



Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 03:38 AM GMT

PDB ID : 4MMT
Title : Crystal Structure of Prefusion-stabilized RSV F Variant DS-Cav1 at pH 9.5
Authors : Joyce, M.G.; Mclellan, J.S.; Stewart-Jones, G.B.E.; Sastry, M.; Yang, Y.;
Graham, B.S.; Kwong, P.D.
Deposited on : 2013-09-09
Resolution : 3.05 Å(reported)

This is a full wwPDB validation report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

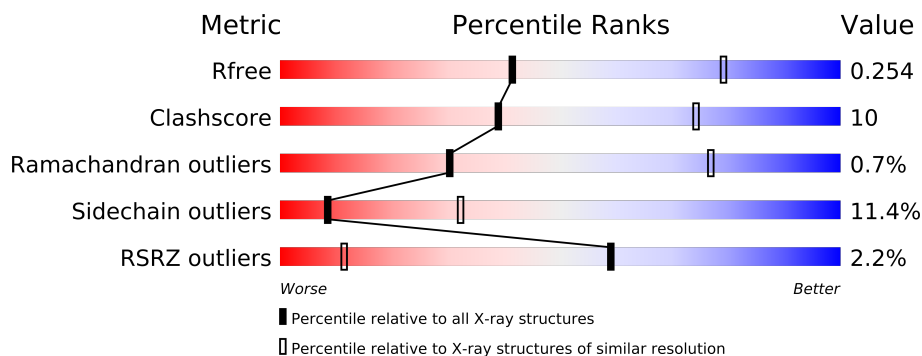
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable22639
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance

The reported resolution of this entry is 3.05 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}	66092	2079 (3.12-3.00)
Clashscore	79885	2629 (3.12-3.00)
Ramachandran outliers	78287	2536 (3.12-3.00)
Sidechain outliers	78261	2539 (3.12-3.00)
RSRZ outliers	66119	2081 (3.12-3.00)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	82	
2	B	414	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3510 atoms, of which 0 are hydrogen and 0 are deuterium.

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 F2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	73	Total	C	N	O	S	0	0	0
			580	366	95	116	3			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	102	ALA	PRO	engineered mutation	UNP P03420

- Molecule 2 is a protein called Fusion glycoprotein F1 fused with Fibrin trimerization domain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	380	Total	C	N	O	S	0	0	0
			2930	1853	483	574	20			

There are 17 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	155	CYS	SER	engineered mutation	UNP P03420
B	190	PHE	SER	engineered mutation	UNP P03420
B	207	LEU	VAL	engineered mutation	UNP P03420
B	290	CYS	SER	engineered mutation	UNP P03420
B	379	VAL	ILE	engineered mutation	UNP P03420
B	447	VAL	MET	engineered mutation	UNP P03420
B	514	SER	-	LINKER	UNP P03420
B	515	ALA	-	LINKER	UNP P03420
B	516	ILE	-	LINKER	UNP P03420
B	517	GLY	-	LINKER	UNP P03420
B	539	LEU	PHE	VARIANT	UNP P10104
B	545	GLY	-	EXPRESSION TAG	UNP P10104
B	546	GLY	-	EXPRESSION TAG	UNP P10104
B	547	LEU	-	EXPRESSION TAG	UNP P10104

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	548	VAL	-	EXPRESSION TAG	UNP P10104
B	549	PRO	-	EXPRESSION TAG	UNP P10104
B	550	ARG	-	EXPRESSION TAG	UNP P10104

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 errors 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.

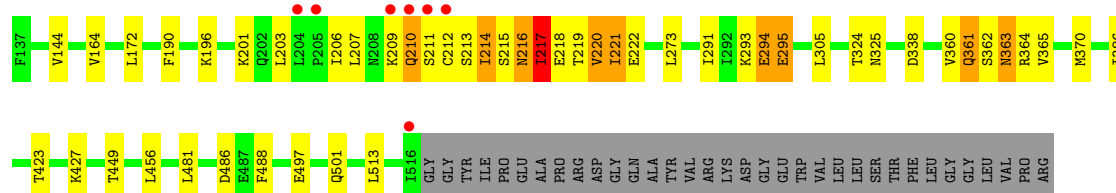
- Molecule 1: Fusion glycoprotein F2

Chain A: 



- Molecule 2: Fusion glycoprotein F1 fused with Fibrin trimerization domain

Chain B: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 41 3 2	Depositor
Cell constants a, b, c, α , β , γ	170.73 Å 170.73 Å 170.73 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.63 – 3.05 45.63 – 3.05	Depositor EDS
% Data completeness (in resolution range)	96.6 (45.63-3.05) 96.8 (45.63-3.05)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	0.14	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.00 (at 3.06 Å)	Xtriage
Refinement program	BUSTER 2.10.0	Depositor
R, R_{free}	0.231 , 0.243 0.243 , 0.254	Depositor DCC
R_{free} test set	816 reflections (5.31%)	DCC
Wilson B-factor (Å ²)	89.1	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 61.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 16192 reflections	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	3510	wwPDB-VP
Average B, all atoms (Å ²)	104.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.47% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

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	0.34	0/586	0.61	0/789
2	B	0.39	0/2974	0.61	0/4033
All	All	0.38	0/3560	0.61	0/4822

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	580	0	0	12	0
2	B	2930	0	0	25	0
All	All	3510	0	0	34	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 10.

All (34) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:79:ILE:CD1	2:B:220:VAL:CG2	2.22	1.16
2:B:361:GLN:NE2	2:B:361:GLN:CA	2.30	0.90

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:B:203:LEU:O	2:B:207:LEU:CG	2.27	0.82
2:B:196:LYS:CD	2:B:295:GLU:OE2	2.27	0.82
2:B:294:GLU:CA	2:B:294:GLU:OE2	2.30	0.80
1:A:26:GLN:CB	2:B:363:ASN:ND2	2.45	0.80
1:A:67:ASN:O	1:A:68:LYS:CD	2.30	0.79
1:A:67:ASN:O	1:A:68:LYS:CE	2.30	0.78
2:B:216:ASN:C	2:B:217:ILE:CG1	2.51	0.78
2:B:216:ASN:O	2:B:217:ILE:CG1	2.37	0.72
2:B:210:GLN:O	2:B:210:GLN:CD	2.30	0.70
1:A:67:ASN:C	1:A:68:LYS:CD	2.64	0.66
2:B:294:GLU:OE2	2:B:295:GLU:N	2.30	0.65
1:A:73:ASP:OD2	1:A:74:ALA:N	2.30	0.64
1:A:67:ASN:O	1:A:68:LYS:NZ	2.30	0.64
2:B:216:ASN:N	2:B:216:ASN:OD1	2.30	0.63
1:A:72:THR:CG2	1:A:74:ALA:O	2.48	0.62
2:B:210:GLN:NE2	2:B:210:GLN:O	2.33	0.62
2:B:210:GLN:CA	2:B:210:GLN:NE2	2.64	0.60
2:B:210:GLN:C	2:B:210:GLN:CD	2.61	0.59
1:A:72:THR:CG2	1:A:73:ASP:N	2.66	0.57
1:A:72:THR:CG2	1:A:74:ALA:N	2.67	0.57
2:B:214:ILE:CG1	2:B:215:SER:N	2.67	0.57
2:B:361:GLN:O	2:B:362:SER:CB	2.54	0.54
2:B:221:ILE:CG2	2:B:222:GLU:N	2.69	0.52
2:B:325:ASN:O	2:B:325:ASN:OD1	2.30	0.50
2:B:360:VAL:O	2:B:361:GLN:NE2	2.45	0.50
2:B:325:ASN:C	2:B:325:ASN:OD1	2.50	0.49
2:B:164:VAL:CG2	2:B:293:LYS:CD	2.91	0.48
2:B:220:VAL:CG1	2:B:221:ILE:N	2.80	0.45
2:B:360:VAL:C	2:B:361:GLN:NE2	2.71	0.44
2:B:361:GLN:NE2	2:B:361:GLN:N	2.65	0.43
1:A:68:LYS:CA	1:A:68:LYS:NZ	2.83	0.42
1:A:49:ARG:NH2	2:B:370:MET:CG	2.83	0.41

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	71/82 (87%)	66 (93%)	5 (7%)	0	100	100
2	B	378/414 (91%)	359 (95%)	16 (4%)	3 (1%)	27	74
All	All	449/496 (90%)	425 (95%)	21 (5%)	3 (1%)	30	77

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	217	ILE
2	B	338	ASP
2	B	363	ASN

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	66/73 (90%)	56 (85%)	10 (15%)	4	18
2	B	347/373 (93%)	310 (89%)	37 (11%)	10	35
All	All	413/446 (93%)	366 (89%)	47 (11%)	8	32

All (47) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	26	GLN
1	A	64	ILE
1	A	65	LYS
1	A	68	LYS
1	A	72	THR
1	A	73	ASP
1	A	79	ILE
1	A	80	LYS
1	A	83	LEU
1	A	96	LEU
2	B	144	VAL
2	B	172	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	190	PHE
2	B	201	LYS
2	B	206	ILE
2	B	209	LYS
2	B	210	GLN
2	B	211	SER
2	B	212	CYS
2	B	213	SER
2	B	214	ILE
2	B	216	ASN
2	B	217	ILE
2	B	218	GLU
2	B	219	THR
2	B	220	VAL
2	B	221	ILE
2	B	273	LEU
2	B	291	ILE
2	B	294	GLU
2	B	295	GLU
2	B	305	LEU
2	B	324	THR
2	B	361	GLN
2	B	364	ARG
2	B	365	VAL
2	B	386	ILE
2	B	423	THR
2	B	427	LYS
2	B	449	THR
2	B	456	LEU
2	B	481	LEU
2	B	486	ASP
2	B	488	PHE
2	B	497	GLU
2	B	501	GLN
2	B	513	LEU

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 chains 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, 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	A	73/82 (89%)	0.19	3 (4%) 35 6	66, 110, 181, 193	0
2	B	380/414 (91%)	-0.01	7 (1%) 65 14	60, 91, 178, 210	0
All	All	453/496 (91%)	0.03	10 (2%) 59 11	60, 92, 178, 210	0

All (10) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	210	GLN	5.0
2	B	212	CYS	4.8
2	B	204	LEU	3.3
2	B	516	ILE	2.9
1	A	66	GLU	2.9
2	B	209	LYS	2.8
2	B	205	PRO	2.6
1	A	69	CYS	2.4
2	B	211	SER	2.4
1	A	67	ASN	2.3

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.