



## wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Nov 29, 2017 – 07:08 PM EST

PDB ID : 6B6H  
EMDB ID: : EMD-7059  
Title : The cryo-EM structure of a bacterial class I transcription activation complex  
Authors : Liu, B.; Hong, C.; Huang, R.; Yu, Z.; Steitz, T.A.  
Deposited on : unknown  
Resolution : 3.90 Å(reported)

This is a wwPDB/EMDataBank EM Map/Model Validation Summary Report  
for a publicly released PDB/EMDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/EMValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

MolProbity : 4.02b-467  
Mogul : 1.7.2 (RC1), CSD as538be (2017)  
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) : rb-20030345

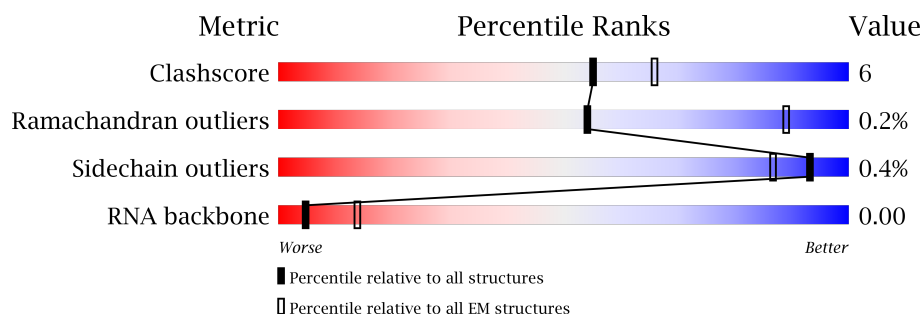
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.90 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	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026
RNA backbone	3398	335

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	329	58% 12% 30%
1	B	329	57% 12% 31%
2	C	1342	81% 18%
3	D	1407	79% 16% 5%
4	E	91	76% 11% 13%
5	F	628	67% 12% 21%
6	G	210	84% 11% .
6	H	210	77% 18% .

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
7	I	75	<div><div></div><div>84%</div><div>16%</div></div>
8	1	88	<div><div></div><div>82%</div><div>18%</div></div>
9	2	88	<div><div></div><div>78%</div><div>18%</div><div></div></div>
10	3	3	<div><div></div><div>33%</div><div>33%</div><div>33%</div></div>

## 2 Entry composition [i](#)

There are 13 unique types of molecules in this entry. The entry contains 72103 atoms, of which 35433 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
1	A	230	Total	C	H	N	O	S	0	0
			3599	1112	1813	317	351	6		
1	B	228	Total	C	H	N	O	S	0	0
			3556	1100	1789	312	349	6		

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

Mol	Chain	Residues	Atoms						AltConf	Trace
2	C	1340	Total	C	H	N	O	S	0	0
			21152	6631	10582	1841	2055	43		

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

Mol	Chain	Residues	Atoms						AltConf	Trace
3	D	1337	Total	C	H	N	O	S	0	0
			21010	6531	10614	1853	1962	50		

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

Mol	Chain	Residues	Atoms						AltConf	Trace
4	E	79	Total	C	H	N	O	S	0	0
			1261	382	634	118	126	1		

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

Mol	Chain	Residues	Atoms						AltConf	Trace
5	F	497	Total	C	H	N	O	S	0	0
			8105	2512	4083	719	768	23		

There are 15 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	-14	MET	-	initiating methionine	UNP P00579

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
F	-13	ARG	-	expression tag	UNP P00579
F	-12	GLY	-	expression tag	UNP P00579
F	-11	SER	-	expression tag	UNP P00579
F	-10	HIS	-	expression tag	UNP P00579
F	-9	HIS	-	expression tag	UNP P00579
F	-8	HIS	-	expression tag	UNP P00579
F	-7	HIS	-	expression tag	UNP P00579
F	-6	HIS	-	expression tag	UNP P00579
F	-5	HIS	-	expression tag	UNP P00579
F	-4	THR	-	expression tag	UNP P00579
F	-3	ASP	-	expression tag	UNP P00579
F	-2	GLN	-	expression tag	UNP P00579
F	-1	PHE	-	expression tag	UNP P00579
F	0	THR	-	expression tag	UNP P00579

- Molecule 6 is a protein called cAMP-activated global transcriptional regulator CRP.

Mol	Chain	Residues	Atoms						AltConf	Trace
6	G	201	Total	C	H	N	O	S	0	0
			3219	1007	1629	279	295	9		
6	H	201	Total	C	H	N	O	S	0	0
			3223	1007	1632	280	295	9		

- Molecule 7 is a protein called DNA-directed RNA polymerase subunit alpha.

Mol	Chain	Residues	Atoms						AltConf	Trace
7	I	75	Total	C	H	N	O	S	0	0
			1198	370	614	102	110	2		

- Molecule 8 is a DNA chain called SYNTHETIC NONTEMPLATE STRAND DNA (88-MER).

Mol	Chain	Residues	Atoms						AltConf	Trace
8	1	88	Total	C	H	N	O	P	0	0
			2796	860	992	331	525	88		

- Molecule 9 is a DNA chain called SYNTHETIC TEMPLATE STRAND DNA (88-MER).

Mol	Chain	Residues	Atoms						AltConf	Trace
9	2	88	Total	C	H	N	O	P	0	0
			2805	863	996	325	533	88		

- Molecule 10 is a RNA chain called NASCENT RNA 3-MER.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	3	3	Total	C	H	N	O	P	
			110	30	33	15	27	5	
								0	0

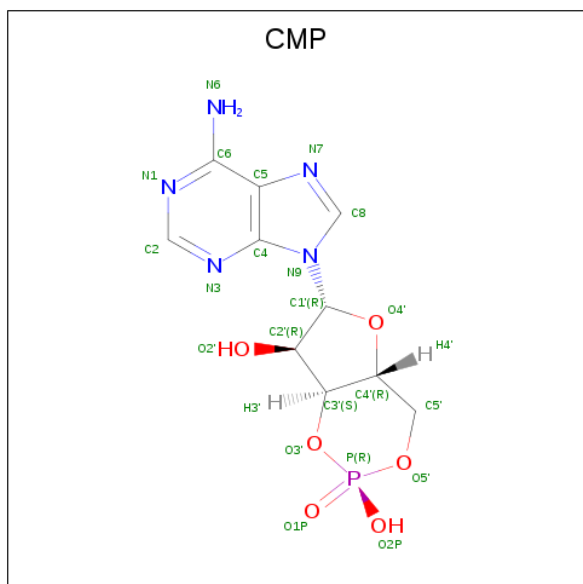
- Molecule 11 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 12 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 13 is ADENOSINE-3',5'-CYCLIC-MONOPHOSPHATE (three-letter code: CMP) (formula: C<sub>10</sub>H<sub>12</sub>N<sub>5</sub>O<sub>6</sub>P).



Mol	Chain	Residues	Atoms					AltConf	
13	G	1	Total 33	C 10	H 11	N 5	O 6	P 1	0
13	H	1	Total 33	C 10	H 11	N 5	O 6	P 1	0



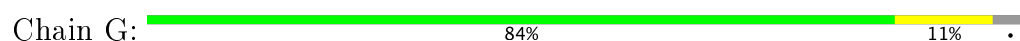
- Molecule 1: DNA-directed RNA polymerase subunit alpha



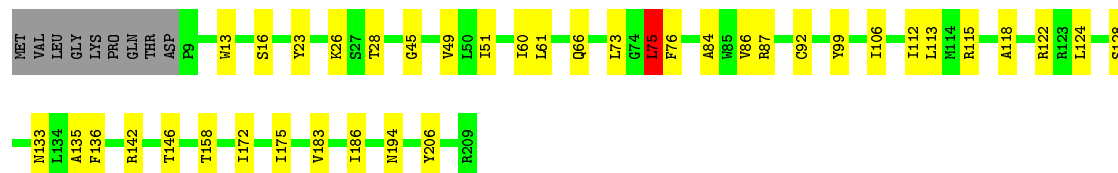
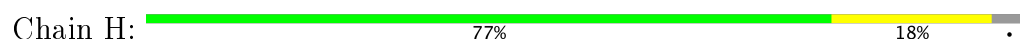




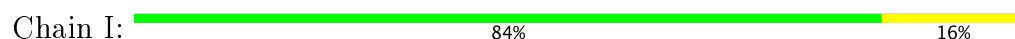
- Molecule 6: cAMP-activated global transcriptional regulator CRP



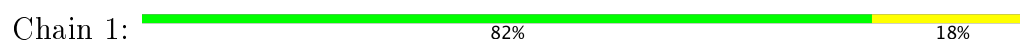
- Molecule 6: cAMP-activated global transcriptional regulator CRP



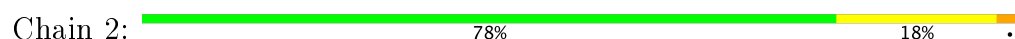
- Molecule 7: DNA-directed RNA polymerase subunit alpha



- Molecule 8: SYNTHETIC NONTEMPLATE STRAND DNA (88-MER)

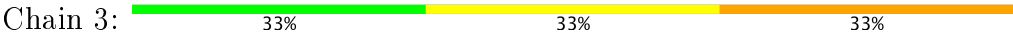


- Molecule 9: SYNTHETIC TEMPLATE STRAND DNA (88-MER)





● Molecule 10: NASCENT RNA 3-MER



## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	161000	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	1.37	Depositor
Minimum defocus (nm)	1200	Depositor
Maximum defocus (nm)	2600	Depositor
Magnification	37037	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

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

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  >2	RMSZ	# Z  >2
1	A	0.40	0/1808	0.61	0/2450
1	B	0.37	0/1789	0.61	0/2425
10	3	0.52	0/50	1.28	0/76
2	C	0.44	0/10739	0.62	0/14489
3	D	0.42	0/10553	0.63	1/14248 (0.0%)
4	E	0.27	0/629	0.53	0/847
5	F	0.34	0/4076	0.60	0/5482
6	G	0.32	0/1614	0.62	1/2170 (0.0%)
6	H	0.33	0/1616	0.66	1/2174 (0.0%)
7	I	0.28	0/592	0.64	0/803
8	1	0.86	1/2024 (0.0%)	1.11	5/3121 (0.2%)
9	2	0.81	0/2028	1.06	6/3130 (0.2%)
All	All	0.47	1/37518 (0.0%)	0.70	14/51415 (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	A	0	1
2	C	0	2
3	D	0	5
5	F	0	3
6	H	0	1
All	All	0	12

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	1	18	DC	C1'-N1	5.90	1.56	1.49

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	1	25	DT	O4'-C1'-N1	7.32	113.12	108.00
8	1	47	DC	O4'-C4'-C3'	-7.06	101.67	104.50
8	1	25	DT	C1'-O4'-C4'	-6.85	103.25	110.10
8	1	50	DA	O4'-C1'-N9	6.79	112.75	108.00
3	D	710	ASP	CB-CG-OD1	5.70	123.43	118.30

There are no chirality outliers.

5 of 12 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	191	ARG	Peptide
2	C	1004	ASP	Peptide
2	C	200	ARG	Peptide
3	D	120	LEU	Peptide
3	D	121	PRO	Peptide

## 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	A	1786	1813	1813	24	0
1	B	1767	1789	1789	24	0
2	C	10570	10582	10582	148	0
3	D	10396	10614	10614	144	0
4	E	627	634	634	9	0
5	F	4022	4083	4083	48	0
6	G	1590	1629	1629	20	0
6	H	1591	1632	1632	26	0
7	I	584	614	614	9	0
8	1	1804	992	992	14	0
9	2	1809	996	996	19	0
10	3	77	33	33	2	0
11	D	2	0	0	0	0
12	D	1	0	0	0	0
13	G	22	11	11	2	0
13	H	22	11	11	4	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	36670	35433	35433	436	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 6.

The worst 5 of 436 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)
13:G:301:CMP:C2	13:G:301:CMP:H2	0.97	1.50
13:H:301:CMP:C2	13:H:301:CMP:H2	0.97	1.49
5:F:551:LEU:HD11	5:F:598:LEU:HD21	1.55	0.86
2:C:494:ASN:ND2	9:2:24:DA:OP1	2.13	0.81
3:D:814:CYS:SG	3:D:883:ARG:NH2	2.54	0.81

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	228/329 (69%)	203 (89%)	25 (11%)	0	100	100
1	B	226/329 (69%)	208 (92%)	18 (8%)	0	100	100
2	C	1338/1342 (100%)	1219 (91%)	115 (9%)	4 (0%)	44	80
3	D	1331/1407 (95%)	1190 (89%)	140 (10%)	1 (0%)	55	88
4	E	77/91 (85%)	76 (99%)	1 (1%)	0	100	100
5	F	493/628 (78%)	438 (89%)	53 (11%)	2 (0%)	38	76
6	G	199/210 (95%)	193 (97%)	6 (3%)	0	100	100
6	H	199/210 (95%)	190 (96%)	9 (4%)	0	100	100
7	I	73/75 (97%)	71 (97%)	2 (3%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	4164/4621 (90%)	3788 (91%)	369 (9%)	7 (0%)	54 84

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	399	ALA
2	C	1004	ASP
2	C	1005	GLU
5	F	156	ALA
2	C	398	SER

### 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 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	198/286 (69%)	197 (100%)	1 (0%)	91 96
1	B	196/286 (68%)	196 (100%)	0	100 100
2	C	1155/1157 (100%)	1152 (100%)	3 (0%)	94 97
3	D	1120/1168 (96%)	1116 (100%)	4 (0%)	93 96
4	E	67/75 (89%)	67 (100%)	0	100 100
5	F	439/554 (79%)	435 (99%)	4 (1%)	82 92
6	G	172/181 (95%)	171 (99%)	1 (1%)	89 95
6	H	173/181 (96%)	171 (99%)	2 (1%)	75 89
7	I	66/66 (100%)	66 (100%)	0	100 100
All	All	3586/3954 (91%)	3571 (100%)	15 (0%)	93 96

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	700	ASN
3	D	798	ARG
6	G	142	ARG
3	D	424	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
5	F	557	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
5	F	128	ASN
7	I	268	ASN
6	G	80	GLN
3	D	1195	GLN
5	F	309	ASN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
10	3	1/3 (33%)	1 (100%)	0

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
10	3	3	G

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](#)

Of 5 ligands modelled in this entry, 3 are monoatomic - leaving 2 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
13	CMP	G	301	-	21,25,25	1.28	3 (14%)	22,39,39	2.20	7 (31%)
13	CMP	H	301	-	21,25,25	1.30	4 (19%)	22,39,39	1.99	4 (18%)

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
13	CMP	G	301	-	-	0/0/31/31	0/4/4/4
13	CMP	H	301	-	-	0/0/31/31	0/4/4/4

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
13	H	301	CMP	O3'-C3'	-2.20	1.41	1.44
13	H	301	CMP	O5'-C5'	-2.18	1.42	1.46
13	G	301	CMP	O5'-C5'	-2.09	1.43	1.46
13	H	301	CMP	C5-C4	2.79	1.46	1.40
13	H	301	CMP	P-O3'	2.79	1.62	1.58

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
13	G	301	CMP	N3-C2-N1	-7.26	122.53	128.86
13	H	301	CMP	N3-C2-N1	-6.70	123.02	128.86
13	G	301	CMP	C4-C5-N7	-2.84	106.67	109.41
13	H	301	CMP	C4-C5-N7	-2.60	106.90	109.41
13	G	301	CMP	C1'-N9-C4	-2.20	122.84	126.64

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
13	G	301	CMP	2	0
13	H	301	CMP	4	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.