



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 7, 2020 – 12:40 AM BST

PDB ID : 6Q0S  
Title : Crystal Structure of RSV strain B18537 Prefusion-stabilized glycoprotein F Variant DS-Cav1  
Authors : Joyce, M.G.; Bao, A.; Kwong, P.D.  
Deposited on : 2019-08-02  
Resolution : 1.94 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 : 2.13.1  
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

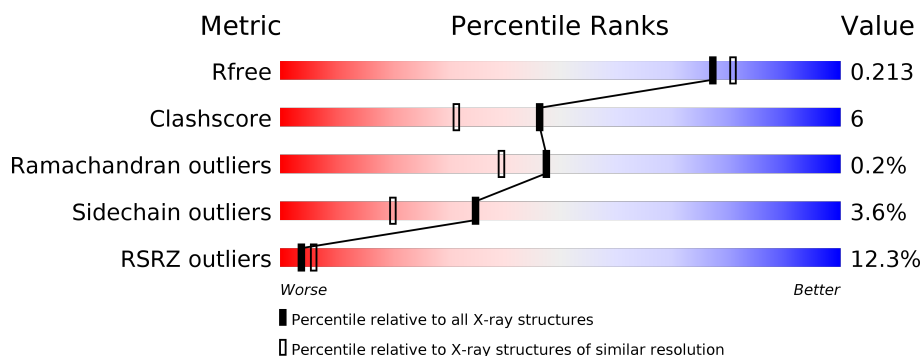
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.94 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	4310 (1.96-1.92)
Clashscore	141614	1023 (1.94-1.94)
Ramachandran outliers	138981	1007 (1.94-1.94)
Sidechain outliers	138945	1007 (1.94-1.94)
RSRZ outliers	127900	4250 (1.96-1.92)

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	F	550	<div> <div>10%</div> <div>71%</div> <div>9%</div> <div>17%</div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	NAG	F	601	-	-	-	X
2	NAG	F	602	-	-	-	X

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 7442 atoms, of which 3590 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
			Total	C	H	N	O	S			
1	F	454	7109	2237	3563	592	692	25	0	3	0

There are 41 discrepancies between the modelled and reference sequences:

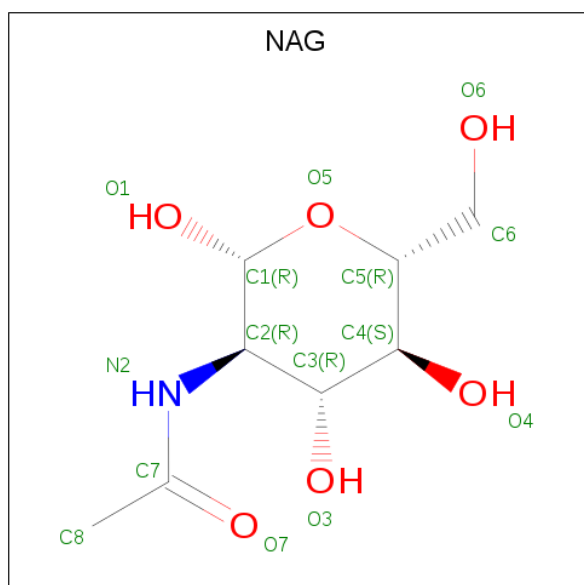
Chain	Residue	Modelled	Actual	Comment	Reference
F	155	CYS	SER	engineered mutation	UNP P13843
F	190	PHE	SER	engineered mutation	UNP P13843
F	207	LEU	VAL	engineered mutation	UNP P13843
F	290	CYS	SER	engineered mutation	UNP P13843
F	514	SER	-	expression tag	UNP P13843
F	515	ALA	-	expression tag	UNP P13843
F	516	ILE	-	expression tag	UNP P13843
F	517	GLY	-	expression tag	UNP P13843
F	518	GLY	-	expression tag	UNP P13843
F	519	TYR	-	expression tag	UNP P13843
F	520	ILE	-	expression tag	UNP P13843
F	521	PRO	-	expression tag	UNP P13843
F	522	GLU	-	expression tag	UNP P13843
F	523	ALA	-	expression tag	UNP P13843
F	524	PRO	-	expression tag	UNP P13843
F	525	ARG	-	expression tag	UNP P13843
F	526	ASP	-	expression tag	UNP P13843
F	527	GLY	-	expression tag	UNP P13843
F	528	GLN	-	expression tag	UNP P13843
F	529	ALA	-	expression tag	UNP P13843
F	530	TYR	-	expression tag	UNP P13843
F	531	VAL	-	expression tag	UNP P13843
F	532	ARG	-	expression tag	UNP P13843
F	533	LYS	-	expression tag	UNP P13843
F	534	ASP	-	expression tag	UNP P13843
F	535	GLY	-	expression tag	UNP P13843
F	536	GLU	-	expression tag	UNP P13843

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	537	TRP	-	expression tag	UNP P13843
F	538	VAL	-	expression tag	UNP P13843
F	539	LEU	-	expression tag	UNP P13843
F	540	LEU	-	expression tag	UNP P13843
F	541	SER	-	expression tag	UNP P13843
F	542	THR	-	expression tag	UNP P13843
F	543	PHE	-	expression tag	UNP P13843
F	544	LEU	-	expression tag	UNP P13843
F	545	GLY	-	expression tag	UNP P13843
F	546	GLY	-	expression tag	UNP P13843
F	547	LEU	-	expression tag	UNP P13843
F	548	VAL	-	expression tag	UNP P13843
F	549	PRO	-	expression tag	UNP P13843
F	550	ARG	-	expression tag	UNP P13843

- 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
2	F	1	Total	C	H	N	O	0	0
			28	8	14	1	5		
2	F	1	Total	C	H	N	O	0	0
			27	8	13	1	5		

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula:  $O_4S$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	F	1	Total	O	S	0	0
			5	4	1		
3	F	1	Total	O	S	0	0
			5	4	1		
3	F	1	Total	O	S	0	0
			5	4	1		

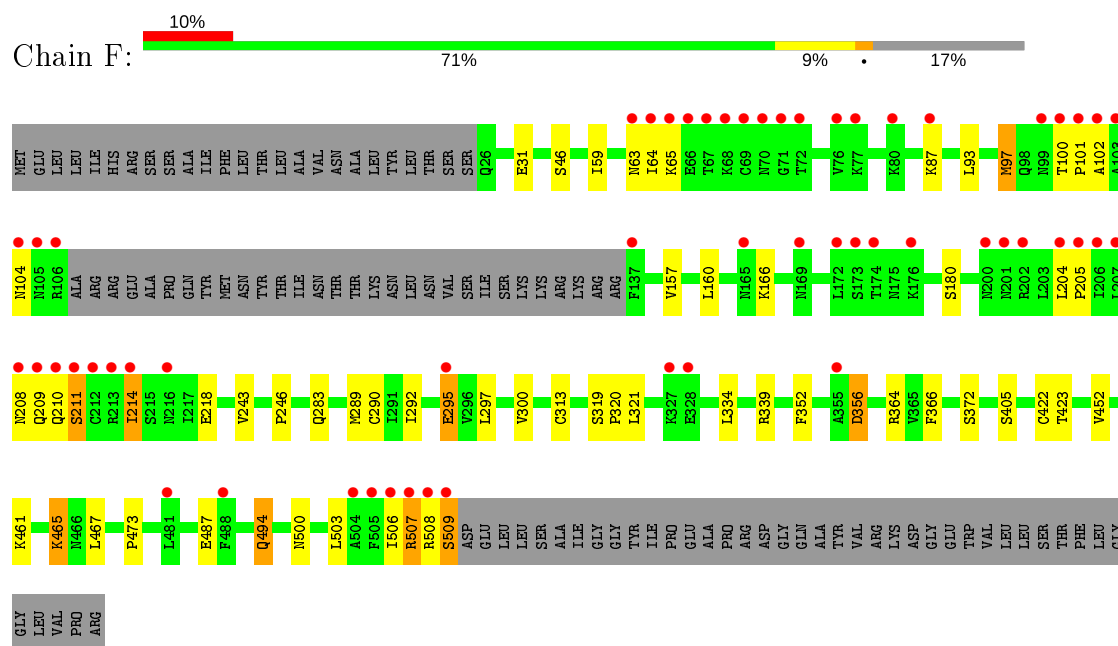
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	F	263	Total	O	0	0
			263	263		

### 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: Fusion glycoprotein F0



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 3 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	167.87Å 167.87Å 167.87Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	34.27 – 1.94 34.27 – 1.94	Depositor EDS
% Data completeness (in resolution range)	95.9 (34.27-1.94) 95.9 (34.27-1.94)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.39 (at 1.94Å)	Xtriage
Refinement program	PHENIX 1.14_3260	Depositor
R, $R_{free}$	0.187 , 0.212 0.191 , 0.213	Depositor DCC
$R_{free}$ test set	2921 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	32.2	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 66.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	7442	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.30% 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, SO4

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	F	0.82	0/3606	0.83	0/4885

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	F	3546	3563	3590	45	0
2	F	28	27	26	2	0
3	F	15	0	0	1	0
4	F	263	0	0	2	0
All	All	3852	3590	3616	45	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 6.

All (45) 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:F:63:ASN:HB3	1:F:295:GLU:OE1	1.38	1.20
1:F:100:THR:HG21	4:F:722:HOH:O	1.69	0.92
1:F:102:ALA:HB1	1:F:243:VAL:CG2	2.08	0.82
1:F:500:ASN:HD21	2:F:601:NAG:C1	1.93	0.80
1:F:295:GLU:N	1:F:295:GLU:OE2	2.21	0.74
1:F:500:ASN:ND2	2:F:601:NAG:C1	2.55	0.69
1:F:487:GLU:OE2	1:F:494:GLN:NE2	2.25	0.68
1:F:507:ARG:HH21	1:F:507:ARG:HG3	1.59	0.67
1:F:166:LYS:HE2	1:F:180:SER:O	1.95	0.66
1:F:97[B]:MET:HA	1:F:97[B]:MET:HE2	1.79	0.65
1:F:97[B]:MET:CE	1:F:97[B]:MET:HA	2.30	0.62
1:F:93:LEU:O	1:F:97[B]:MET:HG2	2.02	0.59
1:F:102:ALA:HB2	4:F:810:HOH:O	2.05	0.57
1:F:503:LEU:HD23	1:F:506:ILE:HD12	1.87	0.57
1:F:339:ARG:HD2	3:F:604:SO4:O3	2.05	0.56
1:F:97[B]:MET:CE	1:F:289:MET:HG3	2.36	0.55
1:F:102:ALA:HB1	1:F:243:VAL:HG23	1.87	0.55
1:F:214:ILE:HG23	1:F:214:ILE:O	2.06	0.55
1:F:63:ASN:CB	1:F:295:GLU:OE1	2.33	0.52
1:F:100:THR:HG23	1:F:101:PRO:HA	1.92	0.52
1:F:31:GLU:HG2	1:F:467:LEU:HB3	1.92	0.52
1:F:210:GLN:HG3	1:F:211:SER:N	2.26	0.51
1:F:97[B]:MET:HE1	1:F:289:MET:HG3	1.93	0.50
1:F:352:PHE:CE2	1:F:372:SER:HB3	2.46	0.50
1:F:290:CYS:SG	1:F:300:VAL:HG23	2.54	0.47
1:F:100:THR:HA	1:F:101:PRO:C	2.35	0.47
1:F:422:CYS:C	1:F:423:THR:HG23	2.34	0.47
1:F:59:ILE:HD12	1:F:297:LEU:HD23	1.97	0.47
1:F:46:SER:HB3	1:F:313:CYS:SG	2.55	0.46
1:F:65:LYS:HA	1:F:87:LYS:HD3	1.97	0.46
1:F:208:ASN:O	1:F:210:GLN:N	2.50	0.45
1:F:321:LEU:HD11	1:F:473:PRO:HB3	1.99	0.45
1:F:507:ARG:HD2	1:F:508:ARG:HG3	1.98	0.45
1:F:204:LEU:O	1:F:208:ASN:ND2	2.51	0.43
1:F:319:SER:OG	1:F:320:PRO:HD2	2.19	0.43
1:F:465:LYS:HD3	1:F:465:LYS:N	2.35	0.42
1:F:405:SER:HB2	1:F:452:VAL:HG21	2.00	0.42
1:F:157:VAL:HA	1:F:160:LEU:HD12	2.02	0.41
1:F:334:LEU:HD12	1:F:334:LEU:C	2.41	0.41
1:F:508:ARG:O	1:F:509:SER:C	2.59	0.41
1:F:204:LEU:N	1:F:205:PRO:CD	2.84	0.40
1:F:246:PRO:HB3	1:F:283[B]:GLN:HA	2.03	0.40

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:364:ARG:HG2	1:F:366:PHE:CE1	2.56	0.40
1:F:97[B]:MET:SD	1:F:290:CYS:O	2.78	0.40
1:F:295:GLU:CA	1:F:295:GLU:OE2	2.70	0.40

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	F	453/550 (82%)	431 (95%)	21 (5%)	1 (0%)	47 39

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	356	ASP

### 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	F	417/496 (84%)	401 (96%)	16 (4%)	33 18

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	64	ILE
1	F	97[A]	MET
1	F	97[B]	MET
1	F	104	ASN
1	F	209	GLN
1	F	211	SER
1	F	214	ILE
1	F	218	GLU
1	F	292	ILE
1	F	295	GLU
1	F	356	ASP
1	F	461	LYS
1	F	465	LYS
1	F	494	GLN
1	F	507	ARG
1	F	509	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	F	494	GLN
1	F	500	ASN

### 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 monosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

5 ligands are 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	F	601	-	14,14,15	0.65	1 (7%)	17,19,21	0.76	1 (5%)
2	NAG	F	602	1	14,14,15	0.47	0	17,19,21	0.80	0
3	SO4	F	605	-	4,4,4	0.14	0	6,6,6	0.18	0
3	SO4	F	603	-	4,4,4	0.17	0	6,6,6	1.05	0
3	SO4	F	604	-	4,4,4	0.26	0	6,6,6	0.37	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. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	F	601	-	-	2/6/23/26	0/1/1/1
2	NAG	F	602	1	-	0/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	601	NAG	O5-C1	2.28	1.47	1.43

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	601	NAG	C1-O5-C5	2.79	115.98	112.19

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	F	601	NAG	O5-C5-C6-O6
2	F	601	NAG	C4-C5-C6-O6

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	F	601	NAG	2	0
3	F	604	SO4	1	0

## 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 ⓘ

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	F	454/550 (82%)	0.58	56 (12%) 4 6	21, 38, 91, 126	0

All (56) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	70	ASN	11.2
1	F	506	ILE	10.8
1	F	508	ARG	9.4
1	F	69	CYS	9.4
1	F	509	SER	8.4
1	F	210	GLN	7.2
1	F	213	ARG	6.9
1	F	209	GLN	6.8
1	F	507	ARG	6.5
1	F	68	LYS	6.3
1	F	505	PHE	6.3
1	F	211	SER	6.2
1	F	106	ARG	6.1
1	F	72	THR	6.0
1	F	65	LYS	6.0
1	F	66	GLU	5.8
1	F	104	ASN	5.8
1	F	207	LEU	5.6
1	F	355	ALA	5.4
1	F	71	GLY	5.3
1	F	102	ALA	5.1
1	F	214	ILE	5.0
1	F	63	ASN	5.0
1	F	504	ALA	4.7
1	F	67	THR	4.5
1	F	172	LEU	4.4
1	F	105	ASN	4.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	F	64	ILE	4.1
1	F	205	PRO	4.0
1	F	212	CYS	4.0
1	F	76	VAL	3.9
1	F	204	LEU	3.8
1	F	208	ASN	3.7
1	F	100	THR	3.5
1	F	103	ALA	3.5
1	F	77	LYS	3.3
1	F	328	GLU	3.1
1	F	206	ILE	3.0
1	F	202	ARG	2.9
1	F	99	ASN	2.9
1	F	488	PHE	2.9
1	F	101	PRO	2.7
1	F	173	SER	2.7
1	F	327	LYS	2.6
1	F	80	LYS	2.6
1	F	201	ASN	2.5
1	F	295	GLU	2.5
1	F	87	LYS	2.5
1	F	169	ASN	2.4
1	F	481	LEU	2.3
1	F	165	ASN	2.3
1	F	200	ASN	2.2
1	F	216	ASN	2.2
1	F	174	THR	2.1
1	F	137	PHE	2.0
1	F	176	LYS	2.0

## 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 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	F	601	14/15	0.37	0.57	133,152,182,182	0
2	NAG	F	602	14/15	0.43	0.46	112,145,179,179	0
3	SO4	F	604	5/5	0.91	0.19	84,86,87,88	0
3	SO4	F	603	5/5	0.92	0.17	73,74,76,77	0
3	SO4	F	605	5/5	0.95	0.23	109,109,110,111	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.