



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 8, 2019 – 01:12 AM EDT

PDB ID : 6H4J  
Title : Usp25 catalytic domain  
Authors : Klemm, T.A.; Sauer, F.; Kisker, C.  
Deposited on : 2018-07-21  
Resolution : 3.07 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  
Xtriage (Phenix) : 1.13  
EDS : rb-20031633  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20031633

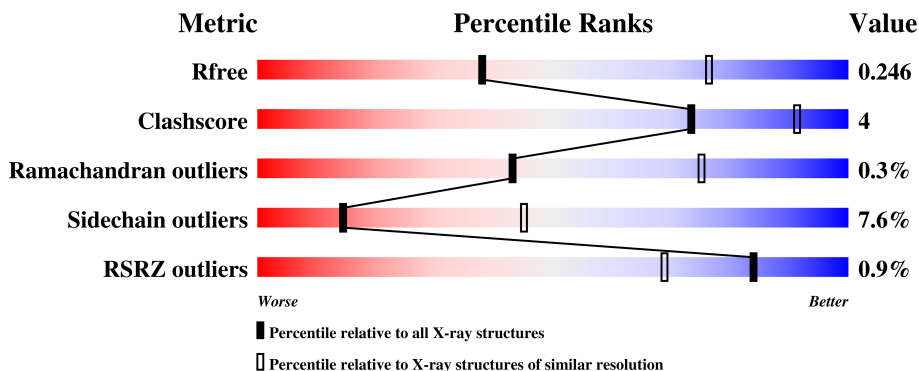
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.07 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))
$R_{free}$	111664	1238 (3.10-3.06)
Clashscore	122126	1332 (3.10-3.06)
Ramachandran outliers	120053	1287 (3.10-3.06)
Sidechain outliers	120020	1286 (3.10-3.06)
RSRZ outliers	108989	1212 (3.10-3.06)

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	552	<div> <div></div> <div>75%13%12%</div> </div>
1	B	552	<div> <div></div> <div>71%12%16%</div> </div>

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 7588 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 25.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	488	Total	C	N	O	S	0	0	0
			3931	2504	681	730	16			
1	B	464	Total	C	N	O	S	0	1	0
			3625	2317	617	677	14			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	155	GLY	-	expression tag	UNP Q9UHP3
A	156	PRO	-	expression tag	UNP Q9UHP3
B	155	GLY	-	expression tag	UNP Q9UHP3
B	156	PRO	-	expression tag	UNP Q9UHP3

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Cl	0	0
			1	1		
2	A	2	Total	Cl	0	0
			2	2		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	11	Total	O	0	0
			11	11		
3	B	18	Total	O	0	0
			18	18		



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	81.94Å 202.63Å 169.59Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.54 – 3.07 48.54 – 3.07	Depositor EDS
% Data completeness (in resolution range)	100.0 (48.54-3.07) 100.0 (48.54-3.07)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.47 (at 3.07Å)	Xtriage
Refinement program	BUSTER 2.10.2	Depositor
R, $R_{free}$	0.181 , 0.230 0.200 , 0.246	Depositor DCC
$R_{free}$ test set	1365 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	86.6	Xtriage
Anisotropy	0.248	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 76.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	7588	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	104.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.06% 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: CL

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	0.51	0/4031	0.67	0/5466
1	B	0.49	0/3718	0.65	0/5059
All	All	0.50	0/7749	0.66	0/10525

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	3931	0	3699	29	0
1	B	3625	0	3302	23	0
2	A	2	0	0	0	0
2	B	1	0	0	0	0
3	A	11	0	0	0	0
3	B	18	0	0	0	0
All	All	7588	0	7001	52	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 4.

The worst 5 of 52 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:A:328:ASN:CG	1:A:329:GLY:H	1.72	0.91
1:A:661:PHE:HA	1:A:664:GLN:HG2	1.66	0.78
1:A:328:ASN:CG	1:A:329:GLY:N	2.45	0.69
1:A:215:GLU:HG3	1:A:217:ARG:H	1.58	0.67
1:A:602:GLN:HB2	1:A:605:ALA:HB3	1.76	0.66

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	478/552 (87%)	450 (94%)	28 (6%)	0	100	100
1	B	451/552 (82%)	420 (93%)	28 (6%)	3 (1%)	24	59
All	All	929/1104 (84%)	870 (94%)	56 (6%)	3 (0%)	43	75

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	625	ASP
1	B	235	THR
1	B	689	VAL

### 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	412/501 (82%)	387 (94%)	25 (6%)	20	54
1	B	361/501 (72%)	327 (91%)	34 (9%)	9	34
All	All	773/1002 (77%)	714 (92%)	59 (8%)	14	45

5 of 59 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	161	ARG
1	B	316	ASN
1	B	665	GLU
1	B	163	ARG
1	B	190	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	359	GLN
1	B	594	HIS
1	A	408	ASN
1	A	185	GLN
1	B	279	GLN

### 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 ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 3 ligands modelled in this entry, 3 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	488/552 (88%)	-0.36	5 (1%) 82 66	50, 92, 149, 182	0
1	B	464/552 (84%)	-0.30	4 (0%) 84 69	54, 109, 172, 202	0
All	All	952/1104 (86%)	-0.33	9 (0%) 84 69	50, 99, 166, 202	0

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	257	ASP	3.8
1	A	513	SER	3.6
1	A	258	SER	2.7
1	B	309	LEU	2.6
1	B	311	GLY	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 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
2	CL	B	801	1/1	0.81	0.12	125,125,125,125	0
2	CL	A	801	1/1	0.83	0.18	111,111,111,111	0
2	CL	A	802	1/1	0.91	0.35	104,104,104,104	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.