



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 03:51 PM GMT

PDB ID : 4DO4  
Title : Pharmacological chaperones for human alpha-N-acetylgalactosaminidase  
Authors : Clark, N.E.; Garman, S.C.  
Deposited on : 2012-02-09  
Resolution : 1.40 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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

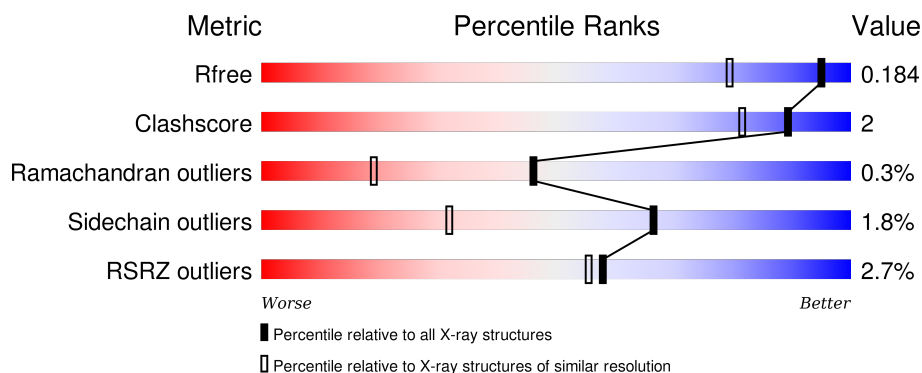
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.40 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))
$R_{free}$	91344	1199 (1.40-1.40)
Clashscore	102246	1295 (1.40-1.40)
Ramachandran outliers	100387	1259 (1.40-1.40)
Sidechain outliers	100360	1258 (1.40-1.40)
RSRZ outliers	91569	1198 (1.40-1.40)

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	400	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 1%, orange 1%, orange 89%, yellow 89%, yellow 95%, green 95%, green 100%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>%</span> <span>89%</span> <span>8%</span> <span>•</span> </div> </div>
1	B	400	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 0%, red 4%, orange 4%, orange 95%, yellow 95%, yellow 99%, green 99%, green 100%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>4%</span> <span>95%</span> <span>5%</span> <span>•</span> </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
7	GOL	A	514	-	-	-	X
7	GOL	A	516	-	-	-	X
7	GOL	A	518	-	-	-	X
7	GOL	B	515	-	-	-	X
7	GOL	B	516	-	-	-	X
7	GOL	B	518	-	-	-	X

## 2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 7713 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 Alpha-N-acetylgalactosaminidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	387	Total	C	N	O	S	0	7	0
			3136	2005	530	573	28			
1	B	400	Total	C	N	O	S	0	5	0
			3223	2059	547	588	29			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	201	GLN	ASN	ENGINEERED MUTATION	UNP P17050
A	412	HIS	-	EXPRESSION TAG	UNP P17050
A	413	HIS	-	EXPRESSION TAG	UNP P17050
A	414	HIS	-	EXPRESSION TAG	UNP P17050
A	415	HIS	-	EXPRESSION TAG	UNP P17050
A	416	HIS	-	EXPRESSION TAG	UNP P17050
A	417	HIS	-	EXPRESSION TAG	UNP P17050
B	201	GLN	ASN	ENGINEERED MUTATION	UNP P17050
B	412	HIS	-	EXPRESSION TAG	UNP P17050
B	413	HIS	-	EXPRESSION TAG	UNP P17050
B	414	HIS	-	EXPRESSION TAG	UNP P17050
B	415	HIS	-	EXPRESSION TAG	UNP P17050
B	416	HIS	-	EXPRESSION TAG	UNP P17050
B	417	HIS	-	EXPRESSION TAG	UNP P17050

- 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
			39	22	2	15		

- Molecule 3 is a polymer of unknown type called SUGAR (5-MER).

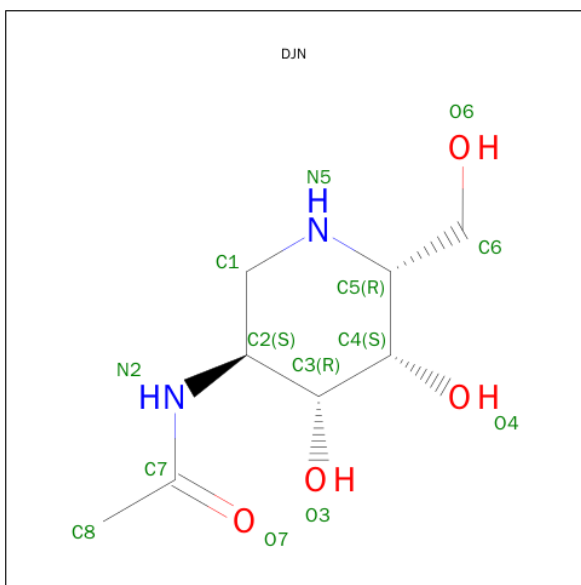
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	5	Total	C	N	O	0	0
			61	34	2	25		
3	B	5	Total	C	N	O	0	0
			61	34	2	25		

- Molecule 4 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



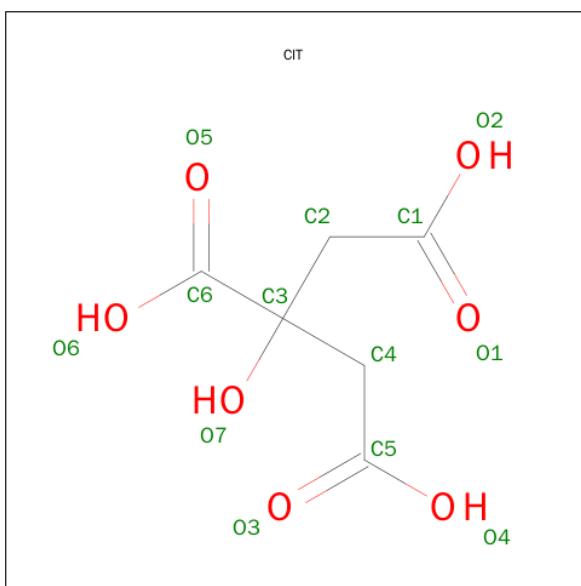
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 5 is N-[(3S,4R,5S,6R)-4,5-DIHYDROXY-6-(HYDROXYMETHYL)PIPERIDIN-3-YL]ACETAMIDE (three-letter code: DJN) (formula:  $C_8H_{16}N_2O_4$ ).



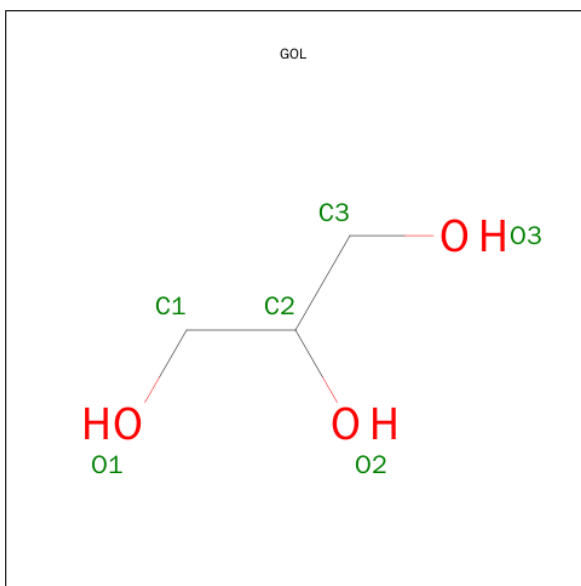
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	0
			14	8	2	4		
5	B	1	Total	C	N	O	0	0
			14	8	2	4		

- Molecule 6 is CITRIC ACID (three-letter code: CIT) (formula:  $C_6H_8O_7$ ).



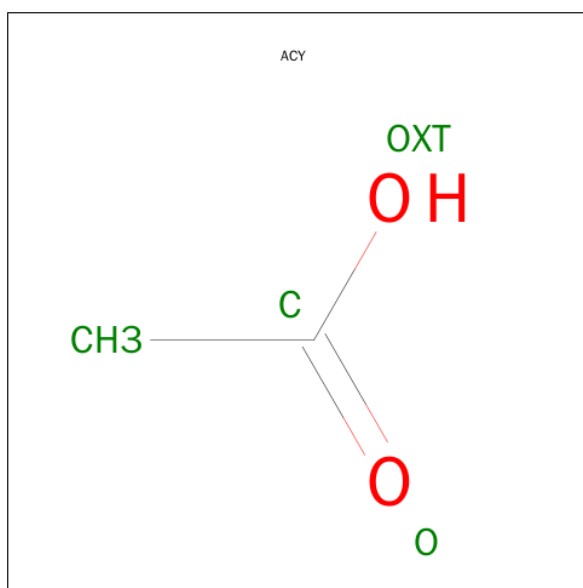
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			13	6	7		
6	B	1	Total	C	O	0	0
			13	6	7		

- Molecule 7 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



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

- Molecule 8 is ACETIC ACID (three-letter code: ACY) (formula:  $C_2H_4O_2$ ).



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

- Molecule 9 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
9	B	3	Total	C	N	O	0	0
			38	22	2	14		

- Molecule 10 is water.

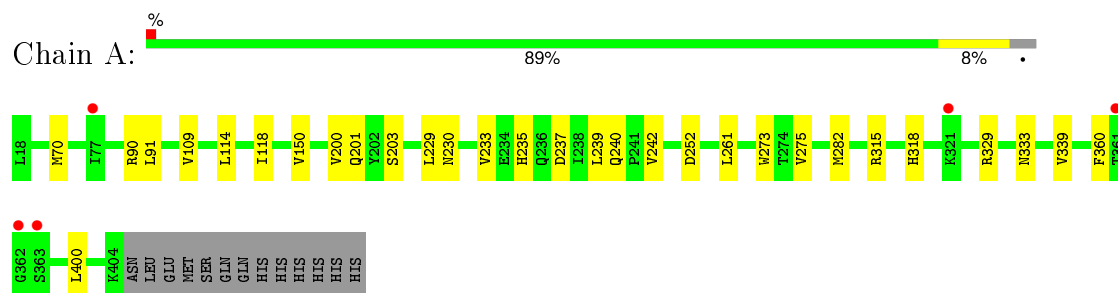
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	531	Total	O	0	0
			531	531		
10	B	460	Total	O	0	0
			460	460		



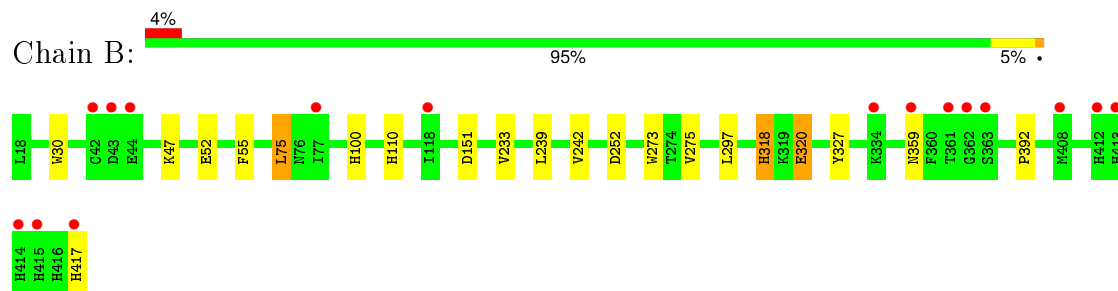
### 3 Residue-property plots [i](#)

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 ( $\text{RSRZ} > 2$ ). 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: Alpha-N-acetylgalactosaminidase



#### • Molecule 1: Alpha-N-acetylgalactosaminidase



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	154.06 Å   114.62 Å   68.63 Å 90.00°   95.82°   90.00°	Depositor
Resolution (Å)	26.63 – 1.40 26.63 – 1.40	Depositor EDS
% Data completeness (in resolution range)	98.6 (26.63-1.40) 98.6 (26.63-1.40)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.08 (at 1.40 Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.141   ,   0.178 0.151   ,   0.184	Depositor DCC
$R_{free}$ test set	11645 reflections (5.37%)	DCC
Wilson B-factor (Å <sup>2</sup> )	18.0	Xtriage
Anisotropy	0.266	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 48.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtriage
Outliers	0 of 228268 reflections	Xtriage
$F_o, F_c$ correlation	0.98	EDS
Total number of atoms	7713	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 4.61% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, BMA, NAG, FUC, CIT, ACY, DJN, 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	0.58	0/3227	0.74	3/4388 (0.1%)
1	B	0.54	0/3326	0.69	0/4523
All	All	0.56	0/6553	0.71	3/8911 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	329	ARG	NE-CZ-NH2	-5.40	117.60	120.30
1	A	282	MET	CA-CB-CG	5.30	122.31	113.30
1	A	261	LEU	CB-CG-CD2	5.13	119.72	111.00

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	A	3136	0	3044	19	0
1	B	3223	0	3116	13	0
2	A	39	0	34	0	0
3	A	61	0	52	0	0
3	B	61	0	52	0	0
4	A	14	0	13	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	14	0	13	0	0
5	A	14	0	16	0	0
5	B	14	0	16	0	0
6	A	13	0	5	0	0
6	B	13	0	5	0	0
7	A	36	0	48	1	0
7	B	42	0	56	0	0
8	A	4	0	3	0	0
9	B	38	0	34	0	0
10	A	531	0	0	9	0
10	B	460	0	0	3	0
All	All	7713	0	6507	32	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 2.

All (32) 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:200:VAL:HG23	10:A:1105:HOH:O	1.43	1.17
1:A:233[A]:VAL:HG23	10:A:715:HOH:O	1.44	1.15
1:A:229[B]:LEU:O	1:A:233[B]:VAL:HG13	1.70	0.89
1:A:233[B]:VAL:HG12	10:A:715:HOH:O	1.76	0.84
1:B:55:PHE:HD1	1:B:75[A]:LEU:HD11	1.55	0.72
1:B:110:HIS:HD2	10:B:640:HOH:O	1.74	0.70
1:B:110:HIS:HE1	1:B:151:ASP:OD2	1.77	0.67
1:A:275[A]:VAL:HG11	1:A:339:VAL:HG23	1.79	0.65
1:B:52:GLU:OE2	1:B:100:HIS:HD2	1.83	0.61
1:A:109:VAL:HG13	1:A:114:LEU:HB2	1.82	0.61
1:A:230:ASN:HB2	7:A:517:GOL:O3	2.01	0.60
1:A:237:ASP:OD1	1:A:315[A]:ARG:NH2	2.37	0.57
1:A:233[B]:VAL:CG1	10:A:715:HOH:O	2.44	0.56
1:A:240:GLN:HG3	1:A:315[A]:ARG:HG2	1.92	0.51
1:B:75[A]:LEU:HD13	1:B:75[A]:LEU:C	2.32	0.49
1:A:118:ILE:HG22	1:A:150:VAL:HG11	1.94	0.49
1:B:417:HIS:C	10:B:997:HOH:O	2.50	0.49
1:B:75[A]:LEU:CD1	1:B:75[A]:LEU:C	2.81	0.49
1:A:275[A]:VAL:HG11	1:A:339:VAL:CG2	2.43	0.49
10:A:794:HOH:O	1:B:392:PRO:HB2	2.13	0.47
1:B:297:LEU:HD11	10:B:1036:HOH:O	2.14	0.47
1:A:235:HIS:HE1	10:A:1015:HOH:O	1.99	0.46

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:30:TRP:CE3	1:B:75[B]:LEU:HD12	2.52	0.45
1:B:318:HIS:ND1	1:B:320:GLU:OE2	2.50	0.45
1:A:201:GLN:NE2	10:A:811:HOH:O	2.35	0.44
1:B:239:LEU:O	1:B:242:VAL:HG22	2.18	0.44
1:A:360:PHE:CE1	1:A:400:LEU:CD2	3.01	0.43
1:A:239:LEU:O	1:A:242:VAL:HG22	2.19	0.42
1:B:233[A]:VAL:HG21	1:B:327:TYR:OH	2.19	0.41
1:A:360:PHE:CE1	1:A:400:LEU:HD23	2.54	0.41
1:A:70:MET:HE3	10:A:1002:HOH:O	2.21	0.41
1:A:90:ARG:HD3	10:A:682:HOH:O	2.20	0.40

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	392/400 (98%)	379 (97%)	12 (3%)	1 (0%)	46	19
1	B	403/400 (101%)	390 (97%)	12 (3%)	1 (0%)	52	22
All	All	795/800 (99%)	769 (97%)	24 (3%)	2 (0%)	46	19

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	252	ASP
1	B	252	ASP

### 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	340/346 (98%)	335 (98%)	5 (2%)	72	41
1	B	351/346 (101%)	343 (98%)	8 (2%)	58	21
All	All	691/692 (100%)	678 (98%)	13 (2%)	66	29

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	91	LEU
1	A	203	SER
1	A	273	TRP
1	A	318	HIS
1	A	333	ASN
1	B	47	LYS
1	B	75[A]	LEU
1	B	75[B]	LEU
1	B	273	TRP
1	B	275	VAL
1	B	318	HIS
1	B	320	GLU
1	B	359	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	143	GLN
1	A	235	HIS
1	B	100	HIS
1	B	110	HIS
1	B	143	GLN
1	B	416	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 ⓘ

16 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	501	1,2	14,14,15	0.70	0	15,19,21	1.51	3 (20%)
2	NAG	A	502	2	14,14,15	0.54	0	15,19,21	1.03	1 (6%)
2	BMA	A	503	2	11,11,12	0.40	0	14,15,17	0.76	1 (7%)
3	NAG	A	504	1,3	14,14,15	0.59	0	15,19,21	0.72	0
3	NAG	A	505	3	14,14,15	0.60	0	15,19,21	0.63	0
3	BMA	A	506	3	11,11,12	0.40	0	14,15,17	0.85	0
3	MAN	A	507	3	11,11,12	0.56	0	14,15,17	0.86	1 (7%)
3	MAN	A	508	3	11,11,12	0.72	0	14,15,17	1.19	1 (7%)
9	NAG	B	501	9,1	14,14,15	0.45	0	15,19,21	1.02	1 (6%)
9	NAG	B	502	9	14,14,15	0.53	0	15,19,21	1.00	1 (6%)
9	FUC	B	503	9	10,10,11	0.57	0	14,14,16	1.08	2 (14%)
3	NAG	B	504	1,3	14,14,15	0.62	0	15,19,21	0.86	0
3	NAG	B	505	3	14,14,15	0.64	0	15,19,21	0.79	1 (6%)
3	BMA	B	506	3	11,11,12	0.44	0	14,15,17	0.81	0
3	MAN	B	507	3	11,11,12	0.52	0	14,15,17	1.07	2 (14%)
3	MAN	B	508	3	11,11,12	0.53	0	14,15,17	1.19	2 (14%)

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	501	1,2	-	0/6/23/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	A	502	2	-	0/6/23/26	0/1/1/1
2	BMA	A	503	2	-	0/2/19/22	0/1/1/1
3	NAG	A	504	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	505	3	-	0/6/23/26	0/1/1/1
3	BMA	A	506	3	-	0/2/19/22	0/1/1/1
3	MAN	A	507	3	-	0/2/19/22	0/1/1/1
3	MAN	A	508	3	-	0/2/19/22	0/1/1/1
9	NAG	B	501	9,1	-	0/6/23/26	0/1/1/1
9	NAG	B	502	9	-	0/6/23/26	0/1/1/1
9	FUC	B	503	9	-	0/0/17/20	0/1/1/1
3	NAG	B	504	1,3	-	0/6/23/26	0/1/1/1
3	NAG	B	505	3	-	0/6/23/26	0/1/1/1
3	BMA	B	506	3	-	0/2/19/22	0/1/1/1
3	MAN	B	507	3	-	0/2/19/22	0/1/1/1
3	MAN	B	508	3	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	501	NAG	O7-C7-C8	-2.81	116.91	122.06
3	B	505	NAG	O4-C4-C3	-2.06	105.69	110.34
2	A	503	BMA	C1-O5-C5	2.00	114.79	112.25
9	B	503	FUC	C1-O5-C5	2.04	115.53	112.38
9	B	502	NAG	C4-C3-C2	2.07	114.44	111.23
3	B	507	MAN	C1-C2-C3	2.12	112.05	109.54
9	B	503	FUC	O5-C5-C6	2.16	109.70	106.13
2	A	501	NAG	O7-C7-N2	2.42	126.81	121.86
2	A	502	NAG	C1-O5-C5	2.43	115.33	112.25
3	A	507	MAN	C1-C2-C3	2.45	112.44	109.54
3	B	508	MAN	C1-C2-C3	2.68	112.71	109.54
9	B	501	NAG	C1-O5-C5	3.01	116.07	112.25
3	B	507	MAN	C1-O5-C5	3.07	116.14	112.25
3	B	508	MAN	C1-O5-C5	3.27	116.40	112.25
2	A	501	NAG	C2-N2-C7	3.34	127.33	123.04
3	A	508	MAN	C1-C2-C3	3.56	113.75	109.54

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.



## 5.6 Ligand geometry

20 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$
4	NAG	A	509	1	14,14,15	0.47	0	15,19,21	0.86	1 (6%)
5	DJN	A	510	-	13,14,14	0.50	0	14,19,19	1.18	2 (14%)
6	CIT	A	511	-	3,12,12	0.78	0	3,17,17	0.84	0
7	GOL	A	512	-	5,5,5	0.40	0	5,5,5	0.30	0
7	GOL	A	513	-	5,5,5	0.40	0	5,5,5	0.28	0
7	GOL	A	514	-	5,5,5	0.28	0	5,5,5	0.46	0
8	ACY	A	515	-	1,3,3	1.46	0	0,3,3	0.00	-
7	GOL	A	516	-	5,5,5	0.29	0	5,5,5	0.37	0
7	GOL	A	517	-	5,5,5	0.40	0	5,5,5	0.26	0
7	GOL	A	518	-	5,5,5	0.39	0	5,5,5	0.28	0
4	NAG	B	509	1	14,14,15	0.45	0	15,19,21	1.36	1 (6%)
5	DJN	B	510	-	13,14,14	0.66	0	14,19,19	1.54	4 (28%)
6	CIT	B	511	-	3,12,12	0.55	0	3,17,17	1.16	0
7	GOL	B	512	-	5,5,5	0.30	0	5,5,5	0.29	0
7	GOL	B	513	-	5,5,5	0.36	0	5,5,5	0.34	0
7	GOL	B	514	-	5,5,5	0.36	0	5,5,5	0.28	0
7	GOL	B	515	-	5,5,5	0.49	0	5,5,5	0.45	0
7	GOL	B	516	-	5,5,5	0.38	0	5,5,5	0.46	0
7	GOL	B	517	-	5,5,5	0.35	0	5,5,5	0.24	0
7	GOL	B	518	-	5,5,5	0.30	0	5,5,5	0.38	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
4	NAG	A	509	1	-	0/6/23/26	0/1/1/1
5	DJN	A	510	-	-	0/6/23/23	0/1/1/1
6	CIT	A	511	-	-	0/6/16/16	0/0/0/0
7	GOL	A	512	-	-	0/4/4/4	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
7	GOL	A	513	-	-	0/4/4/4	0/0/0/0
7	GOL	A	514	-	-	0/4/4/4	0/0/0/0
8	ACY	A	515	-	-	0/0/0/0	0/0/0/0
7	GOL	A	516	-	-	0/4/4/4	0/0/0/0
7	GOL	A	517	-	-	0/4/4/4	0/0/0/0
7	GOL	A	518	-	-	0/4/4/4	0/0/0/0
4	NAG	B	509	1	-	0/6/23/26	0/1/1/1
5	DJN	B	510	-	-	0/6/23/23	0/1/1/1
6	CIT	B	511	-	-	0/6/16/16	0/0/0/0
7	GOL	B	512	-	-	0/4/4/4	0/0/0/0
7	GOL	B	513	-	-	0/4/4/4	0/0/0/0
7	GOL	B	514	-	-	0/4/4/4	0/0/0/0
7	GOL	B	515	-	-	0/4/4/4	0/0/0/0
7	GOL	B	516	-	-	0/4/4/4	0/0/0/0
7	GOL	B	517	-	-	0/4/4/4	0/0/0/0
7	GOL	B	518	-	-	0/4/4/4	0/0/0/0

There are no bond length outliers.

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	510	DJN	C4-C3-C2	-3.54	105.73	111.23
5	A	510	DJN	C4-C3-C2	-2.65	107.11	111.23
5	B	510	DJN	C2-N2-C7	-2.25	120.15	123.04
5	B	510	DJN	C1-C2-N2	-2.11	107.18	110.40
4	A	509	NAG	C1-O5-C5	2.05	114.85	112.25
5	A	510	DJN	C3-C2-N2	2.06	115.49	110.56
5	B	510	DJN	C3-C2-N2	2.34	116.17	110.56
4	B	509	NAG	C1-O5-C5	4.73	118.25	112.25

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
7	A	517	GOL	1	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

### 6.1 Protein, DNA and RNA chains

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	387/400 (96%)	-0.52	5 (1%) 79 78	15, 21, 39, 63	0
1	B	400/400 (100%)	-0.24	16 (4%) 42 39	14, 26, 52, 93	0
All	All	787/800 (98%)	-0.38	21 (2%) 58 55	14, 23, 47, 93	0

All (21) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	361	THR	5.0
1	B	408	MET	4.0
1	B	361	THR	3.5
1	B	413	HIS	3.4
1	B	42	CYS	3.2
1	A	363	SER	3.2
1	B	415	HIS	3.1
1	B	412	HIS	3.0
1	B	363	SER	2.9
1	B	118	ILE	2.8
1	A	362	GLY	2.7
1	A	321	LYS	2.7
1	B	77	ILE	2.5
1	B	417	HIS	2.4
1	B	359	ASN	2.4
1	B	414	HIS	2.1
1	B	43	ASP	2.1
1	B	362	GLY	2.1
1	B	44	GLU	2.0
1	A	77	ILE	2.0
1	B	334	LYS	2.0

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates ⓘ

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
3	NAG	A	504	14/15	0.95	0.07	0.53	25,31,40,45	0
3	NAG	B	504	14/15	0.97	0.04	-1.97	26,28,33,38	0
3	BMA	A	506	11/12	0.92	0.10	-	36,58,70,80	0
9	NAG	B	501	14/15	0.90	0.18	-	51,68,97,100	0
3	NAG	A	505	14/15	0.95	0.11	-	32,38,55,63	0
9	FUC	B	503	10/11	0.77	0.43	-	99,118,124,127	0
3	NAG	B	505	14/15	0.97	0.09	-	29,40,53,61	0
2	BMA	A	503	11/12	0.54	0.39	-	87,122,123,128	0
2	NAG	A	502	14/15	0.82	0.33	-	82,94,107,116	0
3	MAN	B	508	11/12	0.37	0.61	-	124,126,130,130	0
3	MAN	B	507	11/12	0.73	0.30	-	87,104,112,113	0
2	NAG	A	501	14/15	0.88	0.17	-	51,64,78,83	0
3	MAN	A	507	11/12	0.82	0.23	-	79,98,103,103	0
9	NAG	B	502	14/15	0.79	0.43	-	97,113,118,121	0
3	MAN	A	508	11/12	0.27	0.38	-	103,108,115,115	0
3	BMA	B	506	11/12	0.90	0.23	-	54,68,96,111	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	GOL	A	514	6/6	0.89	0.12	8.91	26,46,47,50	0
7	GOL	B	516	6/6	0.95	0.12	8.15	29,52,65,76	0
7	GOL	A	516	6/6	0.97	0.09	5.89	30,53,66,78	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
7	GOL	B	515	6/6	0.91	0.12	4.83	32,52,55,62	0
7	GOL	B	518	6/6	0.93	0.11	3.83	26,47,51,59	0
7	GOL	A	518	6/6	0.91	0.13	2.48	28,65,71,74	0
6	CIT	A	511	13/13	0.94	0.17	1.10	33,43,53,55	0
7	GOL	A	517	6/6	0.90	0.09	0.78	44,57,65,65	0
6	CIT	B	511	13/13	0.90	0.21	0.77	48,65,73,74	0
7	GOL	B	512	6/6	0.84	0.09	0.33	43,52,66,67	0
5	DJN	A	510	14/14	0.98	0.07	0.29	15,17,19,20	0
5	DJN	B	510	14/14	0.98	0.07	0.23	21,23,26,27	0
7	GOL	B	513	6/6	0.75	0.14	-	86,93,97,98	0
8	ACY	A	515	4/4	0.92	0.10	-	27,42,54,59	0
7	GOL	A	513	6/6	0.91	0.19	-	64,83,83,85	0
7	GOL	B	517	6/6	0.89	0.25	-	41,71,96,100	0
4	NAG	A	509	14/15	0.84	0.31	-	79,97,107,109	0
7	GOL	B	514	6/6	0.81	0.16	-	73,89,90,91	0
7	GOL	A	512	6/6	0.94	0.11	-	35,60,78,80	0
4	NAG	B	509	14/15	0.79	0.35	-	85,103,114,115	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.