wwPDB X-ray Structure Validation Summary Report (i)

Jan 23, 2021 – 02:32 PM EST

PDB ID	:	1VOL
Title	:	TFIIB (HUMAN CORE DOMAIN)/TBP (A.THALIANA)/TATA ELEMENT TERNARY COMPLEX
Authors	:	Nikolov, D.B.; Chen, H.; Halay, E.D.; Usheva, A.U.; Hisatake, K.; Lee, D.K.; Roeder, R.G.; Burley, S.K.
Deposited on	:	1996-04-29
Resolution	:	2.70 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 (i) symbol.

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

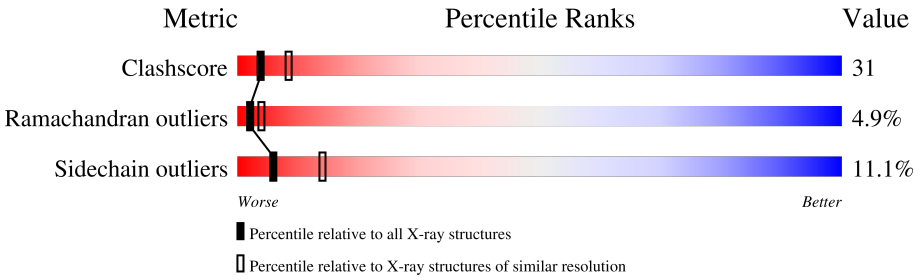
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.16

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 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))
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (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 of 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 <=5%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	C	16	38% 25% 38%
2	D	16	31% 63% 6%
3	A	204	38% 49% 12% •
4	B	200	46% 41% 6% • 7%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 3715 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(*GP*GP*CP*TP*AP*TP*AP*AP*AP*AP*GP*GP*GP*CP*TP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	C	16	332	158	67	92	15	0	0	0

- Molecule 2 is a DNA chain called DNA (5'-D(*CP*AP*GP*CP*CP*CP*TP*TP*TP*TP*AP*TP*AP*GP*CP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	D	16	318	154	53	96	15	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	A	204	1590	1002	285	291	12	0	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	134	LYS	ILE	conflict	UNP Q00403
A	143	ARG	LYS	conflict	UNP Q00403
A	145	ALA	VAL	conflict	UNP Q00403

- Molecule 4 is a protein called PROTEIN (TATA BINDING PROTEIN (TBP)).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	B	187	1475	962	251	254	8	0	0	0

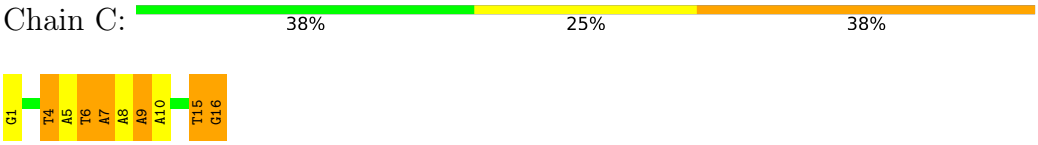
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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. 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. 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.

Note EDS was not executed.

• Molecule 1: DNA (5'-D(*GP*GP*CP*TP*AP*TP*AP*AP*AP*AP*GP*GP*GP*CP*TP*G)-3')

Chain C:

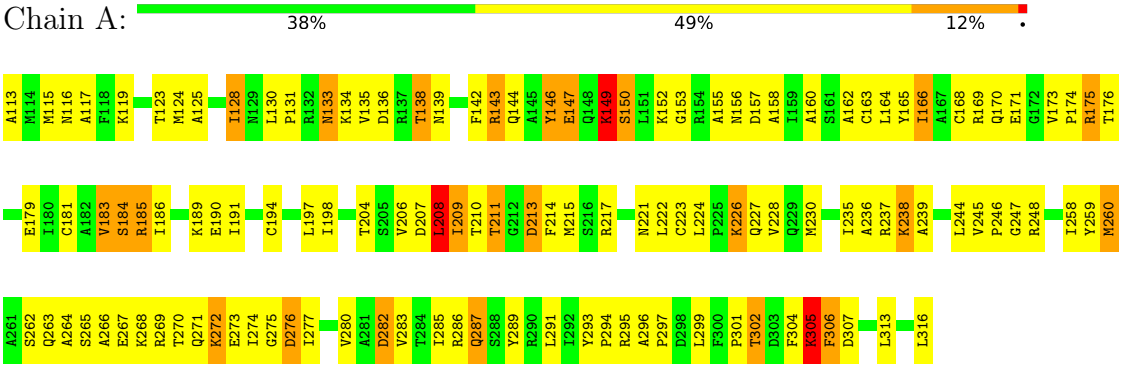


• Molecule 2: DNA (5'-D(*CP*AP*GP*CP*CP*CP*TP*TP*TP*TP*AP*TP*AP*GP*CP*C)-3')

Chain D:

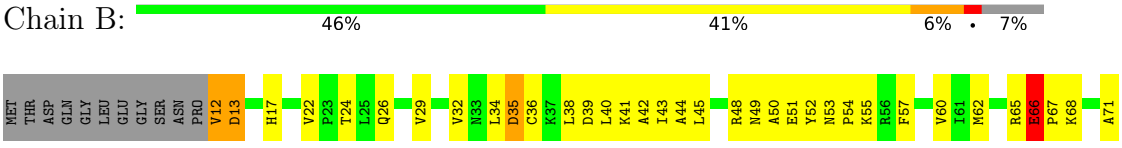
• Molecule 3: PROTEIN (TRANSCRIPTION FACTOR IIB (TFIIB))

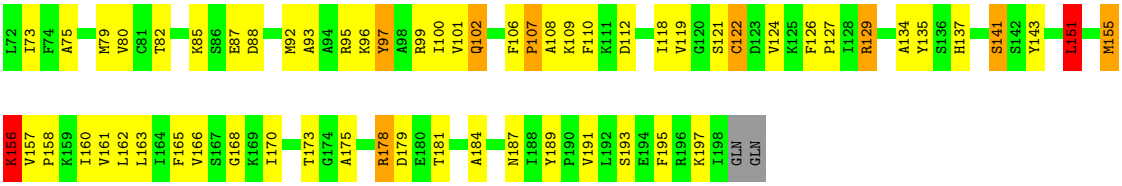
Chain A:



• Molecule 4: PROTEIN (TATA BINDING PROTEIN (TBP))

Chain B:




L72
I73
F74
A75
M79
V80
C81
T82
K85
S86
E87
D88
M92
A93
A94
R95
K96
Y97
A98
R99
I100
V101
Q102
F106
P107
A108
K109
F110
K111
D112
I118
V119
G120
S121
C122
D123
V124
K125
F126
P127
I128
R129
A134
Y135
S136
H137
S141
S142
Y143
L151
M155
K156
V157
P158
K159
I160
V161
L162
L163
I164
F165
V166
S167
G168
K169
I170
T173
G174
A175
R178
D179
E180
T181
A184
N187
I188
Y189
P190
V191
L192
S193
E194
F195
R196
K197
I198
GLN
GLN

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	69.90Å 78.00Å 134.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	8.00 – 2.70	Depositor
% Data completeness (in resolution range)	(Not available) (8.00-2.70)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
Refinement program	X-PLOR	Depositor
R, R_{free}	0.215 , 0.318	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	3715	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

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	C	1.57	3/374 (0.8%)	1.84	9/577 (1.6%)
2	D	1.60	1/354 (0.3%)	2.42	26/543 (4.8%)
3	A	0.64	0/1614	0.84	1/2175 (0.0%)
4	B	0.73	1/1505 (0.1%)	0.91	4/2026 (0.2%)
All	All	0.92	5/3847 (0.1%)	1.25	40/5321 (0.8%)

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	C	0	1
2	D	0	1
3	A	0	1
4	B	0	1
All	All	0	4

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	122	CYS	CB-SG	-6.64	1.71	1.82
1	C	6	DT	C4'-O4'	-5.58	1.39	1.45
2	D	108	DT	C4'-C3'	-5.41	1.47	1.52
1	C	15	DT	C5-C7	5.39	1.53	1.50
1	C	7	DA	C4'-O4'	-5.01	1.40	1.45

The worst 5 of 40 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	108	DT	O4'-C4'-C3'	-11.78	98.93	106.00
2	D	101	DC	O4'-C1'-N1	11.63	116.14	108.00
2	D	115	DC	O4'-C1'-C2'	-11.27	96.88	105.90

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	104	DC	O4'-C1'-N1	9.78	114.85	108.00
1	C	7	DA	O4'-C4'-C3'	-9.35	100.39	106.00

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	A	146	TYR	Sidechain
4	B	97	TYR	Sidechain
1	C	1	DG	Sidechain
2	D	112	DT	Sidechain

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	C	332	0	181	11	0
2	D	318	0	183	20	0
3	A	1590	0	1646	123	0
4	B	1475	0	1558	84	0
All	All	3715	0	3568	223	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 31.

The worst 5 of 223 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)
3:A:116:ASN:HA	3:A:119:LYS:HD3	1.39	0.99
4:B:102:GLN:HG3	4:B:108:ALA:HB3	1.47	0.96
3:A:224:LEU:HD22	3:A:228:VAL:HG11	1.49	0.95
3:A:155:ALA:HB3	3:A:158:ALA:HB2	1.51	0.92
3:A:134:LYS:O	3:A:138:THR:HG22	1.76	0.86

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	202/204 (99%)	165 (82%)	25 (12%)	12 (6%)	1 2
4	B	185/200 (92%)	153 (83%)	25 (14%)	7 (4%)	3 7
All	All	387/404 (96%)	318 (82%)	50 (13%)	19 (5%)	2 4

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	152	LYS
3	A	185	ARG
3	A	267	GLU
3	A	302	THR
3	A	305	LYS

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	172/172 (100%)	149 (87%)	23 (13%)	4 9
4	B	160/171 (94%)	146 (91%)	14 (9%)	10 23
All	All	332/343 (97%)	295 (89%)	37 (11%)	6 14

5 of 37 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	260	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	A	282	ASP
4	B	156	LYS
3	A	272	LYS
3	A	276	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
3	A	227	GLN
3	A	271	GLN
4	B	17	HIS
3	A	156	ASN
3	A	287	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 monosaccharides 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

6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.