Full wwPDB X-ray Structure Validation Report ⓘ

Feb 28, 2014 – 09:42 AM GMT

PDB ID	:	4CCC
Title	:	STRUCTURE OF MOUSE GALACTOCEREBROSIDASE WITH 4NBDG: ENZYME-SUBSTRATE COMPLEX
Authors	:	Hill, C.H.; Graham, S.C.; Read, R.J.; Deane, J.E.
Deposited on	:	2013-10-21
Resolution	:	2.09 Å(reported)

---

1 Overall quality at a glance ⓘ

The reported resolution of this entry is 2.09 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.

Metric	Percentile Ranks	Value
Rfree		0.224
Clashscore		3
Ramachandran outliers		0.5%
Sidechain outliers		0.9%
RSRZ outliers		3.7%

Worse   Better

▮ Percentile relative to all X-ray structures

◻ Percentile relative to X-ray structures of similar resolution
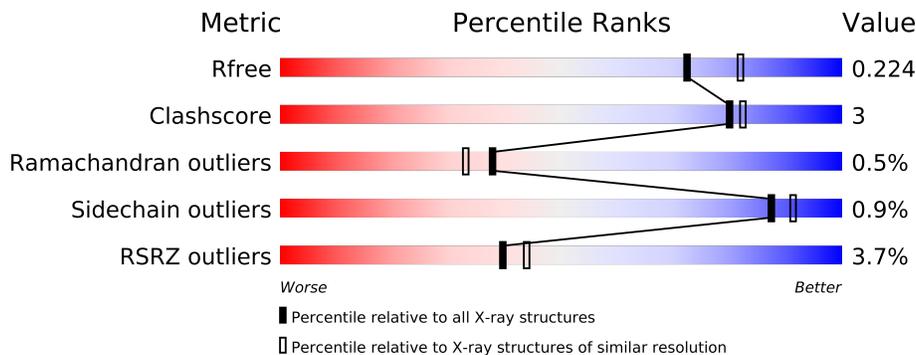
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	66092	3012 (2.10-2.10)
Clashscore	79885	3649 (2.10-2.10)
Ramachandran outliers	78287	3610 (2.10-2.10)
Sidechain outliers	78261	3611 (2.10-2.10)
RSRZ outliers	66119	3013 (2.10-2.10)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	654	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
4	NAG	A	1387	-	X

2 Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 5511 atoms, of which 0 are hydrogen and 0 are deuterium.

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 GALACTOCEREBROSIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	640	5119	3313	847	944	15	0	1	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	15	HIS	-	EXPRESSION TAG	UNP P54818
A	16	HIS	-	EXPRESSION TAG	UNP P54818
A	17	HIS	-	EXPRESSION TAG	UNP P54818
A	18	HIS	-	EXPRESSION TAG	UNP P54818
A	19	HIS	-	EXPRESSION TAG	UNP P54818
A	20	HIS	-	EXPRESSION TAG	UNP P54818
A	21	ILE	-	EXPRESSION TAG	UNP P54818
A	22	GLU	-	EXPRESSION TAG	UNP P54818
A	23	GLY	-	EXPRESSION TAG	UNP P54818
A	24	ARG	-	EXPRESSION TAG	UNP P54818

- Molecule 2 is SUGAR (1-O-[P-NITROPHENYL]-BETA-D-GALACTOPYRANOSE) (three-letter code: 147) (formula: $C_{12}H_{15}NO_8$).
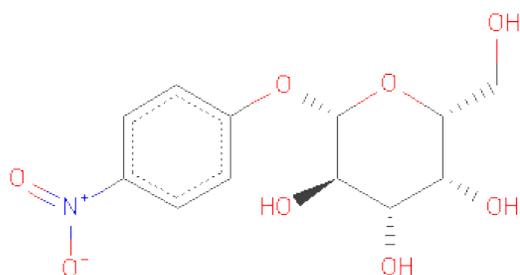
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total 21	C 12	N 1	O 8	0	0

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

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

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	15	HIS	-	EXPRESSION TAG	UNP P54818
A	16	HIS	-	EXPRESSION TAG	UNP P54818
A	17	HIS	-	EXPRESSION TAG	UNP P54818
A	18	HIS	-	EXPRESSION TAG	UNP P54818
A	19	HIS	-	EXPRESSION TAG	UNP P54818
A	20	HIS	-	EXPRESSION TAG	UNP P54818
A	21	ILE	-	EXPRESSION TAG	UNP P54818
A	22	GLU	-	EXPRESSION TAG	UNP P54818
A	23	GLY	-	EXPRESSION TAG	UNP P54818
A	24	ARG	-	EXPRESSION TAG	UNP P54818
A	15	HIS	-	EXPRESSION TAG	UNP P54818

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	16	HIS	-	EXPRESSION TAG	UNP P54818
A	17	HIS	-	EXPRESSION TAG	UNP P54818
A	18	HIS	-	EXPRESSION TAG	UNP P54818
A	19	HIS	-	EXPRESSION TAG	UNP P54818
A	20	HIS	-	EXPRESSION TAG	UNP P54818
A	21	ILE	-	EXPRESSION TAG	UNP P54818
A	22	GLU	-	EXPRESSION TAG	UNP P54818
A	23	GLY	-	EXPRESSION TAG	UNP P54818
A	24	ARG	-	EXPRESSION TAG	UNP P54818
A	15	HIS	-	EXPRESSION TAG	UNP P54818
A	16	HIS	-	EXPRESSION TAG	UNP P54818
A	17	HIS	-	EXPRESSION TAG	UNP P54818
A	18	HIS	-	EXPRESSION TAG	UNP P54818
A	19	HIS	-	EXPRESSION TAG	UNP P54818
A	20	HIS	-	EXPRESSION TAG	UNP P54818
A	21	ILE	-	EXPRESSION TAG	UNP P54818
A	22	GLU	-	EXPRESSION TAG	UNP P54818
A	23	GLY	-	EXPRESSION TAG	UNP P54818
A	24	ARG	-	EXPRESSION TAG	UNP P54818

- Molecule 4 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).
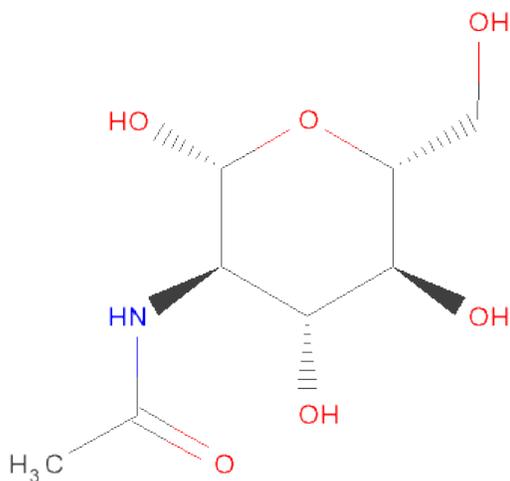
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total Ca 1 1		0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	272	Total O 272 272		0	0

3 Residue-property plots ⓘ

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: GALACTOCEREBROSIDASE

Chain A:
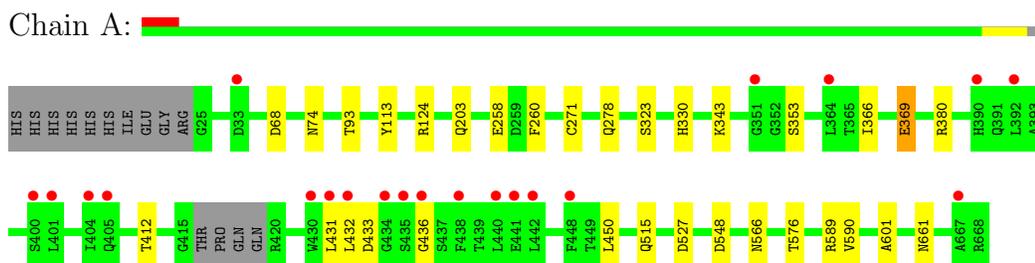
4 Data and refinement statistics (i)

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, α, β, γ	250.01Å 250.01Å 77.79Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	56.38 – 2.09 56.38 – 2.09	Depositor EDS
% Data completeness (in resolution range)	99.9 (56.38-2.09) 99.9 (56.38-2.09)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.17 (at 2.08Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.187 , 0.224 0.188 , 0.224	Depositor DCC
R_{free} test set	2776 reflections (5.08%)	DCC
Wilson B-factor (Å^2)	32.0	Xtriage
Anisotropy	0.351	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 51.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$< |L| > = 0.42, < L^2 > = 0.24$	Xtriage
Outliers	0 of 54656 reflections	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	5511	wwPDB-VP
Average B, all atoms (Å^2)	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.36% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

5   Model quality ⓘ

5.1   Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: 147, CA, 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	A	0.41	0/5282	0.54	0/7204

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2   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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	5119	0	0	17	0
2	A	21	0	15	2	0
3	A	84	0	75	0	0
4	A	14	0	13	0	0
5	A	1	0	0	0	0
6	A	272	0	0	7	1
All	All	5511	0	103	18	1

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 3.

All (18) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:278:GLN:NE2	6:A:2170:HOH:O	2.30	0.64
1:A:330:HIS:NE2	1:A:412:THR:OG1	2.33	0.61
2:A:1001:147:H5'	6:A:2271:HOH:O	2.02	0.60
1:A:566:ASN:ND2	1:A:576:THR:OG1	2.37	0.57
1:A:589:ARG:NH1	6:A:2261:HOH:O	2.38	0.57
1:A:74:ASN:N	6:A:2020:HOH:O	2.42	0.52
1:A:323:SER:N	6:A:2161:HOH:O	2.42	0.52
1:A:258:GLU:OE2	2:A:1001:147:H1	2.14	0.48
1:A:380:ARG:N	6:A:2158:HOH:O	2.46	0.48
1:A:353:SER:OG	1:A:369:GLU:OE1	2.32	0.47
1:A:548:ASP:OD2	1:A:661:ASN:ND2	2.50	0.45
1:A:343:LYS:N	6:A:2001:HOH:O	2.49	0.45
1:A:590:VAL:N	1:A:601:ALA:O	2.51	0.44
1:A:431:LEU:O	1:A:433:ASP:N	2.53	0.41
1:A:397:LEU:N	1:A:436:GLY:O	2.53	0.41
1:A:203:GLN:OE1	1:A:203:GLN:N	2.54	0.41
1:A:68:ASP:OD1	1:A:124:ARG:NE	2.54	0.41
1:A:366:ILE:N	1:A:450:LEU:O	2.55	0.40

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	Distance(Å)	Clash(Å)
6:A:2096:HOH:O	6:A:2096:HOH:O[4_556]	2.16	0.04

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	637/654 (97%)	607 (95%)	27 (4%)	3 (0%)	38	33

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	93	THR
1	A	260	PHE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	432	LEU

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	532/555 (96%)	527 (99%)	5 (1%)	87	91

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	113	TYR
1	A	271	CYS
1	A	369	GLU
1	A	515	GLN
1	A	527	ASP

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 chains 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 ⓘ

6 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	#$|Z| > 2$
3	NAG	A	1284	1,3	12,14,15	0.83	1 (8%)	15,19,21	1.27	2 (13%)
3	NAG	A	1285	3	12,14,15	0.74	1 (8%)	15,19,21	1.21	3 (20%)
3	NAG	A	1363	1,3	12,14,15	0.61	0	15,19,21	1.00	0
3	NAG	A	1364	3	12,14,15	0.58	0	15,19,21	0.99	1 (6%)
3	NAG	A	1542	1,3	12,14,15	0.73	1 (8%)	15,19,21	0.95	0
3	NAG	A	1543	3	12,14,15	0.56	0	15,19,21	1.06	1 (6%)

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	A	1284	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	1285	3	-	0/6/23/26	0/1/1/1
3	NAG	A	1363	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	1364	3	-	0/6/23/26	0/1/1/1
3	NAG	A	1542	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	1543	3	-	0/6/23/26	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1284	NAG	O5-C5	-2.48	1.40	1.45
3	A	1542	NAG	O5-C5	-2.28	1.41	1.45
3	A	1285	NAG	O5-C5	-2.27	1.41	1.45

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	1543	NAG	O5-C5-C6	3.29	110.43	106.98
3	A	1285	NAG	O5-C5-C4	-2.55	107.42	110.65
3	A	1284	NAG	C4-C3-C2	2.48	117.38	111.32
3	A	1285	NAG	O5-C5-C6	2.42	109.52	106.98
3	A	1364	NAG	C2-N2-C7	-2.28	119.26	123.09
3	A	1285	NAG	C3-C2-N2	-2.21	108.39	111.76
3	A	1284	NAG	C3-C2-N2	-2.11	108.55	111.76

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.6 Ligand geometry (i)

Of 3 ligands modelled in this entry, 1 is monoatomic - leaving 2 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	#$|Z| > 2$
2	147	A	1001	-	22,22,22	1.54	6 (27%)	31,31,31	1.34	5 (16%)
4	NAG	A	1387	1	12,14,15	0.59	0	15,19,21	0.94	1 (6%)

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	147	A	1001	-	-	0/10/30/30	0/2/2/2
4	NAG	A	1387	1	-	0/6/23/26	0/1/1/1

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1001	147	O2'-N1'	2.88	1.31	1.25
2	A	1001	147	C3-C2	-2.71	1.45	1.52
2	A	1001	147	C6'-C5'	2.22	1.43	1.38
2	A	1001	147	C4-C3	-2.22	1.46	1.52
2	A	1001	147	C3'-C2'	2.20	1.43	1.38
2	A	1001	147	O1'-C1'	2.08	1.41	1.38

All (6) bond angle outliers are listed below:

Full wwPDB X-ray Structure Validation Report

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	1387	NAG	O5-C5-C6	3.04	110.17	106.98
2	A	1001	147	O2'-N1'-C4'	2.90	119.99	114.42
2	A	1001	147	O5-C1-O1'	2.65	115.40	108.29
2	A	1001	147	O4-C4-C3	-2.64	104.44	110.35
2	A	1001	147	O2-C2-C1	-2.48	104.62	110.04
2	A	1001	147	O3-C3-C4	-2.08	105.69	110.35

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6　Fit of model and data ⓘ

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, 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	A	640/654 (97%)	-0.15	24 (3%)	38	43	29, 53, 86, 121	0

All (24) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	404	ILE	6.7
1	A	432	LEU	5.3
1	A	431	LEU	4.9
1	A	436	GLY	4.2
1	A	390	HIS	3.8
1	A	405	GLN	3.4
1	A	438	PHE	3.3
1	A	401	LEU	3.2
1	A	442	LEU	3.0
1	A	392	LEU	2.9
1	A	33	ASP	2.8
1	A	441	GLU	2.8
1	A	397	LEU	2.7
1	A	351	GLY	2.7
1	A	434	GLY	2.6
1	A	448	PHE	2.6
1	A	440	LEU	2.6
1	A	430	TRP	2.5
1	A	400	SER	2.4
1	A	394	THR	2.3
1	A	435	SER	2.3
1	A	364	LEU	2.3
1	A	667	ALA	2.2
1	A	395	PHE	2.1

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, 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	RSR	LLDF	B-factors(Å^2)	Q<0.9
3	NAG	A	1284	14/15	0.16	4.00	70,83,87,98	0
3	NAG	A	1543	14/15	0.18	2.17	85,88,91,92	0
3	NAG	A	1363	14/15	0.20	1.18	71,80,86,93	0
3	NAG	A	1542	14/15	0.11	-0.39	59,68,74,79	0
3	NAG	A	1285	14/15	0.33	-	107,111,113,114	0
3	NAG	A	1364	14/15	0.31	-	95,99,102,102	0

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. 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, 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	RSR	LLDF	B-factors(Å^2)	Q<0.9
4	NAG	A	1387	14/15	0.32	3.19	74,82,88,89	0
2	147	A	1001	21/21	0.09	-1.45	34,48,72,73	0
5	CA	A	1669	1/1	0.04	-4.90	72,72,72,72	0

6.5 Other polymers (i)

There are no such residues in this entry.