



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 10, 2020 – 08:20 AM BST

PDB ID : 5TDL  
Title : Crystal structure of prefusion-stabilized bovine RSV fusion glycoprotein  
(single-chain DS2-v1 variant: strain 391-2 sc9 DS-Cav1 Q98C Q361C)  
Authors : Chen, L.; Zhang, B.; Kwong, P.D.  
Deposited on : 2016-09-19  
Resolution : 3.50 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) : 1.13  
EDS : 2.13.1  
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) : 2.13.1

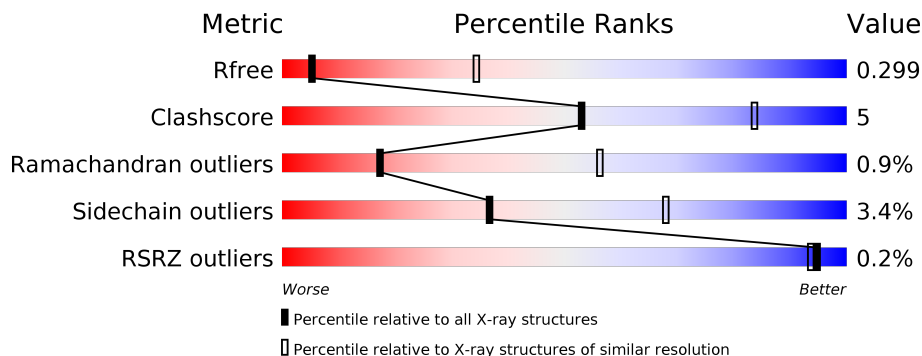
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.50 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	1659 (3.60-3.40)
Clashscore	141614	1036 (3.58-3.42)
Ramachandran outliers	138981	1005 (3.58-3.42)
Sidechain outliers	138945	1006 (3.58-3.42)
RSRZ outliers	127900	1559 (3.60-3.40)

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	531	 69% 15% 15%

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 3498 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 Bovine prefusion RSV F glycoprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	450	3484	2188	582	689	25	0	0	0

There are 36 discrepancies between the modelled and reference sequences:

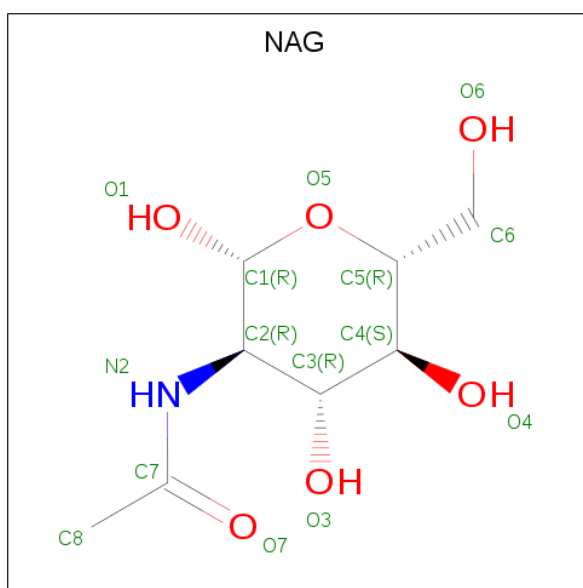
Chain	Residue	Modelled	Actual	Comment	Reference
A	98	CYS	GLN	engineered mutation	UNP P22167
A	144	SER	ILE	engineered mutation	UNP P22167
A	155	CYS	SER	engineered mutation	UNP P22167
A	190	PHE	SER	engineered mutation	UNP P22167
A	207	LEU	VAL	engineered mutation	UNP P22167
A	290	CYS	SER	engineered mutation	UNP P22167
A	361	CYS	GLN	engineered mutation	UNP P22167
A	514	SER	-	linker	UNP P22167
A	515	ALA	-	linker	UNP P22167
A	516	ILE	-	linker	UNP P22167
A	517	GLY	ALA	linker	UNP P10104
A	539	LEU	PHE	variant	UNP P10104
A	545	GLY	-	expression tag	UNP P10104
A	546	GLY	-	expression tag	UNP P10104
A	547	LEU	-	expression tag	UNP P10104
A	548	VAL	-	expression tag	UNP P10104
A	549	PRO	-	expression tag	UNP P10104
A	550	ARG	-	expression tag	UNP P10104
A	551	GLY	-	expression tag	UNP P10104
A	552	SER	-	expression tag	UNP P10104
A	553	HIS	-	expression tag	UNP P10104
A	554	HIS	-	expression tag	UNP P10104
A	555	HIS	-	expression tag	UNP P10104
A	556	HIS	-	expression tag	UNP P10104
A	557	HIS	-	expression tag	UNP P10104
A	558	HIS	-	expression tag	UNP P10104
A	559	SER	-	expression tag	UNP P10104

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	560	ALA	-	expression tag	UNP P10104
A	561	TRP	-	expression tag	UNP P10104
A	562	SER	-	expression tag	UNP P10104
A	563	HIS	-	expression tag	UNP P10104
A	564	PRO	-	expression tag	UNP P10104
A	565	GLN	-	expression tag	UNP P10104
A	566	PHE	-	expression tag	UNP P10104
A	567	GLU	-	expression tag	UNP P10104
A	568	LYS	-	expression tag	UNP P10104

- Molecule 2 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).

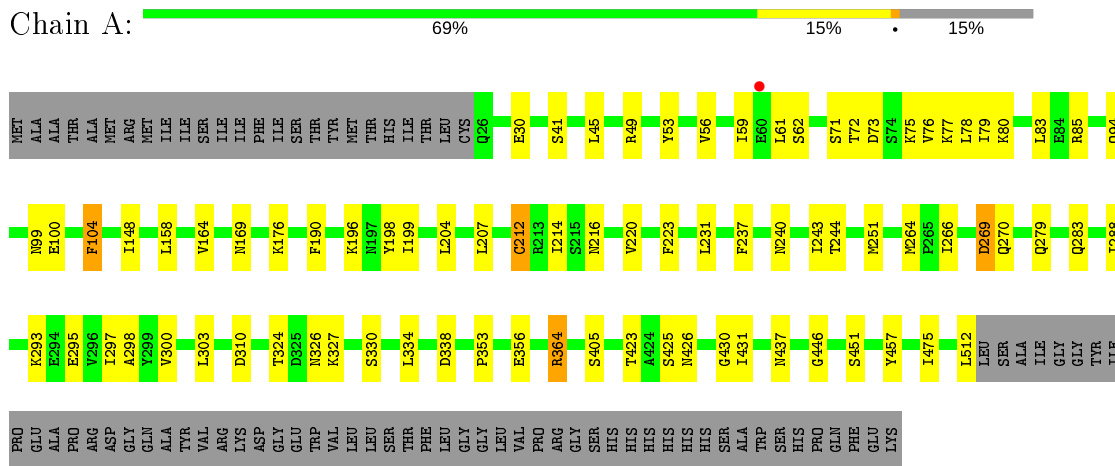


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	14	8	1	5	0	0

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

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: Bovine prefusion RSV F glycoprotein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 3 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	172.07Å 172.07Å 172.07Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.56 – 3.50 47.72 – 3.50	Depositor EDS
% Data completeness (in resolution range)	99.5 (40.56-3.50) 99.5 (47.72-3.50)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.70 (at 3.48Å)	Xtrriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.256 , 0.272 0.273 , 0.299	Depositor DCC
$R_{free}$ test set	573 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	158.9	Xtrriage
Anisotropy	0.000	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 132.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	3498	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	162.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.52% 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: 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	A	0.25	0/3535	0.43	0/4793

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	3484	0	3491	38	0
2	A	14	0	13	0	0
All	All	3498	0	3504	38	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 5.

The worst 5 of 38 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:56:VAL:HG22	1:A:300:VAL:HG22	1.72	0.70
1:A:324:THR:HG21	1:A:437:ASN:HB3	1.76	0.67

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:266:ILE:HD12	1:A:270:GLN:HB2	1.79	0.63
1:A:77:LYS:HD3	1:A:80:LYS:HD2	1.80	0.62
1:A:426:ASN:ND2	1:A:446:GLY:O	2.32	0.62

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	448/531 (84%)	393 (88%)	51 (11%)	4 (1%)	17 56

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	72	THR
1	A	338	ASP
1	A	169	ASN
1	A	353	PRO

### 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	407/476 (86%)	393 (97%)	14 (3%)	37 68

5 of 14 residues with a non-rotameric sidechain are listed below:



Mol	Chain	Res	Type
1	A	237	PHE
1	A	251	MET
1	A	356	GLU
1	A	216	ASN
1	A	327	LYS

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

### 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 monosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

1 ligand is 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	$\# Z  > 2$
2	NAG	A	5001	1	14,14,15	0.32	0	17,19,21	0.49	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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	A	5001	1	-	0/6/23/26	0/1/1/1

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	450/531 (84%)	-0.06	1 (0%) <b>95</b> <b>93</b>	100, 160, 214, 244	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	60	GLU	2.6

### 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 monosaccharides 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, 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	B-factors(Å <sup>2</sup> )	Q<0.9
2	NAG	A	5001	14/15	0.78	0.27	173,188,197,200	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.