wwPDB X-ray Structure Validation Summary Report (i)

Sep 23, 2023 – 03:36 PM EDT

PDB ID	:	5KQV
Title	:	Insulin receptor ectodomain construct comprising domains L1,CR,L2, FnIII-1 and alphaCT peptide in complex with bovine insulin and FAB 83-14 (REVISED STRUCTURE)
Authors	:	Lawrence, M.C.; Smith, B.J.; Croll, T.I.
Deposited on	:	2016-07-06
Resolution	:	4.40 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 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:

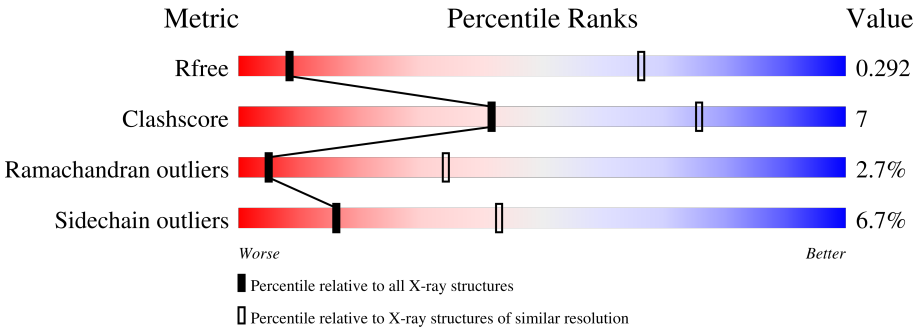
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 4.40 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	1043 (5.00-3.80)
Clashscore	141614	1111 (5.00-3.80)
Ramachandran outliers	138981	1059 (5.00-3.80)
Sidechain outliers	138945	1041 (5.00-3.80)

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	21	48% 43% 10%
1	I	21	48% 33% 19%
2	B	30	43% 13% 43%
2	J	30	27% 23% 7% 43%
3	C	220	75% 18%
3	P	220	75% 21%
4	D	214	86% 12%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	Q	214	
5	E	609	
5	F	609	
6	G	2	
6	H	2	
7	K	3	

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 16735 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 Insulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	21	Total	C	N	O	S	0	0	0
			159	97	25	33	4			
1	I	21	Total	C	N	O	S	0	0	0
			159	97	25	33	4			

- Molecule 2 is a protein called Insulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	17	Total	C	N	O	S	0	0	0
			127	81	21	23	2			
2	J	17	Total	C	N	O	S	0	0	0
			127	81	21	23	2			

- Molecule 3 is a protein called MONOCLONAL ANTIBODY FAB 83-14 - HEAVY CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	211	Total	C	N	O	S	0	0	0
			1589	1014	260	309	6			
3	P	220	Total	C	N	O	S	0	0	0
			1662	1060	273	322	7			

- Molecule 4 is a protein called MONOCLONAL ANTIBODY FAB 83-14 - LIGHT CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	214	Total	C	N	O	S	0	0	0
			1650	1022	277	344	7			
4	Q	214	Total	C	N	O	S	0	0	0
			1650	1022	277	344	7			

- Molecule 5 is a protein called Insulin receptor,Insulin receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	593	Total	C	N	O	S	0	0	0
			4760	3015	824	879	42			
5	F	592	Total	C	N	O	S	0	0	0
			4766	3023	821	880	42			

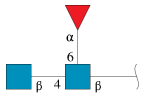
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	144	HIS	TYR	variant	UNP P06213
F	144	HIS	TYR	variant	UNP P06213

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
6	G	2	Total	C	N	O	0	0	0
			24	14	1	9			
6	H	2	Total	C	N	O	0	0	0
			24	14	1	9			

- Molecule 7 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.



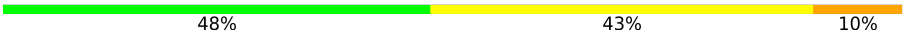
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
7	K	3	Total	C	N	O	0	0	0
			38	22	2	14			

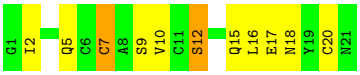
3 Residue-property plots (i)

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: Insulin

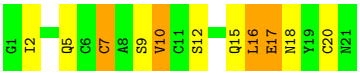
Chain A:





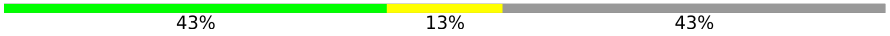
• Molecule 1: Insulin

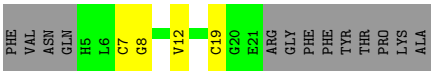
Chain I:



• Molecule 2: Insulin

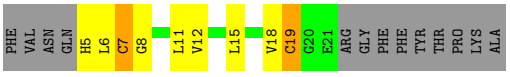
Chain B:





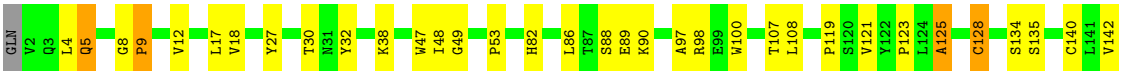
• Molecule 2: Insulin

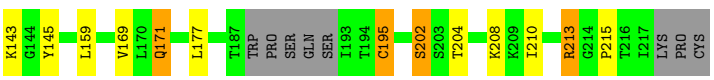
Chain J:



• Molecule 3: MONOCLONAL ANTIBODY FAB 83-14 - HEAVY CHAIN

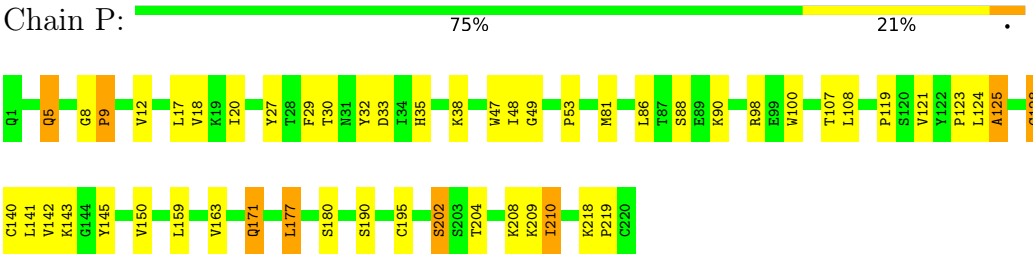
Chain C:





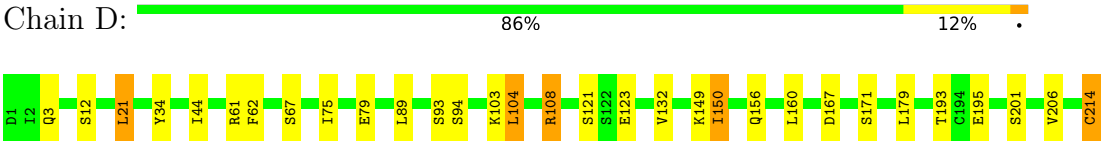
• Molecule 3: MONOCLONAL ANTIBODY FAB 83-14 - HEAVY CHAIN

Chain P:



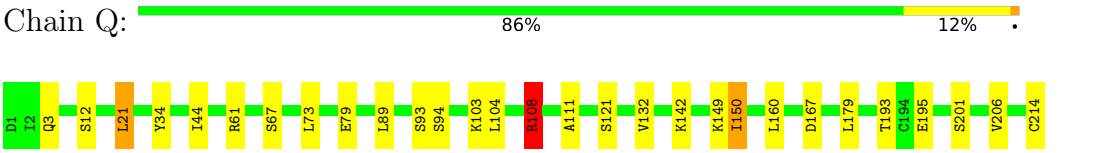
• Molecule 4: MONOCLONAL ANTIBODY FAB 83-14 - LIGHT CHAIN

Chain D:



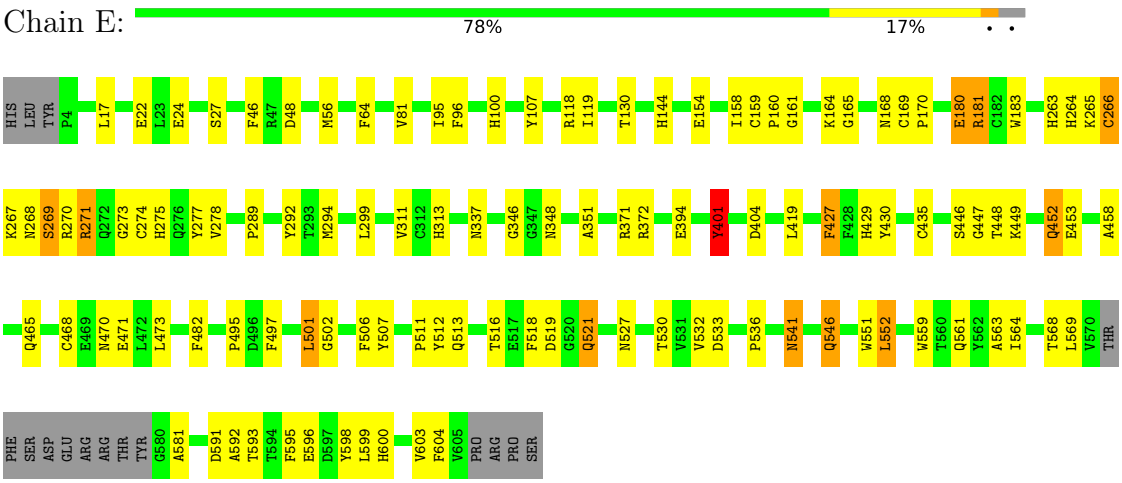
• Molecule 4: MONOCLONAL ANTIBODY FAB 83-14 - LIGHT CHAIN

Chain Q:



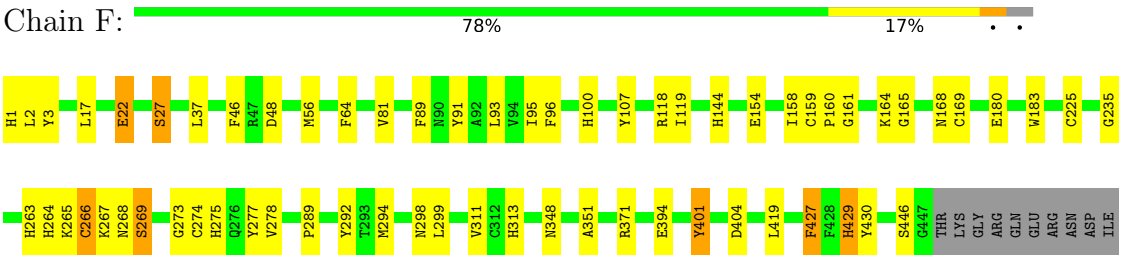
• Molecule 5: Insulin receptor,Insulin receptor

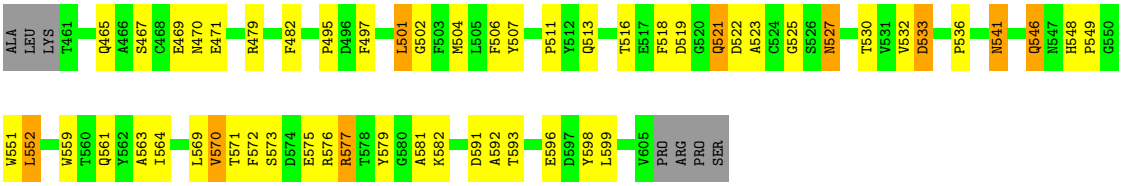
Chain E:



• Molecule 5: Insulin receptor,Insulin receptor

Chain F:





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

Chain G:

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

Chain H:

• Molecule 7: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain K:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	118.15Å 140.10Å 190.02Å 90.00° 95.04° 90.00°	Depositor
Resolution (Å)	42.49 – 4.40 39.95 – 4.40	Depositor EDS
% Data completeness (in resolution range)	88.1 (42.49-4.40) 88.1 (39.95-4.40)	Depositor EDS
R_{merge}	0.21	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.23 (at 4.44Å)	Xtriage
Refinement program	BUSTER 2.10.2	Depositor
R, R_{free}	0.229 , 0.266 0.255 , 0.292	Depositor DCC
R_{free} test set	886 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å^2)	172.2	Xtriage
Anisotropy	0.161	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.23 , 220.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.39$, $< L^2 > = 0.22$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	16735	wwPDB-VP
Average B, all atoms (Å^2)	247.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.37% 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: 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	0.72	0/160	1.03	0/216
1	I	0.83	0/160	1.12	0/216
2	B	0.72	0/129	0.98	0/174
2	J	0.54	0/129	1.00	0/174
3	C	0.54	0/1631	0.93	1/2227 (0.0%)
3	P	0.58	0/1709	0.95	0/2336
4	D	0.43	0/1685	0.73	1/2284 (0.0%)
4	Q	0.44	0/1685	0.73	2/2284 (0.1%)
5	E	0.50	0/4874	0.82	10/6605 (0.2%)
5	F	0.50	0/4884	0.80	5/6623 (0.1%)
All	All	0.51	0/17046	0.83	19/23139 (0.1%)

There are no bond length outliers.

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	E	264	HIS	C-N-CA	7.59	140.68	121.70
5	F	264	HIS	C-N-CA	7.06	139.36	121.70
5	F	264	HIS	CA-C-N	-5.95	104.10	117.20
5	F	502	GLY	N-CA-C	5.92	127.89	113.10
5	E	468	CYS	C-N-CA	5.82	136.26	121.70

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	159	0	145	3	0
1	I	159	0	145	7	0
2	B	127	0	120	0	0
2	J	127	0	120	7	0
3	C	1589	0	1568	25	0
3	P	1662	0	1640	29	0
4	D	1650	0	1572	19	0
4	Q	1650	0	1572	16	0
5	E	4760	0	4626	75	0
5	F	4766	0	4614	85	0
6	G	24	0	22	1	0
6	H	24	0	22	2	0
7	K	38	0	34	1	0
All	All	16735	0	16200	232	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 7.

The worst 5 of 232 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)
5:E:372:ARG:HH21	5:F:533:ASP:CB	1.66	1.08
5:E:372:ARG:NH2	5:F:533:ASP:CB	2.20	1.03
5:E:372:ARG:NH2	5:F:533:ASP:HB3	1.78	0.98
5:E:559:TRP:CE3	5:E:592:ALA:HB2	2.05	0.92
1:I:5:GLN:HE21	1:I:15:GLN:HE21	1.18	0.91

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	19/21 (90%)	13 (68%)	4 (21%)	2 (10%)	0 8
1	I	19/21 (90%)	13 (68%)	4 (21%)	2 (10%)	0 8
2	B	15/30 (50%)	12 (80%)	2 (13%)	1 (7%)	1 18
2	J	15/30 (50%)	12 (80%)	1 (7%)	2 (13%)	0 4
3	C	207/220 (94%)	184 (89%)	15 (7%)	8 (4%)	3 26
3	P	218/220 (99%)	189 (87%)	21 (10%)	8 (4%)	3 28
4	D	212/214 (99%)	204 (96%)	6 (3%)	2 (1%)	17 56
4	Q	212/214 (99%)	204 (96%)	6 (3%)	2 (1%)	17 56
5	E	589/609 (97%)	533 (90%)	42 (7%)	14 (2%)	6 36
5	F	588/609 (97%)	536 (91%)	37 (6%)	15 (3%)	5 35
All	All	2094/2188 (96%)	1900 (91%)	138 (7%)	56 (3%)	5 34

5 of 56 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	18	ASN
3	C	9	PRO
3	C	159	LEU
4	D	94	SER
5	E	267	LYS

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	19/19 (100%)	13 (68%)	6 (32%)	0 2
1	I	19/19 (100%)	13 (68%)	6 (32%)	0 2
2	B	14/25 (56%)	11 (79%)	3 (21%)	1 7
2	J	14/25 (56%)	11 (79%)	3 (21%)	1 7
3	C	178/187 (95%)	164 (92%)	14 (8%)	12 38
3	P	187/187 (100%)	171 (91%)	16 (9%)	10 35

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
4	D	190/190 (100%)	177 (93%)	13 (7%)	16 43
4	Q	190/190 (100%)	178 (94%)	12 (6%)	18 44
5	E	537/553 (97%)	512 (95%)	25 (5%)	26 53
5	F	538/553 (97%)	510 (95%)	28 (5%)	23 50
All	All	1886/1948 (97%)	1760 (93%)	126 (7%)	16 43

5 of 126 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	E	552	LEU
3	P	202	SER
5	F	427	PHE
3	P	195	CYS
4	Q	67	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
5	F	34	GLN
5	F	275	HIS
1	I	5	GLN
5	F	264	HIS
5	F	313	HIS

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)

7 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	$\#\|Z\| > 2$
6	NAG	G	1	5,6	14,14,15	0.40	0	17,19,21	1.16	2 (11%)
6	FUC	G	2	6	10,10,11	0.54	0	14,14,16	0.90	1 (7%)
6	NAG	H	1	5,6	14,14,15	0.41	0	17,19,21	2.06	3 (17%)
6	FUC	H	2	6	10,10,11	0.59	0	14,14,16	1.30	1 (7%)
7	NAG	K	1	5,7	14,14,15	0.43	0	17,19,21	2.02	1 (5%)
7	NAG	K	2	7	14,14,15	0.37	0	17,19,21	1.12	1 (5%)
7	FUC	K	3	7	10,10,11	0.57	0	14,14,16	0.99	1 (7%)

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
6	NAG	G	1	5,6	-	5/6/23/26	0/1/1/1
6	FUC	G	2	6	-	-	0/1/1/1
6	NAG	H	1	5,6	-	2/6/23/26	0/1/1/1
6	FUC	H	2	6	-	-	0/1/1/1
7	NAG	K	1	5,7	-	2/6/23/26	0/1/1/1
7	NAG	K	2	7	-	1/6/23/26	0/1/1/1
7	FUC	K	3	7	-	-	0/1/1/1

There are no bond length outliers.

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	K	1	NAG	O5-C1-C2	-7.79	98.99	111.29
6	H	1	NAG	O5-C1-C2	-6.60	100.87	111.29
6	H	1	NAG	C1-O5-C5	4.58	118.40	112.19
7	K	2	NAG	C1-O5-C5	4.29	118.01	112.19
6	H	2	FUC	C1-C2-C3	3.67	114.17	109.67

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

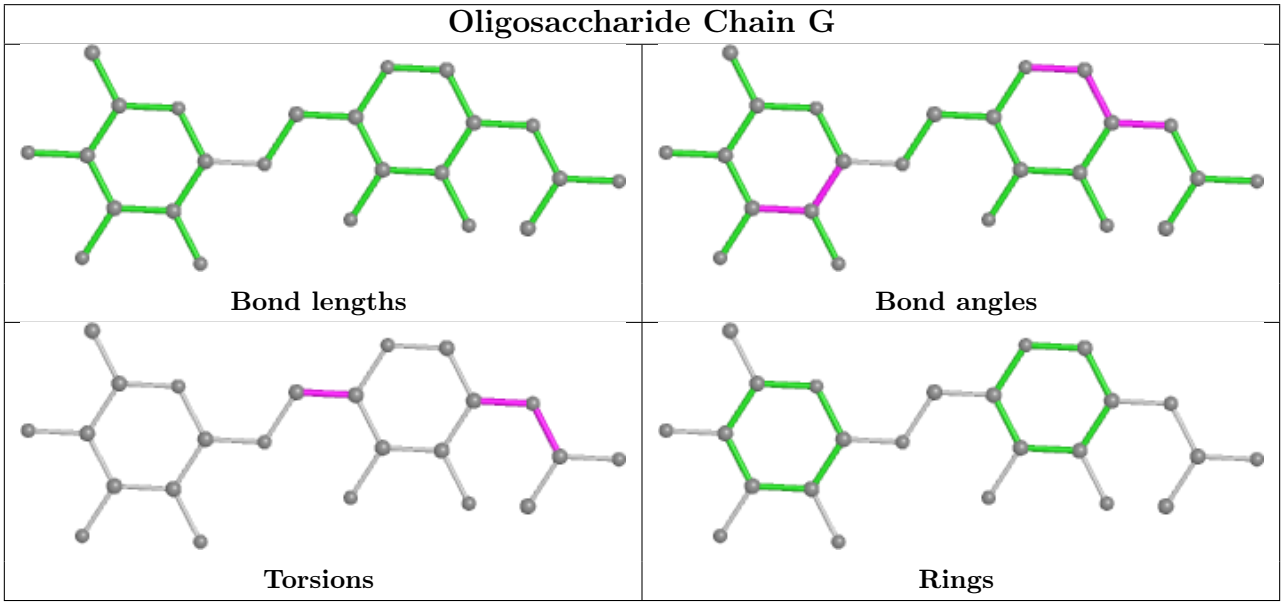
Mol	Chain	Res	Type	Atoms
6	G	1	NAG	C8-C7-N2-C2
6	G	1	NAG	O7-C7-N2-C2
6	H	1	NAG	C8-C7-N2-C2
6	H	1	NAG	O7-C7-N2-C2
7	K	1	NAG	O5-C5-C6-O6

There are no ring outliers.

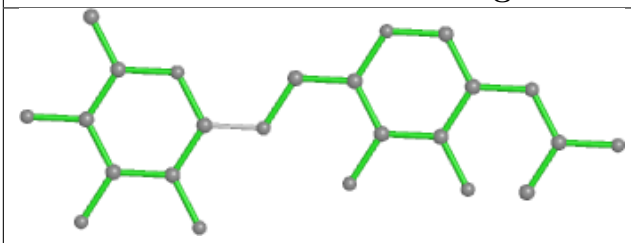
4 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	G	1	NAG	1	0
7	K	2	NAG	1	0
6	H	1	NAG	2	0
7	K	1	NAG	1	0

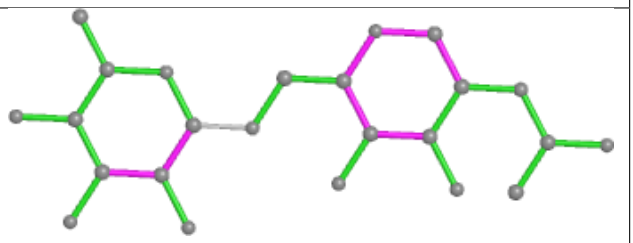
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.



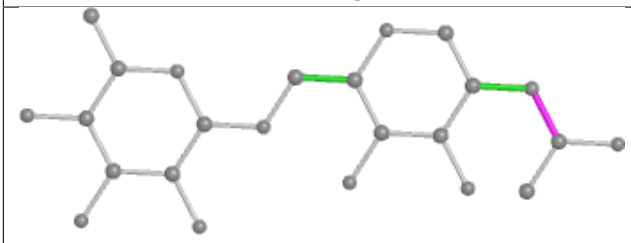
Oligosaccharide Chain H



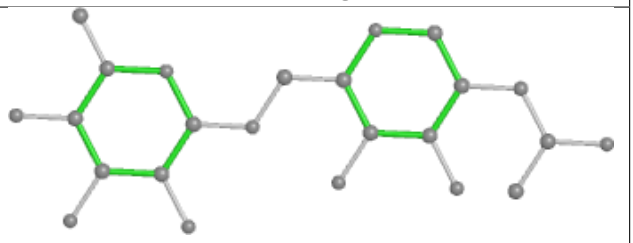
Bond lengths



Bond angles

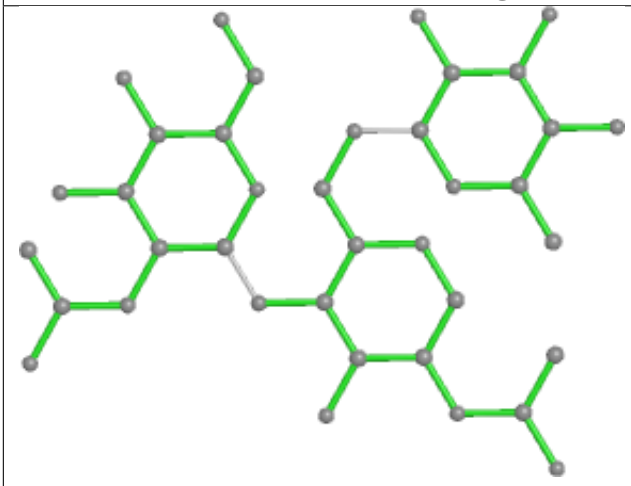


Torsions

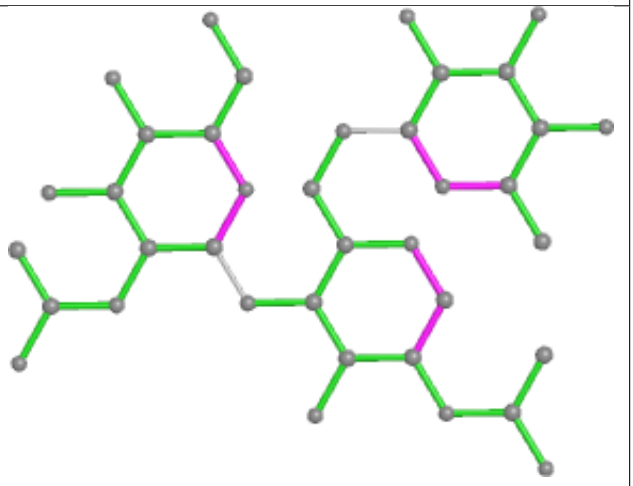


Rings

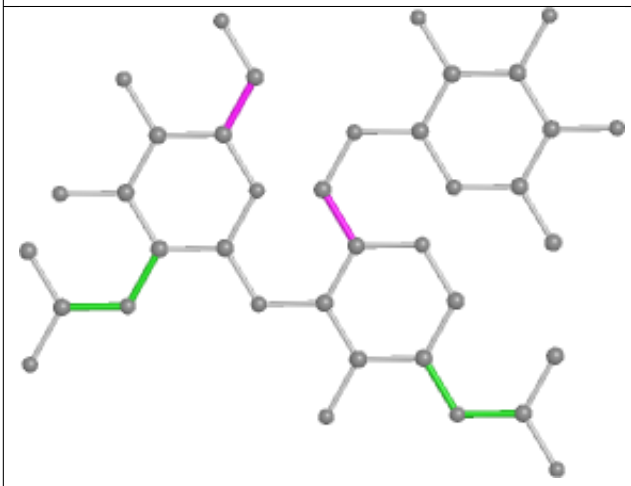
Oligosaccharide Chain K



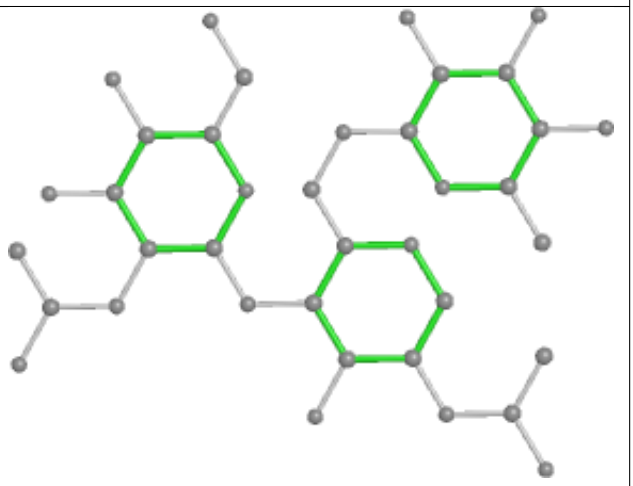
Bond lengths



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.

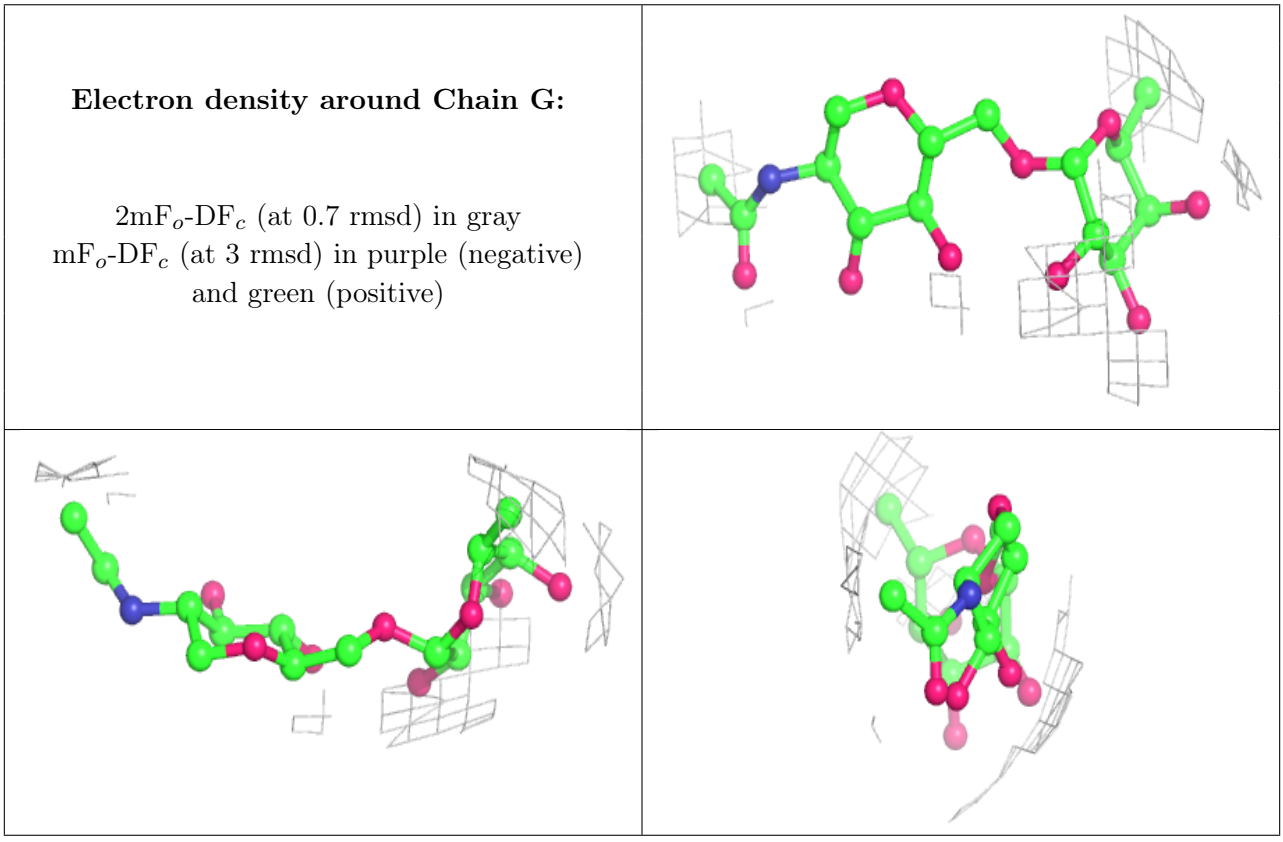
6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates

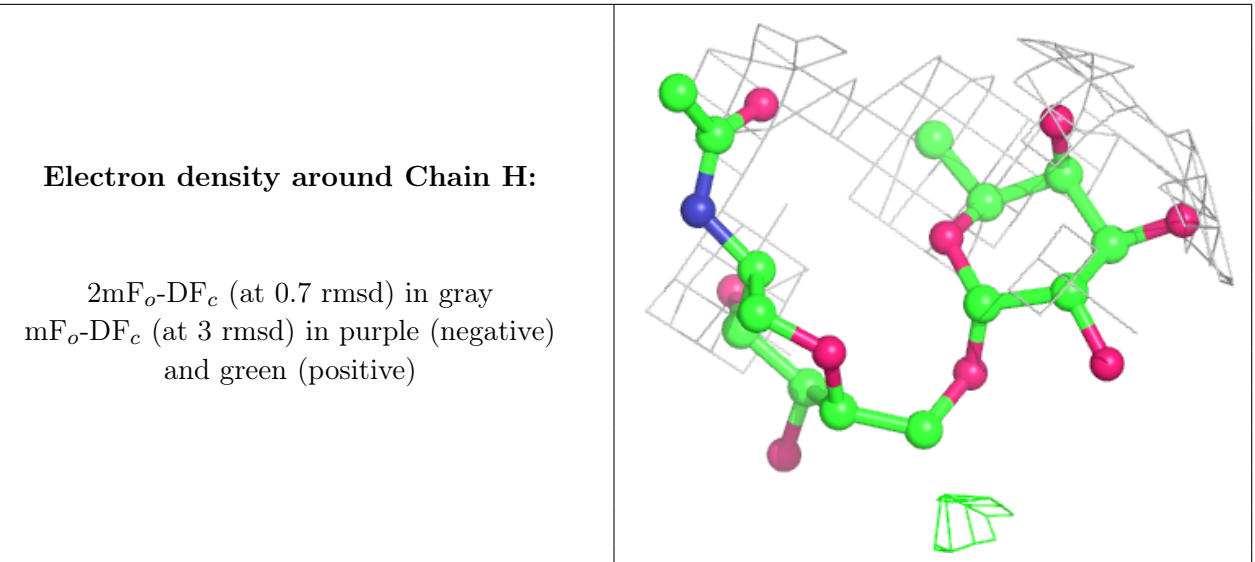
Unable to reproduce the depositors R factor - this section is therefore empty.

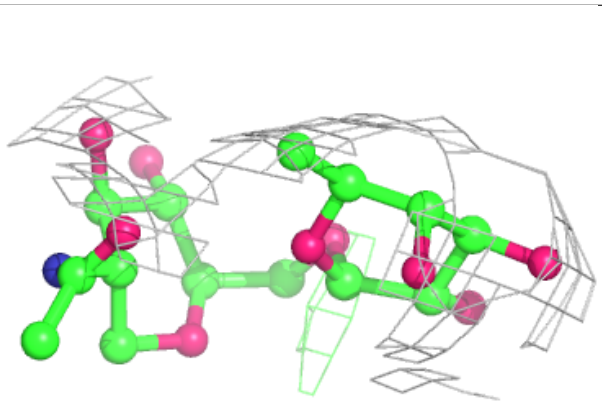
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

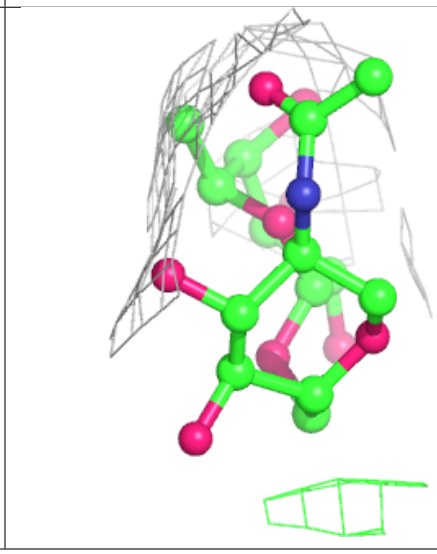


Electron density around Chain H:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

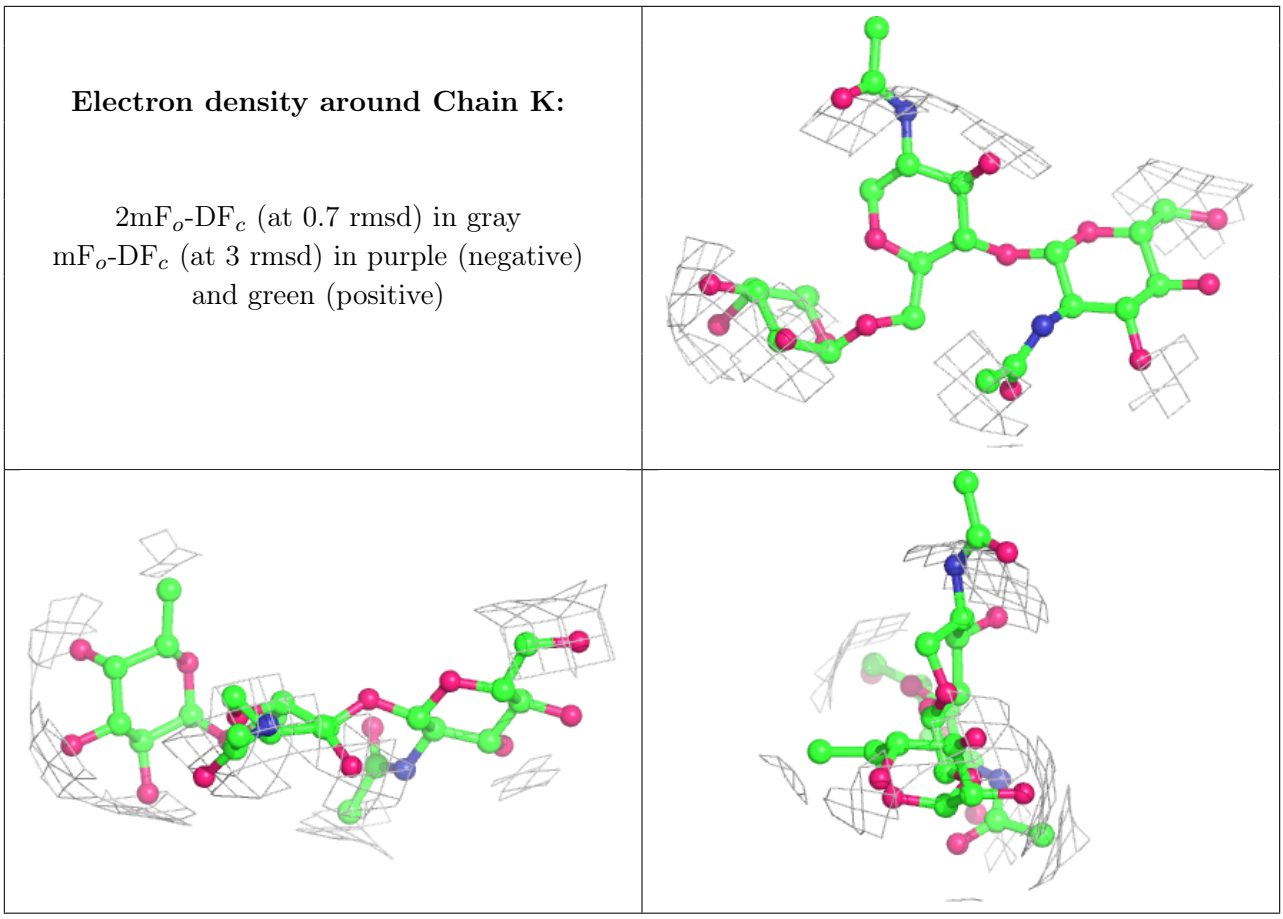






Electron density around Chain K:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



6.4 Ligands

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

6.5 Other polymers

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