



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 24, 2024 – 05:08 AM EDT

PDB ID : 6AF0  
Title : Structure of Ctr9, Paf1 and Cdc73 ternary complex from *Myceliophthora thermophila*  
Authors : Wang, Z.; Deng, P.; Zhou, Y.  
Deposited on : 2018-08-07  
Resolution : 2.88 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

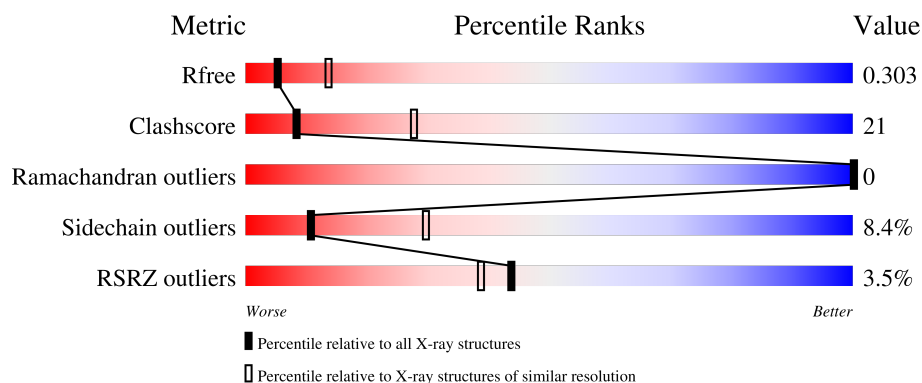
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.88 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	2691 (2.90-2.86)
Clashscore	141614	2947 (2.90-2.86)
Ramachandran outliers	138981	2868 (2.90-2.86)
Sidechain outliers	138945	2871 (2.90-2.86)
RSRZ outliers	127900	2629 (2.90-2.86)

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	939	 4% 61% 31% • •
2	P	121	 % 62% 21% • • 12%
3	C	74	 4% 34% 22% • 43%

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 8372 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 Ctr9 protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	905	Total	C	N	O	S	0	0	0
			7228	4562	1276	1362	28			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	29	GLY	-	expression tag	UNP G2QC65

- Molecule 2 is a protein called Paf1 protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	P	107	Total	C	N	O	S	0	0	0
			810	517	141	149	3			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
P	0	SER	-	expression tag	UNP G2QDK9

- Molecule 3 is a protein called Cdc73 protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	42	Total	C	N	O	S	0	0	0
			334	209	66	58	1			

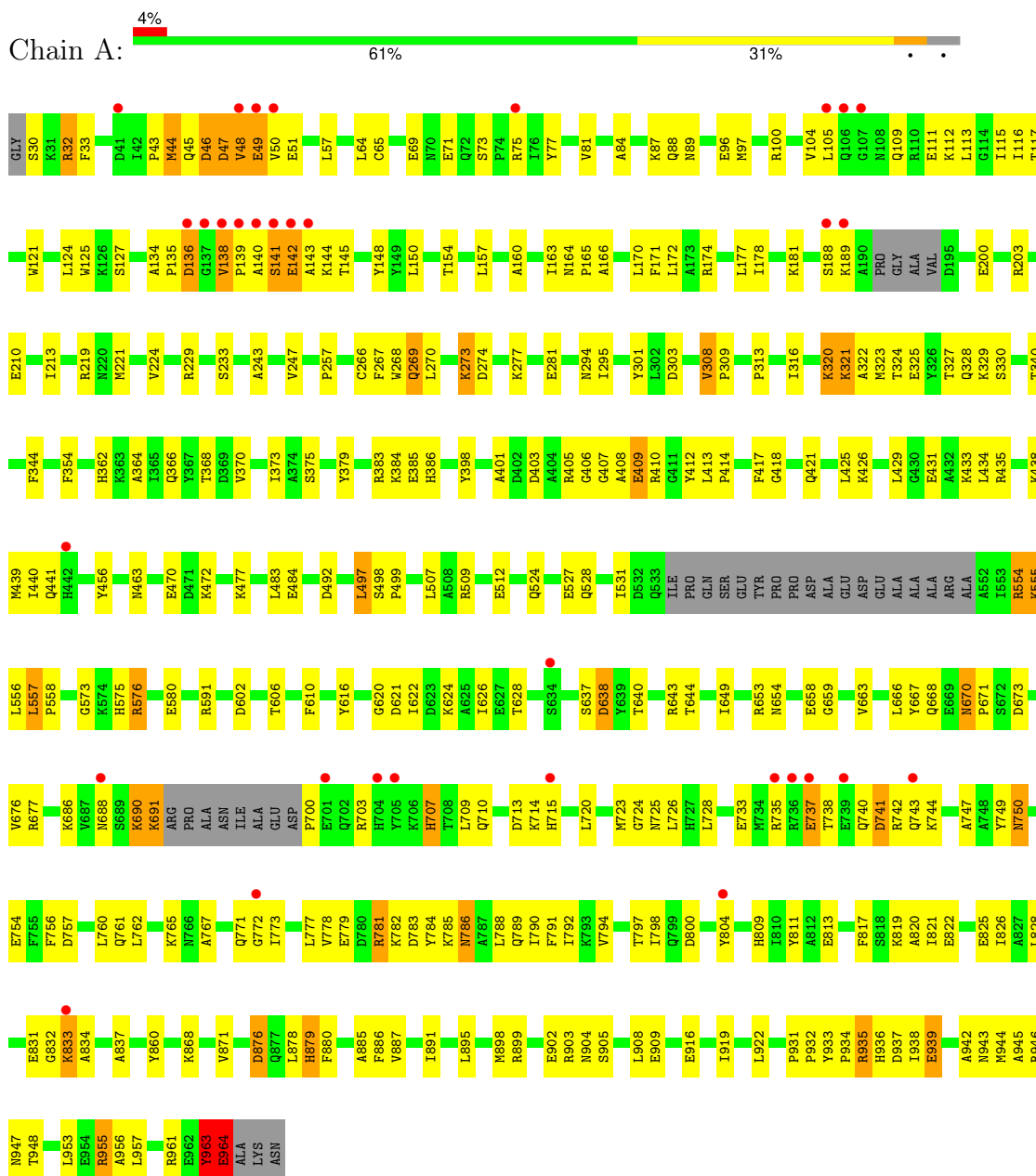
There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	154	SER	-	expression tag	UNP G2Q3X1

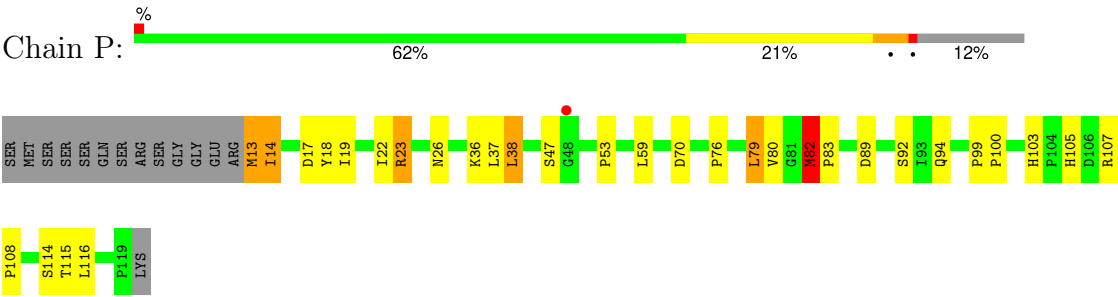
### 3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

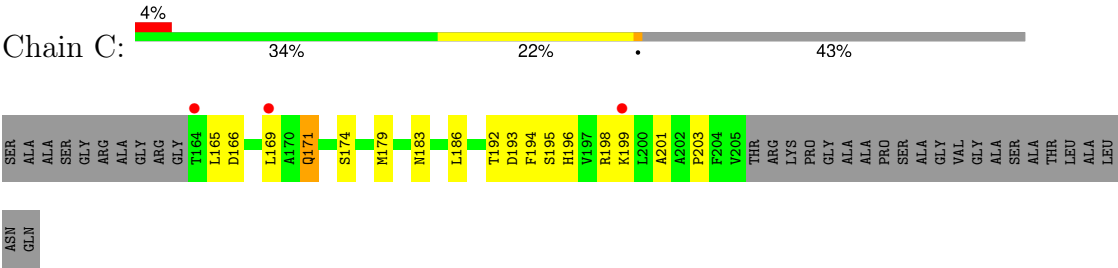
#### • Molecule 1: Ctr9 protein



#### • Molecule 2: Paf1 protein



● Molecule 3: Cdc73 protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	99.87Å 237.79Å 62.74Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	118.89 – 2.88 31.37 – 2.88	Depositor EDS
% Data completeness (in resolution range)	99.3 (118.89-2.88) 99.4 (31.37-2.88)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.53 (at 2.90Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, $R_{free}$	0.245 , 0.309 0.244 , 0.303	Depositor DCC
$R_{free}$ test set	1729 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	90.3	Xtriage
Anisotropy	0.143	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 58.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	8372	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	93.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.11% 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	0.61	7/7365 (0.1%)	0.70	3/9941 (0.0%)
2	P	0.53	0/836	0.76	1/1148 (0.1%)
3	C	0.43	0/340	0.69	0/457
All	All	0.59	7/8541 (0.1%)	0.70	4/11546 (0.0%)

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
2	P	0	1

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	49	GLU	CD-OE2	15.42	1.42	1.25
1	A	49	GLU	CD-OE1	13.52	1.40	1.25
1	A	964	GLU	CD-OE1	10.17	1.36	1.25
1	A	964	GLU	CD-OE2	8.22	1.34	1.25
1	A	963	TYR	CG-CD2	7.97	1.49	1.39

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	832	GLY	N-CA-C	-6.69	96.38	113.10
1	A	44	MET	CG-SD-CE	5.51	109.02	100.20
1	A	269	GLN	N-CA-C	-5.05	97.37	111.00
2	P	70	ASP	CB-CG-OD1	5.04	122.84	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	P	82	MET	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	7228	0	7201	326	0
2	P	810	0	811	37	0
3	C	334	0	343	13	0
All	All	8372	0	8355	349	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 21.

The worst 5 of 349 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:44:MET:HE2	1:A:49:GLU:HG2	1.35	1.04
1:A:49:GLU:HG3	1:A:50:VAL:HG23	1.39	1.01
1:A:44:MET:CE	1:A:49:GLU:HG2	1.91	1.01
1:A:431:GLU:HA	1:A:434:LEU:HG	1.44	0.97
1:A:84:ALA:O	1:A:88:GLN:NE2	2.01	0.94

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	897/939 (96%)	842 (94%)	55 (6%)	0	100	100
2	P	105/121 (87%)	100 (95%)	5 (5%)	0	100	100
3	C	40/74 (54%)	39 (98%)	1 (2%)	0	100	100
All	All	1042/1134 (92%)	981 (94%)	61 (6%)	0	100	100

There are no Ramachandran outliers to report.

### 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	759/782 (97%)	698 (92%)	61 (8%)	12	32
2	P	89/101 (88%)	78 (88%)	11 (12%)	4	12
3	C	35/52 (67%)	33 (94%)	2 (6%)	20	49
All	All	883/935 (94%)	809 (92%)	74 (8%)	11	30

5 of 74 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	944	MET
2	P	116	LEU
1	A	963	TYR
2	P	47	SER
1	A	528	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	715	HIS
1	A	943	ASN
2	P	103	HIS
2	P	26	ASN
1	A	463	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	905/939 (96%)	0.08	33 (3%) 42 38	51, 92, 132, 172	0
2	P	107/121 (88%)	0.09	1 (0%) 84 84	57, 78, 116, 128	0
3	C	42/74 (56%)	0.21	3 (7%) 16 12	71, 105, 127, 144	0
All	All	1054/1134 (92%)	0.08	37 (3%) 44 39	51, 91, 131, 172	0

The worst 5 of 37 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	49	GLU	5.0
1	A	143	ALA	4.5
1	A	136	ASP	4.5
1	A	736	ARG	4.3
1	A	138	VAL	3.7

### 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.