



# wwPDB X-ray Structure Validation Summary Report

May 13, 2020 – 02:37 pm BST

PDB ID : 1CEZ  
Title : CRYSTAL STRUCTURE OF A T7 RNA POLYMERASE-T7 PROMOTER  
COMPLEX  
Authors : Cheetham, G.M.T.; Jeruzalmi, D.; Steitz, T.A.  
Deposited on : 1999-03-11  
Resolution : 2.40 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) : 2.11

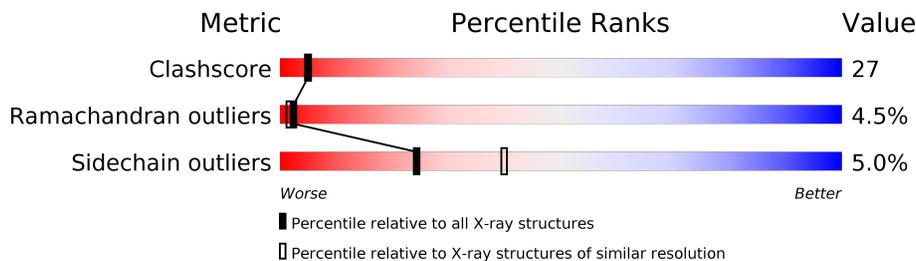
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)

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 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	T	17	24% (green), 65% (yellow), 12% (orange)
2	N	15	27% (green), 60% (yellow), 13% (orange)
3	A	883	59% (green), 33% (yellow), 6% (orange), 2% (grey)

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 7766 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(P\*TP\*AP\*TP\*AP\*GP\*TP\*GP\*AP\*GP\*TP\*CP\*GP\*TP\*AP\*TP\*TP\*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	T	17	352	169	62	104	17	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	N	15	304	146	55	88	15	0	0	0

- Molecule 3 is a protein called PROTEIN (BACTERIOPHAGE T7 RNA POLYMERASE).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	A	862	6639	4228	1164	1211	36	0	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	T	14	Total 14	O 14	0	0
4	N	18	Total 18	O 18	0	0
4	A	439	Total 439	O 439	0	0

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

These plots are drawn for all protein, RNA and DNA 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(P\*TP\*AP\*TP\*AP\*GP\*TP\*GP\*AP\*GP\*TP\*CP\*GP\*TP\*AP\*TP\*T P\*A)-3')

Chain T: 



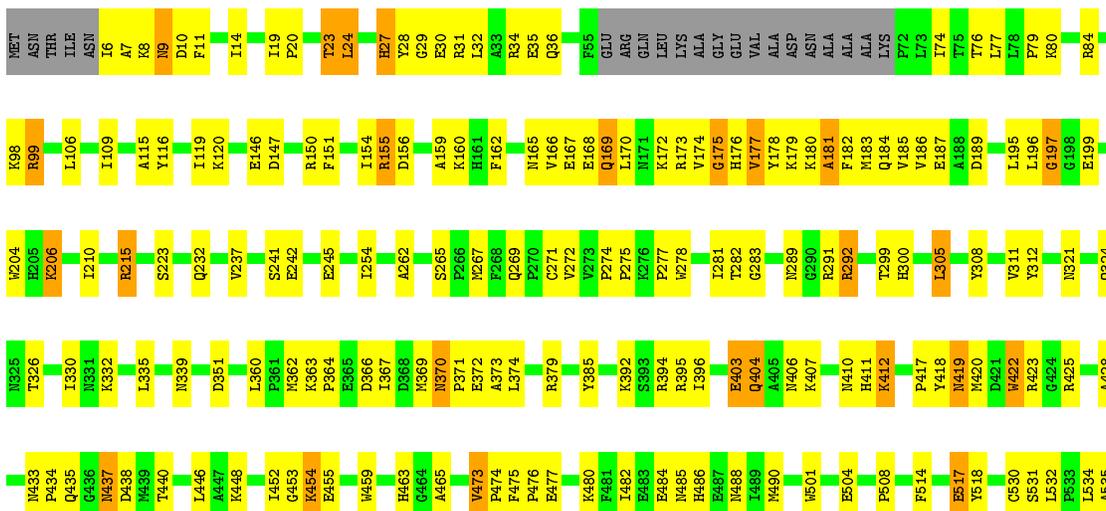
- Molecule 2: DNA (5'-D(P\*TP\*AP\*AP\*TP\*AP\*CP\*GP\*AP\*CP\*TP\*CP\*AP\*CP\*TP\*A)-3')

Chain N: 



- Molecule 3: PROTEIN (BACTERIOPHAGE T7 RNA POLYMERASE)

Chain A: 



F536	D537	G538	S539	C540	Q544	H545	M549	L550	R551	D552	E553	V554	R557	A558	V559	M560	L561	L562	D569	L570	I573	V574	K577	T590	D591	N592	E593	V594	V595	T596	V597	T598	D599	E600	N601	T602	G603	E604	I605	S606	E607	K608	V609	K610	L611	G612	T613	K614	A615	L616	Q619
V625	T626	R627	T630	K631	R632	S633	M634	M635	K642	E643	F644	G645	F646	R647	Q648	Q649	V650	L651	E652	D653	T654	I655	Q656	P657	A658	I659	D660	K663	G664	F667	T668	Q669	P670	N671	A674	G675	T688	A695	M696	K700	S701	L705	L706	V710	K711	D712	K713	K714			
T715	G716	E717	I718	L719	R720	K721	R722	C723	A724	V725	H726	T729	P730	F733	F734	V735	W736	Q737	E738	I743	M750	G753	Q754	F755	R756	L757	I761	T762	T763	N764	K765	E768	E775	S789	H790	K793	T794	V795	W796	W797	A798	H799	E800	K801	I804	E805	S806				
H811	D812	S813	F814	G815	T816	I817	P818	K826	A827	V828	R829	E830	T831	M832	V833	E837	L842	Y846	D847	Q848	F849	A850	D851	L853	H854	E855	S856	Q857	L858	M861	P862	A863	G868	N869	L870	M871	R872	D873	D874	A881	F882	A883									

## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	220.10Å 73.30Å 80.90Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 2.40	Depositor
% Data completeness (in resolution range)	81.3 (40.00-2.40)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.08	Depositor
Refinement program	CNS 0.5	Depositor
R, $R_{free}$	0.224 , 0.270	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	7766	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	53.0	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	T	0.60	0/394	0.76	0/607
2	N	1.29	3/340 (0.9%)	1.05	0/521
3	A	0.49	1/6791 (0.0%)	0.69	3/9207 (0.0%)
All	All	0.56	4/7525 (0.1%)	0.72	3/10335 (0.0%)

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	T	0	2
2	N	0	1
All	All	0	3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	805	GLU	CD-OE2	7.42	1.33	1.25
2	N	114	DT	C4-C5	6.42	1.50	1.45
2	N	114	DT	N1-C2	5.57	1.42	1.38
2	N	113	DC	O3'-P	5.44	1.67	1.61

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	559	VAL	N-CA-C	-5.55	96.01	111.00
3	A	181	ALA	N-CA-C	-5.50	96.16	111.00
3	A	540	CYS	N-CA-C	-5.26	96.80	111.00

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	N	106	DC	Sidechain
1	T	5	DG	Sidechain
1	T	6	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	T	352	0	195	22	0
2	N	304	0	170	20	0
3	A	6639	0	6470	350	0
4	A	439	0	0	29	0
4	N	18	0	0	0	0
4	T	14	0	0	0	0
All	All	7766	0	6835	384	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 27.

The worst 5 of 384 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:281:ILE:HG23	3:A:282:THR:HG23	1.30	1.13
2:N:113:DC:H3'	2:N:115:DA:OP1	1.49	1.11
3:A:710:VAL:HG11	3:A:720:ARG:H	1.10	1.08
3:A:829:ARG:HG3	3:A:829:ARG:HH11	1.16	1.07
3:A:763:THR:HG22	3:A:765:LYS:H	1.15	1.05

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	858/883 (97%)	754 (88%)	65 (8%)	39 (4%)	2 2

5 of 39 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	168	GLU
3	A	178	TYR
3	A	180	LYS
3	A	539	SER
3	A	594	VAL

### 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	666/729 (91%)	633 (95%)	33 (5%)	24 40

5 of 33 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	412	LYS
3	A	437	ASN
3	A	775	GLU
3	A	419	ASN
3	A	422	TRP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 34 such sidechains are listed below:

Mol	Chain	Res	Type
3	A	410	ASN
3	A	437	ASN
3	A	786	GLN
3	A	419	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
3	A	171	ASN

### 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 carbohydrates 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.