



# wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 11:08 PM GMT

PDB ID : 1VTO  
Title : 1.9 Å RESOLUTION REFINED STRUCTURE OF TBP RECOGNIZING  
THE MINOR GROOVE OF TATAAAAG  
Authors : Kim, J.L.; Burley, S.K.  
Deposited on : 1996-09-06  
Resolution : 1.90 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>

---

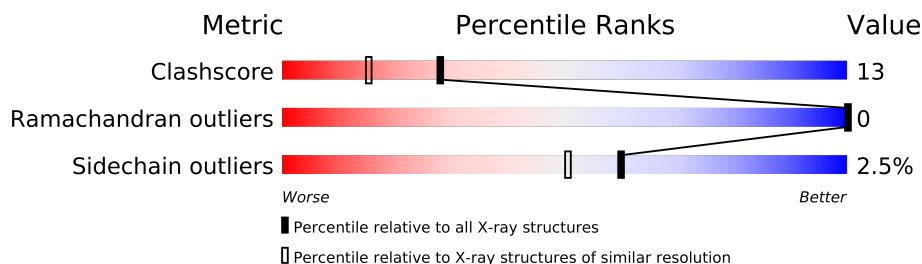
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	21963
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 1.90 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	79885	4465 (1.90-1.90)
Ramachandran outliers	78287	4413 (1.90-1.90)
Sidechain outliers	78261	4414 (1.90-1.90)

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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	C	14	
1	E	14	
2	D	14	
2	F	14	
3	A	190	
3	B	190	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 6447 atoms, of which 1830 are hydrogens 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 DNA chain called DNA (5'-D(\*GP\*CP\*TP\*AP\*TP\*AP\*AP\*AP\*GP\*GP\*GP\*CP\*A)-3').

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	C	14	Total	C	H	N	O	P	0	0	0
			318	137	31	60	77	13			
1	E	14	Total	C	H	N	O	P	0	0	0
			321	138	32	60	78	13			

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

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	D	14	Total	C	H	N	O	P	0	0	0
			305	136	26	44	86	13			
2	F	14	Total	C	H	N	O	P	0	0	0
			305	136	26	44	86	13			

- Molecule 3 is a protein called TATA BINDING PROTEIN.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	A	187	Total	C	H	N	O	S	0	0	0
			1806	960	333	250	255	8			
3	B	188	Total	C	H	N	O	S	0	1	0
			1831	970	342	255	256	8			

- Molecule 4 is water.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	172	Total	H	O	0	0
			516	344	172		
4	B	185	Total	H	O	0	0
			553	368	185		
4	C	38	Total	H	O	0	0
			114	76	38		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	D	39	Total	H	O	0	0
			117	78	39		
4	E	37	Total	H	O	0	0
			111	74	37		
4	F	50	Total	H	O	0	0
			150	100	50		

### 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.

Note EDS was not executed.

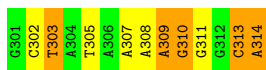
- Molecule 1: DNA (5'-D(\*GP\*CP\*TP\*AP\*TP\*AP\*AP\*AP\*AP\*GP\*GP\*GP\*CP\*A)-3')

Chain C: 



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

Chain E: 



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

Chain D: 



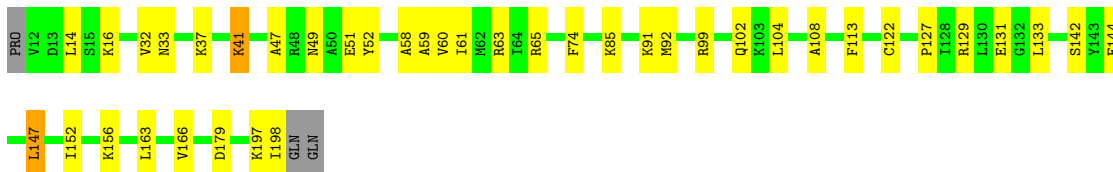
- Molecule 2: DNA (5'-D(\*TP\*GP\*CP\*CP\*CP\*TP\*TP\*TP\*TP\*AP\*TP\*AP\*GP\*C)-3')

Chain F: 



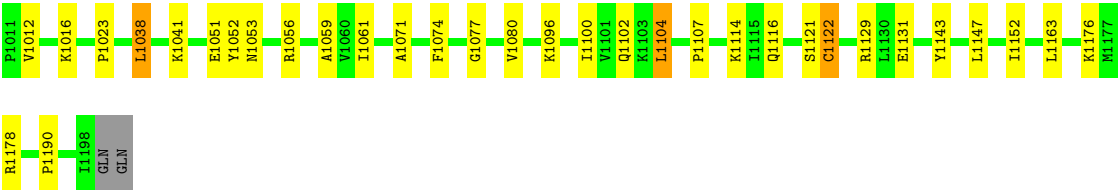
- Molecule 3: TATA BINDING PROTEIN

Chain A: 



- Molecule 3: TATA BINDING PROTEIN

Chain B: 



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	42.00Å 147.00Å 57.00Å 90.00° 90.50° 90.00°	Depositor
Resolution (Å)	6.00 – 1.90	Depositor
% Data completeness (in resolution range)	94.4 (6.00-1.90)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	5.20	Depositor
Refinement program	X-PLOR	Depositor
R, $R_{free}$	0.194 , 0.264	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	6447	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	20.0	wwPDB-VP

## 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	C	1.37	0/324	1.68	7/499 (1.4%)
1	E	1.43	0/326	1.82	13/502 (2.6%)
2	D	1.46	2/310 (0.6%)	2.02	15/476 (3.2%)
2	F	1.53	5/310 (1.6%)	2.42	22/476 (4.6%)
3	A	0.65	1/1503 (0.1%)	0.77	0/2024
3	B	0.68	1/1525 (0.1%)	0.79	0/2052
All	All	0.96	9/4298 (0.2%)	1.31	57/6029 (0.9%)

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
2	F	0	1
All	All	0	3

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	1122	CYS	CB-SG	-8.31	1.68	1.82
2	D	215	DT	C5-C7	6.84	1.54	1.50
2	F	322	DT	C5-C7	6.69	1.54	1.50
2	F	315	DT	C5-C7	5.93	1.53	1.50
2	D	221	DT	C5-C7	5.75	1.53	1.50

The worst 5 of 57 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	317	DC	O4'-C4'-C3'	-13.58	97.85	106.00
2	D	227	DG	O4'-C1'-N9	13.48	117.44	108.00

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	320	DT	C6-C5-C7	-10.79	116.43	122.90
2	F	317	DC	O4'-C1'-N1	10.71	115.49	108.00
2	F	316	DG	O4'-C1'-N9	8.91	114.24	108.00

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	207	DA	Sidechain
2	D	227	DG	Sidechain
2	F	318	DC	Sidechain

## 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	C	287	31	123	6	0
1	E	289	32	126	5	3
2	D	279	26	136	8	0
2	F	279	26	136	15	0
3	A	1473	333	1216	36	2
3	B	1489	342	1234	31	0
4	A	172	344	0	13	2
4	B	185	368	0	14	1
4	C	38	76	0	3	0
4	D	39	78	0	1	0
4	E	37	74	0	2	0
4	F	50	100	0	1	0
All	All	4617	1830	2971	95	4

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 13.

The worst 5 of 95 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
3:A:156:LYS:HA	3:A:156:LYS:HE2	1.43	0.97
3:B:1041:LYS:HD3	3:B:1052:TYR:HE2	1.33	0.89
2:F:317:DC:H2'	2:F:318:DC:C6	2.09	0.87
4:C:924:HOH:O	3:A:58:ALA:HB3	1.75	0.87
3:A:198:ILE:HD11	3:B:1190:PRO:HG2	1.56	0.86

All (4) 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	Distance(Å)	Clash(Å)
4:A:693:HOH:O	4:B:613:HOH:O[2_647]	2.03	0.17
1:E:308:DA:OP1	3:A:49:ASN:ND2[2_657]	2.09	0.11
1:E:311:DG:N7	3:A:65:ARG:NH2[2_657]	2.13	0.07
1:E:309:DA:OP2	4:A:824:HOH:O[2_657]	2.14	0.06

## 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	
3	A	185/190 (97%)	177 (96%)	8 (4%)	0	100	100
3	B	187/190 (98%)	179 (96%)	8 (4%)	0	100	100
All	All	372/380 (98%)	356 (96%)	16 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 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	
3	A	159/163 (98%)	154 (97%)	5 (3%)	52	41

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	B	162/163 (99%)	159 (98%)	3 (2%)	69	63
All	All	321/326 (98%)	313 (98%)	8 (2%)	60	50

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	147	LEU
3	B	1104	LEU
3	B	1038	LEU
3	A	133	LEU
3	A	179	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
3	A	49	ASN
3	B	1053	ASN
3	B	1137	HIS

### 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

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

### 6.3 Carbohydrates

EDS was not executed - this section will therefore be empty.

### 6.4 Ligands

EDS was not executed - this section will therefore be empty.

### 6.5 Other polymers

EDS was not executed - this section will therefore be empty.