



wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 10:32 pm BST

PDB ID : 6XWT
Title : drosophila melanogaster CENP-A/H4 bound to N-terminal CAL1 fragment
Authors : Jeyaprakash, A.A.; Medina-Pritchard, B.; Lazou, V.; Zou, J.; Byron, O.; Abad, M.A.; Rappsilber, J.; Heun, P.
Deposited on : 2020-01-24
Resolution : 3.47 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 ⓘ 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

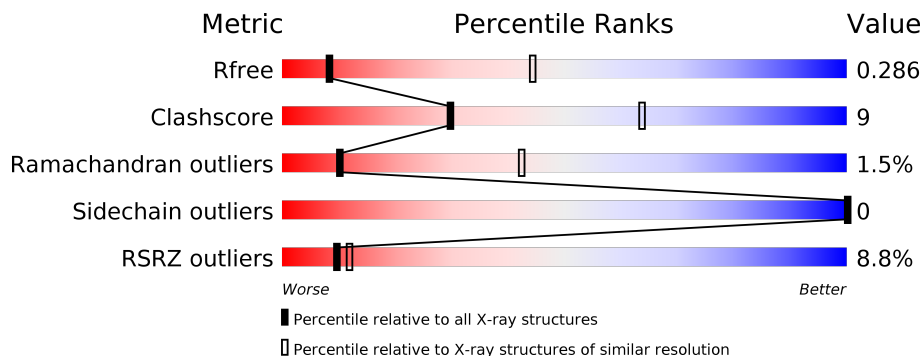
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.47 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	1379 (3.56-3.40)
Clashscore	141614	1461 (3.56-3.40)
Ramachandran outliers	138981	1424 (3.56-3.40)
Sidechain outliers	138945	1425 (3.56-3.40)
RSRZ outliers	127900	1289 (3.56-3.40)

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 $\leq 5\%$. 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	225	
1	C	225	
2	B	103	
2	D	103	
3	E	979	
3	F	979	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2803 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 Histone H3-like centromeric protein cid.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	74	599	380	103	109	7	0	0	0
1	C	74	594	377	103	107	7	0	0	0

- Molecule 2 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	72	569	360	110	98	1	0	0	0
2	D	71	564	358	108	97	1	0	0	0

- Molecule 3 is a protein called Chromosome alignment defect 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	E	30	236	146	37	51	2	0	0	0
3	F	31	241	149	38	52	2	0	0	0

PRO
PRO
ALA
SER
GLU
ARG
GLN
GLU
ILE
TYR
PRO
ALA
THR
ILE
LYS
ILE
THR
LYS
ARG
GLY
ILE
LEU
LEU
HIS
SER
PRO
SER
ALA
GLN
GLY
GLY
ALA
SER
PHE
THR
LEU
SER
THR
GLU
GLN
GLY
LEU
GLY
LYS
ILE
ILE
GLU
ARG
TRP
ALA
ARG
LYS
TYR
TYR
LEU
LYS
TYR
HIS
ILE
GLY
ASN
SER
ARG

SER
PHE
ASP
SER
ARG
HIS
SER
VAL
TYR
TYR
GLN
PRO
THR
ILE
PRO
GLN
LEU
LEU
ALA
ALA
ALA
LEU
SER
ALA
PRO
HIS
GLN
SER
ASP
ALA
GLN
ASN
PRO
ILE
GLY
ASN
ILE
SER
PHE
GLY
LEU
SER
SER
ALA
SER
ASP
ASP
ILE
PHE
GLU
GLN
ILE
GLU
ASN
ARG
TRP
TYR
GLY
THR
VAL
TYR
SER
ILE
LEU
LEU
ASN
GLY
ASN
SER
ARG

GLY
ASP
LYS

4 Data and refinement statistics

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, α , β , γ	178.25Å 178.25Å 133.28Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	100.88 – 3.47 100.88 – 3.47	Depositor EDS
% Data completeness (in resolution range)	99.9 (100.88-3.47) 92.0 (100.88-3.47)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.05 (at 3.49Å)	Xtrriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.272 , 0.286 0.272 , 0.286	Depositor DCC
R_{free} test set	542 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å ²)	125.0	Xtrriage
Anisotropy	0.120	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 108.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	2803	wwPDB-VP
Average B, all atoms (Å ²)	136.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.34% 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, 0.333 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	0.27	0/608	0.53	0/819
1	C	0.29	0/603	0.55	0/812
2	B	0.27	0/574	0.48	0/770
2	D	0.29	0/570	0.53	0/765
3	E	0.26	0/239	0.48	0/319
3	F	0.28	0/244	0.52	0/326
All	All	0.28	0/2838	0.52	0/3811

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	2
1	C	0	1
2	D	0	1
3	E	0	2
3	F	0	1
All	All	0	7

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	70	LEU	Peptide
1	A	71	ARG	Peptide
1	C	70	LEU	Peptide
2	D	31	THR	Peptide
3	E	33	GLU	Peptide

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	599	0	608	16	0
1	C	594	0	599	20	0
2	B	569	0	615	13	0
2	D	564	0	605	15	0
3	E	236	0	215	7	0
3	F	241	0	217	6	0
All	All	2803	0	2859	53	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 9.

The worst 5 of 53 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:90:GLN:NE2	2:B:98:LEU:O	2.21	0.73
1:C:90:GLN:NE2	2:D:98:LEU:O	2.22	0.72
1:C:65:SER:OG	1:C:68:GLU:O	2.11	0.69
2:B:98:LEU:HD22	3:F:19:SER:HA	1.75	0.69
2:B:27:ILE:HG23	2:B:56:ARG:HB3	1.75	0.68

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	72/225 (32%)	67 (93%)	3 (4%)	2 (3%)	5 30

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	72/225 (32%)	66 (92%)	5 (7%)	1 (1%)	11	43
2	B	70/103 (68%)	67 (96%)	3 (4%)	0	100	100
2	D	69/103 (67%)	62 (90%)	5 (7%)	2 (3%)	4	29
3	E	28/979 (3%)	24 (86%)	4 (14%)	0	100	100
3	F	29/979 (3%)	21 (72%)	8 (28%)	0	100	100
All	All	340/2614 (13%)	307 (90%)	28 (8%)	5 (2%)	10	42

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	31	THR
1	A	71	ARG
1	A	72	VAL
2	D	32	LYS
1	C	72	VAL

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	66/198 (33%)	66 (100%)	0	100	100
1	C	64/198 (32%)	64 (100%)	0	100	100
2	B	59/79 (75%)	59 (100%)	0	100	100
2	D	58/79 (73%)	58 (100%)	0	100	100
3	E	25/861 (3%)	25 (100%)	0	100	100
3	F	25/861 (3%)	25 (100%)	0	100	100
All	All	297/2276 (13%)	297 (100%)	0	100	100

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

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](#)

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](#)

There are no ligands in this entry.

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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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	74/225 (32%)	1.09	5 (6%) 17 18	108, 132, 156, 163	0
1	C	74/225 (32%)	0.83	4 (5%) 25 25	111, 129, 157, 163	0
2	B	72/103 (69%)	1.04	11 (15%) 2 3	104, 132, 159, 169	0
2	D	71/103 (68%)	0.72	3 (4%) 36 34	102, 127, 155, 175	0
3	E	30/979 (3%)	0.89	3 (10%) 7 9	119, 144, 195, 214	0
3	F	31/979 (3%)	0.80	5 (16%) 1 2	134, 153, 183, 191	0
All	All	352/2614 (13%)	0.91	31 (8%) 10 12	102, 133, 166, 214	0

The worst 5 of 31 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	27	ILE	4.9
3	E	42	GLU	4.2
2	B	35	ILE	4.0
1	A	70	LEU	3.9
2	B	47	ILE	3.2

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](#)

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.