



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 11, 2024 – 03:45 PM EST

PDB ID : 8UQE  
Title : Crystal structure of RNF168 (RING)-UbcH5c fused to H2A-H2B via a 26-residue linker (RING not modeled in density)  
Authors : Hu, Q.; Botuyan, M.V.; Mer, G.  
Deposited on : 2023-10-23  
Resolution : 3.56 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 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 : 2.36  
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) : 2.36

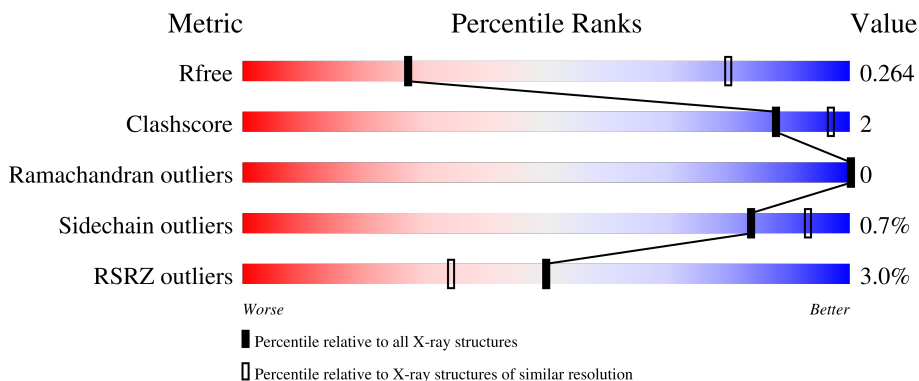
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.56 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	1020 (3.62-3.50)
Clashscore	141614	1100 (3.62-3.50)
Ramachandran outliers	138981	1065 (3.62-3.50)
Sidechain outliers	138945	1066 (3.62-3.50)
RSRZ outliers	127900	1009 (3.64-3.48)

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	B	461	

## 2 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 2553 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 E3 ubiquitin-protein ligase RNF168, Ubiquitin-conjugating enzyme E2 D3, Histone H2B type 2-E, Histone H2A type 1-B/E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	B	331	2553	1618	451	476	8	10	0	0

There are 38 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-1	GLY	-	expression tag	UNP Q8IYW5
B	0	HIS	-	expression tag	UNP Q8IYW5
B	95	SER	-	linker	UNP Q8IYW5
B	96	GLY	-	linker	UNP Q8IYW5
B	97	SER	-	linker	UNP Q8IYW5
B	98	GLY	-	linker	UNP Q8IYW5
B	99	SER	-	linker	UNP Q8IYW5
B	100	GLY	-	linker	UNP Q8IYW5
B	101	SER	-	linker	UNP Q8IYW5
B	1022	ARG	SER	engineered mutation	UNP P61077
B	1085	LYS	CYS	engineered mutation	UNP P61077
B	1148	GLY	-	linker	UNP P61077
B	1149	SER	-	linker	UNP P61077
B	1150	GLY	-	linker	UNP P61077
B	1151	SER	-	linker	UNP P61077
B	1152	GLY	-	linker	UNP P61077
B	1153	SER	-	linker	UNP P61077
B	1154	GLY	-	linker	UNP P61077
B	1155	GLY	-	linker	UNP P61077
B	1156	SER	-	linker	UNP P61077
B	1157	GLY	-	linker	UNP P61077
B	1158	GLY	-	linker	UNP P61077
B	1159	GLY	-	linker	UNP P61077
B	1160	SER	-	linker	UNP P61077
B	1161	GLY	-	linker	UNP P61077
B	1162	GLY	-	linker	UNP P61077

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	1163	SER	-	linker	UNP P61077
B	1164	GLY	-	linker	UNP P61077
B	1165	GLY	-	linker	UNP P61077
B	1166	SER	-	linker	UNP P61077
B	1167	GLY	-	linker	UNP P61077
B	1168	GLY	-	linker	UNP P61077
B	1169	SER	-	linker	UNP P61077
B	1170	GLY	-	linker	UNP P61077
B	1171	SER	-	linker	UNP P61077
B	1172	GLY	-	linker	UNP P61077
B	1173	SER	-	linker	UNP P61077
B	3011	SER	-	linker	UNP Q16778



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	109.28Å 109.28Å 86.78Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	41.54 – 3.56 41.54 – 3.56	Depositor EDS
% Data completeness (in resolution range)	99.4 (41.54-3.56) 99.4 (41.54-3.56)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.53 (at 3.57Å)	Xtrriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, $R_{free}$	0.249 , 0.268 0.250 , 0.264	Depositor DCC
$R_{free}$ test set	724 reflections (9.79%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	110.0	Xtrriage
Anisotropy	0.704	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 75.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	0.062 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	2553	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	135.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.07% 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](#)

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	B	0.25	0/2608	0.47	0/3545

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	B	2553	0	2532	10	0
All	All	2553	0	2532	10	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.

The worst 5 of 10 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:B:2098:VAL:HG13	1:B:2102:LEU:HD22	1.83	0.58
1:B:2069:ILE:HD13	1:B:3058:LEU:HD13	1.87	0.56
1:B:1013:LEU:O	1:B:1013:LEU:HD23	2.14	0.47
1:B:1087:ASP:OD1	1:B:1088:ILE:N	2.51	0.44
1:B:1098:THR:O	1:B:1102:VAL:HG23	2.17	0.44

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	B	327/461 (71%)	323 (99%)	4 (1%)	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	B	271/386 (70%)	269 (99%)	2 (1%)	84 93

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	2042	TYR
1	B	3031	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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](#)

There are no ligands in this entry.

#### 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	B	331/461 (71%)	0.16	10 (3%) 50 34	110, 134, 157, 180	5 (1%)

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1173	SER	4.0
1	B	1050	PHE	2.8
1	B	1069	PHE	2.6
1	B	1045	TYR	2.6
1	B	1018	PRO	2.6

### 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](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.