



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 22, 2022 – 01:25 pm BST

PDB ID : 7ZWI  
Title : Pfs48/45 C-terminal domain bound to fab fragment of monoclonal antibody 32F3  
Authors : Ko, K.T.; Lennartz, F.L.; Higgins, M.K.  
Deposited on : 2022-05-19  
Resolution : 1.90 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 types of validation reports are described at <http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
Xtrriage (Phenix) : 1.13  
EDS : 2.29  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0267  
CCP4 : 7.1.010 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.29

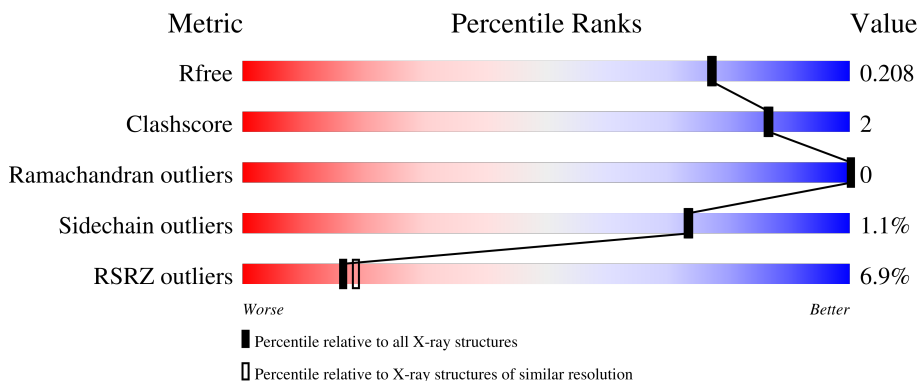
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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.



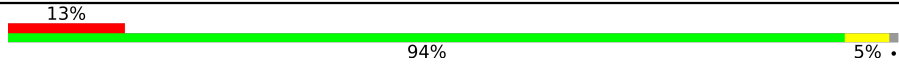
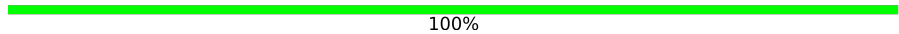
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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 of 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	429	 29% 68%
1	D	429	 31% 68%
2	B	444	 45% 52%
2	E	444	 45% 52%
3	C	213	 11% 94%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	F	213	 13% 94% 5%
4	G	3	 100%

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 9591 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 Gametocyte surface protein P45/48.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	137	1062	670	167	218	7	0	0	0
1	D	136	1055	665	166	217	7	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	429	PRO	TYR	conflict	UNP Q8I6T1
D	429	PRO	TYR	conflict	UNP Q8I6T1

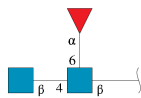
- Molecule 2 is a protein called 32F3 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	215	1636	1041	269	319	7	0	0	0
2	E	214	1631	1038	268	318	7	0	0	0

- Molecule 3 is a protein called 32F3 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	210	1597	995	271	325	6	0	0	0
3	F	211	1609	1001	273	329	6	0	0	0

- Molecule 4 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose.



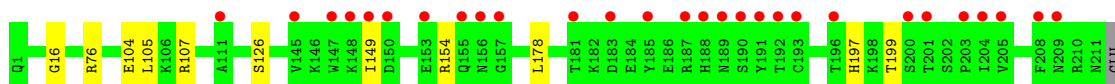
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
4	G	3	38	22	2	14	0	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	146	Total	O	0	0
			146	146		
5	B	169	Total	O	0	0
			169	169		
5	C	177	Total	O	0	0
			177	177		
5	D	158	Total	O	0	0
			158	158		
5	E	149	Total	O	0	0
			149	149		
5	F	164	Total	O	0	0
			164	164		







CYS

- Molecule 4: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain G:  100%MAG1  
MAG2  
FUCC



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	78.66Å 87.10Å 110.62Å 90.00° 99.54° 90.00°	Depositor
Resolution (Å)	77.57 – 1.90 77.57 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.8 (77.57-1.90) 99.9 (77.57-1.90)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.94 (at 1.90Å)	Xtriage
Refinement program	BUSTER 2.10.4 (20-APR-2021)	Depositor
R, $R_{free}$	0.189 , 0.217 0.183 , 0.208	Depositor DCC
$R_{free}$ test set	5743 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	27.4	Xtriage
Anisotropy	0.511	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	(Not available) , (Not available)	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	9591	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.39% 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, FUC

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	0.46	0/1085	0.66	0/1473
1	D	0.49	0/1077	0.67	0/1461
2	B	0.41	0/1678	0.60	0/2290
2	E	0.39	0/1673	0.59	0/2283
3	C	0.41	0/1636	0.59	0/2228
3	F	0.40	0/1648	0.59	0/2244
All	All	0.42	0/8797	0.61	0/11979

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	1062	0	1013	7	0
1	D	1055	0	1005	2	0
2	B	1636	0	1605	10	0
2	E	1631	0	1601	10	0
3	C	1597	0	1531	7	0
3	F	1609	0	1541	6	0
4	G	38	0	34	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	146	0	0	0	0
5	B	169	0	0	0	0
5	C	177	0	0	0	0
5	D	158	0	0	1	0
5	E	149	0	0	0	0
5	F	164	0	0	0	0
All	All	9591	0	8330	36	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 2.

All (36) 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)
3:C:197:HIS:HD2	3:C:199:THR:HG22	1.42	0.84
3:C:197:HIS:CD2	3:C:199:THR:HG22	2.13	0.83
1:A:376:ILE:HG22	1:A:406:THR:HG21	1.63	0.80
3:C:111:ALA:HA	3:C:199:THR:HG21	1.74	0.70
3:F:197:HIS:CE1	3:F:199:THR:HG23	2.33	0.64
3:C:197:HIS:HD2	3:C:199:THR:CG2	2.08	0.64
1:A:376:ILE:HG22	1:A:406:THR:CG2	2.30	0.59
1:A:357:TYR:CE1	2:B:104:LEU:HD13	2.38	0.58
3:C:149:ILE:HD11	3:C:178:LEU:HD21	1.86	0.58
3:F:16:GLY:HA2	3:F:76:ARG:HG3	1.86	0.57
3:F:149:ILE:HD11	3:F:178:LEU:HD21	1.85	0.56
3:C:16:GLY:HA2	3:C:76:ARG:HG3	1.88	0.56
2:B:126:PRO:HB3	2:B:152:TYR:HB3	1.90	0.53
2:B:214:ASP:HB2	2:E:212:LYS:HB2	1.91	0.53
2:E:199:THR:HG22	2:E:216:LYS:HE3	1.92	0.50
2:E:126:PRO:HB3	2:E:152:TYR:HB3	1.93	0.49
1:A:297:GLY:HA2	1:A:329:VAL:HG21	1.96	0.48
2:B:210:SER:HB3	2:E:216:LYS:HB3	1.94	0.48
2:B:216:LYS:HB3	2:E:210:SER:HB3	1.97	0.46
2:E:30:SER:O	2:E:53:ASN:HB2	2.15	0.46
3:F:197:HIS:CE1	3:F:199:THR:CG2	2.98	0.46
3:F:104:GLU:HG2	3:F:105:LEU:N	2.31	0.45
2:B:212:LYS:HB2	2:E:214:ASP:HB2	1.99	0.45
1:D:297:GLY:HA2	1:D:329:VAL:HG21	1.99	0.45
2:B:30:SER:O	2:B:53:ASN:HB2	2.18	0.44
2:B:14:LEU:HD23	2:B:118:VAL:HG12	2.00	0.44
1:A:324:HIS:CD2	2:B:103:ASN:HB3	2.53	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:E:14:LEU:HD23	2:E:118:VAL:HG12	2.00	0.43
1:A:346:GLY:HA3	1:A:413:LYS:O	2.19	0.43
3:F:154:ARG:HD3	3:F:178:LEU:HD11	2.01	0.43
3:C:154:ARG:HD3	3:C:178:LEU:HD11	2.01	0.42
1:A:300:PHE:HA	1:A:314:LEU:O	2.21	0.41
2:B:191:PRO:O	2:B:194:PRO:HD2	2.21	0.41
1:D:328:ASN:ND2	5:D:506:HOH:O	2.51	0.41
2:E:193:SER:O	2:E:197:SER:OG	2.39	0.41
2:E:191:PRO:O	2:E:194:PRO:HD2	2.21	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	A	135/429 (32%)	131 (97%)	4 (3%)	0	100	100
1	D	134/429 (31%)	130 (97%)	4 (3%)	0	100	100
2	B	211/444 (48%)	211 (100%)	0	0	100	100
2	E	210/444 (47%)	210 (100%)	0	0	100	100
3	C	208/213 (98%)	203 (98%)	5 (2%)	0	100	100
3	F	209/213 (98%)	205 (98%)	4 (2%)	0	100	100
All	All	1107/2172 (51%)	1090 (98%)	17 (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	126/403 (31%)	124 (98%)	2 (2%)	62	60
1	D	125/403 (31%)	123 (98%)	2 (2%)	62	60
2	B	186/398 (47%)	183 (98%)	3 (2%)	62	60
2	E	186/398 (47%)	185 (100%)	1 (0%)	88	89
3	C	181/186 (97%)	180 (99%)	1 (1%)	86	87
3	F	183/186 (98%)	181 (99%)	2 (1%)	73	73
All	All	987/1974 (50%)	976 (99%)	11 (1%)	73	73

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	342	LEU
1	A	364	LEU
2	B	104	LEU
2	B	193	SER
2	B	197	SER
3	C	126	SER
1	D	342	LEU
1	D	364	LEU
2	E	104	LEU
3	F	107	ARG
3	F	126	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	328	ASN
3	C	137	ASN
3	C	160	ASN
3	C	197	HIS
3	F	160	ASN

### 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](#)

3 monosaccharides 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
4	NAG	G	1	4,1	14,14,15	0.22	0	17,19,21	0.41	0
4	NAG	G	2	4	14,14,15	0.28	0	17,19,21	0.47	0
4	FUC	G	3	4	10,10,11	0.29	0	14,14,16	0.62	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
4	NAG	G	1	4,1	-	0/6/23/26	0/1/1/1
4	NAG	G	2	4	-	0/6/23/26	0/1/1/1
4	FUC	G	3	4	-	-	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

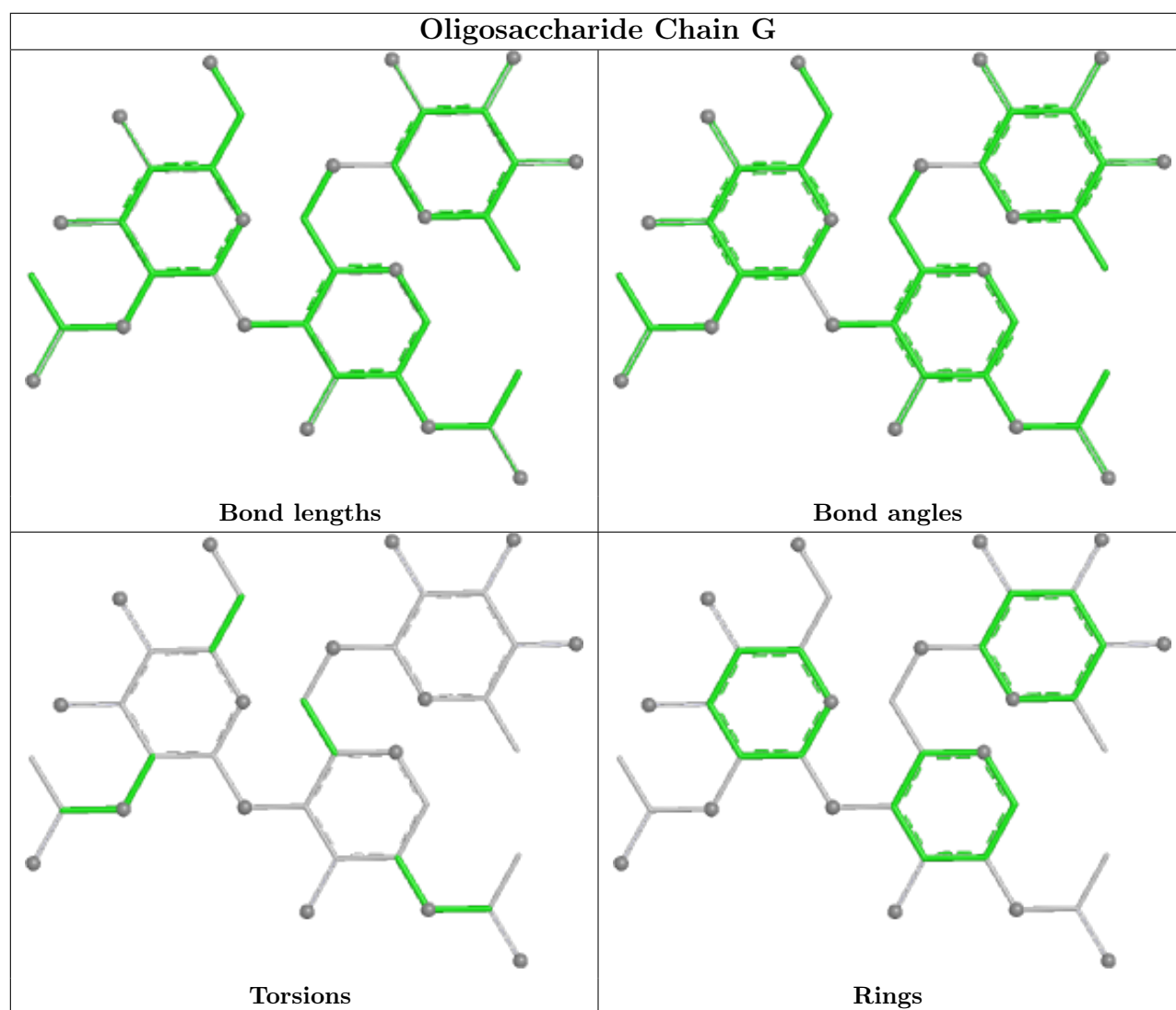
There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.



## 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

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	137/429 (31%)	0.01	4 (2%) 51 54	23, 31, 53, 64	0
1	D	136/429 (31%)	-0.09	0 100 100	21, 30, 49, 56	0
2	B	215/444 (48%)	0.27	14 (6%) 18 21	20, 37, 77, 87	0
2	E	214/444 (48%)	0.14	7 (3%) 46 49	22, 45, 69, 77	0
3	C	210/213 (98%)	0.53	24 (11%) 5 5	18, 42, 89, 99	0
3	F	211/213 (99%)	0.47	28 (13%) 3 3	21, 44, 85, 94	0
All	All	1123/2172 (51%)	0.25	77 (6%) 16 19	18, 36, 80, 99	0

All (77) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	305	SER	7.4
3	C	149	ILE	6.1
3	C	208	PHE	5.4
2	B	134	GLY	5.2
2	B	218	VAL	5.2
3	F	147	TRP	5.2
1	A	304	VAL	4.5
1	A	429	PRO	4.4
3	F	149	ILE	4.2
3	C	191	TYR	4.0
2	B	198	GLU	3.7
3	C	180	LEU	3.6
2	B	179	SER	3.5
2	E	177	LEU	3.5
3	C	190	SER	3.4
3	F	153	GLU	3.4
2	B	140	ASN	3.3
2	B	132	ALA	3.3
2	B	220	ARG	3.2

*Continued on next page...*



*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
3	F	156	ASN	3.2
3	F	187	ARG	3.2
3	C	193	CYS	3.2
3	F	205	VAL	3.1
3	C	207	SER	3.1
3	F	200	SER	3.1
2	E	194	PRO	3.1
3	F	150	ASP	3.0
3	C	155	GLN	3.0
3	F	201	THR	3.0
3	F	188	HIS	3.0
2	B	143	VAL	3.0
3	C	186	GLU	2.9
3	F	209	ASN	2.9
3	C	210	ARG	2.9
3	F	185	TYR	2.9
3	F	190	SER	2.9
2	B	197	SER	2.8
3	F	192	THR	2.8
3	C	147	TRP	2.8
3	C	121	SER	2.8
3	C	184	GLU	2.7
3	F	157	GLY	2.7
2	B	133	PRO	2.7
3	C	150	ASP	2.7
2	B	194	PRO	2.7
3	C	156	ASN	2.7
3	F	111	ALA	2.6
2	B	196	PRO	2.6
3	C	205	VAL	2.6
3	F	196	THR	2.6
3	F	181	THR	2.6
2	E	180	ASP	2.6
3	F	204	ILE	2.6
3	C	126	SER	2.5
3	F	155	GLN	2.5
3	F	189	ASN	2.4
2	E	120	SER	2.3
3	C	201	THR	2.3
3	F	208	PHE	2.3
3	C	187	ARG	2.3
3	F	148	LYS	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
3	F	191	TYR	2.3
3	C	209	ASN	2.3
3	F	193	CYS	2.3
2	E	166	LEU	2.2
3	C	157	GLY	2.2
3	F	183	ASP	2.2
3	F	203	PRO	2.2
3	F	145	VAL	2.2
2	E	179	SER	2.1
2	B	192	SER	2.1
3	C	202	SER	2.1
2	B	191	PRO	2.1
3	C	127	GLY	2.1
3	C	128	GLY	2.1
2	E	192	SER	2.0
1	A	428	ALA	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

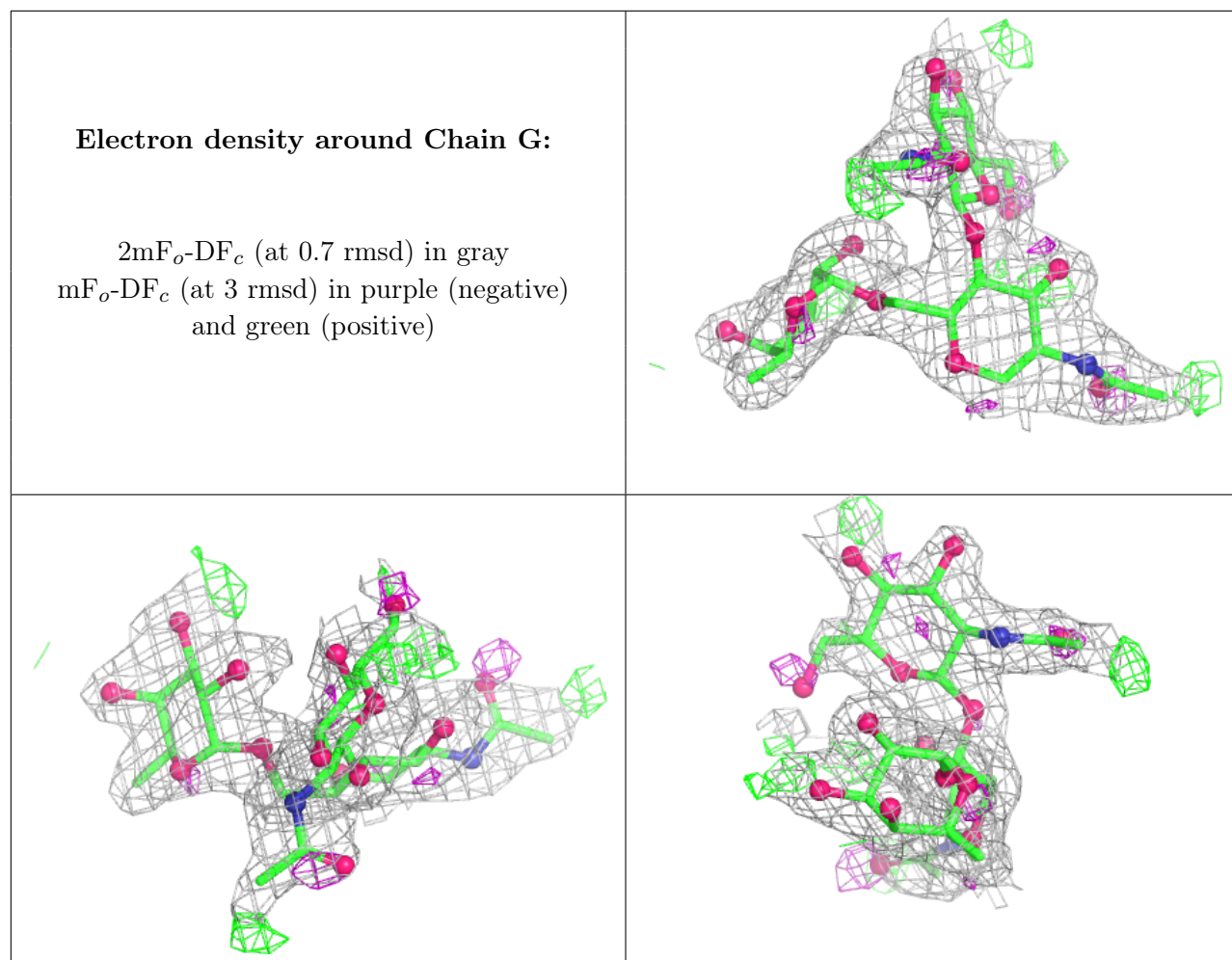
There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates [i](#)

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
4	NAG	G	2	14/15	0.74	0.21	56,58,59,60	0
4	NAG	G	1	14/15	0.88	0.10	47,53,55,55	0
4	FUC	G	3	10/11	0.92	0.11	45,46,46,46	0

The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.



## 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.