



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 10:08 PM GMT

PDB ID : 1S76
Title : T7 RNA polymerase alpha beta methylene ATP elongation complex
Authors : Yin, Y.W.; Steitz, T.A.
Deposited on : 2004-01-29
Resolution : 2.88 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
<http://wwpdb.org/validation/2016/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
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

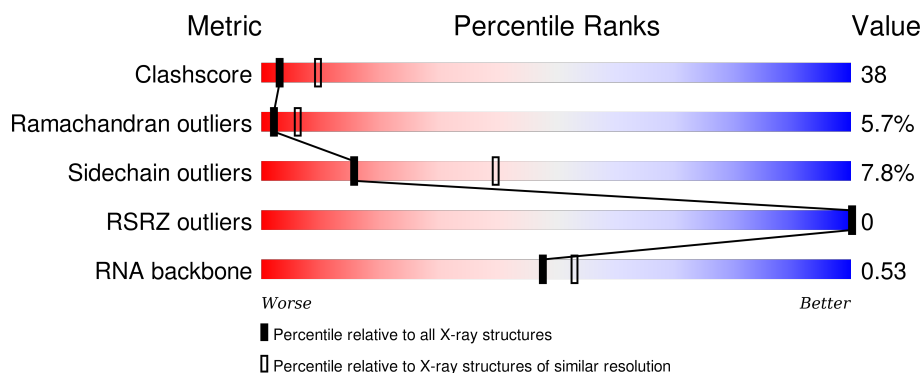
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.88 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	102246	2202 (2.90-2.86)
Ramachandran outliers	100387	2149 (2.90-2.86)
Sidechain outliers	100360	2152 (2.90-2.86)
RSRZ outliers	91569	1950 (2.90-2.86)
RNA backbone	2183	1033 (3.26-2.50)

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 ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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 $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	T	21	
2	N	17	
3	R	9	
4	D	883	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard

residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
6	APC	R	901	-	-	X	-

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 7561 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*GP*CP*CP*GP*TP*GP*CP*GP*CP*AP*TP*TP*CP*GP*CP*CP*GP*TP*GP*TP*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	T	21	Total	C	N	O	P	0	0	0
			429	203	73	132	21			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	N	17	Total	C	N	O	P	0	0	0
			348	165	60	106	17			

- Molecule 3 is a RNA chain called RNA (5'-R(P*AP*CP*AP*CP*GP*GP*CP*GP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	R	9	Total	C	N	O	P	0	0	0
			196	87	39	61	9			

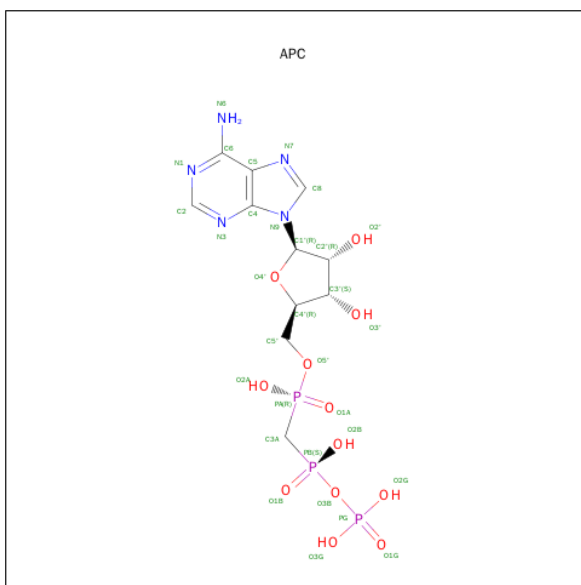
- Molecule 4 is a protein called DNA-directed RNA polymerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	829	Total	C	N	O	S	0	0	0
			6555	4178	1143	1198	36			

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	D	2	Total	Mg	0	0
			2	2		

- Molecule 6 is DIPHOSPHOMETHYLPHOSPHONIC ACID ADENOSYL ESTER (three-letter code: APC) (formula: C₁₁H₁₈N₅O₁₂P₃).

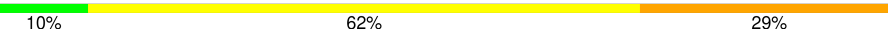


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	R	1	Total	C	N	O	P	0	0
			31	11	5	12	3		

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.

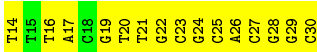
- Molecule 1: DNA (5'-D(P*GP*CP*CP*GP*TP*GP*CP*GP*CP*AP*TP*TP*CP*GP*CP*C P*GP*TP*GP*TP*T)-3')

Chain T: 



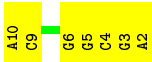
- Molecule 2: DNA (5'-D(P*TP*TP*TP*AP*CP*GP*TP*TP*GP*CP*GP*CP*AP*CP*GP*G P*C)-3')

Chain N: 



- Molecule 3: RNA (5'-R(P*AP*CP*AP*CP*GP*GP*CP*GP*A)-3')

Chain R: 

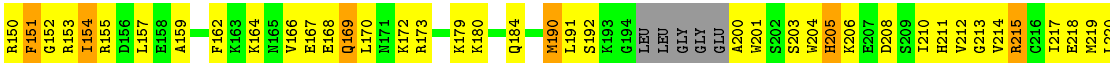


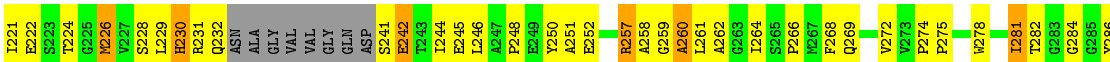
- Molecule 4: DNA-directed RNA polymerase

Chain D: 









H811	B746	A674	SER	Q538	K450	RET	K287
D812	L747	G675	GLU	S539	P451	ASN	A286
S813	N748	Y676	LYS	Q540	L452	PRO	
F814	L749	M677	VAL		G453	GLU	R292
G815	M750		LYS	I543		ALA	
T816	F751	L660	LEU	I544	G456		L296
I817	L752	I681		H545		T375	V297
P818		W682	G612	F546		A376	R296
	F755	E683	T613	S547		K377	
L824	R756	S684	K614	L550		K378	K303
F825	L757	V685	A615	R551			
K826	Q758	S686	L616	D552		A361	R307
K827	P759		V625	E553			Y308
V828	T760	V689	T626	V554		R386	E309
R829	I761	A692	R627			C467	D310
E830	N762		S628	R557		K387	D310
T831	T763	A695	V629	A558		D388	V311
	N764		K631	V574			V312
A843	K765	W698	R632	V574		R395	V313
D844	T766		S633	A576			
F845	S767	S701	V634	K576		L398	Y317
Y846	E768	K704	W635	K577		E399	
D847	I769	L705	T636	V578		F475	I320
Q848	D770	L706	L637	N579		P476	M401
F849	A771	A707	A638	E580		E477	L402
A850		E709	V639	L582		L478	Q324
		D712	S641	Q583		T566	R325
D851				A584		C464	
Q852	I778	T715	G645	D585		A405	
L853	A779	G716	G646	I587			W328
H854	P780		F646	Q588		A409	I330
E855	N781	R720	F647	T590		N410	
S856	F782	K721	R647			H411	
Q857	W783	L722	Q648	E593		F416	L336
D859	H784	R722	L651	VAL		P417	A336
K860	S785	R720	Q649	THR		F432	V337
M861	Q786	K721	V650	VAL		M433	
P862	D787	R722	L651	THR		P434	R357
A863	G788	G723	E652	THR		Q435	E359
L864	S789	A724	D653	ASP		L360	L360
	H790	V725	T654	GLU		D438	P361
L870	L791	H726	Q656	ASN		M439	K362
	R792	T729	P657	LYS			LYS
R873	K793	F730	A658	PRO			PRO
	T794	D731	I659	GLU			GLU
F880	V795	G732	D660	ASP			ASP
A881	V796	F733	S661	ILE			ILE
F882	A797	P734	Q662				ASP
A883	H799	V735	K663				
	E800	W736	E664				
	R801	Q737	L665				
	G802	E738	W666				
	G803	Y739	F667				
	I804	K740	T668				
	E805	K741	Q669				
	S806	F742	P670				
	L809	I743	H671				
	I810	Q744	Q672				
		T745	A673				

4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	138.72Å 143.32Å 146.65Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.88 29.53 – 2.88	Depositor EDS
% Data completeness (in resolution range)	99.8 (30.00-2.88) 91.2 (29.53-2.88)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	14.95 (at 2.90Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.248 , 0.290 0.224 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å ²)	37.1	Xtriage
Anisotropy	0.114	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.06 , 458.1	EDS
Estimated twinning fraction	0.228 for -k,-h,-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.28$, $\langle L^2 \rangle = 0.12$	Xtriage
Outliers	0 of 33374 reflections	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	7561	wwPDB-VP
Average B, all atoms (Å ²)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.94% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: APC, MG

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	0.57	1/478 (0.2%)	0.85	0/734
2	N	0.50	1/388 (0.3%)	0.69	0/595
3	R	0.72	1/219 (0.5%)	0.83	0/338
4	D	0.36	2/6706 (0.0%)	0.58	1/9068 (0.0%)
All	All	0.40	5/7791 (0.1%)	0.62	1/10735 (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	5

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	N	14	DT	OP3-P	-7.30	1.52	1.61
3	R	10	A	OP3-P	-7.28	1.52	1.61
1	T	130	DG	OP3-P	-7.24	1.52	1.61
4	D	292	ARG	C-N	-6.38	1.22	1.34
4	D	251	ALA	C-N	5.80	1.47	1.34

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	D	828	VAL	N-CA-C	-5.01	97.48	111.00

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	T	116	DC	Sidechain
1	T	117	DG	Sidechain
1	T	118	DC	Sidechain
1	T	119	DT	Sidechain
1	T	120	DT	Sidechain

5.2 Too-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 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	429	0	238	28	0
2	N	348	0	193	26	0
3	R	196	0	100	15	0
4	D	6555	0	6518	509	0
5	D	2	0	0	0	0
6	R	31	0	14	11	0
All	All	7561	0	7063	561	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 38.

The worst 5 of 561 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:R:2:A:O3'	6:R:901:APC:H5'1	1.34	1.28
4:D:759:PRO:HG2	4:D:764:ASN:ND2	1.61	1.15
4:D:724:ALA:HB2	4:D:738:GLU:HB3	1.36	1.06
4:D:763:THR:O	4:D:764:ASN:CG	1.96	1.04
3:R:2:A:O3'	6:R:901:APC:C5'	2.09	1.01

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
4	D	821/883 (93%)	663 (81%)	111 (14%)	47 (6%)	2 6

5 of 47 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	D	15	GLU
4	D	149	ALA
4	D	226	MET
4	D	360	LEU
4	D	422	TRP

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
4	D	683/729 (94%)	630 (92%)	53 (8%)	16 40

5 of 53 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	D	401	MET
4	D	472	LYS
4	D	793	LYS
4	D	402	LEU
4	D	423	ARG

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

sidechains are listed below:

Mol	Chain	Res	Type
4	D	485	ASN
4	D	648	GLN
4	D	852	GLN
4	D	579	ASN
4	D	184	GLN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	R	8/9 (88%)	0	0

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

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 ⓘ

Of 3 ligands modelled in this entry, 2 are monoatomic - leaving 1 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
6	APC	R	901	5	25,33,33	1.53	3 (12%)	30,52,52	2.39	9 (30%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	APC	R	901	5	-	0/15/38/38	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	R	901	APC	PA-O2A	-3.11	1.48	1.56
6	R	901	APC	PB-O2B	-3.09	1.48	1.56
6	R	901	APC	PB-O3B	3.88	1.62	1.58

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	R	901	APC	O1B-PB-C3A	-5.30	95.69	109.02
6	R	901	APC	O5'-PA-O1A	-2.70	106.81	113.98
6	R	901	APC	O2B-PB-C3A	2.13	116.16	106.88
6	R	901	APC	C2'-C1'-N9	2.36	117.89	114.29
6	R	901	APC	O5'-PA-C3A	2.64	111.81	104.42

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	R	901	APC	11	0

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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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, 95th 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 ²)	Q<0.9
1	T	21/21 (100%)	-0.65	0 100 100	43, 97, 136, 138	0
2	N	17/17 (100%)	-0.58	0 100 100	121, 128, 138, 142	0
3	R	9/9 (100%)	-0.21	0 100 100	72, 96, 103, 104	0
4	D	829/883 (93%)	-1.29	0 100 100	0, 26, 74, 97	0
All	All	876/930 (94%)	-1.25	0 100 100	0, 29, 93, 142	0

There are no RSRZ outliers to report.

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å ²)	Q<0.9
6	APC	R	901	31/31	0.98	0.14	1.38	13,23,44,67	0
5	MG	D	902	1/1	0.97	0.11	-	23,23,23,23	0
5	MG	D	903	1/1	0.96	0.18	-	31,31,31,31	0

6.5 Other polymers [i](#)

There are no such residues in this entry.