



## wwPDB EM Map/Model Validation Report

Apr 10, 2016 – 01:40 PM BST

PDB ID : 3DNN  
EMDB ID: : EMD-5019  
Title : Molecular structure for the HIV-1 gp120 trimer in the unliganded state  
Authors : Borgnia, M.J.; Liu, J.; Bartesaghi, A.; Sapiro, G.; Subramaniam, S.  
Deposited on : 2008-07-02  
Resolution : 20.00 Å(reported)

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](mailto: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

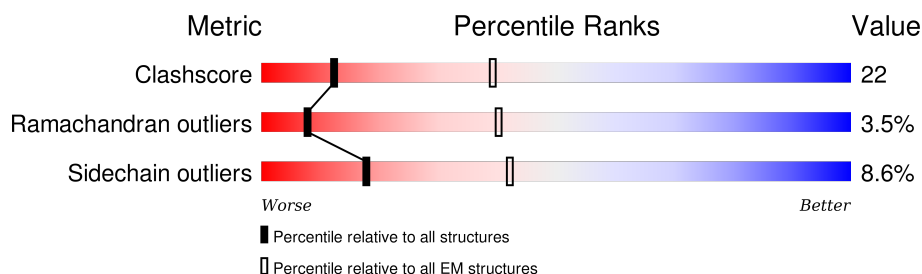
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 20.00 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  $\geq 3$ , 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	35	49% 40% 9% .
1	D	35	46% 43% 9% .
1	G	35	46% 43% 9% .
2	B	170	54% 39% 5% .
2	E	170	54% 40% 5% .
2	H	170	54% 39% 5% .
3	C	83	58% 35% 7%
3	F	83	58% 36% 6%
3	I	83	58% 35% 7%

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 6738 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 HIV-1 envelope glycoprotein gp120.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	35	Total	C	N	O	S	0	0
			293	186	47	56	4		
1	D	35	Total	C	N	O	S	0	0
			293	186	47	56	4		
1	G	35	Total	C	N	O	S	0	0
			293	186	47	56	4		

- Molecule 2 is a protein called HIV-1 envelope glycoprotein gp120.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	170	Total	C	N	O	S	0	0
			1301	818	226	248	9		
2	E	170	Total	C	N	O	S	0	0
			1301	818	226	248	9		
2	H	170	Total	C	N	O	S	0	0
			1301	818	226	248	9		

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	298	GLY	-	LINKER	UNP P04578
B	299	ALA	-	LINKER	UNP P04578
B	329	GLY	-	LINKER	UNP P04578
E	298	GLY	-	LINKER	UNP P04578
E	299	ALA	-	LINKER	UNP P04578
E	329	GLY	-	LINKER	UNP P04578
H	298	GLY	-	LINKER	UNP P04578
H	299	ALA	-	LINKER	UNP P04578
H	329	GLY	-	LINKER	UNP P04578

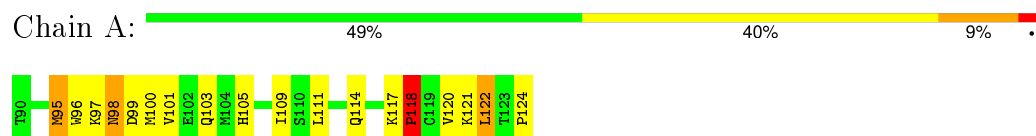
- Molecule 3 is a protein called HIV-1 envelope glycoprotein gp120.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	83	Total 652	C 406	N 118	O 123	S 5	0	0
3	F	83	Total 652	C 406	N 118	O 123	S 5	0	0
3	I	83	Total 652	C 406	N 118	O 123	S 5	0	0

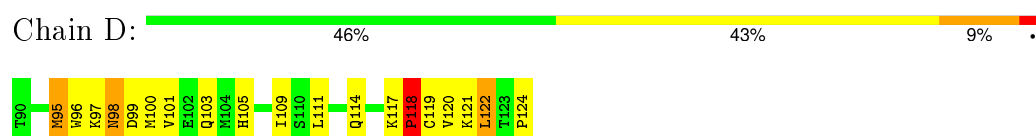
### 3 Residue-property plots

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.

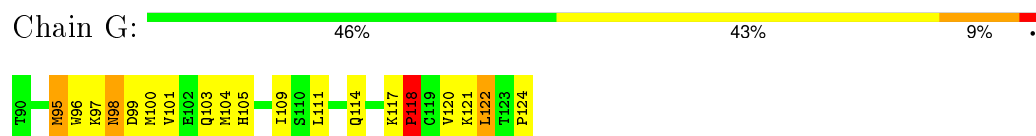
- Molecule 1: HIV-1 envelope glycoprotein gp120



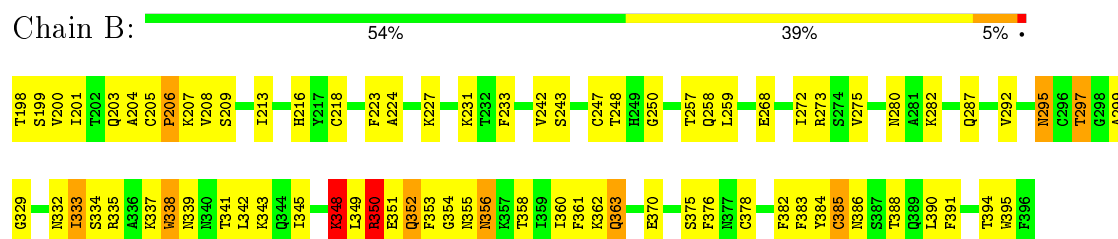
- Molecule 1: HIV-1 envelope glycoprotein gp120



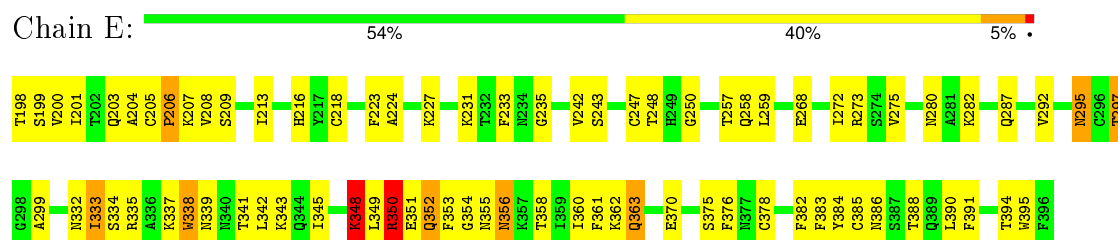
- Molecule 1: HIV-1 envelope glycoprotein gp120



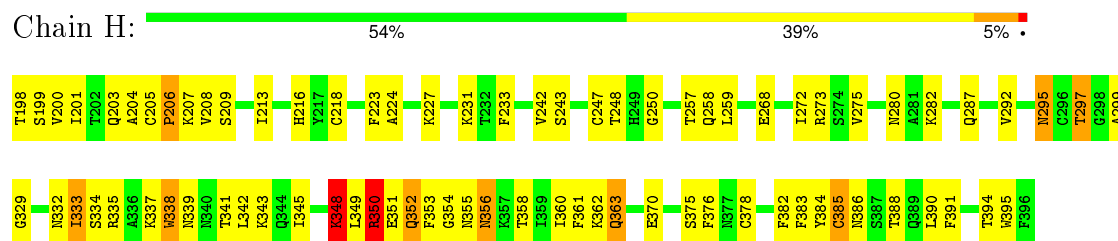
- Molecule 2: HIV-1 envelope glycoprotein gp120



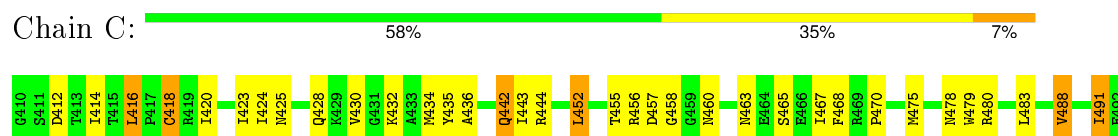
- Molecule 2: HIV-1 envelope glycoprotein gp120



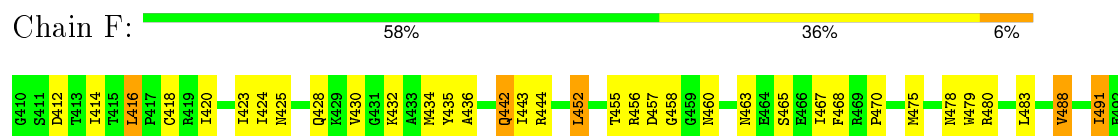
- Molecule 2: HIV-1 envelope glycoprotein gp120



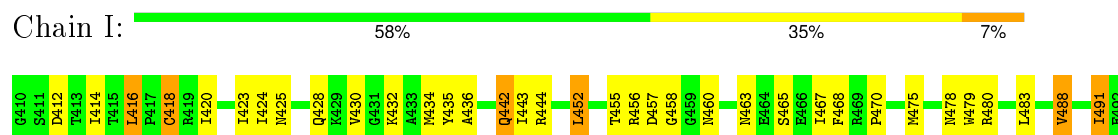
- Molecule 3: HIV-1 envelope glycoprotein gp120



- Molecule 3: HIV-1 envelope glycoprotein gp120



- Molecule 3: HIV-1 envelope glycoprotein gp120



## 4 Experimental information

Property	Value	Source
Reconstruction method	Not provided	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of images	Not provided	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	No CTF correction applied	Depositor
Microscope	FEI Polara G2	Depositor
Voltage (kV)	200	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	8000	Depositor
Minimum defocus (nm)	2000	Depositor
Maximum defocus (nm)	4000	Depositor
Magnification	34000	Depositor
Image detector	CCD	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	0.41	0/300	0.82	0/406
1	D	0.42	0/300	0.81	0/406
1	G	0.42	0/300	0.81	0/406
2	B	0.40	0/1328	0.73	0/1805
2	E	0.40	0/1328	0.73	0/1805
2	H	0.40	0/1328	0.73	0/1805
3	C	0.37	0/663	0.70	0/890
3	F	0.38	0/663	0.70	0/890
3	I	0.38	0/663	0.70	0/890
All	All	0.39	0/6873	0.73	0/9303

There are no bond length outliers.

There are no bond angle outliers.

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	A	293	0	280	18	0
1	D	293	0	280	21	0
1	G	293	0	280	21	0
2	B	1301	0	1271	75	0
2	E	1301	0	1271	73	0
2	H	1301	0	1271	73	0
3	C	652	0	655	40	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	F	652	0	655	38	0
3	I	652	0	655	39	0
All	All	6738	0	6618	293	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 22.

The worst 5 of 293 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:D:124:PRO:HD2	2:E:198:THR:HA	1.52	0.91
1:A:124:PRO:HD2	2:B:198:THR:HA	1.52	0.90
1:G:124:PRO:HD2	2:H:198:THR:HA	1.52	0.90
2:B:363:GLN:HE21	3:C:470:PRO:HG3	1.43	0.83
2:H:363:GLN:HE21	3:I:470:PRO:HG3	1.44	0.83

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	33/35 (94%)	22 (67%)	9 (27%)	2 (6%)	2	26
1	D	33/35 (94%)	22 (67%)	9 (27%)	2 (6%)	2	26
1	G	33/35 (94%)	21 (64%)	10 (30%)	2 (6%)	2	26
2	B	168/170 (99%)	139 (83%)	21 (12%)	8 (5%)	3	32
2	E	168/170 (99%)	139 (83%)	21 (12%)	8 (5%)	3	32
2	H	168/170 (99%)	139 (83%)	21 (12%)	8 (5%)	3	32
3	C	81/83 (98%)	74 (91%)	7 (9%)	0	100	100
3	F	81/83 (98%)	74 (91%)	7 (9%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	I	81/83 (98%)	74 (91%)	7 (9%)	0	100	100
All	All	846/864 (98%)	704 (83%)	112 (13%)	30 (4%)	8	39

5 of 30 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	99	ASP
1	A	118	PRO
2	B	258	GLN
2	B	350	ARG
2	B	352	GLN

### 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	35/35 (100%)	30 (86%)	5 (14%)	4	25
1	D	35/35 (100%)	30 (86%)	5 (14%)	4	25
1	G	35/35 (100%)	30 (86%)	5 (14%)	4	25
2	B	149/149 (100%)	139 (93%)	10 (7%)	20	57
2	E	149/149 (100%)	139 (93%)	10 (7%)	20	57
2	H	149/149 (100%)	139 (93%)	10 (7%)	20	57
3	C	72/72 (100%)	65 (90%)	7 (10%)	10	40
3	F	72/72 (100%)	65 (90%)	7 (10%)	10	40
3	I	72/72 (100%)	65 (90%)	7 (10%)	10	40
All	All	768/768 (100%)	702 (91%)	66 (9%)	18	47

5 of 66 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	E	297	THR
2	E	385	CYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
3	I	418	CYS
2	E	332	ASN
2	E	348	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 30 such sidechains are listed below:

Mol	Chain	Res	Type
2	E	352	GLN
2	E	389	GLN
3	I	442	GLN
2	E	363	GLN
3	F	442	GLN

### 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.