



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2020 – 07:56 am BST

PDB ID : 5OHN  
Title : Crystal structure of USP30 in covalent complex with ubiquitin propargylamide (low resolution)  
Authors : Gersch, M.; Komander, D.  
Deposited on : 2017-07-17  
Resolution : 3.60 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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
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) : 2.11

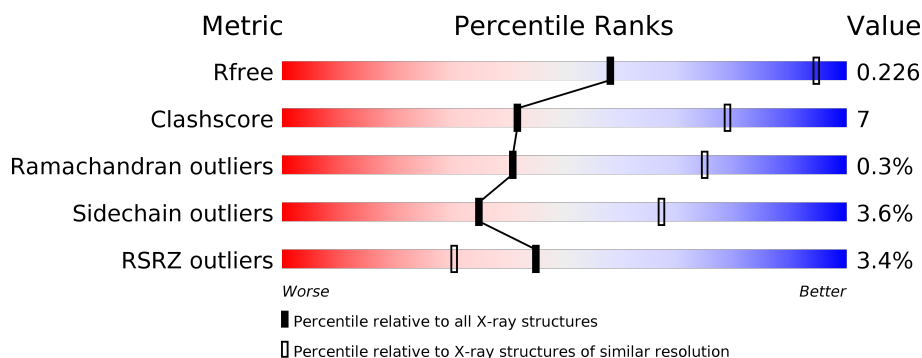
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.60 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	1257 (3.70-3.50)
Clashscore	141614	1353 (3.70-3.50)
Ramachandran outliers	138981	1307 (3.70-3.50)
Sidechain outliers	138945	1307 (3.70-3.50)
RSRZ outliers	127900	1161 (3.70-3.50)

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 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	370	<div> <div>3%</div> <div> <div></div> <div>68%</div> <div>12%</div> <div>19%</div> </div> </div>
1	C	370	<div> <div>3%</div> <div> <div></div> <div>66%</div> <div>13%</div> <div>20%</div> </div> </div>
2	B	76	<div> <div></div> <div> <div></div> <div>89%</div> <div>9%</div> </div> </div>
2	D	76	<div> <div>5%</div> <div> <div></div> <div>86%</div> <div>13%</div> </div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 5789 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 carboxyl-terminal hydrolase 30, Ubiquitin carboxyl-terminal hydrolase 30.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	299	Total	C	N	O	S	0	0	0
			2331	1480	404	430	17			
1	C	297	Total	C	N	O	S	0	0	0
			2312	1462	404	429	17			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	62	GLY	-	expression tag	UNP Q70CQ3
A	63	PRO	-	expression tag	UNP Q70CQ3
A	348	ASP	PHE	engineered mutation	UNP Q70CQ3
A	350	ASP	MET	engineered mutation	UNP Q70CQ3
A	353	GLU	ILE	engineered mutation	UNP Q70CQ3
A	358	SER	-	linker	UNP Q70CQ3
A	359	ASN	-	linker	UNP Q70CQ3
A	360	ALA	-	linker	UNP Q70CQ3
C	62	GLY	-	expression tag	UNP Q70CQ3
C	63	PRO	-	expression tag	UNP Q70CQ3
C	348	ASP	PHE	engineered mutation	UNP Q70CQ3
C	350	ASP	MET	engineered mutation	UNP Q70CQ3
C	353	GLU	ILE	engineered mutation	UNP Q70CQ3
C	358	SER	-	linker	UNP Q70CQ3
C	359	ASN	-	linker	UNP Q70CQ3
C	360	ALA	-	linker	UNP Q70CQ3

- Molecule 2 is a protein called Polyubiquitin-B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	76	Total	C	N	O	S	0	0	0
			572	360	99	112	1			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	76	Total	C	N	O	S	0	0	0
			572	360	99	112	1			

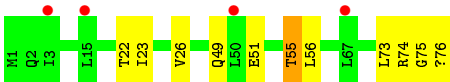
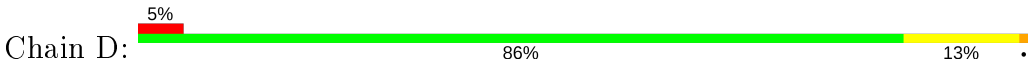
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	76	AYE	GLY	engineered mutation	UNP P0CG47
D	76	AYE	GLY	engineered mutation	UNP P0CG47

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Zn	0	0
			1	1		
3	C	1	Total	Zn	0	0
			1	1		





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	181.83 Å 181.83 Å 94.96 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	157.47 – 3.60 157.47 – 3.60	Depositor EDS
% Data completeness (in resolution range)	99.3 (157.47-3.60) 99.3 (157.47-3.60)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.33 (at 3.58 Å)	Xtriage
Refinement program	PHENIX (1.12rc2_2821: ???)	Depositor
R, $R_{free}$	0.211 , 0.253 0.204 , 0.226	Depositor DCC
$R_{free}$ test set	972 reflections (4.68%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	147.1	Xtriage
Anisotropy	0.229	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 140.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.38$ , $\langle L^2 \rangle = 0.20$	Xtriage
Estimated twinning fraction	0.398 for h,-h-k,-l	Xtriage
Reported twinning fraction	0.521 for H, K, L 0.479 for K, H, -L	Depositor
Outliers	0 of 20774 reflections	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	5789	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	165.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.15% 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: AYE, 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	0.59	0/2388	0.75	0/3245
1	C	0.56	0/2365	0.75	0/3210
2	B	0.43	0/574	0.69	0/778
2	D	0.42	0/574	0.69	0/778
All	All	0.55	0/5901	0.74	0/8011

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	2331	0	2144	32	3
1	C	2312	0	2139	42	3
2	B	572	0	569	5	0
2	D	572	0	569	17	0
3	A	1	0	0	0	0
3	C	1	0	0	0	0
All	All	5789	0	5421	78	3

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 7.



The worst 5 of 78 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:159:GLU:O	2:D:74:ARG:CD	1.80	1.26
1:C:159:GLU:O	2:D:74:ARG:HD3	1.11	1.26
1:C:262:HIS:CB	1:C:263:PRO:HD2	1.69	1.21
1:C:262:HIS:CB	1:C:263:PRO:CD	2.22	1.17
1:C:262:HIS:HB3	1:C:263:PRO:CD	1.73	1.17

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:356:TYR:OH	1:C:356:TYR:OH[5_555]	1.67	0.53
1:A:148:ARG:NH1	1:C:145:ASP:OD2[4_654]	2.16	0.04
1:A:148:ARG:NH2	1:C:145:ASP:CG[4_654]	2.19	0.01

## 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	287/370 (78%)	268 (93%)	18 (6%)	1 (0%)	41	75
1	C	285/370 (77%)	267 (94%)	17 (6%)	1 (0%)	34	71
2	B	73/76 (96%)	72 (99%)	1 (1%)	0	100	100
2	D	73/76 (96%)	72 (99%)	1 (1%)	0	100	100
All	All	718/892 (80%)	679 (95%)	37 (5%)	2 (0%)	41	75

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	262	HIS
1	C	262	HIS

### 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	247/343 (72%)	238 (96%)	9 (4%)	35	67
1	C	247/343 (72%)	236 (96%)	11 (4%)	27	62
2	B	61/68 (90%)	60 (98%)	1 (2%)	62	83
2	D	61/68 (90%)	60 (98%)	1 (2%)	62	83
All	All	616/822 (75%)	594 (96%)	22 (4%)	35	67

5 of 22 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	55	THR
1	C	232	MET
1	C	454	VAL
1	C	76	THR
1	C	170	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	487	GLN
2	B	25	ASN
1	C	286	ASN
1	C	487	GLN
2	D	25	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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 2 ligands modelled in this entry, 2 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	299/370 (80%)	0.14	11 (3%) 41 27	101, 157, 225, 277	0
1	C	297/370 (80%)	0.18	10 (3%) 45 30	112, 157, 226, 267	0
2	B	75/76 (98%)	-0.03	0 100 100	127, 165, 209, 221	0
2	D	75/76 (98%)	0.08	4 (5%) 26 16	123, 159, 199, 233	0
All	All	746/892 (83%)	0.13	25 (3%) 45 30	101, 158, 222, 277	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	152	TRP	4.7
1	A	502	VAL	3.9
1	A	120	LEU	3.8
1	C	105	ARG	3.4
1	A	317	LEU	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 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( $\text{\AA}^2$ )	Q<0.9
3	ZN	A	800	1/1	0.99	0.33	107,107,107,107	0
3	ZN	C	800	1/1	0.99	0.32	116,116,116,116	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.