



wwPDB EM Map/Model Validation Report

Apr 10, 2016 – 02:11 PM BST

PDB ID : 5FJ7
EMDB ID: : EMD-3187
Title : Structure of the P2 polymerase inside in vitro assembled bacteriophage phi6 polymerase complex, with P1 included
Authors : Ilca, S.; Kotecha, A.; Sun, X.; Poranen, M.P.; Stuart, D.I.; Huiskonen, J.T.
Deposited on : 2015-10-06
Resolution : 7.90 Å(reported)
Based on PDB ID : 4K7H,1HHS

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

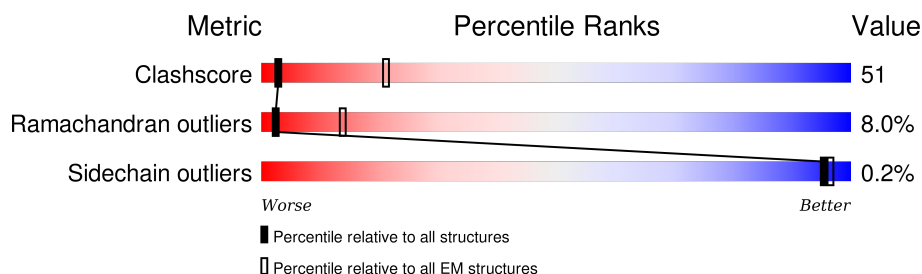
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 7.90 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	761	
1	B	761	
2	C	664	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 17106 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 MAJOR INNER PROTEIN P1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	761	Total	C	N	O	S	0	0
			5920	3741	1048	1109	22		
1	B	761	Total	C	N	O	S	0	0
			5920	3741	1048	1109	22		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	EXPRESSION TAG	UNP P11126
B	1	GLY	-	EXPRESSION TAG	UNP P11126

- Molecule 2 is a protein called RNA-DIRECTED RNA POLYMERASE.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	664	Total	C	N	O	S	0	0
			5265	3342	914	977	32		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	456	MET	ILE	CONFLICT	UNP P11124

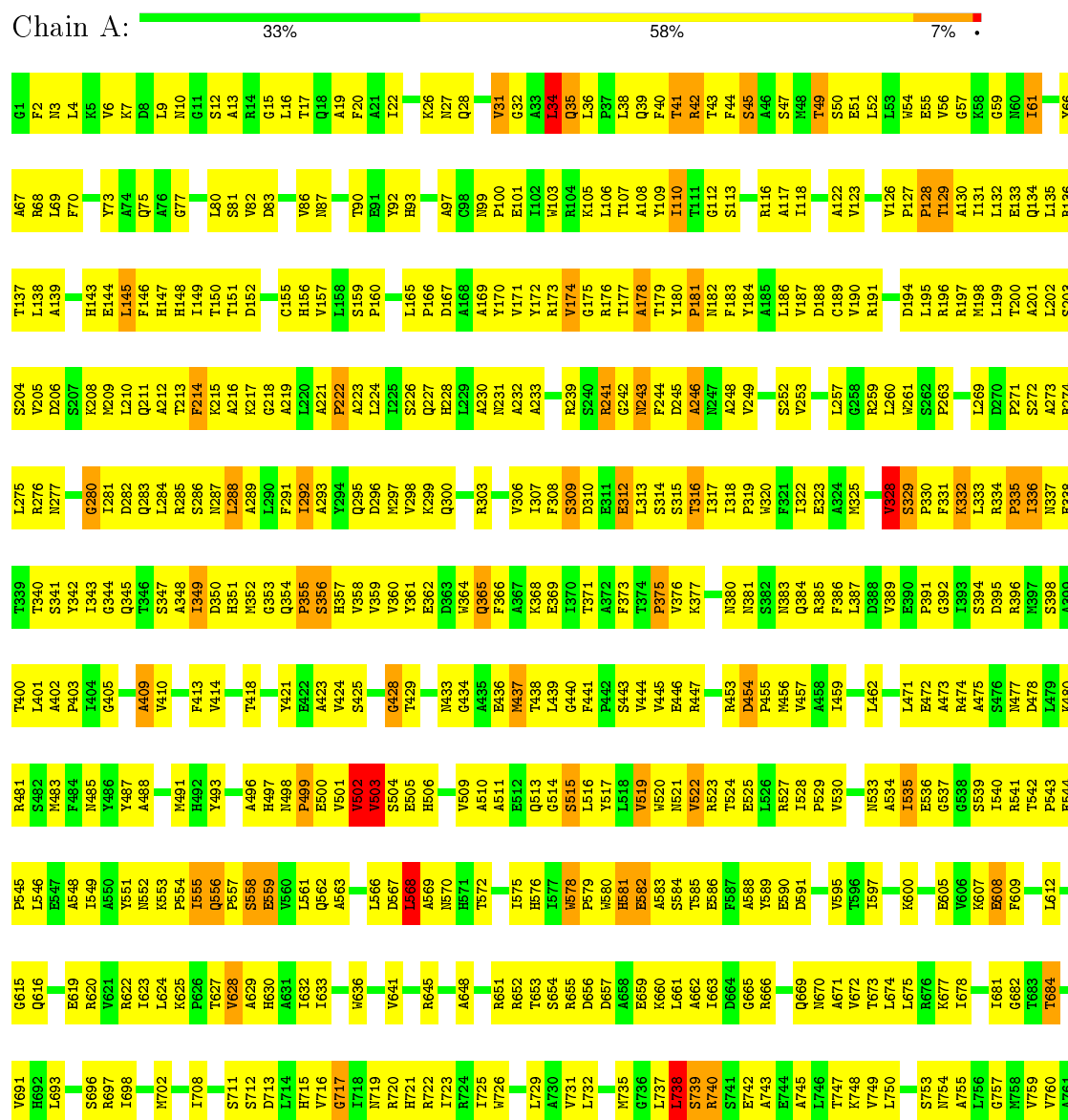
- Molecule 3 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		AltConf
3	C	1	Total	Mn	0
			1	1	

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: MAJOR INNER PROTEIN P1



• Molecule 1: MAJOR INNER PROTEIN P1



L693	R620	L546	S482	L401	T340	N277	K208	E144	F70	G1
L694	R621	E547	N483	A402	S341	T278	M209	L145	F71	F2
L695	R622	E548	F484	P403	Y342	T279	L210	L146	Y73	N3
L696	R623	I549	N485	I404	I343	G280	Q211	H147	A74	L4
L697	R624	E550	Y486	G405	G344	I281	A212	H148	Q75	V5
L698	K625	Y551	Y487	A409	Q345	D282	T213	I149	A76	V6
L702	R626	N552	A488	A409	T346	Q283	F214	T150	G77	
L703	R627	K553	P554	V410	S347	L284	K215	T151		L9
L708	R628	P554	N491	F413	I349	R285	A216	D152	L80	N10
L711	R629	E555	Y492	V414	D350	I287	Q218	C155	S81	G11
L712	R630	Q556	Y493		R351	L288	A219	H156	V82	S12
L713	R631	P557		T418	H351	L289		V157	D83	A13
L714	R632	S558	A496		G352	L290	A220	L158		G15
L715	R633	E559	N497	Y421	G353	F291	P222	S159	H66	L16
L716	R634	E560	P499	E422	R354	L292	A223	P160		T17
L717	R635	Q562	E501	A423	H357	L293	L224	Y92		G18
L718	R636	A563	V501	V424		L294	L225	H93		A19
L719	R637	E566	V502	S425		Q295	S226	L165		F20
L720	R638	D567	V503	G434		R358	Q227	P166	A97	
L721	R639	E568	S504	G435		D296	P227	C98		A21
L722	R640	E569	E505	T429		N297	E228	H99		I22
L723	R641	N570	H506		Y361	T298	L229	P100		
L724	R642	E571		N433	E362	K299	A230	E101		K26
L725	R643	T572		G434		Q300	N231	V171		
L726	R644			A435			A232	Y172		V31
L727	R645	E575	A510	R365			A233	H173		G32
L728	R646	H576	A511	E436				V174		A23
L729	R647	E577	E512	M437		Y306	P237	G175		
L730	R648	L578	O513	T438		I307	E238	H176		L34
L731	R649	M579	G514	L439		F308	E239	T177		Q35
L732	R650	P579	S515	G440		S309	S240	A178		L36
L733	R651	E580	L516	F441		D310	Q241	T179		P37
L734	R652	H581	V517	P442		E311	G242	Y180		Q39
L735	R653	E582	E518	S443		L313	Q243	N242		L41
L736	R654	E583	V519	V444		T374	P244	N322		T41
L737	R655	S584	H520	V445		S314	D245	F133		R42
L738	R656	E585	N521	E446		S315	A246	Y194		T43
L739	R657	E586	V522	R447		T316	N247	A185		F44
L740	R658	F587	R523		K377	I317	A248	L186		S45
L741	R659	A588	T524	R453		L318	V249	V187		A46
L742	R660	Y589	E525	D454		P319		D188		S47
L743	R661	E590	V526	P455		K320	S252	C189		A48
L744	R662	D591	H527	N456		F321	V253	V190		T49
L745	R663	E592	L528	Q384		L322	L257	R191		S50
L746	R664	V595	P529	R385		E323	G258	A192		E51
L747	R665	E596	V530	F386		A324	R259	P128		L52
L748	R666	L597		L387		K325	L260	D194		L53
L749	R667		N533	D388			L261	L195		K54
L750	R668		A534	V389		V328	G261	R196		E55
L751	R669	K600	E535	E390		S329	S262	L197		S56
L752	R670	E605	E536	P391		P330	P263	L198		G57
L753	R671	V606	G537	G392		F331		L199		K58
L754	R672	K607	E538	R474		K332	L269	T200		G59
L755	R673	E608	V539	A475		L333	D270	A201		N60
L756	R674	F609	L540	S476		R334	S271	L202		L61
L757	R675	E610	H541	N477		P335	S272	S203		
L758	R676	L611	T542	D478		L336	A273	S204		Y66
L759	R677	L612	P543	L479		R337	R274	V205		A67
L760	R678		E544	K398		E338	L275	D206		R68
L761	R679		E545	A480		T339	E276	S207		L40
L762	R680	E613	E546	R481		F340		H432		

• Molecule 2: RNA-DIRECTED RNA POLYMERASE

Chain C:  23%  74%

P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	P25	P26	P27	P28	P29	P30	P31	P32	P33	P34	P35	P36	P37	P38	P39	P40	P41	P42	P43	P44	P45	P46	P47	P48	P49	P50	P51	P52	P53	P54	P55	P56	P57	P58	P59	P60	P61	P62	P63	P64	P65	P66	P67	P68	P69	P70	P71	P72	P73	P74	P75	P76	P77	P78	P79	P80	P81	P82	P83	P84	P85	P86	P87	P88	P89	P90	P91	P92	P93	P94	P95	P96	P97	P98	P99	1000
----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------

K631	L502	Y439	L379	S318
W632	Y503	W440	G380	L319
E633	D604	Q441	D381	C320
E634	S505	G442	P382	V321
	R506	R443	S383	A322
V637	R507	E444	N384	T323
S638	E508	E445	P385	D324
A639	P509	I446	D386	V325
N640	G510	R447	L387	S326
I641	S511	Q448	E388	D327
H642	A512	I449	V389	H328
E643	I513	S450	G390	D329
V644	F514	K451	L391	T330
L645	V515	S452	S392	F331
M646	G516	D453	S393	W332
H647	N517	D454	G394	P333
G648	I518	A455	Q395	G334
V649	N519	N456	G396	V335
S650	S520	L457	A397	L336
V651	M521	G458	T398	R337
E652	L522	W459	D399	D338
	N523	T460	L400	L339
E655	N524	K461	M401	T340
R656	Q525	G462	G402	C341
F657	F526	R463	T403	D342
L658	S527	A464	L404	E343
R659	P528	L465	M405	L344
S660	E529	Y466	M406	L345
V661	S534		S407	N346
M662	W535	H469	I408	N347
P663	V536	R470	T409	G348
R664	R537	L471	Y410	V349
	D538	F472	L411	A350
	R539		V412	P351
	S540	K476	M413	W352
	K541	E477	Q414	W353
	R542	G478	L415	V354
	R543	K479	D416	T355
W605	R544	Y480	H417	L356
A606	P545	N481	T418	F357
R607	F546	S482	A419	E358
	P547	P483	P420	T359
B613	Q548	Y484	H421	S360
L614	L549	W485	L422	L361
P616	A550	K487	M423	K362
I617	W551	T488	S424	L363
D618	A552	S489	R425	P364
L619	S553	Y490	I426	V365
B620	M554	E491	K427	Y366
V621	K555	H492	D428	V367
L622	D556	G493	M429	G368
A623	T557	G494	P430	A369
D624	Y558	A495	S431	P370
P625	G559	F496	A432	A371
V626	A560	L497	C433	
K627	C561	G498	R434	Q374
L628	P562	D499	F435	G375
D629	I563	Y500	L436	H376
Y630	Y564	L501	D437	T377
			S438	L378

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	EACH PARTICLE, Not provided	Depositor
Microscope	FEI TECNAI F30	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	0.16	Depositor
Minimum defocus (nm)	1100	Depositor
Maximum defocus (nm)	2600	Depositor
Magnification	160000	Depositor
Image detector	GATAN K2 (4K X 4K)	Depositor

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MN

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	0.43	0/6040	0.70	7/8206 (0.1%)
1	B	0.43	0/6040	0.70	7/8206 (0.1%)
2	C	1.17	7/5396 (0.1%)	0.66	2/7297 (0.0%)
All	All	0.74	7/17476 (0.0%)	0.69	16/23709 (0.1%)

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	5
1	B	0	5
2	C	0	1
All	All	0	11

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	361	LEU	CG-CD1	52.88	3.47	1.51
2	C	357	PHE	CD1-CE1	26.88	1.93	1.39
2	C	357	PHE	CD2-CE2	26.52	1.92	1.39
2	C	357	PHE	CE2-CZ	25.99	1.86	1.37
2	C	357	PHE	CE1-CZ	25.62	1.86	1.37

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	361	LEU	CB-CG-CD1	11.31	130.23	111.00
1	B	128	PRO	CA-C-N	6.76	132.06	117.20

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	128	PRO	CA-C-N	6.75	132.05	117.20
1	A	128	PRO	C-N-CA	6.09	136.93	121.70
1	B	128	PRO	C-N-CA	6.09	136.93	121.70

There are no chirality outliers.

5 of 11 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	328	VAL	Peptide
1	A	34	LEU	Peptide
1	A	35	GLN	Peptide
1	A	365	GLN	Peptide
1	A	437	MET	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	5920	0	5913	504	0
1	B	5920	0	5907	641	0
2	C	5265	0	5154	718	0
3	C	1	0	0	0	0
All	All	17106	0	16974	1735	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 51.

The worst 5 of 1735 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:357:PHE:CZ	2:C:357:PHE:CE1	1.86	1.58
2:C:357:PHE:CE2	2:C:357:PHE:CD2	1.92	1.58
2:C:357:PHE:CE2	2:C:357:PHE:CZ	1.86	1.56
2:C:357:PHE:CD1	2:C:357:PHE:CE1	1.93	1.53
1:B:237:PHE:O	2:C:425:ARG:NH2	1.58	1.36

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	759/761 (100%)	594 (78%)	91 (12%)	74 (10%)	1	14
1	B	759/761 (100%)	594 (78%)	91 (12%)	74 (10%)	1	14
2	C	662/664 (100%)	575 (87%)	61 (9%)	26 (4%)	4	36
All	All	2180/2186 (100%)	1763 (81%)	243 (11%)	174 (8%)	2	19

5 of 174 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	31	VAL
1	A	41	THR
1	A	129	THR
1	A	145	LEU
1	A	174	VAL

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	629/629 (100%)	627 (100%)	2 (0%)	94	96
1	B	629/629 (100%)	627 (100%)	2 (0%)	94	96
2	C	557/557 (100%)	557 (100%)	0	100	100
All	All	1815/1815 (100%)	1811 (100%)	4 (0%)	95	97

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	503	VAL
1	A	519	VAL
1	B	503	VAL
1	B	519	VAL

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	B	345	GLN
1	B	670	ASN
2	C	448	GLN
1	B	485	ASN
1	B	492	HIS

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 ⓘ

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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.