



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 08:57 PM GMT

PDB ID : 1MSW
Title : Structural basis for the transition from initiation to elongation transcription
in T7 RNA polymerase
Authors : Yin, Y.W.; Steitz, T.A.
Deposited on : 2002-09-19
Resolution : 2.10 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) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

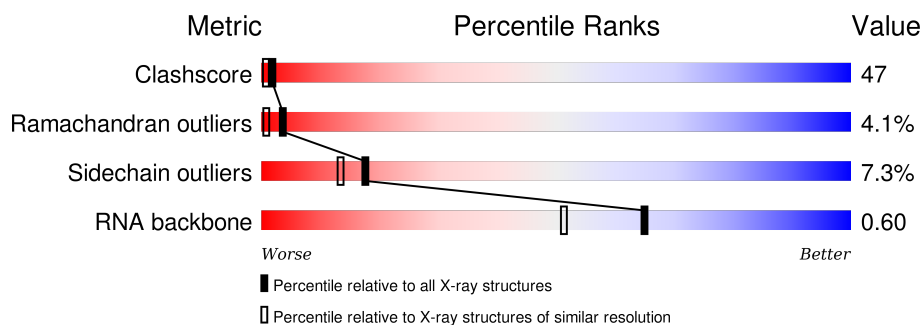
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 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	4460 (2.10-2.10)
Ramachandran outliers	100387	4413 (2.10-2.10)
Sidechain outliers	100360	4414 (2.10-2.10)
RNA backbone	2183	1118 (2.80-1.40)

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 ≥ 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	T	20	<div> <div>5%</div> <div>60%</div> <div>30%</div> <div>5%</div> </div>
2	N	17	<div> <div>18%</div> <div>82%</div> </div>
3	R	10	<div> <div>20%</div> <div>80%</div> </div>
4	D	883	<div> <div>44%</div> <div>46%</div> <div>7%</div> <div>.</div> </div>

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 7956 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 Template DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	T	20	Total	C	N	O	P	0	0	0
			405	193	71	122	19			

- Molecule 2 is a DNA chain called Non-Template DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	N	17	Total	C	N	O	P	0	0	0
			344	165	60	103	16			

- Molecule 3 is a RNA chain called RNA message.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	R	10	Total	C	N	O	P	0	0	0
			215	97	44	65	9			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	863	Total	C	N	O	S	0	0	0
			6802	4333	1182	1251	36			

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	D	166	Total	O	0	0
			166	166		
5	N	2	Total	O	0	0
			2	2		
5	R	9	Total	O	0	0
			9	9		
5	T	13	Total	O	0	0
			13	13		

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 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.

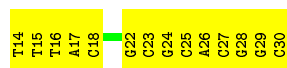
- Molecule 1: Template DNA

Chain T: 



- Molecule 2: Non-Template DNA

Chain N: 



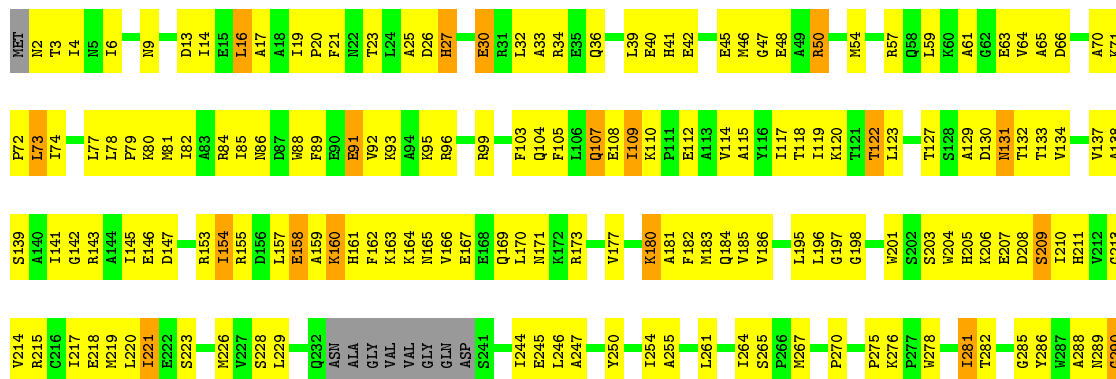
- Molecule 3: RNA message

Chain R: 



- Molecule 4: DNA-directed RNA polymerase

Chain D: 



S856	Q857	L858	D859	G868	N869	L870	N871	L872	R873	D874	I875	L876	S877	S878	F882	A883	A779	P780	N781	F782	V783	H784	D787	G788	S789	H790	L791	T794	V795	V796	W797	K801	Y802	G803	I804	E805	S806	P807	I810	H811	D812	S813	F814	G815	I816	I817	P818	A819	D820	K826	A827	V828	R829	E830	V833	D834	E837	P845	Y846	D847	Q848	F849	F849	A850	D851	Q852	L853	H854	E855
R291	R292	P293	L294	L295	L296	R300	S301	K302	K303	A304	Y308	V311	Y312	K318	K319	I320	I321	I322	A323	Q324	I325	I330	K333	V334	L335	A336	V337	A338	V339	K407	I341	W344	K345	C347	P348	V349	I352	P353	A354	I355	E356	E357	E358	E359	L360	P361	K362	K363	P364	G1U																			
ASP	ILE	ASP	MET	ASN	PRO	GLU	ALA	LEU	T375	K376	K377	K378	R379	A380	A381	A382	A383	V384	K385	R386	R391	R394	R395	I396	S397	L398	E399	F400	K401	L402	E403	Q404	A405	N406	K407	F408	A409	N410	W415	F416	P417	Y418	N419	M420	D421	K422	G423	G424	A428	V429	M433	P434	Q435																
M439	T440	K441	G442	L443	L444	T445	L446	P451	L452	L550	G453	K454	Y458	K461	L462	H463	G464	A468	D471	K472	V473	P474	E477	K478	I479	K480	F481	N488	L489	K490	A491	S495	N499	T500	Q505	P508	L512	A513	F514	C515	F516	E517	E517	L532	P533	L534	D537																						
Q538	S539	C540	I543	G544	H545	F546	M549	L550	R551	G553	D552	E553	R557	N560	L561	L562	T566	Q568	D569	I570	V574	A575	K576	K577	V578	N579	L582	Q583	A584	D585	A586	I587	N588	G589	T590	D591	E593	V594	V595	T602	G603	E604	I605	S606	E607	K608	V609	K610	L611	G612																			
H613	H614	A615	L616	Q619	H620	L621	A622	V623	K631	H632	S633	H634	H635	H647	Q648	Q649	V650	L651	E652	D653	T654	H655	Q656	P657	A658	L659	D660	S661	Q662	K663	Q664	L665	H666	H667	H668	Q669	Q672	V676	K679	H682	H683	S684	H685	S686	V687	T688	V689	V690	A691	A695	H696																		
N697	W698	L699	K700	L706	A707	A708	E709	V710	K711	G716	R720	K721	R722	C723	A724	V725	H726	W727	V728	T729	P730	F733	P734	V735	W736	K741	F742	I743	Q744	T745	H746	L747	N748	L749	H750	F751	L752	Q753	Q754	F755	R756	L757	Q758	P759	T763	N764	K765	D766	S767	Q774	I778																		
A779	P780	N781	F782	V783	H784	D787	G788	S789	H790	L791	T794	V795	V796	W797	K801	Y802	G803	I804	E805	S806	P807	I810	H811	D812	S813	F814	G815	I816	I817	P818	A819	D820	K826	A827	V828	R829	E830	V833	D834	E837	P845	Y846	D847	Q848	F849	F849	A850	D851	Q852	L853	H854	E855																	

4 Data and refinement statistics

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

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	142.91Å 145.46Å 145.60Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.10	Depositor
% Data completeness (in resolution range)	85.0 (20.00-2.10)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.241 , 0.275	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	7956	wwPDB-VP
Average B, all atoms (Å ²)	75.0	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	0.46	0/452	0.90	2/696 (0.3%)
2	N	0.31	0/384	0.67	0/591
3	R	0.47	0/241	0.80	0/375
4	D	0.48	1/6955 (0.0%)	0.65	0/9407
All	All	0.47	1/8032 (0.0%)	0.67	2/11069 (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	7

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	407	LYS	CD-CE	5.13	1.64	1.51

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	T	119	DT	N1-C1'-C2'	-5.28	102.57	112.60
1	T	119	DT	C5'-C4'-C3'	-5.10	104.92	114.10

There are no chirality outliers.

5 of 7 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

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	T	119	DT	Sidechain
1	T	121	DA	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	405	0	227	49	0
2	N	344	0	194	17	0
3	R	215	0	112	15	0
4	D	6802	0	6770	645	0
5	D	166	0	0	246	0
5	N	2	0	0	0	0
5	R	9	0	0	8	0
5	T	13	0	0	27	0
All	All	7956	0	7303	715	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 47.

The worst 5 of 715 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)
4:D:360:LEU:HB2	5:D:968:HOH:O	1.29	1.29
4:D:104:GLN:HB3	5:D:979:HOH:O	1.23	1.27
4:D:276:LYS:HE2	5:D:1032:HOH:O	1.22	1.26
4:D:217:ILE:O	4:D:221:ILE:HD13	1.24	1.24
4:D:428:ALA:HA	5:D:944:HOH:O	1.39	1.22

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	857/883 (97%)	736 (86%)	86 (10%)	35 (4%)	3 1

5 of 35 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	D	16	LEU
4	D	158	GLU
4	D	302	LYS
4	D	303	LYS
4	D	353	PRO

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	711/729 (98%)	659 (93%)	52 (7%)	17 13

5 of 52 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	D	358	GLU
4	D	443	LEU
4	D	847	ASP
4	D	360	LEU
4	D	378	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such

sidechains are listed below:

Mol	Chain	Res	Type
4	D	568	GLN
4	D	697	ASN
4	D	848	GLN
4	D	649	GLN
4	D	656	GLN

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	R	9/10 (90%)	1 (11%)	0

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
3	R	8	C

There are no RNA pucker outliers to report.

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 [i](#)

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

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

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

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

6.3 Carbohydrates [i](#)

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

6.4 Ligands [i](#)

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

6.5 Other polymers [i](#)

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