



# wwPDB X-ray Structure Validation Summary Report

Jun 12, 2014 – 06:07 PM EDT

PDB ID : 4C93  
Title : Crystal structure of the carboxy-terminal domain of yeast Ctf4 bound to Pol alpha.  
Authors : Simon, A.C.; Pellegrini, L.  
Deposited on : 2013-10-02  
Resolution : 2.69 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

---

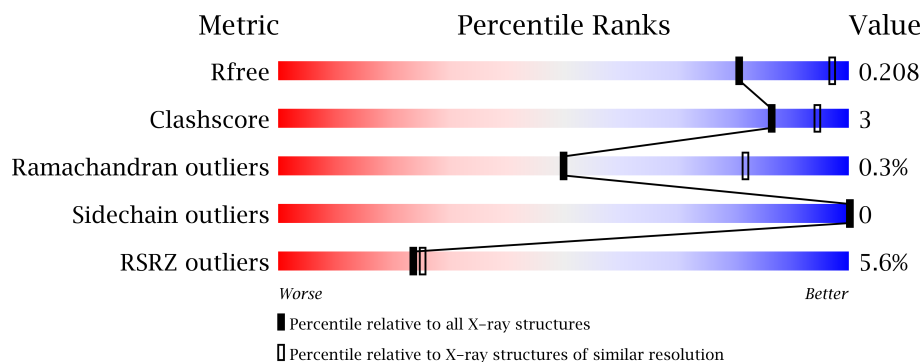
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.16 November 2013  
Xtriage (Phenix) : dev-1439  
EDS : stable23161  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable23161

# 1 Overall quality at a glance

The reported resolution of this entry is 2.69 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))
$R_{free}$	66092	1557 (2.70-2.70)
Clashscore	79885	1939 (2.70-2.70)
Ramachandran outliers	78287	1905 (2.70-2.70)
Sidechain outliers	78261	1905 (2.70-2.70)
RSRZ outliers	66119	1559 (2.70-2.70)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	478	
1	B	478	
1	C	478	
2	D	13	
2	E	13	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 9617 atoms, of which 0 are hydrogen and 0 are deuterium.

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 DNA POLYMERASE ALPHA-BINDING PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	424	Total	C	N	O	S	0	1	0
			3416	2193	566	642	15			
1	B	431	Total	C	N	O	S	0	1	0
			3472	2227	576	653	16			
1	C	296	Total	C	N	O	S	0	1	0
			2405	1562	392	440	11			

- Molecule 2 is a protein called DNA POLYMERASE ALPHA CATALYTIC SUBUNIT A.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	D	10	Total	C	N	O	0	0	0
			82	53	10	19			
2	E	10	Total	C	N	O	0	0	0
			82	53	10	19			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	68	Total	O	0	0
			68	68		
3	B	64	Total	O	0	0
			64	64		
3	C	28	Total	O	0	0
			28	28		



THR	ALA	ALA	VAL	LYS	ILE	SER	GLU	ARG	ALA	GLU	LEU	PRO	SER	LEU	VAL	LYS	LYS	ILE	ASN	ASN	ILE	ARG	GLU	ALA	ARG	TYR	GLU	GLN	GLN	LEU	LYS
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

● Molecule 2: DNA POLYMERASE ALPHA CATALYTIC SUBUNIT A

Chain D: 

ILE	ASP	ASN	F140	D141	D142	I143	L144	G145	E146	F147	E148	S149

● Molecule 2: DNA POLYMERASE ALPHA CATALYTIC SUBUNIT A

Chain E: 

ILE	ASP	ASN	F140	D141	D142	I143	L144	G145	E146	F147	E148	S149

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	88.98Å 100.00Å 219.55Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.20 – 2.69 49.20 – 2.69	Depositor EDS
% Data completeness (in resolution range)	99.5 (49.20-2.69) 99.5 (49.20-2.69)	Depositor EDS
$R_{merge}$	0.16	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.23 (at 2.69Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, $R_{free}$	0.172 , 0.210 0.170 , 0.208	Depositor DCC
$R_{free}$ test set	2762 reflections (5.05%)	DCC
Wilson B-factor (Å <sup>2</sup> )	54.3	Xtriage
Anisotropy	0.221	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 37.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	1 of 54707 reflections (0.002%)	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	9617	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.55% of the height of the origin peak. No significant pseudotranslation is detected.*

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

## 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	0.23	0/3501	0.40	0/4741
1	B	0.24	0/3558	0.40	0/4817
1	C	0.23	0/2481	0.39	0/3370
2	D	0.24	0/83	0.30	0/110
2	E	0.25	0/83	0.32	0/110
All	All	0.23	0/9706	0.40	0/13148

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3416	0	3356	18	0
1	B	3472	0	3418	18	0
1	C	2405	0	2326	19	0
2	D	82	0	67	1	0
2	E	82	0	67	2	0
3	A	68	0	0	1	0
3	B	64	0	0	0	0
3	C	28	0	0	0	0
All	All	9617	0	9234	49	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 3.

The worst 5 of 49 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:789:GLU:HG3	1:B:818:PRO:HG3	1.79	0.64
1:A:904:ARG:NH1	2:D:142:ASP:OD1	2.37	0.58
1:A:648:LYS:HB3	1:B:717:LYS:HD3	1.86	0.57
1:A:778:MET:HG3	1:A:828:LEU:HB3	1.88	0.56
1:A:511:TYR:HB2	1:A:530:ASP:HB3	1.88	0.56

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

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	419/478 (88%)	404 (96%)	14 (3%)	1 (0%)	56	86
1	B	428/478 (90%)	414 (97%)	13 (3%)	1 (0%)	56	86
1	C	293/478 (61%)	275 (94%)	17 (6%)	1 (0%)	50	82
2	D	8/13 (62%)	8 (100%)	0	0	100	100
2	E	8/13 (62%)	8 (100%)	0	0	100	100
All	All	1156/1460 (79%)	1109 (96%)	44 (4%)	3 (0%)	50	82

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	749	TYR
1	B	749	TYR
1	C	749	TYR

### 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	377/422 (89%)	377 (100%)	0	100	100
1	B	384/422 (91%)	384 (100%)	0	100	100
1	C	267/422 (63%)	267 (100%)	0	100	100
2	D	9/12 (75%)	9 (100%)	0	100	100
2	E	9/12 (75%)	9 (100%)	0	100	100
All	All	1046/1290 (81%)	1046 (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. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	510	GLN
1	A	634	HIS
1	B	559	HIS

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

There are no ligands in this entry.

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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	424/478 (88%)	0.01	8 (1%) 64 70	29, 48, 93, 123	0
1	B	431/478 (90%)	-0.04	10 (2%) 57 64	29, 50, 97, 142	0
1	C	296/478 (61%)	0.30	32 (10%) 6 6	34, 59, 115, 157	0
2	D	10/13 (76%)	1.92	6 (60%) 0 0	81, 100, 116, 117	0
2	E	10/13 (76%)	3.92	10 (100%) 0 0	105, 118, 134, 139	0
All	All	1171/1460 (80%)	0.11	66 (5%) 24 25	29, 52, 105, 157	0

The worst 5 of 66 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	E	149	SER	5.0
1	C	733	GLY	4.9
2	E	140	PHE	4.9
1	C	736	GLU	4.8
2	E	144	LEU	4.6

### 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

### 6.4 Ligands ⓘ

There are no ligands in this entry.

## 6.5 Other polymers ⓘ

There are no such residues in this entry.