wwPDB EM Map/Model Validation Report (i)

Jun 1, 2016 – 10:36 PM EDT

PDB ID	:	3JD8
EMDB ID:	:	EMD-6640
Title	:	cryo-EM structure of the full-length human NPC1 at 4.4 angstrom
Authors	:	Gong, X.; Qian, H.W.; Zhou, X.H.; Wu, J.P.; Zhou, Q.; Yan, N.
Deposited on	:	2016-05-01
Resolution	:	4.43 Å(reported)
Based on PDB ID	:	5F1B, 3GKI

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could
stem from errors in the original structure(s) used in the fitting.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
http://wwpdb.org/validation/2016/EMValidationReportHelp

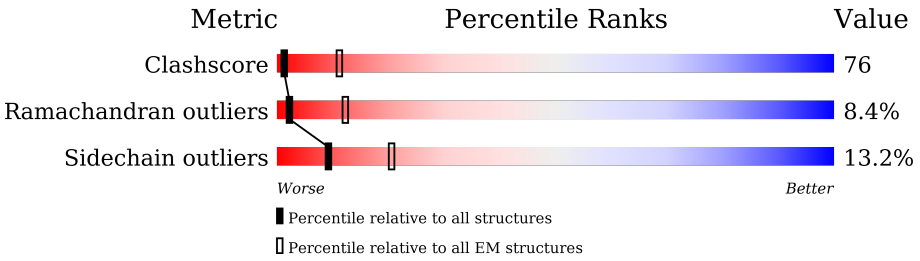
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027674

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 4.43 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)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the 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 <=5%

Mol	Chain	Length	Quality of chain
1	A	1278	48% 30% 10% · 11%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	NAG	A	1306	-	-	X	-
4	NAG	A	1307	-	-	X	-
5	NAG	A	1312	-	-	X	-

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 8050 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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.

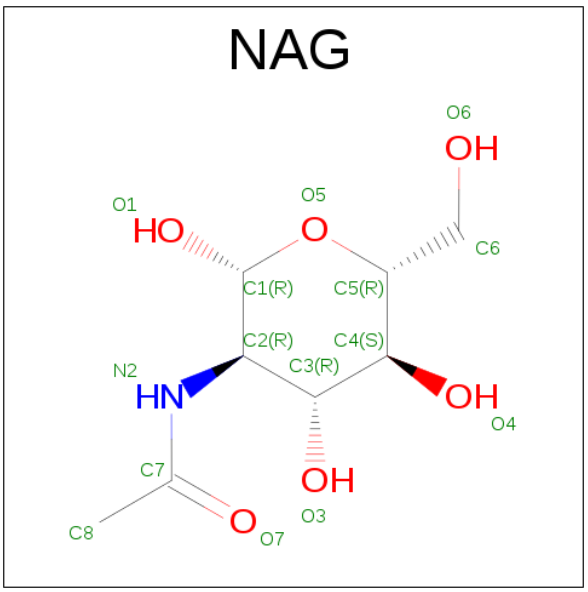
- Molecule 1 is a protein called Niemann-Pick C1 protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	1133	Total	C	N	O	S	1	0
			7695	4862	1315	1476	42		

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

Mol	Chain	Residues	Atoms				AltConf
2	A	2	Total	C	N	O	0
			112	64	8	40	
2	A	2	Total	C	N	O	0
			112	64	8	40	
2	A	2	Total	C	N	O	0
			112	64	8	40	
2	A	2	Total	C	N	O	0
			112	64	8	40	

- Molecule 3 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms				AltConf
3	A	1	Total 126	C 72	N 9	O 45	0
3	A	1	Total 126	C 72	N 9	O 45	0
3	A	1	Total 126	C 72	N 9	O 45	0
3	A	1	Total 126	C 72	N 9	O 45	0
3	A	1	Total 126	C 72	N 9	O 45	0
3	A	1	Total 126	C 72	N 9	O 45	0
3	A	1	Total 126	C 72	N 9	O 45	0
3	A	1	Total 126	C 72	N 9	O 45	0
3	A	1	Total 126	C 72	N 9	O 45	0

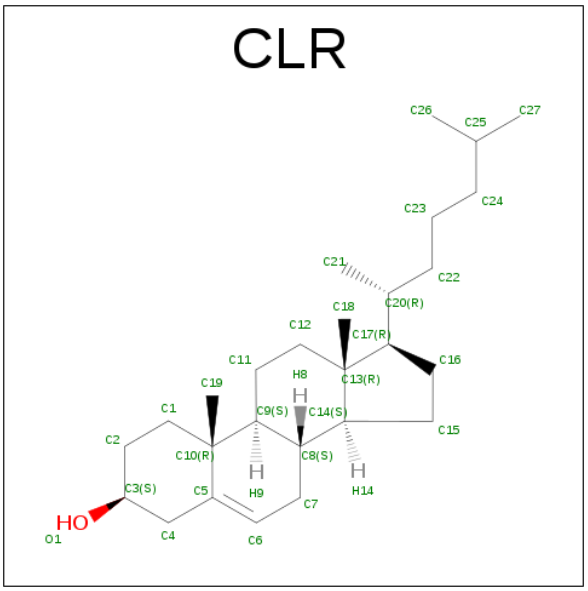
- Molecule 4 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				AltConf
4	A	3	Total 39	C 22	N 2	O 15	0

- Molecule 5 is a polymer of unknown type called SUGAR (4-MER).

Mol	Chain	Residues	Atoms				AltConf
5	A	4	Total 50	C 28	N 2	O 20	0

- Molecule 6 is CHOLESTEROL (three-letter code: CLR) (formula: $C_{27}H_{46}O$).



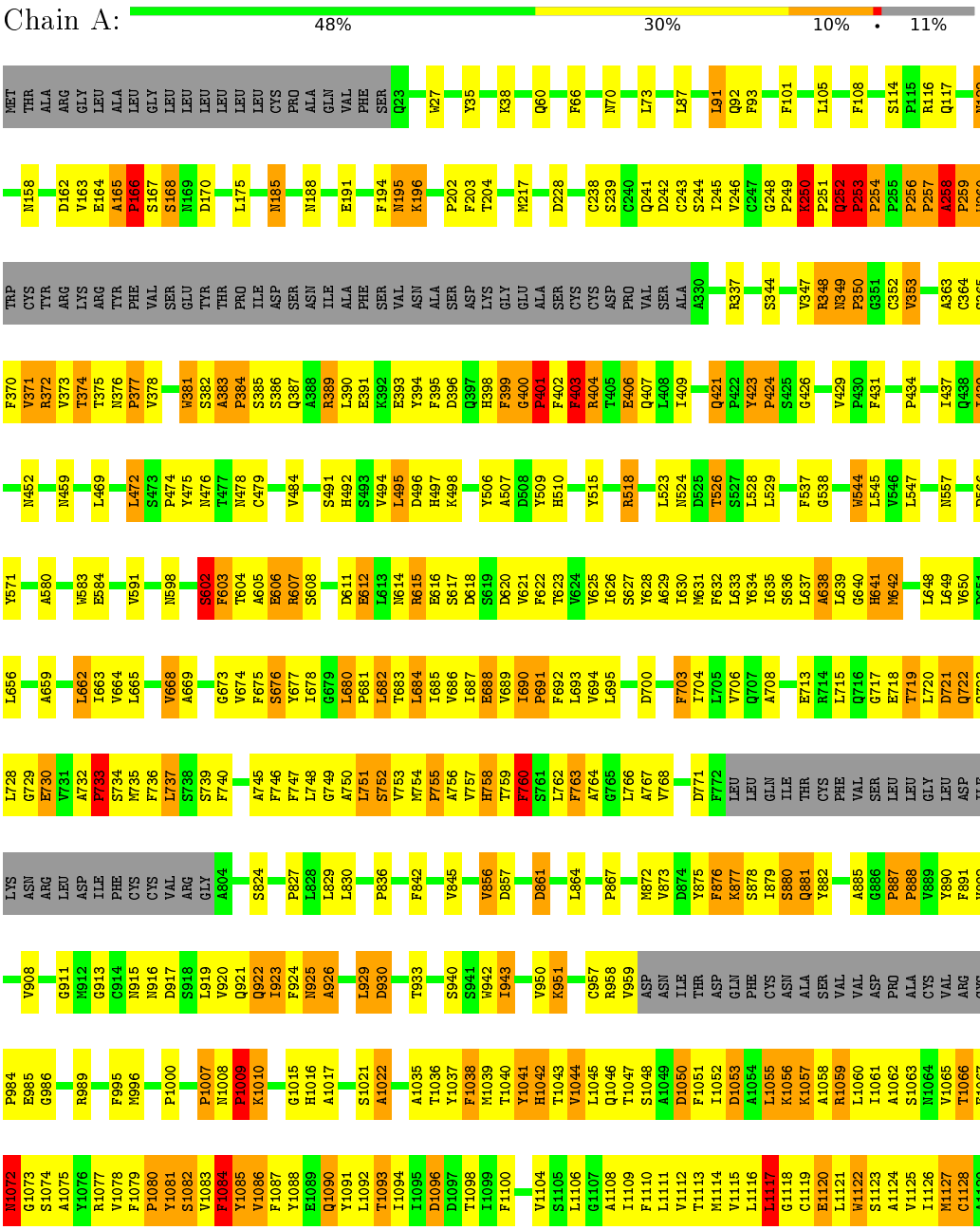
Mol	Chain	Residues	Atoms			AltConf
			Total	C	O	
6	A	1	28	27	1	0

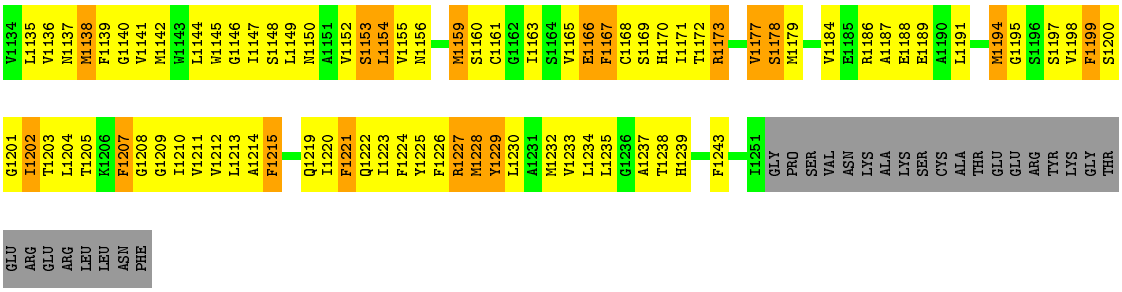
3 Residue-property plots (i)

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. 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. 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: Niemann-Pick C1 protein

Chain A:




V1134 L1135 V1136 N1137 M1138 F1139 G1140 V1141 M1142 W1143 L1144 W1145 G1146 I1147 S1148 L1149 N1150 A1151 V1152 S1153 L1154 V1155 N1156 M1159 S1160 C1161 G1162 I1163 S1164 V1165 E1166 F1167 C1168 S1169 H1170 I1171 T1172 R1173 V1177 S1178 M1179 V1184 E1185 R1186 A1187 E1188 E1189 A1190 L1191 M1194 G1195 S1196 S1197 V1198 F1199 S1200
G1201 I1202 T1203 L1204 T1205 K1206 F1207 G1208 G1209 I1210 V1211 V1212 L1213 A1214 F1215 Q1219 I1220 F1221 Q1222 I1223 F1224 Y1225 F1226 R1227 M1228 Y1229 L1230 A1231 M1232 V1233 L1234 L1235 G1236 A1237 T1238 H1239 F1243 I1251 GLY PRO SER VAL ASN LYS ALA LYS SER CYS ALA THR GLU GLU ARG TYR LYS GLY THR
GLU ARG GLU ARG LEU LEU ASN PHE

4 Experimental information (i)

Property	Value	Source
Reconstruction method	Not provided	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of images used	Not provided	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	Each micrograph	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	50	Depositor
Minimum defocus (nm)	1500	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	22500	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: BMA, NAG, CLR, 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	#$\|Z\|>2$	RMSZ	#$\|Z\|>2$
1	A	0.52	8/7843 (0.1%)	0.67	31/10748 (0.3%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	2

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	940	SER	C-N	13.96	1.66	1.34
1	A	429	VAL	C-N	-5.49	1.23	1.34
1	A	166	PRO	N-CD	5.27	1.55	1.47
1	A	469	LEU	C-N	-5.22	1.22	1.34
1	A	424	PRO	N-CD	5.12	1.55	1.47

The worst 5 of 31 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	602	SER	CB-CA-C	-9.99	91.11	110.10
1	A	755	PRO	CA-N-CD	-8.22	99.99	111.50
1	A	996	MET	C-N-CA	-6.63	105.11	121.70
1	A	887	PRO	N-CA-CB	6.57	111.18	103.30
1	A	836	PRO	N-CA-CB	6.43	111.02	103.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	602	SER	Peptide
1	A	603	PHE	Peptide

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	7695	0	6476	1101	0
2	A	112	0	100	12	0
3	A	126	0	117	13	0
4	A	39	0	34	20	0
5	A	50	0	43	10	0
6	A	28	0	46	1	0
All	All	8050	0	6816	1126	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 76.

The worst 5 of 1126 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:598:ASN:HD21	2:A:1310:NAG:C1	0.99	1.62
1:A:724:LEU:HD22	1:A:1170:HIS:CE1	1.29	1.62
1:A:656:LEU:CD1	1:A:685:ILE:HG13	1.17	1.56
1:A:524:ASN:ND2	2:A:1304:NAG:C1	1.68	1.54
1:A:693:LEU:HD11	1:A:763:PHE:CE2	1.41	1.52

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 PDB entries followed by that with respect to all EM entries.

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	1126/1278 (88%)	898 (80%)	134 (12%)	94 (8%)	1 18

5 of 94 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	250	LYS
1	A	256	PRO
1	A	350	PRO
1	A	372	ARG
1	A	374	THR

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 PDB entries followed by that with respect to all EM entries.

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	643/1109 (58%)	558 (87%)	85 (13%)	5 30

5 of 85 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	719	THR
1	A	763	PHE
1	A	1189	GLU
1	A	721	ASP
1	A	751	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	452	ASN
1	A	459	ASN
1	A	722	GLN
1	A	241	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	447	GLN

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

15 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	1301	1,2	14,14,15	0.62	0	15,19,21	0.87	0
2	NAG	A	1302	2	14,14,15	0.56	0	15,19,21	0.73	1 (6%)
2	NAG	A	1304	2	14,14,15	0.46	0	15,19,21	0.98	2 (13%)
2	NAG	A	1305	2	14,14,15	0.28	0	15,19,21	0.54	0
4	NAG	A	1306	4	14,14,15	0.29	0	15,19,21	0.52	0
4	NAG	A	1307	4	14,14,15	0.29	0	15,19,21	0.53	0
4	BMA	A	1308	4	11,11,12	0.26	0	15,15,17	0.57	0
2	NAG	A	1310	2	14,14,15	0.33	0	15,19,21	1.47	2 (13%)
2	NAG	A	1311	2	14,14,15	0.64	0	15,19,21	1.11	2 (13%)
5	NAG	A	1312	5	14,14,15	0.27	0	15,19,21	0.53	0
5	NAG	A	1313	5	14,14,15	0.26	0	15,19,21	0.84	0
5	BMA	A	1314	5	11,11,12	0.28	0	15,15,17	0.63	0
5	MAN	A	1315	5	11,11,12	0.36	0	15,15,17	1.21	1 (6%)
2	NAG	A	1318	-	14,14,15	0.39	0	15,19,21	1.15	2 (13%)
2	NAG	A	1319	-	14,14,15	0.37	0	15,19,21	1.16	2 (13%)

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	1301	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	1302	2	-	0/6/23/26	0/1/1/1
2	NAG	A	1304	2	-	0/6/23/26	0/1/1/1
2	NAG	A	1305	2	-	0/6/23/26	0/1/1/1
4	NAG	A	1306	4	-	2/6/23/26	0/1/1/1
4	NAG	A	1307	4	-	0/6/23/26	0/1/1/1
4	BMA	A	1308	4	-	0/2/19/22	0/1/1/1
2	NAG	A	1310	2	-	0/6/23/26	0/1/1/1
2	NAG	A	1311	2	-	0/6/23/26	0/1/1/1
5	NAG	A	1312	5	-	0/6/23/26	0/1/1/1
5	NAG	A	1313	5	-	0/6/23/26	0/1/1/1
5	BMA	A	1314	5	-	0/2/19/22	0/1/1/1
5	MAN	A	1315	5	-	0/2/19/22	0/1/1/1
2	NAG	A	1318	-	-	0/6/23/26	0/1/1/1
2	NAG	A	1319	-	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	1310	NAG	O4-C4-C3	-3.38	102.75	110.36
2	A	1318	NAG	C2-N2-C7	-2.42	119.95	123.11
2	A	1319	NAG	C2-N2-C7	-2.41	119.97	123.11
2	A	1311	NAG	O5-C5-C4	-2.01	106.81	110.13
2	A	1318	NAG	C8-C7-N2	2.02	119.97	116.10

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	1306	NAG	C8-C7-N2-C2
4	A	1306	NAG	O7-C7-N2-C2

There are no ring outliers.

10 monomers are involved in 42 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1304	NAG	2	0
2	A	1305	NAG	2	0
4	A	1306	NAG	16	0
4	A	1307	NAG	14	0
4	A	1308	BMA	4	0
2	A	1310	NAG	4	0
5	A	1312	NAG	8	0
5	A	1313	NAG	6	0
5	A	1314	BMA	2	0
2	A	1319	NAG	4	0

5.6 Ligand geometry (i)

10 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	$\#\|Z\| > 2$
3	NAG	A	1303	-	14,14,15	0.30	0	15,19,21	0.53	0
3	NAG	A	1309	-	14,14,15	0.48	0	15,19,21	0.98	0
3	NAG	A	1316	1	14,14,15	0.40	0	15,19,21	1.16	2 (13%)
3	NAG	A	1317	-	14,14,15	0.28	0	15,19,21	0.54	0
3	NAG	A	1320	1	14,14,15	0.29	0	15,19,21	0.53	0
3	NAG	A	1321	-	14,14,15	0.28	0	15,19,21	0.54	0
3	NAG	A	1322	-	14,14,15	0.28	0	15,19,21	0.53	0
3	NAG	A	1323	-	14,14,15	0.28	0	15,19,21	0.54	0
3	NAG	A	1324	-	14,14,15	0.41	0	15,19,21	1.17	2 (13%)
6	CLR	A	1325	-	31,31,31	0.55	0	48,48,48	1.45	8 (16%)

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	1303	-	-	0/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	A	1309	-	-	0/6/23/26	0/1/1/1
3	NAG	A	1316	1	-	0/6/23/26	0/1/1/1
3	NAG	A	1317	-	-	0/6/23/26	0/1/1/1
3	NAG	A	1320	1	-	0/6/23/26	0/1/1/1
3	NAG	A	1321	-	-	0/6/23/26	0/1/1/1
3	NAG	A	1322	-	-	0/6/23/26	0/1/1/1
3	NAG	A	1323	-	-	0/6/23/26	0/1/1/1
3	NAG	A	1324	-	-	0/6/23/26	0/1/1/1
6	CLR	A	1325	-	-	0/10/68/68	0/4/4/4

There are no bond length outliers.

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	A	1325	CLR	C1-C2-C3	-3.46	105.81	110.41
6	A	1325	CLR	C8-C7-C6	-2.83	108.39	112.76
6	A	1325	CLR	C4-C5-C6	-2.52	116.44	120.60
3	A	1324	NAG	C2-N2-C7	-2.44	119.93	123.11
3	A	1316	NAG	C2-N2-C7	-2.41	119.97	123.11

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1303	NAG	2	0
3	A	1317	NAG	2	0
3	A	1322	NAG	4	0
3	A	1323	NAG	2	0
3	A	1324	NAG	3	0
6	A	1325	CLR	1	0

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.