



wwPDB X-ray Structure Validation Summary Report ⓘ

May 26, 2020 – 11:36 pm BST

PDB ID : 4C8S
Title : Crystal structure of the C-terminal region of yeast Ctf4
Authors : Simon, A.C.; Pellegrini, L.
Deposited on : 2013-10-01
Resolution : 3.00 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

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

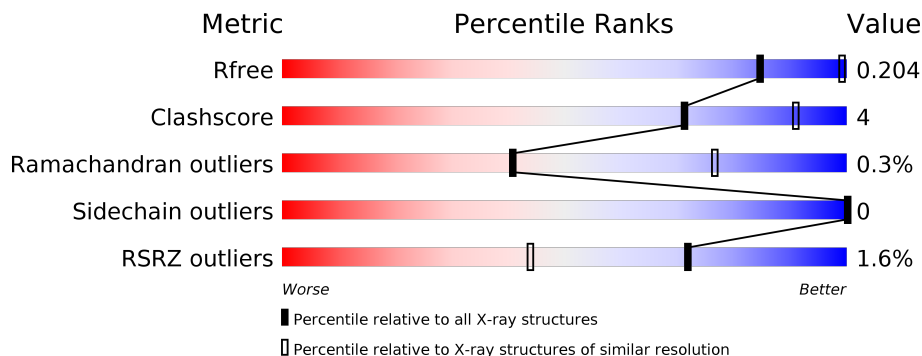
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 ≥ 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	478	 2% 81% 8% 11%
1	B	478	 2% 83% 8% 9%
1	C	478	 2% 53% 9% 38%

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 9447 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 DNA POLYMERASE ALPHA-BINDING PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	425	Total 3425	C 2199	N 568	O 643	S 15	0	1	0
1	B	437	Total 3519	C 2258	N 585	O 660	S 16	0	1	0
1	C	296	Total 2405	C 1562	N 392	O 440	S 11	0	1	0

There are 63 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	450	MET	-	expression tag	UNP Q01454
A	451	GLY	-	expression tag	UNP Q01454
A	452	SER	-	expression tag	UNP Q01454
A	453	SER	-	expression tag	UNP Q01454
A	454	HIS	-	expression tag	UNP Q01454
A	455	HIS	-	expression tag	UNP Q01454
A	456	HIS	-	expression tag	UNP Q01454
A	457	HIS	-	expression tag	UNP Q01454
A	458	HIS	-	expression tag	UNP Q01454
A	459	HIS	-	expression tag	UNP Q01454
A	460	SER	-	expression tag	UNP Q01454
A	461	GLN	-	expression tag	UNP Q01454
A	462	ASP	-	expression tag	UNP Q01454
A	463	PRO	-	expression tag	UNP Q01454
A	464	GLU	-	expression tag	UNP Q01454
A	465	ASN	-	expression tag	UNP Q01454
A	466	LEU	-	expression tag	UNP Q01454
A	467	TYR	-	expression tag	UNP Q01454
A	468	PHE	-	expression tag	UNP Q01454
A	469	GLN	-	expression tag	UNP Q01454
A	470	GLY	-	expression tag	UNP Q01454
B	450	MET	-	expression tag	UNP Q01454
B	451	GLY	-	expression tag	UNP Q01454

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	452	SER	-	expression tag	UNP Q01454
B	453	SER	-	expression tag	UNP Q01454
B	454	HIS	-	expression tag	UNP Q01454
B	455	HIS	-	expression tag	UNP Q01454
B	456	HIS	-	expression tag	UNP Q01454
B	457	HIS	-	expression tag	UNP Q01454
B	458	HIS	-	expression tag	UNP Q01454
B	459	HIS	-	expression tag	UNP Q01454
B	460	SER	-	expression tag	UNP Q01454
B	461	GLN	-	expression tag	UNP Q01454
B	462	ASP	-	expression tag	UNP Q01454
B	463	PRO	-	expression tag	UNP Q01454
B	464	GLU	-	expression tag	UNP Q01454
B	465	ASN	-	expression tag	UNP Q01454
B	466	LEU	-	expression tag	UNP Q01454
B	467	TYR	-	expression tag	UNP Q01454
B	468	PHE	-	expression tag	UNP Q01454
B	469	GLN	-	expression tag	UNP Q01454
B	470	GLY	-	expression tag	UNP Q01454
C	450	MET	-	expression tag	UNP Q01454
C	451	GLY	-	expression tag	UNP Q01454
C	452	SER	-	expression tag	UNP Q01454
C	453	SER	-	expression tag	UNP Q01454
C	454	HIS	-	expression tag	UNP Q01454
C	455	HIS	-	expression tag	UNP Q01454
C	456	HIS	-	expression tag	UNP Q01454
C	457	HIS	-	expression tag	UNP Q01454
C	458	HIS	-	expression tag	UNP Q01454
C	459	HIS	-	expression tag	UNP Q01454
C	460	SER	-	expression tag	UNP Q01454
C	461	GLN	-	expression tag	UNP Q01454
C	462	ASP	-	expression tag	UNP Q01454
C	463	PRO	-	expression tag	UNP Q01454
C	464	GLU	-	expression tag	UNP Q01454
C	465	ASN	-	expression tag	UNP Q01454
C	466	LEU	-	expression tag	UNP Q01454
C	467	TYR	-	expression tag	UNP Q01454
C	468	PHE	-	expression tag	UNP Q01454
C	469	GLN	-	expression tag	UNP Q01454
C	470	GLY	-	expression tag	UNP Q01454

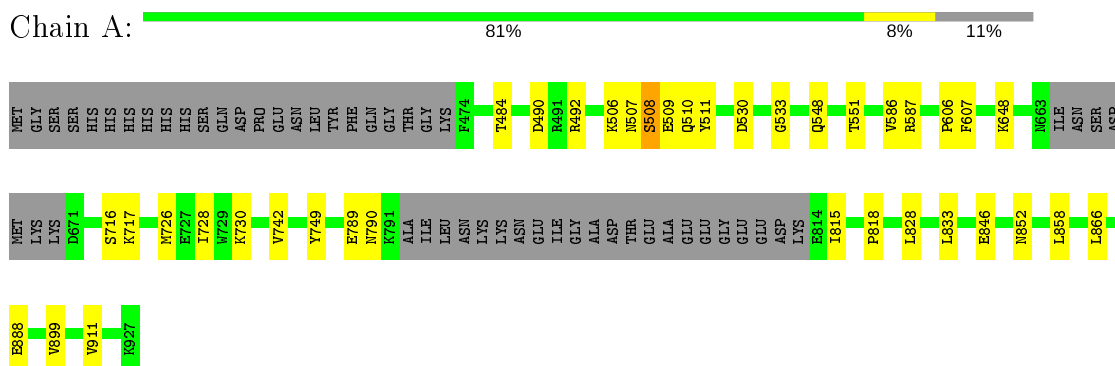
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	39	Total 39	O 39	0	0
2	B	42	Total 42	O 42	0	0
2	C	17	Total 17	O 17	0	0

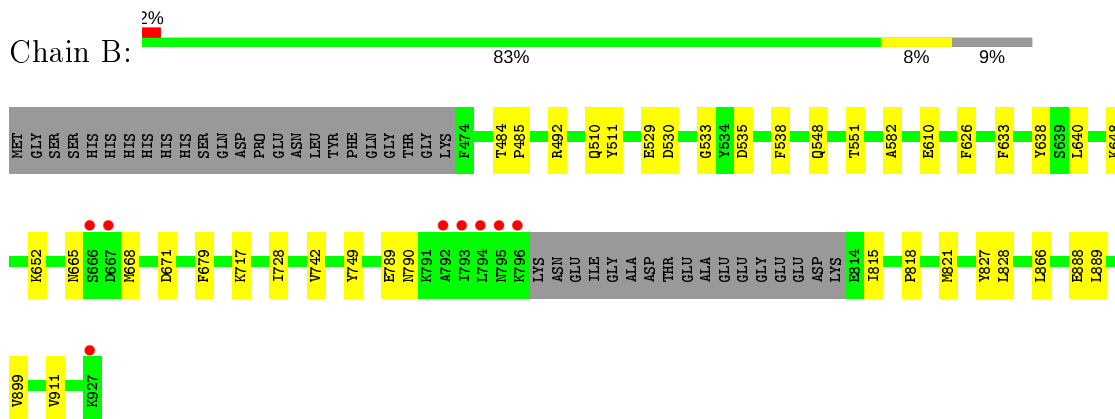
3 Residue-property plots [i](#)

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.

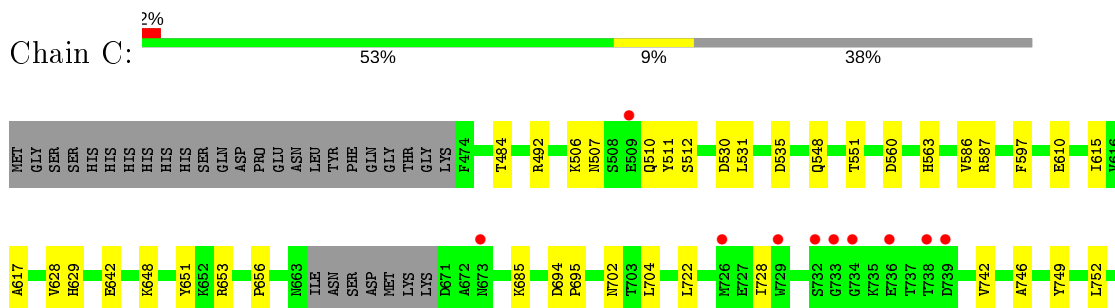
- Molecule 1: DNA POLYMERASE ALPHA-BINDING PROTEIN



- Molecule 1: DNA POLYMERASE ALPHA-BINDING PROTEIN



- Molecule 1: DNA POLYMERASE ALPHA-BINDING PROTEIN



I755	LYS
P769	VAL
M774	SER
E775	GLU
I776	LEU
	ARG
	MET
	PRO
	VAL
	THR
	PHE
	VAL
	LEU
	VAL
	LYS
	SER
	ASN
	LYS
	GLU
	LEU
	LEU
	LEU
	LEU
	LEU
	GLY
	GLU
	ASN
	ASN
	LYS
	LYS
	ALA
	ALA
	ILE
	LEU
	LEU
	ASN
	ASN
	LYS
	LYS
	ASN
	ASN
	GLU
	GLY
	ILE
	GLY
	ALA
	ALA
	ASP
	ASP
	THR
	THR
	LYS
	GLU
	ALA
	ALA
	GLU
	GLY
	LEU
	LEU
	LEU
	ARG
	LEU
	LEU
	LEU
	PHE
	LEU
	LEU
	ALA
	SER
	ALA
	ALA
	CYS
	SER
	ILE
	ILE
	GLN
	ASP
	GLN
	ILE
	ILE
	GLN
	ASN
	ASN
	VAL
	VAL
	GLU
	SER
	MET
	SER
	VAL
	ALA
	ALA
	GLU
	GLU
	LEU
	LEU
	LEU
	ALA
	ALA
	LEU
	SER
	LEU
	LEU
	ALA
	ALA
	HIS
	TYR
	LEU
	LEU
	ARG

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

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

4 Data and refinement statistics

Property	Value	Source
Space group	P 2 21 21	Depositor
Cell constants a, b, c, α , β , γ	89.06Å 100.25Å 219.52Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.87 – 3.00 48.87 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.9 (48.87-3.00) 99.9 (48.87-3.00)	Depositor EDS
R_{merge}	0.19	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.37 (at 3.01Å)	Xtrriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.167 , 0.205 0.167 , 0.204	Depositor DCC
R_{free} test set	2014 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å ²)	55.1	Xtrriage
Anisotropy	0.223	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 50.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	9447	wwPDB-VP
Average B, all atoms (Å ²)	51.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.48% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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	0.23	0/3510	0.40	0/4752
1	B	0.23	0/3605	0.40	0/4879
1	C	0.23	0/2481	0.39	0/3370
All	All	0.23	0/9596	0.40	0/13001

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	3425	0	3369	22	0
1	B	3519	0	3477	25	0
1	C	2405	0	2326	23	0
2	A	39	0	0	0	1
2	B	42	0	0	0	1
2	C	17	0	0	0	0
All	All	9447	0	9172	65	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 4.

The worst 5 of 65 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:C:510:GLN:HG2	1:C:531:LEU:HD23	1.74	0.69
1:C:728:ILE:HD11	1:C:742:VAL:HG23	1.79	0.65
1:C:722:LEU:HD21	1:C:774:MET:HE2	1.84	0.59
1:A:728:ILE:HD11	1:A:742:VAL:HG23	1.83	0.58
1:B:899:VAL:HG13	1:B:911:VAL:HG13	1.86	0.57

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:A:2038:HOH:O	2:B:2027:HOH:O[3_554]	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	420/478 (88%)	405 (96%)	13 (3%)	2 (0%)	29	68
1	B	434/478 (91%)	418 (96%)	15 (4%)	1 (0%)	47	82
1	C	293/478 (61%)	277 (94%)	15 (5%)	1 (0%)	41	76
All	All	1147/1434 (80%)	1100 (96%)	43 (4%)	4 (0%)	41	76

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	749	TYR
1	B	749	TYR
1	C	749	TYR
1	A	508	SER

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	378/422 (90%)	378 (100%)	0	100	100
1	B	389/422 (92%)	389 (100%)	0	100	100
1	C	267/422 (63%)	267 (100%)	0	100	100
All	All	1034/1266 (82%)	1034 (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 (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	507	ASN
1	C	555	GLN

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

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 [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, 95th 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 ²)	Q<0.9
1	A	425/478 (88%)	-0.39	0 100 100	23, 42, 89, 118	0
1	B	437/478 (91%)	-0.35	8 (1%) 68 40	22, 44, 97, 141	0
1	C	296/478 (61%)	-0.27	10 (3%) 45 19	28, 49, 108, 145	0
All	All	1158/1434 (80%)	-0.34	18 (1%) 72 44	22, 45, 98, 145	0

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	795	ASN	3.6
1	B	794	LEU	3.6
1	B	666	SER	3.4
1	C	732	SER	3.3
1	B	927	LYS	3.2

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

There are no such residues in this entry.