



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 17, 2017 – 10:26 PM EDT

PDB ID : 5U3L  
Title : Crystal Structure of DH511.2 Fab in Complex with HIV-1 gp41 MPER 670-683 Peptide  
Authors : Ofek, G.; Wu, L.; Loughheed, C.S.; Williams, L.D.; Nicely, N.I.; Haynes, B.F.  
Deposited on : unknown  
Resolution : 2.17 Å(reported)

This is a wwPDB X-ray Structure 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

<http://wwpdb.org/validation/2016/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  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20029824  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20029824

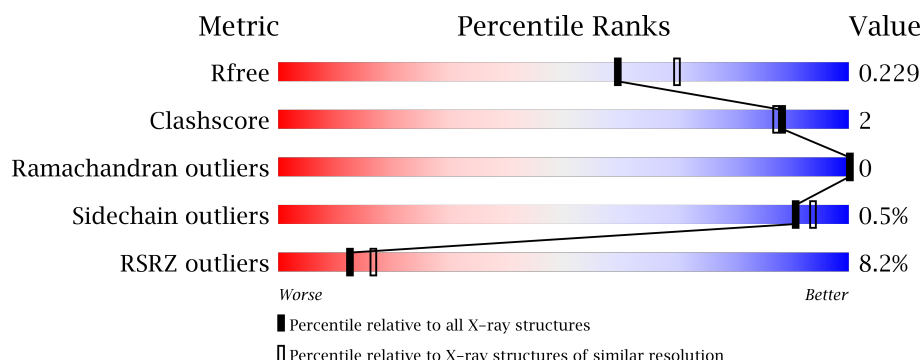
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.17 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.



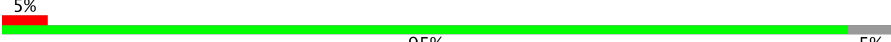
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	100719	1170 (2.16-2.16)
Clashscore	112137	1278 (2.16-2.16)
Ramachandran outliers	110173	1256 (2.16-2.16)
Sidechain outliers	110143	1255 (2.16-2.16)
RSRZ outliers	101464	1175 (2.16-2.16)

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. 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  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	232	<div> <div>6%</div> <div>91%</div> <div>6%</div> </div>
1	H	232	<div> <div>10%</div> <div>86%</div> <div>10%</div> </div>
2	B	214	<div> <div>12%</div> <div>98%</div> <div></div> </div>
2	L	214	<div> <div>5%</div> <div>94%</div> <div>6%</div> </div>
3	C	20	<div> <div></div> <div>95%</div> <div>5%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain	
			5%	5%
3	P	20		

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 13856 atoms, of which 6660 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 DH511.2 Fab Heavy Chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	H	223	Total	C	H	N	O	S	0	0	0
			3253	1060	1583	283	318	9			
1	A	224	Total	C	H	N	O	S	0	0	0
			3362	1084	1659	287	323	9			

- Molecule 2 is a protein called DH511.2 Fab Light Chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	L	214	Total	C	H	N	O	S	0	0	0
			3238	1029	1588	283	333	5			
2	B	213	Total	C	H	N	O	S	0	0	0
			3125	1000	1517	277	326	5			

- Molecule 3 is a protein called gp41 MPER peptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	P	19	Total	C	H	N	O	0	0	0
			336	123	158	30	25			
3	C	19	Total	C	H	N	O	0	0	0
			329	120	155	29	25			

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	H	60	Total	O	0	0
			60	60		
4	L	44	Total	O	0	0
			44	44		
4	A	63	Total	O	0	0
			63	63		
4	B	39	Total	O	0	0
			39	39		

*Continued on next page...*

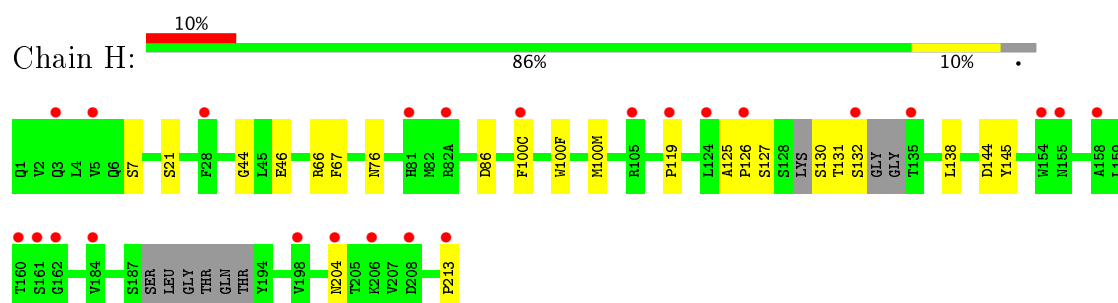
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	P	1	Total	O	0	0
			1	1		
4	C	6	Total	O	0	0
			6	6		

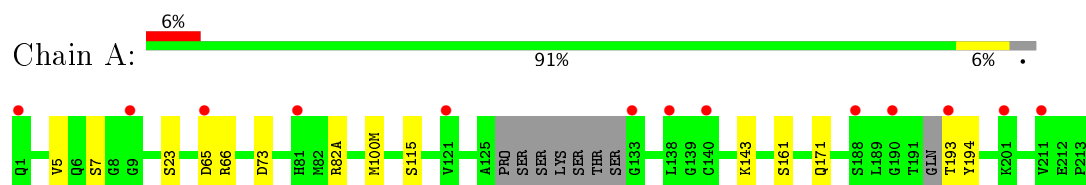
### 3 Residue-property plots [i](#)

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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). 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.

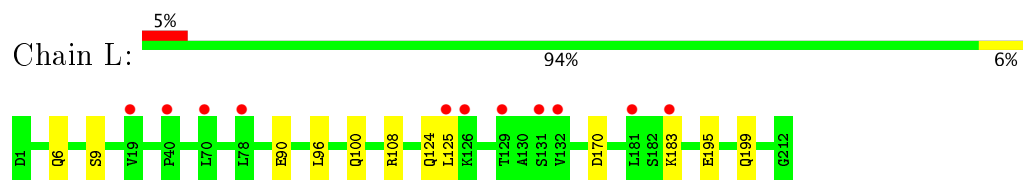
- Molecule 1: DH511.2 Fab Heavy Chain



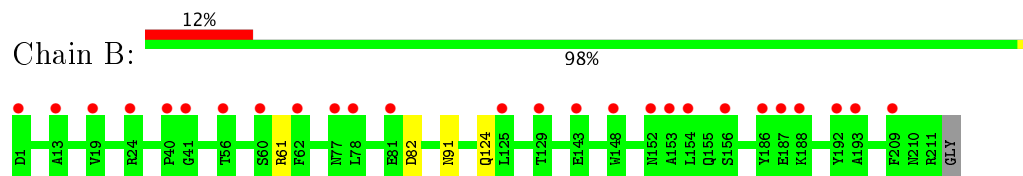
- Molecule 1: DH511.2 Fab Heavy Chain



- Molecule 2: DH511.2 Fab Light Chain

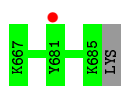


- Molecule 2: DH511.2 Fab Light Chain



- Molecule 3: gp41 MPER peptide





- Molecule 3: gp41 MPER peptide

Chain C:  95% 5%



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	110.07Å 75.64Å 123.43Å 90.00° 94.11° 90.00°	Depositor
Resolution (Å)	42.51 – 2.17 42.51 – 2.17	Depositor EDS
% Data completeness (in resolution range)	94.3 (42.51-2.17) 89.4 (42.51-2.17)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.58 (at 2.16Å)	Xtriage
Refinement program	PHENIX 1.11.1 _2575	Depositor
R, $R_{free}$	0.205 , 0.231 0.206 , 0.229	Depositor DCC
$R_{free}$ test set	2465 reflections (5.11%)	DCC
Wilson B-factor (Å <sup>2</sup> )	34.5	Xtriage
Anisotropy	0.441	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 48.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	13856	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	61.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.89% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.



## 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  >5	RMSZ	# Z  >5
1	A	0.27	0/1746	0.48	0/2372
1	H	0.29	0/1713	0.49	0/2329
2	B	0.27	0/1642	0.46	0/2237
2	L	0.30	0/1685	0.49	0/2292
3	C	0.62	0/183	0.37	0/252
3	P	0.57	0/187	0.45	0/256
All	All	0.31	0/7156	0.48	0/9738

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	1703	1659	1658	10	0
1	H	1670	1583	1580	15	9
2	B	1608	1517	1520	2	0
2	L	1650	1588	1602	6	2
3	C	174	155	147	0	0
3	P	178	158	158	0	0
4	A	63	0	0	6	0
4	B	39	0	0	1	0
4	C	6	0	0	0	0
4	H	60	0	0	6	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	L	44	0	0	2	0
4	P	1	0	0	0	0
All	All	7196	6660	6665	32	11

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 2.

The worst 5 of 32 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:H:130:SER:N	4:H:301:HOH:O	2.04	0.90
2:B:124:GLN:OE1	4:B:301:HOH:O	1.97	0.81
1:A:161:SER:O	4:A:301:HOH:O	2.02	0.77
2:L:90:GLU:OE2	4:L:301:HOH:O	2.06	0.72
1:H:144:ASP:OD1	4:H:303:HOH:O	2.11	0.68

The worst 5 of 11 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:H:100(C):PHE:CE2	1:H:100(C):PHE:CZ[2_456]	1.29	0.91
1:H:100(C):PHE:CD2	1:H:100(C):PHE:CZ[2_456]	1.42	0.78
1:H:100(C):PHE:CZ	1:H:100(C):PHE:HE2[2_456]	1.21	0.39
1:H:100(C):PHE:CD2	1:H:100(C):PHE:CE1[2_456]	1.92	0.28
2:L:9:SER:OG	2:L:195:GLU:OE2[2_556]	1.99	0.21

## 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 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	A	218/232 (94%)	215 (99%)	3 (1%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	H	215/232 (93%)	214 (100%)	1 (0%)	0	100	100
2	B	211/214 (99%)	203 (96%)	8 (4%)	0	100	100
2	L	212/214 (99%)	205 (97%)	7 (3%)	0	100	100
3	C	17/20 (85%)	17 (100%)	0	0	100	100
3	P	17/20 (85%)	17 (100%)	0	0	100	100
All	All	890/932 (96%)	871 (98%)	19 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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 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	A	185/194 (95%)	184 (100%)	1 (0%)	91	94
1	H	176/194 (91%)	175 (99%)	1 (1%)	89	92
2	B	177/189 (94%)	176 (99%)	1 (1%)	89	92
2	L	189/189 (100%)	188 (100%)	1 (0%)	91	94
3	C	15/20 (75%)	15 (100%)	0	100	100
3	P	16/20 (80%)	16 (100%)	0	100	100
All	All	758/806 (94%)	754 (100%)	4 (0%)	91	94

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

Mol	Chain	Res	Type
1	H	100(M)	MET
2	L	199	GLN
1	A	100(M)	MET
2	B	91	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	L	199	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.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	224/232 (96%)	0.64	13 (5%) 24 30	29, 48, 89, 118	0
1	H	223/232 (96%)	0.90	24 (10%) 6 9	33, 57, 86, 112	0
2	B	213/214 (99%)	0.89	26 (12%) 5 7	32, 59, 96, 110	0
2	L	214/214 (100%)	0.67	11 (5%) 29 36	35, 54, 77, 101	0
3	C	19/20 (95%)	0.33	0 100 100	35, 47, 71, 76	0
3	P	19/20 (95%)	0.66	1 (5%) 27 34	28, 46, 68, 77	0
All	All	912/932 (97%)	0.76	75 (8%) 12 17	28, 54, 89, 118	0

The worst 5 of 75 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	119	PRO	4.0
1	H	213	PRO	3.9
2	B	193	ALA	3.8
1	H	204	ASN	3.8
2	L	125	LEU	3.7

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

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

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.