



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Aug 17, 2017 – 04:33 PM EDT

PDB ID : 4BIJ
EMDB ID: : EMD-2355
Title : Threading model of T7 large terminase
Authors : Dauden, M.I.; Martin-Benito, J.; Sanchez-Ferrero, J.C.; Pulido-Cid, M.;
Valpuesta, J.M.; Carrascosa, J.L.
Deposited on : unknown
Resolution : 16.00 Å(reported)
Based on PDB ID : 3CPE

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
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-20029824

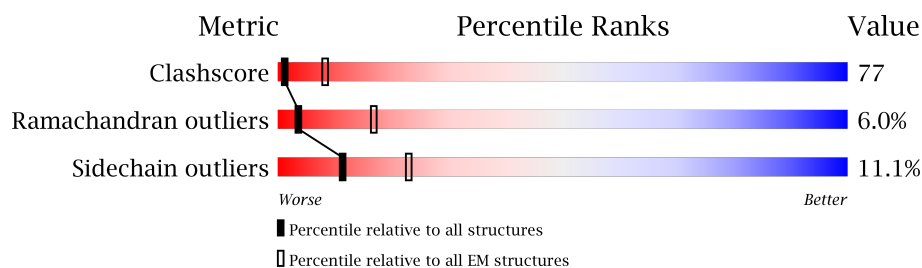
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 16.00 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	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026

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	476	
1	B	476	
1	C	476	
1	D	476	
1	E	476	

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 18855 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.

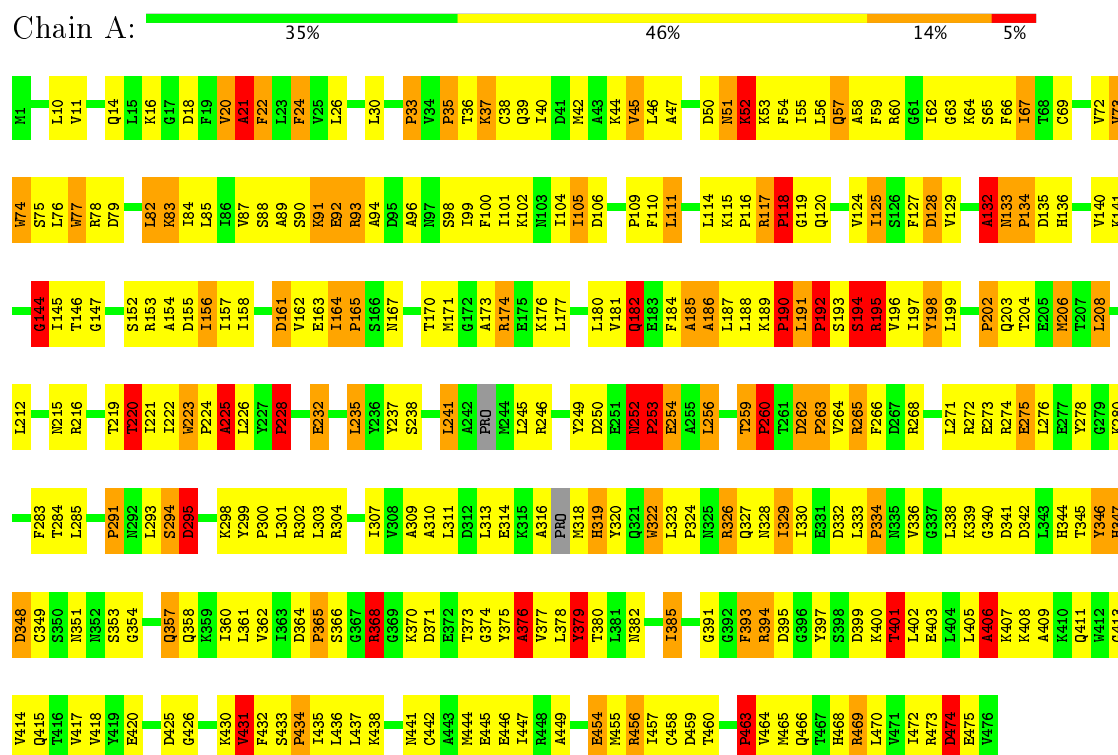
- Molecule 1 is a protein called DNA MATURASE B.

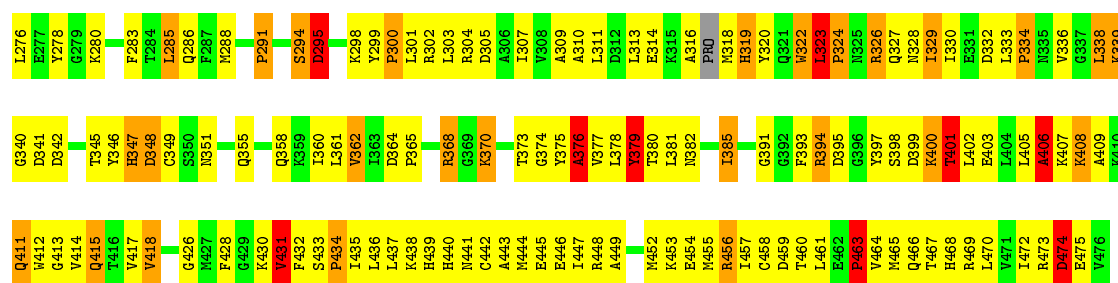
Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	474	Total	C	N	O	S	0	0
			3771	2402	648	703	18		
1	B	474	Total	C	N	O	S	0	0
			3771	2402	648	703	18		
1	C	474	Total	C	N	O	S	0	0
			3771	2402	648	703	18		
1	D	474	Total	C	N	O	S	0	0
			3771	2402	648	703	18		
1	E	474	Total	C	N	O	S	0	0
			3771	2402	648	703	18		

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 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.

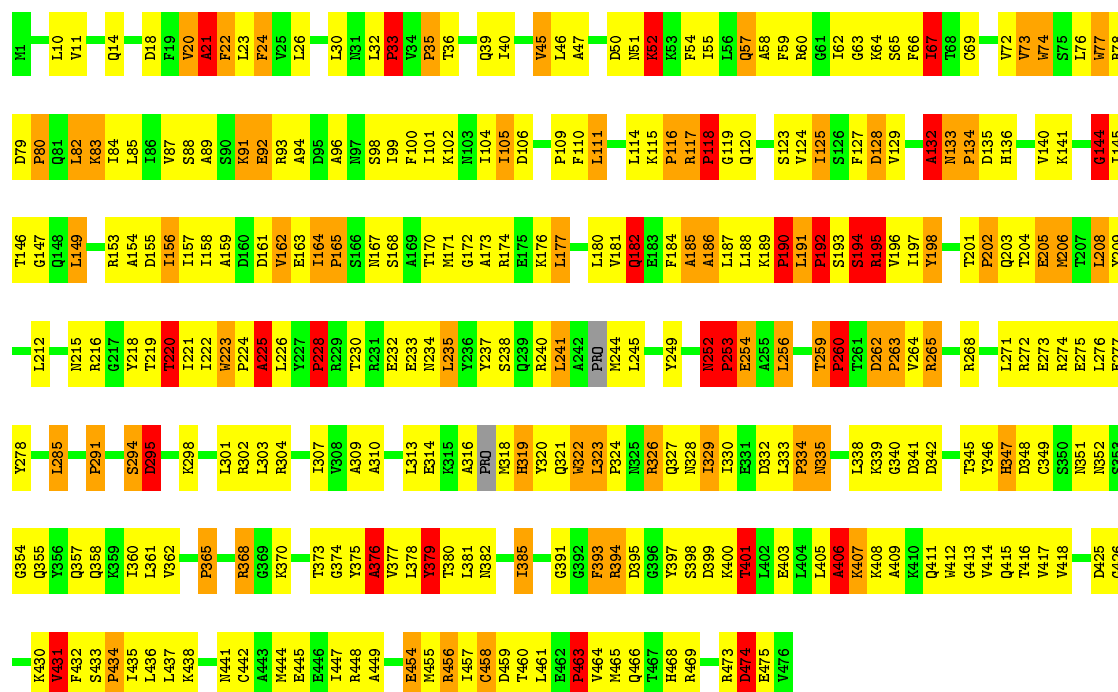
• Molecule 1: DNA MATURASE B



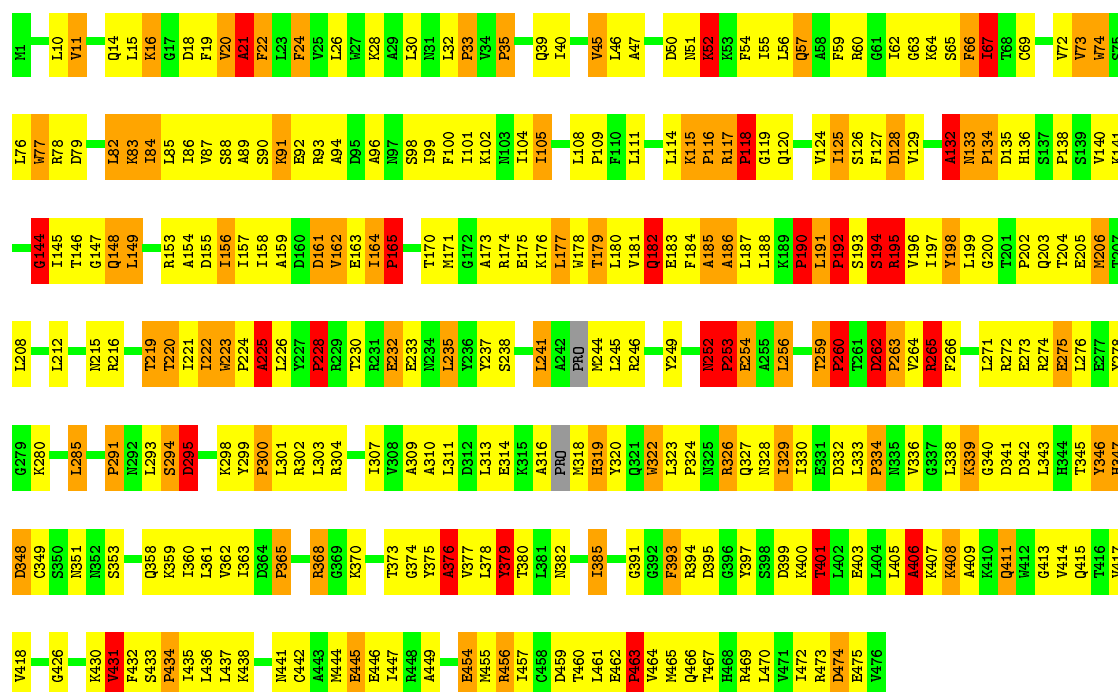
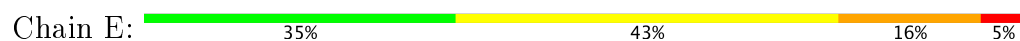


• Molecule 1: DNA MATURASE B

Chain C: 36% 45% 14% 5%



- Molecule 1: DNA MATURASE B



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C5	Depositor
Number of particles used	3650	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	EACH PLATE	Depositor
Microscope	FEI TECNAI F20	Depositor
Voltage (kV)	100	Depositor
Electron dose ($e^-/\text{\AA}^2$)	Not provided	Depositor
Minimum defocus (nm)	1500	Depositor
Maximum defocus (nm)	3200	Depositor
Magnification	67000	Depositor
Image detector	FEI EAGLE (4k x 4k)	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	1.07	22/3850 (0.6%)	1.57	92/5212 (1.8%)
1	B	1.08	22/3850 (0.6%)	1.57	96/5212 (1.8%)
1	C	1.06	23/3850 (0.6%)	1.56	92/5212 (1.8%)
1	D	1.03	22/3850 (0.6%)	1.56	96/5212 (1.8%)
1	E	1.11	22/3850 (0.6%)	1.57	95/5212 (1.8%)
All	All	1.07	111/19250 (0.6%)	1.57	471/26060 (1.8%)

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	13
1	B	0	14
1	C	0	14
1	D	0	14
1	E	0	13
All	All	0	68

The worst 5 of 111 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	291	PRO	N-CD	-23.77	1.14	1.47
1	B	134	PRO	N-CD	20.15	1.76	1.47
1	E	300	PRO	N-CD	19.45	1.75	1.47
1	C	291	PRO	N-CD	-19.44	1.20	1.47
1	B	291	PRO	N-CD	-19.19	1.21	1.47

The worst 5 of 471 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	291	PRO	N-CD-CG	15.74	126.81	103.20

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	291	PRO	N-CD-CG	15.70	126.75	103.20
1	E	77	TRP	CH2-CZ2-CE2	14.73	132.13	117.40
1	A	77	TRP	CH2-CZ2-CE2	14.71	132.12	117.40
1	B	77	TRP	CH2-CZ2-CE2	14.71	132.10	117.40

There are no chirality outliers.

5 of 68 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	144	GLY	Peptide
1	A	194	SER	Mainchain
1	A	21	ALA	Mainchain
1	A	82	LEU	Peptide
1	A	91	LYS	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	3771	0	3778	650	0
1	B	3771	0	3776	649	0
1	C	3771	0	3778	670	0
1	D	3771	0	3778	685	0
1	E	3771	0	3778	661	0
All	All	18855	0	18888	2904	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 77.

The worst 5 of 2904 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)
1:A:346:TYR:N	1:E:173:ALA:CB	1.69	1.56
1:D:173:ALA:CB	1:E:346:TYR:N	1.69	1.54
1:D:346:TYR:CE1	1:D:349:CYS:HB3	1.41	1.52

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:173:ALA:CB	1:C:346:TYR:N	1.69	1.52
1:C:173:ALA:CB	1:D:346:TYR:N	1.69	1.51

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	468/476 (98%)	394 (84%)	46 (10%)	28 (6%)	2	23
1	B	468/476 (98%)	394 (84%)	46 (10%)	28 (6%)	2	23
1	C	468/476 (98%)	394 (84%)	46 (10%)	28 (6%)	2	23
1	D	468/476 (98%)	394 (84%)	46 (10%)	28 (6%)	2	23
1	E	468/476 (98%)	394 (84%)	46 (10%)	28 (6%)	2	23
All	All	2340/2380 (98%)	1970 (84%)	230 (10%)	140 (6%)	3	23

5 of 140 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	83	LYS
1	A	92	GLU
1	A	125	ILE
1	A	132	ALA
1	A	182	GLN

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	409/411 (100%)	372 (91%)	37 (9%)	11	38
1	B	409/411 (100%)	347 (85%)	62 (15%)	3	19
1	C	409/411 (100%)	370 (90%)	39 (10%)	10	36
1	D	409/411 (100%)	363 (89%)	46 (11%)	7	29
1	E	409/411 (100%)	365 (89%)	44 (11%)	7	31
All	All	2045/2055 (100%)	1817 (89%)	228 (11%)	11	29

5 of 228 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	67	ILE
1	C	352	ASN
1	E	228	PRO
1	C	111	LEU
1	C	204	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	352	ASN
1	C	415	GLN
1	E	351	ASN
1	C	357	GLN
1	C	358	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

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.