Full wwPDB X-ray Structure Validation Report ⓘ

May 13, 2020 – 04:13 am BST

PDB ID	:	5WWO
Title	:	Crystal structure of Enp1
Authors	:	Ye, K.; Zhang, W.
Deposited on	:	2017-01-03
Resolution	:	2.40 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
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:

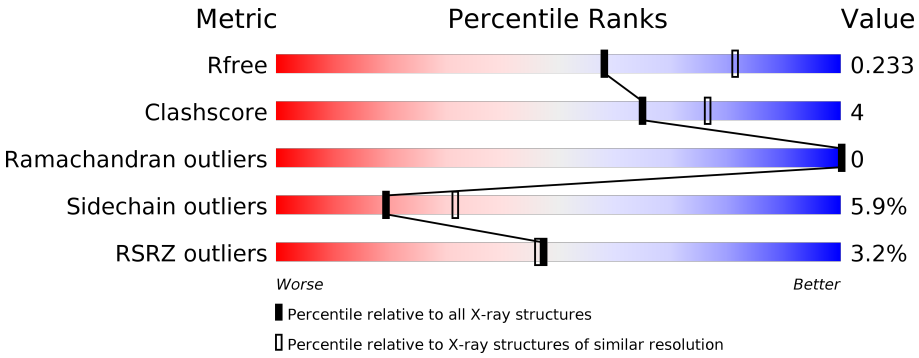
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

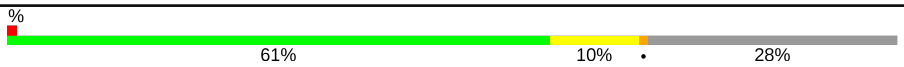
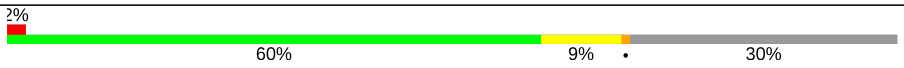
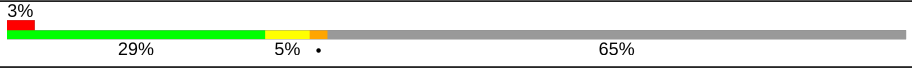
The reported resolution of this entry is 2.40 Å.

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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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 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 <=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	363	
1	B	363	
2	C	131	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 4680 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 Essential nuclear protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	261	Total	C	N	O	S	0	0	0
			2125	1384	363	375	3			
1	B	254	Total	C	N	O	S	0	0	0
			2082	1361	355	363	3			

- Molecule 2 is a protein called Protein LTV1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	46	Total	C	N	O	S	0	0	0
			396	242	57	95	2			

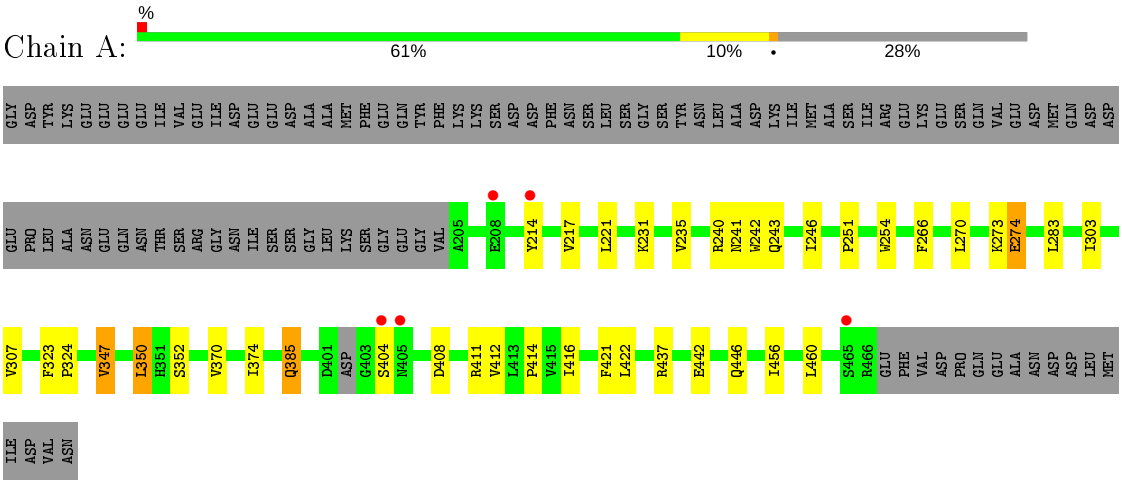
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	39	Total	O	0	0
			39	39		
3	B	37	Total	O	0	0
			37	37		
3	C	1	Total	O	0	0
			1	1		

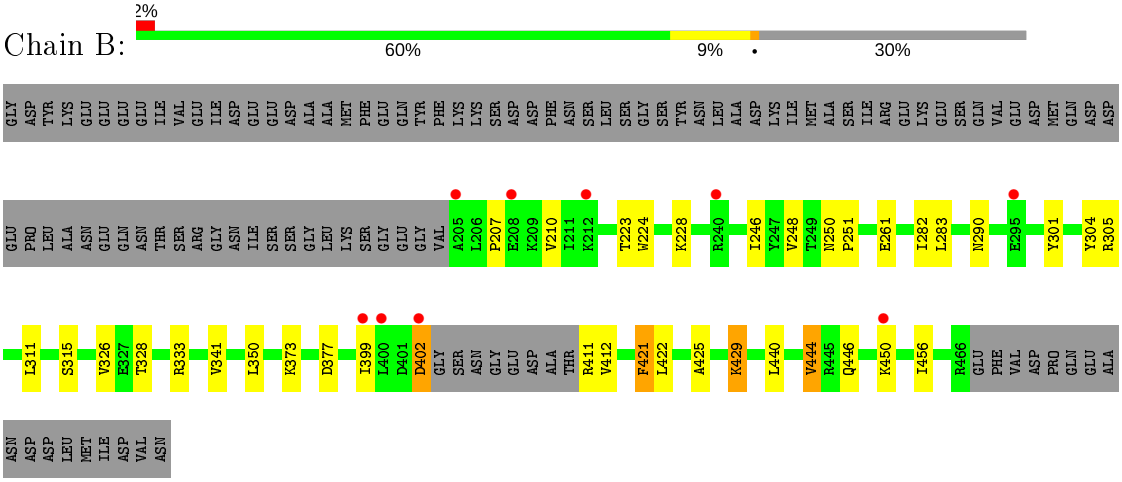
3 Residue-property plots (i)

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

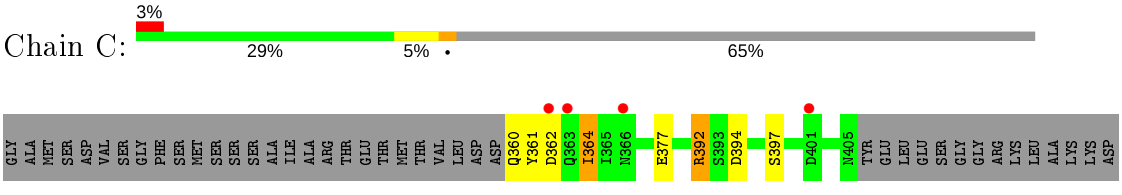
- Molecule 1: Essential nuclear protein 1



- Molecule 1: Essential nuclear protein 1



- Molecule 2: Protein LTV1



LYS GLU ILE GLU ARG LEU LYS GLU ALA ALA ASP GLU VAL SER LYS GLY LYS LEU SER GLN ARG ARG ASN ARG GLU ARG GLN GLU LYS LYS LYS LEU GLU LYS VAL THR ASN THR LEU SER SER LEU LYS PHE

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	69.22Å 98.61Å 109.73Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.75 – 2.40 19.75 – 2.40	Depositor EDS
% Data completeness (in resolution range)	96.7 (19.75-2.40) 96.7 (19.75-2.40)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	6.37 (at 2.41Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.167 , 0.230 0.176 , 0.233	Depositor DCC
R_{free} test set	1476 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å^2)	37.6	Xtriage
Anisotropy	0.087	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 43.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	4680	wwPDB-VP
Average B, all atoms (Å^2)	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.42% 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

5.1 Standard geometry

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	0.44	0/2179	0.54	0/2956
1	B	0.44	0/2136	0.56	0/2898
2	C	0.38	0/403	0.47	0/543
All	All	0.44	0/4718	0.54	0/6397

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	2125	0	2167	22	0
1	B	2082	0	2132	15	0
2	C	396	0	322	5	0
3	A	39	0	0	0	0
3	B	37	0	0	0	0
3	C	1	0	0	0	0
All	All	4680	0	4621	38	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.

All (38) 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:214:TYR:HA	1:A:217:VAL:HG12	1.64	0.78
1:A:242:TRP:HE1	1:A:274:GLU:HG3	1.57	0.69
2:C:392:ARG:HD3	2:C:394:ASP:OD1	1.96	0.65
1:A:270:LEU:HB3	1:A:274:GLU:HB3	1.80	0.63
1:B:328:THR:O	1:B:328:THR:HG22	1.99	0.62
1:A:442:GLU:HG3	1:A:446:GLN:HE21	1.64	0.62
1:B:246:ILE:HG21	1:B:282:ILE:HD11	1.82	0.62
1:A:414:PRO:HB2	1:A:416:ILE:HG22	1.85	0.59
1:A:242:TRP:NE1	1:A:274:GLU:HG3	2.19	0.56
1:A:347:VAL:HG22	1:A:352:SER:HB2	1.88	0.55
1:A:350:LEU:HD13	2:C:364:ILE:HD12	1.89	0.53
1:A:350:LEU:HD21	2:C:361:TYR:HA	1.92	0.52
1:B:399:ILE:HB	1:B:446:GLN:OE1	2.12	0.49
1:B:301:TYR:CZ	1:B:305:ARG:HD2	2.48	0.48
1:A:385:GLN:HG3	2:C:361:TYR:CE2	2.47	0.48
1:A:370:VAL:O	1:A:374:ILE:HG12	2.14	0.47
1:B:373:LYS:NZ	1:B:377:ASP:OD2	2.45	0.47
1:B:207:PRO:HB2	1:B:210:VAL:HG23	1.96	0.47
1:B:402:ASP:N	1:B:402:ASP:OD1	2.47	0.47
1:B:425:ALA:O	1:B:429:LYS:HB3	2.16	0.45
1:B:228:LYS:NZ	2:C:377:GLU:HG3	2.33	0.44
1:B:304:TYR:CE1	1:B:341:VAL:HG22	2.52	0.44
1:A:266:PHE:HD2	1:A:270:LEU:HD11	1.82	0.44
1:A:251:PRO:HA	1:A:254:TRP:CE2	2.53	0.44
1:A:231:LYS:O	1:A:235:VAL:HG13	2.18	0.43
1:A:408:ASP:HA	1:A:411:ARG:HD2	2.00	0.43
1:A:243:GLN:O	1:A:246:ILE:HG22	2.19	0.43
1:B:440:LEU:O	1:B:444:VAL:HG22	2.19	0.43
1:A:240:ARG:HD2	1:A:241:ASN:OD1	2.18	0.43
1:A:323:PHE:HB2	1:A:324:PRO:HD3	2.00	0.43
1:A:437:ARG:HD3	1:A:460:LEU:O	2.19	0.42
1:B:224:TRP:CH2	1:B:261:GLU:HG3	2.55	0.42
1:A:303:ILE:O	1:A:307:VAL:HG23	2.18	0.42
1:A:422:LEU:HA	1:A:456:ILE:HG12	2.02	0.41
1:B:250:ASN:N	1:B:251:PRO:HD3	2.36	0.41
1:A:217:VAL:O	1:A:221:LEU:HG	2.21	0.40
1:B:421:PHE:CE2	1:B:444:VAL:HG13	2.56	0.40
1:B:422:LEU:HA	1:B:456:ILE:HG12	2.02	0.40

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	257/363 (71%)	252 (98%)	5 (2%)	0	100	100
1	B	250/363 (69%)	242 (97%)	8 (3%)	0	100	100
2	C	44/131 (34%)	44 (100%)	0	0	100	100
All	All	551/857 (64%)	538 (98%)	13 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	233/323 (72%)	224 (96%)	9 (4%)	32	50
1	B	229/323 (71%)	213 (93%)	16 (7%)	15	24
2	C	44/118 (37%)	39 (89%)	5 (11%)	5	7
All	All	506/764 (66%)	476 (94%)	30 (6%)	19	32

All (30) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	273	LYS
1	A	274	GLU
1	A	283	LEU
1	A	347	VAL
1	A	350	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	385	GLN
1	A	404	SER
1	A	412	VAL
1	A	421	PHE
1	B	223	THR
1	B	248	VAL
1	B	283	LEU
1	B	290	ASN
1	B	311	LEU
1	B	315	SER
1	B	326	VAL
1	B	333	ARG
1	B	350	LEU
1	B	402	ASP
1	B	411	ARG
1	B	412	VAL
1	B	421	PHE
1	B	429	LYS
1	B	444	VAL
1	B	450	LYS
2	C	360	GLN
2	C	362	ASP
2	C	364	ILE
2	C	392	ARG
2	C	397	SER

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

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, 95^{th} 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	261/363 (71%)	-0.23	5 (1%)	66	64	26, 38, 61, 76	0
1	B	254/363 (69%)	-0.22	9 (3%)	44	43	27, 39, 67, 89	0
2	C	46/131 (35%)	0.60	4 (8%)	10	9	41, 61, 74, 90	0
All	All	561/857 (65%)	-0.15	18 (3%)	47	46	26, 39, 66, 90	0

All (18) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	205	ALA	5.1
2	C	363	GLN	3.7
1	A	405	ASN	3.4
1	B	208	GLU	3.0
1	B	240	ARG	2.6
1	A	465	SER	2.5
1	B	400	LEU	2.4
1	B	399	ILE	2.4
1	B	295	GLU	2.3
2	C	362	ASP	2.3
1	B	450	LYS	2.3
1	A	208	GLU	2.3
2	C	366	ASN	2.2
1	B	212	LYS	2.1
1	A	404	SER	2.1
1	B	402	ASP	2.1
2	C	401	ASP	2.1
1	A	214	TYR	2.0

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 (i)

There are no carbohydrates 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.