



# Full wwPDB X-ray Structure Validation Report ⓘ

May 26, 2020 – 11:52 pm BST

PDB ID : 5GPL  
Title : Crystal structure of Ccp1  
Authors : Yin, F.; Gao, F.; Chen, Y.  
Deposited on : 2016-08-03  
Resolution : 2.10 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](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 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

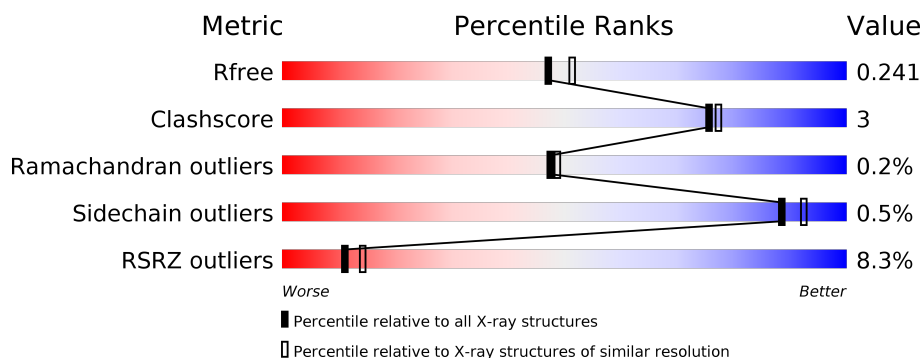
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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  $\geq 3$ , 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	274	<div> <div>8%</div> <div> <div></div> <div>70%</div> <div>5%</div> <div>25%</div> </div> </div>
1	B	274	<div> <div>4%</div> <div> <div></div> <div>68%</div> <div>7%</div> <div>25%</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3736 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 Putative nucleosome assembly protein C36B7.08c.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	206	Total	C	N	O	S	0	3	0
			1729	1108	273	347	1			
1	B	206	Total	C	N	O	S	0	6	0
			1749	1119	278	351	1			

There are 60 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	245	ALA	-	expression tag	UNP Q9HGN2
A	246	ALA	-	expression tag	UNP Q9HGN2
A	247	ALA	-	expression tag	UNP Q9HGN2
A	248	GLU	-	expression tag	UNP Q9HGN2
A	249	ASN	-	expression tag	UNP Q9HGN2
A	250	LEU	-	expression tag	UNP Q9HGN2
A	251	TYR	-	expression tag	UNP Q9HGN2
A	252	PHE	-	expression tag	UNP Q9HGN2
A	253	GLN	-	expression tag	UNP Q9HGN2
A	254	GLY	-	expression tag	UNP Q9HGN2
A	255	LEU	-	expression tag	UNP Q9HGN2
A	256	GLU	-	expression tag	UNP Q9HGN2
A	257	ASP	-	expression tag	UNP Q9HGN2
A	258	TYR	-	expression tag	UNP Q9HGN2
A	259	LYS	-	expression tag	UNP Q9HGN2
A	260	ASP	-	expression tag	UNP Q9HGN2
A	261	ASP	-	expression tag	UNP Q9HGN2
A	262	ASP	-	expression tag	UNP Q9HGN2
A	263	ASP	-	expression tag	UNP Q9HGN2
A	264	LYS	-	expression tag	UNP Q9HGN2
A	265	HIS	-	expression tag	UNP Q9HGN2
A	266	HIS	-	expression tag	UNP Q9HGN2
A	267	HIS	-	expression tag	UNP Q9HGN2
A	268	HIS	-	expression tag	UNP Q9HGN2
A	269	HIS	-	expression tag	UNP Q9HGN2

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	270	HIS	-	expression tag	UNP Q9HGN2
A	271	HIS	-	expression tag	UNP Q9HGN2
A	272	HIS	-	expression tag	UNP Q9HGN2
A	273	HIS	-	expression tag	UNP Q9HGN2
A	274	HIS	-	expression tag	UNP Q9HGN2
B	245	ALA	-	expression tag	UNP Q9HGN2
B	246	ALA	-	expression tag	UNP Q9HGN2
B	247	ALA	-	expression tag	UNP Q9HGN2
B	248	GLU	-	expression tag	UNP Q9HGN2
B	249	ASN	-	expression tag	UNP Q9HGN2
B	250	LEU	-	expression tag	UNP Q9HGN2
B	251	TYR	-	expression tag	UNP Q9HGN2
B	252	PHE	-	expression tag	UNP Q9HGN2
B	253	GLN	-	expression tag	UNP Q9HGN2
B	254	GLY	-	expression tag	UNP Q9HGN2
B	255	LEU	-	expression tag	UNP Q9HGN2
B	256	GLU	-	expression tag	UNP Q9HGN2
B	257	ASP	-	expression tag	UNP Q9HGN2
B	258	TYR	-	expression tag	UNP Q9HGN2
B	259	LYS	-	expression tag	UNP Q9HGN2
B	260	ASP	-	expression tag	UNP Q9HGN2
B	261	ASP	-	expression tag	UNP Q9HGN2
B	262	ASP	-	expression tag	UNP Q9HGN2
B	263	ASP	-	expression tag	UNP Q9HGN2
B	264	LYS	-	expression tag	UNP Q9HGN2
B	265	HIS	-	expression tag	UNP Q9HGN2
B	266	HIS	-	expression tag	UNP Q9HGN2
B	267	HIS	-	expression tag	UNP Q9HGN2
B	268	HIS	-	expression tag	UNP Q9HGN2
B	269	HIS	-	expression tag	UNP Q9HGN2
B	270	HIS	-	expression tag	UNP Q9HGN2
B	271	HIS	-	expression tag	UNP Q9HGN2
B	272	HIS	-	expression tag	UNP Q9HGN2
B	273	HIS	-	expression tag	UNP Q9HGN2
B	274	HIS	-	expression tag	UNP Q9HGN2

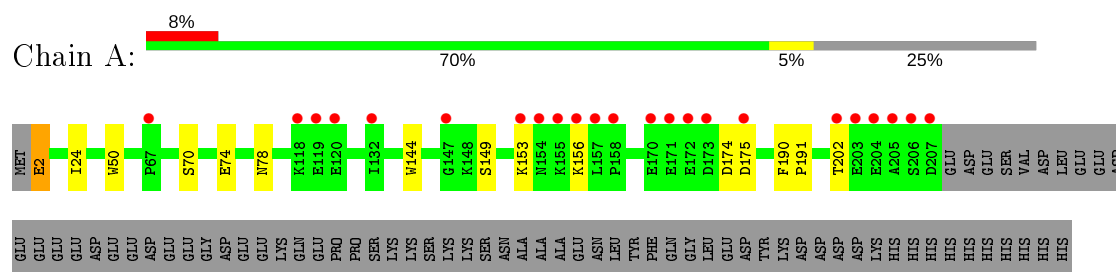
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	121	Total O 121 121	0	0
2	B	137	Total O 137 137	0	0

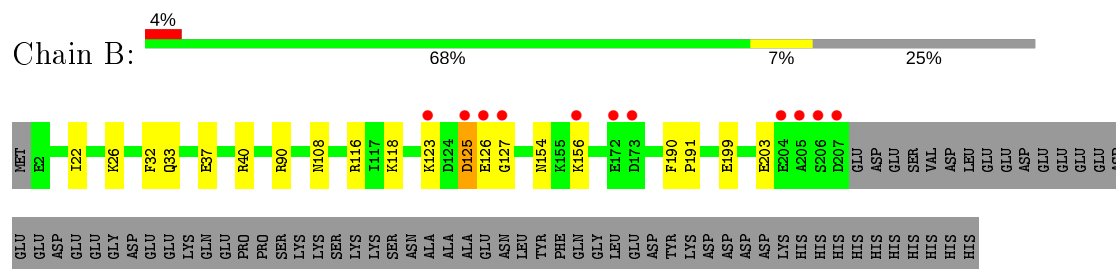
### 3 Residue-property plots [i](#)

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: Putative nucleosome assembly protein C36B7.08c



- Molecule 1: Putative nucleosome assembly protein C36B7.08c



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	86.66 Å 86.66 Å 158.77 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.76 – 2.10 38.76 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.7 (38.76-2.10) 99.7 (38.76-2.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.80 (at 2.10 Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.193 , 0.236 0.199 , 0.241	Depositor DCC
$R_{free}$ test set	1743 reflections (4.84%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	23.9	Xtriage
Anisotropy	0.042	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 46.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	3736	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.00% 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

### 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	0.24	0/1774	0.40	0/2397
1	B	0.24	0/1797	0.41	0/2427
All	All	0.24	0/3571	0.40	0/4824

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	1729	0	1675	9	1
1	B	1749	0	1696	13	0
2	A	121	0	0	0	0
2	B	137	0	0	1	0
All	All	3736	0	3371	21	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 3.

All (21) 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:B:125:ASP:OD1	1:B:126:GLU:N	2.18	0.72

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:123:LYS:HG2	1:B:127:GLY:HA2	1.91	0.52
1:B:37[B]:GLU:OE1	1:B:40[B]:ARG:NH1	2.44	0.51
1:A:50:TRP:CD1	1:A:78:ASN:HA	2.49	0.47
1:A:24:ILE:HG13	1:B:32:PHE:HE2	1.79	0.47
1:A:144:TRP:CD2	1:A:149:SER:HA	2.50	0.47
1:B:90:ARG:NH1	2:B:303:HOH:O	2.41	0.45
1:A:70:SER:O	1:A:74:GLU:HG2	2.17	0.45
1:A:190:PHE:HB3	1:A:191:PRO:HD3	1.97	0.45
1:A:153:LYS:HA	1:A:153:LYS:HD2	1.74	0.44
1:B:190:PHE:HB3	1:B:191:PRO:HD3	1.98	0.44
1:A:174:ASP:OD1	1:A:175:ASP:N	2.51	0.43
1:B:125:ASP:OD1	1:B:126:GLU:HG2	2.18	0.43
1:A:175:ASP:N	1:A:175:ASP:OD1	2.49	0.43
1:B:199[B]:GLU:O	1:B:203:GLU:HG3	2.19	0.43
1:B:33:GLN:O	1:B:37[B]:GLU:HG2	2.19	0.43
1:B:116:ARG:HH21	1:B:118:LYS:HD2	1.85	0.42
1:B:22:ILE:HG22	1:B:26:LYS:HE2	2.01	0.41
1:A:156:LYS:HA	1:A:156:LYS:HD3	1.59	0.41
1:B:154:ASN:ND2	1:B:156:LYS:HB3	2.36	0.41
1:B:125:ASP:CG	1:B:126:GLU:H	2.14	0.41

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:2:GLU:OE1	1:A:202:THR:OG1[4_664]	2.17	0.03

## 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	207/274 (76%)	201 (97%)	6 (3%)	0	100 100

Continued on next page...



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	210/274 (77%)	204 (97%)	5 (2%)	1 (0%)	29	26
All	All	417/548 (76%)	405 (97%)	11 (3%)	1 (0%)	47	49

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	125	ASP

### 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	190/250 (76%)	189 (100%)	1 (0%)	88	92
1	B	193/250 (77%)	192 (100%)	1 (0%)	88	92
All	All	383/500 (77%)	381 (100%)	2 (0%)	88	92

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2	GLU
1	B	108	ASN

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 [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 ⓘ

### 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<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	206/274 (75%)	0.37	23 (11%) 5 6	14, 32, 84, 115	0
1	B	206/274 (75%)	-0.10	11 (5%) 26 32	13, 23, 69, 129	0
All	All	412/548 (75%)	0.14	34 (8%) 11 14	13, 27, 81, 129	0

All (34) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	172	GLU	5.8
1	A	207	ASP	5.5
1	A	173	ASP	5.3
1	A	206	SER	5.2
1	A	157	LEU	4.8
1	B	206	SER	4.3
1	B	125	ASP	4.1
1	A	156	LYS	4.0
1	A	158	PRO	4.0
1	A	153	LYS	4.0
1	A	205	ALA	3.8
1	B	205	ALA	3.7
1	A	154	ASN	3.7
1	A	155	LYS	3.6
1	A	203	GLU	3.4
1	A	119	GLU	3.1
1	A	118	LYS	3.0
1	A	67	PRO	2.8
1	B	172	GLU	2.7
1	B	207	ASP	2.6
1	B	126	GLU	2.6
1	A	147	GLY	2.5
1	A	171	GLU	2.5
1	A	204	GLU	2.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	127	GLY	2.4
1	A	132	ILE	2.4
1	B	123	LYS	2.3
1	A	175	ASP	2.2
1	A	202	THR	2.2
1	B	173	ASP	2.1
1	A	120	GLU	2.1
1	B	204	GLU	2.1
1	A	170	GLU	2.1
1	B	156	LYS	2.0

## 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 [i](#)

There are no such residues in this entry.