



wwPDB EM Validation Summary Report ⓘ

Oct 16, 2024 – 12:30 AM JST

PDB ID : 8K29
EMDB ID : EMD-36835
Title : ICP1 Csy-dsDNA complex (form 2)
Authors : Zhang, L.X.; Feng, Y.
Deposited on : 2023-07-12
Resolution : 3.18 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:

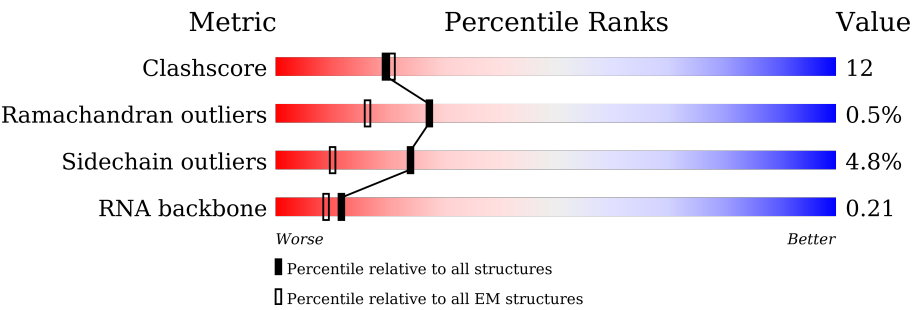
EMDB validation analysis : **FAILED**
MolProbity : 4.02b-467
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ : **FAILED**
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.39

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 3.18 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.






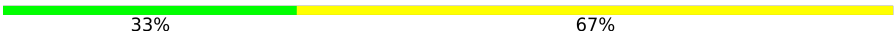
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415
RNA backbone	6643	2191

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	P	60	<div><div>13%53%32%. .</div></div>
2	A	179	<div><div>65%28%. .</div></div>
3	B	248	<div><div>73%24%. .</div></div>
4	C	306	<div><div>77%21%. .</div></div>
4	D	306	<div><div>76%22%. .</div></div>
4	E	306	<div><div>82%16%. .</div></div>
4	F	306	<div><div>78%20%. .</div></div>
4	G	306	<div><div>74%23%. .</div></div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	H	306	 73%24%..
5	I	168	 83%14%..
6	Q	43	 35%63%.
7	R	24	 33%67%

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 21055 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 RNA chain called RNA (60-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
1	P	60	Total	C	N	O	P	0	0
			1260	565	210	425	60		

- Molecule 2 is a protein called Csy1.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	A	174	Total	C	N	O	S	0	0
			1390	881	231	267	11		

- Molecule 3 is a protein called Csy2.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	B	246	Total	C	N	O	S	0	0
			1903	1215	319	355	14		

- Molecule 4 is a protein called Csy3.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	C	302	Total	C	N	O	S	0	0
			2305	1453	391	454	7		
4	D	302	Total	C	N	O	S	0	0
			2305	1453	391	454	7		
4	E	302	Total	C	N	O	S	0	0
			2305	1453	391	454	7		
4	F	302	Total	C	N	O	S	0	0
			2305	1453	391	454	7		
4	G	302	Total	C	N	O	S	0	0
			2305	1453	391	454	7		
4	H	302	Total	C	N	O	S	0	0
			2305	1453	391	454	7		

- Molecule 5 is a protein called Csy4.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	I	165	Total	C	N	O	S	0	0
			1290	818	221	245	6		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
I	51	ILE	VAL	conflict	UNP F1D5V5

- Molecule 6 is a DNA chain called DNA (43-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
6	Q	43	Total	C	N	O	P	0	0
			898	425	184	246	43		


- Molecule 7 is a DNA chain called DNA (5'-D(P*GP*GP*CP*TP*TP*TP*CP*GP*TP*CP*AP*AP*CP*CP*CP*TP*TP*TP*GP*CP*TP*TP*AP*T)-3').

Mol	Chain	Residues	Atoms					AltConf	Trace
7	R	24	Total	C	N	O	P	0	0
			484	233	76	151	24		

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.

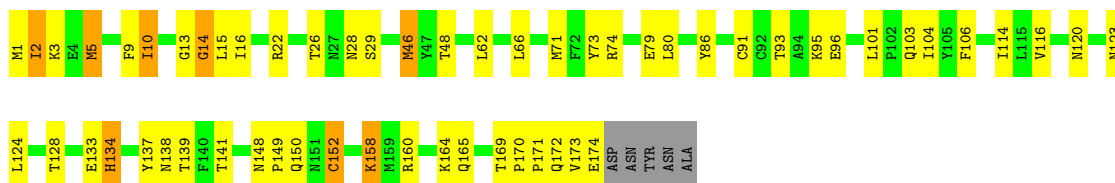
- Molecule 1: RNA (60-MER)

Chain P: 



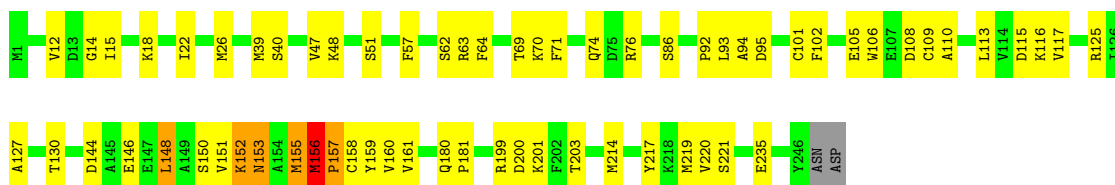
- Molecule 2: Csy1

Chain A: 




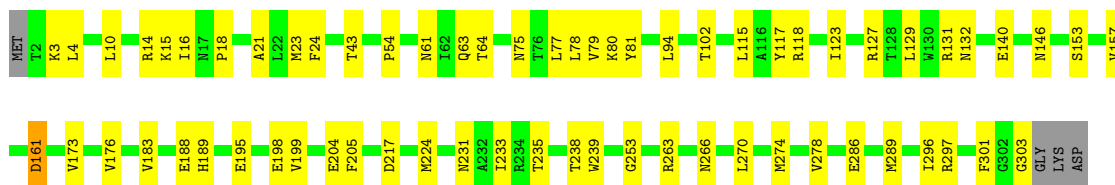
- Molecule 3: Csy2

Chain B: 



- Molecule 4: Csy3

Chain C: 



- Molecule 4: Csy3

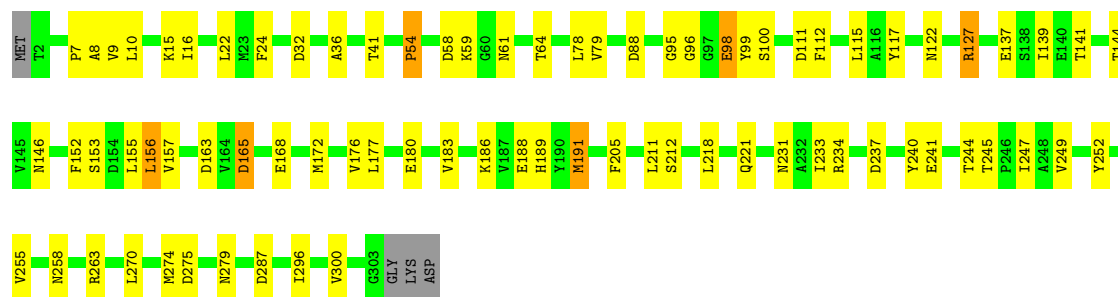
- Molecule 4: Csy3

- Molecule 4: Csy3

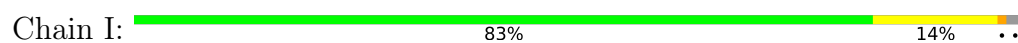
- Molecule 4: Csy3

- Molecule 4: Csy3

Chain H: 73% 24% ..



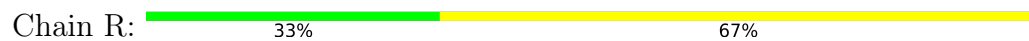
- Molecule 5: Csy4



- Molecule 6: DNA (43-MER)



- Molecule 7: DNA (5'-D(P*GP*GP*CP*TP*TP*TP*CP*GP*TP*CP*AP*AP*CP*CP*CP*TP*TP*TP*GP*CP*TP*TP*AP*T)-3')



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	168154	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 ($e^-/\text{\AA}^2$)	50	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	1500	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor

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	P	0.50	0/1404	1.06	7/2181 (0.3%)
2	A	0.48	0/1422	0.62	0/1919
3	B	0.50	0/1940	0.75	5/2623 (0.2%)
4	C	0.34	0/2344	0.63	1/3183 (0.0%)
4	D	0.32	0/2344	0.59	1/3183 (0.0%)
4	E	0.32	0/2344	0.61	1/3183 (0.0%)
4	F	0.30	0/2344	0.57	1/3183 (0.0%)
4	G	0.44	0/2344	0.66	1/3183 (0.0%)
4	H	0.37	0/2344	0.69	1/3183 (0.0%)
5	I	0.28	0/1308	0.60	0/1757
6	Q	0.64	0/1014	0.99	1/1565 (0.1%)
7	R	0.64	0/538	1.15	0/827
All	All	0.41	0/21690	0.72	19/29970 (0.1%)

There are no bond length outliers.

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	156	MET	CG-SD-CE	-7.54	88.13	100.20
1	P	41	C	P-O3'-C3'	6.71	127.76	119.70
1	P	44	C	C2-N1-C1'	5.96	125.36	118.80
3	B	153	ASN	N-CA-CB	-5.86	100.05	110.60
4	E	54	PRO	N-CA-CB	5.86	110.33	103.30

There are no chirality outliers.

There are no planarity outliers.

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	P	1260	0	639	71	0
2	A	1390	0	1339	58	0
3	B	1903	0	1925	69	0
4	C	2305	0	2239	38	0
4	D	2305	0	2239	55	0
4	E	2305	0	2239	42	0
4	F	2305	0	2238	65	0
4	G	2305	0	2237	52	0
4	H	2305	0	2239	68	0
5	I	1290	0	1336	15	0
6	Q	898	0	475	146	0
7	R	484	0	275	64	0
All	All	21055	0	19420	492	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 12.

The worst 5 of 492 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:H:9:VAL:HG23	6:Q:24:DT:C2	1.26	1.60
1:P:29:A:N6	6:Q:22:DA:N6	1.61	1.41
4:H:300:VAL:CG2	6:Q:23:DT:O2	1.66	1.40
4:H:9:VAL:HG23	6:Q:24:DT:N3	1.39	1.37
2:A:150:GLN:NE2	6:Q:51:DG:N2	1.72	1.32

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	
2	A	172/179 (96%)	161 (94%)	8 (5%)	3 (2%)	7	33
3	B	244/248 (98%)	215 (88%)	27 (11%)	2 (1%)	16	49
4	C	300/306 (98%)	279 (93%)	21 (7%)	0	100	100
4	D	300/306 (98%)	285 (95%)	15 (5%)	0	100	100
4	E	300/306 (98%)	285 (95%)	15 (5%)	0	100	100
4	F	300/306 (98%)	285 (95%)	14 (5%)	1 (0%)	37	67
4	G	300/306 (98%)	281 (94%)	17 (6%)	2 (1%)	19	52
4	H	300/306 (98%)	272 (91%)	25 (8%)	3 (1%)	13	45
5	I	163/168 (97%)	144 (88%)	17 (10%)	2 (1%)	11	41
All	All	2379/2431 (98%)	2207 (93%)	159 (7%)	13 (0%)	27	58

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	B	156	MET
3	B	157	PRO
4	F	98	GLU
4	H	54	PRO
4	H	96	GLY

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	
2	A	153/157 (98%)	137 (90%)	16 (10%)	5	22
3	B	209/211 (99%)	192 (92%)	17 (8%)	9	33
4	C	242/251 (96%)	234 (97%)	8 (3%)	33	62
4	D	242/251 (96%)	235 (97%)	7 (3%)	37	65
4	E	242/251 (96%)	237 (98%)	5 (2%)	48	72
4	F	242/251 (96%)	238 (98%)	4 (2%)	56	77
4	G	242/251 (96%)	226 (93%)	16 (7%)	14	41

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	H	242/251 (96%)	228 (94%)	14 (6%)	17	45
5	I	148/151 (98%)	141 (95%)	7 (5%)	22	52
All	All	1962/2025 (97%)	1868 (95%)	94 (5%)	24	52

5 of 94 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	G	23	MET
4	G	239	TRP
4	G	51	ARG
4	G	172	MET
4	H	98	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
2	A	150	GLN
4	C	63	GLN
4	F	61	ASN
4	G	48	GLN
5	I	47	ASN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	P	59/60 (98%)	30 (50%)	5 (8%)

5 of 30 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	P	-5	U
1	P	-2	A
1	P	1	G
1	P	2	U
1	P	7	C

All (5) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	P	19	U
1	P	20	U
1	P	33	G
1	P	36	A
1	P	41	C

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.