



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 21, 2016 – 02:47 PM EST

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

<http://wwpdb.org/validation/2016/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	:	unknown
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20028320
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20028320

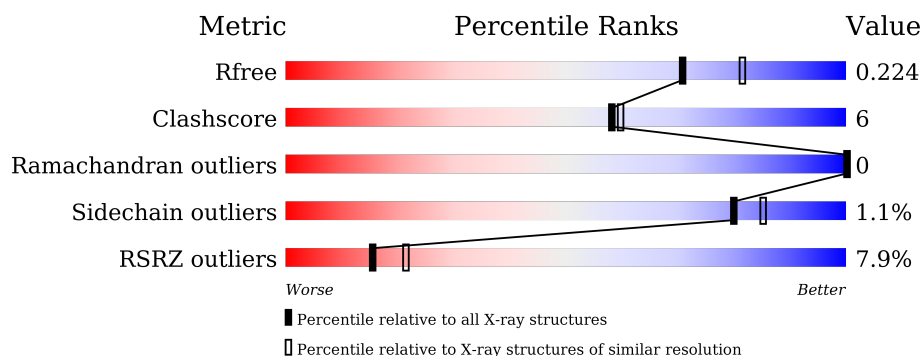
# 1 Overall quality at a glance

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))
$R_{free}$	91344	3939 (2.10-2.10)
Clashscore	102246	4460 (2.10-2.10)
Ramachandran outliers	100387	4413 (2.10-2.10)
Sidechain outliers	100360	4414 (2.10-2.10)
RSRZ outliers	91569	3948 (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. 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>7%</div> <div> <div></div> <div>64%</div> <div>11%</div> <div>25%</div> </div> </div>
1	B	274	<div> <div>5%</div> <div> <div></div> <div>64%</div> <div>11%</div> <div>24%</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3810 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	Se	0	1	0
			1715	1096	275	341	3			
1	B	207	Total	C	N	O	Se	0	3	0
			1742	1112	279	347	4			

There are 64 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	117	MSE	ILE	engineered mutation	UNP Q9HGN2
A	150	MSE	LEU	engineered mutation	UNP Q9HGN2
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

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	268	HIS	-	expression tag	UNP Q9HGN2
A	269	HIS	-	expression tag	UNP Q9HGN2
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	117	MSE	ILE	engineered mutation	UNP Q9HGN2
B	150	MSE	LEU	engineered mutation	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	171	Total 171	O 171	0	0
2	B	182	Total 182	O 182	0	0



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	86.81Å 86.81Å 158.78Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.82 – 2.10 38.82 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.7 (38.82-2.10) 99.7 (38.82-2.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	6.27 (at 2.10Å)	Xtriage
Refinement program	PHENIX (1.10_2155: ???)	Depositor
R, $R_{free}$	0.176 , 0.225 0.174 , 0.224	Depositor DCC
$R_{free}$ test set	1799 reflections (5.00%)	DCC
Wilson B-factor (Å <sup>2</sup> )	24.0	Xtriage
Anisotropy	0.056	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 55.4	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	0.95	EDS
Total number of atoms	3810	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.17% 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	0.45	0/1748	0.61	1/2356 (0.0%)
1	B	0.50	2/1778 (0.1%)	0.61	1/2394 (0.0%)
All	All	0.48	2/3526 (0.1%)	0.61	2/4750 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	171	GLU	CD-OE1	-6.49	1.18	1.25
1	B	171	GLU	CD-OE2	-5.72	1.19	1.25

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	117	MSE	CB-CG-SE	-7.87	89.09	112.70
1	A	157	LEU	CA-CB-CG	5.49	127.93	115.30

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	1715	0	1654	21	0
1	B	1742	0	1687	21	0
2	A	171	0	0	2	0
2	B	182	0	0	3	1
All	All	3810	0	3341	40	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 6.

The worst 5 of 40 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:68:GLU:HG2	1:A:150:MSE:HE1	1.26	1.14
1:A:108:ASN:HD22	1:A:143:HIS:H	1.15	0.95
1:A:68:GLU:CG	1:A:150:MSE:HE1	2.08	0.83
1:A:68:GLU:HG2	1:A:150:MSE:CE	2.06	0.82
1:B:117:MSE:HE2	1:B:135:TYR:CZ	2.18	0.78

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)
2:B:363:HOH:O	2:B:380:HOH:O[5_655]	2.18	0.02

## 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	205/274 (75%)	200 (98%)	5 (2%)	0	100	100
1	B	208/274 (76%)	206 (99%)	2 (1%)	0	100	100
All	All	413/548 (75%)	406 (98%)	7 (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	187/246 (76%)	187 (100%)	0	100	100
1	B	191/246 (78%)	187 (98%)	4 (2%)	61	66
All	All	378/492 (77%)	374 (99%)	4 (1%)	80	85

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	1	MSE
1	B	87	LYS
1	B	108	ASN
1	B	171	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	5	GLN
1	A	12	ASN
1	A	108	ASN
1	B	5	GLN
1	B	9	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 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, 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	203/274 (74%)	0.22	19 (9%) 11 14	13, 30, 77, 137	0
1	B	203/274 (74%)	-0.14	13 (6%) 23 30	13, 24, 74, 111	0
All	All	406/548 (74%)	0.04	32 (7%) 15 21	13, 28, 77, 137	0

The worst 5 of 32 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	172	GLU	7.9
1	B	125	ASP	5.1
1	A	153	LYS	5.0
1	A	207	ASP	4.6
1	A	157	LEU	4.5

### 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 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.