



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 24, 2024 – 02:14 PM EDT

PDB ID : 6XWU  
Title : Crystal structure of drosophila melanogaster CENP-C cumin domain  
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 : 1.82 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](mailto: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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
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)	:	2.37.1

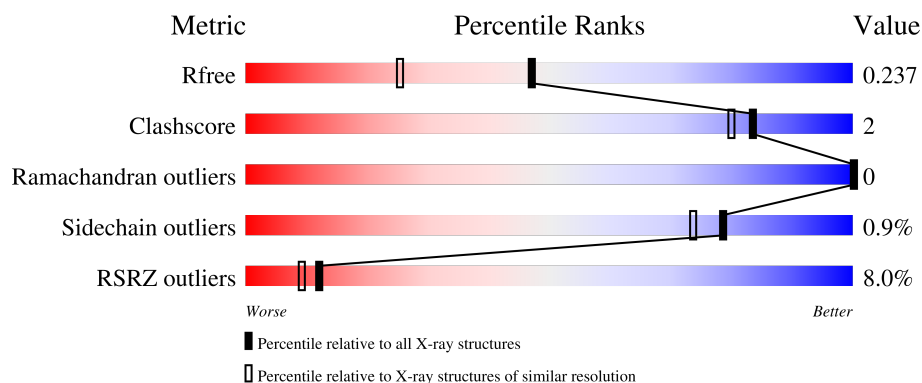
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.82 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	7484 (1.84-1.80)
Clashscore	141614	8401 (1.84-1.80)
Ramachandran outliers	138981	8290 (1.84-1.80)
Sidechain outliers	138945	8290 (1.84-1.80)
RSRZ outliers	127900	7371 (1.84-1.80)

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  $\geq 3$ , 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	1411	<div> <div style="display: flex; align-items: center;"> <div style="width: 90%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="margin-left: 5px;">%</div> </div> <div style="display: flex; justify-content: space-between; margin-top: 2px;"> <span>9%</span> <span>90%</span> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 1107 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 RE68959p.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	138	Total	C	N	O	S	0	0	0
			1065	670	184	203	8			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	42	Total	O	0	0
			42	42		



ALA	ALA	ALA	GLU	SER	LEU	LEU	ALA	GLY	ASP	PRO	ASN	VAL	GLY
ALA	LEU	ALA	ALA	LYS	ALA	LYS	ASP	LYS	ASP	GLN	THR	THR	ALA
GLY	GLU	GLU	GLU	LYS	LYS	LYS	MET	PRO	ASP	GLN	ASN	LYS	ILE
PRO	LYS	VAL	VAL	ALA	ALA	ALA	VAL	PRO	VAL	ALA	SER	SER	LEU
VAL	VAL	PRO	PRO	SER	GLU	GLY	GLU	SER	VAL	VAL	ASN	ARG	ARG
PHE	LYS	LYS	LYS	ALA	ALA	ALA	THR	THR	GLY	ALA	THR	LYS	LYS
THR	THR	THR	ARG	THR	GLU	GLU	GLU	GLU	SER	THR	THR	ASP	SER
PRO	GLY	ARG	GLY	PRO	LYS	LYS	LEU	ALA	ARG	GLU	GLU	GLY	GLY
LEU	ARG	ARG	PRO	LYS	ASN	ALA	ALA	ALA	SER	LYS	SER	MET	LYS
ARG	PRO	PRO	PRO	PRO	THR	THR	THR	THR	LEU	VAL	VAL	ALA	LYS
D1274	D1274	D1274	LYS	VAL	THR	GLY	THR	GLY	PHE	GLU	GLY	HIS	LYS
E1275	E1275	E1275	LYS	VAL	GLY	ILE	ILE	LEU	THR	LEU	LEU	THR	GLN
Q1276	Q1276	Q1276	ALA	GLU	GLY	GLY	ARG	GLN	GLY	ASN	ASN	SER	LYS
G1289	G1289	G1289	VAL	LYS	ARG	ARG	GLN	GLN	SER	THR	THR	ASP	LYS
D1292	D1292	D1292	GLY	ASN	SER	SER	VAL	PRO	SER	GLY	ARG	CYS	GLN
A1293	A1293	A1293	VAL	PRO	PRO	PRO	LYS	GLN	GLU	GLU	GLU	ILE	SER
P1294	P1294	P1294	GLN	PRO	ARG	ARG	ARG	ASN	ASN	GLY	MET	GLY	GLN
P1295	P1295	P1295	THR	ARG	GLY	GLN	GLN	SER	GLN	SER	THR	ASP	LYS
S1296	S1296	S1296	THR	GLY	THR	VAL	VAL	VAL	VAL	GLY	GLY	PRO	ASN
A1297	A1297	A1297	GLU	PRO	PRO	PRO	PRO	PRO	SER	LEU	GLY	ASN	LYS
S1308	S1308	S1308	ALA	GLY	ALA	ALA	ALA	ALA	LEU	LEU	GLU	TRP	GLU
A1309	A1309	A1309	GLU	CYS	THR	CYS	PRO	GLN	GLN	SER	SER	ASN	GLY
M1310	M1310	M1310	GLU	SER	SER	SER	PRO	MET	MET	PRO	PRO	ASP	ASN
E1311	E1311	E1311	PRO	THR	THR	THR	PRO	TRP	TRP	VAL	VAL	MET	CYS
M1326	M1326	M1326	GLU	PRO	PRO	PRO	GLY	CYS	SER	GLY	GLY	ASN	ALA
K1330	K1330	K1330	PRO	ILE	ILE	ILE	GLU	THR	THR	ALA	ALA	SER	THR
L1356	L1356	L1356	ILE	LYS	LYS	LYS	ASP	PRO	PRO	VAL	VAL	ASN	ASN
L1357	L1357	L1357	SER	LEU	LEU	PRO	PRO	PRO	PHE	THR	THR	HIS	GLY
V1358	V1358	V1358	VAL	PRO	PRO	GLY	GLY	PHE	LYS	LYS	THR	ASN	ASP
V1370	V1370	V1370	ALA	ALA	ALA	ALA	ASN	PRO	ASN	PRO	PRO	ALA	SER
G1371	G1371	G1371	PRO	ILE	ILE	ILE	PHE	LEU	LEU	PRO	ILE	LYS	ASP
R1393	R1393	R1393	LEU	PRO	PRO	PRO	MET	PRO	PRO	PRO	GLY	SER	ASP
K1402	K1402	K1402	SER	SER	SER	HIS	GLY	SER	GLY	ALA	GLN	LYS	GLU
I1409	I1409	I1409	ASP	GLN	GLN	GLN	THR	THR	THR	THR	SER	ASN	THR
R1410	R1410	R1410	GLU	LEU	LEU	LEU	GLY	GLY	ILE	ALA	ASP	ARG	GLY
S1411	S1411	S1411	PRO	PRO	PRO	PRO	ARG	LYS	PRO	VAL	VAL	LYS	ASN
			GLU	THR	THR	THR	THR	LYS	SER	SER	PRO	SER	GLU
			GLN	GLU	GLU	GLU	THR	THR	VAL	GLU	GLY	LYS	LYS
			ALA	ALA	ALA	ALA	VAL	VAL	VAL	PRO	PRO	SER	SER
			THR	THR	THR	THR	THR	THR	THR	GLN	GLN	ASN	ASN
			PRO	PRO	PRO	PRO	GLY	GLY	GLY	LYS	ARG	LYS	LYS
			VAL	VAL	VAL	VAL	ASN	LEU	THR	THR	THR	LEU	ASN
			THR	THR	THR	THR	LEU	THR	THR	THR	THR	THR	ASN

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	57.20Å 57.20Å 92.96Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	28.60 – 1.82 28.60 – 1.82	Depositor EDS
% Data completeness (in resolution range)	98.6 (28.60-1.82) 97.9 (28.60-1.82)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.97 (at 1.82Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660, PHENIX 1.17.1_3660	Depositor
R, $R_{free}$	0.194 , 0.235 0.197 , 0.237	Depositor DCC
$R_{free}$ test set	1429 reflections (9.90%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	32.6	Xtriage
Anisotropy	0.115	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 47.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	1107	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.08% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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	0.36	0/1082	0.55	0/1458

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 [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	1065	0	1043	5	1
2	A	42	0	0	2	0
All	All	1107	0	1043	5	1

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 2.

All (5) 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:1393:ARG:NH1	2:A:1501:HOH:O	2.23	0.72
1:A:1294:PRO:O	1:A:1310:ASN:ND2	2.28	0.66
1:A:1370:VAL:HG22	2:A:1502:HOH:O	2.19	0.42
1:A:1308:SER:O	1:A:1311:GLU:HB2	2.20	0.42
1:A:1326:ASN:HD22	1:A:1326:ASN:HA	1.75	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1276:GLN:OE1	1:A:1402:LYS:NZ[7_555]	2.14	0.06

## 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	136/1411 (10%)	133 (98%)	3 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	113/1275 (9%)	112 (99%)	1 (1%)	78	74

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1330	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1326	ASN

### 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 monosaccharides 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	138/1411 (9%)	0.32	11 (7%) 12 9	24, 38, 66, 74	0

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1358	VAL	3.0
1	A	1371	GLY	2.7
1	A	1292	ASP	2.7
1	A	1370	VAL	2.6
1	A	1357	LEU	2.5

### 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 monosaccharides 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.