



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 10:23 PM GMT

PDB ID : 1TF6  
Title : CO-CRYSTAL STRUCTURE OF XENOPUS TFIID ZINC FINGER DOMAIN BOUND TO THE 5S RIBOSOMAL RNA GENE INTERNAL CONTROL REGION  
Authors : Nolte, R.T.; Conlin, R.M.; Harrison, S.C.; Brown, R.S.  
Deposited on : 1998-03-02  
Resolution : 3.10 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 : 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

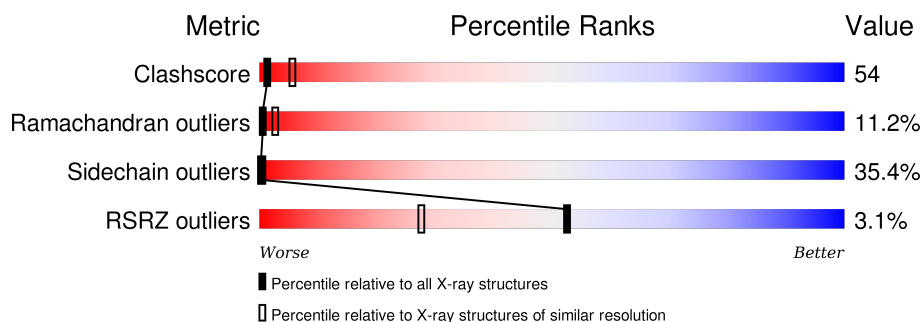
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.10 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))
Clashscore	102246	1222 (3.14-3.06)
Ramachandran outliers	100387	1174 (3.14-3.06)
Sidechain outliers	100360	1174 (3.14-3.06)
RSRZ outliers	91569	1119 (3.14-3.06)

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	B	31	100%
1	E	31	100%
2	C	31	16% 84%
2	F	31	13% 87%
3	A	190	2% 23% 46% 22% 6%
3	D	190	6% 25% 42% 25% . .

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 5495 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 DNA chain called DNA (5'-D(\*AP\*CP\*GP\*GP\*GP\*CP\*CP\*TP\*GP\*GP\*TP\*TP\*AP\*GP\*TP\*AP\*C P\*CP\*TP\*GP\*GP\*AP\* TP\*GP\*GP\*GP\*AP\*GP\*AP\*CP\*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	31	Total	C	N	O	P	0	0	0
			640	303	123	184	30			
1	E	31	Total	C	N	O	P	0	0	0
			640	303	123	184	30			

- Molecule 2 is a DNA chain called DNA (5'-D(\*TP\*GP\*GP\*TP\*CP\*TP\*CP\*CP\*AP\*TP\*CP\*CP\*AP\*GP\*GP\*T P\*AP\*CP\*TP\*AP\*AP\* CP\*CP\*AP\*GP\*GP\*CP\*CP\*CP\*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	31	Total	C	N	O	P	0	0	0
			625	298	113	184	30			
2	F	31	Total	C	N	O	P	0	0	0
			625	298	113	184	30			

- Molecule 3 is a protein called PROTEIN (TRANSCRIPTION FACTOR IIIA).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	179	Total	C	N	O	S	0	0	0
			1466	925	272	253	16			
3	D	182	Total	C	N	O	S	0	0	0
			1487	940	275	256	16			

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	6	Total	Zn	0	0
			6	6		
4	D	6	Total	Zn	0	0
			6	6		

### 3 Residue-property plots

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 ( $RSRZ > 2$ ). 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 (5'-D(\*AP\*CP\*GP\*GP\*GP\*CP\*CP\*TP\*GP\*GP\*TP\*TP\*AP\*GP\*TP\*AP\*C P\*CP\*TP\*GP\*GP\*AP\* TP\*GP\*GP\*GP\*AP\*GP\*AP\*CP\*C)-3')

Chain B:  100%

A1 C2 G3 G4 G5 G6 G7 T8 G9 G10 T11 T12 T13 A14 G15 T16 A17 C18 T19 G20 G21 A22 T23 G24 G25 G26 A27 G28 A29 C30 C31

- Molecule 1: DNA (5'-D(\*AP\*CP\*GP\*GP\*GP\*CP\*CP\*TP\*GP\*GP\*TP\*TP\*AP\*GP\*TP\*AP\*C P\*CP\*TP\*GP\*GP\*AP\* TP\*GP\*GP\*GP\*AP\*GP\*AP\*CP\*C)-3')

Chain E:  100%

A1 C2 G3 G4 G5 G6 G7 T8 G9 G10 T11 T12 T13 A14 G15 T16 A17 C18 T19 G20 G21 A22 T23 G24 G25 G26 A27 G28 A29 C30 C31

- Molecule 2: DNA (5'-D(\*TP\*GP\*GP\*TP\*CP\*TP\*CP\*CP\*AP\*TP\*CP\*CP\*AP\*GP\*GP\*T P\*AP\*CP\*TP\*AP\*AP\* CP\*CP\*AP\*GP\*GP\*CP\*CP\*CP\*G)-3')

Chain C:  16% 84%

T33 G34 G35 T36 C37 T38 C39 C40 C41 A42 T43 T44 G47 G48 T49 A50 C51 T52 A53 A54 C55 C56 A57 G58 G59 C60 C61 G62 G63

- Molecule 2: DNA (5'-D(\*TP\*GP\*GP\*TP\*CP\*TP\*CP\*CP\*AP\*TP\*CP\*CP\*AP\*GP\*GP\*T P\*AP\*CP\*TP\*AP\*AP\* CP\*CP\*AP\*GP\*GP\*CP\*CP\*CP\*G)-3')

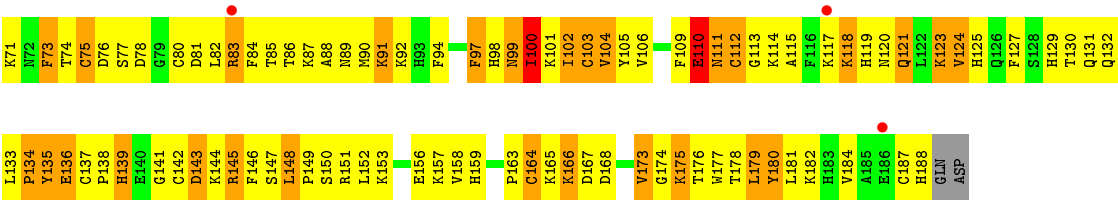
Chain F:  13% 87%

T33 G34 G35 T36 C37 T38 C39 C40 C41 A42 T43 T44 G47 G48 T49 A50 C51 T52 A53 A54 C55 C56 A57 G58 G59 C60 C61 G62 G63

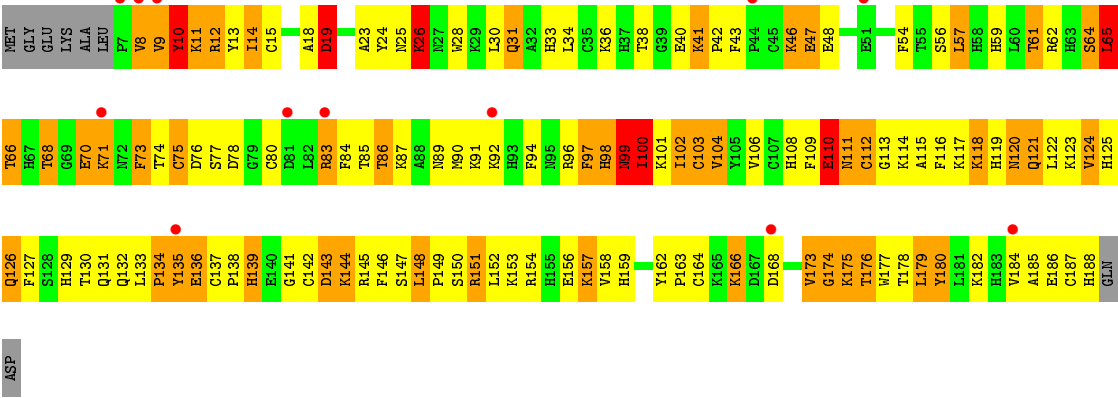
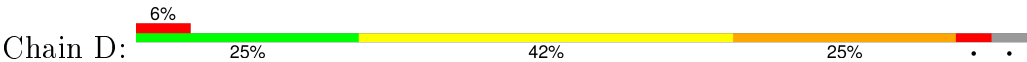
- Molecule 3: PROTEIN (TRANSCRIPTION FACTOR IIIA)

Chain A:  2% 23% 46% 22% 6%

MET GLY GLU LYS ALA LEU PRO VAL Y10 Y11 Y12 Y13 Y14 C15 A18 A19 Y24 N25 K26 N27 N28 K29 L30 Q31 L34 C35 T38 G39 E40 K41 P42 F43 F44 C45 K46 E47 E48 F54 T55 S56 L57 H58 H59 L60 T61 R62 H63 S64 S65 L66 T66 H67 T68 G69 E70



● Molecule 3: PROTEIN (TRANSCRIPTION FACTOR IIIA)



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	64.18 Å   64.71 Å   78.03 Å 90.07°   92.98°   102.95°	Depositor
Resolution (Å)	8.00 – 3.10 24.83 – 3.13	Depositor EDS
% Data completeness (in resolution range)	81.4 (8.00-3.10) 88.7 (24.83-3.13)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.79 (at 3.11 Å)	Xtriage
Refinement program	X-PLOR 3.851	Depositor
R, $R_{free}$	0.308 ,      0.363 0.322 ,      (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	DCC
Wilson B-factor (Å <sup>2</sup> )	68.8	Xtriage
Anisotropy	0.534	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.23 , 49.4	EDS
Estimated twinning fraction	0.000 for -k,-h,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 19062 reflections	Xtriage
$F_o, F_c$ correlation	0.88	EDS
Total number of atoms	5495	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	65.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.35% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN

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	B	0.43	0/719	0.76	0/1110
1	E	0.42	0/719	0.75	0/1110
2	C	0.45	0/699	0.76	0/1075
2	F	0.45	0/699	0.75	0/1075
3	A	0.32	0/1511	0.46	0/2027
3	D	0.32	0/1533	0.48	0/2058
All	All	0.38	0/5880	0.63	0/8455

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	B	640	0	349	67	0
1	E	640	0	349	73	0
2	C	625	0	349	58	0
2	F	625	0	349	52	0
3	A	1466	0	1400	139	0
3	D	1487	0	1426	159	0
4	A	6	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	6	0	0	0	0
All	All	5495	0	4222	529	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 54.

The worst 5 of 529 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:E:27:DA:H2''	1:E:28:DG:H5'	1.26	1.15
1:B:27:DA:H2''	1:B:28:DG:H5'	1.29	1.13
1:B:25:DG:H2''	1:B:26:DG:H5''	1.22	1.10
2:F:40:DC:H2''	2:F:41:DC:H5''	1.29	1.10
2:C:40:DC:H2''	2:C:41:DC:H5''	1.33	1.08

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	
3	A	177/190 (93%)	109 (62%)	50 (28%)	18 (10%)	1	4
3	D	180/190 (95%)	114 (63%)	44 (24%)	22 (12%)	0	2
All	All	357/380 (94%)	223 (62%)	94 (26%)	40 (11%)	0	3

5 of 40 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	100	ILE
3	A	173	VAL
3	D	10	TYR

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
3	D	11	LYS
3	D	12	ARG

### 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	
3	A	161/170 (95%)	106 (66%)	55 (34%)	0	0
3	D	164/170 (96%)	104 (63%)	60 (37%)	0	0
All	All	325/340 (96%)	210 (65%)	115 (35%)	0	0

5 of 115 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	179	LEU
3	D	41	LYS
3	D	159	HIS
3	A	180	TYR
3	D	14	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
3	A	59	HIS
3	A	120	ASN
3	D	59	HIS
3	D	120	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules 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 ⓘ

Of 12 ligands modelled in this entry, 12 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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, 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	B	31/31 (100%)	-0.32	0 100 100	15, 65, 85, 95	0
1	E	31/31 (100%)	-0.47	0 100 100	18, 66, 86, 95	0
2	C	31/31 (100%)	-0.32	0 100 100	32, 64, 88, 90	0
2	F	31/31 (100%)	-0.40	0 100 100	34, 65, 88, 90	0
3	A	179/190 (94%)	-0.17	3 (1%) 73 52	8, 69, 99, 100	0
3	D	182/190 (95%)	0.09	12 (6%) 22 8	9, 71, 99, 100	0
All	All	485/504 (96%)	-0.12	15 (3%) 52 28	8, 69, 99, 100	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	D	168	ASP	5.1
3	D	51	GLU	3.9
3	D	7	PRO	3.7
3	A	186	GLU	3.5
3	D	8	VAL	3.3

### 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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
4	ZN	D	202	1/1	0.85	0.18	-0.60	91,91,91,91	0
4	ZN	A	202	1/1	0.94	0.11	-1.17	64,64,64,64	0
4	ZN	A	206	1/1	0.70	0.08	-1.21	100,100,100,100	0
4	ZN	D	201	1/1	0.99	0.12	-1.22	12,12,12,12	0
4	ZN	A	204	1/1	0.98	0.10	-1.28	24,24,24,24	0
4	ZN	A	201	1/1	0.99	0.13	-1.37	27,27,27,27	0
4	ZN	D	204	1/1	0.93	0.10	-2.32	40,40,40,40	0
4	ZN	A	205	1/1	0.99	0.16	-	32,32,32,32	0
4	ZN	D	205	1/1	0.96	0.13	-	49,49,49,49	0
4	ZN	D	206	1/1	0.98	0.02	-	100,100,100,100	0
4	ZN	D	203	1/1	0.97	0.11	-	100,100,100,100	0
4	ZN	A	203	1/1	0.91	0.12	-	84,84,84,84	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.