



## wwPDB EM Validation Summary Report ⓘ

Mar 4, 2024 – 09:46 AM EST

PDB ID : 6EDT  
EMDB ID : EMD-9037  
Title : Mycobacterium tuberculosis RNAP open promoter complex with RbpA/CarD and AP3 promoter  
Authors : Darst, S.A.; Campbell, E.A.; Boyaci Selcuk, H.; Chen, J.  
Deposited on : 2018-08-10  
Resolution : Not provided

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 : 0.0.1.dev70  
MolProbity : 4.02b-467  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
MapQ : **FAILED**  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

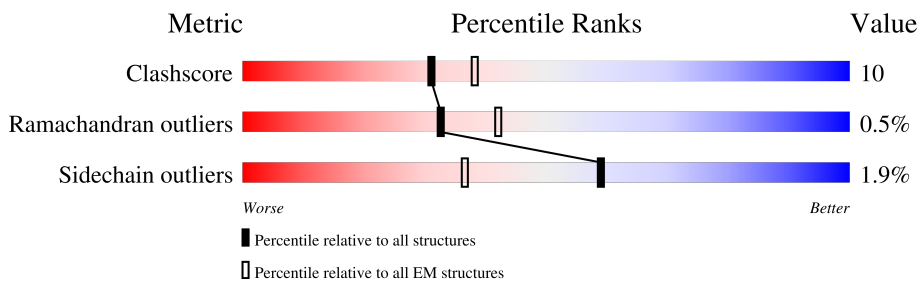
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is unknown.

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	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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	347	57% 7% . 35%
1	B	347	56% 12% 32%
2	C	1134	81% 16% ..
3	D	1371	75% 16% . 7%
4	E	110	62% 13% . 25%
5	F	531	54% 6% 39%
6	J	111	84% 13% ..
7	O	90	40% 32% 28%
8	P	90	44% 24% . 28%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
9	M	162	 69% 26% . .

## 2 Entry composition [i](#)

There are 11 unique types of molecules in this entry. The entry contains 29979 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	S		
1	A	225	Total	C	N	O	S	0	0
			1716	1080	296	338	2		
1	B	237	Total	C	N	O	S	0	0
			1765	1115	301	346	3		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	C	1111	Total	C	N	O	S	0	0
			8606	5392	1511	1664	39		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	D	1270	Total	C	N	O	S	0	0
			9914	6208	1802	1862	42		

There are 55 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	-46	ASP	-	expression tag	UNP A5U053
D	-45	GLY	-	expression tag	UNP A5U053
D	-44	ALA	-	expression tag	UNP A5U053
D	-43	ALA	-	expression tag	UNP A5U053
D	-42	ILE	-	expression tag	UNP A5U053
D	-41	GLU	-	expression tag	UNP A5U053
D	-40	LEU	-	expression tag	UNP A5U053
D	-39	ARG	-	expression tag	UNP A5U053
D	-38	GLU	-	expression tag	UNP A5U053
D	-37	GLY	-	expression tag	UNP A5U053
D	-36	GLU	-	expression tag	UNP A5U053
D	-35	ASP	-	expression tag	UNP A5U053
D	-34	GLU	-	expression tag	UNP A5U053

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	-33	ASP	-	expression tag	UNP A5U053
D	-32	LEU	-	expression tag	UNP A5U053
D	-31	GLU	-	expression tag	UNP A5U053
D	-30	ARG	-	expression tag	UNP A5U053
D	-29	ALA	-	expression tag	UNP A5U053
D	-28	ALA	-	expression tag	UNP A5U053
D	-27	ALA	-	expression tag	UNP A5U053
D	-26	ASN	-	expression tag	UNP A5U053
D	-25	LEU	-	expression tag	UNP A5U053
D	-24	GLY	-	expression tag	UNP A5U053
D	-23	ILE	-	expression tag	UNP A5U053
D	-22	ASN	-	expression tag	UNP A5U053
D	-21	LEU	-	expression tag	UNP A5U053
D	-20	SER	-	expression tag	UNP A5U053
D	-19	ARG	-	expression tag	UNP A5U053
D	-18	ASN	-	expression tag	UNP A5U053
D	-17	GLU	-	expression tag	UNP A5U053
D	-16	SER	-	expression tag	UNP A5U053
D	-15	ALA	-	expression tag	UNP A5U053
D	-14	SER	-	expression tag	UNP A5U053
D	-13	VAL	-	expression tag	UNP A5U053
D	-12	GLU	-	expression tag	UNP A5U053
D	-11	ASP	-	expression tag	UNP A5U053
D	-10	LEU	-	expression tag	UNP A5U053
D	-9	ALA	-	expression tag	UNP A5U053
D	-8	LEU	-	expression tag	UNP A5U053
D	-7	ALA	-	expression tag	UNP A5U053
D	-6	ARG	-	expression tag	UNP A5U053
D	-5	HIS	-	expression tag	UNP A5U053
D	-4	GLY	-	expression tag	UNP A5U053
D	-3	GLY	-	expression tag	UNP A5U053
D	-2	SER	-	expression tag	UNP A5U053
D	-1	GLY	-	expression tag	UNP A5U053
D	0	ALA	-	expression tag	UNP A5U053
D	1317	HIS	-	expression tag	UNP A5U053
D	1318	HIS	-	expression tag	UNP A5U053
D	1319	HIS	-	expression tag	UNP A5U053
D	1320	HIS	-	expression tag	UNP A5U053
D	1321	HIS	-	expression tag	UNP A5U053
D	1322	HIS	-	expression tag	UNP A5U053
D	1323	HIS	-	expression tag	UNP A5U053
D	1324	HIS	-	expression tag	UNP A5U053

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

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
4	E	83	649	414	108	127	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	1	GLY	-	expression tag	UNP A0A0T9N9K3

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	F	322	2540	1583	459	489	9	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	-2	GLY	-	expression tag	UNP P9WGI0
F	-1	PRO	-	expression tag	UNP P9WGI0
F	0	HIS	-	expression tag	UNP P9WGI0

- Molecule 6 is a protein called RNA polymerase-binding protein RbpA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
6	J	109	880	543	166	168	3	0	0

- Molecule 7 is a DNA chain called DNA (65-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
7	O	65	1336	633	243	395	65	0	0

- Molecule 8 is a DNA chain called DNA (65-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
8	P	65	1329	629	250	385	65	0	0

- Molecule 9 is a protein called RNA polymerase-binding transcription factor CarD.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
9	M	159	1241	777	224	239	1	0	0

- Molecule 10 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
			Total	Zn	
10	D	2	2	2	0

- Molecule 11 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

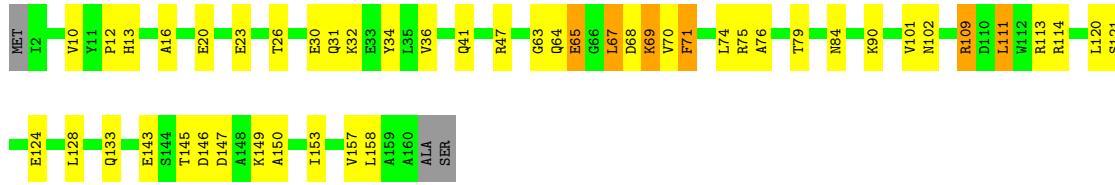
Mol	Chain	Residues	Atoms		AltConf
			Total	Mg	
11	D	1	1	1	0











## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	211381	Depositor
Resolution determination method	Not provided	
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	69.9	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	Not provided	
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, 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	A	0.44	0/1742	0.59	0/2370
1	B	0.45	0/1792	0.62	0/2442
2	C	0.50	0/8765	0.64	1/11885 (0.0%)
3	D	0.52	0/10078	0.63	4/13624 (0.0%)
4	E	0.40	0/662	0.64	0/901
5	F	0.37	0/2571	0.55	0/3468
6	J	0.38	0/896	0.62	0/1210
7	O	0.84	0/1497	0.96	0/2310
8	P	0.87	0/1491	1.04	5/2297 (0.2%)
9	M	0.39	0/1257	0.64	0/1700
All	All	0.53	0/30751	0.67	10/42207 (0.0%)

There are no bond length outliers.

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	P	77	DC	OP1-P-O3'	8.55	124.02	105.20
8	P	77	DC	P-O3'-C3'	7.37	128.54	119.70
3	D	583	THR	CB-CA-C	-6.80	93.24	111.60
2	C	231	ARG	N-CA-C	6.78	129.30	111.00
3	D	1007	GLY	N-CA-C	-6.60	96.60	113.10

There are no chirality outliers.

There are no planarity outliers.

### 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	1716	0	1756	18	0
1	B	1765	0	1794	27	0
2	C	8606	0	8544	218	0
3	D	9914	0	9987	206	0
4	E	649	0	645	10	0
5	F	2540	0	2552	33	0
6	J	880	0	852	12	0
7	O	1336	0	732	37	0
8	P	1329	0	727	36	0
9	M	1241	0	1259	38	0
10	D	2	0	0	0	0
11	D	1	0	0	0	0
All	All	29979	0	28848	572	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 572 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)
2:C:776:ILE:CD1	2:C:780:VAL:HG21	1.23	1.59
2:C:231:ARG:CD	2:C:232:GLN:H	0.90	1.53
2:C:278:TYR:CE2	2:C:282:ARG:HG3	1.55	1.41
2:C:231:ARG:HD2	2:C:232:GLN:N	1.06	1.39
3:D:1010:LEU:CD1	3:D:1145:GLN:HG3	1.52	1.37

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	223/347 (64%)	207 (93%)	16 (7%)	0	100	100
1	B	235/347 (68%)	210 (89%)	24 (10%)	1 (0%)	34	34
2	C	1109/1134 (98%)	991 (89%)	107 (10%)	11 (1%)	15	15
3	D	1264/1371 (92%)	1183 (94%)	77 (6%)	4 (0%)	41	41
4	E	81/110 (74%)	73 (90%)	8 (10%)	0	100	100
5	F	320/531 (60%)	303 (95%)	17 (5%)	0	100	100
6	J	107/111 (96%)	95 (89%)	11 (10%)	1 (1%)	17	17
9	M	157/162 (97%)	147 (94%)	9 (6%)	1 (1%)	25	25
All	All	3496/4113 (85%)	3209 (92%)	269 (8%)	18 (0%)	32	29

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	231	ARG
3	D	1008	THR
3	D	1011	THR
2	C	230	ARG
2	C	775	ASN

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

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	
1	A	194/297 (65%)	191 (98%)	3 (2%)	65	65
1	B	195/297 (66%)	194 (100%)	1 (0%)	88	88
2	C	937/965 (97%)	920 (98%)	17 (2%)	59	59
3	D	1048/1135 (92%)	1029 (98%)	19 (2%)	59	59
4	E	69/89 (78%)	68 (99%)	1 (1%)	67	67
5	F	266/429 (62%)	261 (98%)	5 (2%)	57	57
6	J	92/97 (95%)	89 (97%)	3 (3%)	38	38
9	M	129/131 (98%)	122 (95%)	7 (5%)	22	22

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	2930/3440 (85%)	2874 (98%)	56 (2%)	59 57

5 of 56 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	733	MET
9	M	111	LEU
3	D	1013	ARG
9	M	109	ARG
9	M	65	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 29 such sidechains are listed below:

Mol	Chain	Res	Type
3	D	307	ASN
9	M	13	HIS
3	D	687	GLN
5	F	322	GLN
3	D	657	GLN

### 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 monosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

Of 3 ligands modelled in this entry, 3 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.



There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

There are no chain breaks in this entry.

## 6 Map visualisation

This section contains visualisations of the EMDB entry EMD-9037. These allow visual inspection of the internal detail of the map and identification of artifacts.

No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

### 6.1 Orthogonal projections

This section was not generated.

### 6.2 Central slices

This section was not generated.

### 6.3 Largest variance slices

This section was not generated.

### 6.4 Orthogonal standard-deviation projections (False-color)

This section was not generated.

### 6.5 Orthogonal surface views

This section was not generated.

### 6.6 Mask visualisation

This section was not generated. No masks/segmentation were deposited.

## 7 Map analysis

This section contains the results of statistical analysis of the map.

### 7.1 Map-value distribution

This section was not generated.

### 7.2 Volume estimate versus contour level

This section was not generated.

### 7.3 Rotationally averaged power spectrum

This section was not generated. The rotationally averaged power spectrum had issues being displayed.

## 8 Fourier-Shell correlation

This section was not generated. No FSC curve or half-maps provided.

## 9 Map-model fit

This section was not generated.