



# wwPDB X-ray Structure Validation Summary Report

Feb 27, 2014 – 03:23 PM GMT

PDB ID : 1F3U  
Title : CRYSTAL STRUCTURE OF THE RAP30/74 INTERACTION DOMAINS  
OF HUMAN TFIIF  
Authors : Gaiser, F.; Tan, S.; Richmond, T.J.  
Deposited on : 2000-06-06  
Resolution : 1.70 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>

---

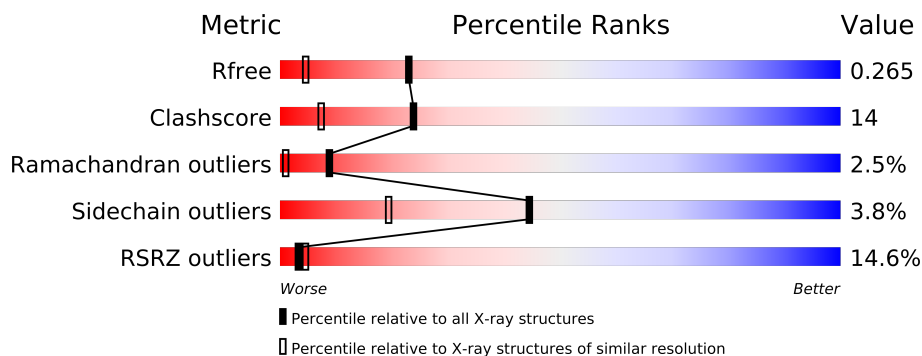
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
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)	:	stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 1.70 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	2456 (1.70-1.70)
Clashscore	79885	2929 (1.70-1.70)
Ramachandran outliers	78287	2878 (1.70-1.70)
Sidechain outliers	78261	2878 (1.70-1.70)
RSRZ outliers	66119	2456 (1.70-1.70)

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	118	
1	C	118	
1	E	118	
1	G	118	
2	B	171	
2	D	171	
2	F	171	
2	H	171	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 9192 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 TRANSCRIPTION INITIATION FACTOR IIF, BETA SUB-UNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	118	Total	C	N	O	S	16	0	0
			895	557	163	174	1			
1	C	118	Total	C	N	O	S	21	0	0
			894	557	163	173	1			
1	E	118	Total	C	N	O	S	0	0	0
			895	557	163	174	1			
1	G	118	Total	C	N	O	S	5	0	0
			895	557	163	174	1			

- Molecule 2 is a protein called TRANSCRIPTION INITIATION FACTOR IIF, ALPHA SUB-UNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	139	Total	C	N	O	S	22	0	0
			1158	737	213	205	3			
2	D	153	Total	C	N	O	S	14	0	0
			1284	812	235	234	3			
2	F	154	Total	C	N	O	S	90	8	0
			1310	827	240	240	3			
2	H	137	Total	C	N	O	S	52	6	0
			1169	743	215	208	3			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	95	Total	O	0	0
			95	95		
3	B	89	Total	O	0	0
			89	89		
3	C	52	Total	O	0	0
			52	52		

*Continued on next page...*

*Continued from previous page...*

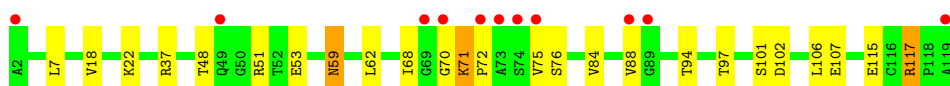
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	D	91	Total 91	O 91	0	0
3	E	118	Total 118	O 118	0	0
3	F	119	Total 119	O 119	0	0
3	G	71	Total 71	O 71	0	0
3	H	57	Total 57	O 57	0	0

### 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.

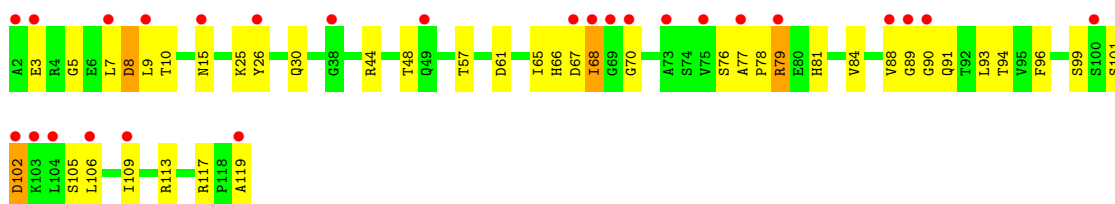
- Molecule 1: TRANSCRIPTION INITIATION FACTOR IIF, BETA SUBUNIT

Chain A: 



- Molecule 1: TRANSCRIPTION INITIATION FACTOR IIF, BETA SUBUNIT

Chain C: 



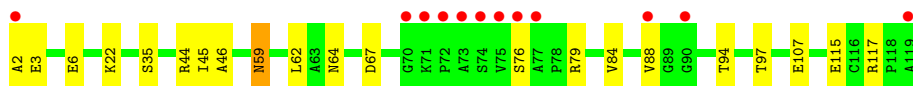
- Molecule 1: TRANSCRIPTION INITIATION FACTOR IIF, BETA SUBUNIT

Chain E: 



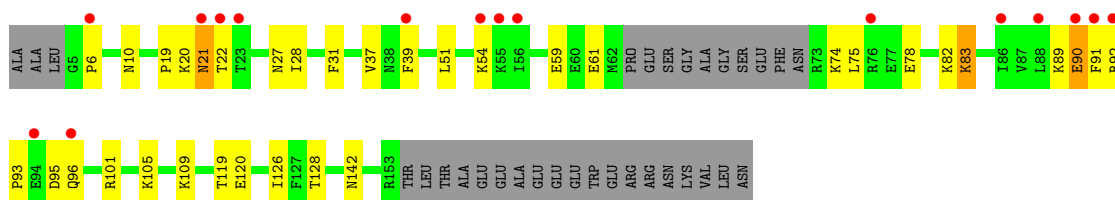
- Molecule 1: TRANSCRIPTION INITIATION FACTOR IIF, BETA SUBUNIT

Chain G: 



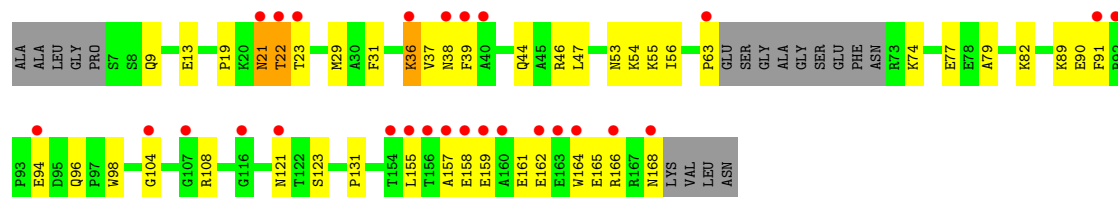
- Molecule 2: TRANSCRIPTION INITIATION FACTOR IIF, ALPHA SUBUNIT

Chain B: 



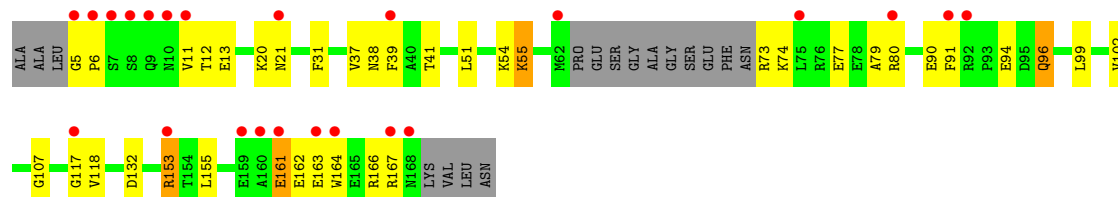
• Molecule 2: TRANSCRIPTION INITIATION FACTOR IIF, ALPHA SUBUNIT

Chain D: 



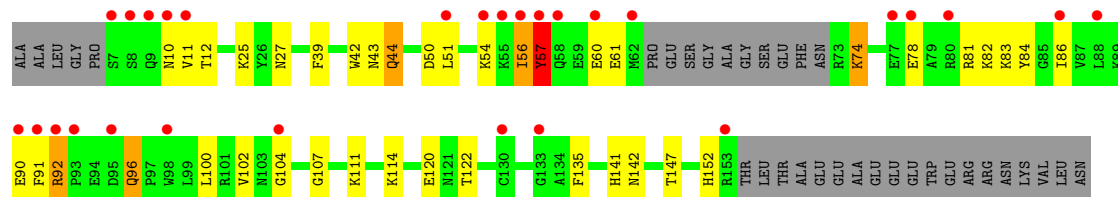
• Molecule 2: TRANSCRIPTION INITIATION FACTOR IIF, ALPHA SUBUNIT

Chain F: 



• Molecule 2: TRANSCRIPTION INITIATION FACTOR IIF, ALPHA SUBUNIT

Chain H: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	48.02Å 72.29Å 82.27Å 104.51° 93.32° 104.32°	Depositor
Resolution (Å)	6.00 – 1.70 67.37 – 1.70	Depositor EDS
% Data completeness (in resolution range)	96.4 (6.00-1.70) 96.7 (67.37-1.70)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.03 (at 1.70Å)	Xtriage
Refinement program	CNS 0.9A	Depositor
R, $R_{free}$	0.225 , 0.260 0.233 , 0.265	Depositor DCC
$R_{free}$ test set	4381 reflections (4.19%)	DCC
Wilson B-factor (Å <sup>2</sup> )	29.7	Xtriage
Anisotropy	0.145	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 58.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 108913 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	9192	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 46.23 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 1.1801e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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	0.33	0/909	0.65	0/1227
1	C	0.30	0/908	0.67	0/1226
1	E	0.38	0/909	0.71	1/1227 (0.1%)
1	G	0.33	0/909	0.63	0/1227
2	B	0.35	0/1186	0.68	0/1594
2	D	0.31	0/1314	0.60	0/1768
2	F	0.38	0/1340	0.65	0/1804
2	H	0.34	0/1196	0.61	0/1608
All	All	0.34	0/8671	0.65	1/11681 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	38	GLY	N-CA-C	5.22	126.16	113.10

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	895	0	904	25	0
1	C	894	0	901	33	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	895	0	904	29	0
1	G	895	0	904	31	0
2	B	1158	0	1151	35	0
2	D	1284	0	1261	42	0
2	F	1310	0	1283	34	0
2	H	1169	0	1162	47	0
3	A	95	0	0	1	0
3	B	89	0	0	7	0
3	C	52	0	0	2	0
3	D	91	0	0	2	0
3	E	118	0	0	4	0
3	F	119	0	0	2	0
3	G	71	0	0	2	0
3	H	57	0	0	1	0
All	All	9192	0	8470	236	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 14.

The worst 5 of 236 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:D:56:ILE:HD11	2:D:161:GLU:HG2	1.52	0.88
1:A:59:ASN:HD21	1:A:62:LEU:HD13	1.40	0.86
2:B:51:LEU:HD22	2:B:91:PHE:HZ	1.43	0.83
2:D:21:ASN:HD22	2:D:21:ASN:C	1.84	0.80
1:G:59:ASN:HD22	1:G:59:ASN:C	1.87	0.78

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	116/118 (98%)	105 (90%)	6 (5%)	5 (4%)	4 0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	116/118 (98%)	103 (89%)	7 (6%)	6 (5%)	3	0
1	E	116/118 (98%)	114 (98%)	2 (2%)	0	100	100
1	G	116/118 (98%)	109 (94%)	7 (6%)	0	100	100
2	B	135/171 (79%)	128 (95%)	5 (4%)	2 (2%)	15	2
2	D	149/171 (87%)	139 (93%)	8 (5%)	2 (1%)	18	3
2	F	153/171 (90%)	140 (92%)	7 (5%)	6 (4%)	5	0
2	H	136/171 (80%)	121 (89%)	10 (7%)	5 (4%)	5	0
All	All	1037/1156 (90%)	959 (92%)	52 (5%)	26 (2%)	9	1

5 of 26 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	72	PRO
1	A	76	SER
2	B	119	THR
1	C	102	ASP
2	F	77	GLU

### 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	96/96 (100%)	94 (98%)	2 (2%)	66	45
1	C	95/96 (99%)	94 (99%)	1 (1%)	84	72
1	E	96/96 (100%)	92 (96%)	4 (4%)	40	16
1	G	96/96 (100%)	92 (96%)	4 (4%)	40	16
2	B	122/147 (83%)	117 (96%)	5 (4%)	41	17
2	D	135/147 (92%)	131 (97%)	4 (3%)	53	29
2	F	138/147 (94%)	130 (94%)	8 (6%)	28	9
2	H	124/147 (84%)	118 (95%)	6 (5%)	35	13
All	All	902/972 (93%)	868 (96%)	34 (4%)	44	19

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	48	THR
2	F	94	GLU
2	H	91	PHE
2	F	55	LYS
2	B	95	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	21	ASN
2	D	53	ASN
2	H	121	ASN
2	D	38	ASN
2	D	44	GLN

### 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	118/118 (100%)	0.79	11 (9%) 9 11	22, 39, 69, 80	6 (5%)
1	C	118/118 (100%)	1.27	26 (22%) 1 2	29, 52, 76, 85	8 (6%)
1	E	118/118 (100%)	0.58	8 (6%) 17 21	20, 32, 58, 75	0
1	G	118/118 (100%)	0.99	12 (10%) 7 10	21, 42, 69, 77	1 (0%)
2	B	135/171 (78%)	0.79	16 (11%) 5 7	21, 42, 68, 82	0
2	D	151/171 (88%)	1.06	27 (17%) 2 3	27, 49, 72, 79	0
2	F	145/171 (84%)	1.03	23 (15%) 3 4	19, 38, 70, 78	10 (6%)
2	H	133/171 (77%)	1.30	28 (21%) 1 2	25, 48, 76, 83	10 (7%)
All	All	1036/1156 (89%)	0.98	151 (14%) 3 4	19, 42, 73, 85	35 (3%)

The worst 5 of 151 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	74	SER	9.5
1	A	75	VAL	8.8
2	F	164	TRP	8.6
2	B	21	ASN	8.1
1	G	73	ALA	8.1

### 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.