Full wwPDB X-ray Structure Validation Report ⓘ

Nov 9, 2024 – 01:16 pm GMT

PDB ID	:	5I8O
Title	:	HMM5 Fab in complex with disaccharide
Authors	:	Andersen, G.R.; Tjerrild, L.M.; Spillner, E.
Deposited on	:	2016-02-19
Resolution	:	1.80 Å(reported)

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

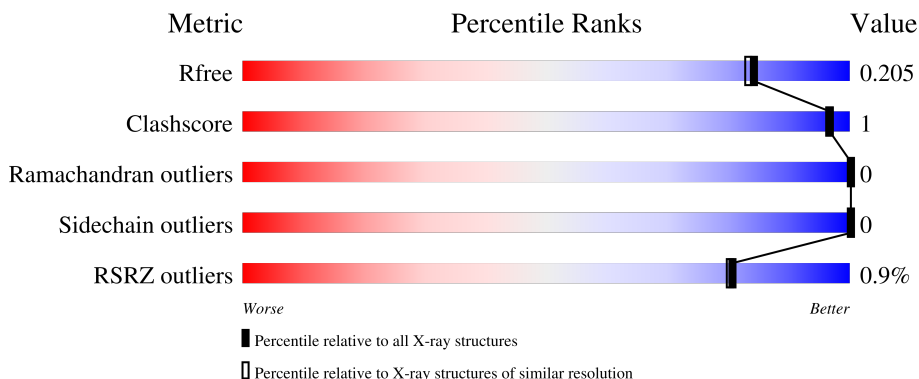|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 1.80 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	Percentile Ranks	Value
Rfree		0.205
Clashscore		1
Ramachandran outliers		0
Sidechain outliers		0
RSRZ outliers		0.9%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	7108 (1.80-1.80)
Clashscore	180529	8162 (1.80-1.80)
Ramachandran outliers	177936	8077 (1.80-1.80)
Sidechain outliers	177891	8076 (1.80-1.80)
RSRZ outliers	164620	7108 (1.80-1.80)

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% 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	H	219	% 98% ·
2	L	217	% 99% ·
3	A	2	100%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 3940 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 HMM5 antibody heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	H	219	Total 1607	C 1016	N 266	O 319	S 6	0	0	0

- Molecule 2 is a protein called HMM5 antibody light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	L	217	Total 1643	C 1025	N 274	O 338	S 6	0	0	0

- Molecule 3 is an oligosaccharide called alpha-L-fucopyranose-(1-3)-methyl 2-acetamido-2-de oxy-alpha-D-glucopyranoside.
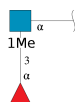
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	A	2	Total 26	C 15	N 1	O 10	0	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	H	330	Total 330	O 330	0	0
4	L	334	Total 334	O 334	0	0

3 Residue-property plots ⓘ

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: HMM5 antibody heavy chain
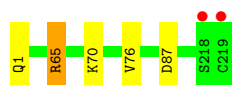
Chain H:

• Molecule 2: HMM5 antibody light chain

Chain L:

• Molecule 3: alpha-L-fucopyranose-(1-3)-methyl 2-acetamido-2-deoxy-alpha-D-glucopyranoside

Chain A:

4 Data and refinement statistics ⓘ

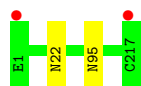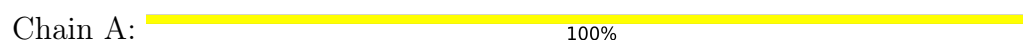| Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	63.45Å 63.59Å 108.77Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.91 – 1.80 44.91 – 1.80	Depositor EDS
% Data completeness (in resolution range)	100.0 (44.91-1.80) 100.0 (44.91-1.80)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.50 (at 1.79Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.164 , 0.205 0.165 , 0.205	Depositor DCC
R_{free} test set	1991 reflections (4.80%)	wwPDB-VP
Wilson B-factor (Å^2)	16.2	Xtriage
Anisotropy	0.171	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 46.3	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.019 for k,h,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	3940	wwPDB-VP
Average B, all atoms (Å^2)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.08% 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: 2F8, 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	H	0.37	0/1647	0.57	1/2251 (0.0%)
2	L	0.37	0/1677	0.53	0/2279
All	All	0.37	0/3324	0.55	1/4530 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	H	65	ARG	NE-CZ-NH2	-5.73	117.43	120.30

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	H	1607	0	1586	3	0
2	L	1643	0	1581	3	0
3	A	26	0	26	0	0
4	H	330	0	0	0	1
4	L	334	0	0	3	2
All	All	3940	0	3193	6	2

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 1.

All (6) 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:H:1:GLN:OE1	1:H:1:GLN:N	2.21	0.70
2:L:95:ASN:HB3	4:L:303:HOH:O	1.90	0.69
2:L:22:ASN:ND2	4:L:302:HOH:O	2.18	0.56
2:L:95:ASN:CB	4:L:303:HOH:O	2.53	0.50
1:H:65:ARG:HH22	1:H:87:ASP:CG	2.20	0.45
1:H:70:LYS:HG2	1:H:76:VAL:HG22	2.00	0.44

All (2) 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)
4:H:652:HOH:O	4:L:560:HOH:O[3_555]	1.88	0.32
4:L:568:HOH:O	4:L:588:HOH:O[4_545]	1.99	0.21

5.3   Torsion angles ⓘ

5.3.1   Protein backbone ⓘ

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	H	217/219 (99%)	214 (99%)	3 (1%)	0	100	100
2	L	215/217 (99%)	208 (97%)	7 (3%)	0	100	100
All	All	432/436 (99%)	422 (98%)	10 (2%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2   Protein sidechains ⓘ

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	H	182/182 (100%)	182 (100%)	0	100	100
2	L	185/185 (100%)	185 (100%)	0	100	100
All	All	367/367 (100%)	367 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 ⓘ

2 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	2F8	A	1	3	16,16,16	0.86	1 (6%)	22,22,22	0.89	1 (4%)
3	FUC	A	2	3	10,10,11	0.83	0	14,14,16	0.88	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
3	2F8	A	1	3	-	0/8/28/28	0/1/1/1
3	FUC	A	2	3	-	-	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1	2F8	O1-C1	2.34	1.44	1.40

All (2) bond angle outliers are listed below:

| Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	2	FUC	C6-C5-C4	-2.50	108.45	113.07
3	A	1	2F8	CM-O1-C1	-2.13	109.99	113.27

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.

Oligosaccharide Chain A

Bond lengths

Bond angles

Torsions

Rings
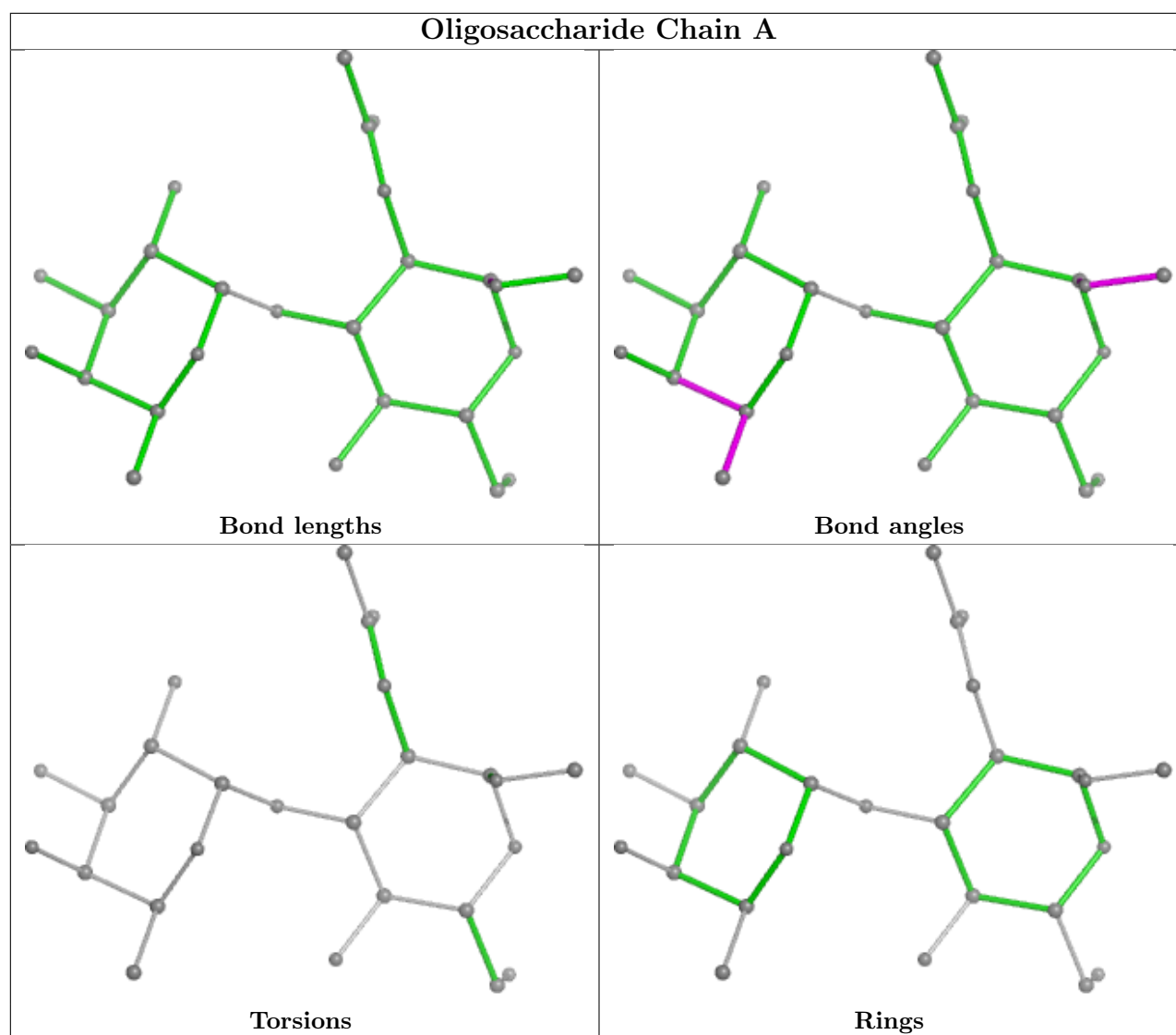
5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

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, 95th 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^2)	Q<0.9
1	H	219/219 (100%)	-0.51	2 (0%)	81	80	11, 16, 27, 86	0
2	L	217/217 (100%)	-0.52	2 (0%)	81	80	10, 16, 28, 97	0
All	All	436/436 (100%)	-0.52	4 (0%)	81	80	10, 16, 28, 97	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	219	CYS	3.9
2	L	217	CYS	3.9
1	H	218	SER	2.7
2	L	1	GLU	2.4

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)

SUGAR-RSR INFOmissingINFO

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.