



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 26, 2020 – 08:45 am BST

PDB ID : 2SIV  
Title : SIV GP41 CORE STRUCTURE  
Authors : Malashkevich, V.N.; Chan, D.C.; Chutkowski, C.T.; Kim, P.S.  
Deposited on : 1998-06-17  
Resolution : 2.20 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

<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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
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

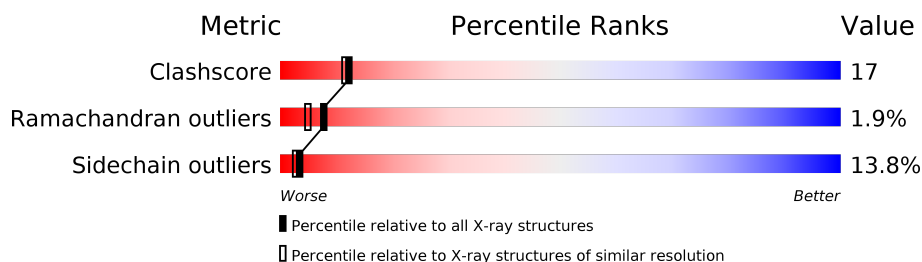
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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))
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)

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  $\geq 3$ , 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	A	38	
1	C	38	
1	E	38	
2	B	36	
2	D	36	
2	F	36	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 1993 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 SIV GP41 GLYCOPROTEIN.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	A	38	Total	C	N	O	0	0	1
			298	186	58	54			
1	C	38	Total	C	N	O	0	0	1
			298	186	58	54			
1	E	38	Total	C	N	O	0	0	1
			298	186	58	54			

- Molecule 2 is a protein called SIV GP41 GLYCOPROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	36	Total	C	N	O	S	7	0	1
			305	193	49	62	1			
2	D	36	Total	C	N	O	S	0	0	1
			305	193	49	62	1			
2	F	36	Total	C	N	O	S	7	0	1
			305	193	49	62	1			

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	634	LYS	GLN	CONFLICT	UNP Q87973
B	640	GLU	ALA	CONFLICT	UNP Q87973
B	644	ALA	GLN	CONFLICT	UNP Q87973
D	634	LYS	GLN	CONFLICT	UNP Q87973
D	640	GLU	ALA	CONFLICT	UNP Q87973
D	644	ALA	GLN	CONFLICT	UNP Q87973
F	634	LYS	GLN	CONFLICT	UNP Q87973
F	640	GLU	ALA	CONFLICT	UNP Q87973
F	644	ALA	GLN	CONFLICT	UNP Q87973

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	33	Total 33	O 33	0	0
3	B	47	Total 47	O 47	0	0
3	C	26	Total 26	O 26	0	0
3	D	30	Total 30	O 30	0	0
3	E	21	Total 21	O 21	0	0
3	F	27	Total 27	O 27	0	0

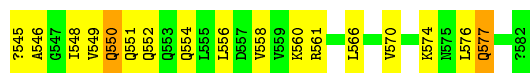
### 3 Residue-property plots [i](#)

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.

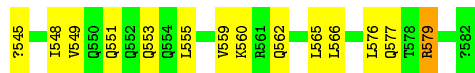
- Molecule 1: SIV GP41 GLYCOPROTEIN

Chain A: 



- Molecule 1: SIV GP41 GLYCOPROTEIN

Chain C: 



- Molecule 1: SIV GP41 GLYCOPROTEIN

Chain E: 



- Molecule 2: SIV GP41 GLYCOPROTEIN

Chain B: 



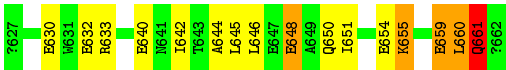
- Molecule 2: SIV GP41 GLYCOPROTEIN

Chain D: 



- Molecule 2: SIV GP41 GLYCOPROTEIN

Chain F: 



## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	55.72Å 57.71Å 66.51Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.20	Depositor
% Data completeness (in resolution range)	95.7 (20.00-2.20)	Depositor
$R_{merge}$	0.06	Depositor
$R_{sym}$	0.06	Depositor
Refinement program	X-PLOR 3.851	Depositor
R, $R_{free}$	0.211 , 0.288	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	1993	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	32.0	wwPDB-VP

## 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: ACE, 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	A	0.37	0/296	0.51	0/400
1	C	0.43	0/296	0.73	1/400 (0.2%)
1	E	0.46	0/296	0.86	1/400 (0.2%)
2	B	0.42	0/307	0.52	0/414
2	D	0.43	0/307	0.53	0/414
2	F	0.38	0/307	0.52	0/414
All	All	0.42	0/1809	0.62	2/2442 (0.1%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	579	ARG	NE-CZ-NH2	6.98	123.79	120.30
1	E	579	ARG	NE-CZ-NH2	6.74	123.67	120.30

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	298	0	319	18	0
1	C	298	0	319	14	0
1	E	298	0	319	17	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	305	0	285	12	0
2	D	305	0	285	14	0
2	F	305	0	285	8	0
3	A	33	0	0	0	0
3	B	47	0	0	2	0
3	C	26	0	0	1	0
3	D	30	0	0	2	0
3	E	21	0	0	1	0
3	F	27	0	0	0	0
All	All	1993	0	1812	61	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 17.

The worst 5 of 61 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:B:660:LEU:HD11	1:C:549:VAL:HG21	1.66	0.76
2:F:660:LEU:O	2:F:661:GLN:HB2	1.91	0.70
1:A:570:VAL:O	1:A:574:LYS:HG3	1.91	0.70
2:D:661:GLN:HG2	2:D:662:NH2:N	2.04	0.70
2:D:660:LEU:CD1	1:E:545:ACE:H1	2.22	0.69

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	36/38 (95%)	35 (97%)	1 (3%)	0	100	100
1	C	36/38 (95%)	36 (100%)	0	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	E	36/38 (95%)	35 (97%)	0	1 (3%)	5	2
2	B	34/36 (94%)	33 (97%)	1 (3%)	0	100	100
2	D	34/36 (94%)	30 (88%)	2 (6%)	2 (6%)	1	0
2	F	34/36 (94%)	33 (97%)	0	1 (3%)	4	2
All	All	210/222 (95%)	202 (96%)	4 (2%)	4 (2%)	8	5

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	661	GLN
2	F	661	GLN
1	E	547	GLY
2	D	660	LEU

### 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	33/33 (100%)	30 (91%)	3 (9%)	9	9
1	C	33/33 (100%)	30 (91%)	3 (9%)	9	9
1	E	33/33 (100%)	29 (88%)	4 (12%)	5	4
2	B	32/32 (100%)	30 (94%)	2 (6%)	18	20
2	D	32/32 (100%)	27 (84%)	5 (16%)	2	2
2	F	32/32 (100%)	22 (69%)	10 (31%)	0	0
All	All	195/195 (100%)	168 (86%)	27 (14%)	3	3

5 of 27 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	657	MET
1	E	561	ARG
2	F	659	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	E	550	GLN
2	B	654	GLU

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	641	ASN
2	F	661	GLN
1	E	563	GLN
2	B	629	GLN
1	C	552	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

### 6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

### 6.4 Ligands

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers

EDS was not executed - this section is therefore empty.