



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 08:29 PM GMT

PDB ID : 4S1Z  
Title : Crystal structure of TRABID NZF1 in complex with K29 linked di-Ubiquitin  
Authors : Kristariyanto, Y.A.; Abdul Rehman, S.A.; Campbell, D.G.; Morrice, N.A.;  
Johnson, C.; Toth, R.; Kulathu, Y.  
Deposited on : 2015-01-16  
Resolution : 3.03 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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
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) : trunk26865

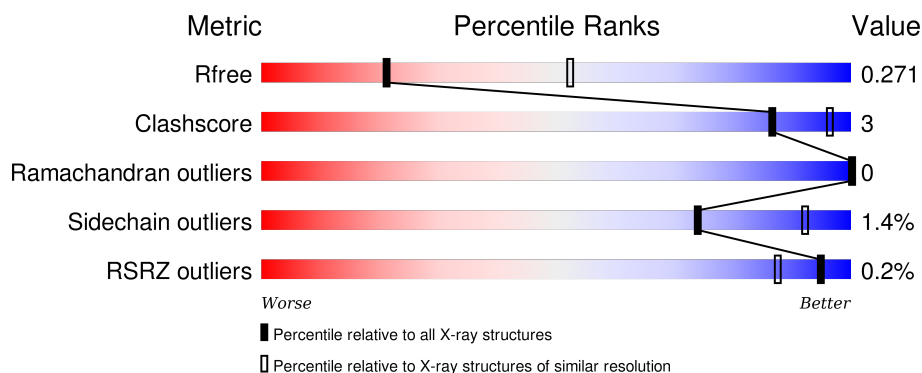
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.03 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	1995 (3.08-3.00)
Clashscore	102246	2351 (3.08-3.00)
Ramachandran outliers	100387	2272 (3.08-3.00)
Sidechain outliers	100360	2275 (3.08-3.00)
RSRZ outliers	91569	2013 (3.08-3.00)

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 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	76	
1	B	76	
1	C	76	
1	D	76	
1	E	76	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	36	<div><div></div><div>53%19%25%</div></div>
2	G	36	<div><div></div><div>61%14%25%</div></div>
2	H	36	<div><div></div><div>3%75%25%</div></div>
2	I	36	<div><div></div><div>69%28%</div></div>
2	J	36	<div><div></div><div>61%6%31%</div></div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3628 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 Ubiquitin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	73	Total	C	N	O	S	0	0	0
			562	355	95	111	1			
1	B	73	Total	C	N	O	S	0	0	0
			567	358	98	110	1			
1	C	73	Total	C	N	O	S	0	0	0
			539	344	92	102	1			
1	D	75	Total	C	N	O		0	0	0
			492	299	93	100				
1	E	72	Total	C	N	O		0	0	0
			511	326	86	99				

- Molecule 2 is a protein called Ubiquitin thioesterase ZRANB1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	27	Total	C	N	O	S	0	0	0
			192	121	31	35	5			
2	G	27	Total	C	N	O	S	0	0	0
			216	137	38	36	5			
2	H	27	Total	C	N	O	S	0	0	0
			197	125	30	37	5			
2	J	25	Total	C	N	O	S	0	0	0
			165	103	28	29	5			
2	I	26	Total	C	N	O	S	0	0	0
			182	115	32	30	5			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	-2	GLY	-	EXPRESSION TAG	UNP A6QP16
F	-1	PRO	-	EXPRESSION TAG	UNP A6QP16
F	0	LEU	-	EXPRESSION TAG	UNP A6QP16
F	1	GLY	-	EXPRESSION TAG	UNP A6QP16
G	-2	GLY	-	EXPRESSION TAG	UNP A6QP16

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
G	-1	PRO	-	EXPRESSION TAG	UNP A6QP16
G	0	LEU	-	EXPRESSION TAG	UNP A6QP16
G	1	GLY	-	EXPRESSION TAG	UNP A6QP16
H	-2	GLY	-	EXPRESSION TAG	UNP A6QP16
H	-1	PRO	-	EXPRESSION TAG	UNP A6QP16
H	0	LEU	-	EXPRESSION TAG	UNP A6QP16
H	1	GLY	-	EXPRESSION TAG	UNP A6QP16
J	-2	GLY	-	EXPRESSION TAG	UNP A6QP16
J	-1	PRO	-	EXPRESSION TAG	UNP A6QP16
J	0	LEU	-	EXPRESSION TAG	UNP A6QP16
J	1	GLY	-	EXPRESSION TAG	UNP A6QP16
I	-2	GLY	-	EXPRESSION TAG	UNP A6QP16
I	-1	PRO	-	EXPRESSION TAG	UNP A6QP16
I	0	LEU	-	EXPRESSION TAG	UNP A6QP16
I	1	GLY	-	EXPRESSION TAG	UNP A6QP16

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

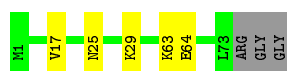
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	H	1	Total Zn 1 1	0	0
3	G	1	Total Zn 1 1	0	0
3	J	1	Total Zn 1 1	0	0
3	I	1	Total Zn 1 1	0	0
3	F	1	Total Zn 1 1	0	0

### 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: Ubiquitin

Chain A:  89% 7% .




- Molecule 1: Ubiquitin

Chain B:  92% . .



- Molecule 1: Ubiquitin

Chain C:  89% 7% .



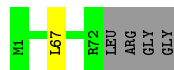
- Molecule 1: Ubiquitin

Chain D:  96% . .



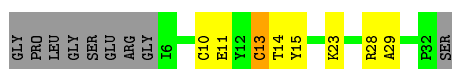
- Molecule 1: Ubiquitin

Chain E:  93% . 5%



- Molecule 2: Ubiquitin thioesterase ZRANB1

Chain F:  53% 19% . 25%



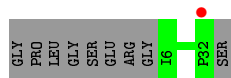
- Molecule 2: Ubiquitin thioesterase ZRANB1

Chain G: 61% 14% 25%



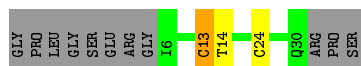
- Molecule 2: Ubiquitin thioesterase ZRANB1

Chain H: 3% 75% 25%



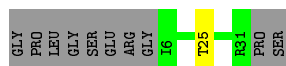
- Molecule 2: Ubiquitin thioesterase ZRANB1

Chain J: 61% 6% 31%



- Molecule 2: Ubiquitin thioesterase ZRANB1

Chain I: 69% 1% 28%



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	99.22Å 123.97Å 78.31Å 90.00° 103.68° 90.00°	Depositor
Resolution (Å)	76.10 – 3.03 29.83 – 3.03	Depositor EDS
% Data completeness (in resolution range)	98.8 (76.10-3.03) 99.0 (29.83-3.03)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.13 (at 3.06Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, $R_{free}$	0.222 , 0.270 0.226 , 0.271	Depositor DCC
$R_{free}$ test set	957 reflections (5.70%)	DCC
Wilson B-factor (Å <sup>2</sup> )	82.2	Xtriage
Anisotropy	0.795	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 78.3	EDS
Estimated twinning fraction	0.019 for -1/2*h+1/2*k-l,1/2*h-1/2*k-l,-1/2 *h-1/2*k 0.003 for -1/2*h-1/2*k-l,-1/2*h-1/2*k+l,-1/ 2*h+1/2*k	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 17755 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	3628	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	89.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.20% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.28	0/568	0.49	0/767
1	B	0.26	0/573	0.45	0/774
1	C	0.26	0/545	0.47	0/740
1	D	0.25	0/495	0.46	0/679
1	E	0.28	0/517	0.46	0/707
2	F	0.29	0/198	0.48	0/272
2	G	0.31	0/223	0.46	0/304
2	H	0.28	0/204	0.42	0/282
2	I	0.29	0/188	0.38	0/259
2	J	0.29	0/170	0.50	0/235
All	All	0.27	0/3681	0.46	0/5019

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	562	0	568	3	0
1	B	567	0	582	2	0
1	C	539	0	534	4	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	492	0	415	3	0
1	E	511	0	480	1	0
2	F	192	0	150	5	0
2	G	216	0	195	3	0
2	H	197	0	151	0	0
2	I	182	0	130	1	0
2	J	165	0	114	3	0
3	F	1	0	0	0	0
3	G	1	0	0	0	0
3	H	1	0	0	0	0
3	I	1	0	0	0	0
3	J	1	0	0	1	0
All	All	3628	0	3319	19	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 3.

The worst 5 of 19 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:C:31:GLN:HE22	1:D:49:GLN:HE22	1.26	0.80
2:J:13:CYS:SG	3:J:101:ZN:ZN	1.86	0.63
2:J:13:CYS:N	2:J:14:THR:HA	2.17	0.59
1:A:63:LYS:HG3	1:A:64:GLU:HG2	1.88	0.55
1:D:68:HIS:HB3	2:J:13:CYS:O	2.07	0.53

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	71/76 (93%)	70 (99%)	1 (1%)	0	100	100
1	B	71/76 (93%)	71 (100%)	0	0	100	100
1	C	71/76 (93%)	70 (99%)	1 (1%)	0	100	100
1	D	73/76 (96%)	70 (96%)	3 (4%)	0	100	100
1	E	70/76 (92%)	70 (100%)	0	0	100	100
2	F	25/36 (69%)	21 (84%)	4 (16%)	0	100	100
2	G	25/36 (69%)	24 (96%)	1 (4%)	0	100	100
2	H	25/36 (69%)	25 (100%)	0	0	100	100
2	I	24/36 (67%)	22 (92%)	2 (8%)	0	100	100
2	J	23/36 (64%)	22 (96%)	1 (4%)	0	100	100
All	All	478/560 (85%)	465 (97%)	13 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains ⓘ

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	61/68 (90%)	61 (100%)	0	100	100
1	B	63/68 (93%)	63 (100%)	0	100	100
1	C	55/68 (81%)	54 (98%)	1 (2%)	66	90
1	D	40/68 (59%)	40 (100%)	0	100	100
1	E	49/68 (72%)	49 (100%)	0	100	100
2	F	17/30 (57%)	16 (94%)	1 (6%)	24	61
2	G	22/30 (73%)	21 (96%)	1 (4%)	34	72
2	H	18/30 (60%)	18 (100%)	0	100	100
2	I	13/30 (43%)	13 (100%)	0	100	100
2	J	12/30 (40%)	10 (83%)	2 (17%)	3	12
All	All	350/490 (71%)	345 (99%)	5 (1%)	74	92

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

Mol	Chain	Res	Type
1	C	40	GLN
2	F	13	CYS
2	G	20	SER
2	J	13	CYS
2	J	24	CYS

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

Mol	Chain	Res	Type
1	B	31	GLN
1	D	49	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 5 ligands modelled in this entry, 5 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	73/76 (96%)	-0.59	0 100 100	62, 73, 86, 93	0
1	B	73/76 (96%)	-0.58	0 100 100	65, 78, 95, 104	0
1	C	73/76 (96%)	-0.55	0 100 100	24, 81, 93, 98	0
1	D	75/76 (98%)	-0.31	0 100 100	80, 102, 122, 128	0
1	E	72/76 (94%)	-0.57	0 100 100	89, 103, 115, 117	0
2	F	27/36 (75%)	0.14	0 100 100	96, 111, 149, 155	0
2	G	27/36 (75%)	-0.49	0 100 100	67, 76, 105, 115	0
2	H	27/36 (75%)	-0.40	1 (3%) 45 19	76, 90, 108, 120	0
2	I	26/36 (72%)	-0.69	0 100 100	70, 76, 93, 102	0
2	J	25/36 (69%)	-0.45	0 100 100	95, 105, 115, 116	0
All	All	498/560 (88%)	-0.48	1 (0%) 95 87	24, 87, 117, 155	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	32	PRO	2.3

### 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(Å <sup>2</sup> )	Q<0.9
3	ZN	G	101	1/1	1.00	0.12	-0.95	69,69,69,69	0
3	ZN	I	101	1/1	0.99	0.10	-1.09	75,75,75,75	0
3	ZN	H	101	1/1	0.99	0.07	-1.32	75,75,75,75	0
3	ZN	F	101	1/1	0.98	0.09	-1.52	106,106,106,106	0
3	ZN	J	101	1/1	0.99	0.09	-2.05	108,108,108,108	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.