wwPDB X-ray Structure Validation Summary Report (i)

May 19, 2020 – 05:43 am BST

PDB ID	:	4N4H
Title	:	Crystal structure of the Bromo-PWWP of the mouse zinc finger MYND-type containing 11 isoform alpha in complex with histone H3.1K36me3
Authors	:	Li, Y.; Ren, Y.; Li, H.
Deposited on	:	2013-10-08
Resolution	:	2.30 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

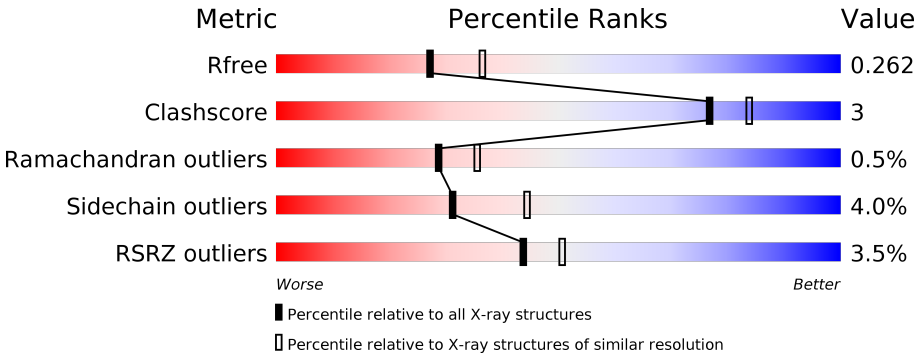
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.30 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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	252	3%, 67%, 6%, 25%
2	B	22	50%, 50%

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 1724 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 Zinc finger MYND domain-containing protein 11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	188	Total	C	N	O	S	0	0	0
			1579	1010	284	273	12			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	120	MET	-	EXPRESSION TAG	UNP Q8R5C8
A	121	GLY	-	EXPRESSION TAG	UNP Q8R5C8
A	122	SER	-	EXPRESSION TAG	UNP Q8R5C8
A	123	SER	-	EXPRESSION TAG	UNP Q8R5C8
A	124	HIS	-	EXPRESSION TAG	UNP Q8R5C8
A	125	HIS	-	EXPRESSION TAG	UNP Q8R5C8
A	126	HIS	-	EXPRESSION TAG	UNP Q8R5C8
A	127	HIS	-	EXPRESSION TAG	UNP Q8R5C8
A	128	HIS	-	EXPRESSION TAG	UNP Q8R5C8
A	129	HIS	-	EXPRESSION TAG	UNP Q8R5C8
A	130	SER	-	EXPRESSION TAG	UNP Q8R5C8
A	131	SER	-	EXPRESSION TAG	UNP Q8R5C8
A	132	GLY	-	EXPRESSION TAG	UNP Q8R5C8
A	133	LEU	-	EXPRESSION TAG	UNP Q8R5C8
A	134	VAL	-	EXPRESSION TAG	UNP Q8R5C8
A	135	PRO	-	EXPRESSION TAG	UNP Q8R5C8
A	136	ARG	-	EXPRESSION TAG	UNP Q8R5C8
A	137	GLY	-	EXPRESSION TAG	UNP Q8R5C8
A	138	SER	-	EXPRESSION TAG	UNP Q8R5C8
A	139	HIS	-	EXPRESSION TAG	UNP Q8R5C8
A	140	MET	-	EXPRESSION TAG	UNP Q8R5C8
A	141	ALA	-	EXPRESSION TAG	UNP Q8R5C8
A	142	SER	-	EXPRESSION TAG	UNP Q8R5C8
A	143	MET	-	EXPRESSION TAG	UNP Q8R5C8
A	144	THR	-	EXPRESSION TAG	UNP Q8R5C8
A	145	GLY	-	EXPRESSION TAG	UNP Q8R5C8
A	146	GLY	-	EXPRESSION TAG	UNP Q8R5C8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	147	GLN	-	EXPRESSION TAG	UNP Q8R5C8
A	148	GLN	-	EXPRESSION TAG	UNP Q8R5C8
A	149	MET	-	EXPRESSION TAG	UNP Q8R5C8
A	150	GLY	-	EXPRESSION TAG	UNP Q8R5C8
A	151	ARG	-	EXPRESSION TAG	UNP Q8R5C8
A	152	GLY	-	EXPRESSION TAG	UNP Q8R5C8
A	153	SER	-	EXPRESSION TAG	UNP Q8R5C8
A	234	ALA	ASP	ENGINEERED MUTATION	UNP Q8R5C8
A	236	ALA	GLU	ENGINEERED MUTATION	UNP Q8R5C8

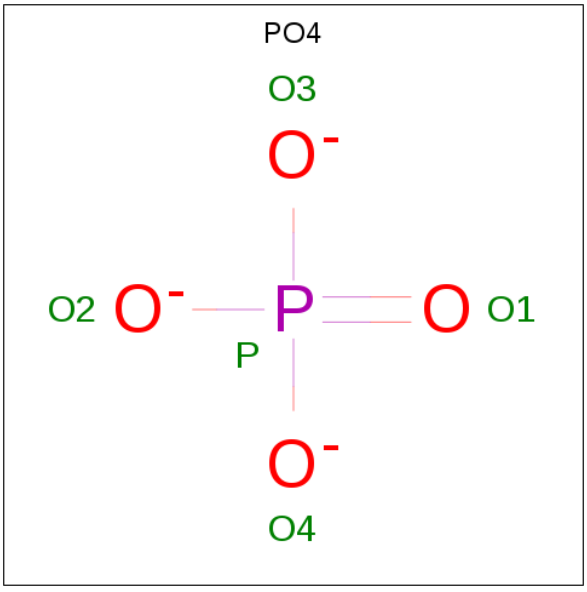
- Molecule 2 is a protein called Peptide from Histone H3.1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	11	Total	C	N	O	0	0	0
			77	50	15	12			

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

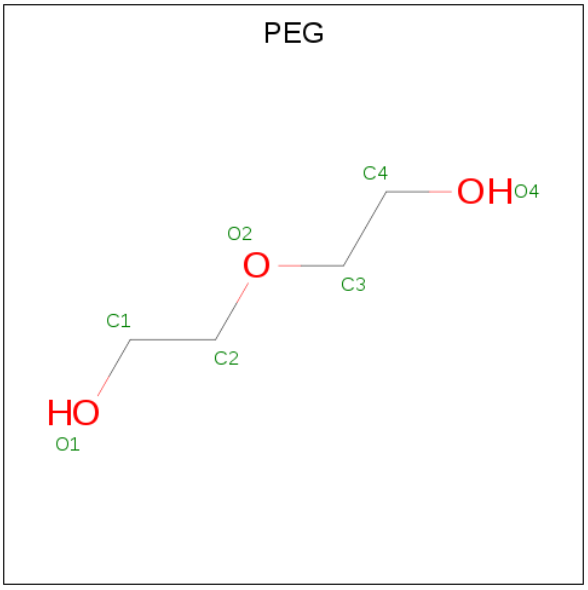
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Zn	0	0
			1	1		

- Molecule 4 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total O P 5 4 1	0	0

- Molecule 5 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 7 4 3	0	0

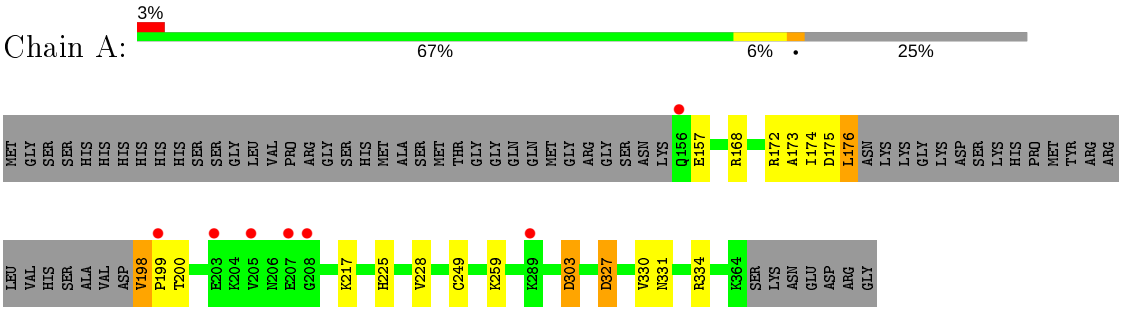
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	52	Total O 52 52	0	0
6	B	3	Total O 3 3	0	0

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: Zinc finger MYND domain-containing protein 11



- Molecule 2: Peptide from Histone H3.1

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	64.36Å 64.36Å 122.06Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	36.48 – 2.30 36.48 – 2.30	Depositor EDS
% Data completeness (in resolution range)	97.1 (36.48-2.30) 96.6 (36.48-2.30)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.60 (at 2.31Å)	Xtriage
Refinement program	PHENIX 1.8_1069	Depositor
R, R_{free}	0.206 , 0.263 0.206 , 0.262	Depositor DCC
R_{free} test set	868 reflections (7.48%)	wwPDB-VP
Wilson B-factor (Å^2)	40.6	Xtriage
Anisotropy	0.088	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 45.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	1724	wwPDB-VP
Average B, all atoms (Å^2)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.20% 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: ZN, PEG, PO4, M3L

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	0.23	0/1622	0.38	0/2185
2	B	0.18	0/67	0.45	0/91
All	All	0.23	0/1689	0.39	0/2276

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	1579	0	1532	10	0
2	B	77	0	84	0	0
3	A	1	0	0	0	0
4	A	5	0	0	0	0
5	A	7	0	10	0	0
6	A	52	0	0	0	0
6	B	3	0	0	0	0
All	All	1724	0	1626	10	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.

The worst 5 of 10 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:327:ASP:N	1:A:327:ASP:OD1	2.35	0.58
1:A:198:VAL:HG13	1:A:199:PRO:HD3	1.85	0.58
1:A:175:ASP:O	1:A:176:LEU:HB2	2.10	0.51
1:A:157:GLU:OE2	1:A:259:LYS:NZ	2.33	0.51
1:A:217:LYS:HE3	1:A:249:CYS:SG	2.55	0.46

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	184/252 (73%)	181 (98%)	2 (1%)	1 (0%)	29 35
2	B	8/22 (36%)	8 (100%)	0	0	100 100
All	All	192/274 (70%)	189 (98%)	2 (1%)	1 (0%)	29 35

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	174	ILE

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	167/220 (76%)	160 (96%)	7 (4%)	30 42
2	B	6/14 (43%)	6 (100%)	0	100 100
All	All	173/234 (74%)	166 (96%)	7 (4%)	31 44

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	198	VAL
1	A	327	ASP
1	A	200	THR
1	A	172	ARG
1	A	303	ASP

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

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

1 non-standard protein/DNA/RNA residue is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	M3L	B	36	2	10,11,12	0.72	0	9,14,16	0.54	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	M3L	B	36	2	-	0/9/10/12	-

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.

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 3 ligands modelled in this entry, 1 is monoatomic - leaving 2 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
5	PEG	A	403	-	6,6,6	0.45	0	5,5,5	0.30	0
4	PO4	A	402	-	4,4,4	0.90	0	6,6,6	0.49	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	PEG	A	403	-	-	1/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	403	PEG	O2-C3-C4-O4

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	188/252 (74%)	0.13	7 (3%) 41 48	20, 38, 71, 84	0
2	B	10/22 (45%)	-0.25	0 100 100	26, 31, 50, 55	0
All	All	198/274 (72%)	0.11	7 (3%) 44 51	20, 38, 71, 84	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	289	LYS	3.7
1	A	203	GLU	3.5
1	A	199	PRO	3.4
1	A	156	GLN	2.8
1	A	207	GLU	2.8

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

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	M3L	B	36	12/13	0.96	0.12	27,34,41,41	0

6.3 Carbohydrates

There are no carbohydrates 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
4	PO4	A	402	5/5	0.78	0.19	52,55,65,87	0
5	PEG	A	403	7/7	0.86	0.27	36,46,58,61	0
3	ZN	A	401	1/1	1.00	0.13	31,31,31,31	0

6.5 Other polymers

There are no such residues in this entry.