



wwPDB EM Map/Model Validation Report ⓘ

Apr 10, 2016 – 01:39 PM BST

PDB ID : 2Y7C
EMDB ID: : EMD-1534
Title : Atomic model of the Ocr-bound methylase complex from the Type I restriction-modification enzyme EcoKI (M2S1). Based on fitting into EM map 1534.
Authors : Kennaway, C.K.; Obarska-Kosinska, A.; White, J.H.; Tuszynska, I.; Cooper, L.P.; Bujnicki, J.M.; Trinick, J.; Dryden, D.T.F.
Deposited on : 2011-01-31
Resolution : 18.00 Å(reported)
Based on PDB ID : 1S7Z, 1YF2, 2AR0

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could stem from errors in the original structure(s) used in the fitting.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/EMValidationReportHelp>

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) : trunk27241

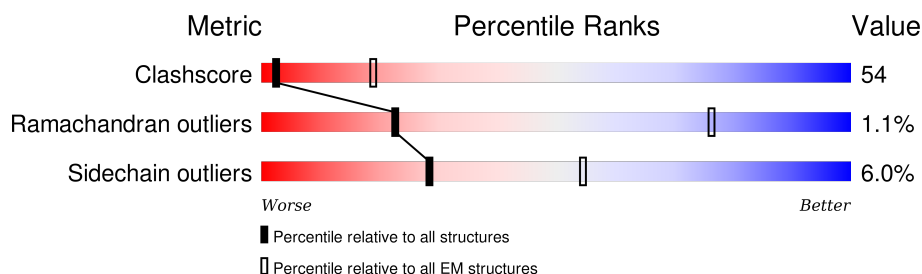
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY






The reported resolution of this entry is 18.00 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	464	
2	B	529	
2	C	529	
3	D	116	
3	E	116	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 13725 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 TYPE-1 RESTRICTION ENZYME ECOKI SPECIFICITY PROTEIN.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	464	Total	C	N	O	S	0	0
			3622	2298	644	671	9		

- Molecule 2 is a protein called TYPE I RESTRICTION ENZYME ECOKI M PROTEIN.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	529	Total	C	N	O	S	0	0
			4175	2612	730	816	17		
2	C	529	Total	C	N	O	S	0	0
			4175	2612	730	816	17		

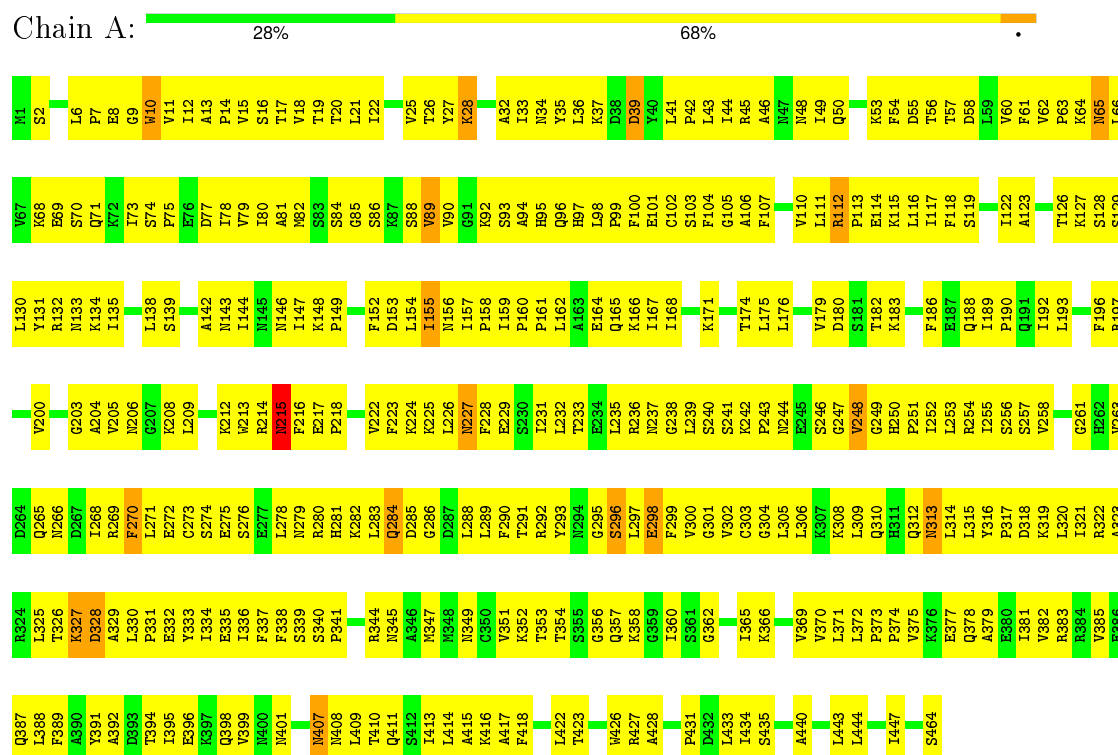
- Molecule 3 is a protein called GENE 0.3 PROTEIN.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	D	107	Total	C	N	O	S	0	1
			877	550	137	185	5		
3	E	106	Total	C	N	O	S	0	0
			876	550	136	185	5		

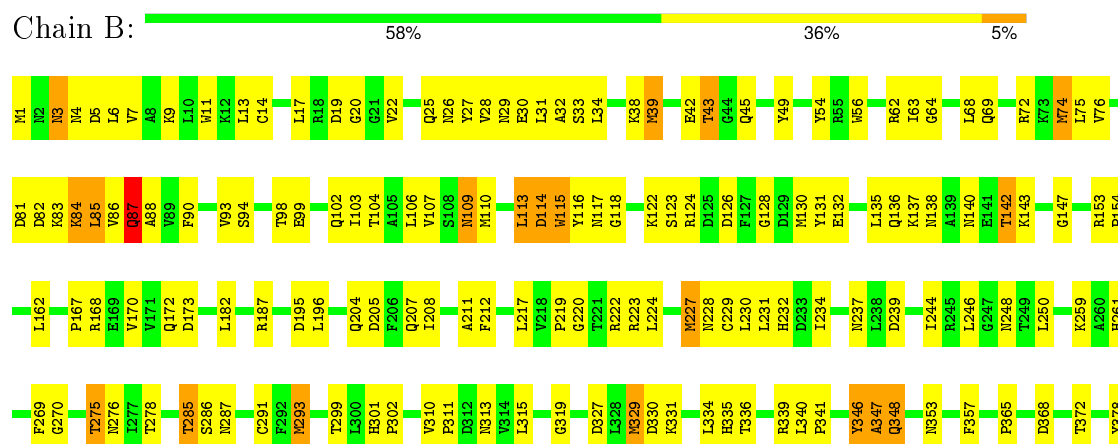
3 Residue-property plots

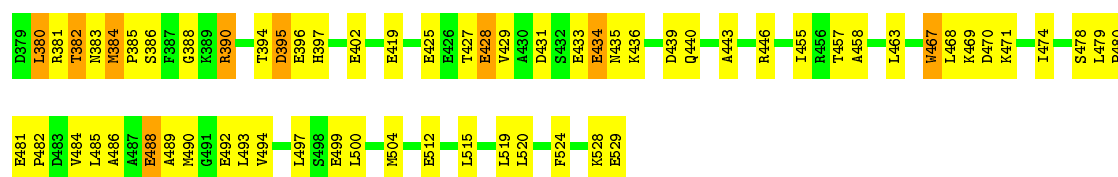
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according to the number of geometric 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: TYPE-1 RESTRICTION ENZYME ECOKI SPECIFICITY PROTEIN

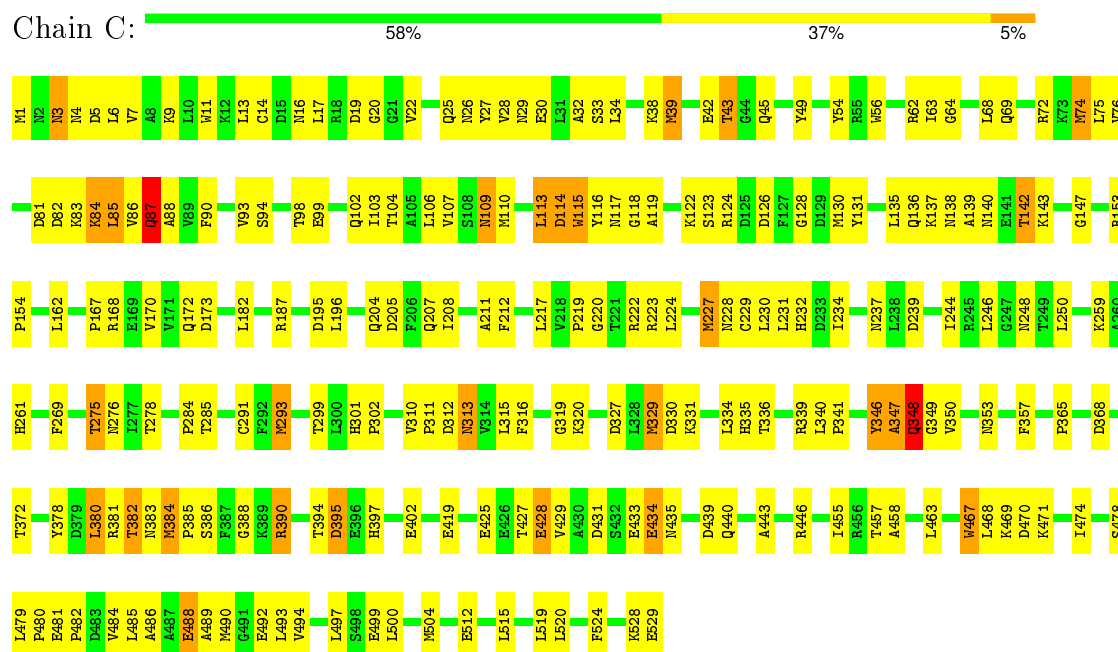


• Molecule 2: TYPE I RESTRICTION ENZYME ECOKI M PROTEIN

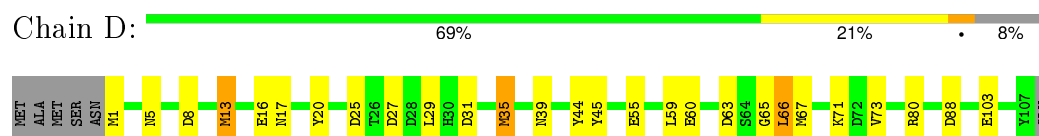




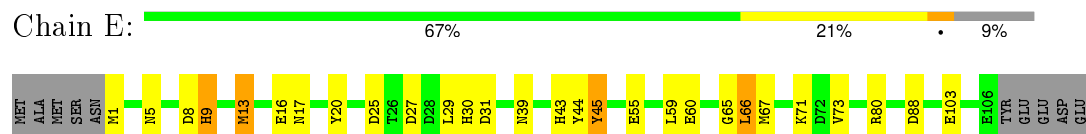
• Molecule 2: TYPE I RESTRICTION ENZYME ECOKI M PROTEIN



• Molecule 3: GENE 0.3 PROTEIN



• Molecule 3: GENE 0.3 PROTEIN



4 Experimental information ⓘ

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of images	Not provided	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	FILTERED AT FIRST ZERO	Depositor
Microscope	OTHER	Depositor
Voltage (kV)	80	Depositor
Electron dose ($e^-/\text{\AA}^2$)	25	Depositor
Minimum defocus (nm)	275	Depositor
Maximum defocus (nm)	870	Depositor
Magnification	40000	Depositor
Image detector	KODAK SO163 FILM	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 > 2$	RMSZ	# $ Z > 2$
1	A	0.94	0/3685	1.04	1/4968 (0.0%)
2	B	0.76	11/4262 (0.3%)	0.77	6/5773 (0.1%)
2	C	0.77	11/4262 (0.3%)	0.76	5/5773 (0.1%)
3	D	1.36	4/894 (0.4%)	1.22	7/1215 (0.6%)
3	E	1.36	4/893 (0.4%)	1.24	8/1213 (0.7%)
All	All	0.91	30/13996 (0.2%)	0.91	27/18942 (0.1%)

The worst 5 of 30 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	467	TRP	C-N	21.00	1.82	1.34
2	B	467	TRP	C-N	19.85	1.79	1.34
2	B	115	TRP	C-N	15.35	1.69	1.34
2	C	115	TRP	C-N	15.29	1.69	1.34
2	B	293	MET	CG-SD	8.37	2.02	1.81

The worst 5 of 27 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	467	TRP	O-C-N	-11.62	104.11	122.70
2	C	467	TRP	O-C-N	-11.29	104.63	122.70
3	D	29	LEU	CB-CG-CD2	9.66	127.42	111.00
3	E	29	LEU	CB-CG-CD2	9.63	127.37	111.00
2	B	467	TRP	CA-C-N	7.08	132.79	117.20

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	A	3622	0	3739	800	0
2	B	4175	0	4048	321	0
2	C	4175	0	4049	333	0
3	D	877	0	803	21	0
3	E	876	0	803	32	0
All	All	13725	0	13442	1480	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 54.

The worst 5 of 1480 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:C:293:MET:SD	2:C:293:MET:CG	2.02	1.45
2:B:293:MET:SD	2:B:293:MET:CG	2.02	1.45
2:B:115:TRP:C	2:B:116:TYR:N	1.69	1.44
2:C:115:TRP:C	2:C:116:TYR:N	1.69	1.43
2:C:383:ASN:HB3	2:C:470:ASP:CB	1.45	1.43

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	462/464 (100%)	431 (93%)	20 (4%)	11 (2%)	7	47
2	B	527/529 (100%)	491 (93%)	32 (6%)	4 (1%)	24	69
2	C	527/529 (100%)	493 (94%)	30 (6%)	4 (1%)	24	69
3	D	105/116 (90%)	105 (100%)	0	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	E	104/116 (90%)	104 (100%)	0	0	100	100
All	All	1725/1754 (98%)	1624 (94%)	82 (5%)	19 (1%)	23	63

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	89	VAL
1	A	139	SER
1	A	296	SER
2	B	84	LYS
2	B	347	ALA

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	398/398 (100%)	384 (96%)	14 (4%)	43	74
2	B	452/452 (100%)	418 (92%)	34 (8%)	17	53
2	C	452/452 (100%)	421 (93%)	31 (7%)	19	56
3	D	95/104 (91%)	90 (95%)	5 (5%)	28	64
3	E	95/104 (91%)	90 (95%)	5 (5%)	28	64
All	All	1492/1510 (99%)	1403 (94%)	89 (6%)	28	60

5 of 89 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	395	ASP
2	C	3	ASN
3	D	66	LEU
2	B	396	GLU
2	B	434	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 74 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	29	ASN
2	B	240	HIS
2	C	383	ASN
2	B	65	GLN
2	B	140	ASN

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.