



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 25, 2023 – 02:17 AM EDT

PDB ID : 5VZW  
Title : TRIM23 RING domain in complex with UbcH5-Ub  
Authors : Pornillos, O.; Dawidziak, D.  
Deposited on : 2017-05-29  
Resolution : 2.28 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 2.35.1  
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) : 2.35.1

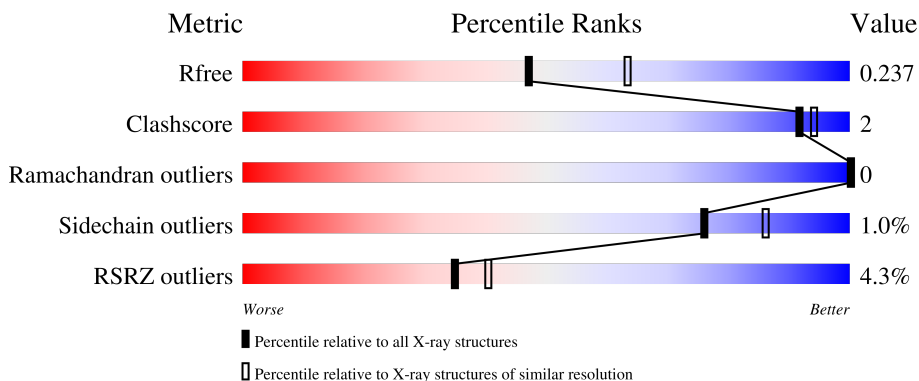
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.28 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.



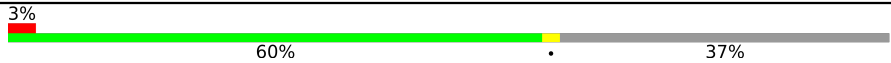
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	6980 (2.30-2.26)
Clashscore	141614	7711 (2.30-2.26)
Ramachandran outliers	138981	7597 (2.30-2.26)
Sidechain outliers	138945	7598 (2.30-2.26)
RSRZ outliers	127900	6849 (2.30-2.26)

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 of 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	147	
1	B	147	
2	C	79	
2	D	79	
3	F	123	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	G	123	 <p>A horizontal bar chart representing the quality of chain. The bar is divided into three segments: a small red segment on the left labeled '3%', a large green segment in the middle labeled '60%', and a grey segment on the right labeled '37%'. A small black dot is visible on the green segment.</p>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 9346 atoms, of which 4610 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 Ubiquitin-conjugating enzyme E2 D2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	146	2264	741	1110	196	211	6	0	0	0
1	B	146	2180	724	1057	188	205	6	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	22	ARG	SER	engineered mutation	UNP P62837
A	85	LYS	CYS	engineered mutation	UNP P62837
B	22	ARG	SER	engineered mutation	UNP P62837
B	85	LYS	CYS	engineered mutation	UNP P62837

- Molecule 2 is a protein called Ubiquitin.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
2	C	78	1255	387	638	109	120	1	0	0	0
2	D	77	1187	372	601	101	112	1	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-2	GLY	-	expression tag	UNP P0CG47
C	-1	SER	-	expression tag	UNP P0CG47
C	0	HIS	-	expression tag	UNP P0CG47
D	-2	GLY	-	expression tag	UNP P0CG47
D	-1	SER	-	expression tag	UNP P0CG47
D	0	HIS	-	expression tag	UNP P0CG47

- Molecule 3 is a protein called E3 ubiquitin-protein ligase TRIM23.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
3	F	77	1199	378	600	109	106	6	0	0	0
3	G	77	1206	379	604	110	107	6	0	0	0

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	F	2	Total	Zn	0	0
			2	2		
4	G	2	Total	Zn	0	0
			2	2		

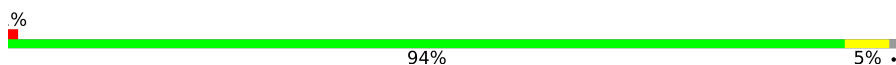
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	13	Total	O	0	0
			13	13		
5	C	8	Total	O	0	0
			8	8		
5	F	15	Total	O	0	0
			15	15		
5	G	5	Total	O	0	0
			5	5		
5	B	4	Total	O	0	0
			4	4		
5	D	6	Total	O	0	0
			6	6		

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

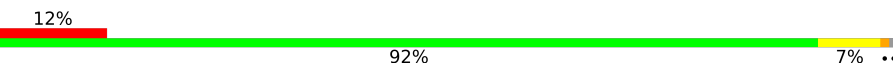
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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.

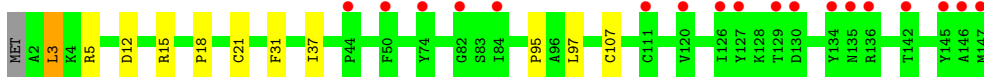
- Molecule 1: Ubiquitin-conjugating enzyme E2 D2

Chain A: 



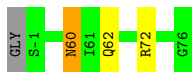
- Molecule 1: Ubiquitin-conjugating enzyme E2 D2

Chain B: 



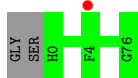
- Molecule 2: Ubiquitin

Chain C: 



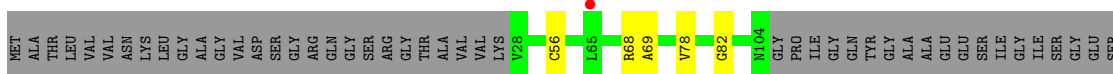
- Molecule 2: Ubiquitin

Chain D: 



- Molecule 3: E3 ubiquitin-protein ligase TRIM23

Chain F: 



- Molecule 3: E3 ubiquitin-protein ligase TRIM23



MET	ALA	THR	LEU	VAL	ASN	LYS	LEU	GLY	GLY	VAL	ASP	SER	GLY	ARG	GLN	GLY	SER	ARG	GLY	THR	ALA	VAL	VAL	LYS	V28	L29	E30	V37	C56	L65	R76	Q103	N104	GLY	PRO	ILE	GLY	GLN	TYR	GLY	ALA	ALA	GLU	GLU	SER	ILE	GLY	ILE	SER	GLY	GLU	SER
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	52.83Å 109.36Å 122.96Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.57 – 2.28 47.57 – 2.28	Depositor EDS
% Data completeness (in resolution range)	99.0 (47.57-2.28) 99.0 (47.57-2.28)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.66 (at 2.27Å)	Xtrriage
Refinement program	PHENIX (1.11.1_2575: ???)	Depositor
R, $R_{free}$	0.208 , 0.237 0.208 , 0.237	Depositor DCC
$R_{free}$ test set	2000 reflections (6.06%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	59.0	Xtrriage
Anisotropy	0.048	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 47.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	9346	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	72.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.26% 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	0.27	0/1190	0.45	0/1627
1	B	0.26	0/1159	0.43	0/1591
2	C	0.26	0/624	0.47	0/839
2	D	0.25	0/592	0.46	0/798
3	F	0.25	0/609	0.46	0/823
3	G	0.24	0/612	0.44	0/827
All	All	0.26	0/4786	0.45	0/6505

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 [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	1154	1110	1111	4	0
1	B	1123	1057	1057	6	0
2	C	617	638	641	2	0
2	D	586	601	604	0	0
3	F	599	600	600	2	0
3	G	602	604	604	2	0
4	F	2	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	G	2	0	0	0	0
5	A	13	0	0	0	0
5	B	4	0	0	0	0
5	C	8	0	0	0	0
5	D	6	0	0	0	0
5	F	15	0	0	0	0
5	G	5	0	0	0	0
All	All	4736	4610	4617	15	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 2.

All (15) 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:F:68:ARG:NH2	3:F:82:GLY:O	2.31	0.60
2:C:60:ASN:OD1	2:C:62:GLN:NE2	2.36	0.58
1:A:12:ASP:OD1	1:A:15:ARG:NH2	2.38	0.56
1:B:12:ASP:OD1	1:B:15:ARG:NH2	2.37	0.56
1:A:5:ARG:NH2	1:A:97:LEU:O	2.38	0.56
1:B:5:ARG:NH1	1:B:95:PRO:O	2.38	0.56
1:B:5:ARG:NH2	1:B:97:LEU:O	2.44	0.51
1:A:5:ARG:NH1	1:A:95:PRO:O	2.45	0.49
1:B:18:PRO:HG2	1:B:21:CYS:SG	2.56	0.46
1:A:37:ILE:HG23	1:A:107:CYS:SG	2.59	0.43
3:G:30:GLU:OE1	3:G:30:GLU:N	2.53	0.42
1:B:37:ILE:HG23	1:B:107:CYS:SG	2.61	0.41
1:B:3:LEU:HD12	1:B:31:PHE:CZ	2.56	0.41
3:F:69:ALA:HB1	3:F:78:VAL:HG13	2.02	0.40
2:C:72:ARG:O	3:G:76:ARG:NH1	2.47	0.40

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	144/147 (98%)	142 (99%)	2 (1%)	0	100	100
1	B	144/147 (98%)	142 (99%)	2 (1%)	0	100	100
2	C	76/79 (96%)	76 (100%)	0	0	100	100
2	D	75/79 (95%)	75 (100%)	0	0	100	100
3	F	75/123 (61%)	72 (96%)	3 (4%)	0	100	100
3	G	75/123 (61%)	73 (97%)	2 (3%)	0	100	100
All	All	589/698 (84%)	580 (98%)	9 (2%)	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	125/131 (95%)	124 (99%)	1 (1%)	81	90
1	B	118/131 (90%)	117 (99%)	1 (1%)	81	90
2	C	70/70 (100%)	69 (99%)	1 (1%)	67	79
2	D	63/70 (90%)	63 (100%)	0	100	100
3	F	67/99 (68%)	66 (98%)	1 (2%)	65	77
3	G	68/99 (69%)	67 (98%)	1 (2%)	65	77
All	All	511/600 (85%)	506 (99%)	5 (1%)	76	86

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	3	LEU
2	C	60	ASN
3	F	56	CYS
3	G	56	CYS
1	B	3	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	C	62	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 4 ligands modelled in this entry, 4 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

### 6.1 Protein, DNA and RNA chains

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	146/147 (99%)	0.29	2 (1%) 75 79	47, 64, 84, 100	0
1	B	146/147 (99%)	0.67	18 (12%) 4 5	48, 72, 93, 105	0
2	C	78/79 (98%)	0.31	0 100 100	45, 61, 79, 86	0
2	D	77/79 (97%)	0.35	1 (1%) 77 81	47, 67, 93, 98	0
3	F	77/123 (62%)	0.31	1 (1%) 77 81	46, 58, 83, 99	0
3	G	77/123 (62%)	0.49	4 (5%) 27 32	44, 57, 85, 95	0
All	All	601/698 (86%)	0.42	26 (4%) 35 40	44, 64, 89, 105	0

All (26) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	126	ILE	4.5
1	B	74	TYR	4.1
1	B	147	MET	4.0
1	B	120	VAL	3.3
1	B	135	ASN	3.2
1	B	134	TYR	3.0
3	G	65	LEU	2.9
3	G	28	VAL	2.9
3	F	65	LEU	2.9
1	B	136	ARG	2.7
1	B	129	THR	2.6
3	G	37	VAL	2.5
1	B	145	TYR	2.4
1	B	82	GLY	2.4
1	B	146	ALA	2.4
1	B	111	CYS	2.4
1	B	84	ILE	2.3
1	B	127	TYR	2.3
3	G	103	GLN	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	107	CYS	2.1
2	D	4	PHE	2.1
1	A	28	ASP	2.0
1	B	142	THR	2.0
1	B	44	PRO	2.0
1	B	130	ASP	2.0
1	B	50	PHE	2.0

## 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<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
4	ZN	F	302	1/1	0.99	0.20	55,55,55,55	0
4	ZN	G	302	1/1	0.99	0.20	55,55,55,55	0
4	ZN	G	301	1/1	1.00	0.20	44,44,44,44	0
4	ZN	F	301	1/1	1.00	0.18	47,47,47,47	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.