



# Full wwPDB X-ray Structure Validation Report ⓘ

May 16, 2020 – 06:29 pm BST

PDB ID : 6OJ7  
Title : Respiratory syncytial virus fusion glycoprotein N-terminal heptad repeat domain+VIQKI I456F  
Authors : Outlaw, V.K.; Gellman, S.H.  
Deposited on : 2019-04-10  
Resolution : 1.45 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

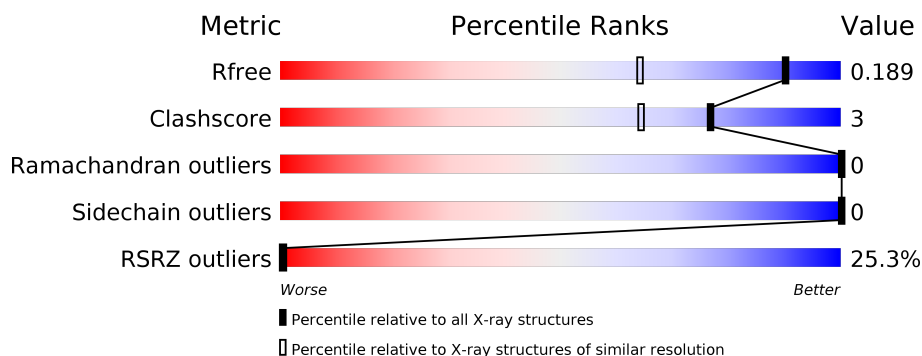
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.45 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))
$R_{free}$	130704	1156 (1.46-1.46)
Clashscore	141614	1202 (1.46-1.46)
Ramachandran outliers	138981	1178 (1.46-1.46)
Sidechain outliers	138945	1178 (1.46-1.46)
RSRZ outliers	127900	1139 (1.46-1.46)

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  $\leq 5\%$ . 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	53	<div> <div>28%</div> <div> <div></div> <div>87%</div> <div>6%</div> <div>8%</div> </div> </div>
2	C	38	<div> <div>16%</div> <div> <div></div> <div>87%</div> <div>5%</div> <div>8%</div> </div> </div>

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 1387 atoms, of which 676 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 Fusion glycoprotein F0.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	49	Total	C	H	N	O	0	1	0
			746	231	383	62	70			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	158	ACE	-	acetylation	UNP A0A1U8ZTH8
A	210	NH2	-	amidation	UNP A0A1U8ZTH8

- Molecule 2 is a protein called Fusion glycoprotein F0.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	35	Total	C	H	N	O	0	0	0
			576	182	293	46	55			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	448	ACE	-	acetylation	UNP A0A1X9QNS5
C	456	PHE	ILE	engineered mutation	UNP A0A1X9QNS5
C	459	VAL	GLU	engineered mutation	UNP A0A1X9QNS5
C	463	ILE	ALA	engineered mutation	UNP A0A1X9QNS5
C	466	GLN	ASP	engineered mutation	UNP A0A1X9QNS5
C	479	LYS	GLN	engineered mutation	UNP A0A1X9QNS5
C	480	ILE	LYS	engineered mutation	UNP A0A1X9QNS5
C	485	NH2	-	amidation	UNP A0A1X9QNS5

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	22	Total	O	0	0
			22	22		

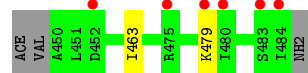
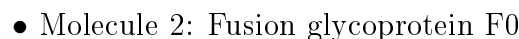
*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	43	Total	O	0	0
			43	43		



- Molecule 1: Fusion glycoprotein F0



## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	32.28Å 32.28Å 203.84Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	26.96 – 1.45 26.96 – 1.45	Depositor EDS
% Data completeness (in resolution range)	98.5 (26.96-1.45) 98.3 (26.96-1.45)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.22 (at 1.45Å)	Xtriage
Refinement program	PHENIX 1.13_2998	Depositor
R, $R_{free}$	0.179 , 0.189 0.179 , 0.189	Depositor DCC
$R_{free}$ test set	1374 reflections (9.90%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	18.3	Xtriage
Anisotropy	0.137	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.43 , 59.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.081 for -h-k,k,-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	1387	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.80% 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

### 5.1 Standard geometry

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

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	0.60	0/366	0.66	0/495
2	C	0.57	0/286	0.65	0/384
All	All	0.59	0/652	0.66	0/879

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

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	363	383	388	3	0
2	C	283	293	293	2	0
3	A	22	0	0	0	0
3	C	43	0	0	0	0
All	All	711	676	681	4	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 3.

All (4) 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:187:VAL:HG12	2:C:463:ILE:HD11	1.89	0.52
2:C:479:LYS:HA	2:C:479:LYS:HD3	1.84	0.45
1:A:201:LYS:HA	1:A:201:LYS:HD3	1.84	0.43
1:A:202:GLN:HG3	1:A:202:GLN:O	2.19	0.41

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	A	48/53 (91%)	46 (96%)	2 (4%)	0	100	100
2	C	33/38 (87%)	33 (100%)	0	0	100	100
All	All	81/91 (89%)	79 (98%)	2 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	A	42/47 (89%)	42 (100%)	0	100	100
2	C	33/35 (94%)	33 (100%)	0	100	100
All	All	75/82 (92%)	75 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.



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

### 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 ⓘ

There are no ligands in this entry.

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

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	48/53 (90%)	1.69	15 (31%) 0 0	15, 26, 73, 113	0
2	C	35/38 (92%)	0.82	6 (17%) 1 2	15, 29, 55, 59	0
All	All	83/91 (91%)	1.32	21 (25%) 0 0	15, 28, 69, 113	0

All (21) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	205	PRO	9.2
1	A	204	LEU	8.6
2	C	483	SER	5.7
1	A	159	HIS	5.5
1	A	160	LEU	5.2
1	A	203	LEU	5.2
1	A	202	GLN	4.9
2	C	479	LYS	3.9
1	A	192	VAL	3.2
1	A	201	LYS	2.9
1	A	178	VAL	2.8
1	A	188	LEU	2.8
2	C	452	ASP	2.5
1	A	206	ILE	2.5
1	A	199	ILE	2.3
2	C	475	ARG	2.3
1	A	185	VAL	2.3
1	A	195	LEU	2.3
2	C	484	ILE	2.2
2	C	480	ILE	2.1
1	A	184	GLY	2.1

## 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](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.