



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Jan 30, 2017 – 07:42 PM EST

PDB ID : 5J0N
EMDB ID: : EMD-3400
Title : Lambda excision HJ intermediate
Authors : Van Duyne, G.; Grigorieff, N.; Landy, A.
Deposited on : 2016-03-28
Resolution : 11.00 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
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 : unknown
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20028442

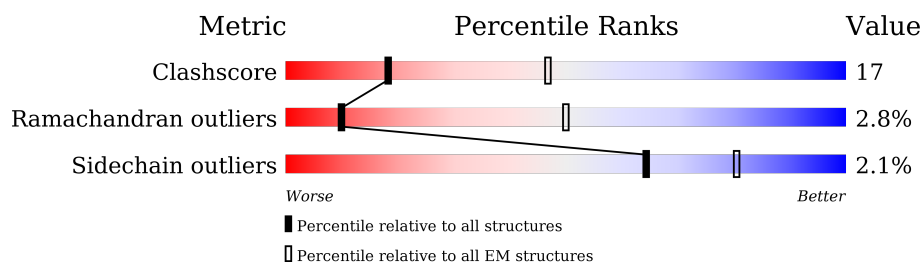
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 11.00 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	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

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 ≥ 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	197	
2	B	139	
3	C	41	
4	D	99	
5	E	356	
5	F	356	
5	G	356	
5	H	356	
6	I	96	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	K	96	 61% 38% .
7	J	94	 73% 24% .
7	L	94	 66% 34%
8	M	55	 51% 44% . .
8	N	55	 58% 35% 5% .
8	O	55	 65% 35%

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 29353 atoms, of which 4591 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 DNA chain called attP(-117 to +79).

Mol	Chain	Residues	Atoms						AltConf	Trace
1	A	197	Total	C	H	N	O	P	0	0
			4377	1944	345	693	1199	196		

- Molecule 2 is a DNA chain called attB(-21) to attP(+117).

Mol	Chain	Residues	Atoms						AltConf	Trace
2	B	139	Total	C	H	N	O	P	0	0
			3109	1367	259	535	810	138		

- Molecule 3 is a DNA chain called attB(-19 to +21).

Mol	Chain	Residues	Atoms						AltConf	Trace
3	C	41	Total	C	H	N	O	P	0	0
			915	401	79	148	247	40		

- Molecule 4 is a DNA chain called attP(-79) to attB(+19).

Mol	Chain	Residues	Atoms						AltConf	Trace
4	D	99	Total	C	H	N	O	P	0	0
			2211	974	183	364	592	98		

- Molecule 5 is a protein called Integrase.

Mol	Chain	Residues	Atoms						AltConf	Trace
5	E	274	Total	C	H	N	O	S	0	0
			2661	1353	510	381	408	9		
5	F	356	Total	C	H	N	O	S	0	0
			3536	1768	703	524	529	12		
5	G	345	Total	C	H	N	O	S	0	0
			3413	1715	673	503	511	11		
5	H	356	Total	C	H	N	O	S	0	0
			3536	1768	703	524	529	12		

- Molecule 6 is a protein called Integration host factor subunit alpha.

Mol	Chain	Residues	Atoms						AltConf	Trace
6	I	96	Total	C	H	N	O	S	0	0
			973	484	199	144	145	1		
6	K	96	Total	C	H	N	O	S	0	0
			973	484	199	144	145	1		

- Molecule 7 is a protein called Integration host factor subunit beta.

Mol	Chain	Residues	Atoms						AltConf	Trace
7	J	94	Total	C	H	N	O	S	0	0
			932	467	183	139	140	3		
7	L	94	Total	C	H	N	O	S	0	0
			932	467	183	139	140	3		

- Molecule 8 is a protein called Excisionase.

Mol	Chain	Residues	Atoms						AltConf	Trace
8	M	53	Total	C	H	N	O	S	0	0
			581	292	120	90	78	1		
8	N	54	Total	C	H	N	O	S	0	0
			598	298	126	94	79	1		
8	O	55	Total	C	H	N	O	S	0	0
			606	303	126	95	81	1		

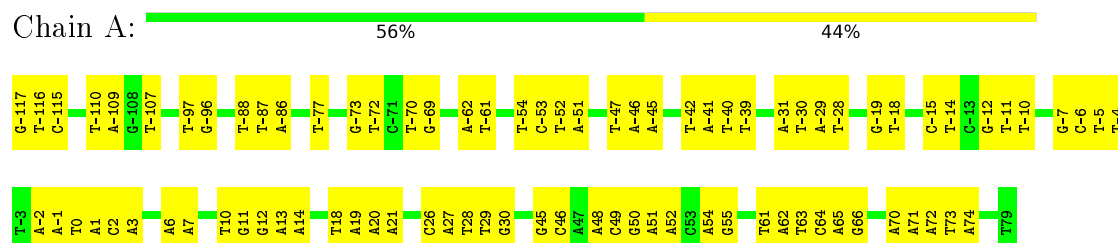
There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
M	28	SER	CYS	conflict	UNP P03699
N	28	SER	CYS	conflict	UNP P03699
O	28	SER	CYS	conflict	UNP P03699

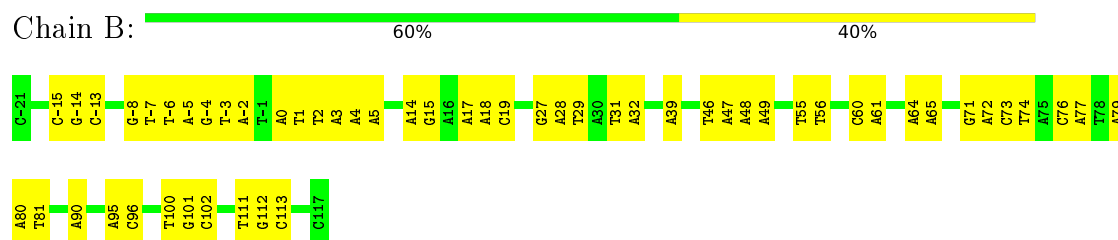
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

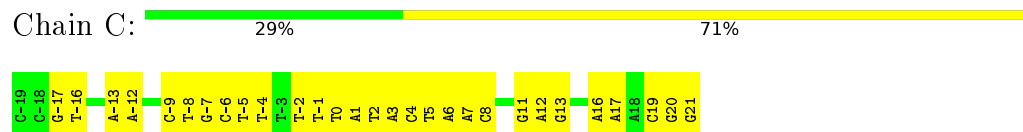
- Molecule 1: attP(-117 to +79)



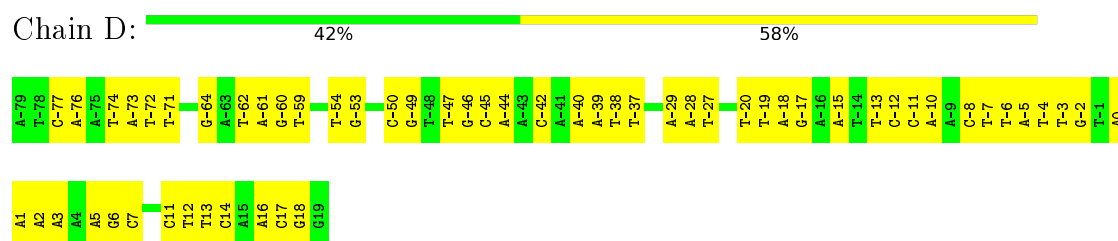
- Molecule 2: attB(-21) to attP(+117)



- Molecule 3: attB(-19 to +21)

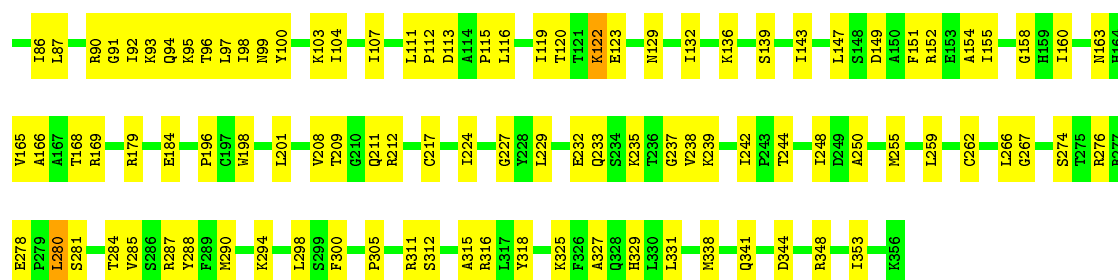


- Molecule 4: attP(-79) to attB(+19)



- Molecule 5: Integrase

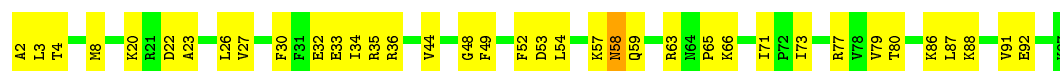




- Molecule 6: Integration host factor subunit alpha



- Molecule 6: Integration host factor subunit alpha



- Molecule 7: Integration host factor subunit beta



- Molecule 7: Integration host factor subunit beta



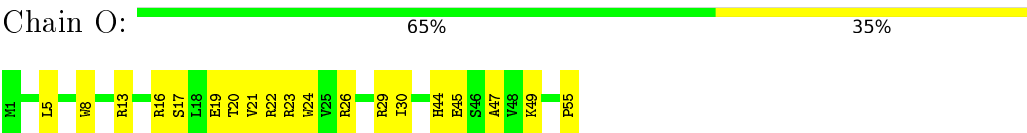
- Molecule 8: Excisionase



- Molecule 8: Excisionase



- Molecule 8: Excisionase



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	10956	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	Not provided	Depositor
Microscope	FEI TECNAI F30	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	Not provided	Depositor
Minimum defocus (nm)	2600.0	Depositor
Maximum defocus (nm)	4150.0	Depositor
Magnification	100000	Depositor
Image detector	Not provided	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 >2	RMSZ	# Z >2
1	A	0.24	0/4516	0.63	0/6971
2	B	0.24	0/3205	0.63	0/4942
3	C	0.24	0/936	0.61	0/1443
4	D	0.24	0/2275	0.63	0/3510
5	E	0.22	0/2185	0.40	0/2940
5	F	0.52	4/2882 (0.1%)	1.92	14/3875 (0.4%)
5	G	0.22	0/2787	0.41	0/3748
5	H	0.22	0/2882	0.40	0/3875
6	I	0.26	0/785	0.38	0/1047
6	K	0.26	0/785	0.39	0/1047
7	J	0.24	0/763	0.39	0/1021
7	L	0.24	0/763	0.40	0/1021
8	M	0.25	0/473	0.42	0/639
8	N	0.25	0/484	0.42	0/653
8	O	0.25	0/493	0.39	0/665
All	All	0.28	4/26214 (0.0%)	0.80	14/37397 (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
5	F	0	1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	F	56	PHE	CG-CD1	13.58	1.59	1.38
5	F	56	PHE	CG-CD2	11.28	1.55	1.38
5	F	56	PHE	CD1-CE1	9.61	1.58	1.39
5	F	56	PHE	CD2-CE2	6.52	1.52	1.39

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	F	56	PHE	CE1-CZ-CE2	-63.52	5.67	120.00
5	F	56	PHE	CD1-CE1-CZ	-53.20	56.26	120.10
5	F	56	PHE	CZ-CE2-CD2	-50.45	59.56	120.10
5	F	56	PHE	CG-CD1-CE1	-34.51	82.83	120.80
5	F	56	PHE	CG-CD2-CE2	-34.10	83.28	120.80

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
5	F	56	PHE	Sidechain

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	4032	345	2249	105	0
2	B	2850	259	1569	56	0
3	C	836	79	466	29	0
4	D	2028	183	1124	72	0
5	E	2151	510	2196	75	0
5	F	2833	703	2871	144	0
5	G	2740	673	2784	127	0
5	H	2833	703	2871	113	0
6	I	774	199	792	23	0
6	K	774	199	792	27	0
7	J	749	183	756	27	0
7	L	749	183	756	25	0
8	M	461	120	464	31	0
8	N	472	126	477	30	0
8	O	480	126	484	14	0
All	All	24762	4591	20651	780	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 17.

The worst 5 of 780 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:B:4:DA:H3'	5:E:93:LYS:HG3	1.32	1.11
1:A:48:DA:H2''	1:A:49:DC:H5''	1.33	1.10
1:A:-7:DG:H2''	1:A:-6:DC:H5''	1.33	1.10
2:B:0:DA:H2''	2:B:1:DT:H5''	1.39	1.03
1:A:11:DG:H2''	1:A:12:DG:H5''	1.43	1.00

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	
5	E	270/356 (76%)	249 (92%)	16 (6%)	5 (2%)	10	52
5	F	354/356 (99%)	303 (86%)	40 (11%)	11 (3%)	5	42
5	G	341/356 (96%)	288 (84%)	38 (11%)	15 (4%)	3	33
5	H	354/356 (99%)	299 (84%)	41 (12%)	14 (4%)	4	35
6	I	94/96 (98%)	87 (93%)	6 (6%)	1 (1%)	17	63
6	K	94/96 (98%)	86 (92%)	7 (7%)	1 (1%)	17	63
7	J	92/94 (98%)	85 (92%)	5 (5%)	2 (2%)	8	49
7	L	92/94 (98%)	87 (95%)	5 (5%)	0	100	100
8	M	51/55 (93%)	45 (88%)	6 (12%)	0	100	100
8	N	52/55 (94%)	42 (81%)	8 (15%)	2 (4%)	4	37
8	O	53/55 (96%)	50 (94%)	3 (6%)	0	100	100
All	All	1847/1969 (94%)	1621 (88%)	175 (10%)	51 (3%)	10	44

5 of 51 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	F	75	THR
5	G	12	LEU
5	H	18	ILE
5	H	55	LEU
5	H	60	LYS

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	
5	E	230/300 (77%)	226 (98%)	4 (2%)	68	87
5	F	300/300 (100%)	293 (98%)	7 (2%)	58	83
5	G	291/300 (97%)	282 (97%)	9 (3%)	47	77
5	H	300/300 (100%)	292 (97%)	8 (3%)	52	79
6	I	83/83 (100%)	83 (100%)	0	100	100
6	K	83/83 (100%)	82 (99%)	1 (1%)	78	90
7	J	79/79 (100%)	79 (100%)	0	100	100
7	L	79/79 (100%)	78 (99%)	1 (1%)	76	89
8	M	50/52 (96%)	49 (98%)	1 (2%)	63	85
8	N	51/52 (98%)	49 (96%)	2 (4%)	39	72
8	O	52/52 (100%)	51 (98%)	1 (2%)	65	86
All	All	1598/1680 (95%)	1564 (98%)	34 (2%)	64	84

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	G	39	ARG
5	G	354	GLU
8	N	39	ARG
5	G	142	LEU
5	F	14	PRO

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
5	G	21	ASN
5	G	52	ASN
6	I	93	ASN
5	F	233	GLN
5	F	308	HIS

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 carbohydrates 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.