



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 09:10 AM GMT

PDB ID : 3HG4  
Title : Human alpha-galactosidase catalytic mechanism 3. Covalent intermediate  
Authors : Guce, A.I.; Clark, N.E.; Garman, S.C.  
Deposited on : 2009-05-13  
Resolution : 2.30 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

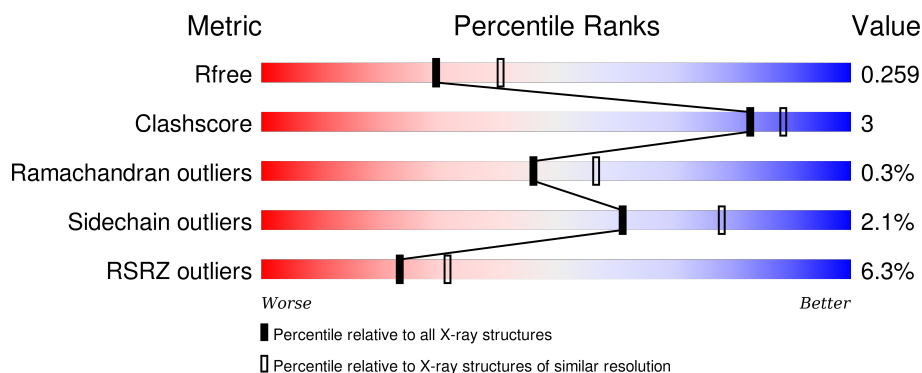
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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	3852 (2.30-2.30)
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)
RSRZ outliers	91569	3857 (2.30-2.30)

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	398	<div> <div>4%</div> <div>88%</div> <div>9%</div> <div>••</div> </div>
1	B	398	<div> <div>8%</div> <div>91%</div> <div>7%</div> <div>•</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
5	7JZ	B	802	X	-	-	-
9	GAL	B	803	X	-	-	-

## 2 Entry composition [i](#)

There are 10 unique types of molecules in this entry. The entry contains 7250 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-galactosidase A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	390	Total	C	N	O	S	0	0	0
			3122	1988	534	574	26			
1	B	392	Total	C	N	O	S	0	0	0
			3139	1998	537	577	27			

- 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
			38	22	2	14		
2	B	3	Total	C	N	O	0	0
			38	22	2	14		

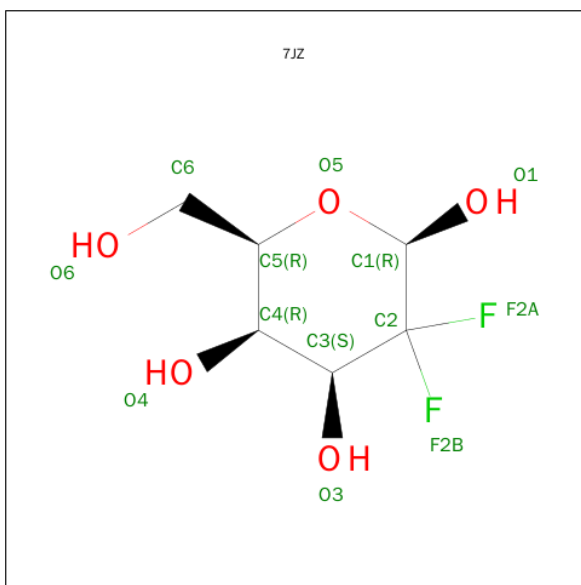
- Molecule 3 is a polymer of unknown type called SUGAR (5-MER).

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	2	Total	C	N	O	0	0
			28	16	2	10		

- Molecule 5 is 2,6-ANHYDRO-5-DEOXY-5,5-DIFLUORO-D-ARABINO-HEXITOL (three-letter code: 7JZ) (formula: C<sub>6</sub>H<sub>10</sub>F<sub>2</sub>O<sub>5</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	F	O	0	0
			12	6	2	4		
5	B	1	Total	C	F	O	0	0
			12	6	2	4		

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

*Continued on next page...*

Continued from previous page...

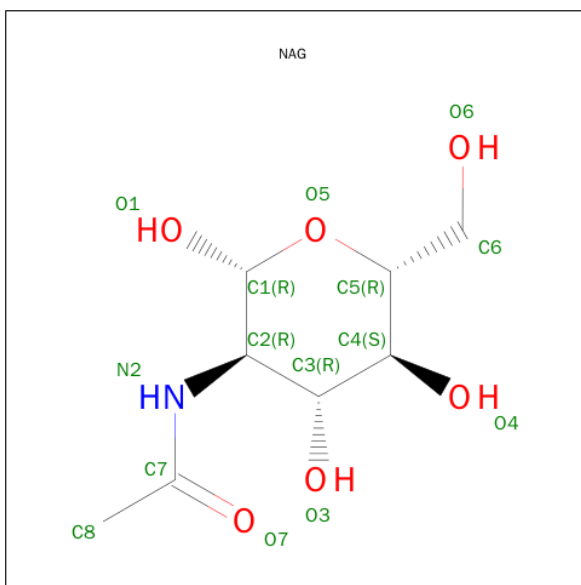
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	B	1	Total	O	S	0	0
			5	4	1		

- Molecule 7 is ACETIC ACID (three-letter code: ACY) (formula:  $C_2H_4O_2$ ).



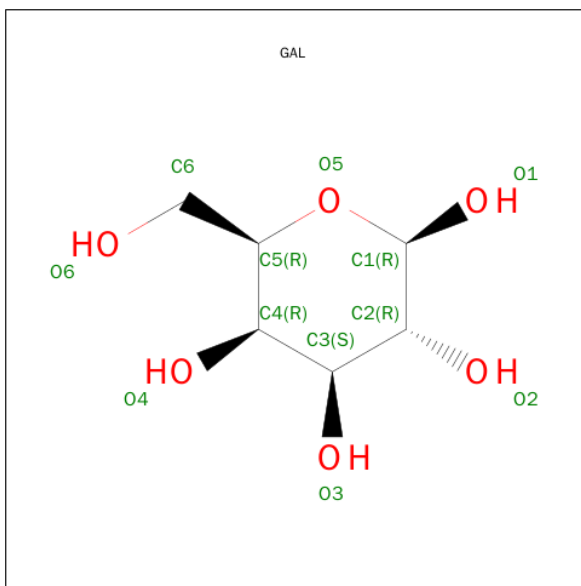
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	C	O	0	0
			4	2	2		
7	A	1	Total	C	O	0	0
			4	2	2		
7	B	1	Total	C	O	0	0
			4	2	2		
7	B	1	Total	C	O	0	0
			4	2	2		

- Molecule 8 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

- Molecule 9 is SUGAR (BETA-D-GALACTOSE) (three-letter code: GAL) (formula:  $C_6H_{12}O_6$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	B	1	Total	C	O	0	0
			12	6	6		

- Molecule 10 is water.

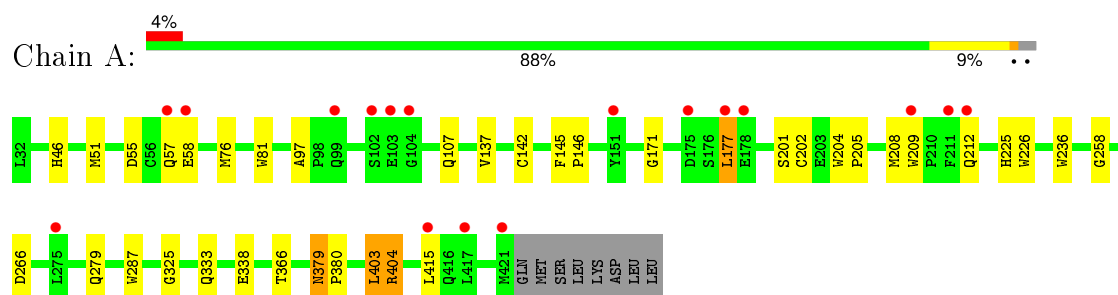
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	370	Total 370	O 370	0	0
10	B	312	Total 312	O 312	0	0



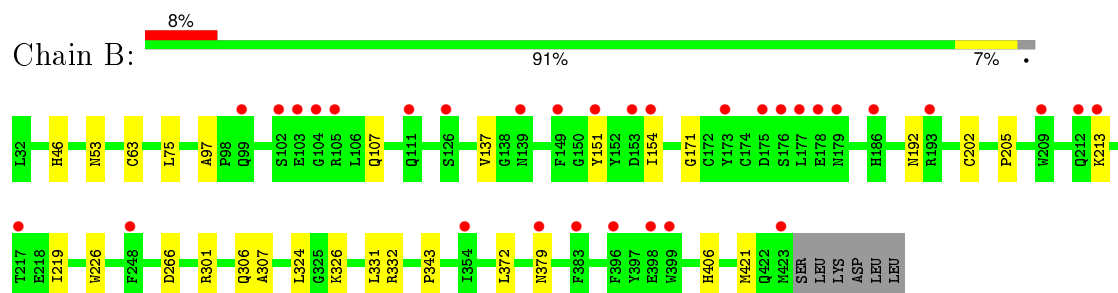
### 3 Residue-property plots [i](#)

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: Alpha-galactosidase A



#### • Molecule 1: Alpha-galactosidase A



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	90.22Å 90.22Å 216.62Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	18.05 – 2.30 18.05 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.9 (18.05-2.30) 99.8 (18.05-2.30)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	0.11	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.17 (at 2.30Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.165 , 0.221 0.211 , 0.259	Depositor DCC
$R_{free}$ test set	945 reflections (2.09%)	DCC
Wilson B-factor (Å <sup>2</sup> )	39.0	Xtriage
Anisotropy	0.004	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 48.1	EDS
Estimated twinning fraction	0.035 for -h,-k,l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 46268 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	7250	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	49.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.48% 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: BMA, NAG, GAL, 7JZ, FUC, ACY, MAN, SO4

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	0.41	0/3209	0.56	1/4358 (0.0%)
1	B	0.40	0/3226	0.53	0/4380
All	All	0.41	0/6435	0.54	1/8738 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	177	LEU	CA-CB-CG	5.72	128.46	115.30

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	3122	0	2981	20	0
1	B	3139	0	2998	15	0
2	A	38	0	34	1	0
2	B	38	0	34	0	0
3	A	61	0	52	0	0
3	B	61	0	52	5	0
4	A	28	0	25	0	0
5	A	12	0	8	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	B	12	0	8	0	0
6	A	5	0	0	0	0
6	B	10	0	0	1	0
7	A	8	0	6	0	0
7	B	8	0	6	0	0
8	B	14	0	13	0	0
9	B	12	0	12	0	0
10	A	370	0	0	2	0
10	B	312	0	0	3	0
All	All	7250	0	6229	40	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 3.

All (40) 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)
3:B:692:NAG:H81	10:B:1431:HOH:O	1.95	0.66
1:A:403:LEU:HD11	1:A:415:LEU:HD13	1.83	0.60
1:A:379:ASN:HD22	1:A:380:PRO:HA	1.67	0.59
1:B:151:TYR:HD1	1:B:154:ILE:HD11	1.68	0.59
1:B:75:LEU:HD11	1:B:301:ARG:HG2	1.85	0.59
1:B:137:VAL:HG12	1:B:171:GLY:HA2	1.86	0.57
1:B:192:ASN:HA	3:B:692:NAG:H83	1.86	0.56
1:B:205:PRO:HG3	1:B:219:ILE:HD13	1.86	0.56
1:A:204:TRP:HB3	1:A:205:PRO:HD3	1.88	0.56
1:A:201:SER:HB2	1:A:225:HIS:CE1	2.41	0.56
3:B:694:BMA:H62	3:B:696:MAN:H5	1.89	0.54
1:B:75:LEU:CD1	1:B:301:ARG:HG2	2.38	0.53
3:B:694:BMA:C6	3:B:696:MAN:H5	2.39	0.53
1:A:366:THR:HG22	1:A:404:ARG:HB3	1.91	0.53
2:A:639:NAG:H61	2:A:640:NAG:H82	1.92	0.51
1:B:202:CYS:O	1:B:226:TRP:HA	2.11	0.50
1:A:177:LEU:HD22	10:A:1374:HOH:O	2.10	0.50
1:A:236:TRP:CE3	1:A:279:GLN:HG2	2.48	0.49
1:A:137:VAL:HG12	1:A:171:GLY:HA2	1.96	0.47
1:B:97:ALA:HB3	1:B:107:GLN:HG3	1.97	0.47
1:B:406:HIS:ND1	6:B:822:SO4:O4	2.33	0.47
1:A:333:GLN:HG3	1:A:338:GLU:HG2	1.97	0.46
1:B:306:GLN:HG2	10:B:1468:HOH:O	2.14	0.46
1:A:97:ALA:HB3	1:A:107:GLN:HG3	1.98	0.45

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:204:TRP:HE1	1:A:208:MET:CE	2.29	0.45
1:B:151:TYR:CD1	1:B:154:ILE:HD11	2.49	0.45
1:B:331:LEU:HG	1:B:332:ARG:HG3	1.99	0.45
1:A:142:CYS:HB2	5:A:801:7JZ:O6	2.18	0.44
1:A:202:CYS:O	1:A:226:TRP:HA	2.18	0.44
1:B:307:ALA:HA	10:B:1670:HOH:O	2.18	0.42
1:A:55:ASP:OD2	1:A:58:GLU:HB2	2.18	0.42
1:A:57:GLN:HG2	10:A:1511:HOH:O	2.21	0.41
1:A:145:PHE:HB3	1:A:146:PRO:CD	2.50	0.41
1:B:326:LYS:HG2	1:B:343:PRO:HG2	2.01	0.41
1:A:379:ASN:HD22	1:A:380:PRO:CA	2.31	0.41
1:A:258:GLY:HA2	1:A:325:GLY:O	2.20	0.41
1:A:76:MET:HA	1:A:81:TRP:HB2	2.03	0.40
1:A:403:LEU:HD11	1:A:415:LEU:CD1	2.49	0.40
3:B:694:BMA:H62	3:B:696:MAN:C5	2.51	0.40
1:B:53:ASN:HB3	1:B:63:CYS:HA	2.04	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	388/398 (98%)	377 (97%)	10 (3%)	1 (0%)	46	57
1	B	390/398 (98%)	378 (97%)	11 (3%)	1 (0%)	46	57
All	All	778/796 (98%)	755 (97%)	21 (3%)	2 (0%)	46	57

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	266	ASP
1	B	266	ASP

### 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	331/339 (98%)	323 (98%)	8 (2%)	57	74
1	B	333/339 (98%)	327 (98%)	6 (2%)	66	82
All	All	664/678 (98%)	650 (98%)	14 (2%)	61	78

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	46	HIS
1	A	51	MET
1	A	209	TRP
1	A	212	GLN
1	A	287	TRP
1	A	379	ASN
1	A	403	LEU
1	A	404	ARG
1	B	46	HIS
1	B	213	LYS
1	B	324	LEU
1	B	372	LEU
1	B	379	ASN
1	B	421	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	122	ASN
1	A	225	HIS
1	A	379	ASN
1	B	179	ASN
1	B	212	GLN
1	B	228	ASN
1	B	379	ASN

### 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 ⓘ

18 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	639	1,2	14,14,15	0.51	0	15,19,21	0.86	0
2	NAG	A	640	2	14,14,15	0.56	0	15,19,21	1.05	1 (6%)
2	FUC	A	644	2	10,10,11	0.59	0	14,14,16	1.63	3 (21%)
3	NAG	A	692	1,3	14,14,15	0.52	0	15,19,21	0.82	1 (6%)
3	NAG	A	693	3	14,14,15	0.49	0	15,19,21	0.91	1 (6%)
3	BMA	A	694	3	11,11,12	0.47	0	14,15,17	1.11	1 (7%)
3	MAN	A	695	3	11,11,12	0.77	0	14,15,17	1.98	3 (21%)
3	MAN	A	696	3	11,11,12	0.59	0	14,15,17	0.89	1 (7%)
4	NAG	A	715	1,4	14,14,15	0.50	0	15,19,21	0.97	1 (6%)
4	NAG	A	716	4	14,14,15	0.42	0	15,19,21	1.10	1 (6%)
2	NAG	B	639	1,2	14,14,15	0.61	0	15,19,21	1.11	2 (13%)
2	NAG	B	640	2	14,14,15	0.50	0	15,19,21	0.67	0
2	FUC	B	644	2	10,10,11	0.64	0	14,14,16	1.09	2 (14%)
3	NAG	B	692	1,3	14,14,15	0.48	0	15,19,21	0.88	1 (6%)
3	NAG	B	693	3	14,14,15	0.53	0	15,19,21	0.77	0
3	BMA	B	694	3	11,11,12	0.46	0	14,15,17	1.80	3 (21%)
3	MAN	B	695	3	11,11,12	0.58	0	14,15,17	0.72	0
3	MAN	B	696	3	11,11,12	0.57	0	14,15,17	1.60	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	639	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	640	2	-	0/6/23/26	0/1/1/1
2	FUC	A	644	2	-	0/0/17/20	0/1/1/1
3	NAG	A	692	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	693	3	-	0/6/23/26	0/1/1/1
3	BMA	A	694	3	-	0/2/19/22	0/1/1/1
3	MAN	A	695	3	-	0/2/19/22	0/1/1/1
3	MAN	A	696	3	-	0/2/19/22	0/1/1/1
4	NAG	A	715	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	716	4	-	0/6/23/26	0/1/1/1
2	NAG	B	639	1,2	-	0/6/23/26	0/1/1/1
2	NAG	B	640	2	-	0/6/23/26	0/1/1/1
2	FUC	B	644	2	-	0/0/17/20	0/1/1/1
3	NAG	B	692	1,3	-	0/6/23/26	0/1/1/1
3	NAG	B	693	3	-	0/6/23/26	0/1/1/1
3	BMA	B	694	3	-	0/2/19/22	0/1/1/1
3	MAN	B	695	3	-	0/2/19/22	0/1/1/1
3	MAN	B	696	3	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

All (23) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	694	BMA	C6-C5-C4	-2.11	107.81	113.02
2	B	644	FUC	C1-C2-C3	2.04	111.96	109.54
2	B	639	NAG	C3-C4-C5	2.09	113.83	110.20
4	A	715	NAG	C4-C3-C2	2.13	114.54	111.23
3	A	692	NAG	C1-O5-C5	2.19	115.03	112.25
3	A	696	MAN	C1-O5-C5	2.20	115.04	112.25
3	A	695	MAN	C3-C4-C5	2.39	114.36	110.20
2	B	644	FUC	O5-C5-C6	2.40	110.10	106.13
2	A	644	FUC	C1-O5-C5	2.44	116.15	112.38
3	A	693	NAG	C1-O5-C5	2.59	115.53	112.25
3	B	692	NAG	C1-O5-C5	2.60	115.55	112.25
4	A	716	NAG	C1-O5-C5	2.61	115.56	112.25
2	A	640	NAG	C4-C3-C2	2.66	115.37	111.23
2	B	639	NAG	C4-C3-C2	2.75	115.50	111.23
2	A	644	FUC	O5-C5-C6	2.88	110.89	106.13

*Continued on next page...*



Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	694	BMA	C1-O5-C5	3.03	116.09	112.25
3	A	695	MAN	C2-C3-C4	3.55	117.07	111.04
3	A	694	BMA	C1-O5-C5	3.58	116.79	112.25
3	B	696	MAN	C1-C2-C3	3.84	114.09	109.54
3	B	696	MAN	C1-O5-C5	3.99	117.31	112.25
2	A	644	FUC	C1-C2-C3	4.01	114.28	109.54
3	B	694	BMA	C3-C4-C5	4.40	117.86	110.20
3	A	695	MAN	C1-C2-C3	5.64	116.21	109.54

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	639	NAG	1	0
2	A	640	NAG	1	0
3	B	692	NAG	2	0
3	B	694	BMA	3	0
3	B	696	MAN	3	0

## 5.6 Ligand geometry [i](#)

11 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	$\# Z  > 2$
5	7JZ	A	801	1	11,12,13	0.65	0	13,18,20	1.42	2 (15%)
6	SO4	A	823	-	4,4,4	0.38	0	6,6,6	0.22	0
7	ACY	A	842	-	1,3,3	1.31	0	0,3,3	0.00	-
7	ACY	A	843	-	1,3,3	1.33	0	0,3,3	0.00	-
8	NAG	B	715	1	14,14,15	0.46	0	15,19,21	1.12	2 (13%)
5	7JZ	B	802	1	11,12,13	0.61	0	13,18,20	1.65	3 (23%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
9	GAL	B	803	-	12,12,12	0.56	0	17,17,17	0.96	1 (5%)
6	SO4	B	821	-	4,4,4	0.26	0	6,6,6	0.12	0
6	SO4	B	822	-	4,4,4	2.72	3 (75%)	6,6,6	0.39	0
7	ACY	B	841	-	1,3,3	1.22	0	0,3,3	0.00	-
7	ACY	B	844	-	1,3,3	1.16	0	0,3,3	0.00	-

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	7JZ	A	801	1	-	0/2/22/26	0/1/1/1
6	SO4	A	823	-	-	0/0/0/0	0/0/0/0
7	ACY	A	842	-	-	0/0/0/0	0/0/0/0
7	ACY	A	843	-	-	0/0/0/0	0/0/0/0
8	NAG	B	715	1	-	0/6/23/26	0/1/1/1
5	7JZ	B	802	1	1/1/4/5	0/2/22/26	0/1/1/1
9	GAL	B	803	-	1/1/5/5	0/2/22/22	0/1/1/1
6	SO4	B	821	-	-	0/0/0/0	0/0/0/0
6	SO4	B	822	-	-	0/0/0/0	0/0/0/0
7	ACY	B	841	-	-	0/0/0/0	0/0/0/0
7	ACY	B	844	-	-	0/0/0/0	0/0/0/0

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	B	822	SO4	O1-S	2.76	1.56	1.47
6	B	822	SO4	O4-S	3.06	1.58	1.47
6	B	822	SO4	O3-S	3.54	1.60	1.47

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	802	7JZ	F2A-C2-C1	-2.27	107.21	110.27
9	B	803	GAL	C3-C4-C5	2.12	113.89	110.20
8	B	715	NAG	C3-C4-C5	2.52	114.58	110.20
8	B	715	NAG	C1-O5-C5	2.65	115.61	112.25
5	A	801	7JZ	C1-O5-C5	2.71	116.19	112.11
5	B	802	7JZ	C1-O5-C5	2.76	116.26	112.11
5	A	801	7JZ	F2B-C2-F2A	3.46	109.63	105.11
5	B	802	7JZ	F2B-C2-F2A	3.95	110.26	105.11

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	B	802	7JZ	C1
9	B	803	GAL	C4

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	801	7JZ	1	0
6	B	822	SO4	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	390/398 (97%)	0.30	17 (4%) 38 47	41, 47, 53, 63	0
1	B	392/398 (98%)	0.38	32 (8%) 14 20	41, 47, 54, 68	0
All	All	782/796 (98%)	0.34	49 (6%) 23 31	41, 47, 53, 68	0

All (49) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	421	MET	7.1
1	B	102	SER	5.8
1	A	211	PHE	5.3
1	A	57	GLN	4.5
1	A	102	SER	4.2
1	B	423	MET	4.2
1	B	173	TYR	3.8
1	B	149	PHE	3.7
1	B	103	GLU	3.7
1	B	399	TRP	3.6
1	B	193	ARG	3.4
1	B	175	ASP	3.4
1	A	177	LEU	3.4
1	B	213	LYS	3.3
1	B	178	GLU	3.2
1	B	209	TRP	3.1
1	A	212	GLN	3.0
1	B	177	LEU	2.8
1	B	99	GLN	2.8
1	A	58	GLU	2.8
1	A	415	LEU	2.8
1	A	175	ASP	2.8
1	B	126	SER	2.7
1	A	178	GLU	2.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	176	SER	2.5
1	B	186	HIS	2.4
1	B	104	GLY	2.4
1	B	383	PHE	2.4
1	B	105	ARG	2.3
1	A	417	LEU	2.3
1	B	154	ILE	2.3
1	B	354	ILE	2.3
1	A	103	GLU	2.2
1	B	151	TYR	2.2
1	B	212	GLN	2.2
1	B	179	ASN	2.2
1	B	398	GLU	2.2
1	A	99	GLN	2.2
1	B	248	PHE	2.1
1	A	151	TYR	2.1
1	A	209	TRP	2.1
1	B	111	GLN	2.1
1	A	275	LEU	2.1
1	A	104	GLY	2.1
1	B	139	ASN	2.1
1	B	379	ASN	2.1
1	B	217	THR	2.0
1	B	396	PHE	2.0
1	B	153	ASP	2.0

## 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	639	14/15	0.83	0.28	1.34	70,74,80,81	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
2	NAG	B	639	14/15	0.83	0.31	0.79	73,77,85,86	0
3	NAG	A	693	14/15	0.86	0.21	-0.08	66,70,73,78	0
3	NAG	B	692	14/15	0.90	0.15	-0.29	61,65,66,70	0
3	NAG	A	692	14/15	0.94	0.11	-0.60	52,54,57,62	0
4	NAG	A	716	14/15	0.87	0.47	-	84,86,87,87	0
2	NAG	B	640	14/15	0.70	0.57	-	87,90,90,91	0
3	BMA	A	694	11/12	0.71	0.42	-	83,87,89,91	0
2	NAG	A	640	14/15	0.70	0.46	-	82,84,84,84	0
2	FUC	B	644	10/11	0.78	0.54	-	88,90,90,90	0
3	BMA	B	694	11/12	0.82	0.38	-	86,88,91,94	0
3	MAN	A	696	11/12	0.76	0.52	-	93,94,94,94	0
2	FUC	A	644	10/11	0.87	0.40	-	83,85,85,85	0
3	MAN	A	695	11/12	0.75	0.45	-	91,91,92,92	0
3	NAG	B	693	14/15	0.88	0.24	-	74,77,79,82	0
4	NAG	A	715	14/15	0.87	0.29	-	72,77,78,81	0
3	MAN	B	695	11/12	0.85	0.29	-	87,88,88,89	0
3	MAN	B	696	11/12	0.64	0.56	-	96,97,97,97	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( $\text{\AA}^2$ )	Q<0.9
5	7JZ	B	802	12/13	0.94	0.18	1.37	45,49,50,50	0
9	GAL	B	803	12/12	0.68	0.26	1.33	74,76,76,77	0
6	SO4	B	821	5/5	0.98	0.15	0.13	70,70,70,71	0
6	SO4	A	823	5/5	0.99	0.10	-0.12	52,53,54,54	0
5	7JZ	A	801	12/13	0.94	0.13	-0.30	42,44,45,45	0
8	NAG	B	715	14/15	0.77	0.41	-	73,77,78,78	0
7	ACY	A	842	4/4	0.69	0.24	-	100,100,100,100	0
7	ACY	B	844	4/4	0.84	0.17	-	83,83,83,83	0
7	ACY	B	841	4/4	0.72	0.44	-	109,109,109,109	0
6	SO4	B	822	5/5	0.49	0.42	-	158,158,158,158	0
7	ACY	A	843	4/4	0.53	0.43	-	100,100,100,100	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.