



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 09:04 PM GMT

PDB ID : 4UTC
Title : Crystal structure of dengue 2 virus envelope glycoprotein
Authors : Kikuti, C.; Rouvinski, A.; Guardado-Calvo, P.; Barba-Spaeth, G.; Duquerroy, S.; Vaney, M.C.; Rey, F.A.
Deposited on : 2014-07-18
Resolution : 3.08 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
<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

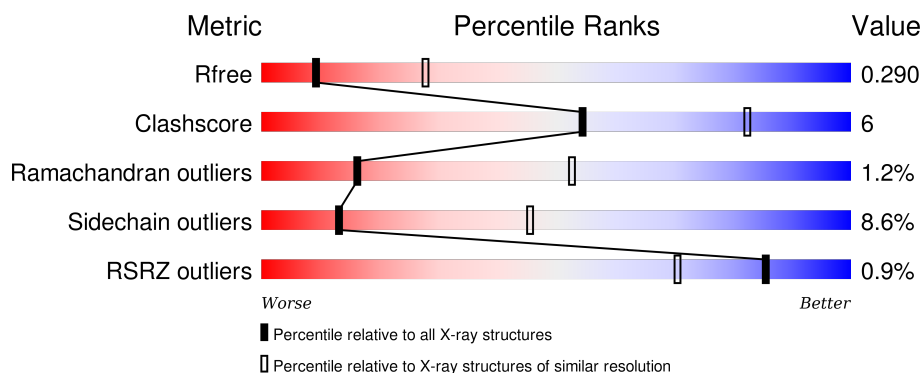
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.08 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	1119 (3.12-3.04)
Clashscore	102246	1098 (3.10-3.06)
Ramachandran outliers	100387	1057 (3.10-3.06)
Sidechain outliers	100360	1057 (3.10-3.06)
RSRZ outliers	91569	1001 (3.10-3.06)

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. 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	422	<div> <div>72%</div> <div>19%</div> <div>7%</div> </div>
1	B	422	<div> <div>72%</div> <div>18%</div> <div>8%</div> </div>

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 6243 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 ENVELOPE GLYCOPROTEIN E.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	391	Total	C	N	O	S	0	0	0
			3035	1912	521	578	24			
1	B	390	Total	C	N	O	S	0	0	0
			3037	1914	523	576	24			

There are 64 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	392	LEU	-	EXPRESSION TAG	UNP Q68Y26
A	393	ARG	-	EXPRESSION TAG	UNP Q68Y26
A	394	PRO	-	EXPRESSION TAG	UNP Q68Y26
A	395	LEU	-	EXPRESSION TAG	UNP Q68Y26
A	396	GLU	-	EXPRESSION TAG	UNP Q68Y26
A	397	SER	-	EXPRESSION TAG	UNP Q68Y26
A	398	ARG	-	EXPRESSION TAG	UNP Q68Y26
A	399	GLY	-	EXPRESSION TAG	UNP Q68Y26
A	400	PRO	-	EXPRESSION TAG	UNP Q68Y26
A	401	PHE	-	EXPRESSION TAG	UNP Q68Y26
A	402	GLU	-	EXPRESSION TAG	UNP Q68Y26
A	403	GLY	-	EXPRESSION TAG	UNP Q68Y26
A	404	LYS	-	EXPRESSION TAG	UNP Q68Y26
A	405	PRO	-	EXPRESSION TAG	UNP Q68Y26
A	406	ILE	-	EXPRESSION TAG	UNP Q68Y26
A	407	PRO	-	EXPRESSION TAG	UNP Q68Y26
A	408	ASN	-	EXPRESSION TAG	UNP Q68Y26
A	409	PRO	-	EXPRESSION TAG	UNP Q68Y26
A	410	LEU	-	EXPRESSION TAG	UNP Q68Y26
A	411	LEU	-	EXPRESSION TAG	UNP Q68Y26
A	412	GLY	-	EXPRESSION TAG	UNP Q68Y26
A	413	LEU	-	EXPRESSION TAG	UNP Q68Y26
A	414	ASP	-	EXPRESSION TAG	UNP Q68Y26
A	415	SER	-	EXPRESSION TAG	UNP Q68Y26
A	416	THR	-	EXPRESSION TAG	UNP Q68Y26

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	417	ARG	-	EXPRESSION TAG	UNP Q68Y26
A	418	THR	-	EXPRESSION TAG	UNP Q68Y26
A	419	GLY	-	EXPRESSION TAG	UNP Q68Y26
A	420	HIS	-	EXPRESSION TAG	UNP Q68Y26
A	421	HIS	-	EXPRESSION TAG	UNP Q68Y26
A	422	HIS	-	EXPRESSION TAG	UNP Q68Y26
A	118	LYS	MET	CONFLICT	UNP Q68Y26
B	392	LEU	-	EXPRESSION TAG	UNP Q68Y26
B	393	ARG	-	EXPRESSION TAG	UNP Q68Y26
B	394	PRO	-	EXPRESSION TAG	UNP Q68Y26
B	395	LEU	-	EXPRESSION TAG	UNP Q68Y26
B	396	GLU	-	EXPRESSION TAG	UNP Q68Y26
B	397	SER	-	EXPRESSION TAG	UNP Q68Y26
B	398	ARG	-	EXPRESSION TAG	UNP Q68Y26
B	399	GLY	-	EXPRESSION TAG	UNP Q68Y26
B	400	PRO	-	EXPRESSION TAG	UNP Q68Y26
B	401	PHE	-	EXPRESSION TAG	UNP Q68Y26
B	402	GLU	-	EXPRESSION TAG	UNP Q68Y26
B	403	GLY	-	EXPRESSION TAG	UNP Q68Y26
B	404	LYS	-	EXPRESSION TAG	UNP Q68Y26
B	405	PRO	-	EXPRESSION TAG	UNP Q68Y26
B	406	ILE	-	EXPRESSION TAG	UNP Q68Y26
B	407	PRO	-	EXPRESSION TAG	UNP Q68Y26
B	408	ASN	-	EXPRESSION TAG	UNP Q68Y26
B	409	PRO	-	EXPRESSION TAG	UNP Q68Y26
B	410	LEU	-	EXPRESSION TAG	UNP Q68Y26
B	411	LEU	-	EXPRESSION TAG	UNP Q68Y26
B	412	GLY	-	EXPRESSION TAG	UNP Q68Y26
B	413	LEU	-	EXPRESSION TAG	UNP Q68Y26
B	414	ASP	-	EXPRESSION TAG	UNP Q68Y26
B	415	SER	-	EXPRESSION TAG	UNP Q68Y26
B	416	THR	-	EXPRESSION TAG	UNP Q68Y26
B	417	ARG	-	EXPRESSION TAG	UNP Q68Y26
B	418	THR	-	EXPRESSION TAG	UNP Q68Y26
B	419	GLY	-	EXPRESSION TAG	UNP Q68Y26
B	420	HIS	-	EXPRESSION TAG	UNP Q68Y26
B	421	HIS	-	EXPRESSION TAG	UNP Q68Y26
B	422	HIS	-	EXPRESSION TAG	UNP Q68Y26
B	118	LYS	MET	CONFLICT	UNP Q68Y26

- 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		

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

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

- Molecule 4 is a polymer of unknown type called SUGAR (6-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	B	6	Total	C	N	O	0	0
			71	40	2	29		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	3	Total	O	0	0
			3	3		
5	B	3	Total	O	0	0
			3	3		

4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α , β , γ	105.46 Å 105.46 Å 165.90 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.08 45.37 – 3.08	Depositor EDS
% Data completeness (in resolution range)	97.5 (20.00-3.08) 97.9 (45.37-3.08)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.07 (at 3.06 Å)	Xtriage
Refinement program	BUSTER 2.11.4	Depositor
R, R_{free}	0.211 , 0.279 0.222 , 0.290	Depositor DCC
R_{free} test set	882 reflections (5.32%)	DCC
Wilson B-factor (Å ²)	86.0	Xtriage
Anisotropy	0.003	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 83.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 17535 reflections	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	6243	wwPDB-VP
Average B, all atoms (Å ²)	85.0	wwPDB-VP

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

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 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, FUC

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	0.53	0/3095	0.76	0/4183
1	B	0.51	0/3098	0.77	1/4188 (0.0%)
All	All	0.52	0/6193	0.76	1/8371 (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	B	145	SER	N-CA-C	-5.96	94.90	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	3035	0	3021	35	0
1	B	3037	0	3032	37	0
2	A	38	0	34	5	0
3	A	28	0	25	0	0
3	B	28	0	25	0	0
4	B	71	0	61	6	0
5	A	3	0	0	0	0
5	B	3	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	6243	0	6198	80	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 80 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:B:502:FUC:H61	4:B:503:NAG:H3	1.60	0.84
2:A:502:FUC:H61	2:A:503:NAG:H3	1.57	0.83
1:A:20:TRP:HA	1:A:287:LEU:O	1.82	0.79
1:B:195:GLU:HG2	1:B:209:HIS:CE1	2.20	0.76
1:A:239:THR:HB	1:A:251:VAL:HG23	1.76	0.66

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	385/422 (91%)	353 (92%)	29 (8%)	3 (1%)	24	63
1	B	386/422 (92%)	357 (92%)	23 (6%)	6 (2%)	12	44
All	All	771/844 (91%)	710 (92%)	52 (7%)	9 (1%)	16	53

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	345	ARG
1	A	176	THR
1	B	146	GLY
1	B	344	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	148	GLU

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	339/366 (93%)	312 (92%)	27 (8%)	15	49
1	B	339/366 (93%)	308 (91%)	31 (9%)	12	40
All	All	678/732 (93%)	620 (91%)	58 (9%)	13	45

5 of 58 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	390	ASN
1	B	134	ASN
1	B	351	LEU
1	B	49	GLU
1	B	84	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	200	GLN
1	A	282	HIS

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

5.5 Carbohydrates ⓘ

13 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	0.38	0	15,19,21	0.75	0
2	FUC	A	502	2	10,10,11	0.43	0	14,14,16	1.85	2 (14%)
2	NAG	A	503	2	14,14,15	0.23	0	15,19,21	0.40	0
3	NAG	A	567	1,3	14,14,15	0.26	0	15,19,21	1.17	1 (6%)
3	NAG	A	569	3	14,14,15	0.31	0	15,19,21	1.51	1 (6%)
4	NAG	B	501	1,4	14,14,15	0.40	0	15,19,21	0.76	0
4	FUC	B	502	4	10,10,11	0.49	0	14,14,16	1.92	3 (21%)
4	NAG	B	503	4	14,14,15	0.27	0	15,19,21	0.72	1 (6%)
4	BMA	B	504	4	11,11,12	0.43	0	14,15,17	1.42	1 (7%)
4	BMA	B	505	4	11,11,12	0.45	0	14,15,17	1.96	1 (7%)
4	BMA	B	506	4	11,11,12	0.51	0	14,15,17	1.10	1 (7%)
3	NAG	B	567	1,3	14,14,15	0.26	0	15,19,21	0.72	1 (6%)
3	NAG	B	569	3	14,14,15	0.32	0	15,19,21	1.50	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	NAG	A	501	1,2	-	0/6/23/26	0/1/1/1
2	FUC	A	502	2	-	0/0/17/20	0/1/1/1
2	NAG	A	503	2	-	0/6/23/26	0/1/1/1
3	NAG	A	567	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	569	3	-	0/6/23/26	0/1/1/1
4	NAG	B	501	1,4	-	0/6/23/26	0/1/1/1
4	FUC	B	502	4	-	0/0/17/20	0/1/1/1
4	NAG	B	503	4	-	0/6/23/26	0/1/1/1
4	BMA	B	504	4	-	0/2/19/22	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	BMA	B	505	4	-	0/2/19/22	1/1/1/1
4	BMA	B	506	4	-	0/2/19/22	0/1/1/1
3	NAG	B	567	1,3	-	0/6/23/26	0/1/1/1
3	NAG	B	569	3	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	567	NAG	C1-O5-C5	2.13	114.95	112.25
4	B	502	FUC	C1-C2-C3	2.24	112.19	109.54
4	B	502	FUC	O2-C2-C1	2.25	113.71	109.21
4	B	503	NAG	C1-O5-C5	2.35	115.23	112.25
2	A	502	FUC	C1-C2-C3	2.80	112.85	109.54

There are no chirality outliers.

There are no torsion outliers.

All (1) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	505	BMA	C1-C2-C3-C4-C5-O5

7 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	501	NAG	3	0
2	A	502	FUC	2	0
2	A	503	NAG	3	0
4	B	501	NAG	3	0
4	B	502	FUC	1	0
4	B	503	NAG	5	0
4	B	504	BMA	2	0

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	391/422 (92%)	-0.00	5 (1%) 79 61	46, 78, 125, 147	0
1	B	390/422 (92%)	0.11	2 (0%) 91 82	40, 86, 118, 134	0
All	All	781/844 (92%)	0.06	7 (0%) 85 70	40, 82, 121, 147	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	135	LEU	4.7
1	B	378	ILE	3.2
1	A	189	THR	2.9
1	A	395	LEU	2.2
1	A	277	LEU	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	RSCC	RSR	LLDF	B-factors(Å ²)	Q<0.9
2	NAG	A	501	14/15	0.82	0.38	1.51	129,136,140,145	0
4	NAG	B	501	14/15	0.94	0.18	-0.71	75,81,88,90	0
2	FUC	A	502	10/11	0.80	0.24	-	123,125,126,127	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
4	BMA	B	505	11/12	0.96	0.12	-	92,93,96,98	0
4	FUC	B	502	10/11	0.93	0.20	-	61,69,72,73	0
4	BMA	B	506	11/12	0.79	0.23	-	95,100,104,108	0
3	NAG	B	567	14/15	0.84	0.26	-	100,106,113,115	0
3	NAG	B	569	14/15	0.75	0.22	-	121,123,126,127	0
2	NAG	A	503	14/15	0.74	0.35	-	145,152,156,156	0
4	BMA	B	504	11/12	0.84	0.20	-	92,94,101,102	0
4	NAG	B	503	14/15	0.93	0.19	-	83,86,88,89	0
3	NAG	A	569	14/15	0.64	0.36	-	138,146,147,148	0
3	NAG	A	567	14/15	0.87	0.15	-	113,119,126,132	0

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.