



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 26, 2021 – 09:01 AM GMT

PDB ID : 1QN9  
Title : Crystal structure of the C(-29) Adenovirus major late promoter TATA box variant bound to wild-type TBP (Arabidopsis thaliana TBP isoform 2). TATA element recognition by the TATA box-binding protein has been conserved throughout evolution.  
Authors : Patikoglou, G.A.; Kim, J.L.; Sun, L.; Yang, S.-H.; Kodadek, T.; Burley, S.K.  
Deposited on : 1999-10-14  
Resolution : 1.90 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

<https://www.wwpdb.org/validation/2017/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  
Xtrriage (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

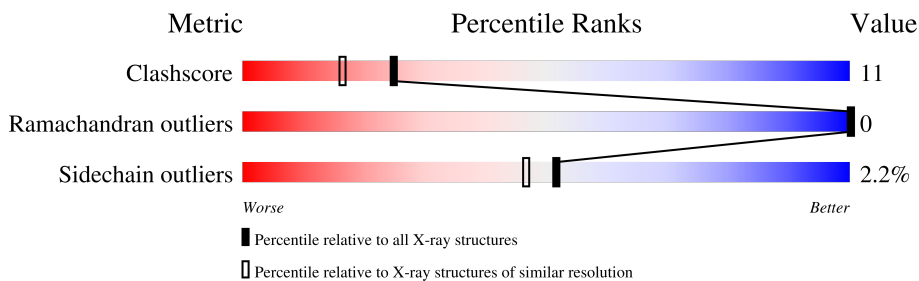
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

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	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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  $\leq 5\%$

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	200	71% (green), 20% (yellow), 9% (orange), 0% (red), 0% (grey)
1	B	200	69% (green), 24% (yellow), 7% (orange), 0% (red), 0% (grey)
2	C	14	21% (green), 57% (yellow), 21% (orange), 0% (red), 0% (grey)
2	E	14	50% (green), 36% (yellow), 14% (orange), 0% (red), 0% (grey)
3	D	14	21% (green), 50% (yellow), 29% (orange), 0% (red), 0% (grey)
3	F	14	14% (green), 64% (yellow), 21% (orange), 0% (red), 0% (grey)

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 4387 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 INITIATION FACTOR TFIID-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	183	1443	941	246	248	8	0	0	0
1	B	186	1464	953	249	254	8	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	C	14	286	136	61	76	13	0	0	0
2	E	14	288	137	61	77	13	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	D	14	278	135	44	86	13	0	0	0
3	F	14	278	135	44	86	13	0	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	110	Total	O	0	0
			110	110		
4	B	127	Total	O	0	0
			127	127		
4	C	28	Total	O	0	0
			28	28		

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>	<b>ZeroOcc</b>	<b>AltConf</b>
4	D	26	Total O 26 26	0	0
4	E	28	Total O 28 28	0	0
4	F	31	Total O 31 31	0	0

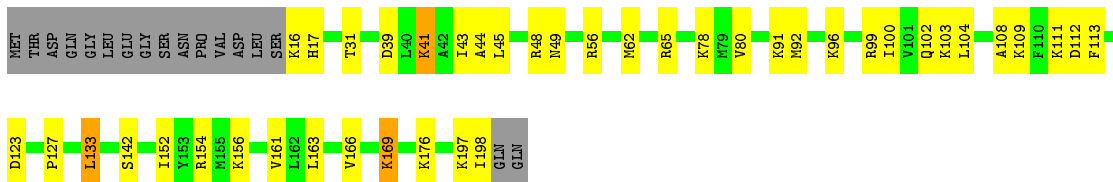
### 3 Residue-property plots [i](#)

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.

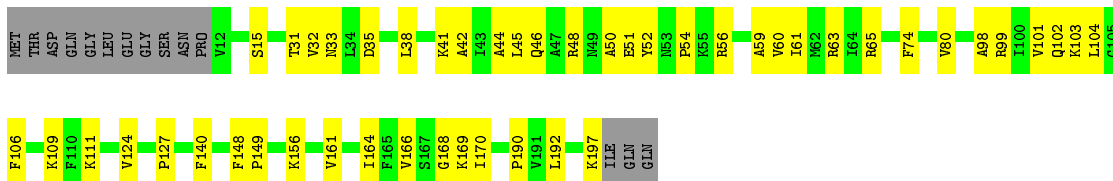
- Molecule 1: TRANSCRIPTION INITIATION FACTOR TFIIID-1

Chain A:  71% 20% 9%



- Molecule 1: TRANSCRIPTION INITIATION FACTOR TFIIID-1

Chain B:  69% 24% 7%



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

Chain C:  21% 57% 21%



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

Chain E:  50% 36% 14%



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

Chain D:  21% 50% 29%

T215	G216	C217	C218	C219	T220	T221	T222	T223	G224	T225	A226	G227	C228
------	------	------	------	------	------	------	------	------	------	------	------	------	------

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

Chain F: 14% 64% 21%

T215	G216	C217	C218	C219	T220	T221	T222	T223	G224	T225	A226	G227	C228
------	------	------	------	------	------	------	------	------	------	------	------	------	------

## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	41.80 Å 146.70 Å 57.40 Å 90.00° 90.50° 90.00°	Depositor
Resolution (Å)	6.00 – 1.90	Depositor
% Data completeness (in resolution range)	86.8 (6.00-1.90)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.04	Depositor
Refinement program	X-PLOR	Depositor
R, $R_{free}$	0.209 , 0.275	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	4387	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	33.0	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	A	0.57	0/1473	0.74	1/1983 (0.1%)
1	B	0.56	0/1494	0.75	0/2012
2	C	1.29	1/323 (0.3%)	1.72	9/497 (1.8%)
2	E	1.26	0/325	1.63	6/500 (1.2%)
3	D	1.34	2/309 (0.6%)	2.14	22/475 (4.6%)
3	F	1.32	3/309 (1.0%)	2.03	18/475 (3.8%)
All	All	0.85	6/4233 (0.1%)	1.24	56/5942 (0.9%)

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	215	DT	C5-C7	6.57	1.53	1.50
3	D	216	DG	P-O5'	5.86	1.65	1.59
3	F	221	DT	C5-C7	5.51	1.53	1.50
3	F	222	DT	C5-C7	5.42	1.53	1.50
3	F	215	DT	C5-C7	5.30	1.53	1.50

The worst 5 of 56 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	227	DG	O4'-C1'-N9	8.91	114.24	108.00
3	D	221	DT	C6-C5-C7	-8.57	117.75	122.90
3	D	216	DG	O4'-C1'-N9	8.45	113.91	108.00
2	C	201	DG	O4'-C1'-N9	8.14	113.70	108.00
3	D	220	DT	C6-C5-C7	-8.07	118.06	122.90

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	A	1443	0	1518	31	0
1	B	1464	0	1536	38	1
2	C	286	0	153	7	1
2	E	288	0	157	5	0
3	D	278	0	158	4	0
3	F	278	0	158	6	0
4	A	110	0	0	7	0
4	B	127	0	0	5	0
4	C	28	0	0	1	0
4	D	26	0	0	0	0
4	E	28	0	0	0	0
4	F	31	0	0	1	0
All	All	4387	0	3680	83	1

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 11.

The worst 5 of 83 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:F:215:DT:H2''	3:F:216:DG:H2'	1.60	0.82
1:A:49:ASN:O	1:A:62:MET:HG3	1.87	0.74
1:A:127:PRO:HB2	1:A:197:LYS:HB3	1.70	0.72
2:C:203:DT:H5'	4:C:2005:HOH:O	1.89	0.71
1:A:198:ILE:HD11	1:B:190:PRO:HG2	1.74	0.70

All (1) 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	Interatomic distance (Å)	Clash overlap (Å)
1:B:65:ARG:NH1	2:C:211:DG:O6[2_656]	2.18	0.02

## 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	181/200 (90%)	175 (97%)	6 (3%)	0	100	100
1	B	184/200 (92%)	179 (97%)	5 (3%)	0	100	100
All	All	365/400 (91%)	354 (97%)	11 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	155/171 (91%)	150 (97%)	5 (3%)	39	30
1	B	158/171 (92%)	156 (99%)	2 (1%)	69	68
All	All	313/342 (92%)	306 (98%)	7 (2%)	52	47

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	133	LEU
1	B	38	LEU
1	A	169	LYS
1	A	48	ARG
1	B	15	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	137	HIS

### 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.