Full wwPDB X-ray Structure Validation Report ⓘ

Nov 17, 2021 – 01:04 pm GMT

PDB ID	:	7NY6
Title	:	Crystal structure of the Capsaspora owczarzaki macroH2A macrodomain
Authors	:	Knobloch, G.; Guberovic, I.; Basquin, J.; Buschbeck, M.; Ladurner, A.G.
Deposited on	:	2021-03-21
Resolution	:	1.34 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:

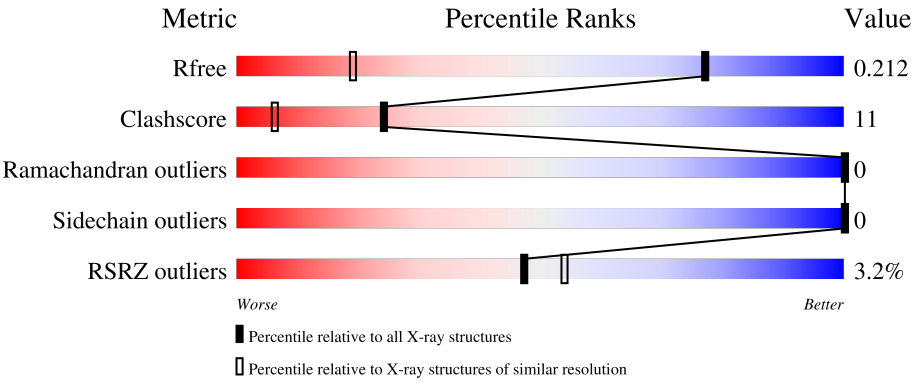
MolProbity	:	4.02b-467
Mogul	:	1.8.4 (270009), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 1.34 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	1385 (1.36-1.32)
Clashscore	141614	1417 (1.36-1.32)
Ramachandran outliers	138981	1397 (1.36-1.32)
Sidechain outliers	138945	1397 (1.36-1.32)
RSRZ outliers	127900	1369 (1.36-1.32)

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 of 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	211	3% 76% 12% • 10%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	EDO	A	408	-	-	X	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	PEG	A	406	-	-	-	X

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 2858 atoms, of which 1218 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.

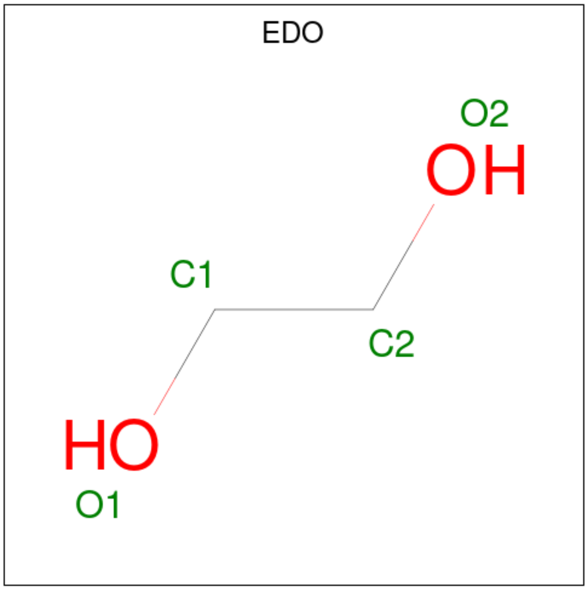
- Molecule 1 is a protein called Histone macroH2A1.1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	189	2585	894	1166	244	276	5	0	1	0

There are 24 discrepancies between the modelled and reference sequences:

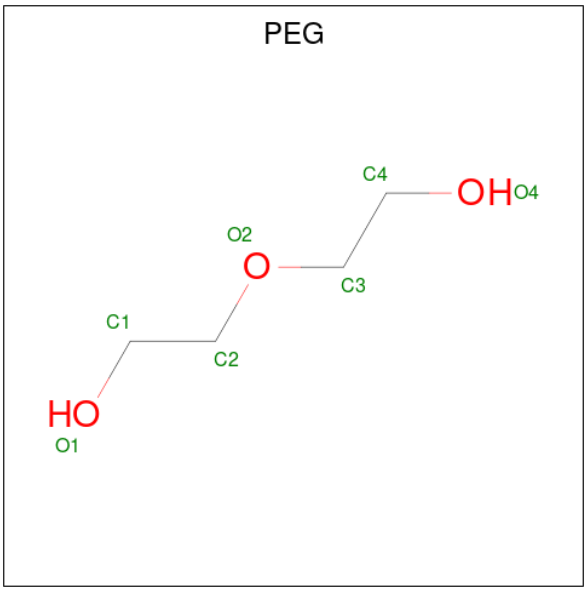
Chain	Residue	Modelled	Actual	Comment	Reference
A	158	MET	-	initiating methionine	UNP A0A0D2UG83
A	159	LYS	-	expression tag	UNP A0A0D2UG83
A	160	HIS	-	expression tag	UNP A0A0D2UG83
A	161	HIS	-	expression tag	UNP A0A0D2UG83
A	162	HIS	-	expression tag	UNP A0A0D2UG83
A	163	HIS	-	expression tag	UNP A0A0D2UG83
A	164	HIS	-	expression tag	UNP A0A0D2UG83
A	165	HIS	-	expression tag	UNP A0A0D2UG83
A	166	PRO	-	expression tag	UNP A0A0D2UG83
A	167	MET	-	expression tag	UNP A0A0D2UG83
A	168	SER	-	expression tag	UNP A0A0D2UG83
A	169	ASP	-	expression tag	UNP A0A0D2UG83
A	170	TYR	-	expression tag	UNP A0A0D2UG83
A	171	ASP	-	expression tag	UNP A0A0D2UG83
A	172	ILE	-	expression tag	UNP A0A0D2UG83
A	173	PRO	-	expression tag	UNP A0A0D2UG83
A	174	THR	-	expression tag	UNP A0A0D2UG83
A	175	THR	-	expression tag	UNP A0A0D2UG83
A	176	GLU	-	expression tag	UNP A0A0D2UG83
A	177	ASN	-	expression tag	UNP A0A0D2UG83
A	178	LEU	-	expression tag	UNP A0A0D2UG83
A	179	TYR	-	expression tag	UNP A0A0D2UG83
A	180	PHE	-	expression tag	UNP A0A0D2UG83
A	181	GLN	-	expression tag	UNP A0A0D2UG83

- Molecule 2 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C H O 10 2 6 2	0	0
2	A	1	Total C H O 10 2 6 2	0	0
2	A	1	Total C H O 10 2 6 2	0	0
2	A	1	Total C H O 10 2 6 2	0	0
2	A	1	Total C H O 10 2 6 2	0	0
2	A	1	Total C H O 10 2 6 2	0	0
2	A	1	Total C H O 10 2 6 2	0	0

- Molecule 3 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	H	O	0	0
			17	4	10	3		

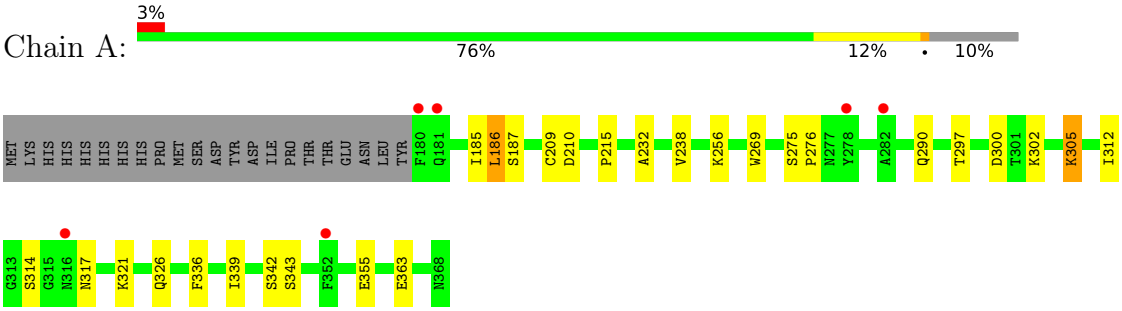
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	186	Total	O	0	0
			186	186		

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: Histone macroH2A1.1



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	41.03Å 44.07Å 44.86Å 90.00° 92.36° 90.00°	Depositor
Resolution (Å)	41.00 – 1.34 44.82 – 1.34	Depositor EDS
% Data completeness (in resolution range)	87.7 (41.00-1.34) 86.8 (44.82-1.34)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.02 (at 1.34Å)	Xtriage
Refinement program	PHENIX (1.18.2_3874: ???)	Depositor
R, R_{free}	0.183 , 0.212 0.183 , 0.212	Depositor DCC
R_{free} test set	1577 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	16.3	Xtriage
Anisotropy	0.208	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.016 for -h,l,k 0.034 for -h,-l,-k 0.047 for h,-k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	2858	wwPDB-VP
Average B, all atoms (Å^2)	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.62% 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

Bond lengths and bond angles in the following residue types are not validated in this section: EDO, PEG

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	0.49	0/1446	0.75	3/1970 (0.2%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	186	LEU	CB-CG-CD1	8.68	125.76	111.00
1	A	300	ASP	CB-CG-OD2	6.00	123.70	118.30
1	A	305	LYS	CG-CD-CE	5.65	128.84	111.90

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	1419	1166	1396	30	0
2	A	28	42	42	13	0
3	A	7	10	10	2	0
4	A	186	0	0	8	0
All	All	1640	1218	1448	32	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 11.

All (32) 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:187:SER:O	4:A:501:HOH:O	1.79	1.00
1:A:290:GLN:NE2	4:A:503:HOH:O	2.06	0.89
1:A:314:SER:HB2	1:A:321:LYS:NZ	1.91	0.86
1:A:355:GLU:HG2	2:A:408:EDO:H22	1.59	0.84
1:A:238:VAL:HG23	4:A:513:HOH:O	1.91	0.71
1:A:314:SER:HB2	1:A:321:LYS:HZ2	1.64	0.62
1:A:305:LYS:HD3	1:A:343:SER:O	1.99	0.61
1:A:355:GLU:HA	2:A:408:EDO:H11	1.83	0.61
1:A:185:ILE:HG23	4:A:501:HOH:O	2.03	0.57
1:A:297:THR:HG22	2:A:404:EDO:C2	2.35	0.56
1:A:314:SER:O	4:A:505:HOH:O	2.18	0.56
1:A:232:ALA:HB2	2:A:401:EDO:H22	1.87	0.55
1:A:317:ASN:OD1	2:A:403:EDO:H22	2.05	0.55
1:A:297:THR:HG22	2:A:404:EDO:H22	1.89	0.55
1:A:256:LYS:NZ	4:A:507:HOH:O	2.19	0.53
1:A:339:ILE:HD11	1:A:342:SER:HB2	1.91	0.53
1:A:355:GLU:HG2	2:A:408:EDO:C2	2.35	0.52
1:A:314:SER:HB2	1:A:321:LYS:HZ3	1.70	0.51
2:A:405:EDO:O2	4:A:506:HOH:O	2.18	0.50
1:A:209:CYS:HA	3:A:406:PEG:H41	1.96	0.47
1:A:339:ILE:CD1	1:A:342:SER:HB2	2.45	0.47
1:A:215:PRO:HG3	1:A:312:ILE:HG12	1.98	0.46
1:A:355:GLU:HG3	2:A:408:EDO:C1	2.46	0.46
1:A:355:GLU:CG	2:A:408:EDO:C1	2.95	0.44
1:A:297:THR:HG22	2:A:404:EDO:H21	2.00	0.43
1:A:269:TRP:CZ2	1:A:302:LYS:HE3	2.54	0.43
1:A:210:ASP:HA	3:A:406:PEG:H22	2.02	0.42
1:A:326:GLN:HG2	1:A:363:GLU:OE1	2.20	0.42
1:A:336:PHE:O	2:A:405:EDO:H11	2.20	0.41
1:A:186:LEU:HD22	1:A:186:LEU:N	2.36	0.41
2:A:407:EDO:H11	4:A:632:HOH:O	2.21	0.41
1:A:275:SER:OG	1:A:276:PRO:HD2	2.21	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	188/211 (89%)	184 (98%)	4 (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	153/177 (86%)	153 (100%)	0	100 100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes 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 (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

8 ligands are modelled in this entry.

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

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	189/211 (89%)	0.27	6 (3%) 47 54	12, 18, 31, 51	0

All (6) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	278	TYR	3.6
1	A	181	GLN	3.5
1	A	282	ALA	3.1
1	A	352	PHE	2.4
1	A	316	ASN	2.3
1	A	180	PHE	2.2

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 monosaccharides in this entry.

6.4 Ligands

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^{th} 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(Å^2)	Q<0.9
2	EDO	A	407	4/4	0.49	0.19	37,44,53,60	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	EDO	A	408	4/4	0.58	0.25	33,49,59,64	0
2	EDO	A	402	4/4	0.66	0.19	38,46,50,59	0
2	EDO	A	404	4/4	0.69	0.37	31,42,53,64	0
2	EDO	A	405	4/4	0.69	0.16	26,37,38,44	0
2	EDO	A	401	4/4	0.79	0.20	31,37,51,61	0
3	PEG	A	406	7/7	0.79	0.41	34,43,59,59	0
2	EDO	A	403	4/4	0.93	0.15	19,33,49,59	0

6.5 Other polymers (i)

There are no such residues in this entry.