



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 05:25 AM GMT

PDB ID : 2QLZ  
Title : Crystal Structure of Transcription Factor PF0095 from *Pyrococcus furiosus*  
Authors : Yang, H.; Lipscomb, G.L.; Scott, R.A.; Rose, J.P.; Wang, B.C.  
Deposited on : 2007-07-13  
Resolution : 2.50 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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
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) : trunk26865

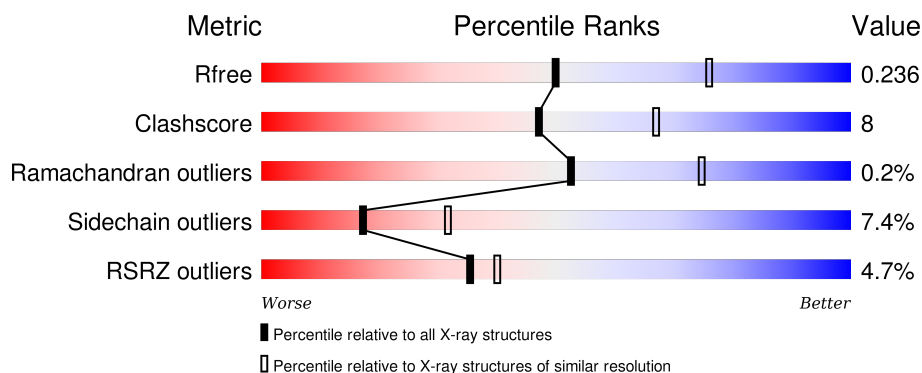
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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	3553 (2.50-2.50)
Clashscore	102246	4242 (2.50-2.50)
Ramachandran outliers	100387	4156 (2.50-2.50)
Sidechain outliers	100360	4158 (2.50-2.50)
RSRZ outliers	91569	3562 (2.50-2.50)

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	232	<div> <div>72% 19% 6%</div> <div>72% 19% 6%</div> </div>
1	B	232	<div> <div>7% 67% 25% 5%</div> <div>7% 67% 25% 5%</div> </div>
1	C	232	<div> <div>6% 75% 16% 8%</div> <div>6% 75% 16% 8%</div> </div>
1	D	232	<div> <div>3% 75% 16% 7%</div> <div>3% 75% 16% 7%</div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 7059 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 Transcription Factor PF0095.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	219	Total	C	N	O	S	Se	0	0	0
			1765	1132	286	338	2	7			
1	B	220	Total	C	N	O	S	Se	0	0	0
			1748	1123	277	338	2	8			
1	C	213	Total	C	N	O	S	Se	0	0	0
			1710	1098	271	332	2	7			
1	D	216	Total	C	N	O	S	Se	0	0	0
			1736	1115	277	334	2	8			

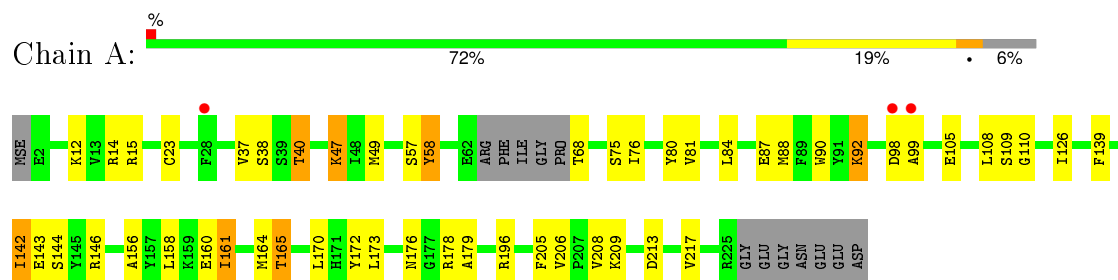
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	27	Total	O	0	0
			27	27		
2	B	29	Total	O	0	0
			29	29		
2	C	23	Total	O	0	0
			23	23		
2	D	21	Total	O	0	0
			21	21		

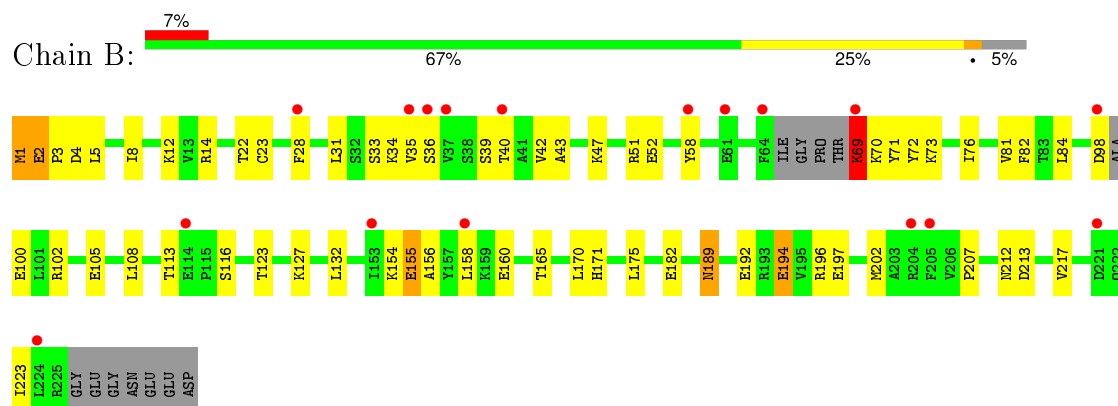
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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 ( $\text{RSRZ} > 2$ ). 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.

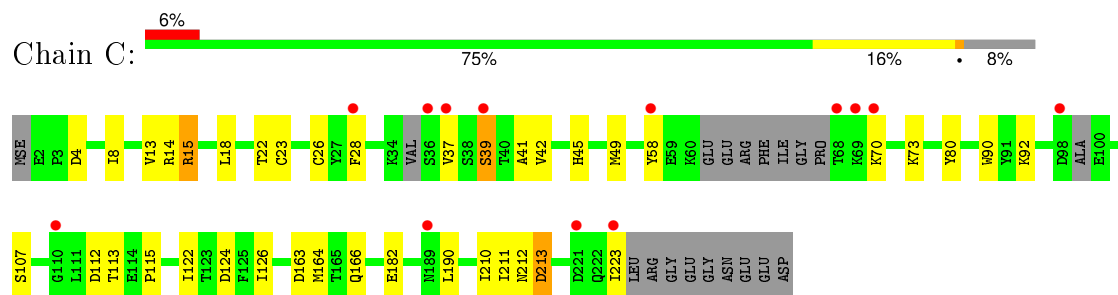
#### • Molecule 1: Transcription Factor PF0095



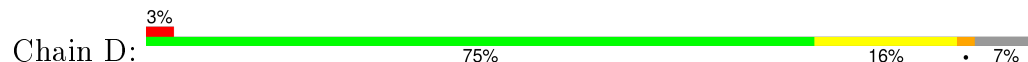
#### • Molecule 1: Transcription Factor PF0095

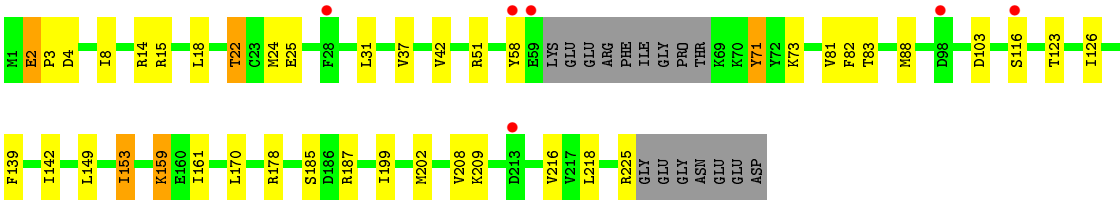


#### • Molecule 1: Transcription Factor PF0095



#### • Molecule 1: Transcription Factor PF0095





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	52.13Å 174.70Å 52.25Å 90.00° 93.11° 90.00°	Depositor
Resolution (Å)	19.43 – 2.50 19.43 – 2.50	Depositor EDS
% Data completeness (in resolution range)	100.0 (19.43-2.50) 94.5 (19.43-2.50)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	11.52 (at 2.49Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.226 , 0.271 0.220 , 0.236	Depositor DCC
$R_{free}$ test set	1534 reflections (5.33%)	DCC
Wilson B-factor (Å <sup>2</sup> )	30.4	Xtriage
Anisotropy	0.032	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 41.8	EDS
Estimated twinning fraction	0.033 for l,-k,h	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	1 of 30316 reflections (0.003%)	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	7059	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.51% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.54	0/1783	0.65	0/2386
1	B	0.58	2/1765 (0.1%)	0.66	2/2365 (0.1%)
1	C	0.57	1/1726 (0.1%)	0.63	2/2309 (0.1%)
1	D	0.45	0/1754	0.61	0/2349
All	All	0.54	3/7028 (0.0%)	0.64	4/9409 (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
1	A	0	2
1	B	0	2
1	C	0	1
All	All	0	5

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	41	ALA	CA-CB	-5.92	1.40	1.52
1	B	72	TYR	CD2-CE2	-5.62	1.30	1.39
1	B	72	TYR	CD1-CE1	-5.05	1.31	1.39

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	2	GLU	CB-CA-C	8.28	126.96	110.40
1	C	213	ASP	CB-CA-C	7.49	125.38	110.40
1	B	1	MSE	N-CA-C	5.81	126.69	111.00
1	C	15	ARG	NE-CZ-NH2	-5.57	117.51	120.30

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	109	SER	Peptide
1	A	110	GLY	Peptide
1	B	69	LYS	Peptide
1	B	70	LYS	Peptide
1	C	39	SER	Peptide

## 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	1765	0	1792	35	0
1	B	1748	0	1739	36	0
1	C	1710	0	1712	30	0
1	D	1736	0	1750	22	0
2	A	27	0	0	2	0
2	B	29	0	0	0	0
2	C	23	0	0	0	0
2	D	21	0	0	0	0
All	All	7059	0	6993	105	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 8.

The worst 5 of 105 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:210:ILE:HD11	1:C:213:ASP:O	1.26	1.32
1:B:1:MSE:O	1:B:2:GLU:HB2	1.45	1.11
1:C:210:ILE:CD1	1:C:213:ASP:O	2.13	0.96
1:C:18:LEU:HD21	1:C:49:MSE:HE1	1.59	0.83
1:A:164:MSE:O	1:A:165:THR:HG23	1.80	0.82

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	215/232 (93%)	203 (94%)	12 (6%)	0	100	100
1	B	214/232 (92%)	201 (94%)	11 (5%)	2 (1%)	21	37
1	C	205/232 (88%)	196 (96%)	9 (4%)	0	100	100
1	D	212/232 (91%)	204 (96%)	8 (4%)	0	100	100
All	All	846/928 (91%)	804 (95%)	40 (5%)	2 (0%)	52	75

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	35	VAL
1	B	207	PRO

### 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	196/204 (96%)	180 (92%)	16 (8%)	14	27
1	B	190/204 (93%)	173 (91%)	17 (9%)	12	23
1	C	189/204 (93%)	181 (96%)	8 (4%)	36	62
1	D	191/204 (94%)	175 (92%)	16 (8%)	14	25
All	All	766/816 (94%)	709 (93%)	57 (7%)	17	31

5 of 57 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	116	SER
1	B	194	GLU
1	D	178	ARG
1	B	154	LYS
1	B	165	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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	212/232 (91%)	0.11	3 (1%) 78 80	9, 29, 44, 48	0
1	B	212/232 (91%)	0.30	17 (8%) 15 16	9, 29, 47, 50	0
1	C	206/232 (88%)	0.32	13 (6%) 23 26	11, 29, 45, 50	0
1	D	208/232 (89%)	0.09	6 (2%) 55 60	11, 29, 45, 48	0
All	All	838/928 (90%)	0.20	39 (4%) 35 40	9, 29, 45, 50	0

The worst 5 of 39 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	36	SER	4.8
1	C	68	THR	4.8
1	B	28	PHE	4.7
1	C	28	PHE	4.7
1	D	98	ASP	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 ⓘ

There are no such residues in this entry.