



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 14, 2018 – 01:15 pm GMT

PDB ID : 4S1R
Title : Crystal structure of a VRC01-lineage antibody, 45-VRC01.H08.F-117225, in complex with clade A/E HIV-1 gp120 core
Authors : Kwon, Y.D.; Yang, Y.; Zhang, B.; Kwong, P.D.
Deposited on : 2015-01-14
Resolution : 3.21 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

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
Mogul : 1.7.3 (157068), CSD as539be (2018)
Xtriage (Phenix) : 1.13
EDS : trunk31020
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk31020

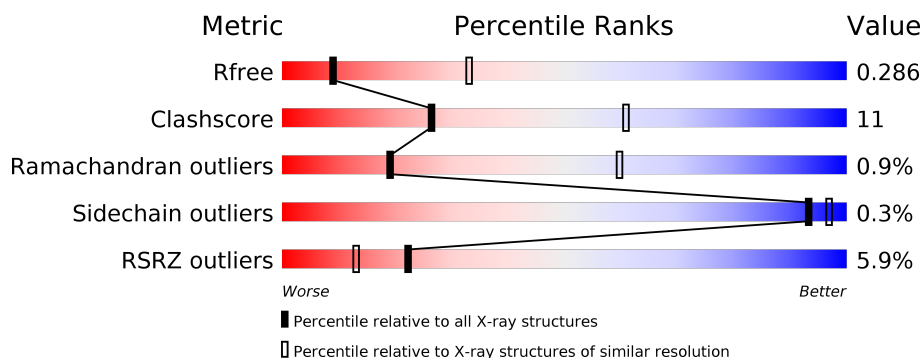
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.21 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}	111664	1145 (3.24-3.20)
Clashscore	122126	1266 (3.24-3.20)
Ramachandran outliers	120053	1245 (3.24-3.20)
Sidechain outliers	120020	1244 (3.24-3.20)
RSRZ outliers	108989	1103 (3.24-3.20)

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 ≥ 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	G	353	<div> <div style="width: 100%; height: 10px; background-color: red; position: relative;"> % </div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> 80% 16% • </div> </div>
2	H	235	<div> <div style="width: 100%; height: 10px; background-color: red; position: relative;"> 7% </div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> 75% 24% </div> </div>
3	L	210	<div> <div style="width: 100%; height: 10px; background-color: red; position: relative;"> 12% </div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> 69% 28% • </div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	NAG	G	503	-	-	-	X
4	NAG	G	510	-	-	-	X

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 6173 atoms, of which 0 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 clade A/E 93TH057 HIV-1 gp120 core.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	G	342	Total	C	N	O	S	0	0	0
			2677	1680	466	509	22			

- Molecule 2 is a protein called Fab of VRC01-lineage antibody,45-VRC01.H08.F-117225 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	235	Total	C	N	O	S	0	0	0
			1761	1102	307	340	12			

- Molecule 3 is a protein called Fab of VRC01 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	203	Total	C	N	O	S	0	0	0
			1581	993	268	315	5			

- Molecule 4 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: C₈H₁₅NO₆).

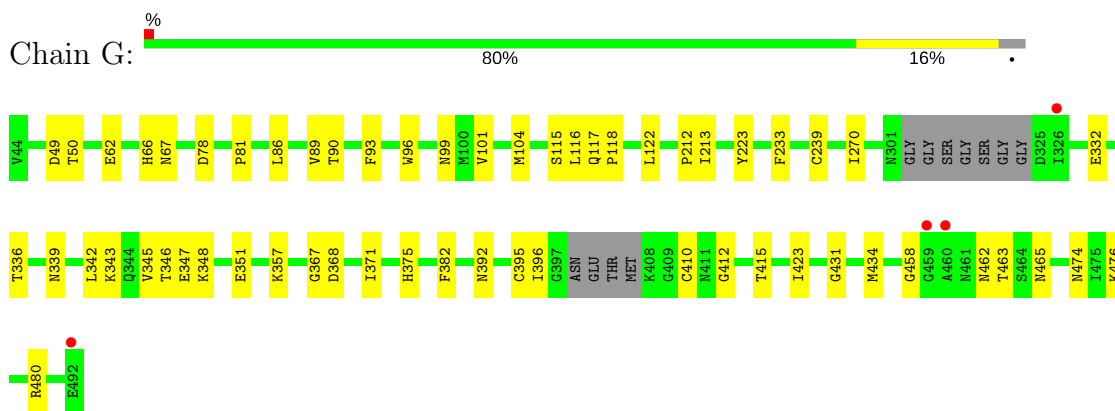


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	G	1	Total	C	N	O	0	0
			14	8	1	5		
4	G	1	Total	C	N	O	0	0
			14	8	1	5		
4	G	1	Total	C	N	O	0	0
			14	8	1	5		
4	G	1	Total	C	N	O	0	0
			14	8	1	5		
4	G	1	Total	C	N	O	0	0
			14	8	1	5		
4	G	1	Total	C	N	O	0	0
			14	8	1	5		
4	G	1	Total	C	N	O	0	0
			14	8	1	5		
4	G	1	Total	C	N	O	0	0
			14	8	1	5		

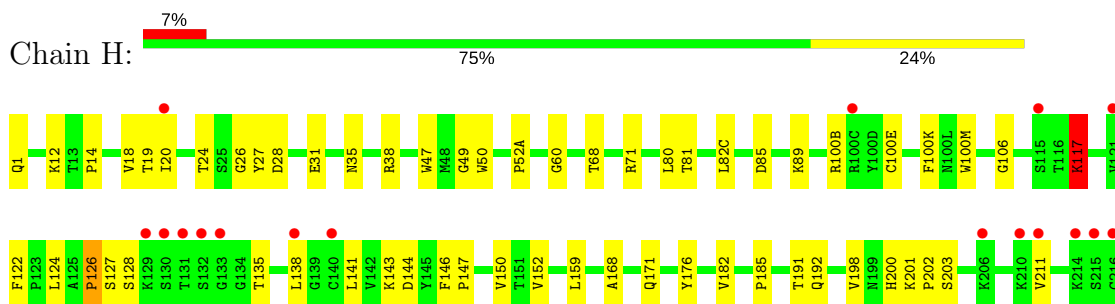
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

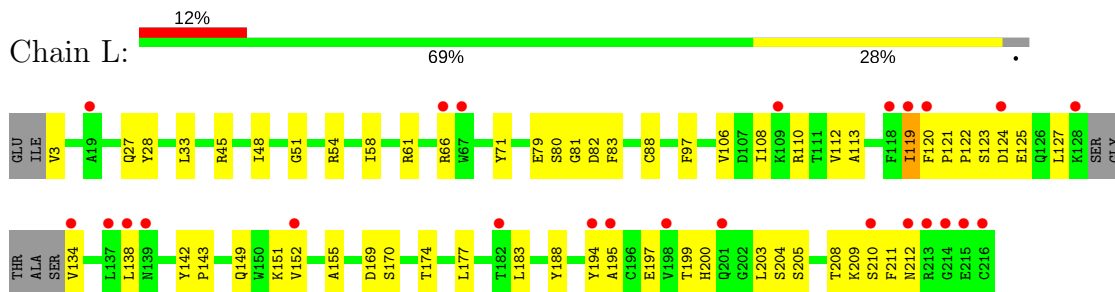
- Molecule 1: clade A/E 93TH057 HIV-1 gp120 core



- Molecule 2: Fab of VRC01-lineage antibody, 45-VRC01.H08.F-117225 heavy chain



- Molecule 3: Fab of VRC01 light chain



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	65.69Å 68.20Å 204.10Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.16 – 3.21 48.16 – 3.21	Depositor EDS
% Data completeness (in resolution range)	93.2 (48.16-3.21) 87.3 (48.16-3.21)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	0.16	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.31 (at 3.19Å)	Xtriage
Refinement program	PHENIX (phenix.refine: dev_1839)	Depositor
R, R_{free}	0.261 , 0.286 0.263 , 0.286	Depositor DCC
R_{free} test set	664 reflections (4.56%)	wwPDB-VP
Wilson B-factor (Å ²)	81.6	Xtriage
Anisotropy	0.201	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 39.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.035 for k,h,-l	Xtriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	6173	wwPDB-VP
Average B, all atoms (Å ²)	102.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.97% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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](#)

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

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	G	0.29	0/2733	0.54	0/3709
2	H	0.24	0/1804	0.45	0/2450
3	L	0.28	0/1617	0.46	0/2194
All	All	0.27	0/6154	0.50	0/8353

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	G	2677	0	2609	51	0
2	H	1761	0	1717	37	0
3	L	1581	0	1521	50	0
4	G	154	0	143	7	0
All	All	6173	0	5990	135	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 11.

The worst 5 of 135 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:G:336:THR:HG21	4:G:508:NAG:H62	1.31	1.10
1:G:336:THR:CG2	4:G:508:NAG:H62	1.83	1.07
2:H:138:LEU:HD22	2:H:211:VAL:HG11	1.35	1.05
2:H:138:LEU:HD22	2:H:211:VAL:CG1	1.99	0.93
1:G:339:ASN:HB3	1:G:396:ILE:HD11	1.58	0.86

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 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	G	336/353 (95%)	322 (96%)	12 (4%)	2 (1%)	27	68
2	H	233/235 (99%)	199 (85%)	30 (13%)	4 (2%)	10	45
3	L	199/210 (95%)	174 (87%)	24 (12%)	1 (0%)	31	71
All	All	768/798 (96%)	695 (90%)	66 (9%)	7 (1%)	19	60

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	116	LEU
2	H	117	LYS
2	H	126	PRO
2	H	144	ASP
1	G	462	ASN

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 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	G	305/311 (98%)	305 (100%)	0	100	100
2	H	197/197 (100%)	196 (100%)	1 (0%)	90	95
3	L	177/182 (97%)	176 (99%)	1 (1%)	87	95
All	All	679/690 (98%)	677 (100%)	2 (0%)	93	97

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	H	117	LYS
3	L	119	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	G	229	ASN
2	H	100(F)	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 ⓘ

11 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
4	NAG	G	501	1	14,14,15	0.69	1 (7%)	17,19,21	0.60	0
4	NAG	G	502	1	14,14,15	0.24	0	17,19,21	0.49	0
4	NAG	G	503	1	14,14,15	0.71	0	17,19,21	0.67	0
4	NAG	G	504	1	14,14,15	0.28	0	17,19,21	0.65	1 (5%)
4	NAG	G	505	1	14,14,15	0.44	0	17,19,21	0.41	0
4	NAG	G	506	1	14,14,15	0.27	0	17,19,21	0.53	0
4	NAG	G	507	1	14,14,15	0.18	0	17,19,21	0.50	0
4	NAG	G	508	1	14,14,15	2.02	1 (7%)	17,19,21	1.55	1 (5%)
4	NAG	G	509	1	14,14,15	0.30	0	17,19,21	0.63	1 (5%)
4	NAG	G	510	1	14,14,15	0.65	0	17,19,21	0.84	1 (5%)
4	NAG	G	511	1	14,14,15	0.21	0	17,19,21	0.53	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	G	501	1	-	0/6/23/26	0/1/1/1
4	NAG	G	502	1	-	0/6/23/26	0/1/1/1
4	NAG	G	503	1	-	0/6/23/26	0/1/1/1
4	NAG	G	504	1	-	0/6/23/26	0/1/1/1
4	NAG	G	505	1	-	0/6/23/26	0/1/1/1
4	NAG	G	506	1	-	0/6/23/26	0/1/1/1
4	NAG	G	507	1	-	0/6/23/26	0/1/1/1
4	NAG	G	508	1	-	0/6/23/26	0/1/1/1
4	NAG	G	509	1	-	0/6/23/26	0/1/1/1
4	NAG	G	510	1	-	0/6/23/26	0/1/1/1
4	NAG	G	511	1	-	0/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	G	508	NAG	O5-C1	-7.40	1.31	1.43
4	G	501	NAG	C1-C2	2.11	1.55	1.52

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	G	508	NAG	C1-O5-C5	-5.44	104.70	112.19
4	G	509	NAG	C1-O5-C5	2.04	115.00	112.19
4	G	504	NAG	C1-O5-C5	2.14	115.13	112.19
4	G	510	NAG	C1-O5-C5	2.97	116.27	112.19

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	G	506	NAG	1	0
4	G	508	NAG	6	0

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, 95th 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 ²)	Q<0.9
1	G	342/353 (96%)	-0.02	4 (1%) 79 67	56, 79, 114, 131	0
2	H	235/235 (100%)	0.50	17 (7%) 15 9	56, 105, 166, 206	0
3	L	203/210 (96%)	0.62	25 (12%) 4 3	76, 131, 183, 203	0
All	All	780/798 (97%)	0.31	46 (5%) 22 13	56, 92, 167, 206	0

The worst 5 of 46 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	130	SER	9.2
2	H	215	SER	7.2
2	H	132	SER	6.4
2	H	131	THR	5.7
2	H	214	LYS	5.5

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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
4	NAG	G	503	14/15	0.65	0.52	105,124,128,128	0
4	NAG	G	510	14/15	0.70	0.44	98,130,134,135	0
4	NAG	G	508	14/15	0.71	0.37	74,103,115,117	0
4	NAG	G	511	14/15	0.80	0.35	80,90,103,103	0
4	NAG	G	501	14/15	0.82	0.42	133,138,141,141	0
4	NAG	G	505	14/15	0.87	0.20	99,104,112,114	0
4	NAG	G	509	14/15	0.91	0.24	71,85,91,93	0
4	NAG	G	507	14/15	0.91	0.15	74,84,94,101	0
4	NAG	G	502	14/15	0.92	0.16	80,86,92,94	0
4	NAG	G	506	14/15	0.93	0.19	70,93,100,102	0
4	NAG	G	504	14/15	0.94	0.14	59,69,78,84	0

6.5 Other polymers [i](#)

There are no such residues in this entry.