W O R L D W I D E

wwPDB

PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 15, 2023 – 07:00 PM JST

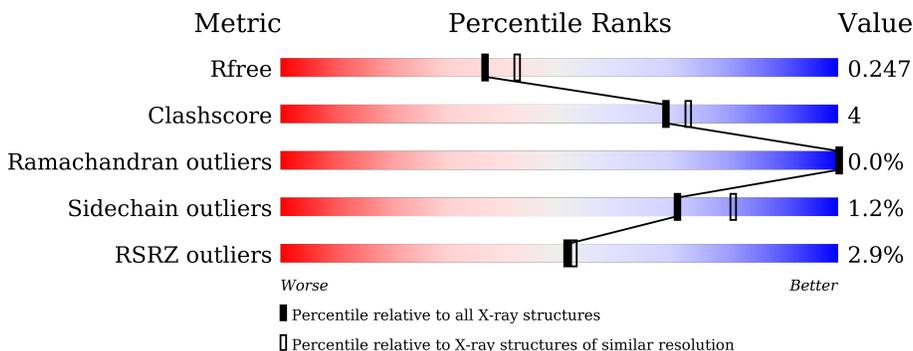|		
PDB ID	:	6JTJ
Title	:	Crystal structure of NagZ from Neisseria gonorrhoeae in complex with N-acetylglucosamine
Authors	:	Chen, Y.
Deposited on	:	2019-04-11
Resolution	:	2.18 Å(reported)

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 2.18 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		0.247
Clashscore		4
Ramachandran outliers		0.0%
Sidechain outliers		1.2%
RSRZ outliers		2.9%

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}	130704	6864 (2.20-2.16)
Clashscore	141614	7689 (2.20-2.16)
Ramachandran outliers	138981	7564 (2.20-2.16)
Sidechain outliers	138945	7564 (2.20-2.16)
RSRZ outliers	127900	6738 (2.20-2.16)

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 of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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% 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	397	2% / 77% / 7% / 16%
1	B	397	2% / 80% / 8% / 13%
1	C	397	4% / 79% / 5% · 15%
1	D	397	2% / 81% / 5% · 13%
1	E	397	4% / 81% / 5% / 15%
1	F	397	2% / 80% / 5% / 14%

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
2	NAG	A	401	-	-	X	-

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 16730 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 Beta-hexosaminidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	334	Total	C	N	O	S	0	0	0
			2560	1617	457	472	14			
1	B	346	Total	C	N	O	S	0	0	0
			2633	1659	471	487	16			
1	C	339	Total	C	N	O	S	0	0	0
			2593	1637	464	478	14			
1	D	344	Total	C	N	O	S	0	0	0
			2625	1656	469	484	16			
1	E	339	Total	C	N	O	S	0	0	0
			2596	1639	464	478	15			
1	F	340	Total	C	N	O	S	0	0	0
			2601	1642	465	479	15			

There are 216 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-35	MET	-	expression tag	UNP Q5FA94
A	-34	GLY	-	expression tag	UNP Q5FA94
A	-33	SER	-	expression tag	UNP Q5FA94
A	-32	SER	-	expression tag	UNP Q5FA94
A	-31	HIS	-	expression tag	UNP Q5FA94
A	-30	HIS	-	expression tag	UNP Q5FA94
A	-29	HIS	-	expression tag	UNP Q5FA94
A	-28	HIS	-	expression tag	UNP Q5FA94
A	-27	HIS	-	expression tag	UNP Q5FA94
A	-26	HIS	-	expression tag	UNP Q5FA94
A	-25	SER	-	expression tag	UNP Q5FA94
A	-24	SER	-	expression tag	UNP Q5FA94
A	-23	GLY	-	expression tag	UNP Q5FA94
A	-22	LEU	-	expression tag	UNP Q5FA94
A	-21	VAL	-	expression tag	UNP Q5FA94
A	-20	PRO	-	expression tag	UNP Q5FA94
A	-19	ARG	-	expression tag	UNP Q5FA94

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-18	GLY	-	expression tag	UNP Q5FA94
A	-17	SER	-	expression tag	UNP Q5FA94
A	-16	HIS	-	expression tag	UNP Q5FA94
A	-15	MET	-	expression tag	UNP Q5FA94
A	-14	ALA	-	expression tag	UNP Q5FA94
A	-13	SER	-	expression tag	UNP Q5FA94
A	-12	MET	-	expression tag	UNP Q5FA94
A	-11	THR	-	expression tag	UNP Q5FA94
A	-10	GLY	-	expression tag	UNP Q5FA94
A	-9	GLY	-	expression tag	UNP Q5FA94
A	-8	GLN	-	expression tag	UNP Q5FA94
A	-7	GLN	-	expression tag	UNP Q5FA94
A	-6	MET	-	expression tag	UNP Q5FA94
A	-5	GLY	-	expression tag	UNP Q5FA94
A	-4	ARG	-	expression tag	UNP Q5FA94
A	-3	GLY	-	expression tag	UNP Q5FA94
A	-2	SER	-	expression tag	UNP Q5FA94
A	-1	GLU	-	expression tag	UNP Q5FA94
A	0	PHE	-	expression tag	UNP Q5FA94
B	-35	MET	-	expression tag	UNP Q5FA94
B	-34	GLY	-	expression tag	UNP Q5FA94
B	-33	SER	-	expression tag	UNP Q5FA94
B	-32	SER	-	expression tag	UNP Q5FA94
B	-31	HIS	-	expression tag	UNP Q5FA94
B	-30	HIS	-	expression tag	UNP Q5FA94
B	-29	HIS	-	expression tag	UNP Q5FA94
B	-28	HIS	-	expression tag	UNP Q5FA94
B	-27	HIS	-	expression tag	UNP Q5FA94
B	-26	HIS	-	expression tag	UNP Q5FA94
B	-25	SER	-	expression tag	UNP Q5FA94
B	-24	SER	-	expression tag	UNP Q5FA94
B	-23	GLY	-	expression tag	UNP Q5FA94
B	-22	LEU	-	expression tag	UNP Q5FA94
B	-21	VAL	-	expression tag	UNP Q5FA94
B	-20	PRO	-	expression tag	UNP Q5FA94
B	-19	ARG	-	expression tag	UNP Q5FA94
B	-18	GLY	-	expression tag	UNP Q5FA94
B	-17	SER	-	expression tag	UNP Q5FA94
B	-16	HIS	-	expression tag	UNP Q5FA94
B	-15	MET	-	expression tag	UNP Q5FA94
B	-14	ALA	-	expression tag	UNP Q5FA94
B	-13	SER	-	expression tag	UNP Q5FA94

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-12	MET	-	expression tag	UNP Q5FA94
B	-11	THR	-	expression tag	UNP Q5FA94
B	-10	GLY	-	expression tag	UNP Q5FA94
B	-9	GLY	-	expression tag	UNP Q5FA94
B	-8	GLN	-	expression tag	UNP Q5FA94
B	-7	GLN	-	expression tag	UNP Q5FA94
B	-6	MET	-	expression tag	UNP Q5FA94
B	-5	GLY	-	expression tag	UNP Q5FA94
B	-4	ARG	-	expression tag	UNP Q5FA94
B	-3	GLY	-	expression tag	UNP Q5FA94
B	-2	SER	-	expression tag	UNP Q5FA94
B	-1	GLU	-	expression tag	UNP Q5FA94
B	0	PHE	-	expression tag	UNP Q5FA94
C	-35	MET	-	expression tag	UNP Q5FA94
C	-34	GLY	-	expression tag	UNP Q5FA94
C	-33	SER	-	expression tag	UNP Q5FA94
C	-32	SER	-	expression tag	UNP Q5FA94
C	-31	HIS	-	expression tag	UNP Q5FA94
C	-30	HIS	-	expression tag	UNP Q5FA94
C	-29	HIS	-	expression tag	UNP Q5FA94
C	-28	HIS	-	expression tag	UNP Q5FA94
C	-27	HIS	-	expression tag	UNP Q5FA94
C	-26	HIS	-	expression tag	UNP Q5FA94
C	-25	SER	-	expression tag	UNP Q5FA94
C	-24	SER	-	expression tag	UNP Q5FA94
C	-23	GLY	-	expression tag	UNP Q5FA94
C	-22	LEU	-	expression tag	UNP Q5FA94
C	-21	VAL	-	expression tag	UNP Q5FA94
C	-20	PRO	-	expression tag	UNP Q5FA94
C	-19	ARG	-	expression tag	UNP Q5FA94
C	-18	GLY	-	expression tag	UNP Q5FA94
C	-17	SER	-	expression tag	UNP Q5FA94
C	-16	HIS	-	expression tag	UNP Q5FA94
C	-15	MET	-	expression tag	UNP Q5FA94
C	-14	ALA	-	expression tag	UNP Q5FA94
C	-13	SER	-	expression tag	UNP Q5FA94
C	-12	MET	-	expression tag	UNP Q5FA94
C	-11	THR	-	expression tag	UNP Q5FA94
C	-10	GLY	-	expression tag	UNP Q5FA94
C	-9	GLY	-	expression tag	UNP Q5FA94
C	-8	GLN	-	expression tag	UNP Q5FA94
C	-7	GLN	-	expression tag	UNP Q5FA94

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-6	MET	-	expression tag	UNP Q5FA94
C	-5	GLY	-	expression tag	UNP Q5FA94
C	-4	ARG	-	expression tag	UNP Q5FA94
C	-3	GLY	-	expression tag	UNP Q5FA94
C	-2	SER	-	expression tag	UNP Q5FA94
C	-1	GLU	-	expression tag	UNP Q5FA94
C	0	PHE	-	expression tag	UNP Q5FA94
D	-35	MET	-	expression tag	UNP Q5FA94
D	-34	GLY	-	expression tag	UNP Q5FA94
D	-33	SER	-	expression tag	UNP Q5FA94
D	-32	SER	-	expression tag	UNP Q5FA94
D	-31	HIS	-	expression tag	UNP Q5FA94
D	-30	HIS	-	expression tag	UNP Q5FA94
D	-29	HIS	-	expression tag	UNP Q5FA94
D	-28	HIS	-	expression tag	UNP Q5FA94
D	-27	HIS	-	expression tag	UNP Q5FA94
D	-26	HIS	-	expression tag	UNP Q5FA94
D	-25	SER	-	expression tag	UNP Q5FA94
D	-24	SER	-	expression tag	UNP Q5FA94
D	-23	GLY	-	expression tag	UNP Q5FA94
D	-22	LEU	-	expression tag	UNP Q5FA94
D	-21	VAL	-	expression tag	UNP Q5FA94
D	-20	PRO	-	expression tag	UNP Q5FA94
D	-19	ARG	-	expression tag	UNP Q5FA94
D	-18	GLY	-	expression tag	UNP Q5FA94
D	-17	SER	-	expression tag	UNP Q5FA94
D	-16	HIS	-	expression tag	UNP Q5FA94
D	-15	MET	-	expression tag	UNP Q5FA94
D	-14	ALA	-	expression tag	UNP Q5FA94
D	-13	SER	-	expression tag	UNP Q5FA94
D	-12	MET	-	expression tag	UNP Q5FA94
D	-11	THR	-	expression tag	UNP Q5FA94
D	-10	GLY	-	expression tag	UNP Q5FA94
D	-9	GLY	-	expression tag	UNP Q5FA94
D	-8	GLN	-	expression tag	UNP Q5FA94
D	-7	GLN	-	expression tag	UNP Q5FA94
D	-6	MET	-	expression tag	UNP Q5FA94
D	-5	GLY	-	expression tag	UNP Q5FA94
D	-4	ARG	-	expression tag	UNP Q5FA94
D	-3	GLY	-	expression tag	UNP Q5FA94
D	-2	SER	-	expression tag	UNP Q5FA94
D	-1	GLU	-	expression tag	UNP Q5FA94

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	0	PHE	-	expression tag	UNP Q5FA94
E	-35	MET	-	expression tag	UNP Q5FA94
E	-34	GLY	-	expression tag	UNP Q5FA94
E	-33	SER	-	expression tag	UNP Q5FA94
E	-32	SER	-	expression tag	UNP Q5FA94
E	-31	HIS	-	expression tag	UNP Q5FA94
E	-30	HIS	-	expression tag	UNP Q5FA94
E	-29	HIS	-	expression tag	UNP Q5FA94
E	-28	HIS	-	expression tag	UNP Q5FA94
E	-27	HIS	-	expression tag	UNP Q5FA94
E	-26	HIS	-	expression tag	UNP Q5FA94
E	-25	SER	-	expression tag	UNP Q5FA94
E	-24	SER	-	expression tag	UNP Q5FA94
E	-23	GLY	-	expression tag	UNP Q5FA94
E	-22	LEU	-	expression tag	UNP Q5FA94
E	-21	VAL	-	expression tag	UNP Q5FA94
E	-20	PRO	-	expression tag	UNP Q5FA94
E	-19	ARG	-	expression tag	UNP Q5FA94
E	-18	GLY	-	expression tag	UNP Q5FA94
E	-17	SER	-	expression tag	UNP Q5FA94
E	-16	HIS	-	expression tag	UNP Q5FA94
E	-15	MET	-	expression tag	UNP Q5FA94
E	-14	ALA	-	expression tag	UNP Q5FA94
E	-13	SER	-	expression tag	UNP Q5FA94
E	-12	MET	-	expression tag	UNP Q5FA94
E	-11	THR	-	expression tag	UNP Q5FA94
E	-10	GLY	-	expression tag	UNP Q5FA94
E	-9	GLY	-	expression tag	UNP Q5FA94
E	-8	GLN	-	expression tag	UNP Q5FA94
E	-7	GLN	-	expression tag	UNP Q5FA94
E	-6	MET	-	expression tag	UNP Q5FA94
E	-5	GLY	-	expression tag	UNP Q5FA94
E	-4	ARG	-	expression tag	UNP Q5FA94
E	-3	GLY	-	expression tag	UNP Q5FA94
E	-2	SER	-	expression tag	UNP Q5FA94
E	-1	GLU	-	expression tag	UNP Q5FA94
E	0	PHE	-	expression tag	UNP Q5FA94
F	-35	MET	-	expression tag	UNP Q5FA94
F	-34	GLY	-	expression tag	UNP Q5FA94
F	-33	SER	-	expression tag	UNP Q5FA94
F	-32	SER	-	expression tag	UNP Q5FA94
F	-31	HIS	-	expression tag	UNP Q5FA94

Continued on next page...

Continued from previous page...

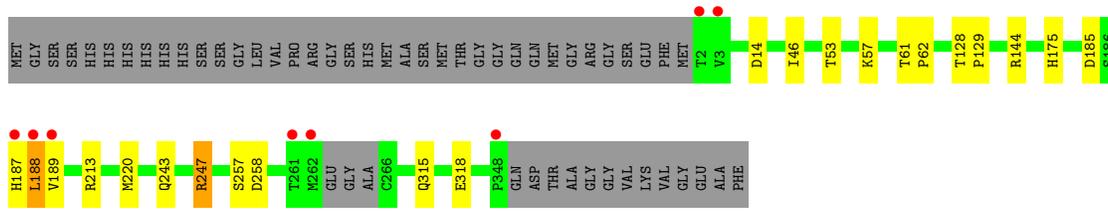| Chain	Residue	Modelled	Actual	Comment	Reference
F	-30	HIS	-	expression tag	UNP Q5FA94
F	-29	HIS	-	expression tag	UNP Q5FA94
F	-28	HIS	-	expression tag	UNP Q5FA94
F	-27	HIS	-	expression tag	UNP Q5FA94
F	-26	HIS	-	expression tag	UNP Q5FA94
F	-25	SER	-	expression tag	UNP Q5FA94
F	-24	SER	-	expression tag	UNP Q5FA94
F	-23	GLY	-	expression tag	UNP Q5FA94
F	-22	LEU	-	expression tag	UNP Q5FA94
F	-21	VAL	-	expression tag	UNP Q5FA94
F	-20	PRO	-	expression tag	UNP Q5FA94
F	-19	ARG	-	expression tag	UNP Q5FA94
F	-18	GLY	-	expression tag	UNP Q5FA94
F	-17	SER	-	expression tag	UNP Q5FA94
F	-16	HIS	-	expression tag	UNP Q5FA94
F	-15	MET	-	expression tag	UNP Q5FA94
F	-14	ALA	-	expression tag	UNP Q5FA94
F	-13	SER	-	expression tag	UNP Q5FA94
F	-12	MET	-	expression tag	UNP Q5FA94
F	-11	THR	-	expression tag	UNP Q5FA94
F	-10	GLY	-	expression tag	UNP Q5FA94
F	-9	GLY	-	expression tag	UNP Q5FA94
F	-8	GLN	-	expression tag	UNP Q5FA94
F	-7	GLN	-	expression tag	UNP Q5FA94
F	-6	MET	-	expression tag	UNP Q5FA94
F	-5	GLY	-	expression tag	UNP Q5FA94
F	-4	ARG	-	expression tag	UNP Q5FA94
F	-3	GLY	-	expression tag	UNP Q5FA94
F	-2	SER	-	expression tag	UNP Q5FA94
F	-1	GLU	-	expression tag	UNP Q5FA94
F	0	PHE	-	expression tag	UNP Q5FA94

- Molecule 2 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$) (labeled as "Ligand of Interest" by depositor).
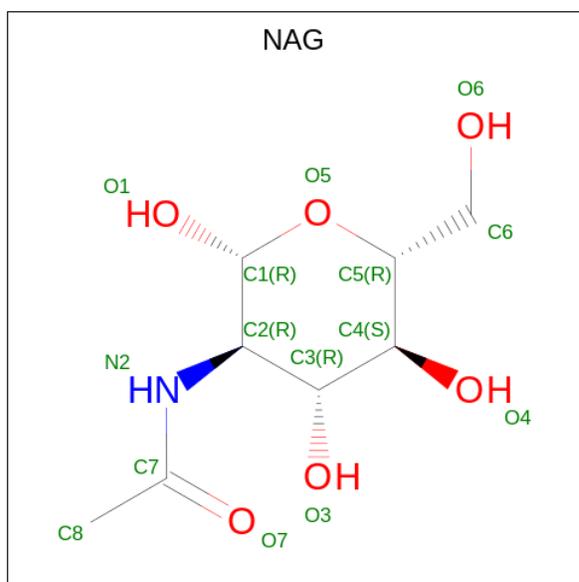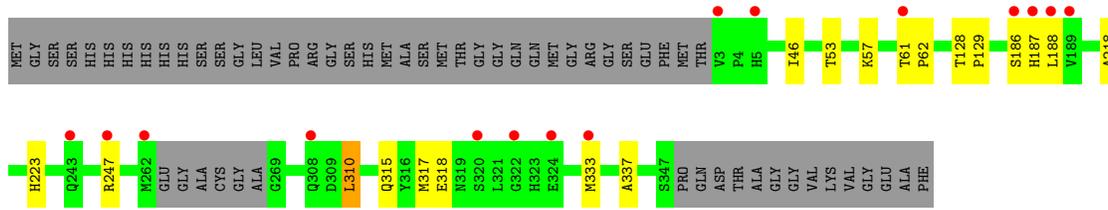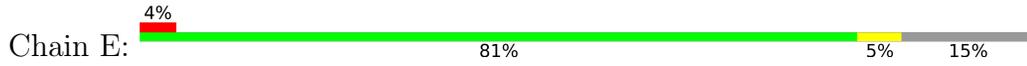
NAG

O6
OH

O5
O

O1
HO

C6

C1(R) C5(R)

C2(R) C4(S)
C3(R)

N2
HN

OH
O4

C7

OH

C8

O
O7

O3

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			15	8	1	6		
2	B	1	Total	C	N	O	0	0
			15	8	1	6		
2	C	1	Total	C	N	O	0	0
			15	8	1	6		
2	D	1	Total	C	N	O	0	0
			15	8	1	6		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	183	Total	O	0	0
			183	183		
3	B	267	Total	O	0	0
			267	267		
3	C	153	Total	O	0	0
			153	153		
3	D	172	Total	O	0	0
			172	172		
3	E	98	Total	O	0	0
			98	98		
3	F	189	Total	O	0	0
			189	189		

WORLDWIDE
PDB
PROTEIN DATA BANK

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes 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: Beta-hexosaminidase

Chain A:

- Molecule 1: Beta-hexosaminidase

Chain B:

- Molecule 1: Beta-hexosaminidase

Chain C:

- Molecule 1: Beta-hexosaminidase

Chain D:

MET GLY SER SER HIS HIS HIS HIS HIS HIS SER SER GLY LEU VAL PRO ARG GLY SER HIS MET ALA SER MET THR GLY GLY GLN GLN MET GLY ARG GLY SER GLU PHE MET **T2** V3 D14 I46 T53 K57 T61 P62 T128 P129 R144 H175 D185 S186

H187 L188 V189 R213 M220 Q243 R247 S257 D258 T261 M262 GLU GLY ALA C266 Q315 E318 P348 GLN ASP THR ALA GLY GLY VAL LYS VAL GLY GLU ALA PHE

- Molecule 1: Beta-hexosaminidase

Chain E:

MET GLY SER SER HIS HIS HIS HIS HIS HIS SER SER GLY LEU VAL PRO ARG GLY SER HIS MET ALA SER MET THR GLY GLY GLN GLN MET GLY ARG GLY SER GLU PHE MET THR V3 P4 H5 I46 T53 K57 T61 P62 T128 P129 S186 H187 L188 V189 A218

H223 Q243 R247 M262 GLU GLY ALA CYS GLY ALA G269 Q308 D309 L310 Q315 Y316 M317 E318 N319 S320 L321 G322 H323 E324 M333 A337 S347 PRO GLN ASP THR ALA GLY GLY VAL LYS VAL GLY GLU ALA PHE

- Molecule 1: Beta-hexosaminidase

Chain F:

MET GLY SER SER HIS HIS HIS HIS HIS HIS SER SER GLY LEU VAL PRO ARG GLY SER HIS MET ALA SER MET THR GLY GLY GLN GLN MET GLY ARG GLY SER GLU PHE MET THR V3 M12 I37 Q44 N45 I46 T61 P62 E80 T128 P129 D185 S186 H187 L188

V189 A218 H223 R247 D258 T261 M262 GLU GLY ALA CYS GLY A268 G269 T303 L310 Q315 Y316 M317 E318 S347 PRO GLN ASP THR ALA GLY GLY VAL LYS VAL GLY GLU ALA PHE

4   Data and refinement statistics   ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	102.18Å 125.31Å 190.61Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	104.71 – 2.18 49.56 – 2.18	Depositor EDS
% Data completeness (in resolution range)	99.7 (104.71-2.18) 99.7 (49.56-2.18)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	2.47 (at 2.18Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.214 , 0.244 0.219 , 0.247	Depositor DCC
R_{free} test set	6401 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	29.0	Xtriage
Anisotropy	0.024	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 38.0	EDS
L-test for twinning[2]	$<|L|> = 0.49, <L^2> = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	16730	wwPDB-VP
Average B, all atoms (Å^2)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.76% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $<|L|>$, $<L^2>$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: 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	0.50	0/2607	0.66	1/3523 (0.0%)
1	B	0.52	0/2684	0.68	1/3630 (0.0%)
1	C	0.49	0/2642	0.69	3/3572 (0.1%)
1	D	0.47	0/2675	0.66	0/3617
1	E	0.44	0/2645	0.66	2/3575 (0.1%)
1	F	0.51	0/2650	0.67	1/3582 (0.0%)
All	All	0.49	0/15903	0.67	8/21499 (0.0%)

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	D	0	1
1	F	0	1
All	All	0	2

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	330	MET	CG-SD-CE	7.83	112.73	100.20
1	C	188	LEU	CB-CA-C	-5.46	99.83	110.20
1	E	310	LEU	CA-CB-CG	5.45	127.84	115.30
1	F	310	LEU	CA-CB-CG	5.28	127.44	115.30
1	C	274	ARG	NE-CZ-NH2	5.25	122.92	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Mol	Chain	Res	Type	Group
1	D	188	LEU	Peptide
1	F	185	ASP	Peptide

5.2 Too-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 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	2560	0	2555	28	0
1	B	2633	0	2617	24	0
1	C	2593	0	2587	24	0
1	D	2625	0	2618	17	0
1	E	2596	0	2591	17	0
1	F	2601	0	2596	18	0
2	A	15	0	15	7	0
2	B	15	0	15	5	0
2	C	15	0	15	6	0
2	D	15	0	15	3	0
3	A	183	0	0	6	0
3	B	267	0	0	5	0
3	C	153	0	0	0	0
3	D	172	0	0	4	0
3	E	98	0	0	2	0
3	F	189	0	0	4	0
All	All	16730	0	15624	117	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 4.

The worst 5 of 117 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)
2:A:401:NAG:H83	1:C:185:ASP:O	1.27	1.27
1:A:218:ALA:HB2	1:A:317:MET:HE3	1.29	1.15
1:E:218:ALA:HB2	1:E:317:MET:HE3	1.27	1.14
1:F:218:ALA:HB2	1:F:317:MET:HE3	1.28	1.14
2:A:401:NAG:C8	1:C:185:ASP:O	2.07	1.02

There are no symmetry-related clashes.

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	328/397 (83%)	318 (97%)	10 (3%)	0	100	100
1	B	344/397 (87%)	329 (96%)	15 (4%)	0	100	100
1	C	335/397 (84%)	322 (96%)	12 (4%)	1 (0%)	41	43
1	D	340/397 (86%)	327 (96%)	13 (4%)	0	100	100
1	E	335/397 (84%)	325 (97%)	10 (3%)	0	100	100
1	F	336/397 (85%)	326 (97%)	10 (3%)	0	100	100
All	All	2018/2382 (85%)	1947 (96%)	70 (4%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	188	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	264/309 (85%)	262 (99%)	2 (1%)	81	89
1	B	270/309 (87%)	267 (99%)	3 (1%)	73	83
1	C	267/309 (86%)	264 (99%)	3 (1%)	73	83

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	271/309 (88%)	267 (98%)	4 (2%)	65	76
1	E	268/309 (87%)	266 (99%)	2 (1%)	84	91
1	F	268/309 (87%)	263 (98%)	5 (2%)	57	68
All	All	1608/1854 (87%)	1589 (99%)	19 (1%)	71	81

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	3	VAL
1	F	247	ARG
1	F	310	LEU
1	F	185	ASP
1	D	46	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	187	HIS
1	E	315	GLN
1	F	223	HIS
1	B	323	HIS
1	A	223	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 ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

4 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	$\#|Z| > 2$
2	NAG	D	401	-	15,15,15	1.10	0	21,21,21	3.11	12 (57%)
2	NAG	A	401	-	15,15,15	0.89	0	21,21,21	1.50	2 (9%)
2	NAG	B	401	-	15,15,15	0.75	0	21,21,21	1.76	5 (23%)
2	NAG	C	401	-	15,15,15	0.74	0	21,21,21	1.66	4 (19%)

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	D	401	-	-	4/6/26/26	0/1/1/1
2	NAG	A	401	-	-	2/6/26/26	0/1/1/1
2	NAG	B	401	-	-	2/6/26/26	0/1/1/1
2	NAG	C	401	-	-	2/6/26/26	0/1/1/1

There are no bond length outliers.

The worst 5 of 23 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	401	NAG	C1-C2-N2	-6.13	103.63	110.73
2	D	401	NAG	C8-C7-N2	5.11	124.75	116.10
2	D	401	NAG	O3-C3-C4	-4.81	99.22	110.35
2	B	401	NAG	O5-C1-C2	4.69	114.23	109.52
2	A	401	NAG	C3-C2-N2	4.44	119.00	110.62

There are no chirality outliers.
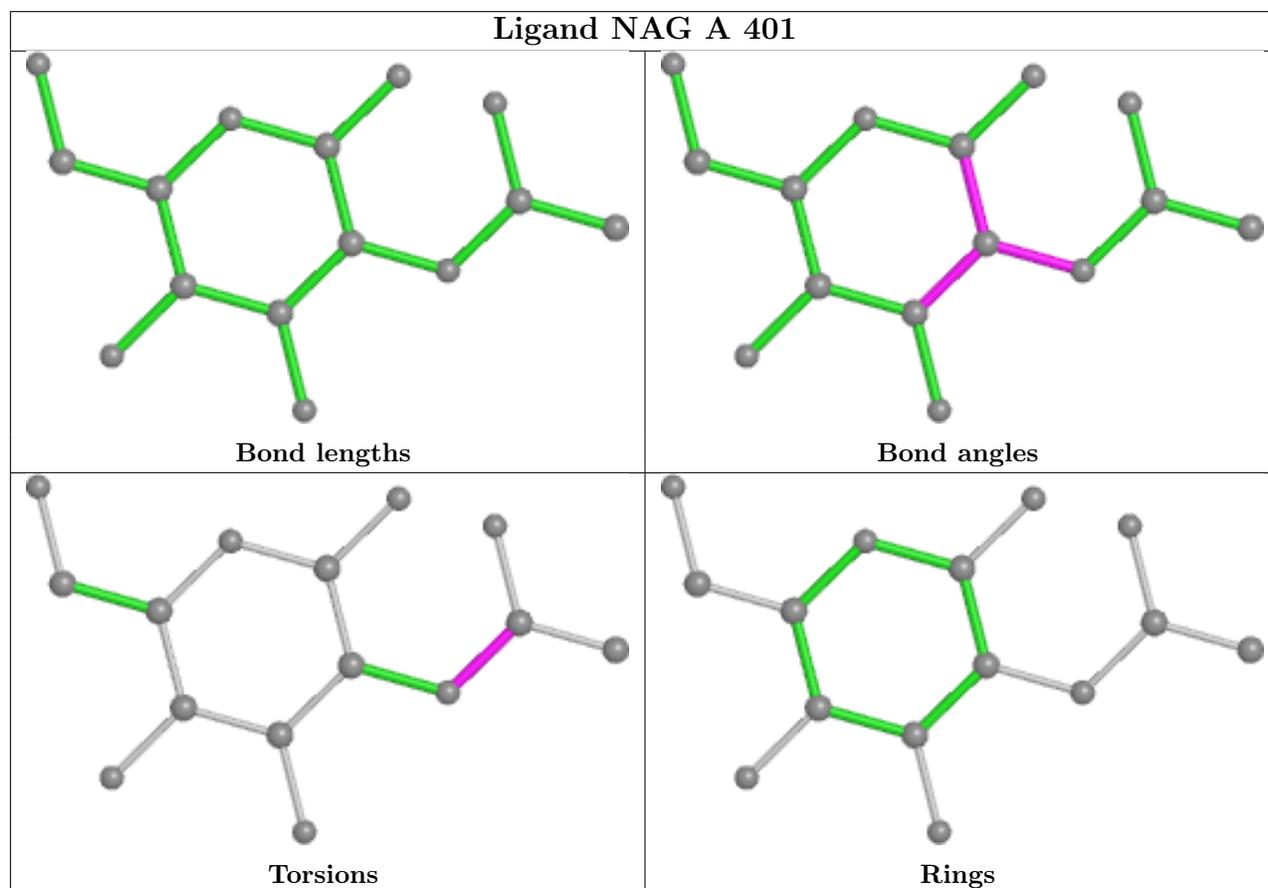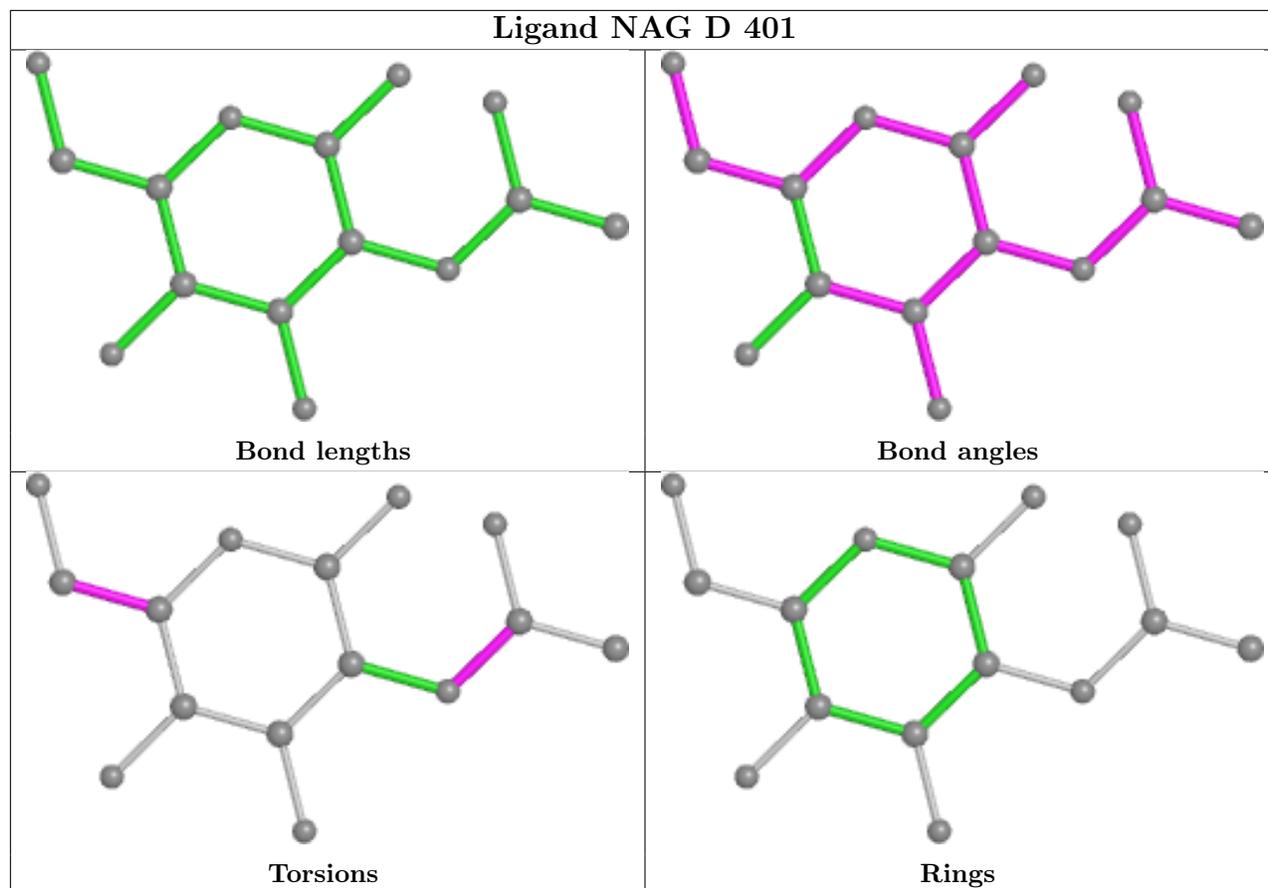
5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	D	401	NAG	C8-C7-N2-C2
2	D	401	NAG	O7-C7-N2-C2
2	D	401	NAG	O5-C5-C6-O6
2	A	401	NAG	C8-C7-N2-C2
2	A	401	NAG	O7-C7-N2-C2

There are no ring outliers.

4 monomers are involved in 21 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	401	NAG	3	0
2	A	401	NAG	7	0
2	B	401	NAG	5	0
2	C	401	NAG	6	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand NAG D 401

Bond lengths

Bond angles

Torsions

Rings

Ligand NAG A 401

Bond lengths

Bond angles

Torsions

Rings

Ligand NAG B 401

Bond lengths

Bond angles

Torsions

Rings

Ligand NAG C 401

Bond lengths

Bond angles

Torsions

Rings
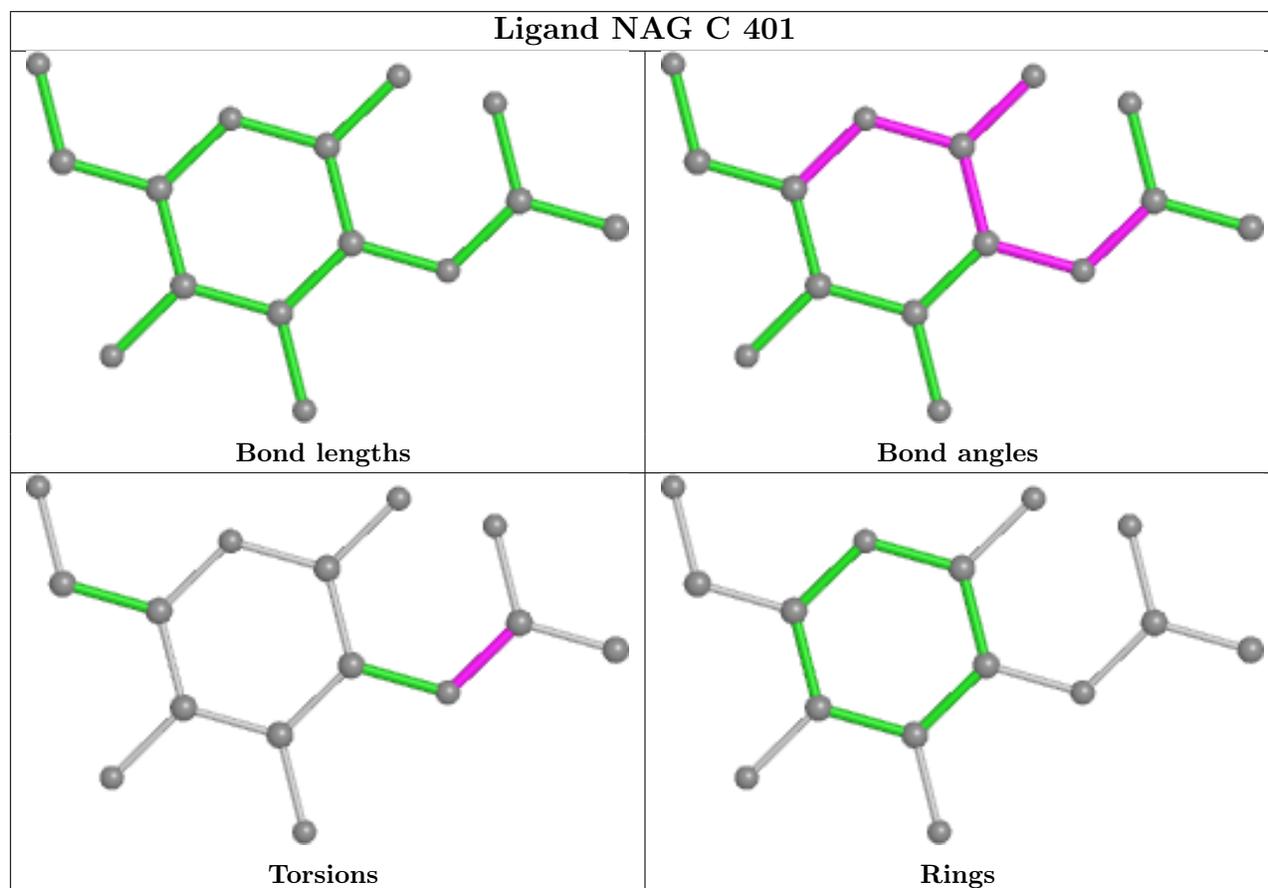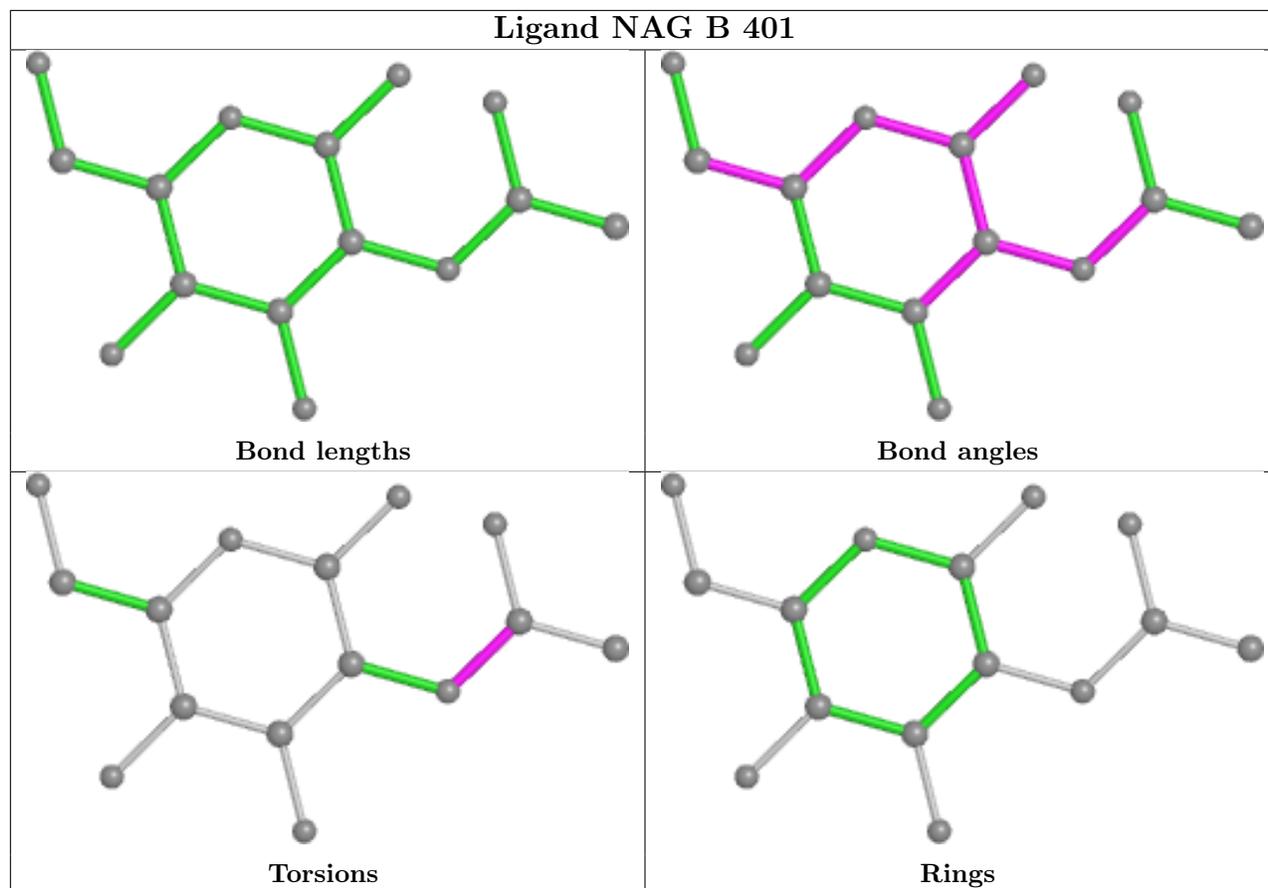
5.7 Other polymers ⓘ

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	334/397 (84%)	-0.02	7 (2%)	63	64	17, 32, 56, 81	0
1	B	346/397 (87%)	-0.21	7 (2%)	65	66	15, 24, 46, 106	0
1	C	339/397 (85%)	0.11	14 (4%)	37	38	16, 34, 61, 91	0
1	D	344/397 (86%)	-0.15	8 (2%)	60	61	22, 33, 53, 103	0
1	E	339/397 (85%)	0.36	15 (4%)	34	35	26, 44, 70, 130	0
1	F	340/397 (85%)	-0.04	9 (2%)	56	56	20, 32, 53, 108	0
All	All	2042/2382 (85%)	0.01	60 (2%)	51	52	15, 33, 60, 130	0

The worst 5 of 60 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	188	LEU	10.2
1	F	188	LEU	7.0
1	B	189	VAL	6.5
1	F	262	MET	6.2
1	A	189	VAL	6.0

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

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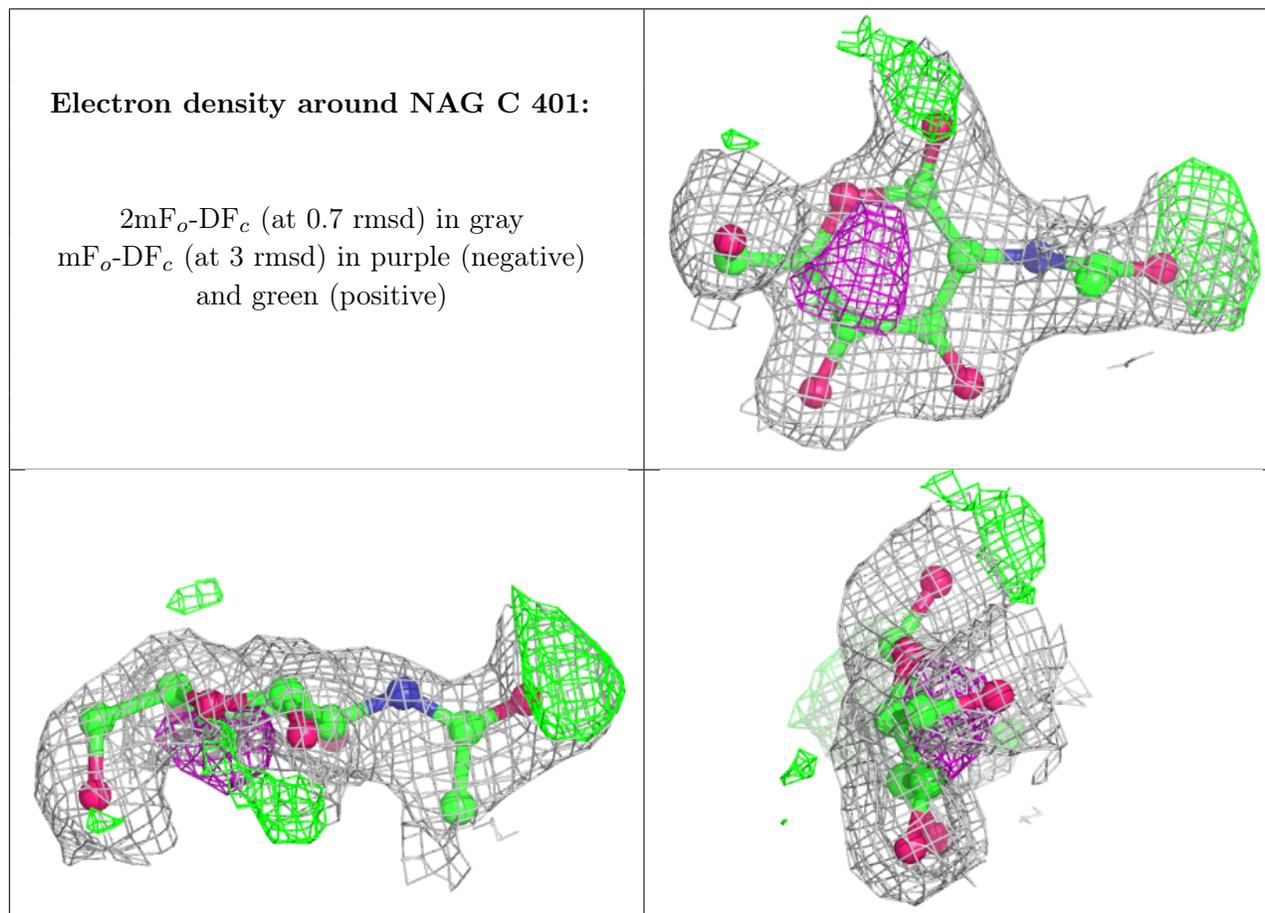
6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

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. 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	B-factors(Å^2)	Q<0.9
2	NAG	C	401	15/15	0.64	0.23	39,47,55,55	0
2	NAG	A	401	15/15	0.72	0.20	36,47,55,56	0
2	NAG	B	401	15/15	0.83	0.18	28,32,38,40	0
2	NAG	D	401	15/15	0.88	0.18	50,58,62,62	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.
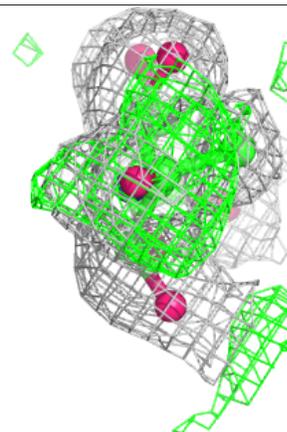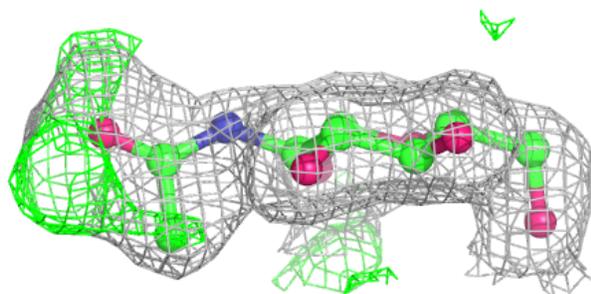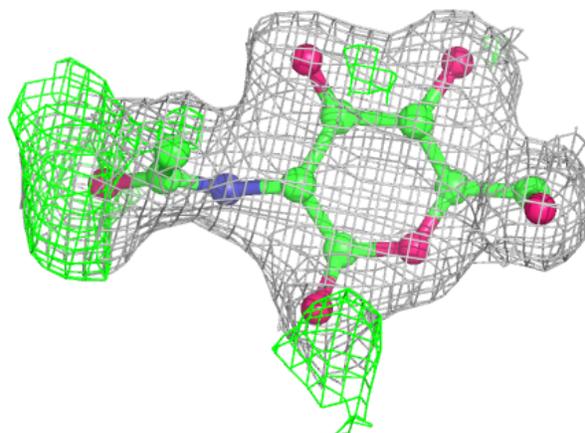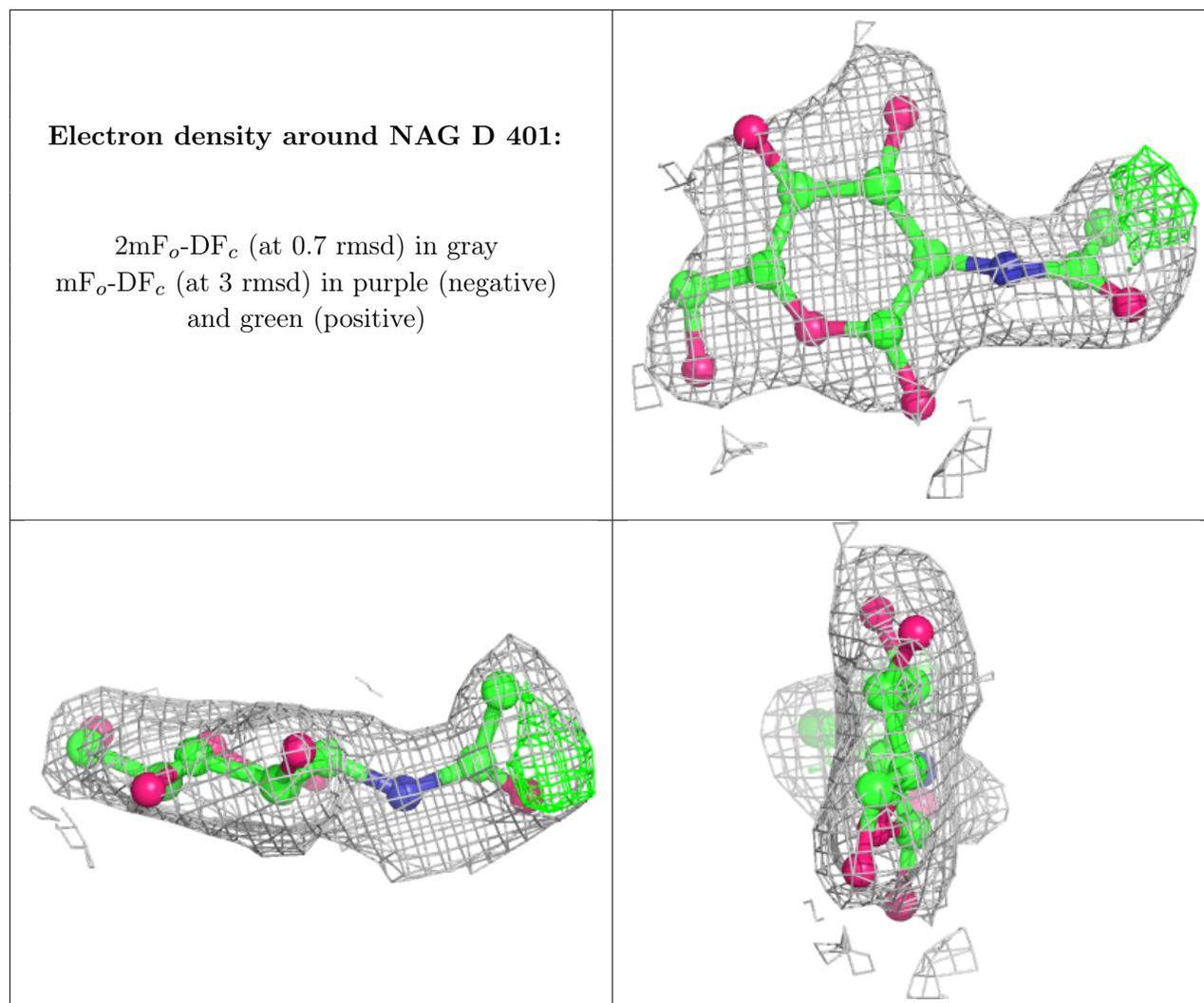

Electron density around NAG C 401:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

Electron density around NAG A 401:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
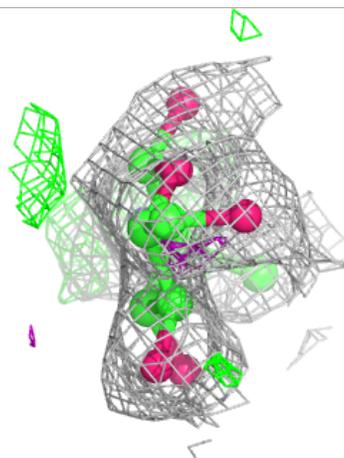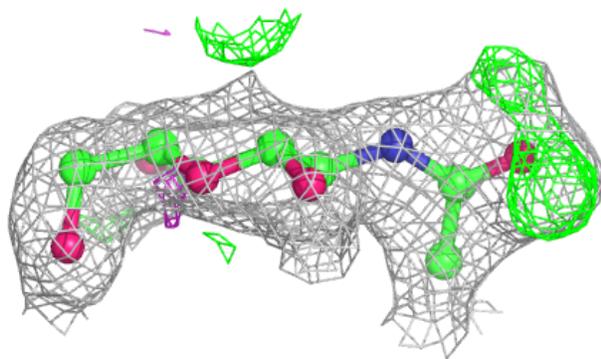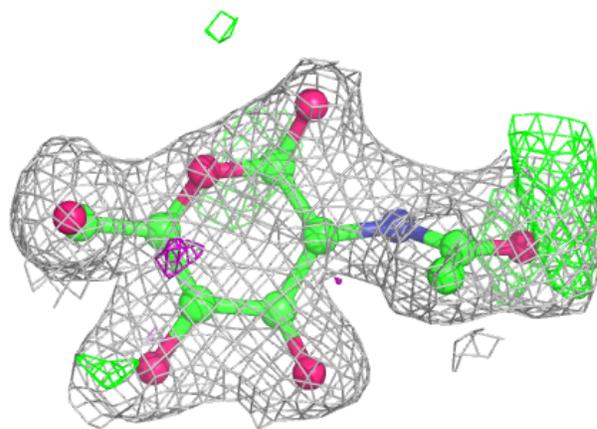
Electron density around NAG B 401:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)

Electron density around NAG D 401:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

6.5 Other polymers ⓘ

There are no such residues in this entry.