



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 12:25 AM GMT

PDB ID : 2AEC  
Title : Crystal Structure of Human M340H-Beta1,4-Galactosyltransferase-I (M340H-B4GAL-T1) in Complex with GlcNAc-beta1,2-Man-alpha1,6-Man-beta-O  
R  
Authors : Ramasamy, V.; Ramakrishnan, B.; Boeggeman, E.; Ratner, D.M.; Seeberger, P.H.; Qasba, P.K.  
Deposited on : 2005-07-22  
Resolution : 2.00 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

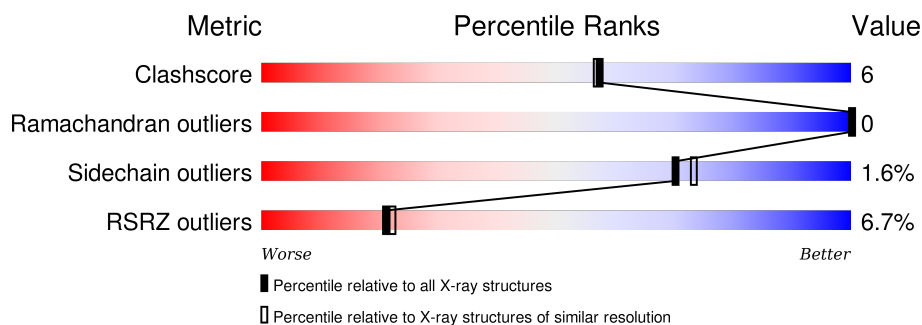
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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)	Similar resolution (#Entries, resolution range(Å))
Clashscore	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (2.00-2.00)
RSRZ outliers	91569	6262 (2.00-2.00)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	A	287	<div> <div>6%</div> <div>82%13%5%</div> </div>
1	B	287	<div> <div>5%</div> <div>84%10%5%</div> </div>
1	C	287	<div> <div>8%</div> <div>79%15%5%</div> </div>

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	SO4	B	428	-	-	-	X
4	SO4	B	433	-	-	-	X
7	MES	C	436	-	-	-	X
8	GOL	A	420	-	-	-	X
8	GOL	B	415	-	-	-	X

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 7464 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 Beta-1,4-galactosyltransferase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	272	Total	C	N	O	S	0	0	0
			2212	1417	385	399	11			
1	B	272	Total	C	N	O	S	0	0	0
			2212	1417	385	399	11			
1	C	272	Total	C	N	O	S	0	0	0
			2212	1417	385	399	11			

There are 51 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	112	ALA	-	SEE REMARK 999	UNP P15291
A	113	SER	-	SEE REMARK 999	UNP P15291
A	114	MET	-	SEE REMARK 999	UNP P15291
A	115	THR	-	SEE REMARK 999	UNP P15291
A	116	GLY	-	SEE REMARK 999	UNP P15291
A	117	GLY	-	SEE REMARK 999	UNP P15291
A	118	GLN	-	SEE REMARK 999	UNP P15291
A	119	GLN	-	SEE REMARK 999	UNP P15291
A	120	MET	-	SEE REMARK 999	UNP P15291
A	121	GLY	-	SEE REMARK 999	UNP P15291
A	122	ARG	-	SEE REMARK 999	UNP P15291
A	123	GLY	-	SEE REMARK 999	UNP P15291
A	124	SER	-	SEE REMARK 999	UNP P15291
A	125	ALA	-	SEE REMARK 999	UNP P15291
A	337	THR	ARG	ENGINEERED	UNP P15291
A	338	THR	CYS	ENGINEERED	UNP P15291
A	340	HIS	MET	ENGINEERED	UNP P15291
B	112	ALA	-	SEE REMARK 999	UNP P15291
B	113	SER	-	SEE REMARK 999	UNP P15291
B	114	MET	-	SEE REMARK 999	UNP P15291
B	115	THR	-	SEE REMARK 999	UNP P15291
B	116	GLY	-	SEE REMARK 999	UNP P15291
B	117	GLY	-	SEE REMARK 999	UNP P15291

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	118	GLN	-	SEE REMARK 999	UNP P15291
B	119	GLN	-	SEE REMARK 999	UNP P15291
B	120	MET	-	SEE REMARK 999	UNP P15291
B	121	GLY	-	SEE REMARK 999	UNP P15291
B	122	ARG	-	SEE REMARK 999	UNP P15291
B	123	GLY	-	SEE REMARK 999	UNP P15291
B	124	SER	-	SEE REMARK 999	UNP P15291
B	125	ALA	-	SEE REMARK 999	UNP P15291
B	337	THR	ARG	ENGINEERED	UNP P15291
B	338	THR	CYS	ENGINEERED	UNP P15291
B	340	HIS	MET	ENGINEERED	UNP P15291
C	112	ALA	-	SEE REMARK 999	UNP P15291
C	113	SER	-	SEE REMARK 999	UNP P15291
C	114	MET	-	SEE REMARK 999	UNP P15291
C	115	THR	-	SEE REMARK 999	UNP P15291
C	116	GLY	-	SEE REMARK 999	UNP P15291
C	117	GLY	-	SEE REMARK 999	UNP P15291
C	118	GLN	-	SEE REMARK 999	UNP P15291
C	119	GLN	-	SEE REMARK 999	UNP P15291
C	120	MET	-	SEE REMARK 999	UNP P15291
C	121	GLY	-	SEE REMARK 999	UNP P15291
C	122	ARG	-	SEE REMARK 999	UNP P15291
C	123	GLY	-	SEE REMARK 999	UNP P15291
C	124	SER	-	SEE REMARK 999	UNP P15291
C	125	ALA	-	SEE REMARK 999	UNP P15291
C	337	THR	ARG	ENGINEERED	UNP P15291
C	338	THR	CYS	ENGINEERED	UNP P15291
C	340	HIS	MET	ENGINEERED	UNP P15291

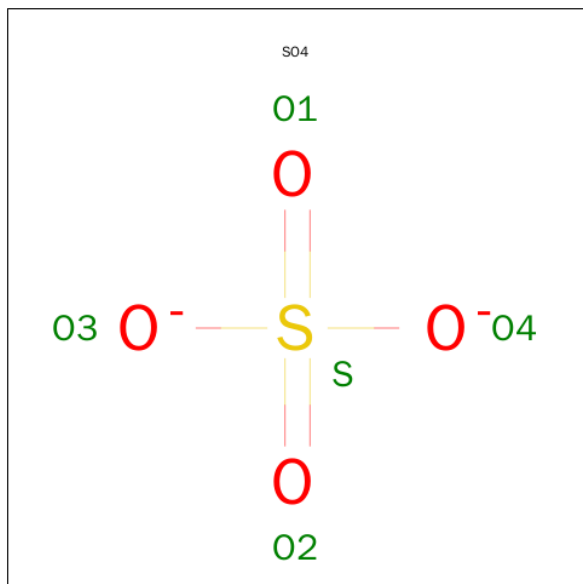
- Molecule 2 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	3	Total	C	N	O	0	0
			37	20	1	16		
2	B	3	Total	C	N	O	0	0
			37	20	1	16		
2	C	3	Total	C	N	O	0	0
			37	20	1	16		

- Molecule 3 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Mn	0	0
			1	1		
3	A	1	Total	Mn	0	0
			1	1		
3	C	1	Total	Mn	0	0
			1	1		

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



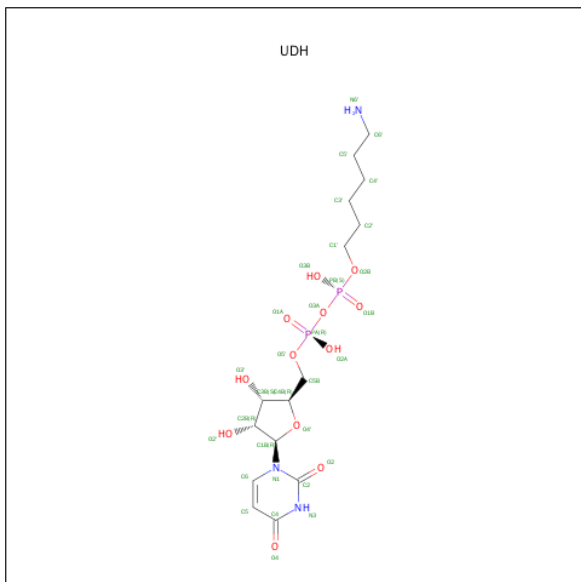
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	O	S	0	0
			5	4	1		
4	A	1	Total	O	S	0	0
			5	4	1		
4	A	1	Total	O	S	0	0
			5	4	1		
4	B	1	Total	O	S	0	0
			5	4	1		
4	B	1	Total	O	S	0	0
			5	4	1		
4	C	1	Total	O	S	0	0
			5	4	1		
4	B	1	Total	O	S	0	0
			5	4	1		
4	B	1	Total	O	S	0	0
			5	4	1		
4	C	1	Total	O	S	0	0
			5	4	1		

*Continued on next page...*

Continued from previous page...

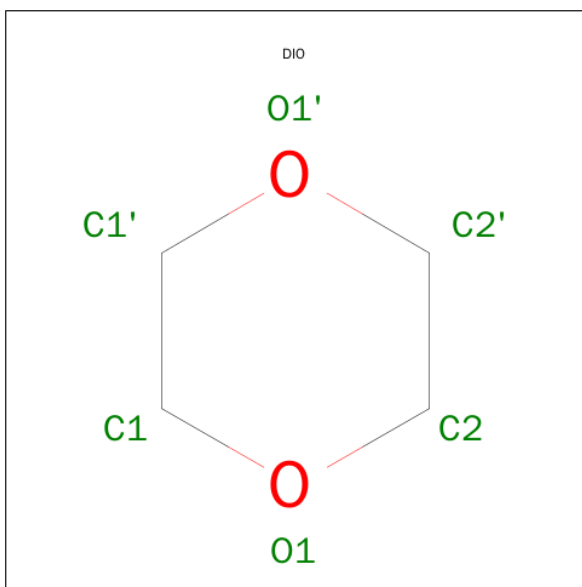
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	C	1	Total	O	S	0	0
			5	4	1		
4	B	1	Total	O	S	0	0
			5	4	1		
4	A	1	Total	O	S	0	0
			5	4	1		
4	B	1	Total	O	S	0	0
			5	4	1		
4	C	1	Total	O	S	0	0
			5	4	1		

- Molecule 5 is 6-AMINOHEXYL-URIDINE-C1,5'-DIPHOSPHATE (three-letter code: UDH) (formula:  $C_{15}H_{27}N_3O_{12}P_2$ ).



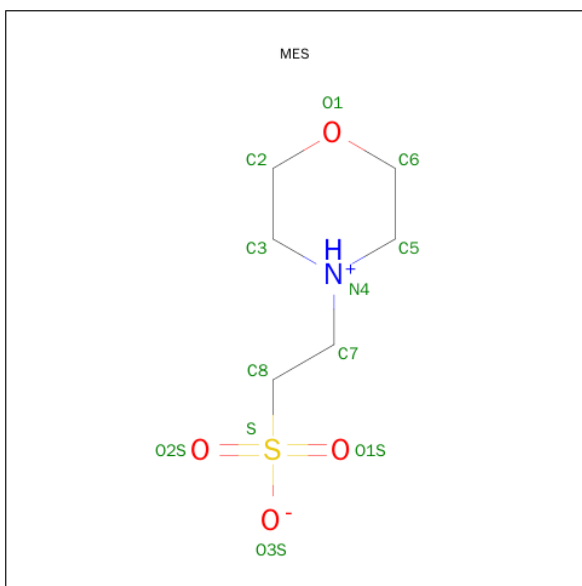
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	N	O	P	0	0
			32	15	3	12	2		
5	B	1	Total	C	N	O	P	0	0
			32	15	3	12	2		
5	C	1	Total	C	N	O	P	0	0
			32	15	3	12	2		

- Molecule 6 is 1,4-DIETHYLENE DIOXIDE (three-letter code: DIO) (formula:  $C_4H_8O_2$ ).



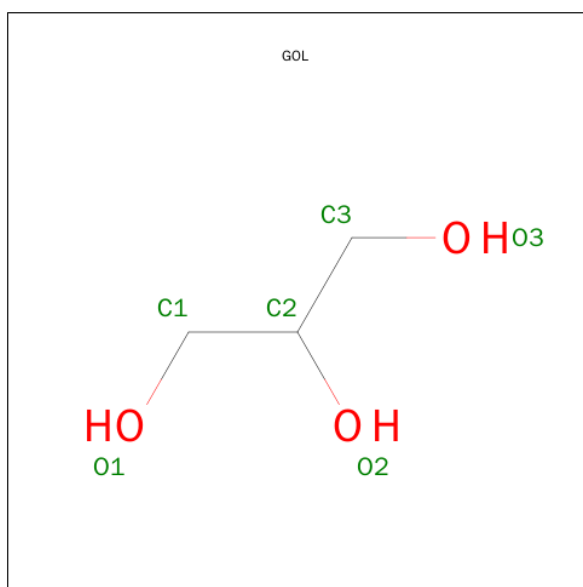
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	C	1	Total	C	O	0	0
			6	4	2		

- Molecule 7 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula:  $C_6H_{13}NO_4S$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
7	C	1	Total	C	N	O	S	0	0
			12	6	1	4	1		

- Molecule 8 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	B	1	Total	C	O	0	0
			6	3	3		
8	B	1	Total	C	O	0	0
			6	3	3		
8	B	1	Total	C	O	0	0
			6	3	3		
8	A	1	Total	C	O	0	0
			6	3	3		
8	C	1	Total	C	O	0	0
			6	3	3		
8	B	1	Total	C	O	0	0
			6	3	3		
8	A	1	Total	C	O	0	0
			6	3	3		

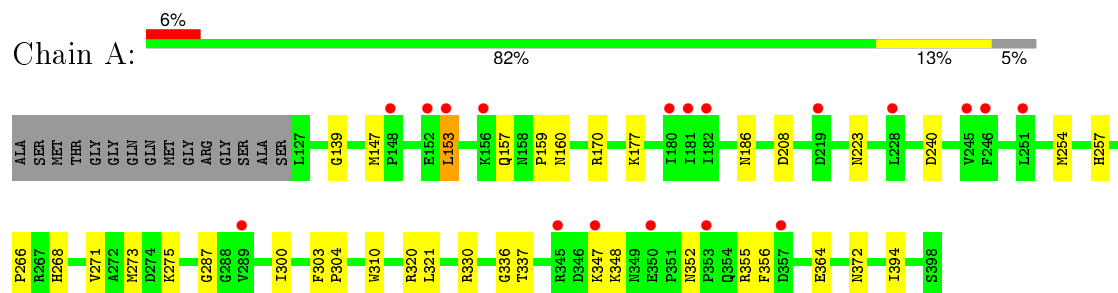
- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	171	Total	O	0	0
			171	171		
9	B	181	Total	O	0	0
			181	181		
9	C	136	Total	O	0	0
			136	136		

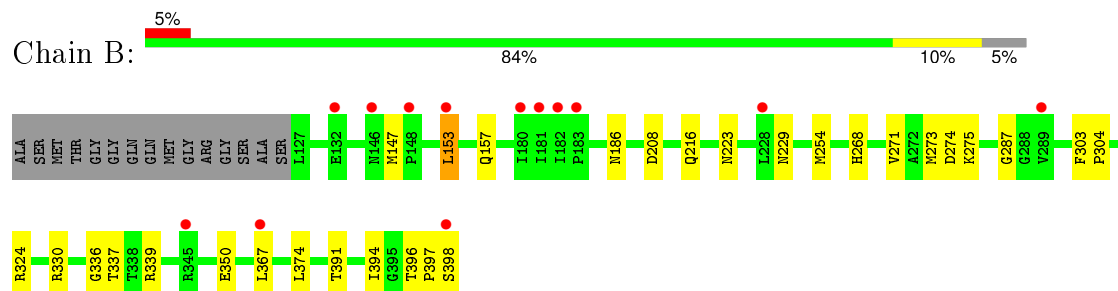
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

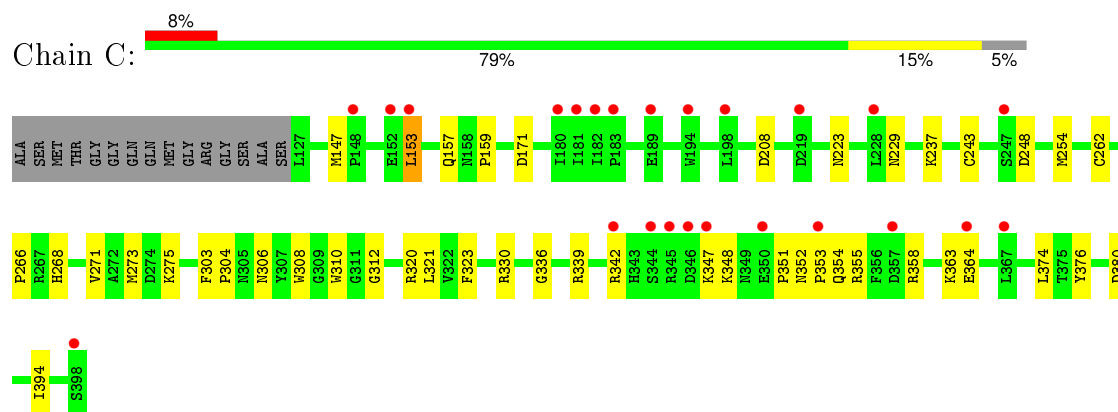
#### • Molecule 1: Beta-1,4-galactosyltransferase 1



#### • Molecule 1: Beta-1,4-galactosyltransferase 1



#### • Molecule 1: Beta-1,4-galactosyltransferase 1



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	107.23Å 194.99Å 143.74Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 2.00 29.85 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.2 (40.00-2.00) 99.4 (29.85-2.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	5.05 (at 2.00Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.208 , 0.239 0.204 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	DCC
Wilson B-factor (Å <sup>2</sup> )	25.3	Xtriage
Anisotropy	0.775	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 51.2	EDS
Estimated twinning fraction	0.000 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.013 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 100651 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	7464	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.07% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: GOL, DIO, NAG, MN, BMA, SO4, MES, UDH, MAN

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	0.45	0/2274	0.69	1/3089 (0.0%)
1	B	0.49	0/2274	0.69	1/3089 (0.0%)
1	C	0.41	0/2274	0.66	2/3089 (0.1%)
All	All	0.45	0/6822	0.68	4/9267 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	B	208	ASP	N-CA-C	-5.29	96.70	111.00
1	C	208	ASP	N-CA-C	-5.22	96.90	111.00
1	C	312	GLY	N-CA-C	5.18	126.05	113.10
1	A	208	ASP	N-CA-C	-5.15	97.10	111.00

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	2212	0	2157	28	0
1	B	2212	0	2157	22	0
1	C	2212	0	2157	33	0
2	A	37	0	33	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	37	0	33	0	0
2	C	37	0	33	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
4	A	20	0	0	0	0
4	B	30	0	0	0	0
4	C	20	0	0	0	0
5	A	32	0	25	0	0
5	B	32	0	25	0	0
5	C	32	0	25	1	0
6	C	6	0	8	2	0
7	C	12	0	13	3	0
8	A	12	0	16	3	0
8	B	24	0	32	3	0
8	C	6	0	8	1	0
9	A	171	0	0	0	0
9	B	181	0	0	1	0
9	C	136	0	0	0	0
All	All	7464	0	6722	82	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 6.

The worst 5 of 82 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:C:268:HIS:HD2	1:C:330:ARG:HE	1.09	0.92
1:A:273:MET:HE1	1:A:275:LYS:HE2	1.52	0.90
1:A:153:LEU:HD22	1:A:157:GLN:HE21	1.36	0.88
1:B:324:ARG:HH11	1:B:367:LEU:HD21	1.41	0.85
1:B:398:SER:H	8:B:419:GOL:H12	1.42	0.84

There are no symmetry-related clashes.

## 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	A	270/287 (94%)	263 (97%)	7 (3%)	0	100	100
1	B	270/287 (94%)	264 (98%)	6 (2%)	0	100	100
1	C	270/287 (94%)	263 (97%)	7 (3%)	0	100	100
All	All	810/861 (94%)	790 (98%)	20 (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	A	243/252 (96%)	239 (98%)	4 (2%)	70	73
1	B	243/252 (96%)	240 (99%)	3 (1%)	78	81
1	C	243/252 (96%)	238 (98%)	5 (2%)	61	63
All	All	729/756 (96%)	717 (98%)	12 (2%)	70	73

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	153	LEU
1	B	350	GLU
1	C	342	ARG
1	B	147	MET
1	C	153	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	377	GLN
1	B	186	ASN
1	C	233	GLN
1	A	268	HIS
1	B	268	HIS

### 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 ⓘ

9 carbohydrates 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	405	2	14,14,15	1.63	4 (28%)	15,19,21	0.95	1 (6%)
2	MAN	A	406	2	11,11,12	2.30	4 (36%)	14,15,17	2.52	4 (28%)
2	BMA	A	407	2	12,12,12	1.81	4 (33%)	17,17,17	0.74	0
2	NAG	B	408	2	14,14,15	1.58	3 (21%)	15,19,21	0.93	1 (6%)
2	MAN	B	409	2	11,11,12	2.29	4 (36%)	14,15,17	2.42	4 (28%)
2	BMA	B	410	2	12,12,12	1.88	4 (33%)	17,17,17	0.75	0
2	NAG	C	411	2	14,14,15	1.93	6 (42%)	15,19,21	1.04	1 (6%)
2	MAN	C	412	2	11,11,12	2.46	4 (36%)	14,15,17	2.55	4 (28%)
2	BMA	C	413	2	12,12,12	1.74	5 (41%)	17,17,17	0.75	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
2	NAG	A	405	2	-	0/6/23/26	0/1/1/1
2	MAN	A	406	2	-	0/2/19/22	0/1/1/1
2	BMA	A	407	2	-	0/2/22/22	0/1/1/1
2	NAG	B	408	2	-	0/6/23/26	0/1/1/1
2	MAN	B	409	2	-	0/2/19/22	0/1/1/1
2	BMA	B	410	2	-	0/2/22/22	0/1/1/1
2	NAG	C	411	2	-	0/6/23/26	0/1/1/1
2	MAN	C	412	2	-	0/2/19/22	0/1/1/1
2	BMA	C	413	2	-	0/2/22/22	0/1/1/1

The worst 5 of 38 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	413	BMA	C4-C3	2.05	1.57	1.52
2	C	411	NAG	C4-C3	2.08	1.57	1.52
2	A	405	NAG	C4-C3	2.13	1.58	1.52
2	B	409	MAN	C4-C5	2.21	1.57	1.53
2	A	407	BMA	O5-C5	2.25	1.50	1.44

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	406	MAN	O2-C2-C3	-4.95	100.17	110.12
2	C	412	MAN	O2-C2-C3	-4.94	100.19	110.12
2	B	409	MAN	O2-C2-C3	-4.38	101.31	110.12
2	B	408	NAG	C4-C3-C2	-2.86	106.78	111.23
2	A	405	NAG	C4-C3-C2	-2.81	106.87	111.23

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	406	MAN	1	0

## 5.6 Ligand geometry

Of 29 ligands modelled in this entry, 3 are monoatomic - leaving 26 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$
5	UDH	A	399	3	25,33,33	1.63	3 (12%)	33,47,47	2.78	9 (27%)
8	GOL	A	417	-	5,5,5	0.16	0	5,5,5	0.64	0
8	GOL	A	420	-	5,5,5	0.31	0	5,5,5	0.59	0
4	SO4	A	421	-	4,4,4	3.16	2 (50%)	6,6,6	1.01	0
4	SO4	A	422	-	4,4,4	3.14	2 (50%)	6,6,6	1.01	0
4	SO4	A	423	-	4,4,4	3.26	2 (50%)	6,6,6	0.95	0
4	SO4	A	432	-	4,4,4	3.28	2 (50%)	6,6,6	0.96	0
5	UDH	B	401	3	25,33,33	1.71	3 (12%)	33,47,47	2.65	9 (27%)
8	GOL	B	414	-	5,5,5	0.43	0	5,5,5	0.73	0
8	GOL	B	415	-	5,5,5	0.27	0	5,5,5	0.69	0
8	GOL	B	416	-	5,5,5	0.52	0	5,5,5	0.50	0
8	GOL	B	419	-	5,5,5	0.36	0	5,5,5	0.55	0
4	SO4	B	424	-	4,4,4	3.21	2 (50%)	6,6,6	1.02	0
4	SO4	B	425	-	4,4,4	3.30	2 (50%)	6,6,6	0.97	0
4	SO4	B	427	-	4,4,4	3.16	2 (50%)	6,6,6	1.01	0
4	SO4	B	428	-	4,4,4	3.17	2 (50%)	6,6,6	0.94	0
4	SO4	B	431	-	4,4,4	3.23	2 (50%)	6,6,6	0.97	0
4	SO4	B	433	-	4,4,4	3.30	2 (50%)	6,6,6	0.96	0
5	UDH	C	403	3	25,33,33	1.72	5 (20%)	33,47,47	2.71	7 (21%)
8	GOL	C	418	-	5,5,5	0.32	0	5,5,5	0.66	0
4	SO4	C	426	-	4,4,4	3.23	2 (50%)	6,6,6	0.92	0
4	SO4	C	429	-	4,4,4	3.21	2 (50%)	6,6,6	0.97	0
4	SO4	C	430	-	4,4,4	3.31	2 (50%)	6,6,6	0.94	0
4	SO4	C	434	-	4,4,4	3.36	2 (50%)	6,6,6	0.94	0
6	DIO	C	435	-	6,6,6	0.92	0	6,6,6	0.70	0
7	MES	C	436	-	11,12,12	1.42	3 (27%)	14,16,16	0.74	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
5	UDH	A	399	3	-	0/21/41/41	0/2/2/2
8	GOL	A	417	-	-	0/4/4/4	0/0/0/0
8	GOL	A	420	-	-	0/4/4/4	0/0/0/0
4	SO4	A	421	-	-	0/0/0/0	0/0/0/0
4	SO4	A	422	-	-	0/0/0/0	0/0/0/0
4	SO4	A	423	-	-	0/0/0/0	0/0/0/0
4	SO4	A	432	-	-	0/0/0/0	0/0/0/0
5	UDH	B	401	3	-	0/21/41/41	0/2/2/2
8	GOL	B	414	-	-	0/4/4/4	0/0/0/0
8	GOL	B	415	-	-	0/4/4/4	0/0/0/0
8	GOL	B	416	-	-	0/4/4/4	0/0/0/0
8	GOL	B	419	-	-	0/4/4/4	0/0/0/0
4	SO4	B	424	-	-	0/0/0/0	0/0/0/0
4	SO4	B	425	-	-	0/0/0/0	0/0/0/0
4	SO4	B	427	-	-	0/0/0/0	0/0/0/0
4	SO4	B	428	-	-	0/0/0/0	0/0/0/0
4	SO4	B	431	-	-	0/0/0/0	0/0/0/0
4	SO4	B	433	-	-	0/0/0/0	0/0/0/0
5	UDH	C	403	3	-	0/21/41/41	0/2/2/2
8	GOL	C	418	-	-	0/4/4/4	0/0/0/0
4	SO4	C	426	-	-	0/0/0/0	0/0/0/0
4	SO4	C	429	-	-	0/0/0/0	0/0/0/0
4	SO4	C	430	-	-	0/0/0/0	0/0/0/0
4	SO4	C	434	-	-	0/0/0/0	0/0/0/0
6	DIO	C	435	-	-	0/0/6/6	0/1/1/1
7	MES	C	436	-	-	0/6/14/14	0/1/1/1

The worst 5 of 42 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	434	SO4	O3-S	-4.59	1.30	1.47
4	B	425	SO4	O3-S	-4.55	1.31	1.47
4	C	430	SO4	O3-S	-4.54	1.31	1.47
4	A	432	SO4	O3-S	-4.53	1.31	1.47
4	B	431	SO4	O3-S	-4.53	1.31	1.47

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	399	UDH	O4'-C4B-C3B	-3.43	98.24	105.15

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	401	UDH	O4'-C4B-C3B	-3.36	98.39	105.15
5	C	403	UDH	C5-C4-N3	-3.16	115.03	123.12
5	A	399	UDH	C5-C4-N3	-3.15	115.05	123.12
5	B	401	UDH	C5-C4-N3	-3.10	115.15	123.12

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

8 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	A	417	GOL	3	0
8	B	414	GOL	1	0
8	B	416	GOL	1	0
8	B	419	GOL	1	0
5	C	403	UDH	1	0
8	C	418	GOL	1	0
6	C	435	DIO	2	0
7	C	436	MES	3	0

## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	272/287 (94%)	0.17	18 (6%) 22 22	18, 27, 44, 55	0
1	B	272/287 (94%)	0.04	13 (4%) 34 36	16, 24, 39, 50	0
1	C	272/287 (94%)	0.42	24 (8%) 12 13	22, 33, 52, 61	0
All	All	816/861 (94%)	0.21	55 (6%) 21 22	16, 28, 47, 61	0

The worst 5 of 55 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	347	LYS	4.4
1	C	398	SER	4.3
1	B	398	SER	4.1
1	C	182	ILE	3.9
1	A	347	LYS	3.4

### 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](#)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	NAG	A	405	14/15	0.94	0.11	0.19	30,34,36,37	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	NAG	C	411	14/15	0.88	0.13	-0.16	48,50,52,53	0
2	NAG	B	408	14/15	0.95	0.09	-0.60	26,29,30,31	0
2	MAN	B	409	11/12	0.93	0.24	-	36,45,47,50	0
2	BMA	A	407	12/12	0.87	0.24	-	41,46,48,49	0
2	MAN	A	406	11/12	0.88	0.22	-	40,47,51,53	0
2	BMA	C	413	12/12	0.80	0.25	-	58,60,61,62	0
2	BMA	B	410	12/12	0.87	0.24	-	40,43,45,47	0
2	MAN	C	412	11/12	0.88	0.24	-	56,60,62,64	0

## 6.4 Ligands ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
7	MES	C	436	12/12	0.43	0.36	16.08	79,82,89,90	0
4	SO4	B	433	5/5	0.84	0.34	6.18	86,87,87,88	0
8	GOL	A	420	6/6	0.85	0.39	5.37	49,55,55,55	0
8	GOL	B	415	6/6	0.84	0.20	4.16	44,48,49,52	0
4	SO4	B	428	5/5	0.89	0.26	3.72	58,62,62,62	0
8	GOL	C	418	6/6	0.93	0.22	1.55	45,45,47,47	0
4	SO4	A	422	5/5	0.91	0.19	1.32	64,66,67,67	0
8	GOL	B	414	6/6	0.94	0.17	0.92	28,34,36,38	0
4	SO4	C	426	5/5	0.94	0.16	0.54	51,54,54,54	0
8	GOL	B	419	6/6	0.76	0.17	0.32	47,49,50,54	0
8	GOL	A	417	6/6	0.94	0.14	0.26	32,36,37,40	0
6	DIO	C	435	6/6	0.67	0.21	0.18	66,66,67,67	0
4	SO4	C	429	5/5	0.94	0.15	0.02	81,81,81,82	0
4	SO4	B	427	5/5	0.93	0.16	-0.10	76,76,77,78	0
5	UDH	B	401	32/32	0.96	0.14	-0.27	17,21,54,59	0
8	GOL	B	416	6/6	0.97	0.10	-0.40	20,24,25,26	0
5	UDH	A	399	32/32	0.95	0.14	-0.45	23,28,51,54	0
4	SO4	A	421	5/5	0.91	0.14	-0.46	74,74,74,74	0
5	UDH	C	403	32/32	0.93	0.16	-0.52	31,36,54,57	0
4	SO4	A	423	5/5	0.84	0.31	-	86,87,87,88	0
4	SO4	B	431	5/5	0.93	0.28	-	87,88,88,88	0
4	SO4	C	430	5/5	0.84	0.29	-	100,100,101,101	0
4	SO4	B	425	5/5	0.86	0.48	-	98,98,98,99	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	SO4	C	434	5/5	0.82	0.48	-	107,107,107,108	0
3	MN	A	400	1/1	0.99	0.07	-	29,29,29,29	0
3	MN	C	404	1/1	0.97	0.05	-	34,34,34,34	0
4	SO4	A	432	5/5	0.85	0.35	-	96,96,96,96	0
3	MN	B	402	1/1	1.00	0.07	-	22,22,22,22	0
4	SO4	B	424	5/5	0.87	0.25	-	75,75,76,76	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.