



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 11, 2017 – 06:53 AM EDT

PDB ID : 5I8E  
Title : Crystal Structure of Broadly Neutralizing HIV-1 Fusion Peptide-Targeting Antibody VRC34.01 Fab  
Authors : Xu, K.; Zhou, T.; Liu, K.; Kwong, P.D.  
Deposited on : unknown  
Resolution : 2.65 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

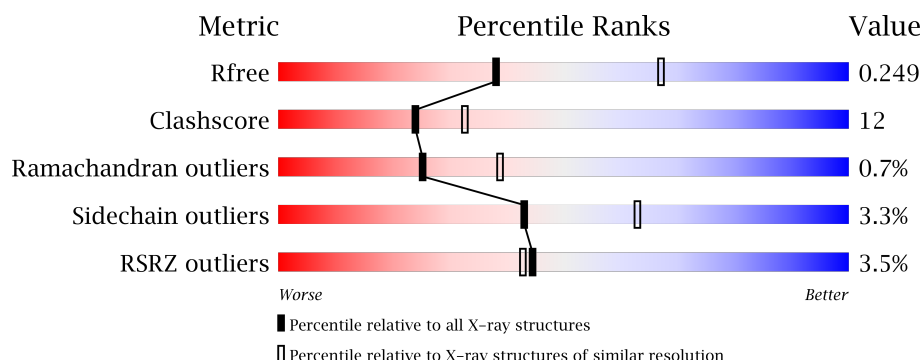
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.65 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	3491 (2.70-2.62)
Clashscore	112137	1026 (2.68-2.64)
Ramachandran outliers	110173	1010 (2.68-2.64)
Sidechain outliers	110143	1010 (2.68-2.64)
RSRZ outliers	101464	3511 (2.70-2.62)

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	222	<div> <div>8%</div> <div>74%</div> <div>23%</div> <div>..</div> </div>
1	C	222	<div> <div>81%</div> <div>15%</div> <div>..</div> </div>
2	B	210	<div> <div>3%</div> <div>70%</div> <div>24%</div> <div>..</div> </div>
2	D	210	<div> <div>2%</div> <div>65%</div> <div>32%</div> <div>.</div> </div>

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 6546 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 VRC34.01 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	217	Total	C	N	O	S	0	0	0
			1633	1031	277	319	6			
1	C	216	Total	C	N	O	S	0	0	0
			1627	1027	276	318	6			

- Molecule 2 is a protein called VRC34.01 Fab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	203	Total	C	N	O	S	0	0	0
			1541	976	258	302	5			
2	D	210	Total	C	N	O	S	0	0	0
			1610	1015	269	321	5			

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Zn	0	0
			1	1		
3	A	4	Total	Zn	0	0
			4	4		
3	D	1	Total	Zn	0	0
			1	1		
3	C	4	Total	Zn	0	0
			4	4		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	33	Total	O	0	0
			33	33		
4	B	31	Total	O	0	0
			31	31		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	C	31	Total	O	0	0
			31	31		
4	D	30	Total	O	0	0
			30	30		





## 4 Data and refinement statistics

Property	Value	Source
Space group	I 4	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	114.61Å 114.61Å 174.08Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	37.31 – 2.65 37.31 – 2.66	Depositor EDS
% Data completeness (in resolution range)	92.9 (37.31-2.65) 92.9 (37.31-2.66)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.40 (at 2.65Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.186 , 0.248 0.187 , 0.249	Depositor DCC
$R_{free}$ test set	1536 reflections (5.11%)	DCC
Wilson B-factor (Å <sup>2</sup> )	64.6	Xtriage
Anisotropy	0.027	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 50.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	0.116 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	6546	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	82.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.69% 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](#)

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

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	0.46	0/1673	0.71	0/2280
1	C	0.51	0/1667	0.66	0/2273
2	B	0.54	0/1577	0.77	0/2145
2	D	0.51	0/1647	0.76	1/2239 (0.0%)
All	All	0.51	0/6564	0.73	1/8937 (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
2	B	0	3
2	D	0	3
All	All	0	6

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
2	D	157	GLY	N-CA-C	5.78	127.56	113.10

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	185	ASP	Peptide
2	B	190	LYS	Peptide
2	B	50	TYR	Peptide

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Group
2	D	150	VAL	Peptide
2	D	50	TYR	Peptide

## 5.2 Too-close contacts

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	1633	0	1593	41	0
1	C	1627	0	1586	22	0
2	B	1541	0	1487	44	0
2	D	1610	0	1569	54	0
3	A	4	0	0	0	0
3	B	1	0	0	0	0
3	C	4	0	0	0	0
3	D	1	0	0	0	0
4	A	33	0	0	0	0
4	B	31	0	0	0	0
4	C	31	0	0	0	0
4	D	30	0	0	1	0
All	All	6546	0	6235	158	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 12.

The worst 5 of 158 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)
2:D:29:VAL:HG13	2:D:92:SER:HB3	1.50	0.92
2:B:186:TYR:HD1	2:B:188:LYS:HD2	1.36	0.88
2:B:129:THR:HG23	2:B:182:SER:HB2	1.57	0.86
1:C:138:LEU:HD13	1:C:211:VAL:HG21	1.57	0.85
2:B:186:TYR:CD1	2:B:188:LYS:HD2	2.11	0.85

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	213/222 (96%)	204 (96%)	9 (4%)	0	100	100
1	C	212/222 (96%)	203 (96%)	9 (4%)	0	100	100
2	B	199/210 (95%)	188 (94%)	7 (4%)	4 (2%)	9	13
2	D	208/210 (99%)	197 (95%)	9 (4%)	2 (1%)	18	28
All	All	832/864 (96%)	792 (95%)	34 (4%)	6 (1%)	25	39

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	51	ALA
2	D	51	ALA
2	B	31	ASN
2	D	190	LYS
2	B	149	LYS

### 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	182/186 (98%)	175 (96%)	7 (4%)	38	57
1	C	181/186 (97%)	175 (97%)	6 (3%)	43	63
2	B	171/183 (93%)	166 (97%)	5 (3%)	48	68
2	D	183/183 (100%)	177 (97%)	6 (3%)	43	63
All	All	717/738 (97%)	693 (97%)	24 (3%)	43	63

5 of 24 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	158	ASN
1	C	59	THR
2	D	158	ASN
2	B	188	LYS
1	C	50	TRP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	38	GLN
2	D	147	GLN
2	B	138	ASN
2	B	6	GLN
2	B	55	GLN

### 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 ⓘ

Of 10 ligands modelled in this entry, 10 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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	217/222 (97%)	0.08	18 (8%) 12 10	40, 73, 167, 214	0
1	C	216/222 (97%)	-0.23	1 (0%) 90 91	39, 69, 109, 134	0
2	B	203/210 (96%)	-0.01	7 (3%) 46 44	40, 71, 166, 229	0
2	D	210/210 (100%)	-0.11	4 (1%) 67 66	41, 79, 152, 210	0
All	All	846/864 (97%)	-0.07	30 (3%) 44 43	39, 74, 156, 229	0

The worst 5 of 30 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	157	GLY	7.8
1	A	129	LYS	4.6
1	A	137	ALA	4.5
1	A	138	LEU	4.3
1	A	158	ALA	4.1

### 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	ZN	A	303	1/1	0.98	0.18	1.23	51,51,51,51	0
3	ZN	C	303	1/1	0.99	0.14	0.23	49,49,49,49	0
3	ZN	C	304	1/1	0.97	0.16	-0.03	103,103,103,103	0
3	ZN	A	304	1/1	0.97	0.14	-0.13	109,109,109,109	0
3	ZN	A	301	1/1	0.92	0.12	-0.81	118,118,118,118	0
3	ZN	C	301	1/1	0.87	0.09	-2.64	113,113,113,113	0
3	ZN	B	301	1/1	1.00	0.22	-	94,94,94,94	0
3	ZN	C	302	1/1	0.99	0.16	-	72,72,72,72	0
3	ZN	A	302	1/1	0.99	0.24	-	81,81,81,81	0
3	ZN	D	301	1/1	0.99	0.23	-	58,58,58,58	0

## 6.5 Other polymers

There are no such residues in this entry.