Full wwPDB X-ray Structure Validation Report (i)

Jan 18, 2022 – 06:31 pm GMT

PDB ID	:	6S2I
Title	:	Anti-tumor antibody 14F7-derived scFv in complex with NeuGc Gm3
Authors	:	Bjerregaard-Andersen, K.; Heggelund, J.E.; Krengel, U.
Deposited on	:	2019-06-20
Resolution	:	2.29 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

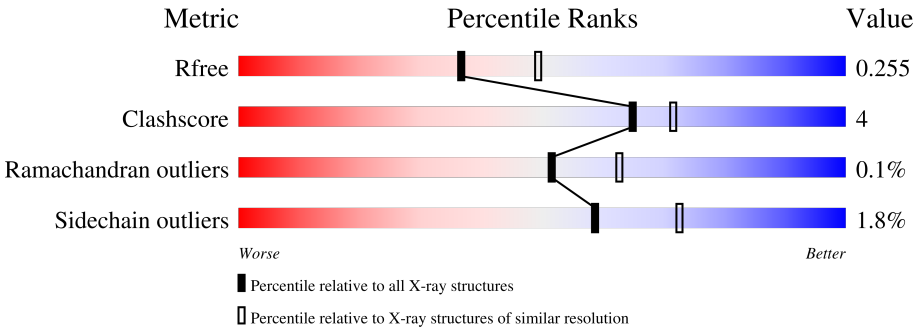
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.24
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.24

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 2.29 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	6980 (2.30-2.26)
Clashscore	141614	7711 (2.30-2.26)
Ramachandran outliers	138981	7597 (2.30-2.26)
Sidechain outliers	138945	7598 (2.30-2.26)

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 >=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 <=5%

Mol	Chain	Length	Quality of chain
1	A	253	43% 6% 51%
1	B	253	37% 5% 58%
1	C	253	41% 8% 51%
1	D	253	38% 5% 57%
1	E	253	41% 7% 52%
1	F	253	40% • 57%
1	G	253	42% 6% 52%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	H	253	36% 7% 57%
2	J	3	100%

2 Entry composition (i)

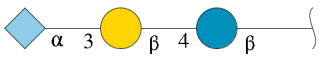
There are 3 unique types of molecules in this entry. The entry contains 7397 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 scFv C1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	125	Total	C	N	O	S	0	0	0
			1000	632	169	193	6			
1	B	106	Total	C	N	O	S	0	1	0
			817	509	136	169	3			
1	C	125	Total	C	N	O	S	0	1	0
			1003	634	169	194	6			
1	D	108	Total	C	N	O	S	0	0	0
			831	518	139	171	3			
1	E	121	Total	C	N	O	S	0	0	0
			966	612	159	189	6			
1	F	108	Total	C	N	O	S	0	0	0
			831	518	139	171	3			
1	G	121	Total	C	N	O	S	0	1	0
			971	615	160	190	6			
1	H	108	Total	C	N	O	S	0	0	0
			830	518	139	170	3			

- Molecule 2 is an oligosaccharide called N-glycolyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose-(1-4)-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	J	3	Total	C	N	O	0	0	0
			44	23	1	20			

- Molecule 3 is water.

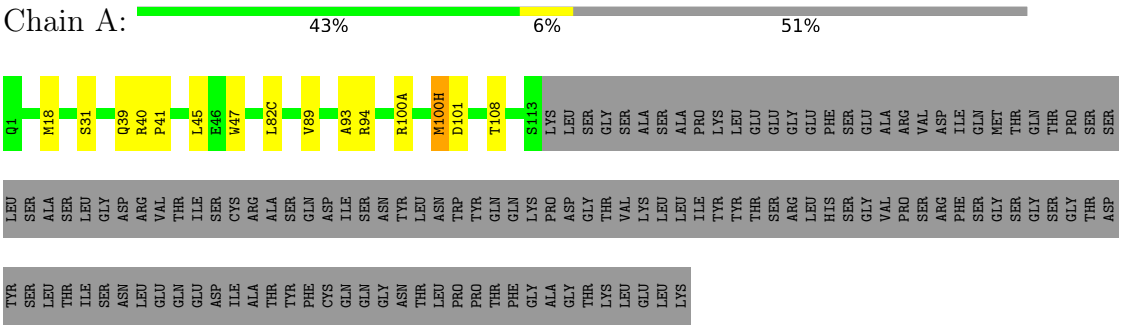
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	27	Total O 27 27	0	0
3	B	13	Total O 13 13	0	0
3	C	22	Total O 22 22	0	0
3	D	8	Total O 8 8	0	0
3	E	11	Total O 11 11	0	0
3	F	13	Total O 13 13	0	0
3	G	2	Total O 2 2	0	0
3	H	8	Total O 8 8	0	0

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. 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. 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.

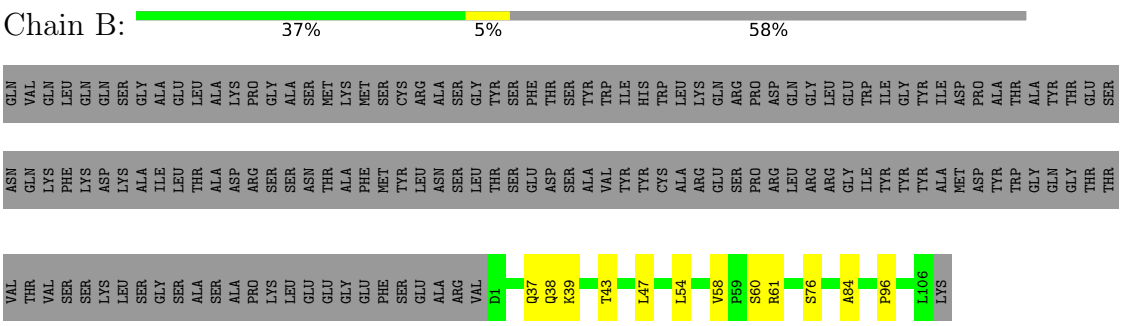
- Molecule 1: scFv C1

Chain A:



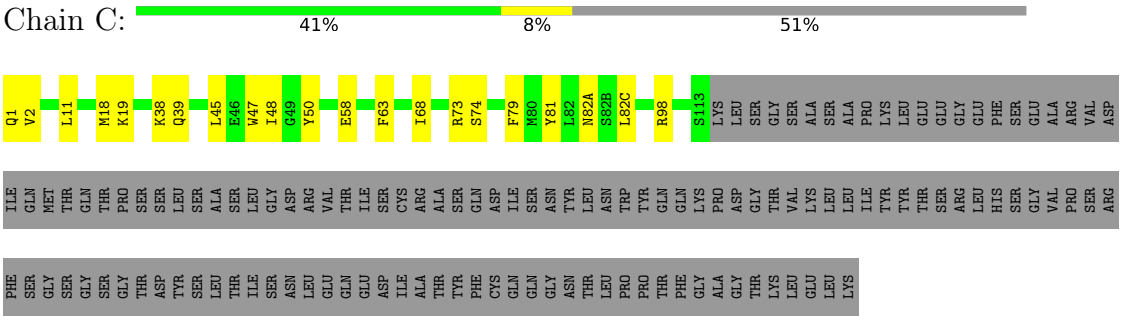
- Molecule 1: scFv C1

Chain B:



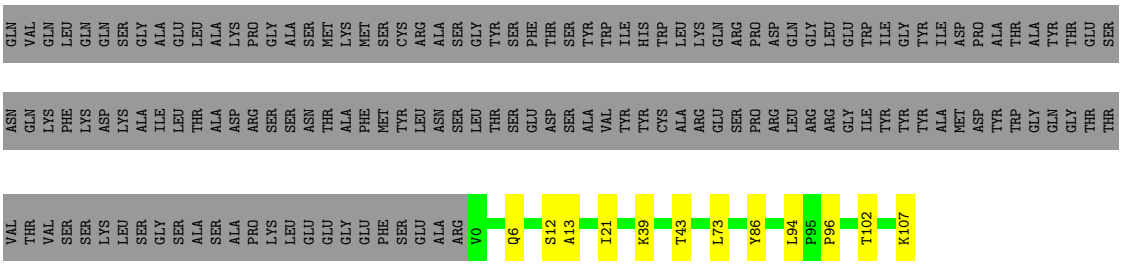
- Molecule 1: scFv C1

Chain C:



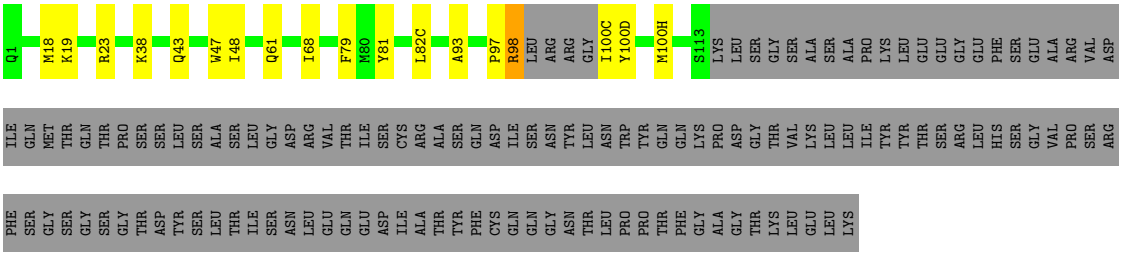
- Molecule 1: scFv C1

Chain D:



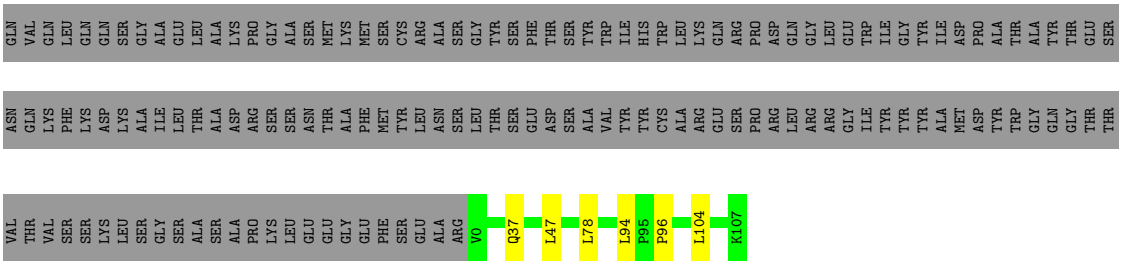
• Molecule 1: scFv C1

Chain E:



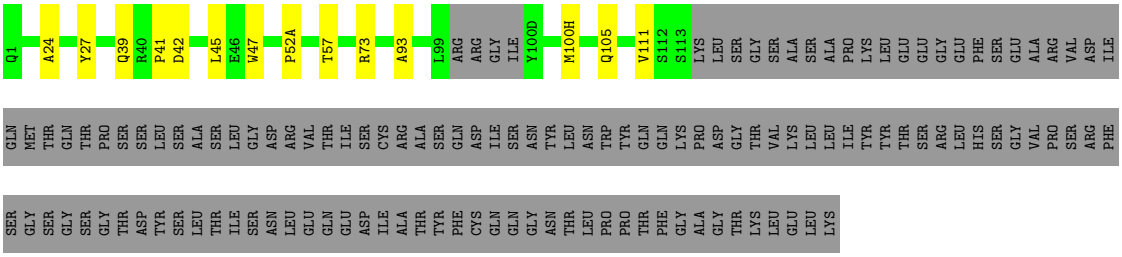
• Molecule 1: scFv C1

Chain F:



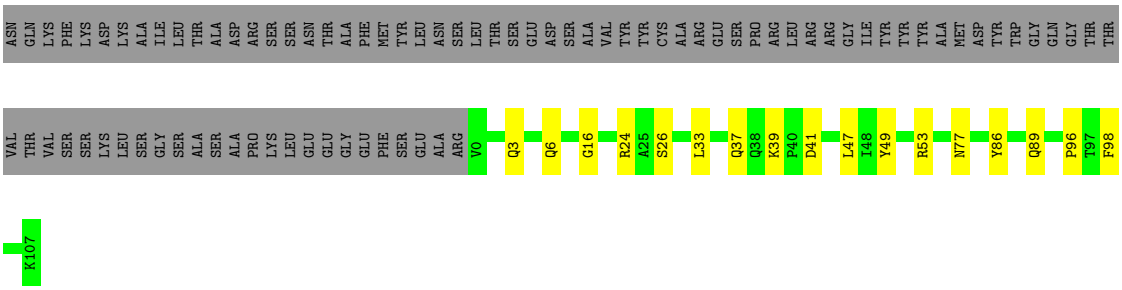
• Molecule 1: scFv C1

Chain G:



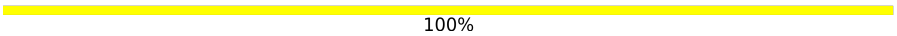
• Molecule 1: scFv C1

Chain H:



• Molecule 2: N-glycolyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose-(1-4)-beta-D-glucopyranose

Chain J:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	62.88Å 113.69Å 66.95Å 90.00° 91.24° 90.00°	Depositor
Resolution (Å)	62.87 – 2.29 62.87 – 2.29	Depositor EDS
% Data completeness (in resolution range)	98.7 (62.87-2.29) 98.8 (62.87-2.29)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.44 (at 2.29Å)	Xtriage
Refinement program	PHENIX (1.13_2998: ???)	Depositor
R, R_{free}	0.222 , 0.255 0.222 , 0.255	Depositor DCC
R_{free} test set	2152 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å^2)	36.5	Xtriage
Anisotropy	0.388	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.000 for h,-k,-l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	7397	wwPDB-VP
Average B, all atoms (Å^2)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.91% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: GAL, NGC, BGC

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	0/1026	0.47	0/1390
1	B	0.25	0/836	0.47	0/1136
1	C	0.25	0/1032	0.45	0/1398
1	D	0.24	0/847	0.46	0/1149
1	E	0.27	0/991	0.48	0/1343
1	F	0.24	0/847	0.46	0/1149
1	G	0.27	0/999	0.49	0/1354
1	H	0.24	0/846	0.46	0/1149
All	All	0.25	0/7424	0.47	0/10068

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	1000	0	953	8	0
1	B	817	0	791	8	0
1	C	1003	0	958	13	0
1	D	831	0	808	9	0
1	E	966	0	912	10	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	F	831	0	808	4	0
1	G	971	0	918	6	0
1	H	830	0	808	8	0
2	J	44	0	28	0	0
3	A	27	0	0	0	0
3	B	13	0	0	0	0
3	C	22	0	0	2	0
3	D	8	0	0	0	0
3	E	11	0	0	0	0
3	F	13	0	0	0	0
3	G	2	0	0	0	0
3	H	8	0	0	0	0
All	All	7397	0	6984	59	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 4.

All (59) 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:B:37:GLN:HB2	1:B:47:LEU:HD11	1.79	0.64
1:A:89:VAL:HG22	1:A:108:THR:HG22	1.82	0.61
1:F:37:GLN:HB2	1:F:47:LEU:HD11	1.83	0.60
1:C:38:LYS:HB2	1:C:48:ILE:HD11	1.86	0.58
1:D:21:ILE:HD11	1:D:73:LEU:HD23	1.86	0.57
1:F:94:LEU:HD23	1:F:96:PRO:HD3	1.87	0.57
1:A:93:ALA:HB1	1:A:100(H):MET:HB3	1.88	0.56
1:G:52(A):PRO:O	1:G:73:ARG:NH1	2.39	0.55
1:H:6:GLN:NE2	1:H:86:TYR:O	2.35	0.55
1:E:47:TRP:CD2	1:F:96:PRO:HD2	2.43	0.54
1:C:39:GLN:HB2	1:C:45:LEU:HD23	1.89	0.53
1:C:1:GLN:OE1	1:C:1:GLN:N	2.40	0.53
1:C:47:TRP:CD2	1:D:96:PRO:HD2	2.44	0.52
1:C:98:ARG:NH1	3:C:301:HOH:O	2.37	0.50
1:A:39:GLN:HB2	1:A:45:LEU:HD23	1.93	0.50
1:D:6:GLN:NE2	1:D:86:TYR:O	2.45	0.49
1:E:98:ARG:HB2	1:E:100(C):ILE:O	2.12	0.49
1:A:94:ARG:NH1	1:A:101:ASP:OD2	2.46	0.49
1:G:105:GLN:OE1	1:G:105:GLN:N	2.41	0.48
1:E:19:LYS:HE2	1:E:79:PHE:CD1	2.48	0.48
1:A:47:TRP:CD2	1:B:96:PRO:HD2	2.48	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:40:ARG:HB3	1:A:41:PRO:HD2	1.94	0.47
1:C:68:ILE:HB	1:C:81:TYR:HB2	1.97	0.47
1:H:39:LYS:HD2	1:H:41:ASP:CG	2.35	0.47
1:D:39:LYS:HB2	1:D:43:THR:HG22	1.96	0.47
1:B:54:LEU:HD11	1:B:60:SER:HA	1.95	0.47
1:E:38:LYS:HB2	1:E:48:ILE:HD11	1.97	0.47
1:E:97:PRO:O	1:E:98:ARG:HG2	2.15	0.47
1:B:39:LYS:HB2	1:B:43:THR:HG22	1.97	0.47
1:E:18:MET:HG2	1:E:82(C):LEU:HD11	1.97	0.47
1:C:48:ILE:HG23	1:C:63:PHE:CG	2.51	0.46
1:F:78:LEU:HD11	1:F:104:LEU:HD21	1.97	0.46
1:C:19:LYS:HE3	1:C:79:PHE:CD1	2.50	0.46
1:H:37:GLN:HB2	1:H:47:LEU:HD11	1.97	0.46
1:C:58:GLU:HG2	1:D:94:LEU:HD13	1.98	0.45
1:H:3:GLN:HB2	1:H:26:SER:HB3	1.98	0.45
1:E:68:ILE:HB	1:E:81:TYR:HB2	1.97	0.45
1:G:93:ALA:HB1	1:G:100(H):MET:HB3	1.98	0.45
1:B:54:LEU:HD22	1:B:58:VAL:HB	1.98	0.45
1:A:18:MET:HG2	1:A:82(C):LEU:HD11	1.99	0.44
1:E:93:ALA:HB1	1:E:100(H):MET:HB3	1.98	0.44
1:B:61:ARG:HB2	1:B:76:SER:O	2.17	0.44
1:D:21:ILE:HG21	1:D:102:THR:HG21	2.00	0.44
1:G:39:GLN:HB2	1:G:45:LEU:HG	2.00	0.44
1:A:31:SER:OG	1:E:23:ARG:NH1	2.51	0.43
1:C:82(A):ASN:ND2	3:C:304:HOH:O	2.51	0.43
1:D:13:ALA:O	1:D:107:LYS:N	2.52	0.42
1:E:98:ARG:HD2	1:E:100(D):TYR:CE2	2.54	0.42
1:C:47:TRP:HZ2	1:C:50:TYR:HD2	1.66	0.42
1:H:16:GLY:HA2	1:H:77:ASN:OD1	2.19	0.42
1:G:24:ALA:HB1	1:G:27:TYR:CE1	2.56	0.41
1:H:49:TYR:O	1:H:53:ARG:HB2	2.20	0.41
1:C:18:MET:HG2	1:C:82(C):LEU:HD11	2.01	0.41
1:B:38:GLN:O	1:B:84:ALA:HB1	2.20	0.41
1:H:89:GLN:HB2	1:H:98:PHE:CD2	2.56	0.41
1:B:54:LEU:HD23	1:B:54:LEU:HA	1.92	0.40
1:D:12:SER:HB2	1:D:107:LYS:OXT	2.21	0.40
1:C:47:TRP:CE3	1:D:96:PRO:HD2	2.56	0.40
1:G:47:TRP:CD2	1:H:96:PRO:HD2	2.56	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	123/253 (49%)	123 (100%)	0	0	100	100
1	B	105/253 (42%)	100 (95%)	5 (5%)	0	100	100
1	C	124/253 (49%)	123 (99%)	1 (1%)	0	100	100
1	D	106/253 (42%)	103 (97%)	3 (3%)	0	100	100
1	E	117/253 (46%)	113 (97%)	4 (3%)	0	100	100
1	F	106/253 (42%)	101 (95%)	5 (5%)	0	100	100
1	G	118/253 (47%)	112 (95%)	5 (4%)	1 (1%)	19	22
1	H	106/253 (42%)	102 (96%)	4 (4%)	0	100	100
All	All	905/2024 (45%)	877 (97%)	27 (3%)	1 (0%)	51	63

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	41	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	105/215 (49%)	103 (98%)	2 (2%)	57	71
1	B	94/215 (44%)	94 (100%)	0	100	100
1	C	106/215 (49%)	102 (96%)	4 (4%)	33	44
1	D	95/215 (44%)	95 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	E	102/215 (47%)	99 (97%)	3 (3%)	42 56
1	F	95/215 (44%)	95 (100%)	0	100 100
1	G	103/215 (48%)	100 (97%)	3 (3%)	42 56
1	H	95/215 (44%)	93 (98%)	2 (2%)	53 68
All	All	795/1720 (46%)	781 (98%)	14 (2%)	59 72

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	100(A)	ARG
1	A	100(H)	MET
1	C	2	VAL
1	C	11	LEU
1	C	73	ARG
1	C	74	SER
1	E	43	GLN
1	E	61	GLN
1	E	98	ARG
1	G	42	ASP
1	G	57	THR
1	G	111	VAL
1	H	24	ARG
1	H	33	LEU

Sometimes 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)

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	$\#\|Z\| > 2$
2	BGC	J	1	2	12,12,12	1.27	1 (8%)	17,17,17	1.10	1 (5%)
2	GAL	J	2	2	11,11,12	1.63	3 (27%)	15,15,17	1.90	3 (20%)
2	NGC	J	3	2	18,21,22	1.85	5 (27%)	22,29,32	1.29	3 (13%)

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	BGC	J	1	2	-	2/2/22/22	0/1/1/1
2	GAL	J	2	2	-	1/2/19/22	0/1/1/1
2	NGC	J	3	2	-	5/15/36/40	0/1/1/1

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	J	3	NGC	C7-C6	4.48	1.58	1.53
2	J	3	NGC	C8-C7	2.50	1.58	1.53
2	J	1	BGC	C4-C5	2.47	1.58	1.53
2	J	2	GAL	C2-C3	2.45	1.56	1.52
2	J	3	NGC	C3-C2	2.41	1.56	1.52
2	J	3	NGC	C6-C5	2.23	1.56	1.53
2	J	2	GAL	O5-C1	-2.15	1.40	1.43
2	J	3	NGC	C10-N5	2.09	1.38	1.34
2	J	2	GAL	C4-C5	2.06	1.57	1.53

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	J	2	GAL	C1-O5-C5	4.43	118.19	112.19
2	J	3	NGC	O11-C11-C10	-3.12	105.00	111.62
2	J	2	GAL	O5-C5-C6	3.07	112.01	107.20

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	J	2	GAL	C1-C2-C3	2.83	113.15	109.67
2	J	3	NGC	C6-O6-C2	2.30	116.25	111.34
2	J	1	BGC	O3-C3-C4	-2.16	105.36	110.35
2	J	3	NGC	C11-C10-N5	2.08	119.99	116.37

There are no chirality outliers.

All (8) torsion outliers are listed below:

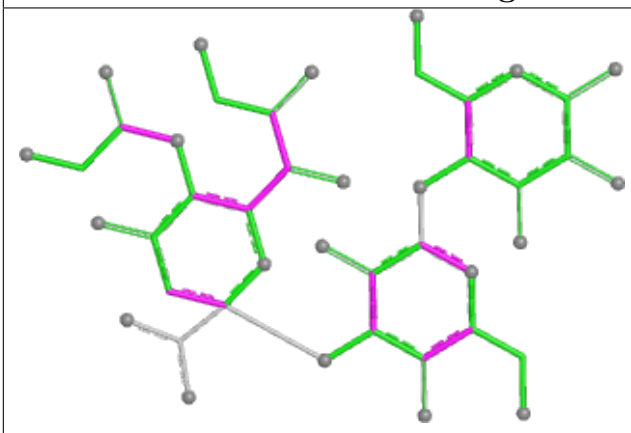
Mol	Chain	Res	Type	Atoms
2	J	1	BGC	C4-C5-C6-O6
2	J	1	BGC	O5-C5-C6-O6
2	J	3	NGC	O7-C7-C8-C9
2	J	3	NGC	C6-C7-C8-C9
2	J	2	GAL	O5-C5-C6-O6
2	J	3	NGC	C6-C7-C8-O8
2	J	3	NGC	O10-C10-C11-O11
2	J	3	NGC	O7-C7-C8-O8

There are no ring outliers.

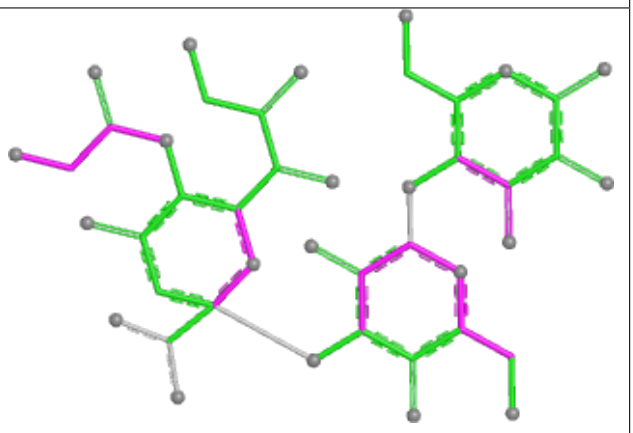
No monomer is involved in short contacts.

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

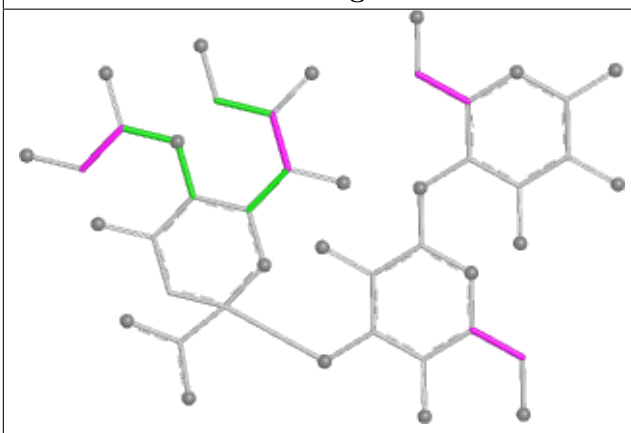
Oligosaccharide Chain J



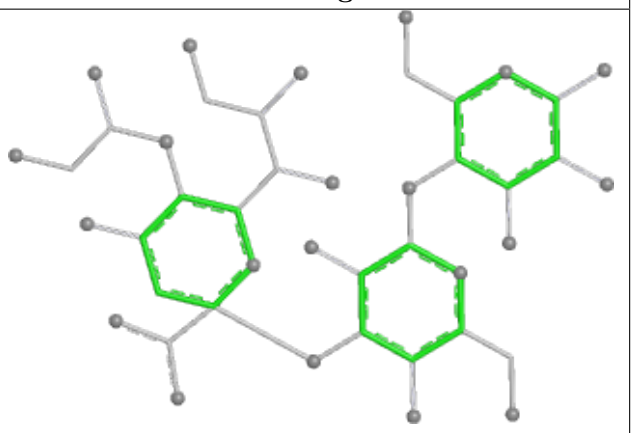
Bond lengths



Bond angles



Torsions



Rings

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

Unable to reproduce the depositors R factor - this section is therefore empty.

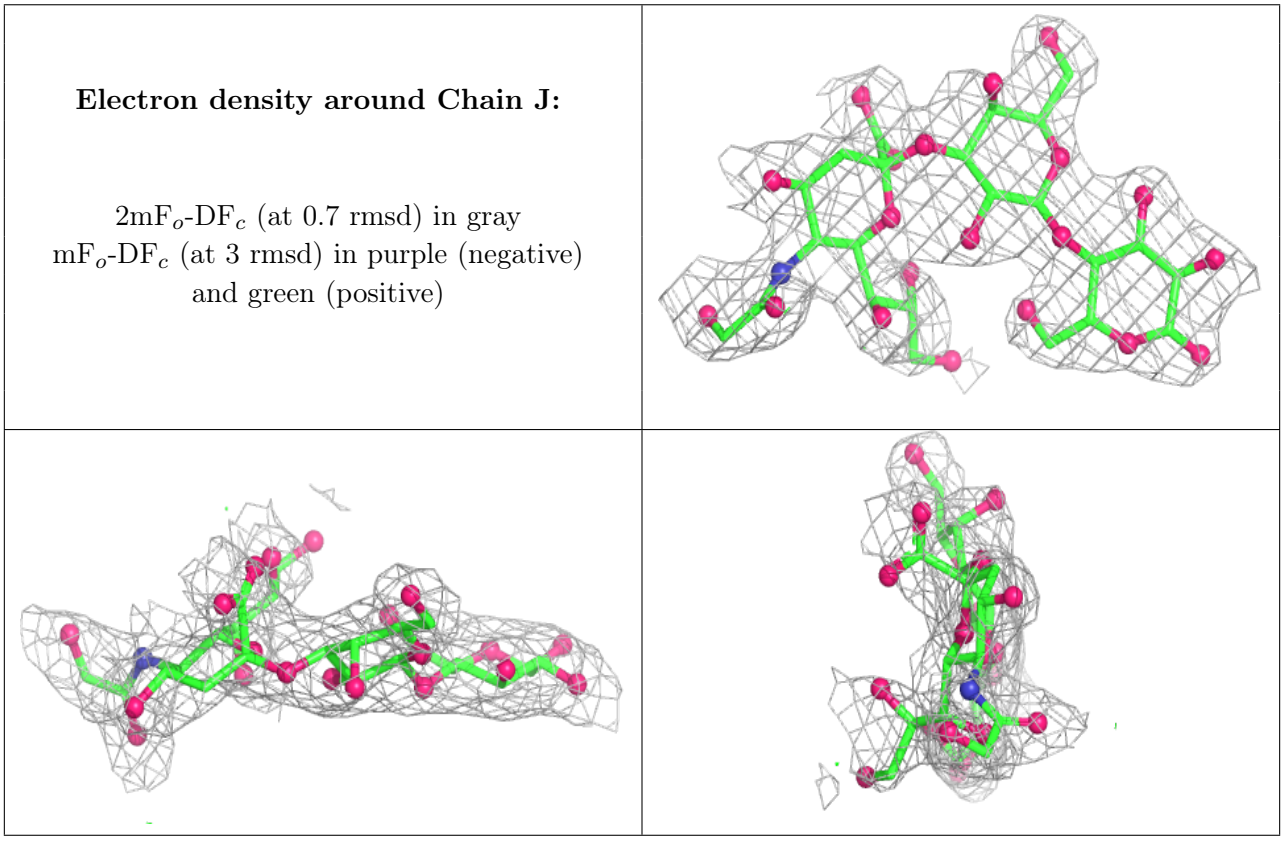
6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers (i)

Unable to reproduce the depositors R factor - this section is therefore empty.