



## wwPDB EM Validation Summary Report ⓘ

Mar 6, 2025 – 03:11 pm GMT

PDB ID : 9Q96  
EMDB ID : EMD-52919  
Title : Cryo-EM Structure of Bacterial RNA polymerase-sigma54 transcription open complex with wild type sigma54, from RPi(-10-1)  
Authors : Gao, F.; Zhang, X.  
Deposited on : 2025-02-26  
Resolution : 4.60 Å(reported)

This is a wwPDB EM 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/EMValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

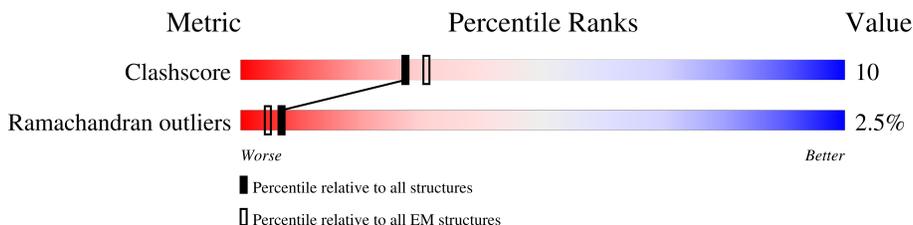
EMDB validation analysis : **FAILED**  
MolProbity : 4.02b-467  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
MapQ : **FAILED**  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.41

# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:  
*ELECTRON MICROSCOPY*

The reported resolution of this entry is 4.60 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)	EM structures (#Entries)
Clashscore	210492	15764
Ramachandran outliers	207382	16835

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	329	84% 9% 6%
1	B	329	67% 5% 29%
2	C	1342	79% 18% .
3	D	1407	77% 15% . . .
4	E	91	74% . . . 18%
5	M	497	52% 9% . 36%
6	N	46	22% 78%
7	T	46	22% 54% 13% 11%

## 2 Entry composition i

There are 7 unique types of molecules in this entry. The entry contains 19633 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 DNA-directed RNA polymerase subunit alpha.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
1	A	309	Total	C	N	O	0	0
			1526	908	309	309		
1	B	235	Total	C	N	O	0	0
			1160	690	235	235		

- Molecule 2 is a protein called DNA-directed RNA polymerase subunit beta.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
2	C	1341	Total	C	N	O	0	0
			6599	3917	1341	1341		

- Molecule 3 is a protein called DNA-directed RNA polymerase subunit beta'.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
3	D	1345	Total	C	N	O	0	0
			6619	3929	1345	1345		

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

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
4	E	75	Total	C	N	O	0	0
			371	221	75	75		

- Molecule 5 is a protein called RNA polymerase sigma-54 factor.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
5	M	316	Total	C	N	O	0	0
			1570	938	316	316		

There are 45 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
M	-19	MET	-	initiating methionine	UNP A0A377VEN9

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
M	-18	GLY	-	expression tag	UNP A0A377VEN9
M	-17	SER	-	expression tag	UNP A0A377VEN9
M	-16	SER	-	expression tag	UNP A0A377VEN9
M	-15	HIS	-	expression tag	UNP A0A377VEN9
M	-14	HIS	-	expression tag	UNP A0A377VEN9
M	-13	HIS	-	expression tag	UNP A0A377VEN9
M	-12	HIS	-	expression tag	UNP A0A377VEN9
M	-11	HIS	-	expression tag	UNP A0A377VEN9
M	-10	HIS	-	expression tag	UNP A0A377VEN9
M	-9	SER	-	expression tag	UNP A0A377VEN9
M	-8	SER	-	expression tag	UNP A0A377VEN9
M	-7	GLY	-	expression tag	UNP A0A377VEN9
M	-6	LEU	-	expression tag	UNP A0A377VEN9
M	-5	VAL	-	expression tag	UNP A0A377VEN9
M	-4	PRO	-	expression tag	UNP A0A377VEN9
M	-3	ARG	-	expression tag	UNP A0A377VEN9
M	-2	GLY	-	expression tag	UNP A0A377VEN9
M	-1	SER	-	expression tag	UNP A0A377VEN9
M	0	HIS	-	expression tag	UNP A0A377VEN9
M	1	MET	-	expression tag	UNP A0A377VEN9
M	2	LYS	-	expression tag	UNP A0A377VEN9
M	3	GLN	-	expression tag	UNP A0A377VEN9
M	4	GLY	-	expression tag	UNP A0A377VEN9
M	5	LEU	-	expression tag	UNP A0A377VEN9
M	6	GLN	-	expression tag	UNP A0A377VEN9
M	7	LEU	-	expression tag	UNP A0A377VEN9
M	8	ARG	-	expression tag	UNP A0A377VEN9
M	9	LEU	-	expression tag	UNP A0A377VEN9
M	10	SER	-	expression tag	UNP A0A377VEN9
M	11	GLN	-	expression tag	UNP A0A377VEN9
M	12	GLN	-	expression tag	UNP A0A377VEN9
M	13	LEU	-	expression tag	UNP A0A377VEN9
M	14	ALA	-	expression tag	UNP A0A377VEN9
M	15	MET	-	expression tag	UNP A0A377VEN9
M	16	THR	-	expression tag	UNP A0A377VEN9
M	17	PRO	-	expression tag	UNP A0A377VEN9
M	18	GLN	-	expression tag	UNP A0A377VEN9
M	19	LEU	-	expression tag	UNP A0A377VEN9
M	20	GLN	-	expression tag	UNP A0A377VEN9
M	21	GLN	-	expression tag	UNP A0A377VEN9
M	22	ALA	-	expression tag	UNP A0A377VEN9
M	23	ILE	-	expression tag	UNP A0A377VEN9

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
M	24	ARG	-	expression tag	UNP A0A377VEN9
M	25	LEU	-	expression tag	UNP A0A377VEN9

- Molecule 6 is a DNA chain called NIFH PROMOTER NON-TEMPLATE DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
6	N	46	946	448	173	279	46	0	0

- Molecule 7 is a DNA chain called NIFH PROMOTER TEMPLATE DNA.

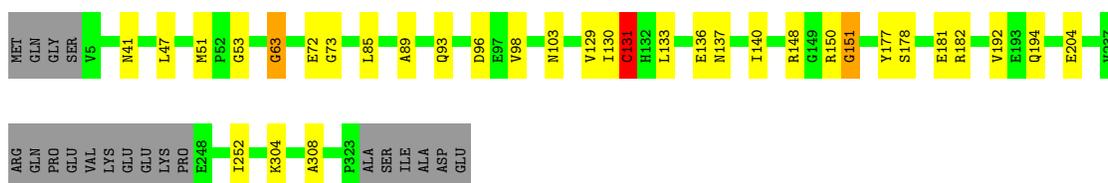
Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
7	T	41	842	396	165	240	41	0	0

### 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.

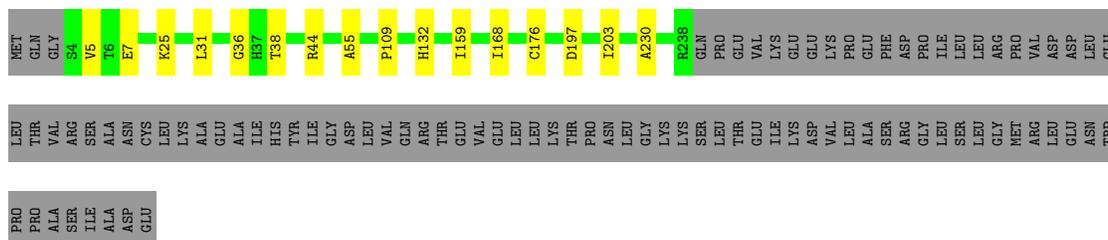
- Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain A:  84% 9% 6%



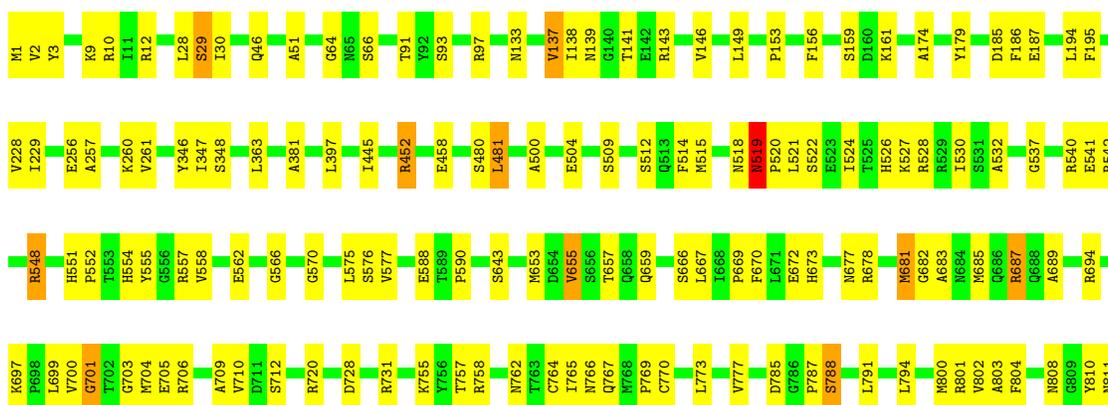
- Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain B:  67% 5% 29%



- Molecule 2: DNA-directed RNA polymerase subunit beta

Chain C:  79% 18% 3%







## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, POINT, POINT, POINT	Depositor
Number of particles used	20233, 20233, 79678, 20233	Depositor
Resolution determination method	FSC 0.143 CUT-OFF, FSC 0.143 CUT-OFF, FSC 0.143 CUT-OFF, FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION, PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	TFS KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	40	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	Not provided	
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor

## 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	1.50	13/1524 (0.9%)	1.31	9/2119 (0.4%)
1	B	1.15	5/1159 (0.4%)	1.05	1/1612 (0.1%)
2	C	1.77	159/6598 (2.4%)	1.50	93/9172 (1.0%)
3	D	1.67	133/6614 (2.0%)	1.48	83/9188 (0.9%)
4	E	1.21	2/370 (0.5%)	1.19	1/514 (0.2%)
5	M	0.94	7/1567 (0.4%)	0.98	9/2183 (0.4%)
6	N	0.50	0/1060	0.82	1/1635 (0.1%)
7	T	0.70	1/945 (0.1%)	1.28	14/1453 (1.0%)
All	All	1.53	320/19837 (1.6%)	1.37	211/27876 (0.8%)

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	A	0	6
2	C	0	24
3	D	0	24
4	E	0	1
5	M	0	1
All	All	0	56

The worst 5 of 320 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	555	TYR	C-O	10.89	1.44	1.23
2	C	526	HIS	C-O	-10.65	1.03	1.23
2	C	1295	SER	CA-CB	10.07	1.68	1.52
3	D	360	TYR	N-CA	9.79	1.66	1.46
2	C	794	LEU	C-O	9.78	1.42	1.23

The worst 5 of 211 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	T	0	DC	O5'-P-OP2	-14.62	92.55	105.70
7	T	-1	DC	O5'-P-OP1	-13.57	93.49	105.70
3	D	353	SER	N-CA-CB	12.88	129.83	110.50
2	C	554	HIS	N-CA-CB	-10.74	91.26	110.60
3	D	765	GLU	C-N-CA	-10.48	100.29	122.30

There are no chirality outliers.

5 of 56 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	131	CYS	Mainchain
1	A	151	GLY	Mainchain
1	A	47	LEU	Mainchain
1	A	63	GLY	Peptide
1	A	93	GLN	Peptide

## 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	1526	0	684	6	0
1	B	1160	0	526	4	0
2	C	6599	0	2971	51	0
3	D	6619	0	3083	75	0
4	E	371	0	172	5	0
5	M	1570	0	692	43	0
6	N	946	0	518	62	0
7	T	842	0	457	60	0
All	All	19633	0	9103	274	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 10.

The worst 5 of 274 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)
6:N:14:DT:H2''	6:N:15:DT:C7	1.31	1.56

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:M:378:GLU:CB	6:N:-16:DT:H73	1.64	1.26
6:N:14:DT:C2'	6:N:15:DT:C7	2.22	1.18
2:C:1262:LYS:O	2:C:1263:ALA:HB3	1.44	1.15
5:M:378:GLU:CB	6:N:-16:DT:C7	2.24	1.15

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 PDB entries followed by that with respect to all EM entries.

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	305/329 (93%)	269 (88%)	32 (10%)	4 (1%)	10	42
1	B	233/329 (71%)	209 (90%)	22 (9%)	2 (1%)	14	51
2	C	1339/1342 (100%)	1183 (88%)	128 (10%)	28 (2%)	5	30
3	D	1335/1407 (95%)	1133 (85%)	163 (12%)	39 (3%)	3	23
4	E	73/91 (80%)	68 (93%)	3 (4%)	2 (3%)	4	26
5	M	310/497 (62%)	254 (82%)	41 (13%)	15 (5%)	2	17
All	All	3595/3995 (90%)	3116 (87%)	389 (11%)	90 (2%)	7	27

5 of 90 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	252	ILE
2	C	29	SER
2	C	347	ILE
2	C	541	GLU
2	C	1155	VAL

### 5.3.2 Protein sidechains [i](#)

There are no protein residues with a non-rotameric sidechain to report in this entry.

### 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 oligosaccharides 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.