WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 11:09 am GMT

PDB ID	:	2QJ0
Title	:	Structure of the yeast U-box-containing ubiquitin ligase Ufd2p
Authors	:	Tu, D.; Brunger, A.T.
Deposited on	:	2007-07-06
Resolution	:	2.65 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*
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:

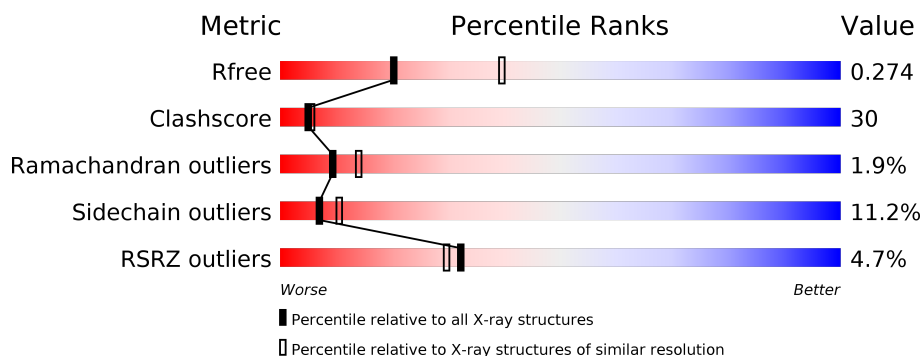|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

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 2.65 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	Percentile Ranks	Value
Rfree		0.274
Clashscore		30
Ramachandran outliers		1.9%
Sidechain outliers		11.2%
RSRZ outliers		4.7%

Worse   *Better*

▌Percentile relative to all X-ray structures

▯Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	100719	3491 (2.70-2.62)
Clashscore	112137	1026 (2.68-2.64)
Ramachandran outliers	110173	1010 (2.68-2.64)
Sidechain outliers	110143	1010 (2.68-2.64)
RSRZ outliers	101464	3511 (2.70-2.62)

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 <=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	982	4% 49% 39% 7% 5%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 7661 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 conjugation factor E4.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	937	Total 7567	C 4867	N 1247	O 1425	S 10	Se 18	0	0	0

There are 41 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-20	GLY	-	EXPRESSION TAG	UNP P54860
A	-19	SER	-	EXPRESSION TAG	UNP P54860
A	-18	HIS	-	EXPRESSION TAG	UNP P54860
A	-17	MSE	-	EXPRESSION TAG	UNP P54860
A	-16	ALA	-	EXPRESSION TAG	UNP P54860
A	-15	SER	-	EXPRESSION TAG	UNP P54860
A	-14	MSE	-	EXPRESSION TAG	UNP P54860
A	-13	THR	-	EXPRESSION TAG	UNP P54860
A	-12	GLY	-	EXPRESSION TAG	UNP P54860
A	-11	GLY	-	EXPRESSION TAG	UNP P54860
A	-10	GLN	-	EXPRESSION TAG	UNP P54860
A	-9	GLN	-	EXPRESSION TAG	UNP P54860
A	-8	MSE	-	EXPRESSION TAG	UNP P54860
A	-7	GLY	-	EXPRESSION TAG	UNP P54860
A	-6	ARG	-	EXPRESSION TAG	UNP P54860
A	-5	GLY	-	EXPRESSION TAG	UNP P54860
A	-4	SER	-	EXPRESSION TAG	UNP P54860
A	-3	GLU	-	EXPRESSION TAG	UNP P54860
A	-2	PHE	-	EXPRESSION TAG	UNP P54860
A	-1	ARG	-	EXPRESSION TAG	UNP P54860
A	0	SER	-	EXPRESSION TAG	UNP P54860
A	1	MSE	MET	MODIFIED RESIDUE	UNP P54860
A	102	LEU	SER	ENGINEERED	UNP P54860
A	109	MSE	MET	MODIFIED RESIDUE	UNP P54860
A	248	MSE	MET	MODIFIED RESIDUE	UNP P54860
A	282	MSE	MET	MODIFIED RESIDUE	UNP P54860
A	312	MSE	MET	MODIFIED RESIDUE	UNP P54860

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	407	MSE	MET	MODIFIED RESIDUE	UNP P54860
A	442	MSE	MET	MODIFIED RESIDUE	UNP P54860
A	564	MSE	MET	MODIFIED RESIDUE	UNP P54860
A	590	MSE	MET	MODIFIED RESIDUE	UNP P54860
A	600	MSE	MET	MODIFIED RESIDUE	UNP P54860
A	601	MSE	MET	MODIFIED RESIDUE	UNP P54860
A	677	VAL	ASP	ENGINEERED	UNP P54860
A	680	MSE	MET	MODIFIED RESIDUE	UNP P54860
A	745	MSE	MET	MODIFIED RESIDUE	UNP P54860
A	771	MSE	MET	MODIFIED RESIDUE	UNP P54860
A	890	MSE	MET	MODIFIED RESIDUE	UNP P54860
A	894	MSE	MET	MODIFIED RESIDUE	UNP P54860
A	905	MSE	MET	MODIFIED RESIDUE	UNP P54860
A	927	MSE	MET	MODIFIED RESIDUE	UNP P54860

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	94	Total 94	O 94	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes 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 conjugation factor E4

M890 Y891 T892 I893 M894 K895 D896 P897 V898 I899 L900 S903 K904 M905 N906 I907 D908 R909 S910 T911 I912 K913 A914 H915 L916 D919 S920 T921 D922 P923 F924 N925 R926 M927 P928 L929 K930 L931 E932 D933 V934 T935 P936 N937 L940 R941 Q942 K943 I944 L945 C946 F947 K948 K949 Q950 K951

LYS GLU GLU ALA LYS HIS LYS ALA SER GLU
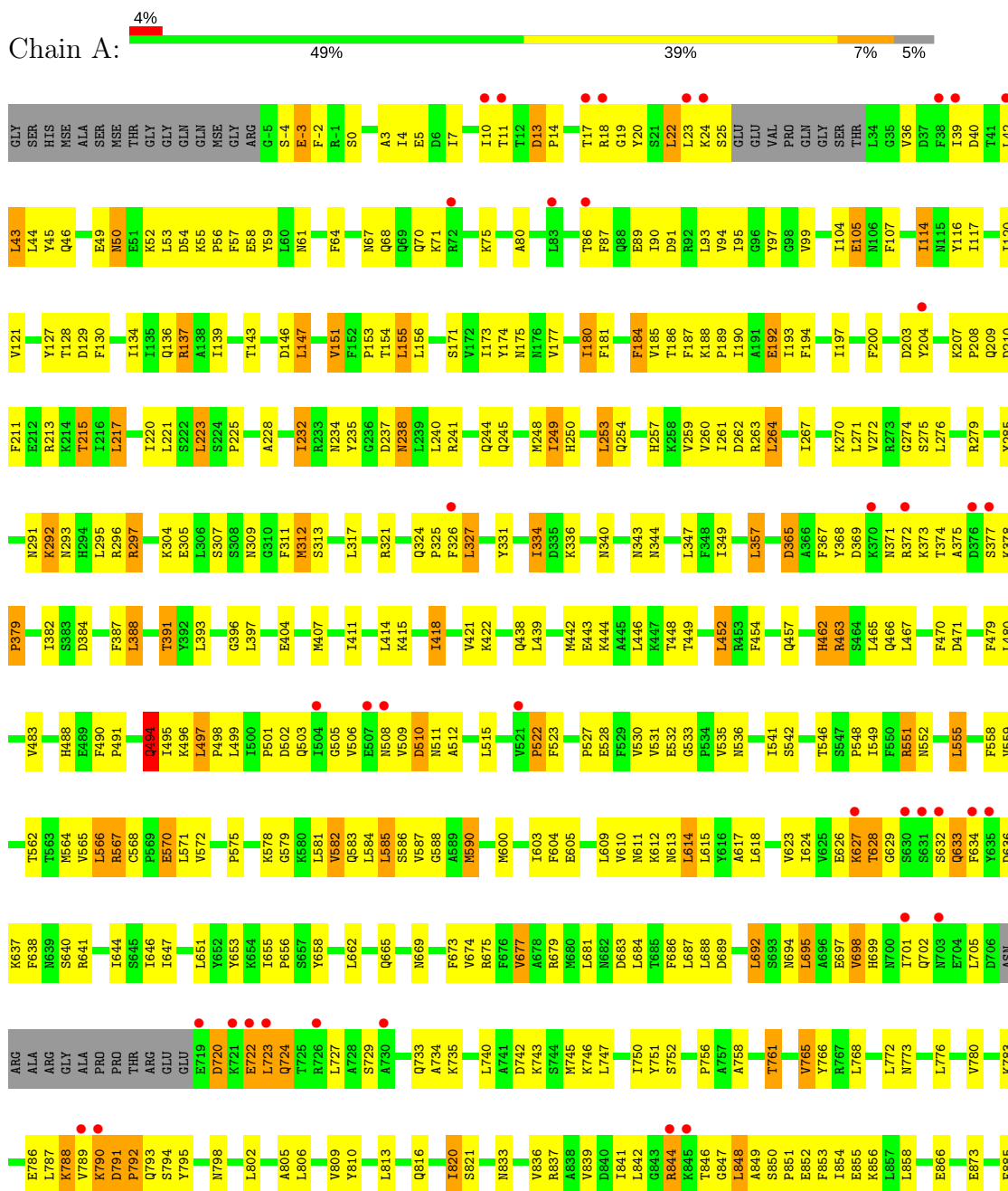
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	65.80Å 122.80Å 178.86Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.65 48.66 – 2.65	Depositor EDS
% Data completeness (in resolution range)	87.4 (50.00-2.65) 84.4 (48.66-2.65)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
$<I/\sigma(I)>$ [1]	2.24 (at 2.65Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.234 , 0.269 0.238 , 0.274	Depositor DCC
R_{free} test set	3580 reflections (9.89%)	DCC
Wilson B-factor (Å^2)	59.9	Xtriage
Anisotropy	0.334	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 36.3	EDS
L-test for twinning[2]	$<|L|> = 0.49, <L^2> = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	7661	wwPDB-VP
Average B, all atoms (Å^2)	76.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.64% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $<|L|>$, $<L^2>$ 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	A	0.42	0/7709	0.65	2/10390 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	633	GLN	N-CA-C	-6.06	94.64	111.00
1	A	723	LEU	CA-CB-CG	5.34	127.58	115.30

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	7567	0	7560	453	0
2	A	94	0	0	12	0
All	All	7661	0	7560	453	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 30.

The worst 5 of 453 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:114:ILE:H	1:A:114:ILE:HD12	1.10	1.13
1:A:575:PRO:HG3	1:A:628:THR:HG21	1.22	1.10
1:A:848:LEU:H	1:A:848:LEU:HD23	1.24	1.02
1:A:14:PRO:HA	1:A:22:LEU:HD21	1.45	0.99
1:A:745:MSE:HE1	1:A:806:LEU:HD13	1.41	0.98

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	931/982 (95%)	792 (85%)	121 (13%)	18 (2%)	9	14

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	849	ALA
1	A	931	LEU
1	A	950	GLN
1	A	105	GLU
1	A	494	GLN

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.

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	848/861 (98%)	753 (89%)	95 (11%)	7	10

5 of 95 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	393	LEU
1	A	497	LEU
1	A	848	LEU
1	A	418	ILE
1	A	463	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	277	ASN
1	A	494	GLN
1	A	906	ASN
1	A	293	ASN
1	A	462	HIS

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 ⓘ

There are no ligands in this entry.

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	919/982 (93%)	0.05	43 (4%)	32	30	38, 71, 125, 151	0

The worst 5 of 43 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	504	ILE	7.5
1	A	24	LYS	5.9
1	A	789	VAL	5.7
1	A	376	ASP	5.6
1	A	38	PHE	5.3

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.