



# wwPDB Geometry-Only Validation Summary Report ⓘ

May 25, 2020 – 03:25 pm BST

PDB ID : 1P5A  
Title : Conformational Mapping of the N-terminal Peptide of HIV-1 GP41 in lipid detergent and aqueous environments using <sup>13</sup>C-enhanced Fourier Transform Infrared Spectroscopy  
Authors : Gordon, L.M.; Mobley, P.W.; Lee, W.; Eskandari, S.; Kaznessis, Y.N.; Sherman, M.A.; Waring, A.J.  
Deposited on : 2003-04-25  
Resolution : Not provided

This is a wwPDB Geometry-Only Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

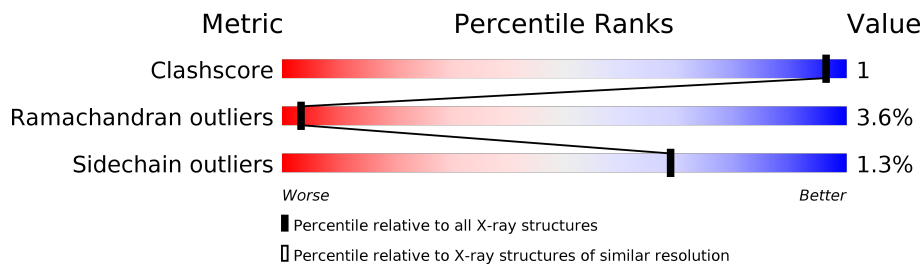
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*INFRARED SPECTROSCOPY*

The reported resolution of this entry is unknown.

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)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	-
Ramachandran outliers	138981	-
Sidechain outliers	138945	-







The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. 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  $\leq 5\%$ .

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	1-A	24	50% 42% 8%
1	10-A	24	83% 13% .
1	11-A	24	71% 29%
1	12-A	24	83% 17%
1	13-A	24	83% 13% .
1	14-A	24	83% 13% .
1	15-A	24	83% 17%
1	16-A	24	83% 8% 8%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	17-A	24	 92% 8%
1	18-A	24	 79% 21%
1	19-A	24	 71% 25% .
1	2-A	24	 88% 8% .
1	3-A	24	 71% 25% .
1	4-A	24	 75% 25%
1	5-A	24	 79% 17% .
1	6-A	24	 88% 8% .
1	7-A	24	 79% 21%
1	8-A	24	 71% 21% 8%
1	9-A	24	 71% 25% .

## 2 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 5814 atoms, of which 2983 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, 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 Envelope polypeptide GP160.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
1	1-A	24	Total 306	C 95	H 157	N 27	O 26	S 1	0	0	1
1	2-A	24	Total 306	C 95	H 157	N 27	O 26	S 1	0	0	1
1	3-A	24	Total 306	C 95	H 157	N 27	O 26	S 1	0	0	1
1	4-A	24	Total 306	C 95	H 157	N 27	O 26	S 1	0	0	1
1	5-A	24	Total 306	C 95	H 157	N 27	O 26	S 1	0	0	1
1	6-A	24	Total 306	C 95	H 157	N 27	O 26	S 1	0	0	1
1	7-A	24	Total 306	C 95	H 157	N 27	O 26	S 1	0	0	1
1	8-A	24	Total 306	C 95	H 157	N 27	O 26	S 1	0	0	1
1	9-A	24	Total 306	C 95	H 157	N 27	O 26	S 1	0	0	1
1	10-A	24	Total 306	C 95	H 157	N 27	O 26	S 1	0	0	1
1	11-A	24	Total 306	C 95	H 157	N 27	O 26	S 1	0	0	1
1	12-A	24	Total 306	C 95	H 157	N 27	O 26	S 1	0	0	1
1	13-A	24	Total 306	C 95	H 157	N 27	O 26	S 1	0	0	1
1	14-A	24	Total 306	C 95	H 157	N 27	O 26	S 1	0	0	1
1	15-A	24	Total 306	C 95	H 157	N 27	O 26	S 1	0	0	1
1	16-A	24	Total 306	C 95	H 157	N 27	O 26	S 1	0	0	1

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>					<b>ZeroOcc</b>	<b>AltConf</b>	<b>Trace</b>	
1	17-A	24	Total	C	H	N	O	S	0	0	1
			306	95	157	27	26	1			
1	18-A	24	Total	C	H	N	O	S	0	0	1
			306	95	157	27	26	1			
1	19-A	24	Total	C	H	N	O	S	0	0	1
			306	95	157	27	26	1			

### 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 the various outlier classes 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.

Note EDS was not executed.

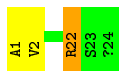
- Molecule 1: Envelope polyprotein GP160

Chain 1-A: 



- Molecule 1: Envelope polyprotein GP160

Chain 2-A: 




- Molecule 1: Envelope polyprotein GP160

Chain 3-A: 




- Molecule 1: Envelope polyprotein GP160

Chain 4-A: 



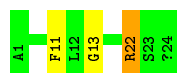
- Molecule 1: Envelope polyprotein GP160

Chain 5-A: 




- Molecule 1: Envelope polyprotein GP160

Chain 6-A:  88% 8%



- Molecule 1: Envelope polyprotein GP160

Chain 7-A:  79% 21%



- Molecule 1: Envelope polyprotein GP160

Chain 8-A:  71% 21% 8%



- Molecule 1: Envelope polyprotein GP160

Chain 9-A:  71% 25%




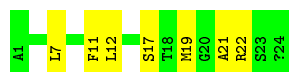
- Molecule 1: Envelope polyprotein GP160

Chain 10-A:  83% 13%




- Molecule 1: Envelope polyprotein GP160

Chain 11-A:  71% 29%




- Molecule 1: Envelope polyprotein GP160

Chain 12-A:  83% 17%




- Molecule 1: Envelope polyprotein GP160

Chain 13-A:  83% 13%




• Molecule 1: Envelope polyprotein GP160

Chain 14-A:  83% 13%




• Molecule 1: Envelope polyprotein GP160

Chain 15-A:  83% 17%



• Molecule 1: Envelope polyprotein GP160

Chain 16-A:  83% 8% 8%




• Molecule 1: Envelope polyprotein GP160

Chain 17-A:  92% 8%




• Molecule 1: Envelope polyprotein GP160

Chain 18-A:  79% 21%



• Molecule 1: Envelope polyprotein GP160

Chain 19-A:  71% 25%





## 4 Model quality

### 4.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NH2

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  >5	RMSZ	# Z  >5
1	1-A	1.67	3/149 (2.0%)	2.25	8/198 (4.0%)
1	2-A	1.80	1/149 (0.7%)	2.30	4/198 (2.0%)
1	3-A	1.76	2/149 (1.3%)	2.18	7/198 (3.5%)
1	4-A	1.67	1/149 (0.7%)	1.97	3/198 (1.5%)
1	5-A	1.73	2/149 (1.3%)	1.98	3/198 (1.5%)
1	6-A	1.61	0/149	2.06	3/198 (1.5%)
1	7-A	1.71	0/149	1.93	3/198 (1.5%)
1	8-A	1.56	1/149 (0.7%)	2.05	4/198 (2.0%)
1	9-A	1.65	1/149 (0.7%)	2.11	5/198 (2.5%)
1	10-A	1.68	0/149	1.81	1/198 (0.5%)
1	11-A	1.59	0/149	2.28	4/198 (2.0%)
1	12-A	1.80	1/149 (0.7%)	2.31	4/198 (2.0%)
1	13-A	1.59	0/149	2.14	3/198 (1.5%)
1	14-A	1.70	0/149	1.85	3/198 (1.5%)
1	15-A	1.69	1/149 (0.7%)	1.97	4/198 (2.0%)
1	16-A	1.67	2/149 (1.3%)	1.99	3/198 (1.5%)
1	17-A	1.62	1/149 (0.7%)	1.73	2/198 (1.0%)
1	18-A	1.73	2/149 (1.3%)	1.95	2/198 (1.0%)
1	19-A	1.64	0/149	2.04	5/198 (2.5%)
All	All	1.68	18/2831 (0.6%)	2.05	71/3762 (1.9%)

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	1-A	0	1
1	2-A	0	1
1	5-A	0	1
1	6-A	0	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	#Chirality outliers	#Planarity outliers
1	9-A	0	2
1	10-A	0	2
1	13-A	0	1
1	14-A	0	2
1	16-A	0	1
1	19-A	0	1
All	All	0	13

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	5-A	17	SER	CA-CB	7.52	1.64	1.52
1	16-A	11	PHE	CG-CD2	6.64	1.48	1.38
1	12-A	8	PHE	CE2-CZ	6.42	1.49	1.37
1	4-A	17	SER	CA-CB	6.36	1.62	1.52
1	18-A	8	PHE	CG-CD2	6.30	1.48	1.38

The worst 5 of 71 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	12-A	22	ARG	NE-CZ-NH2	-16.25	112.17	120.30
1	11-A	22	ARG	NE-CZ-NH2	-16.13	112.24	120.30
1	2-A	22	ARG	NE-CZ-NH2	-16.08	112.26	120.30
1	13-A	22	ARG	NE-CZ-NH1	14.13	127.36	120.30
1	2-A	22	ARG	NE-CZ-NH1	14.08	127.34	120.30

There are no chirality outliers.

5 of 13 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	1-A	11	PHE	Sidechain
1	2-A	22	ARG	Sidechain
1	5-A	11	PHE	Sidechain
1	6-A	22	ARG	Sidechain
1	9-A	8	PHE	Sidechain

## 4.2 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	1-A	149	157	155	1	0
1	2-A	149	157	155	0	0
1	3-A	149	157	155	0	0
1	4-A	149	157	155	1	0
1	5-A	149	157	155	0	0
1	6-A	149	157	155	0	0
1	7-A	149	157	155	0	0
1	8-A	149	157	155	1	0
1	9-A	149	157	155	0	0
1	10-A	149	157	155	1	0
1	11-A	149	157	155	1	0
1	12-A	149	157	155	0	0
1	13-A	149	157	155	0	0
1	14-A	149	157	155	0	0
1	15-A	149	157	155	0	0
1	16-A	149	157	155	1	0
1	17-A	149	157	155	0	0
1	18-A	149	157	155	0	0
1	19-A	149	157	155	0	0
All	All	2831	2983	2945	6	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 1.

The worst 5 of 6 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:12:LEU:HD22	1:A:19:MET:HB2	1.89	0.54
1:A:12:LEU:HD23	1:A:17:SER:HB3	1.98	0.46
1:A:8:PHE:HA	1:A:12:LEU:HD12	1.99	0.44
1:A:12:LEU:HD22	1:A:19:MET:HB2	2.02	0.42
1:A:23:SER:HG	1:A:24:NH2:N	2.19	0.41

There are no symmetry-related clashes.

## 4.3 Torsion angles [i](#)

### 4.3.1 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 X-ray entries followed by that with respect to entries of similar resolution.

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	1-A	22/24 (92%)	14 (64%)	6 (27%)	2 (9%)	1	1
1	2-A	22/24 (92%)	14 (64%)	8 (36%)	0	100	100
1	3-A	22/24 (92%)	18 (82%)	3 (14%)	1 (4%)	2	2
1	4-A	22/24 (92%)	17 (77%)	5 (23%)	0	100	100
1	5-A	22/24 (92%)	15 (68%)	6 (27%)	1 (4%)	2	2
1	6-A	22/24 (92%)	15 (68%)	6 (27%)	1 (4%)	2	2
1	7-A	22/24 (92%)	12 (54%)	9 (41%)	1 (4%)	2	2
1	8-A	22/24 (92%)	15 (68%)	4 (18%)	3 (14%)	0	0
1	9-A	22/24 (92%)	13 (59%)	8 (36%)	1 (4%)	2	2
1	10-A	22/24 (92%)	16 (73%)	6 (27%)	0	100	100
1	11-A	22/24 (92%)	14 (64%)	7 (32%)	1 (4%)	2	2
1	12-A	22/24 (92%)	13 (59%)	9 (41%)	0	100	100
1	13-A	22/24 (92%)	17 (77%)	5 (23%)	0	100	100
1	14-A	22/24 (92%)	19 (86%)	3 (14%)	0	100	100
1	15-A	22/24 (92%)	17 (77%)	5 (23%)	0	100	100
1	16-A	22/24 (92%)	15 (68%)	7 (32%)	0	100	100
1	17-A	22/24 (92%)	16 (73%)	6 (27%)	0	100	100
1	18-A	22/24 (92%)	15 (68%)	6 (27%)	1 (4%)	2	2
1	19-A	22/24 (92%)	11 (50%)	8 (36%)	3 (14%)	0	0
All	All	418/456 (92%)	286 (68%)	117 (28%)	15 (4%)	3	3

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	19-A	23	SER
1	1-A	3	GLY

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	3-A	22	ARG
1	8-A	15	ALA
1	19-A	15	ALA

#### 4.3.2 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 X-ray entries followed by that with respect to entries of similar resolution.

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	1-A	12/12 (100%)	12 (100%)	0	100	100
1	2-A	12/12 (100%)	12 (100%)	0	100	100
1	3-A	12/12 (100%)	12 (100%)	0	100	100
1	4-A	12/12 (100%)	12 (100%)	0	100	100
1	5-A	12/12 (100%)	12 (100%)	0	100	100
1	6-A	12/12 (100%)	12 (100%)	0	100	100
1	7-A	12/12 (100%)	11 (92%)	1 (8%)	11	11
1	8-A	12/12 (100%)	12 (100%)	0	100	100
1	9-A	12/12 (100%)	12 (100%)	0	100	100
1	10-A	12/12 (100%)	12 (100%)	0	100	100
1	11-A	12/12 (100%)	11 (92%)	1 (8%)	11	11
1	12-A	12/12 (100%)	12 (100%)	0	100	100
1	13-A	12/12 (100%)	11 (92%)	1 (8%)	11	11
1	14-A	12/12 (100%)	12 (100%)	0	100	100
1	15-A	12/12 (100%)	12 (100%)	0	100	100
1	16-A	12/12 (100%)	12 (100%)	0	100	100
1	17-A	12/12 (100%)	12 (100%)	0	100	100
1	18-A	12/12 (100%)	12 (100%)	0	100	100
1	19-A	12/12 (100%)	12 (100%)	0	100	100
All	All	228/228 (100%)	225 (99%)	3 (1%)	69	69

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	7-A	2	VAL
1	11-A	17	SER
1	13-A	2	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

#### 4.3.3 RNA [i](#)

There are no RNA molecules in this entry.

#### 4.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

#### 4.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

#### 4.6 Ligand geometry [i](#)

There are no ligands in this entry.

#### 4.7 Other polymers [i](#)

There are no such residues in this entry.

#### 4.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.