tag	UNP Q92990
E	0	SER	-	expression tag	UNP Q92990
F	-1	GLY	-	expression tag	UNP Q92990
F	0	SER	-	expression tag	UNP Q92990

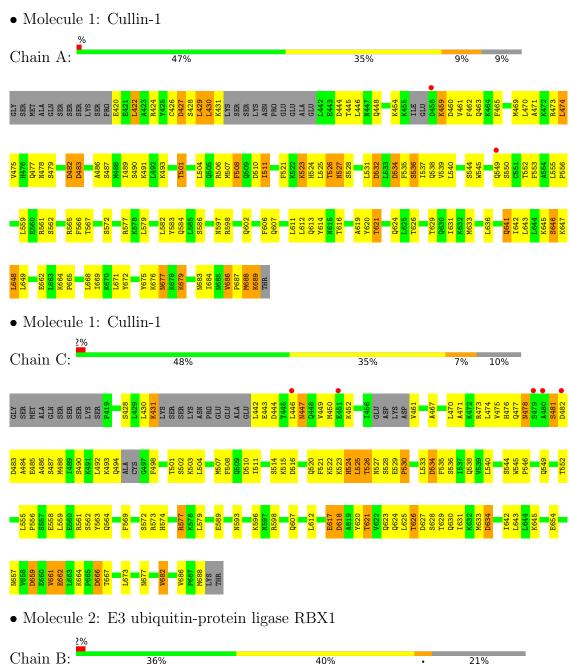
• Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	В	3	Total Zn 3 3	0	0
4	D	3	Total Zn 3 3	0	0

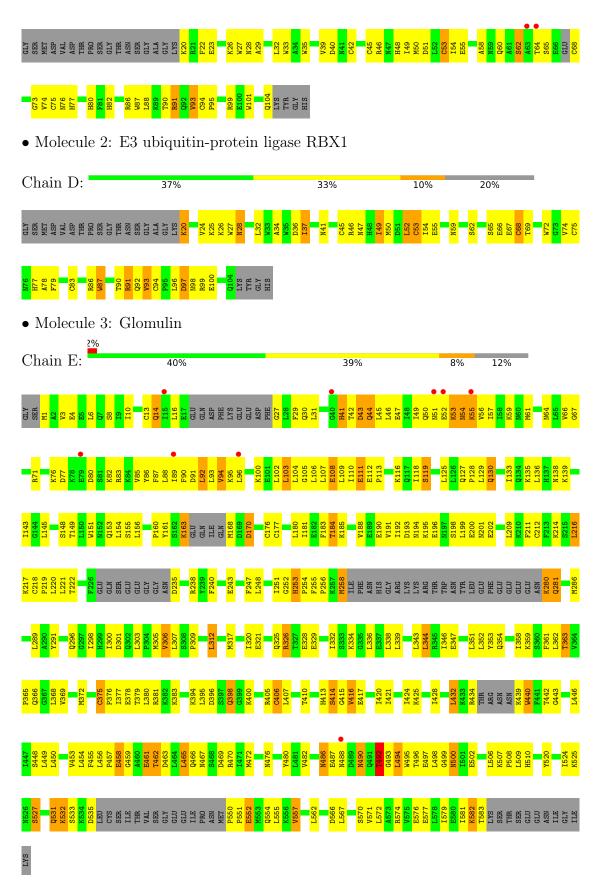


### 3 Residue-property plots (i)

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 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.

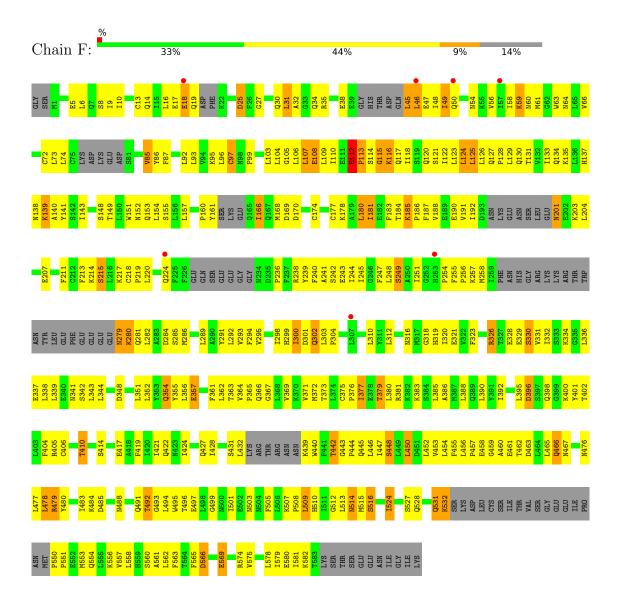






• Molecule 3: Glomulin







## 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants	53.33Å 193.93Å 142.07Å	Donogitor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $98.81^{\circ}$ $90.00^{\circ}$	Depositor
Resolution (Å)	37.39 - 3.00	Depositor
Resolution (A)	45.83 - 3.00	EDS
% Data completeness	98.8 (37.39-3.00)	Depositor
(in resolution range)	98.8 (45.83-3.00)	EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$< I/\sigma(I) > 1$	1.71 (at 3.01Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.1_743)	Depositor
P.P.	0.219 , 0.289	Depositor
$R, R_{free}$	0.229 , $0.295$	DCC
$R_{free}$ test set	2838 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	83.5	Xtriage
Anisotropy	0.590	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$ , $B_{sol}(Å^2)$	0.31 , 85.7	EDS
L-test for twinning <sup>2</sup>	$< L >=0.49, < L^2>=0.32$	Xtriage
Estimated twinning fraction	0.033 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	13897	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	109.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 3.52% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;sup>2</sup>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.



<sup>&</sup>lt;sup>1</sup>Intensities estimated from amplitudes.

# 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		
IVIOI		RMSZ	# Z  > 5	RMSZ	# Z  > 5	
1	A	0.46	0/2136	0.62	0/2870	
1	С	0.43	0/2106	0.61	0/2830	
2	В	0.47	0/711	0.68	0/965	
2	D	0.44	0/721	0.66	0/980	
3	Е	0.44	0/4269	0.63	0/5753	
3	F	0.39	0/4184	0.60	0/5639	
All	All	0.43	0/14127	0.62	0/19037	

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 maintenain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
3	Е	0	4
3	F	0	4
All	All	0	9

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 9 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	501	THR	Peptide
3	Е	111	GLU	Peptide
3	Е	432	LEU	Peptide
3	Е	492	THR	Peptide

Continued on next page...



Continued from previous page...

1	N/L-1	Clasica	Dan	Т	C
	MOI	Chain	Res	туре	Group
	3	Е	493	GLY	Peptide

#### 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)	$\mathbf{H}(\mathbf{added})$	Clashes	Symm-Clashes
1	A	2103	0	2127	117	0
1	С	2073	0	2101	118	0
2	В	692	0	647	48	0
2	D	701	0	654	44	0
3	Е	4202	0	4360	271	0
3	F	4120	0	4271	277	0
4	В	3	0	0	0	0
4	D	3	0	0	0	0
All	All	13897	0	14160	820	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 29.

The worst 5 of 820 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	$\begin{array}{c} {\rm Interatomic} \\ {\rm distance} \ ({\rm \AA}) \end{array}$	Clash overlap (Å)
1:C:525:LEU:O	1:C:529:GLU:HB2	1.21	1.37
1:C:530:PRO:O	1:C:561:ARG:NH2	1.68	1.26
1:A:646:SER:OG	1:A:648:LEU:HD12	1.35	1.21
1:A:647:LYS:HB2	1:A:675:TYR:CE1	1.76	1.20
1:C:527:ASN:HD22	3:F:551:PRO:HB3	1.09	1.10

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	Perce	$_{ m ntiles}$
1	A	252/282~(89%)	225 (89%)	26 (10%)	1 (0%)	34	72
1	С	246/282 (87%)	228 (93%)	17 (7%)	1 (0%)	34	72
2	В	80/106 (76%)	72 (90%)	8 (10%)	0	100	100
2	D	83/106 (78%)	76 (92%)	6 (7%)	1 (1%)	13	48
3	E	509/596 (85%)	464 (91%)	44 (9%)	1 (0%)	47	82
3	F	492/596 (83%)	450 (92%)	40 (8%)	2 (0%)	34	72
All	All	1662/1968 (84%)	1515 (91%)	141 (8%)	6 (0%)	34	72

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	С	530	PRO
3	Е	306	VAL
2	D	65	SER
3	F	113	PRO
1	A	489	ILE

#### 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	$237/258 \ (92\%)$	197 (83%)	40 (17%)	2 11
1	С	234/258 (91%)	199 (85%)	35 (15%)	3 14
2	В	74/90 (82%)	60 (81%)	14 (19%)	1 8
2	D	75/90 (83%)	59 (79%)	16 (21%)	1 5
3	E	481/548 (88%)	400 (83%)	81 (17%)	2 11
3	F	471/548 (86%)	391 (83%)	80 (17%)	2 10
All	All	1572/1792 (88%)	1306 (83%)	266 (17%)	2 11



5 of 266 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	F	330	SER
3	F	379	THR
3	F	560	SER
2	D	69	THR
2	D	53	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 38 such sidechains are listed below:

Mol	Chain	Res	Type
3	F	279	ASN
3	F	486	ASN
3	F	296	GLN
3	F	398	GLN
3	F	510	HIS

#### 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 6 ligands modelled in this entry, 6 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,  $95^{th}$  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>	$\#\mathrm{RSRZ}{>}2$	$\mathbf{OWAB}(\mathrm{\AA}^2)$	Q < 0.9
1	A	258/282 (91%)	-0.09	2 (0%) 86 65	51, 88, 168, 204	0
1	С	254/282 (90%)	0.00	5 (1%) 65 36	60, 99, 174, 207	0
2	В	84/106 (79%)	-0.05	2 (2%) 59 30	52, 114, 156, 217	0
2	D	85/106 (80%)	-0.17	0 100 100	69, 100, 152, 169	0
3	E	523/596 (87%)	-0.10	9 (1%) 70 41	49, 99, 165, 213	0
3	F	512/596 (85%)	-0.08	7 (1%) 75 49	54, 119, 170, 212	0
All	All	1716/1968 (87%)	-0.08	25 (1%) 73 46	49, 104, 169, 217	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	F	46	LEU	5.4
3	Е	488	ASN	3.8
1	С	480	ALA	3.7
3	F	307	LEU	3.1
1	A	549	GLN	3.0

#### 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 monosaccharides 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,  $95^{th}$  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	$\mathbf{B} ext{-}\mathbf{factors}(\mathring{\mathbf{A}}^2)$	Q<0.9
4	ZN	В	201	1/1	0.96	0.21	91,91,91,91	0
4	ZN	В	203	1/1	0.98	0.14	112,112,112,112	0
4	ZN	В	202	1/1	0.99	0.18	96,96,96,96	0
4	ZN	D	201	1/1	0.99	0.20	79,79,79,79	0
4	ZN	D	202	1/1	0.99	0.10	120,120,120,120	0
4	ZN	D	203	1/1	0.99	0.19	101,101,101,101	0

#### 6.5 Other polymers (i)

There are no such residues in this entry.

