



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 19, 2016 – 07:44 PM GMT

PDB ID : 4WPH
Title : Crystal structure of USP7 ubiquitin-like domains in compact conformation
Authors : Pfoh, R.; Lacdao, I.; Saridakis, V.
Deposited on : 2014-10-18
Resolution : 2.92 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
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 : unknown
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026982
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) : rb-20026982

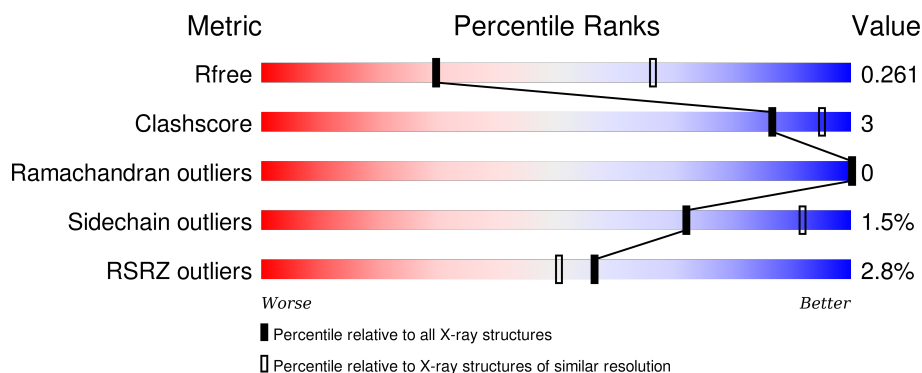
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.92 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	1643 (2.94-2.90)
Clashscore	102246	1871 (2.94-2.90)
Ramachandran outliers	100387	1824 (2.94-2.90)
Sidechain outliers	100360	1826 (2.94-2.90)
RSRZ outliers	91569	1650 (2.94-2.90)

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 ≥ 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	375	<div> <div>2%</div> <div>83%7%10%</div> </div>
1	B	375	<div> <div>3%</div> <div>81%6%13%</div> </div>
2	C	11	<div> <div>73%18%9%</div> </div>
2	D	11	<div> <div>9%</div> <div>73%9%18%</div> </div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 5582 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 7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	337	Total	C	N	O	S	0	2	0
			2748	1743	474	514	17			
1	B	326	Total	C	N	O	S	0	0	0
			2635	1674	445	500	16			

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	514	MET	-	initiating methionine	UNP Q93009
A	515	GLY	-	expression tag	UNP Q93009
A	516	SER	-	expression tag	UNP Q93009
A	517	SER	-	expression tag	UNP Q93009
A	518	HIS	-	expression tag	UNP Q93009
A	519	HIS	-	expression tag	UNP Q93009
A	520	HIS	-	expression tag	UNP Q93009
A	521	HIS	-	expression tag	UNP Q93009
A	522	HIS	-	expression tag	UNP Q93009
A	523	HIS	-	expression tag	UNP Q93009
A	524	SER	-	expression tag	UNP Q93009
A	525	SER	-	expression tag	UNP Q93009
A	526	GLY	-	expression tag	UNP Q93009
A	527	ARG	-	expression tag	UNP Q93009
A	528	GLU	-	expression tag	UNP Q93009
A	529	ASN	-	expression tag	UNP Q93009
A	530	LEU	-	expression tag	UNP Q93009
A	531	TYR	-	expression tag	UNP Q93009
A	532	PHE	-	expression tag	UNP Q93009
A	533	GLN	-	expression tag	UNP Q93009
A	534	GLY	-	expression tag	UNP Q93009
B	514	MET	-	initiating methionine	UNP Q93009
B	515	GLY	-	expression tag	UNP Q93009
B	516	SER	-	expression tag	UNP Q93009
B	517	SER	-	expression tag	UNP Q93009

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	518	HIS	-	expression tag	UNP Q93009
B	519	HIS	-	expression tag	UNP Q93009
B	520	HIS	-	expression tag	UNP Q93009
B	521	HIS	-	expression tag	UNP Q93009
B	522	HIS	-	expression tag	UNP Q93009
B	523	HIS	-	expression tag	UNP Q93009
B	524	SER	-	expression tag	UNP Q93009
B	525	SER	-	expression tag	UNP Q93009
B	526	GLY	-	expression tag	UNP Q93009
B	527	ARG	-	expression tag	UNP Q93009
B	528	GLU	-	expression tag	UNP Q93009
B	529	ASN	-	expression tag	UNP Q93009
B	530	LEU	-	expression tag	UNP Q93009
B	531	TYR	-	expression tag	UNP Q93009
B	532	PHE	-	expression tag	UNP Q93009
B	533	GLN	-	expression tag	UNP Q93009
B	534	GLY	-	expression tag	UNP Q93009

- Molecule 2 is a protein called ICP0.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	10	Total 80	C 47	N 21	O 11	S 1	0	0	0
2	D	9	Total 76	C 45	N 20	O 10	S 1	0	0	0

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Cl	0	0
			1	1		
3	A	3	Total	Cl	0	0
			3	3		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	21	Total	O	0	0
			21	21		
4	B	17	Total	O	0	0
			17	17		

Continued on next page...

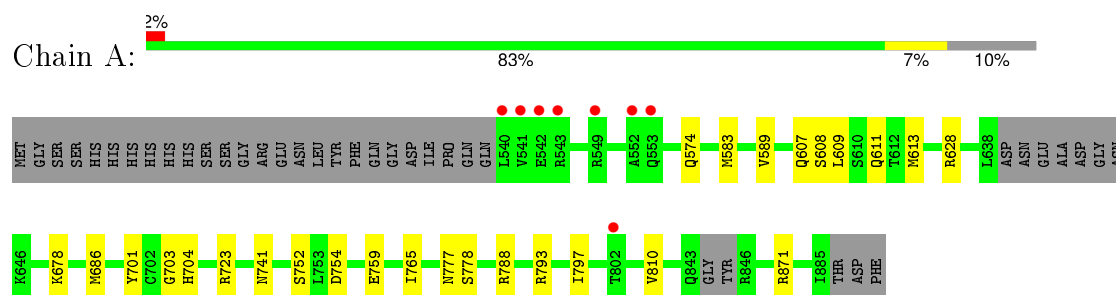
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	C	1	Total	O	0	0
			1	1		

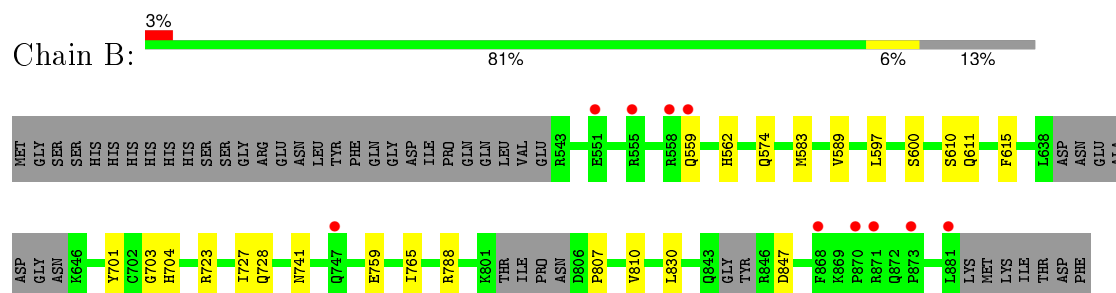
3 Residue-property plots [i](#)

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.

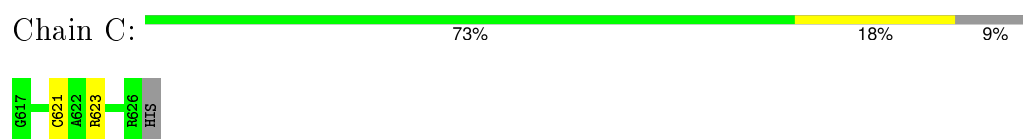
- Molecule 1: Ubiquitin carboxyl-terminal hydrolase 7



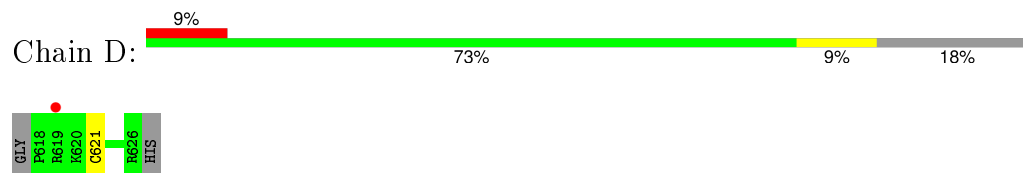
- Molecule 1: Ubiquitin carboxyl-terminal hydrolase 7



- Molecule 2: ICP0



- Molecule 2: ICP0



4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α , β , γ	92.34Å 92.34Å 190.28Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	30.00 – 2.92 30.61 – 2.92	Depositor EDS
% Data completeness (in resolution range)	99.5 (30.00-2.92) 99.6 (30.61-2.92)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.54 (at 2.90Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.233 , 0.262 0.232 , 0.261	Depositor DCC
R_{free} test set	1041 reflections (5.22%)	DCC
Wilson B-factor (Å ²)	65.6	Xtriage
Anisotropy	0.189	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 42.1	EDS
Estimated twinning fraction	0.027 for -h,-k,l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 20978 reflections	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	5582	wwPDB-VP
Average B, all atoms (Å ²)	77.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.48% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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

5.1 Standard geometry

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	0.48	0/2806	0.69	0/3788
1	B	0.47	0/2689	0.66	0/3633
2	C	0.60	0/80	0.85	0/103
2	D	0.42	0/76	0.78	0/97
All	All	0.47	0/5651	0.68	0/7621

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

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	2748	0	2696	18	0
1	B	2635	0	2547	14	0
2	C	80	0	91	1	0
2	D	76	0	89	1	0
3	A	3	0	0	0	0
3	B	1	0	0	0	0
4	A	21	0	0	1	0
4	B	17	0	0	0	0
4	C	1	0	0	0	0
All	All	5582	0	5423	30	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.

All (30) 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:628:ARG:NH2	4:A:2111:HOH:O	2.38	0.56
1:B:704:HIS:O	1:B:723:ARG:NH2	2.41	0.54
1:A:752:SER:OG	1:A:754:ASP:OD1	2.18	0.53
1:A:704:HIS:O	1:A:723:ARG:NH2	2.42	0.52
1:B:701:TYR:CE2	1:B:703:GLY:HA2	2.45	0.51
1:B:759:GLU:HA	2:C:621:CYS:SG	2.50	0.51
1:A:701:TYR:CE2	1:A:703:GLY:HA2	2.47	0.50
1:A:793[B]:ARG:HG2	1:A:793[B]:ARG:HH11	1.79	0.48
1:B:574:GLN:HB3	1:B:589:VAL:HG23	1.95	0.48
1:A:793[A]:ARG:HG2	1:A:793[A]:ARG:HH11	1.80	0.47
1:A:607:GLN:NE2	1:B:807:PRO:O	2.49	0.45
1:A:574:GLN:HB3	1:A:589:VAL:HG23	1.99	0.45
1:B:559:GLN:O	1:B:562:HIS:CD2	2.69	0.45
1:A:609:LEU:HD11	1:A:613:MET:HE3	1.99	0.44
1:B:741:ASN:OD1	1:B:788:ARG:HG2	2.18	0.44
1:A:793[B]:ARG:NH1	1:A:793[B]:ARG:HG2	2.33	0.44
1:A:797:ILE:HG13	1:A:810:VAL:HG22	1.99	0.44
1:B:583:MET:HE2	1:B:765:ILE:HD11	2.00	0.43
1:B:597:LEU:HB2	1:B:600:SER:HB3	2.00	0.43
1:A:793[A]:ARG:NH1	1:A:793[A]:ARG:HG2	2.34	0.43
1:A:741:ASN:OD1	1:A:788:ARG:HG2	2.20	0.42
1:A:611:GLN:HG3	1:B:810:VAL:HG21	2.01	0.41
1:A:759:GLU:HA	2:D:621:CYS:SG	2.60	0.41
1:A:583:MET:HE1	1:A:765:ILE:HG12	2.02	0.41
1:B:610:SER:HB2	1:B:615:PHE:O	2.21	0.41
1:A:871[B]:ARG:HH11	1:A:871[B]:ARG:HG3	1.86	0.41
1:B:701:TYR:CZ	1:B:703:GLY:HA2	2.56	0.41
1:A:583:MET:HE2	1:A:765:ILE:HD11	2.02	0.41
1:B:727:ILE:HG22	1:B:728:GLN:O	2.21	0.40
1:B:583:MET:HE1	1:B:765:ILE:HG12	2.03	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	333/375 (89%)	325 (98%)	8 (2%)	0	100	100
1	B	318/375 (85%)	312 (98%)	6 (2%)	0	100	100
2	C	8/11 (73%)	8 (100%)	0	0	100	100
2	D	7/11 (64%)	6 (86%)	1 (14%)	0	100	100
All	All	666/772 (86%)	651 (98%)	15 (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	303/343 (88%)	298 (98%)	5 (2%)	68	90
1	B	287/343 (84%)	284 (99%)	3 (1%)	82	95
2	C	8/9 (89%)	7 (88%)	1 (12%)	6	16
2	D	8/9 (89%)	8 (100%)	0	100	100
All	All	606/704 (86%)	597 (98%)	9 (2%)	72	92

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	608	SER
1	A	678	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	686	MET
1	A	777	ASN
1	A	778	SER
1	B	611	GLN
1	B	830	LEU
1	B	847	ASP
2	C	623	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	607	GLN
1	A	843	GLN
1	B	562	HIS
1	B	843	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 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

There are no such residues in this entry.

5.8 Polymer linkage issues

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, 95th 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 ²)	Q<0.9
1	A	337/375 (89%)	-0.05	8 (2%) 62 57	35, 73, 126, 181	0
1	B	326/375 (86%)	0.01	10 (3%) 52 46	38, 71, 133, 157	0
2	C	10/11 (90%)	-0.31	0 100 100	54, 61, 78, 99	0
2	D	9/11 (81%)	0.27	1 (11%) 7 4	88, 93, 112, 117	0
All	All	682/772 (88%)	-0.02	19 (2%) 56 50	35, 73, 133, 181	0

All (19) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	542	GLU	4.3
1	B	873	PRO	4.2
1	B	747	GLN	4.1
1	B	871	ARG	3.6
1	B	555	ARG	3.1
1	B	870	PRO	3.1
1	A	549	ARG	2.9
1	A	543	ARG	2.7
1	B	559	GLN	2.5
1	A	553	GLN	2.3
1	B	881	LEU	2.3
1	B	551	GLU	2.3
1	A	541	VAL	2.2
1	B	558	ARG	2.2
2	D	619	ARG	2.2
1	A	552	ALA	2.2
1	B	868	PHE	2.0
1	A	540	LEU	2.0
1	A	802	THR	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 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, 95th 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(\AA^2)	Q<0.9
3	CL	B	2001	1/1	0.78	0.32	0.42	90,90,90,90	0
3	CL	A	2002	1/1	0.92	0.30	-	74,74,74,74	0
3	CL	A	2001	1/1	0.77	0.31	-	89,89,89,89	0
3	CL	A	2003	1/1	0.92	0.12	-	82,82,82,82	0

6.5 Other polymers [i](#)

There are no such residues in this entry.