wwPDB EM Map/Model Validation Report (i)

Apr 10, 2016 – 01:49 PM BST

PDB ID	:	3J5M
EMDB ID:	:	EMD-5779
Title	:	Cryo-EM structure of the BG505 SOSIP.664 HIV-1 Env trimer with 3 PGV04 Fabs
Authors	:	Lyumkis, D.; Julien, J.-P.; Wilson, I.A.; Ward, A.B.
Deposited on	:	2013-10-26
Resolution	:	5.80 Å(reported)
Based on PDB ID	:	3SE9, 3U2S, 1ENV, 2B4C

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

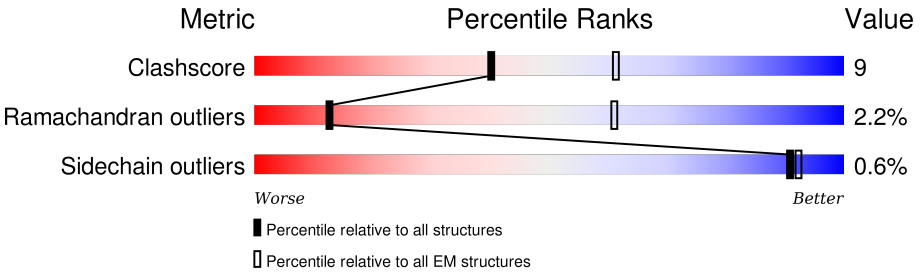
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk27241

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 5.80 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 <=5%

Mol	Chain	Length	Quality of chain
1	A	475	70% 14% • 13%
1	E	475	69% 15% • 13%
1	I	475	70% 13% • 13%
2	B	64	98% •
2	F	64	98% •
2	J	64	98% •
3	C	208	95% •
3	G	208	96% •
3	K	208	96% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	D	228	
4	H	228	
4	L	228	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 20394 atoms, of which 15 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 BG505 SOSIP gp120.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	415	Total	C	N	O	S	0	0
			3140	1964	554	595	27		
1	E	415	Total	C	N	O	S	0	0
			3140	1964	554	595	27		
1	I	415	Total	C	N	O	S	0	0
			3140	1964	554	595	27		

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	332	ASN	THR	ENGINEERED MUTATION	UNP Q2N0S6
A	501	CYS	ALA	ENGINEERED MUTATION	UNP Q2N0S6
E	332	ASN	THR	ENGINEERED MUTATION	UNP Q2N0S6
E	501	CYS	ALA	ENGINEERED MUTATION	UNP Q2N0S6
I	332	ASN	THR	ENGINEERED MUTATION	UNP Q2N0S6
I	501	CYS	ALA	ENGINEERED MUTATION	UNP Q2N0S6

- Molecule 2 is a protein called BG505 SOSIP gp41.

Mol	Chain	Residues	Atoms				AltConf	Trace
2	B	63	Total	C	N	O	0	0
			315	189	63	63		
2	F	63	Total	C	N	O	0	0
			315	189	63	63		
2	J	63	Total	C	N	O	0	0
			315	189	63	63		

- Molecule 3 is a protein called PGV04 light chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	208	Total	C	N	O	S	0	0
			1621	1018	275	321	7		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf	Trace
3	G	208	Total	C	N	O	S	0	0
			1621	1018	275	321	7		
3	K	208	Total	C	N	O	S	0	0
			1621	1018	275	321	7		

- Molecule 4 is a protein called PGV04 heavy chain.

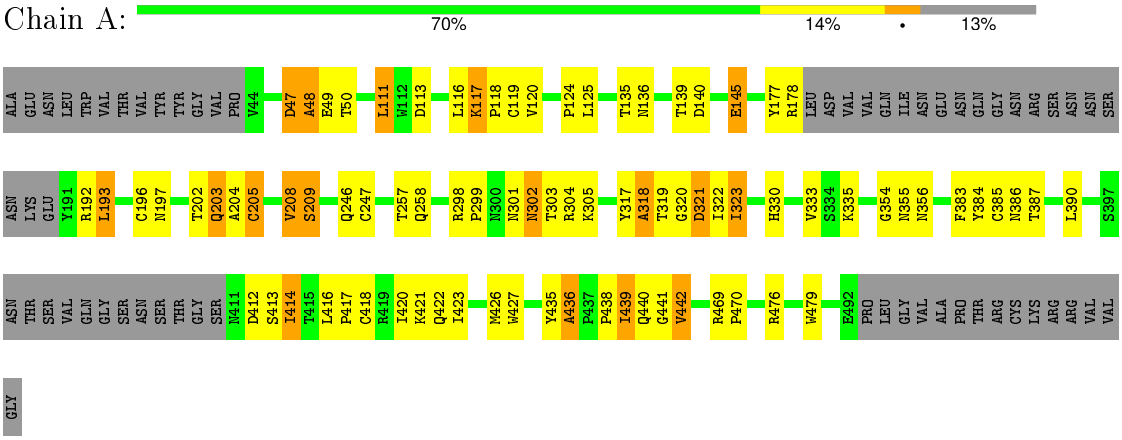
Mol	Chain	Residues	Atoms						AltConf	Trace
4	D	225	Total	C	H	N	O	S	0	0
			1722	1090	5	298	324	5		
4	H	225	Total	C	H	N	O	S	0	0
			1722	1090	5	298	324	5		
4	L	225	Total	C	H	N	O	S	0	0
			1722	1090	5	298	324	5		

3 Residue-property plots (i)

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.

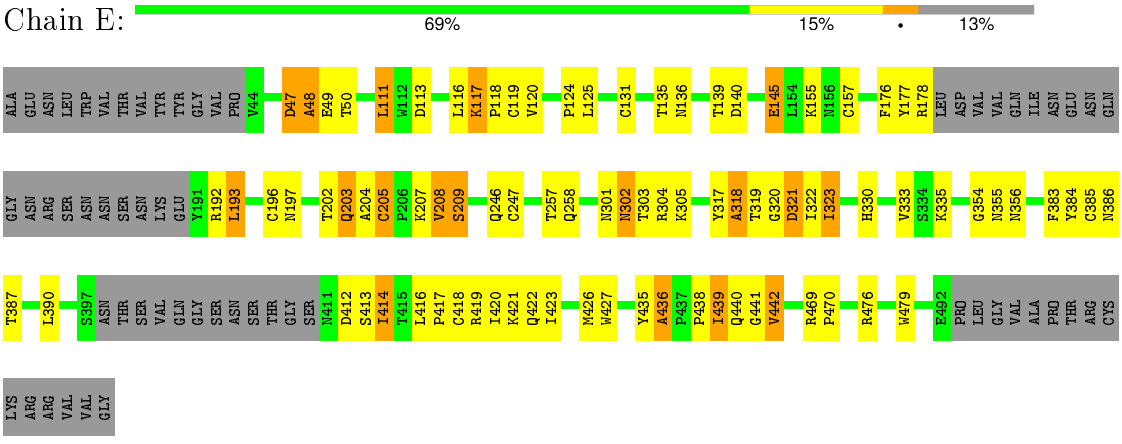
- Molecule 1: BG505 SOSIP gp120

Chain A:



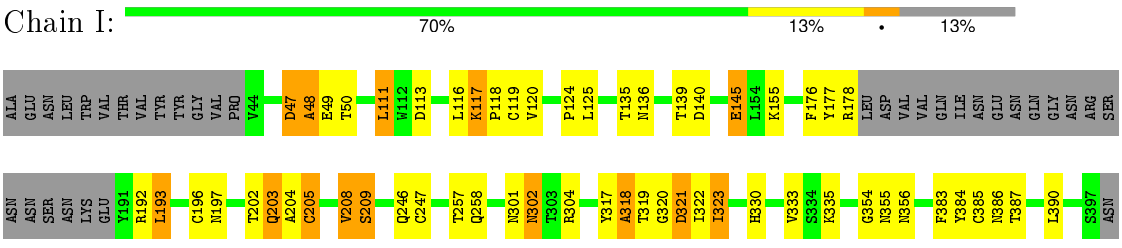
- Molecule 1: BG505 SOSIP gp120

Chain E:



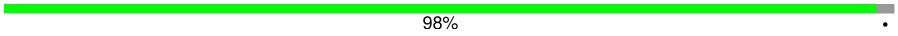
- Molecule 1: BG505 SOSIP gp120

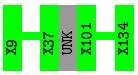
Chain I:



- Molecule 2: BG505 SOSIP gp41

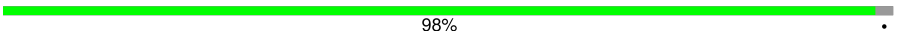
Chain B:





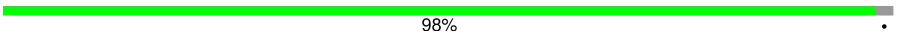
- Molecule 2: BG505 SOSIP gp41

Chain F:



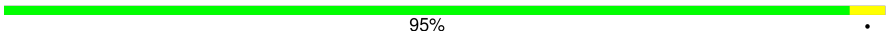
- Molecule 2: BG505 SOSIP gp41

Chain J:



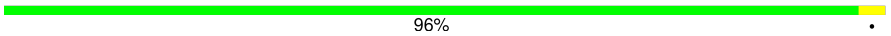
- Molecule 3: PGV04 light chain

Chain C:



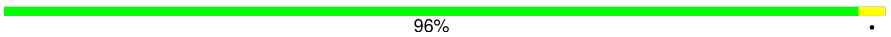
- Molecule 3: PGV04 light chain

Chain G:



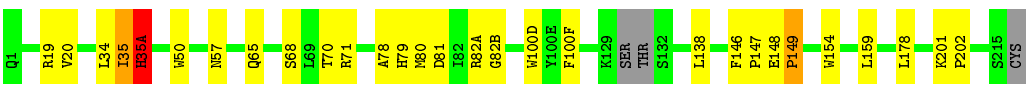
- Molecule 3: PGV04 light chain

Chain K:



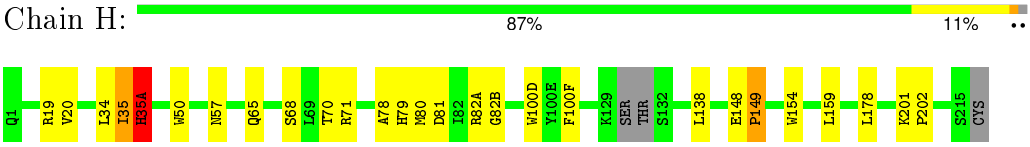
- Molecule 4: PGV04 heavy chain

Chain D:



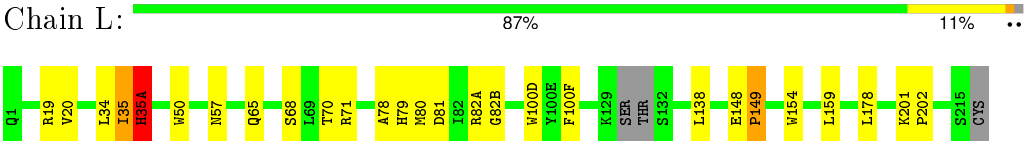
• Molecule 4: PGV04 heavy chain

Chain H:



• Molecule 4: PGV04 heavy chain

Chain L:



4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of images	49572	Depositor
Resolution determination method	FSC at 0.143 cut-off	Depositor
CTF correction method	Frealign	Depositor
Microscope	FEI TECNAI F20	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/Å^2$)	32	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	5000	Depositor
Magnification	29000	Depositor
Image detector	Gatan K2 Summit	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	0.59	0/3206	0.57	1/4358 (0.0%)
1	E	0.60	0/3206	0.57	1/4358 (0.0%)
1	I	0.60	1/3206 (0.0%)	0.57	1/4358 (0.0%)
3	C	0.67	0/1658	0.56	1/2244 (0.0%)
3	G	0.67	0/1658	0.56	1/2244 (0.0%)
3	K	0.67	0/1658	0.56	1/2244 (0.0%)
4	D	0.58	2/1761 (0.1%)	0.57	2/2396 (0.1%)
4	H	0.58	2/1761 (0.1%)	0.57	2/2396 (0.1%)
4	L	0.58	2/1761 (0.1%)	0.57	2/2396 (0.1%)
All	All	0.61	7/19875 (0.0%)	0.57	12/26994 (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
1	A	0	2
1	E	0	2
1	I	0	2
3	C	0	1
3	G	0	1
3	K	0	1
4	D	0	1
4	H	0	1
4	L	0	1
All	All	0	12

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	L	82(B)	GLY	C-N	5.29	1.46	1.34
4	D	82(B)	GLY	C-N	5.27	1.46	1.34

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	H	82(B)	GLY	C-N	5.27	1.46	1.34
4	D	149	PRO	N-CD	5.11	1.55	1.47
1	I	422	GLN	C-N	5.09	1.45	1.34

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	145	GLU	O-C-N	-12.75	102.31	122.70
1	A	145	GLU	O-C-N	-12.73	102.32	122.70
1	I	145	GLU	O-C-N	-12.73	102.33	122.70
3	G	91	LEU	O-C-N	-8.36	109.33	122.70
3	C	91	LEU	O-C-N	-8.34	109.36	122.70

There are no chirality outliers.

5 of 12 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	145	GLU	Mainchain
1	A	47	ASP	Peptide
3	C	91	LEU	Mainchain
4	D	35(A)	HIS	Mainchain
1	E	47	ASP	Peptide

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	A	3140	0	2943	90	0
1	E	3140	0	2943	93	0
1	I	3140	0	2943	90	0
2	B	315	0	67	0	0
2	F	315	0	67	0	0
2	J	315	0	67	0	0
3	C	1621	0	1579	6	0
3	G	1621	0	1579	5	0
3	K	1621	0	1579	5	0
4	D	1717	5	1690	24	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	H	1717	5	1690	25	0
4	L	1717	5	1690	23	0
All	All	20379	15	18837	361	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 9.

The worst 5 of 361 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:384:TYR:CE1	1:A:421:LYS:HG2	1.30	1.66
1:E:384:TYR:CE1	1:E:421:LYS:HG2	1.30	1.63
1:I:384:TYR:CE1	1:I:421:LYS:HG2	1.30	1.60
1:E:384:TYR:CE1	1:E:421:LYS:CG	1.89	1.55
1:I:384:TYR:CE1	1:I:421:LYS:CG	1.89	1.54

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	409/475 (86%)	359 (88%)	33 (8%)	17 (4%)	3 34
1	E	409/475 (86%)	359 (88%)	33 (8%)	17 (4%)	3 34
1	I	409/475 (86%)	359 (88%)	33 (8%)	17 (4%)	3 34
3	C	206/208 (99%)	198 (96%)	7 (3%)	1 (0%)	34 77
3	G	206/208 (99%)	198 (96%)	7 (3%)	1 (0%)	34 77
3	K	206/208 (99%)	198 (96%)	7 (3%)	1 (0%)	34 77
4	D	221/228 (97%)	212 (96%)	9 (4%)	0	100 100
4	H	221/228 (97%)	212 (96%)	9 (4%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
4	L	221/228 (97%)	212 (96%)	9 (4%)	0	100 100
All	All	2508/2733 (92%)	2307 (92%)	147 (6%)	54 (2%)	13 49

5 of 54 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	193	LEU
1	A	302	ASN
1	A	318	ALA
1	A	321	ASP
1	A	323	ILE

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	333/422 (79%)	330 (99%)	3 (1%)	84 93
1	E	333/422 (79%)	330 (99%)	3 (1%)	84 93
1	I	333/422 (79%)	330 (99%)	3 (1%)	84 93
3	C	182/182 (100%)	181 (100%)	1 (0%)	92 96
3	G	182/182 (100%)	181 (100%)	1 (0%)	92 96
3	K	182/182 (100%)	181 (100%)	1 (0%)	92 96
4	D	190/193 (98%)	190 (100%)	0	100 100
4	H	190/193 (98%)	190 (100%)	0	100 100
4	L	190/193 (98%)	190 (100%)	0	100 100
All	All	2115/2391 (88%)	2103 (99%)	12 (1%)	91 95

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	412	ASP
1	E	414	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	I	412	ASP
1	E	111	LEU
1	I	111	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	67	ASN
4	L	171	GLN
4	H	171	GLN
4	D	171	GLN
1	I	67	ASN

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.