



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 08:59 am 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
Xtriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : recalc28949

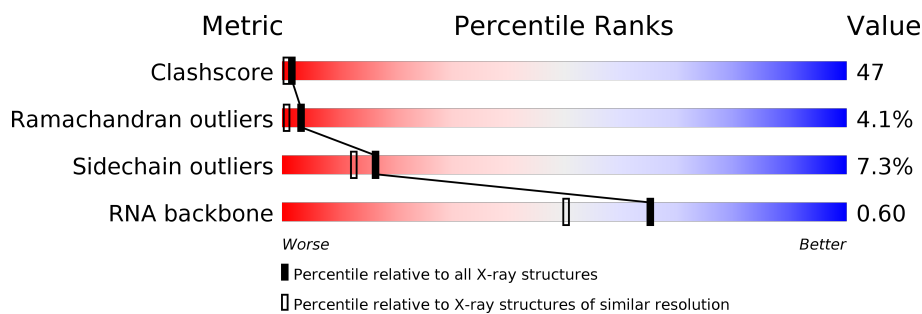
1 Overall quality at a glance

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	112137	4788 (2.10-2.10)
Ramachandran outliers	110173	4740 (2.10-2.10)
Sidechain outliers	110143	4741 (2.10-2.10)
RNA backbone	2435	1063 (2.70-1.50)

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.

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

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

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 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.

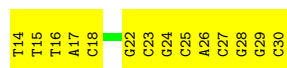
• Molecule 1: Template DNA

Chain T: 



• Molecule 2: Non-Template DNA

Chain N: 



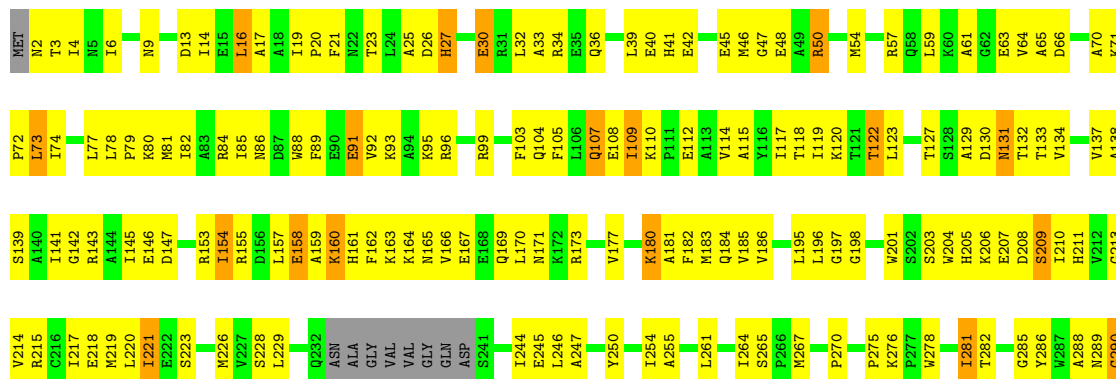
• Molecule 3: RNA message

Chain R: 



• Molecule 4: DNA-directed RNA polymerase

Chain D: 



S856	A779	G697	T613	G538	M439	R291
Q857	P780	W698	K614	S539	T440	ASP
L858	W781	L699	A615	C540	K441	ILE
S859	F782	K700	L616		G442	ASP
	F783			T643	L443	MET
H784			Q619	H544	L444	ASN
D787	A707	L621	W620	H545	T445	PRO
	G788	E709	A622	F546	L446	GLU
S789	V710	Y623				
H790	K711			M649	P451	T375
L791		K631		L550	L452	A376
	G716	R632		R551	G453	W377
T794		S633		D552	K454	K378
V795		W634		E553		R379
W796	R720	M635			Y458	A380
W797	R721			R557		A381
	R722				K461	A382
K801	R723	R647		N560	L462	A383
Y802	A724	Q648		L561	H463	K318
G803	V725	Q649		L562	G464	A319
I804	H726	V650				I320
E805	W727	L651		T566	A468	N321
S806	W728	E652		V567		I322
S806	T729	D653		E568	D471	
F807	P730	T654		D569	K472	R394
		T655		L570	V473	R395
	F733	Q656			P474	I396
H811	F734	P657		V674		S397
D812	V735	A658		A575	E477	L398
S813	W736	T659		K478	R478	E399
F814		D660		X577	I479	V334
G815	K741	S661		V578	K480	L335
T816	P742	G662		N579	F481	A336
I817	I743	K663				V337
P818	Q744	Q664		L582	N488	A438
A819	T745	L665		Q583	L489	A405
D820	R746	M666		A584	M490	N406
	L747	F667		D585	A491	K407
K826	N748	T668		A586		F408
	L749	Q669		I587	S495	A409
V828	W750			N588		N410
R829	F751	Q672		E589	N499	
E830	L752	T670		T590	T500	W415
	G753	Y676		D591		F416
V833	Q754	N592		E592	Q505	P417
D834	F755	E593		V594		Y418
	W756	K679			P508	N419
E837	L757	W682		V595		M420
	Q758	E683			L512	D421
F845	P759	S684		T602	A513	A354
Y846		W685		G603	F514	I355
D847		S686		E604	C515	E356
Q848	N764	V687		I605	F516	R357
F849	K765	T688		S606	E517	E358
A850	D766	V689		E607		A359
D851	S767	V690		K608	L532	L360
Q852		A691		V609	P533	P361
L853	Q774			K610	L534	K363
H854		A695		L611		P434
E855	I778	M696			D537	Q435

4 Data and refinement statistics

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

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	142.91 Å 145.46 Å 145.60 Å 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

5.1 Standard geometry

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

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
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

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%)	16 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

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.