



wwPDB EM Validation Summary Report ⓘ

Aug 7, 2023 – 12:12 PM JST

PDB ID : 8IGR
EMDB ID : EMD-35438
Title : Cryo-EM structure of CII-dependent transcription activation complex
Authors : Zhao, M.; Gao, B.; Wen, A.; Feng, Y.; Lu, Y.
Deposited on : 2023-02-21
Resolution : 3.10 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

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:

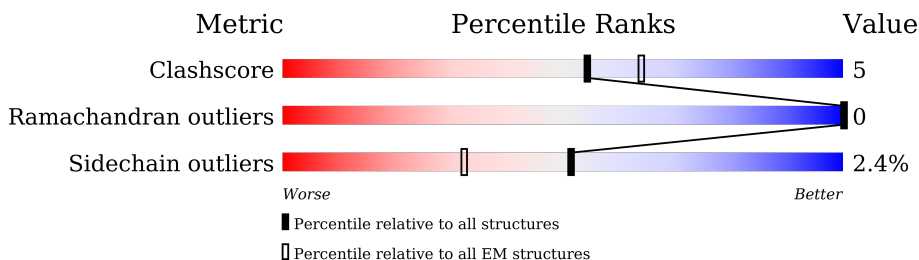
EMDB validation analysis : **FAILED**
MolProbity : 4.02b-467
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ : **FAILED**
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.35

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 3.10 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)	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 ≥ 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	G	329	74% 15% 11%
1	H	329	58% 9% 33%
2	I	1342	73% 13% 13%
3	J	1407	70% 11% 18%
4	K	91	66% 18% 16%
5	L	613	45% 6% 50%
6	A	97	66% 8% 24%
6	B	97	61% 14% 25%
6	C	97	69% 7% 24%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	D	97	
7	N	85	
8	T	85	

2 Entry composition [i](#)

There are 10 unique types of molecules in this entry. The entry contains 30043 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	G	293	Total	C	N	O	S	0	0
			2254	1411	399	437	7		
1	H	219	Total	C	N	O	S	0	0
			1678	1048	295	329	6		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	I	1167	Total	C	N	O	S	0	0
			9176	5757	1597	1782	40		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	J	1154	Total	C	N	O	S	0	0
			9022	5675	1616	1685	46		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	K	76	Total	C	N	O	S	0	0
			605	368	115	121	1		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	L	309	Total	C	N	O	S	0	0
			2500	1561	451	473	15		

- Molecule 6 is a protein called Transcriptional activator II.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	A	74	Total	C	N	O	S	0	0
			583	369	105	106	3		
6	B	73	Total	C	N	O	S	0	0
			577	367	104	103	3		
6	C	74	Total	C	N	O	S	1	0
			587	372	107	105	3		
6	D	71	Total	C	N	O	S	0	0
			557	355	98	101	3		

- Molecule 7 is a DNA chain called nontemplate strand DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	N	61	Total	C	N	O	P	0	0
			1247	599	214	373	61		

- Molecule 8 is a DNA chain called template strand DNA.

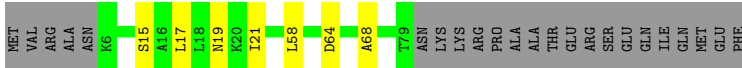
Mol	Chain	Residues	Atoms					AltConf	Trace
8	T	61	Total	C	N	O	P	0	0
			1254	599	235	359	61		

- Molecule 9 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
9	J	1	Total	Mg	0
			1	1	

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

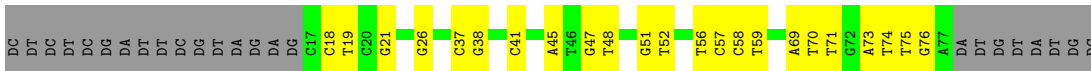
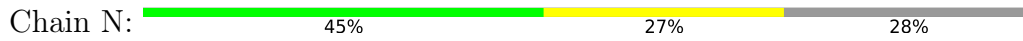
Mol	Chain	Residues	Atoms		AltConf
10	J	2	Total	Zn	0
			2	2	



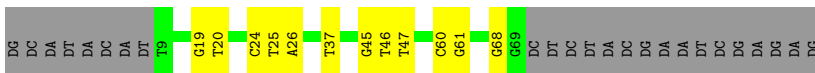
● Molecule 6: Transcriptional activator II



● Molecule 7: nontemplate strand DNA



● Molecule 8: template strand DNA



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	103401	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 ($e^-/\text{\AA}^2$)	51.7	Depositor
Minimum defocus (nm)	800	Depositor
Maximum defocus (nm)	1700	Depositor
Magnification	Not provided	
Image detector	FEI FALCON IV (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: MG, ZN

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	G	0.29	0/2279	0.54	0/3087
1	H	0.28	0/1697	0.55	0/2301
2	I	0.32	0/9325	0.53	0/12584
3	J	0.29	0/9162	0.54	0/12366
4	K	0.26	0/607	0.55	0/817
5	L	0.26	0/2532	0.54	0/3402
6	A	0.24	0/590	0.49	0/796
6	B	0.24	0/584	0.49	0/786
6	C	0.24	0/602	0.50	0/810
6	D	0.25	0/564	0.47	0/761
7	N	0.56	0/1395	0.95	0/2151
8	T	0.55	0/1409	0.91	0/2172
All	All	0.33	0/30746	0.59	0/42033

There are no bond length outliers.

There are no bond angle outliers.

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	G	2254	0	2314	28	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	H	1678	0	1698	17	0
2	I	9176	0	9173	104	0
3	J	9022	0	9213	105	0
4	K	605	0	612	11	0
5	L	2500	0	2554	25	0
6	A	583	0	607	5	0
6	B	577	0	607	8	0
6	C	587	0	611	6	0
6	D	557	0	581	11	0
7	N	1247	0	695	17	0
8	T	1254	0	688	9	0
9	J	1	0	0	0	0
10	J	2	0	0	0	0
All	All	30043	0	29353	321	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 5.

The worst 5 of 321 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:A:36:ASP:OD1	6:A:37:LYS:N	2.22	0.72
6:D:26:THR:HG23	6:D:40:ILE:CD1	2.20	0.71
2:I:890:LYS:NZ	2:I:911:SER:O	2.21	0.71
3:J:652:GLU:N	3:J:652:GLU:OE1	2.24	0.71
2:I:122:VAL:HG21	2:I:493:ILE:HD11	1.74	0.68

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	G	289/329 (88%)	269 (93%)	20 (7%)	0	100	100
1	H	215/329 (65%)	200 (93%)	15 (7%)	0	100	100
2	I	1161/1342 (86%)	1062 (92%)	99 (8%)	0	100	100
3	J	1150/1407 (82%)	1079 (94%)	71 (6%)	0	100	100
4	K	74/91 (81%)	70 (95%)	4 (5%)	0	100	100
5	L	305/613 (50%)	288 (94%)	17 (6%)	0	100	100
6	A	72/97 (74%)	71 (99%)	1 (1%)	0	100	100
6	B	71/97 (73%)	70 (99%)	1 (1%)	0	100	100
6	C	73/97 (75%)	70 (96%)	3 (4%)	0	100	100
6	D	69/97 (71%)	69 (100%)	0	0	100	100
All	All	3479/4499 (77%)	3248 (93%)	231 (7%)	0	100	100

There are no Ramachandran outliers to report.

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	G	249/286 (87%)	242 (97%)	7 (3%)	43	73
1	H	184/286 (64%)	180 (98%)	4 (2%)	52	78
2	I	1005/1157 (87%)	974 (97%)	31 (3%)	40	70
3	J	966/1168 (83%)	946 (98%)	20 (2%)	53	79
4	K	65/75 (87%)	64 (98%)	1 (2%)	65	85
5	L	270/540 (50%)	267 (99%)	3 (1%)	73	89
6	A	61/81 (75%)	58 (95%)	3 (5%)	25	57
6	B	60/81 (74%)	59 (98%)	1 (2%)	60	83
6	C	62/81 (76%)	62 (100%)	0	100	100
6	D	58/81 (72%)	57 (98%)	1 (2%)	60	83
All	All	2980/3836 (78%)	2909 (98%)	71 (2%)	51	76

5 of 71 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	J	1165	PHE
3	J	1258	ARG
5	L	514	ASP
2	I	407	ARG
2	I	377	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.