



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 7, 2020 – 06:38 AM BST

PDB ID : 6YJR
Title : Crystal structure of unliganded MGAT5 (alpha-1,6-mannosylglycoprotein 6-beta-N-acetylglucosaminyltransferase V) luminal domain.
Authors : Wu, L.; Darby, J.F.; Gilio, A.K.; Davies, G.J.
Deposited on : 2020-04-04
Resolution : 2.20 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 ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) 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	:	2.13.1
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)	:	2.13.1

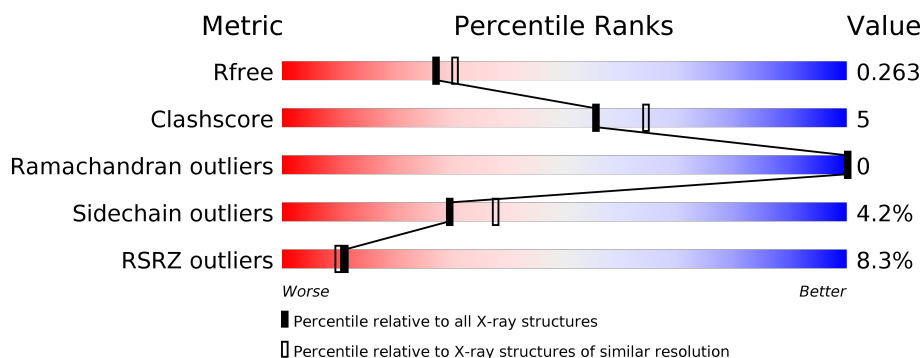
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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 on 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 $\leq 5\%$. 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	AAA	528	<div> <div>5%</div> <div> <div></div> <div>80%</div> <div>13%</div> <div>• 6%</div> </div> </div>
1	BBB	528	<div> <div>11%</div> <div> <div></div> <div>78%</div> <div>14%</div> <div>• 7%</div> </div> </div>
2	A	2	<div> <div></div> <div>100%</div> </div>
2	B	2	<div> <div>50%</div> <div>50%</div> </div>

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 15909 atoms, of which 7857 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 Alpha-1,6-mannosylglycoprotein 6-beta-N-acetylglucosaminyl transferase A.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	AAA	495	Total	C	H	N	O	S	219	0	0
			7881	2541	3916	690	709	25			
1	BBB	493	Total	C	H	N	O	S	223	0	0
			7774	2505	3853	685	706	25			

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



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	2	Total	C	H	N	O	5	0	0
			49	14	25	1	9			
2	B	2	Total	C	H	N	O	5	0	0
			49	14	25	1	9			

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C₈H₁₅NO₆).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	AAA	1	Total	C	H	N	O	3	0
			28	8	14	1	5		

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	AAA	1	Total	C	H	O	1	0
			10	2	6	2		
4	AAA	1	Total	C	H	O	1	0
			10	2	6	2		
4	BBB	1	Total	C	H	O	1	0
			10	2	6	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	BBB	1	Total	C	H	O	1	0
			10	2	6	2		

- Molecule 5 is water.


Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	AAA	55	Total	O	0	0
			55	55		
5	BBB	33	Total	O	0	0
			33	33		

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

Chain A:  100%

MAG1
FUC2

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

Chain B:  50% 50%

MAG1
FUC2

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	59.81Å 100.42Å 207.98Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.81 – 2.20 48.81 – 2.20	Depositor EDS
% Data completeness (in resolution range)	57.1 (48.81-2.20) 57.1 (48.81-2.20)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.07 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.208 , 0.263 0.210 , 0.263	Depositor DCC
R_{free} test set	1866 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å ²)	45.2	Xtriage
Anisotropy	0.949	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.40 , 57.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	15909	wwPDB-VP
Average B, all atoms (Å ²)	76.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.24% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: FUC, EDO, NAG

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	AAA	0.68	0/4067	0.85	2/5500 (0.0%)
1	BBB	0.67	0/4020	0.85	0/5436
All	All	0.67	0/8087	0.85	2/10936 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	AAA	441	ASN	CB-CA-C	-6.31	97.79	110.40
1	AAA	644	GLU	CB-CA-C	5.89	122.18	110.40

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	AAA	3965	3916	3866	37	1
1	BBB	3921	3853	3791	43	1
2	A	24	25	22	0	0
2	B	24	25	22	1	0
3	AAA	14	14	13	0	0
4	AAA	8	12	12	0	0
4	BBB	8	12	12	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	AAA	55	0	0	1	0
5	BBB	33	0	0	2	0
All	All	8052	7857	7738	80	1

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 5.

The worst 5 of 80 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:AAA:433:ASN:O	1:AAA:437:ILE:HG12	1.81	0.79
1:AAA:417:HIS:ND1	1:AAA:418:THR:HG23	1.98	0.78
1:BBB:657:GLN:HG3	1:BBB:657:GLN:O	1.86	0.74
1:BBB:581:LEU:HD23	5:BBB:913:HOH:O	1.93	0.68
1:BBB:295:LEU:HD12	5:BBB:932:HOH:O	1.98	0.64

All (1) 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:AAA:480:TYR:OH	1:BBB:501:ASP:OD2[4_545]	2.00	0.20

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	AAA	487/528 (92%)	464 (95%)	23 (5%)	0	100	100
1	BBB	485/528 (92%)	457 (94%)	28 (6%)	0	100	100
All	All	972/1056 (92%)	921 (95%)	51 (5%)	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	AAA	429/470 (91%)	413 (96%)	16 (4%)	34	43
1	BBB	419/470 (89%)	399 (95%)	20 (5%)	25	32
All	All	848/940 (90%)	812 (96%)	36 (4%)	30	38

5 of 36 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	BBB	231	LYS
1	BBB	295	LEU
1	BBB	629	GLN
1	BBB	272	VAL
1	BBB	318	SER

Some 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 ⓘ

4 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	NAG	A	1	1,2	14,14,15	0.49	0	17,19,21	1.44	2 (11%)
2	FUC	A	2	2	10,10,11	0.89	0	14,14,16	1.95	3 (21%)
2	NAG	B	1	1,2	14,14,15	0.85	0	17,19,21	1.91	5 (29%)
2	FUC	B	2	2	10,10,11	0.81	0	14,14,16	1.56	3 (21%)

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	NAG	A	1	1,2	-	2/6/23/26	0/1/1/1
2	FUC	A	2	2	-	-	0/1/1/1
2	NAG	B	1	1,2	-	2/6/23/26	0/1/1/1
2	FUC	B	2	2	-	-	0/1/1/1

There are no bond length outliers.

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
2	A	2	FUC	C1-C2-C3	-4.51	104.12	109.67
2	A	1	NAG	C1-C2-N2	-4.39	103.00	110.49
2	B	1	NAG	C3-C4-C5	-3.61	103.80	110.24
2	B	1	NAG	O4-C4-C5	3.52	118.05	109.30
2	B	1	NAG	O5-C5-C6	3.48	112.67	107.20

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	1	NAG	O5-C5-C6-O6
2	B	1	NAG	C4-C5-C6-O6
2	A	1	NAG	O5-C5-C6-O6
2	A	1	NAG	C4-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	2	FUC	1	0

5.6 Ligand geometry

5 ligands 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	AAA	803	1	14,14,15	0.78	0	17,19,21	1.71	4 (23%)
4	EDO	AAA	805	-	3,3,3	0.32	0	2,2,2	0.33	0
4	EDO	BBB	803	-	3,3,3	0.17	0	2,2,2	0.15	0
4	EDO	BBB	804	-	3,3,3	0.11	0	2,2,2	0.20	0
4	EDO	AAA	804	-	3,3,3	0.18	0	2,2,2	0.34	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
3	NAG	AAA	803	1	-	2/6/23/26	0/1/1/1
4	EDO	AAA	805	-	-	1/1/1/1	-
4	EDO	BBB	803	-	-	1/1/1/1	-
4	EDO	BBB	804	-	-	1/1/1/1	-
4	EDO	AAA	804	-	-	1/1/1/1	-

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	AAA	803	NAG	O5-C5-C6	4.13	113.68	107.20
3	AAA	803	NAG	O7-C7-C8	-2.47	117.47	122.06
3	AAA	803	NAG	C1-O5-C5	2.18	115.15	112.19
3	AAA	803	NAG	O5-C1-C2	-2.01	108.11	111.29

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	AAA	803	NAG	C4-C5-C6-O6
3	AAA	803	NAG	O5-C5-C6-O6
4	BBB	804	EDO	O1-C1-C2-O2
4	AAA	804	EDO	O1-C1-C2-O2
4	AAA	805	EDO	O1-C1-C2-O2

There are no ring outliers.

No monomer is involved in short contacts.

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 [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	AAA	495/528 (93%)	0.56	25 (5%)	28 26	52, 69, 99, 116	0
1	BBB	493/528 (93%)	0.84	57 (11%)	4 4	57, 79, 109, 130	0
All	All	988/1056 (93%)	0.70	82 (8%)	11 10	52, 74, 106, 130	0

The worst 5 of 82 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	BBB	504	PHE	5.3
1	AAA	321	LEU	5.2
1	BBB	664	PHE	4.6
1	BBB	581	LEU	3.9
1	BBB	343	ASP	3.6

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, 95th 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 ²)	Q<0.9
2	FUC	B	2	10/11	0.78	0.19	30,112,119,122	3
2	NAG	A	1	14/15	0.83	0.12	30,93,98,100	2
2	NAG	B	1	14/15	0.84	0.12	30,101,105,107	2
2	FUC	A	2	10/11	0.90	0.12	30,93,95,97	3

6.4 Ligands [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, 95th 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 ²)	Q<0.9
3	NAG	AAA	803	14/15	0.74	0.14	30,113,123,131	3
4	EDO	AAA	804	4/4	0.90	0.21	30,68,71,72	1
4	EDO	AAA	805	4/4	0.93	0.12	30,64,68,69	1
4	EDO	BBB	804	4/4	0.96	0.21	30,85,87,88	1
4	EDO	BBB	803	4/4	0.96	0.25	30,65,67,67	1

6.5 Other polymers [i](#)

There are no such residues in this entry.