wwPDB EM Validation Summary Report (i)

May 23, 2024 – 09:54 AM EDT

PDB ID	:	3J2Q
Title	:	Model of membrane-bound factor VIII organized in 2D crystals
Authors	:	Stoilova-Mcphie, S.; Lynch, G.C.; Ludtke, S.; Pettitt, B.M.
Deposited on	:	2012-12-11
Resolution	:	15.00 Å(reported)

This is a wwPDB EM Validation Summary Report for a publicly released PDB/EMDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
with specific help available everywhere you see the (i) 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 (1)) were used in the production of this report:

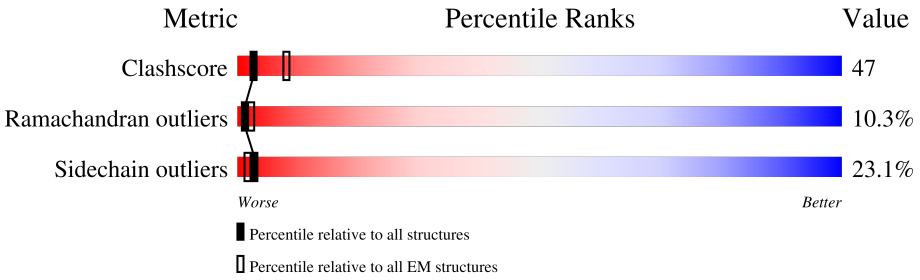
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance

The following experimental techniques were used to determine the structure:
ELECTRON CRYSTALLOGRAPHY

The reported resolution of this entry is 15.00 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)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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	754	15% 47% 19% · 16%
2	B	684	21% 51% 17% · 8%
3	C	2	50% 50%
4	D	4	100%

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
4	BMA	D	2	-	-	X	-
4	BMA	D	4	-	-	X	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
7	NAG	B	2401	X	-	-	-
7	NAG	B	2402	X	-	-	-

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 10317 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 Coagulation factor VIII heavy chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	630	5086	3278	855	928	25	0	0

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	746	PRO	-	expression tag	UNP P00451
A	747	PRO	-	expression tag	UNP P00451
A	748	VAL	-	expression tag	UNP P00451
A	749	LEU	-	expression tag	UNP P00451
A	750	LYS	-	expression tag	UNP P00451
A	751	ARG	-	expression tag	UNP P00451
A	752	HIS	-	expression tag	UNP P00451
A	753	GLN	-	expression tag	UNP P00451
A	754	ARG	-	expression tag	UNP P00451

- Molecule 2 is a protein called Coagulation factor VIII light chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	631	5125	3288	881	924	32	0	0

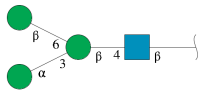
There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1880	LEU	PHE	SEE REMARK 999	UNP P00451

- Molecule 3 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms				AltConf	Trace
3	C	2	Total	C	N	O	0	0
			28	16	2	10		

- Molecule 4 is an oligosaccharide called alpha-D-mannopyranose-(1-3)-[beta-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				AltConf	Trace
4	D	4	Total	C	N	O	0	0
			47	26	1	20		

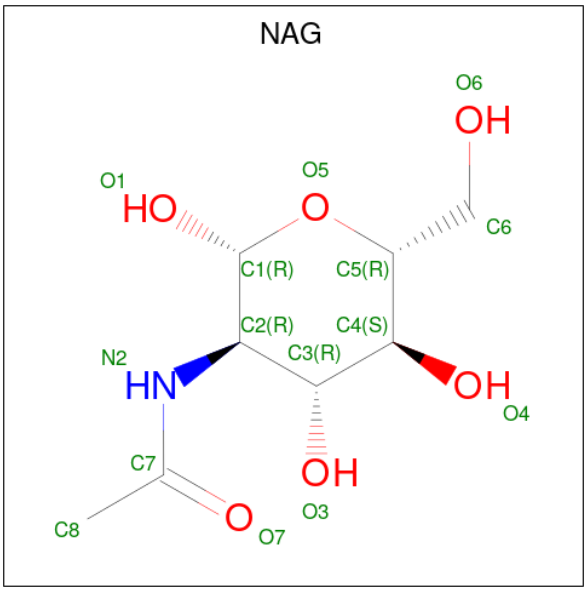
- Molecule 5 is COPPER (II) ION (three-letter code: CU) (formula: Cu).

Mol	Chain	Residues	Atoms		AltConf
5	A	1	Total	Cu	0
			1	1	
5	B	1	Total	Cu	0
			1	1	

- Molecule 6 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		AltConf
6	A	1	Total	Ca	0
			1	1	

- Molecule 7 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



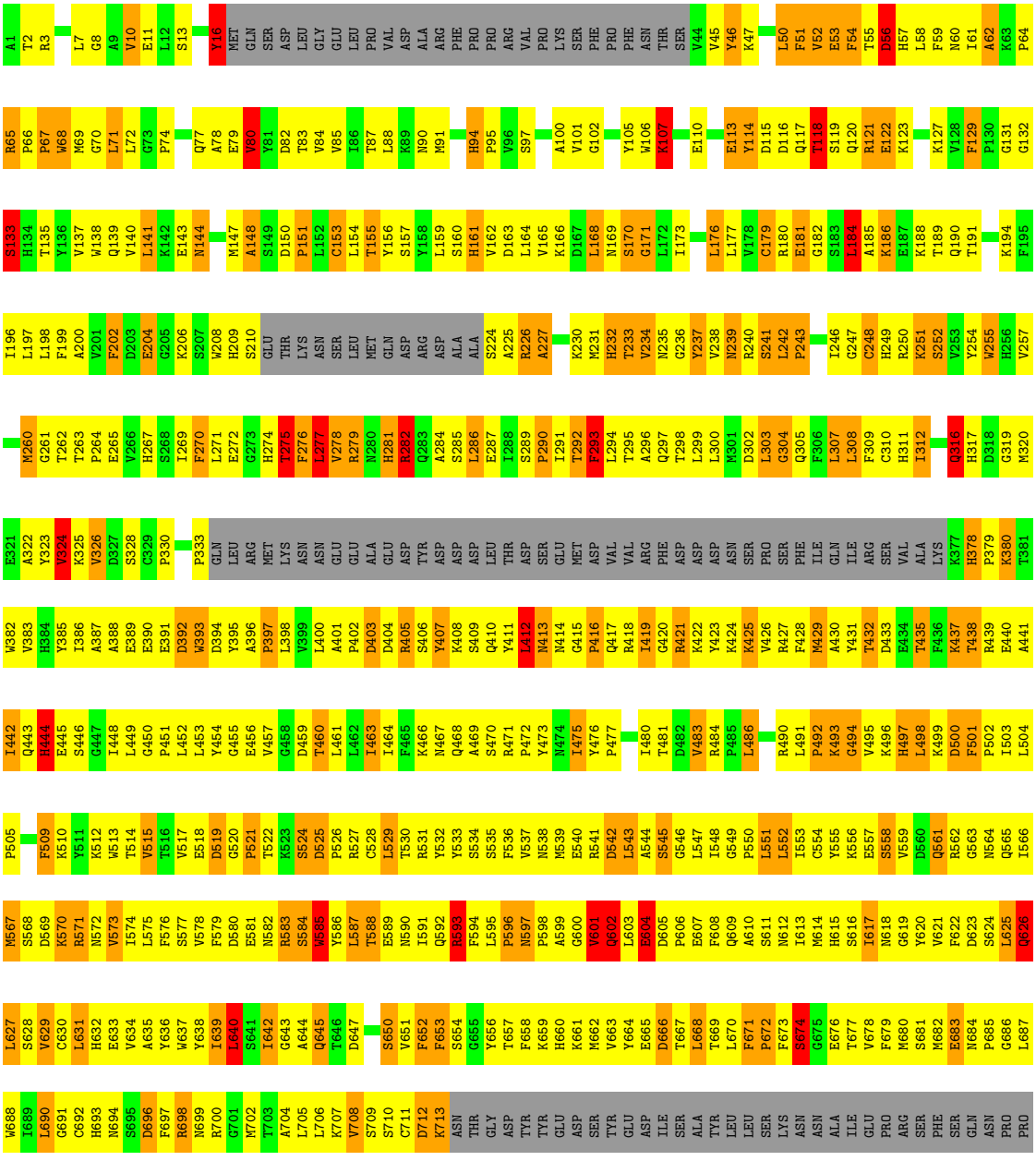
Mol	Chain	Residues	Atoms	AltConf
7	B	1	Total C N O 14 8 1 5	0
7	B	1	Total C N O 14 8 1 5	0

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.

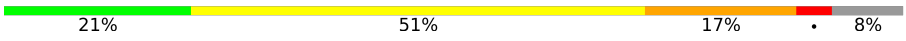
- Molecule 1: Coagulation factor VIII heavy chain

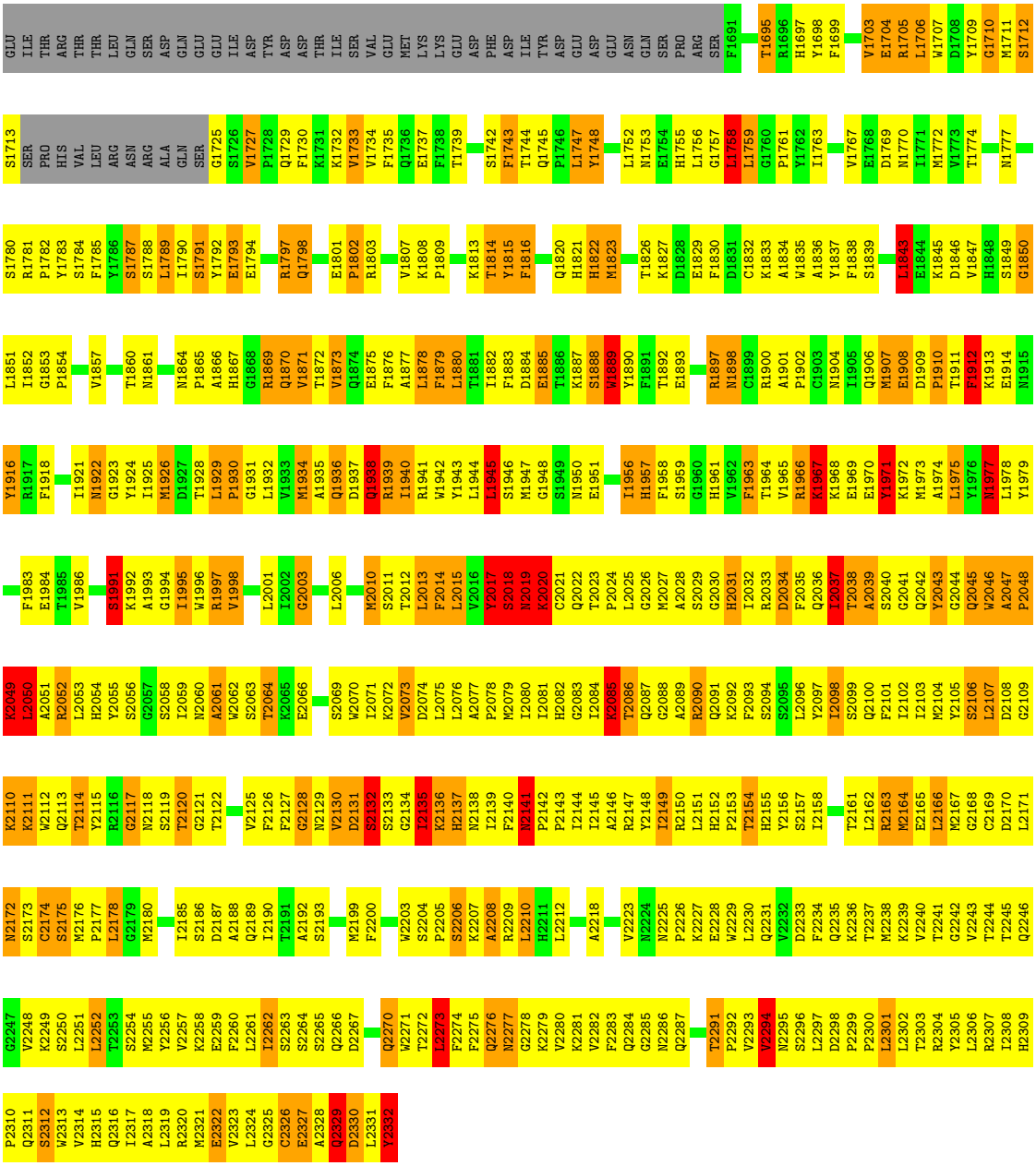
Chain A:



• Molecule 2: Coagulation factor VIII light chain

Chain B:



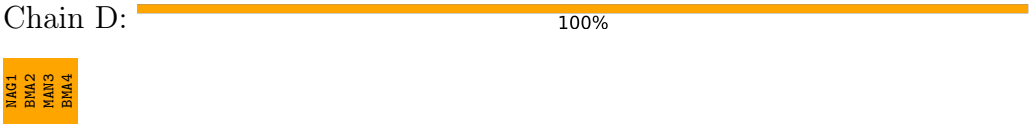


• Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain C:

• Molecule 4: alpha-D-mannopyranose-(1-3)-[beta-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain D:



4 Data and refinement statistics (i)

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	81.00Å 71.00Å 100.00Å 67.00° 60.00° 65.00°	Depositor
Resolution (Å)	(Not available) – 15.00	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-15.00)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$	-	Xtriage
Refinement program	unknown	Depositor
R, R_{free}	(Not available) , (Not available)	Depositor
Wilson B-factor (Å^2)	(Not available)	Xtriage
Anisotropy	(Not available)	Xtriage
L-test for twinning[1]	$< \|L\| >$ = (Not available), $< L^2 >$ = (Not available)	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	10317	wwPDB-VP
Average B, all atoms (Å^2)	2.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *(Not available)*

[1]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: CA, NAG, MAN, BMA, CU

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	0.81	12/5230 (0.2%)	0.95	11/7098 (0.2%)
2	B	0.93	20/5270 (0.4%)	1.31	39/7136 (0.5%)
All	All	0.87	32/10500 (0.3%)	1.15	50/14234 (0.4%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	7
2	B	2	7
All	All	2	14

The worst 5 of 32 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	2132	SER	CA-CB	20.34	1.83	1.52
2	B	2020	LYS	CA-C	14.78	1.91	1.52
2	B	2018	SER	CA-C	14.67	1.91	1.52
2	B	2018	SER	N-CA	13.70	1.73	1.46
1	A	121	ARG	CZ-NH1	13.13	1.50	1.33

The worst 5 of 50 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	2020	LYS	O-C-N	35.80	179.97	122.70
2	B	2020	LYS	CA-C-O	-33.53	49.68	120.10
2	B	2017	TYR	O-C-N	-30.67	73.63	122.70
2	B	2018	SER	N-CA-CB	-26.55	70.67	110.50

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	121	ARG	NE-CZ-NH2	-24.06	108.27	120.30

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	B	2018	SER	CA
2	B	2135	ILE	CB

5 of 14 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	237	TYR	Peptide
1	A	278	VAL	Peptide
1	A	281	HIS	Peptide
1	A	282	ARG	Peptide
1	A	56	ASP	Peptide

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	5086	0	4971	442	3289
2	B	5125	0	4992	575	3230
3	C	28	0	25	1	0
4	D	47	0	40	2	54
5	A	1	0	0	1	0
5	B	1	0	0	0	0
6	A	1	0	0	0	0
7	B	28	0	26	1	0
All	All	10317	0	10054	948	3311

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 47.

The worst 5 of 948 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:107:LYS:CG	1:A:107:LYS:CD	1.74	1.61
2:B:1710:GLY:CA	2:B:2090:ARG:HH11	1.19	1.56
1:A:119:SER:CB	2:B:2139:ILE:HG12	1.32	1.53
2:B:2132:SER:CB	2:B:2132:SER:CA	1.83	1.52
2:B:2019:ASN:CB	2:B:2019:ASN:CG	1.78	1.50

The worst 5 of 3311 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:437:LYS:O	4:D:4:BMA:C5[1_565]	0.18	2.02
1:A:579:PHE:N	2:B:2146:ALA:C[1_565]	0.19	2.01
1:A:572:ASN:ND2	2:B:2243:VAL:O[1_565]	0.22	1.98
1:A:604:GLU:CB	2:B:2126:PHE:N[1_565]	0.25	1.95
1:A:630:CYS:C	2:B:2257:VAL:CG1[1_565]	0.28	1.92

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 PDB entries followed by that with respect to all EM entries.

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	622/754 (82%)	430 (69%)	131 (21%)	61 (10%)	0 10
2	B	627/684 (92%)	434 (69%)	125 (20%)	68 (11%)	0 8
All	All	1249/1438 (87%)	864 (69%)	256 (20%)	129 (10%)	1 8

5 of 129 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	118	THR
1	A	133	SER
1	A	181	GLU
1	A	227	ALA
1	A	232	HIS

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 PDB entries followed by that with respect to all EM entries.

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	559/674 (83%)	414 (74%)	145 (26%)	0 3
2	B	560/612 (92%)	446 (80%)	114 (20%)	1 7
All	All	1119/1286 (87%)	860 (77%)	259 (23%)	3 4

5 of 259 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	2122	THR
2	B	2156	TYR
1	A	497	HIS
1	A	484	ARG
2	B	2178	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 32 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	2129	ASN
2	B	2266	GLN
2	B	1736	GLN
1	A	694	ASN
2	B	2311	GLN

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)

6 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	$\#\|Z\| > 2$
3	NAG	C	1	3,1	14,14,15	0.87	0	17,19,21	1.08	1 (5%)
3	NAG	C	2	3	14,14,15	2.88	3 (21%)	17,19,21	1.74	2 (11%)
4	NAG	D	1	4	14,14,15	0.95	0	17,19,21	1.63	6 (35%)
4	BMA	D	2	4	11,11,12	0.65	0	15,15,17	1.64	2 (13%)
4	MAN	D	3	4	11,11,12	0.70	0	15,15,17	1.99	2 (13%)
4	BMA	D	4	4	11,11,12	0.80	0	15,15,17	1.70	3 (20%)

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
3	NAG	C	1	3,1	-	3/6/23/26	0/1/1/1
3	NAG	C	2	3	-	4/6/23/26	0/1/1/1
4	NAG	D	1	4	-	5/6/23/26	0/1/1/1
4	BMA	D	2	4	-	2/2/19/22	0/1/1/1
4	MAN	D	3	4	-	2/2/19/22	0/1/1/1
4	BMA	D	4	4	-	2/2/19/22	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	2	NAG	O7-C7	8.01	1.41	1.23
3	C	2	NAG	C8-C7	-6.02	1.38	1.50
3	C	2	NAG	C1-C2	2.72	1.56	1.52

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	D	3	MAN	C1-O5-C5	6.48	120.98	112.19
3	C	2	NAG	C4-C3-C2	5.16	118.58	111.02
4	D	2	BMA	C1-O5-C5	4.50	118.29	112.19
4	D	4	BMA	C1-C2-C3	3.87	114.42	109.67
4	D	4	BMA	C2-C3-C4	3.78	117.43	110.89

There are no chirality outliers.

5 of 18 torsion outliers are listed below:

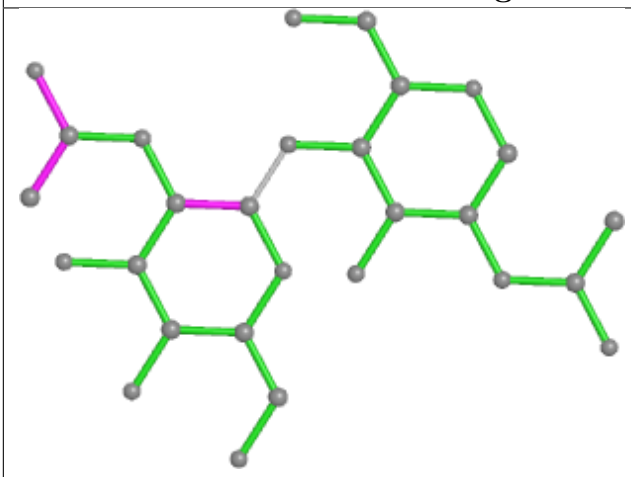
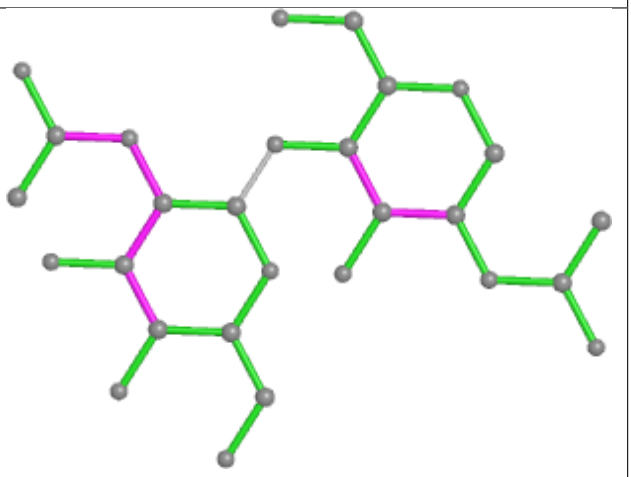
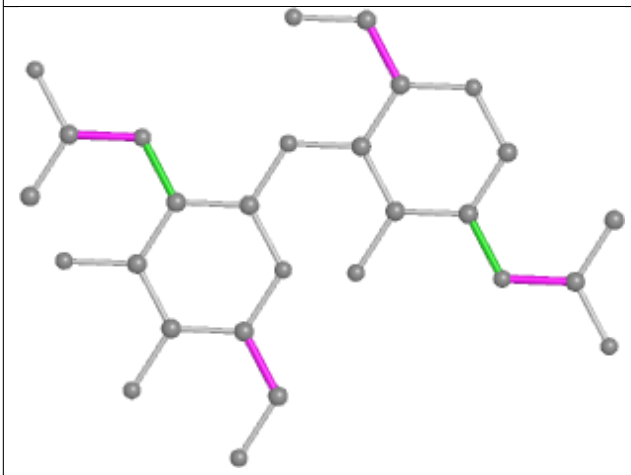
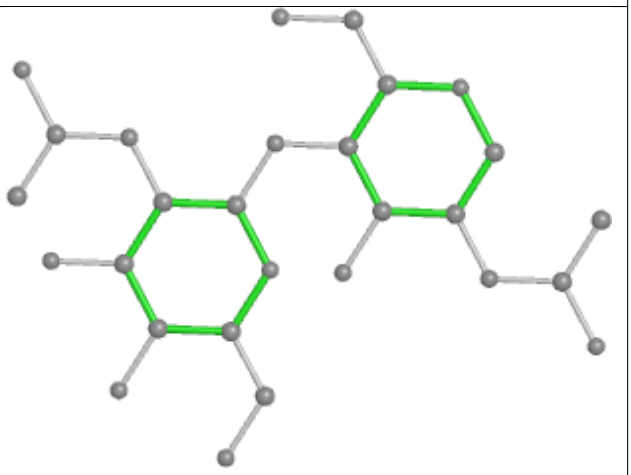
Mol	Chain	Res	Type	Atoms
3	C	1	NAG	C8-C7-N2-C2
3	C	1	NAG	O7-C7-N2-C2
4	D	1	NAG	C3-C2-N2-C7
4	D	1	NAG	C8-C7-N2-C2
4	D	1	NAG	O7-C7-N2-C2

There are no ring outliers.

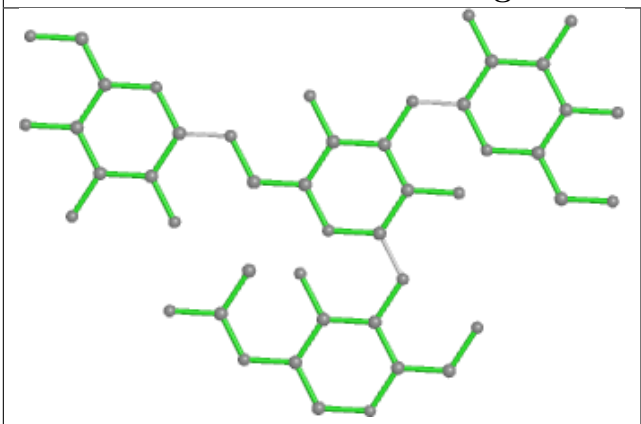
5 monomers are involved in 57 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	D	4	BMA	0	31
4	D	1	NAG	2	3
4	D	3	MAN	0	1
4	D	2	BMA	0	19
3	C	1	NAG	1	0

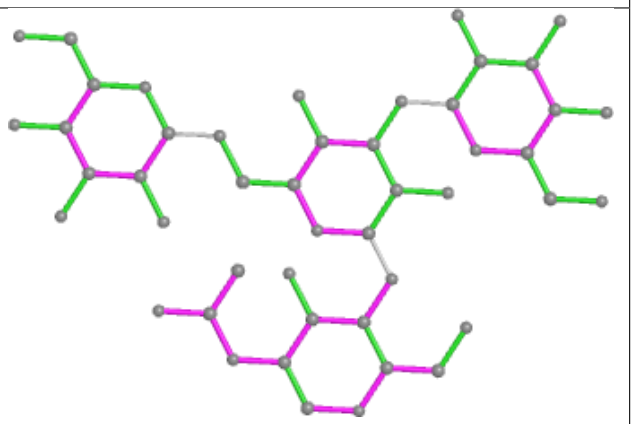
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

Oligosaccharide Chain C	
	
Bond lengths	**Bond angles**
	
Torsions	**Rings**

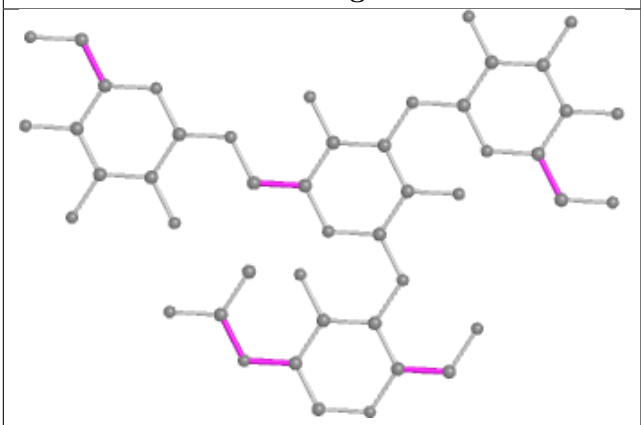
Oligosaccharide Chain D



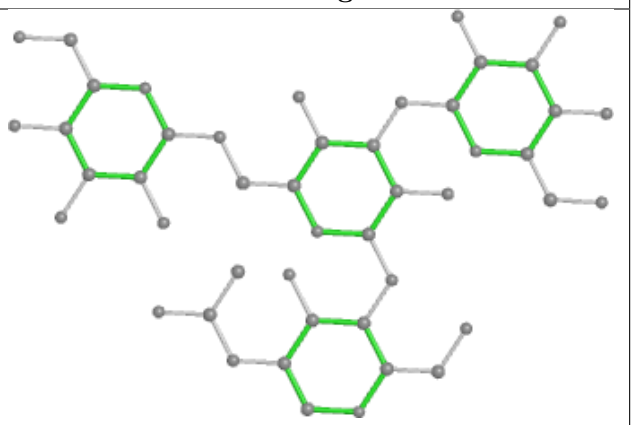
Bond lengths



Bond angles



Torsions



Rings

5.6 Ligand geometry (i)

Of 5 ligands modelled in this entry, 3 are monoatomic - leaving 2 for Mogul analysis.

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	$\#\|Z\| > 2$
7	NAG	B	2402	2	14,14,15	0.85	0	17,19,21	1.25	2 (11%)
7	NAG	B	2401	2	14,14,15	0.76	0	17,19,21	2.03	5 (29%)

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
7	NAG	B	2402	2	1/1/5/7	4/6/23/26	0/1/1/1
7	NAG	B	2401	2	1/1/5/7	3/6/23/26	0/1/1/1

There are no bond length outliers.

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	B	2401	NAG	C1-O5-C5	5.03	119.00	112.19
7	B	2401	NAG	C3-C4-C5	4.41	118.11	110.24
7	B	2402	NAG	C1-O5-C5	2.97	116.22	112.19
7	B	2401	NAG	C2-N2-C7	2.51	126.47	122.90
7	B	2402	NAG	C2-N2-C7	2.49	126.45	122.90

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
7	B	2401	NAG	C1
7	B	2402	NAG	C1

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	B	2401	NAG	C8-C7-N2-C2
7	B	2401	NAG	O7-C7-N2-C2
7	B	2402	NAG	C8-C7-N2-C2
7	B	2402	NAG	O7-C7-N2-C2
7	B	2402	NAG	O5-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	B	2402	NAG	1	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	2017:TYR	C	2018:SER	N	1.62