



wwPDB/EMDatabank EM Map/Model Validation Summary Report ⓘ

Jan 5, 2017 – 01:11 AM EST

PDB ID : 5MBV
EMDB ID: : EMD-3460
Title : Cryo-EM structure of Lambda Phage protein GamS bound to RecBCD
Authors : Wilkinson, M.; Chaban, Y.; Wigley, D.B.
Deposited on : 2016-11-08
Resolution : 3.80 Å(reported)
Based on PDB ID : 5LD2, 2UUZ

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

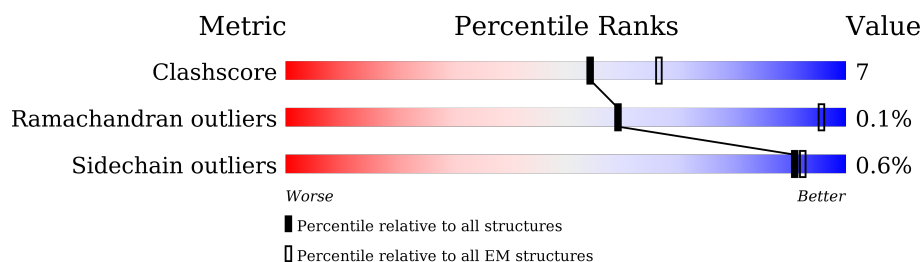
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.80 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	B	1181	79% 16% 5%
2	C	1122	84% 16%
3	D	609	76% 22% .
4	A	98	89% 11%
4	E	98	70% 6% 23%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 24118 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 RecBCD enzyme subunit RecB.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	B	1127	Total	C	N	O	S	0	0
			9037	5704	1601	1694	38		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	GLY	-	expression tag	UNP P08394

- Molecule 2 is a protein called RecBCD enzyme subunit RecC.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	1121	Total	C	N	O	S	0	0
			9078	5783	1568	1684	43		

- Molecule 3 is a protein called RecBCD enzyme subunit RecD.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	D	593	Total	C	N	O	S	0	0
			4570	2854	840	856	20		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	0	MET	-	initiating methionine	UNP P04993
D	1	GLY	-	expression tag	UNP P04993

- Molecule 4 is a protein called Host-nuclease inhibitor protein gam.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	A	98	Total	C	N	O	S	0	0
			817	505	147	160	5		
4	E	75	Total	C	N	O	S	0	0
			616	379	111	122	4		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	79	ILE	LEU	engineered mutation	UNP P03702
E	79	ILE	LEU	engineered mutation	UNP P03702

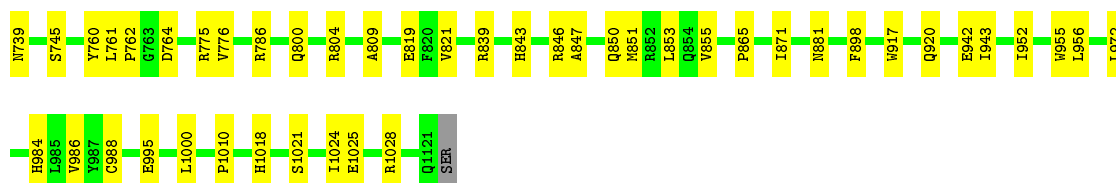
i

- Molecule 1: RecBCD enzyme subunit RecB

S1091	Q944	T831	R434	GLY
A1092	F945	D832	W652	MET
T1093	P946	P933	Y441	SER
T1094	R947	S836	G656	R254
	F964	S836	V657	R255
R1106		G839	W658	K256
Y1111	E978	R840	W660	
L1112	F982	L841		L8
L1113		L842		L11
L1116	W986	Q948	W665	R12
I1126	E992	L860	R668	L13
W993	W993	L860	W669	G17
F1135	I994	A870	W670	E18
V1138	T995	A870	L675	T27
Y1140	V997	W883		F289
	L1002	N889	R682	SER
R1144	K1017	Q894	R683	GLN
G1145	Q1018	R895	L688	ARG
V1146	W1019	W901	L694	PHE
P1160	E1020	R902	L711	LEU
M1168	P1026	Q911	N721	GLU
	I1027	ARG	A722	ASP
G1174	I1032	HIS	Q725	ARG
MET	Q1035	GLY	Q726	THR
LEU		ILE	W727	LYS
GLU	R1041	ALA	R728	ALA
GLU		GLN	L729	ALA
ALA	S1047	ASP	E730	GLY
		LEU	S731	GLY
	P1051	MET	L747	T304
		PRO		
		ARG	L321	R307
	R1059	LEU	W552	P308
	G1060	ASP	W753	L309
	W1061	VAL	L754	F310
	L1062	ASP	F755	E123
		ALA	F756	R135
	I1066	ALA	L757	L321
	D1067	GLY		V380
	L1068	VAL	R761	A381
		ALA		M382
	L1078	SER	R794	R398
	L1079	VAL	L795	H402
	D1080	VAL		K188
	Y1081	GLU	I628	G199
		GLU	R800	P190
		GLU		D414
		PRO	W832	F415
		THR	S809	K416
	N1084	LEU	W638	I427
	W1085	VAL	S814	L409
		GLY	R843	G199
		VAL	L815	P190
		VAL		D414
		VAL		F415
		VAL		K416
		VAL		I427
		VAL		L409
		VAL		G199
		VAL		P190
		VAL		D414
		VAL		F415
		VAL		K416
		VAL		I427
		VAL		L409
		VAL		G199
		VAL		P190
		VAL		D414
		VAL		F415
		VAL		K416
		VAL		I427
		VAL		L409
		VAL		G199
		VAL		P190
		VAL		D414
		VAL		F415
		VAL		K416
		VAL		I427
		VAL		L409
		VAL		G199
		VAL		P190
		VAL		D414
		VAL		F415
		VAL		K416
		VAL		I427
		VAL		L409
		VAL		G199
		VAL		P190
		VAL		D414
		VAL		F415
		VAL		K416
		VAL		I427
		VAL		L409
		VAL		G199
		VAL		P190
		VAL		D414
		VAL		F415
		VAL		K416
		VAL		I427
		VAL		L409
		VAL		G199
		VAL		P190
		VAL		D414
		VAL		F415
		VAL		K416
		VAL		I427
		VAL		L409
		VAL		G199
		VAL		P190
		VAL		D414
		VAL		F415
		VAL		K416
		VAL		I427
		VAL		L409
		VAL		G199
		VAL		P190
		VAL		D414
		VAL		F415
		VAL		K416
		VAL		

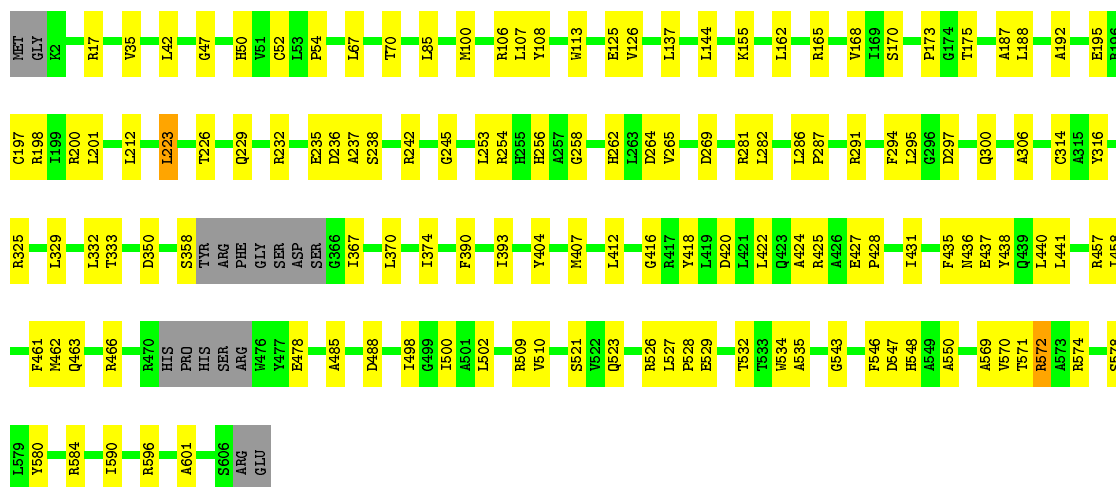
Chain C: 84% 16%

R582	P430	L239	#1
D602	R433	F240	L13
A603	Y434	T241	V21
R604		N242	R25
T605	S440	F243	
		C244	I35
W616	P449	R245	L36
A618	V450		V37
I619	L451	G249	Q38
	L457	D250	
E640		I251	A43
L641	L460	Y256	
L645	F465		L77
	V466	S268	
F654	S467	F269	T80
L655		E270	
A656	L474		Q89
G657	D475	L276	
	V476		N96
M660	P477	S280	T97
L661	V478		
C662		G291	P100
	V497	E292	
P666			T109
M667	I502	R308	L110
	R503		L111
		F326	
W674	I506		Y114
V675	D507	V329	L115
C676	D508		
L677		D332	L124
	R512	F125	F126
M680		N354	L127
M681	N524	E357	
P686	R530		V155
R687		N363	
Q688	L533	K364	Q162
	G534		
K702		L367	H185
R703	S539	H377	R186
G704			A187
D705	W544		R188
	Q545	V386	L189
R708			
		R392	T196
D712	L548		
	P549	T404	P204
L715	V550	P405	P205
		R406	
Y728	L567	D407	R210
I729	L568	I408	V211
S730		I409	F212
R731	L571	V410	T213
I732	N572		C214
Q737	G577	V425	
			P231



• Molecule 3: RecBCD enzyme subunit RecD

Chain D: 76% 22%



• Molecule 4: Host-nuclease inhibitor protein gam

Chain A: 89% 11%



• Molecule 4: Host-nuclease inhibitor protein gam

Chain E: 70% 6% 23%



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	122796	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	Not provided	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	Not provided	Depositor
Minimum defocus (nm)	1200	Depositor
Maximum defocus (nm)	2400	Depositor
Magnification	37313	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	B	0.23	0/9231	0.39	1/12521 (0.0%)
2	C	0.23	0/9305	0.39	0/12644
3	D	0.24	0/4643	0.41	0/6287
4	A	0.22	0/833	0.32	0/1120
4	E	0.22	0/626	0.33	0/839
All	All	0.23	0/24638	0.39	1/33411 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	1080	ASP	CB-CG-OD2	5.14	122.93	118.30

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	B	9037	0	8887	119	0
2	C	9078	0	8877	115	0
3	D	4570	0	4614	85	0
4	A	817	0	764	11	0
4	E	616	0	583	6	0
All	All	24118	0	23725	315	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 7.

The worst 5 of 315 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:37:VAL:HG21	2:C:43:ALA:HB2	1.62	0.81
2:C:503:ARG:HE	2:C:865:PRO:HG2	1.48	0.77
1:B:252:ILE:HG23	1:B:254:ARG:H	1.52	0.75
3:D:144:LEU:HD22	3:D:223:LEU:HD11	1.70	0.72
2:C:539:SER:HA	2:C:549:PRO:HG2	1.72	0.72

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	B	1121/1181 (95%)	1070 (96%)	49 (4%)	2 (0%)	52	86
2	C	1119/1122 (100%)	1054 (94%)	65 (6%)	0	100	100
3	D	587/609 (96%)	561 (96%)	25 (4%)	1 (0%)	52	86
4	A	96/98 (98%)	94 (98%)	2 (2%)	0	100	100
4	E	73/98 (74%)	72 (99%)	1 (1%)	0	100	100
All	All	2996/3108 (96%)	2851 (95%)	142 (5%)	3 (0%)	59	90

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	D	126	VAL
1	B	380	VAL
1	B	1091	SER

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	B	957/999 (96%)	950 (99%)	7 (1%)	88	95
2	C	976/977 (100%)	973 (100%)	3 (0%)	94	98
3	D	478/492 (97%)	473 (99%)	5 (1%)	82	92
4	A	86/86 (100%)	86 (100%)	0	100	100
4	E	66/86 (77%)	66 (100%)	0	100	100
All	All	2563/2640 (97%)	2548 (99%)	15 (1%)	91	96

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	1094	THR
2	C	109	THR
3	D	226	THR
1	B	947	ARG
3	D	223	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 32 such sidechains are listed below:

Mol	Chain	Res	Type
2	C	572	ASN
2	C	739	ASN
4	A	131	GLN
2	C	660	ASN
2	C	817	HIS

5.3.3 RNA [i](#)

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.