wwPDB X-ray Structure Validation Summary Report (i)

Jun 26, 2024 – 04:38 AM EDT

PDB ID	:	6Y5F
Title	:	Crystal structure of the envelope glycoprotein prefusion complex of Andes virus - Mutant H953F
Authors	:	Serris, A.; Rey, F.A.; Guardado-Calvo, P.
Deposited on	:	2020-02-25
Resolution	:	3.20 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

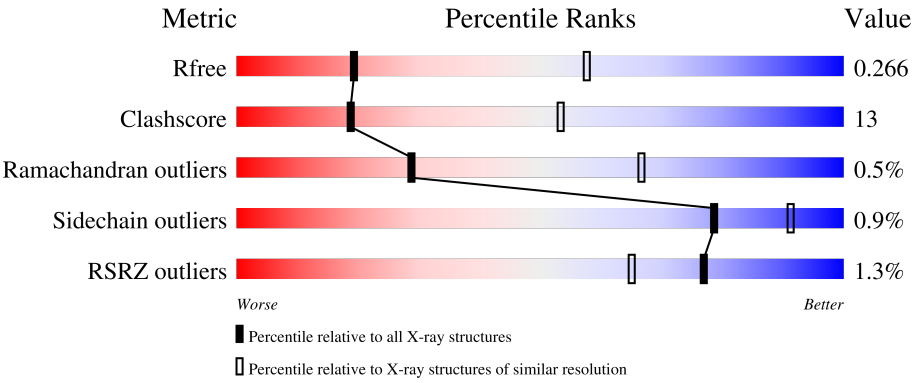
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.20 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}	130704	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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	893	29% 11% 60%
1	B	893	% 33% 13% 53%
2	C	4	50% 50%
2	D	4	25% 75%

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	C	1	-	-	X	-
2	NAG	D	1	X	-	-	-

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 6021 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 polyprotein,Envelope polyprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	353	Total	C	N	O	S	0	0	0
			2709	1709	447	532	21			
1	B	417	Total	C	N	O	S	0	0	0
			3187	2003	531	618	35			

There are 168 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	375	GLY	-	linker	UNP A0A3Q8TA45
A	376	GLY	-	linker	UNP A0A3Q8TA45
A	377	SER	-	linker	UNP A0A3Q8TA45
A	378	GLY	-	linker	UNP A0A3Q8TA45
A	379	LEU	-	linker	UNP A0A3Q8TA45
A	380	VAL	-	linker	UNP A0A3Q8TA45
A	381	PRO	-	linker	UNP A0A3Q8TA45
A	382	ARG	-	linker	UNP A0A3Q8TA45
A	383	GLY	-	linker	UNP A0A3Q8TA45
A	384	SER	-	linker	UNP A0A3Q8TA45
A	385	GLY	-	linker	UNP A0A3Q8TA45
A	386	GLY	-	linker	UNP A0A3Q8TA45
A	387	GLY	-	linker	UNP A0A3Q8TA45
A	388	SER	-	linker	UNP A0A3Q8TA45
A	389	GLY	-	linker	UNP A0A3Q8TA45
A	390	GLY	-	linker	UNP A0A3Q8TA45
A	391	GLY	-	linker	UNP A0A3Q8TA45
A	392	SER	-	linker	UNP A0A3Q8TA45
A	393	TRP	-	linker	UNP A0A3Q8TA45
A	394	SER	-	linker	UNP A0A3Q8TA45
A	395	HIS	-	linker	UNP A0A3Q8TA45
A	396	PRO	-	linker	UNP A0A3Q8TA45
A	397	GLN	-	linker	UNP A0A3Q8TA45
A	398	PHE	-	linker	UNP A0A3Q8TA45
A	399	GLU	-	linker	UNP A0A3Q8TA45

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	400	LYS	-	linker	UNP A0A3Q8TA45
A	401	GLY	-	linker	UNP A0A3Q8TA45
A	402	GLY	-	linker	UNP A0A3Q8TA45
A	403	GLY	-	linker	UNP A0A3Q8TA45
A	404	THR	-	linker	UNP A0A3Q8TA45
A	405	GLY	-	linker	UNP A0A3Q8TA45
A	406	GLY	-	linker	UNP A0A3Q8TA45
A	407	GLY	-	linker	UNP A0A3Q8TA45
A	408	THR	-	linker	UNP A0A3Q8TA45
A	409	LEU	-	linker	UNP A0A3Q8TA45
A	410	VAL	-	linker	UNP A0A3Q8TA45
A	411	PRO	-	linker	UNP A0A3Q8TA45
A	412	ARG	-	linker	UNP A0A3Q8TA45
A	413	GLY	-	linker	UNP A0A3Q8TA45
A	414	SER	-	linker	UNP A0A3Q8TA45
A	415	GLY	-	linker	UNP A0A3Q8TA45
A	416	THR	-	linker	UNP A0A3Q8TA45
A	417	GLY	-	linker	UNP A0A3Q8TA45
A	418	GLY	-	linker	UNP A0A3Q8TA45
A	720	PHE	HIS	conflict	UNP A0A3Q8TA45
A	875	GLY	-	expression tag	UNP A0A3Q8TA45
A	876	PRO	-	expression tag	UNP A0A3Q8TA45
A	877	PHE	-	expression tag	UNP A0A3Q8TA45
A	878	GLU	-	expression tag	UNP A0A3Q8TA45
A	879	ASP	-	expression tag	UNP A0A3Q8TA45
A	880	ASP	-	expression tag	UNP A0A3Q8TA45
A	881	ASP	-	expression tag	UNP A0A3Q8TA45
A	882	ASP	-	expression tag	UNP A0A3Q8TA45
A	883	LYS	-	expression tag	UNP A0A3Q8TA45
A	884	ALA	-	expression tag	UNP A0A3Q8TA45
A	885	GLY	-	expression tag	UNP A0A3Q8TA45
A	886	TRP	-	expression tag	UNP A0A3Q8TA45
A	887	SER	-	expression tag	UNP A0A3Q8TA45
A	888	HIS	-	expression tag	UNP A0A3Q8TA45
A	889	PRO	-	expression tag	UNP A0A3Q8TA45
A	890	GLN	-	expression tag	UNP A0A3Q8TA45
A	891	PHE	-	expression tag	UNP A0A3Q8TA45
A	892	GLU	-	expression tag	UNP A0A3Q8TA45
A	893	LYS	-	expression tag	UNP A0A3Q8TA45
A	894	GLY	-	expression tag	UNP A0A3Q8TA45
A	895	GLY	-	expression tag	UNP A0A3Q8TA45
A	896	GLY	-	expression tag	UNP A0A3Q8TA45

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	897	SER	-	expression tag	UNP A0A3Q8TA45
A	898	GLY	-	expression tag	UNP A0A3Q8TA45
A	899	GLY	-	expression tag	UNP A0A3Q8TA45
A	900	GLY	-	expression tag	UNP A0A3Q8TA45
A	901	SER	-	expression tag	UNP A0A3Q8TA45
A	902	GLY	-	expression tag	UNP A0A3Q8TA45
A	903	GLY	-	expression tag	UNP A0A3Q8TA45
A	904	GLY	-	expression tag	UNP A0A3Q8TA45
A	905	SER	-	expression tag	UNP A0A3Q8TA45
A	906	TRP	-	expression tag	UNP A0A3Q8TA45
A	907	SER	-	expression tag	UNP A0A3Q8TA45
A	908	HIS	-	expression tag	UNP A0A3Q8TA45
A	909	PRO	-	expression tag	UNP A0A3Q8TA45
A	910	GLN	-	expression tag	UNP A0A3Q8TA45
A	911	PHE	-	expression tag	UNP A0A3Q8TA45
A	912	GLU	-	expression tag	UNP A0A3Q8TA45
A	913	LYS	-	expression tag	UNP A0A3Q8TA45
B	608	GLY	-	linker	UNP A0A3Q8TA45
B	609	GLY	-	linker	UNP A0A3Q8TA45
B	610	SER	-	linker	UNP A0A3Q8TA45
B	611	GLY	-	linker	UNP A0A3Q8TA45
B	612	LEU	-	linker	UNP A0A3Q8TA45
B	613	VAL	-	linker	UNP A0A3Q8TA45
B	614	PRO	-	linker	UNP A0A3Q8TA45
B	615	ARG	-	linker	UNP A0A3Q8TA45
B	616	GLY	-	linker	UNP A0A3Q8TA45
B	617	SER	-	linker	UNP A0A3Q8TA45
B	618	GLY	-	linker	UNP A0A3Q8TA45
B	619	GLY	-	linker	UNP A0A3Q8TA45
B	620	GLY	-	linker	UNP A0A3Q8TA45
B	621	SER	-	linker	UNP A0A3Q8TA45
B	622	GLY	-	linker	UNP A0A3Q8TA45
B	623	GLY	-	linker	UNP A0A3Q8TA45
B	624	GLY	-	linker	UNP A0A3Q8TA45
B	625	SER	-	linker	UNP A0A3Q8TA45
B	626	TRP	-	linker	UNP A0A3Q8TA45
B	627	SER	-	linker	UNP A0A3Q8TA45
B	628	HIS	-	linker	UNP A0A3Q8TA45
B	629	PRO	-	linker	UNP A0A3Q8TA45
B	630	GLN	-	linker	UNP A0A3Q8TA45
B	631	PHE	-	linker	UNP A0A3Q8TA45
B	632	GLU	-	linker	UNP A0A3Q8TA45

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	633	LYS	-	linker	UNP A0A3Q8TA45
B	634	GLY	-	linker	UNP A0A3Q8TA45
B	635	GLY	-	linker	UNP A0A3Q8TA45
B	636	GLY	-	linker	UNP A0A3Q8TA45
B	637	THR	-	linker	UNP A0A3Q8TA45
B	638	GLY	-	linker	UNP A0A3Q8TA45
B	639	GLY	-	linker	UNP A0A3Q8TA45
B	640	GLY	-	linker	UNP A0A3Q8TA45
B	641	THR	-	linker	UNP A0A3Q8TA45
B	642	LEU	-	linker	UNP A0A3Q8TA45
B	643	VAL	-	linker	UNP A0A3Q8TA45
B	644	PRO	-	linker	UNP A0A3Q8TA45
B	645	ARG	-	linker	UNP A0A3Q8TA45
B	646	GLY	-	linker	UNP A0A3Q8TA45
B	647	SER	-	linker	UNP A0A3Q8TA45
B	648	GLY	-	linker	UNP A0A3Q8TA45
B	649	THR	-	linker	UNP A0A3Q8TA45
B	650	GLY	-	linker	UNP A0A3Q8TA45
B	651	GLY	-	linker	UNP A0A3Q8TA45
B	953	PHE	HIS	conflict	UNP A0A3Q8TA45
B	1108	GLY	-	expression tag	UNP A0A3Q8TA45
B	1109	PRO	-	expression tag	UNP A0A3Q8TA45
B	1110	PHE	-	expression tag	UNP A0A3Q8TA45
B	1111	GLU	-	expression tag	UNP A0A3Q8TA45
B	1112	ASP	-	expression tag	UNP A0A3Q8TA45
B	1113	ASP	-	expression tag	UNP A0A3Q8TA45
B	1114	ASP	-	expression tag	UNP A0A3Q8TA45
B	1115	ASP	-	expression tag	UNP A0A3Q8TA45
B	1116	LYS	-	expression tag	UNP A0A3Q8TA45
B	1117	ALA	-	expression tag	UNP A0A3Q8TA45
B	1118	GLY	-	expression tag	UNP A0A3Q8TA45
B	1119	TRP	-	expression tag	UNP A0A3Q8TA45
B	1120	SER	-	expression tag	UNP A0A3Q8TA45
B	1121	HIS	-	expression tag	UNP A0A3Q8TA45
B	1122	PRO	-	expression tag	UNP A0A3Q8TA45
B	1123	GLN	-	expression tag	UNP A0A3Q8TA45
B	1124	PHE	-	expression tag	UNP A0A3Q8TA45
B	1125	GLU	-	expression tag	UNP A0A3Q8TA45
B	1126	LYS	-	expression tag	UNP A0A3Q8TA45
B	1127	GLY	-	expression tag	UNP A0A3Q8TA45
B	1128	GLY	-	expression tag	UNP A0A3Q8TA45
B	1129	GLY	-	expression tag	UNP A0A3Q8TA45

Continued on next page...

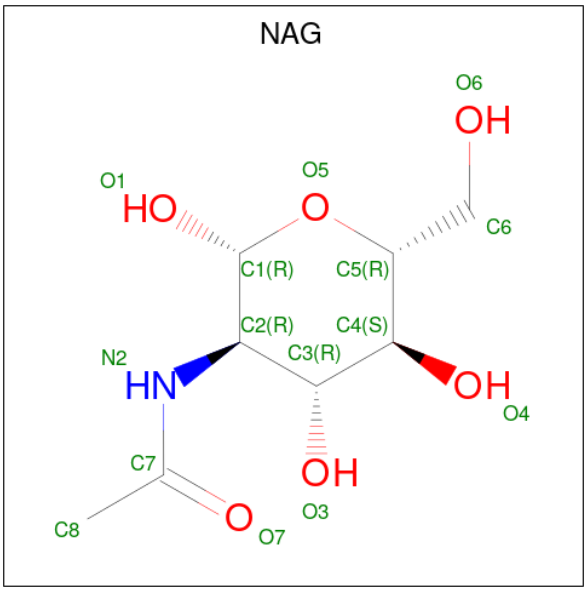
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	1130	SER	-	expression tag	UNP A0A3Q8TA45
B	1131	GLY	-	expression tag	UNP A0A3Q8TA45
B	1132	GLY	-	expression tag	UNP A0A3Q8TA45
B	1133	GLY	-	expression tag	UNP A0A3Q8TA45
B	1134	SER	-	expression tag	UNP A0A3Q8TA45
B	1135	GLY	-	expression tag	UNP A0A3Q8TA45
B	1136	GLY	-	expression tag	UNP A0A3Q8TA45
B	1137	GLY	-	expression tag	UNP A0A3Q8TA45
B	1138	SER	-	expression tag	UNP A0A3Q8TA45
B	1139	TRP	-	expression tag	UNP A0A3Q8TA45
B	1140	SER	-	expression tag	UNP A0A3Q8TA45
B	1141	HIS	-	expression tag	UNP A0A3Q8TA45
B	1142	PRO	-	expression tag	UNP A0A3Q8TA45
B	1143	GLN	-	expression tag	UNP A0A3Q8TA45
B	1144	PHE	-	expression tag	UNP A0A3Q8TA45
B	1145	GLU	-	expression tag	UNP A0A3Q8TA45
B	1146	LYS	-	expression tag	UNP A0A3Q8TA45

- Molecule 2 is an oligosaccharide called alpha-D-mannopyranose-(1-3)-beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.

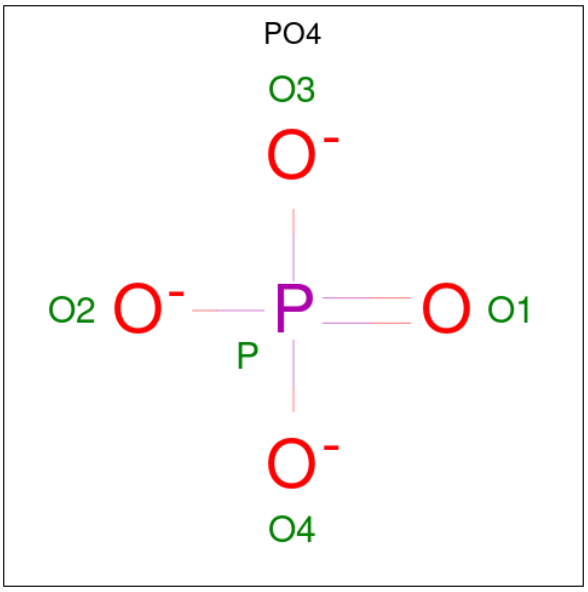
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	4	Total	C	N	O	0	0	0
			50	28	2	20			
2	D	4	Total	C	N	O	0	0	0
			50	28	2	20			

- Molecule 3 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).



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

- Molecule 4 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P) (labeled as "Ligand of Interest" by depositor).



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

- Molecule 5 is water.

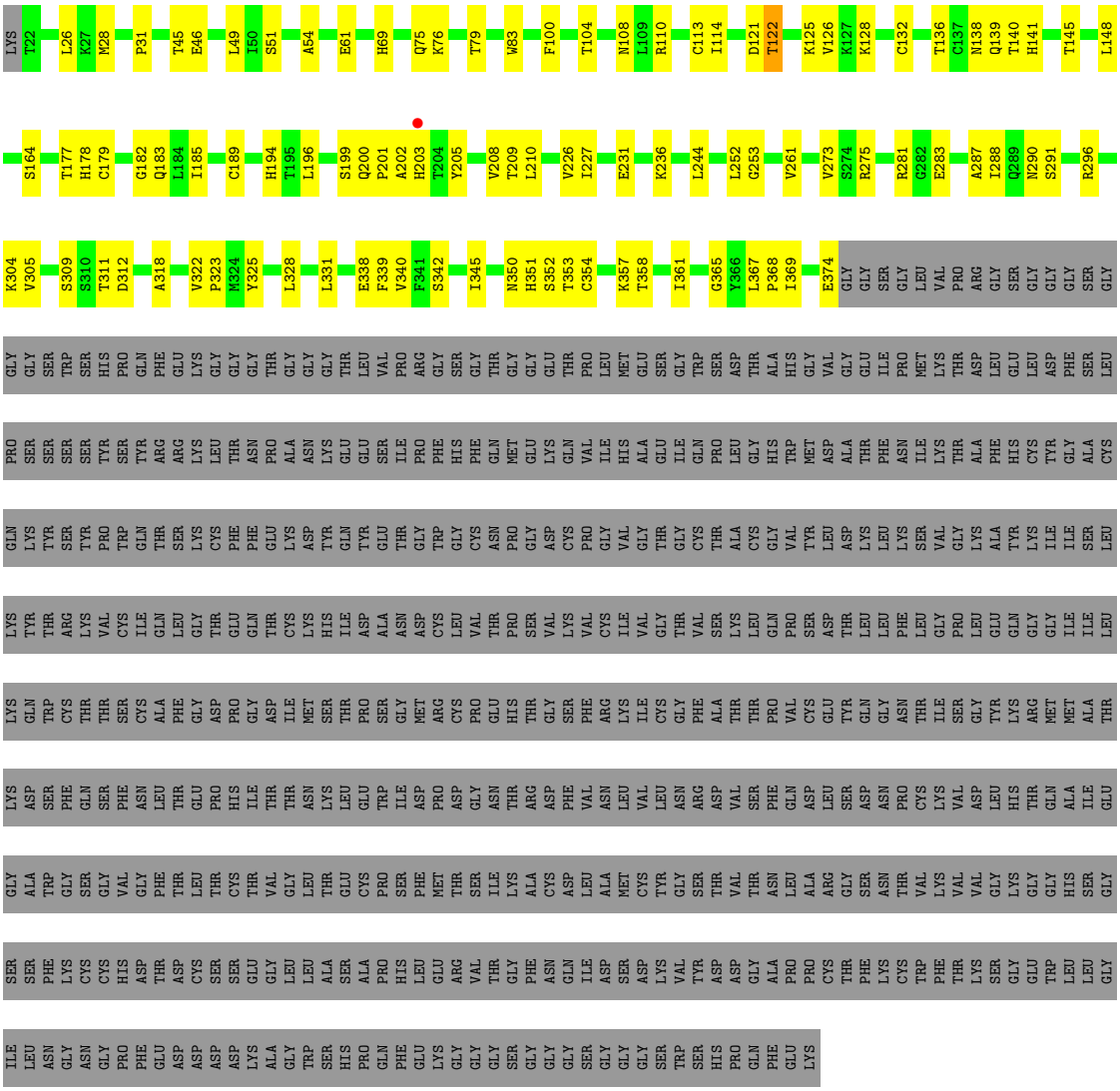
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total O 1 1	0	0

3 Residue-property plots (i)

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.

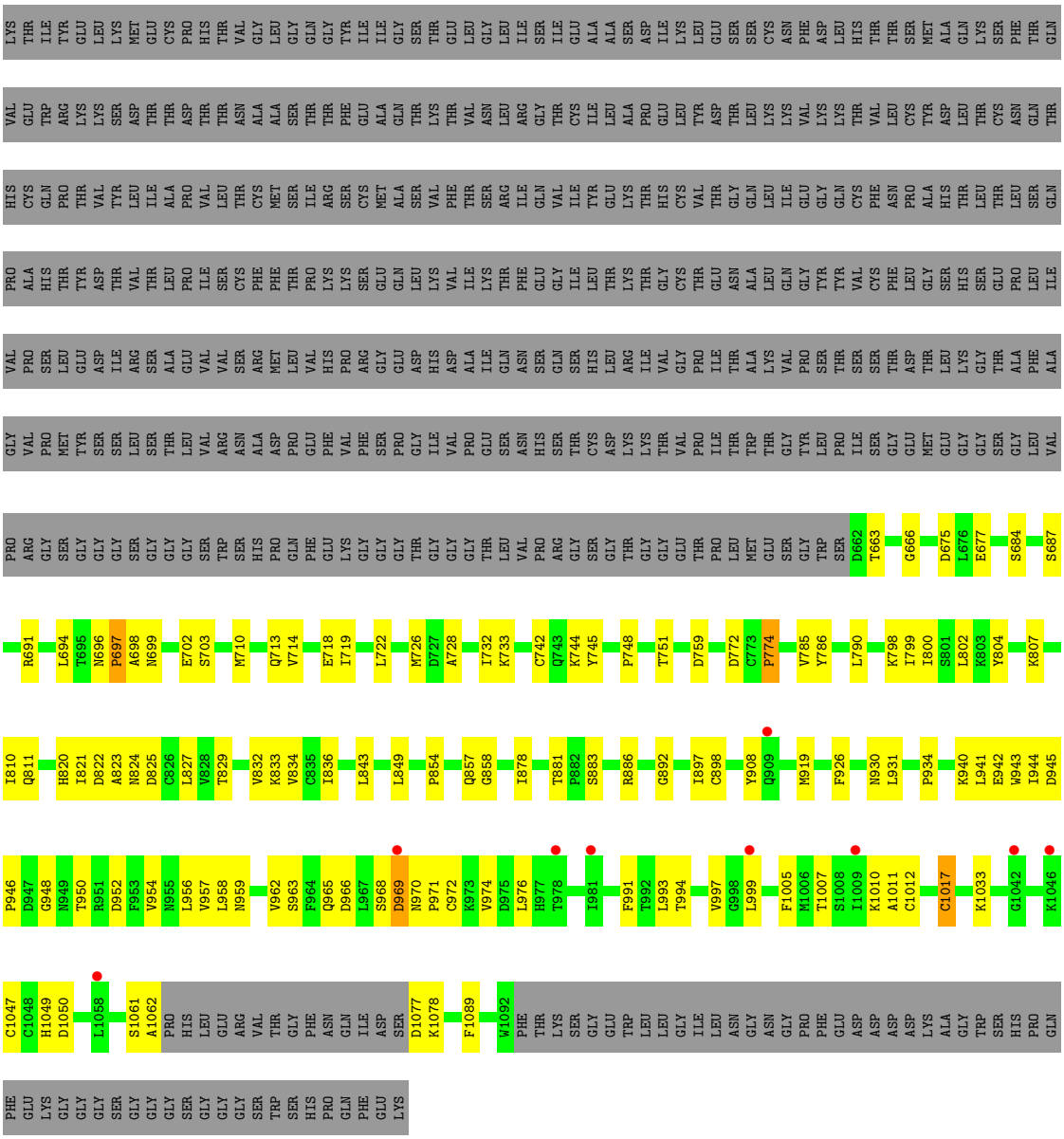
• Molecule 1: Envelope polyprotein,Envelope polyprotein

Chain A:



• Molecule 1: Envelope polyprotein,Envelope polyprotein

Chain B:



• Molecule 2: alpha-D-mannopyranose-(1-3)-beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain C:

• Molecule 2: alpha-D-mannopyranose-(1-3)-beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain D:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 63 2 2	Depositor
Cell constants a, b, c, α, β, γ	188.50Å 188.50Å 117.28Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.36 – 3.20 49.36 – 3.20	Depositor EDS
% Data completeness (in resolution range)	99.5 (49.36-3.20) 99.8 (49.36-3.20)	Depositor EDS
R_{merge}	0.20	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.20 (at 3.19Å)	Xtriage
Refinement program	PHENIX 1.14rc3	Depositor
R, R_{free}	0.211 , 0.266 0.211 , 0.266	Depositor DCC
R_{free} test set	1037 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	107.4	Xtriage
Anisotropy	0.154	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 58.7	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	0.92	EDS
Total number of atoms	6021	wwPDB-VP
Average B, all atoms (Å^2)	104.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.19% 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: PO4, NAG, MAN, BMA

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	0.38	0/2765	0.61	0/3764
1	B	0.39	0/3264	0.60	0/4424
All	All	0.38	0/6029	0.61	0/8188

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	2709	0	2692	80	0
1	B	3187	0	3052	74	0
2	C	50	0	43	7	0
2	D	50	0	43	4	0
3	B	14	0	13	0	0
4	B	10	0	0	0	0
5	A	1	0	0	0	0
All	All	6021	0	5843	155	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 13.

The worst 5 of 155 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:B:696:ASN:O	1:B:698:ALA:N	2.11	0.83
1:B:663:THR:HG23	1:B:1017:CYS:HB3	1.64	0.79
1:A:304:LYS:NZ	1:A:311:THR:O	2.19	0.76
1:A:185:ILE:HD12	1:A:210:LEU:HD23	1.68	0.75
1:B:687:SER:OG	1:B:1078:LYS:HE2	1.90	0.71

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	351/893 (39%)	327 (93%)	23 (7%)	1 (0%)	41 74
1	B	413/893 (46%)	374 (91%)	36 (9%)	3 (1%)	22 61
All	All	764/1786 (43%)	701 (92%)	59 (8%)	4 (0%)	29 67

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	697	PRO
1	B	969	ASP
1	A	122	THR
1	B	774	PRO

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	316/761 (42%)	315 (100%)	1 (0%)	92 96
1	B	359/761 (47%)	354 (99%)	5 (1%)	67 86
All	All	675/1522 (44%)	669 (99%)	6 (1%)	78 91

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	1017	CYS
1	B	1047	CYS
1	B	1089	PHE
1	B	699	ASN
1	A	354	CYS

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

Mol	Chain	Res	Type
1	A	75	GLN
1	B	824	ASN

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 (i)

8 monosaccharides 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	C	1	2,1	14,14,15	0.81	1 (7%)	17,19,21	1.61	3 (17%)
2	NAG	C	2	2	14,14,15	0.67	1 (7%)	17,19,21	0.68	0
2	BMA	C	3	2	11,11,12	1.45	3 (27%)	15,15,17	0.83	0
2	MAN	C	4	2	11,11,12	1.32	2 (18%)	15,15,17	1.31	2 (13%)
2	NAG	D	1	2,1	14,14,15	1.28	1 (7%)	17,19,21	1.12	0
2	NAG	D	2	2	14,14,15	0.58	1 (7%)	17,19,21	0.63	0
2	BMA	D	3	2	11,11,12	1.50	2 (18%)	15,15,17	1.54	2 (13%)
2	MAN	D	4	2	11,11,12	1.40	2 (18%)	15,15,17	1.43	3 (20%)

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	C	1	2,1	-	2/6/23/26	0/1/1/1
2	NAG	C	2	2	-	3/6/23/26	0/1/1/1
2	BMA	C	3	2	-	2/2/19/22	0/1/1/1
2	MAN	C	4	2	-	2/2/19/22	0/1/1/1
2	NAG	D	1	2,1	1/1/5/7	3/6/23/26	0/1/1/1
2	NAG	D	2	2	-	2/6/23/26	0/1/1/1
2	BMA	D	3	2	-	2/2/19/22	0/1/1/1
2	MAN	D	4	2	-	0/2/19/22	0/1/1/1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	1	NAG	O5-C1	-4.30	1.36	1.43
2	D	3	BMA	C1-C2	3.46	1.60	1.52
2	D	4	MAN	C1-C2	2.95	1.58	1.52
2	C	4	MAN	C1-C2	2.91	1.58	1.52
2	C	1	NAG	O5-C1	2.59	1.47	1.43

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	1	NAG	C1-O5-C5	4.81	118.71	112.19
2	D	3	BMA	C1-C2-C3	4.22	114.85	109.67
2	C	4	MAN	C1-O5-C5	3.60	117.07	112.19
2	D	4	MAN	C1-O5-C5	3.37	116.76	112.19
2	C	1	NAG	C2-N2-C7	2.52	126.50	122.90

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	D	1	NAG	C1

5 of 16 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	1	NAG	C3-C2-N2-C7
2	C	2	NAG	O5-C5-C6-O6
2	D	2	NAG	O5-C5-C6-O6
2	C	4	MAN	O5-C5-C6-O6
2	D	3	BMA	O5-C5-C6-O6

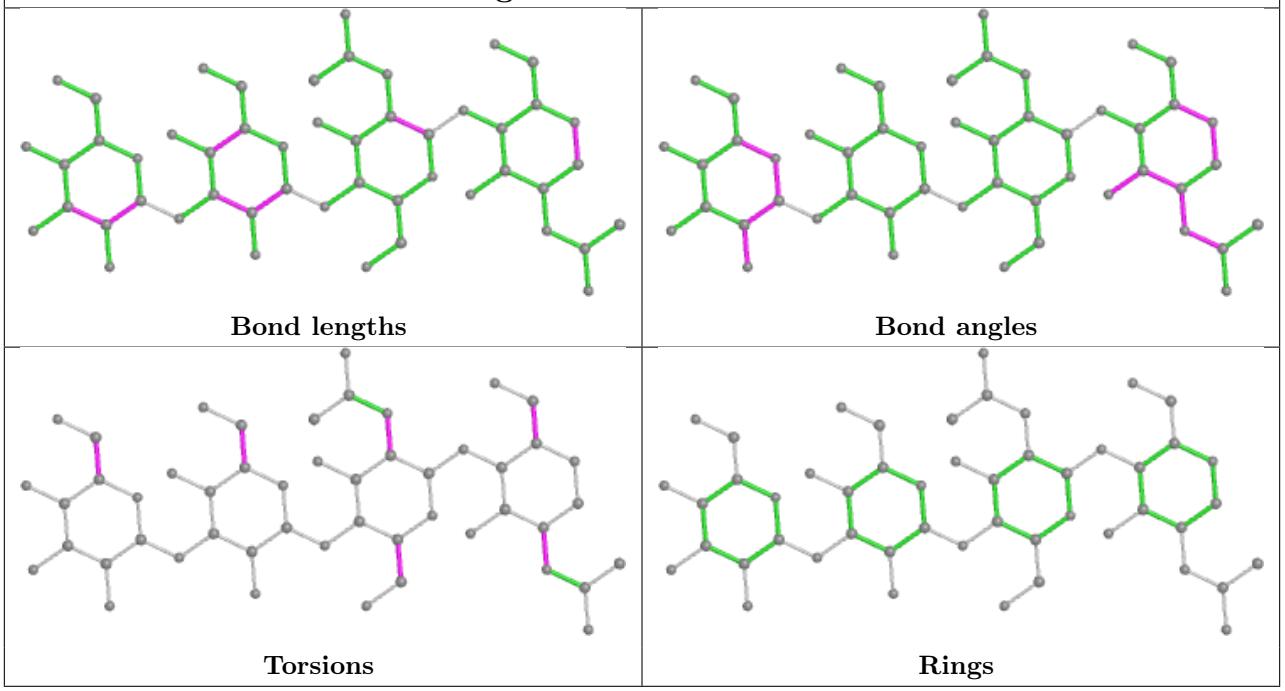
There are no ring outliers.

5 monomers are involved in 11 short contacts:

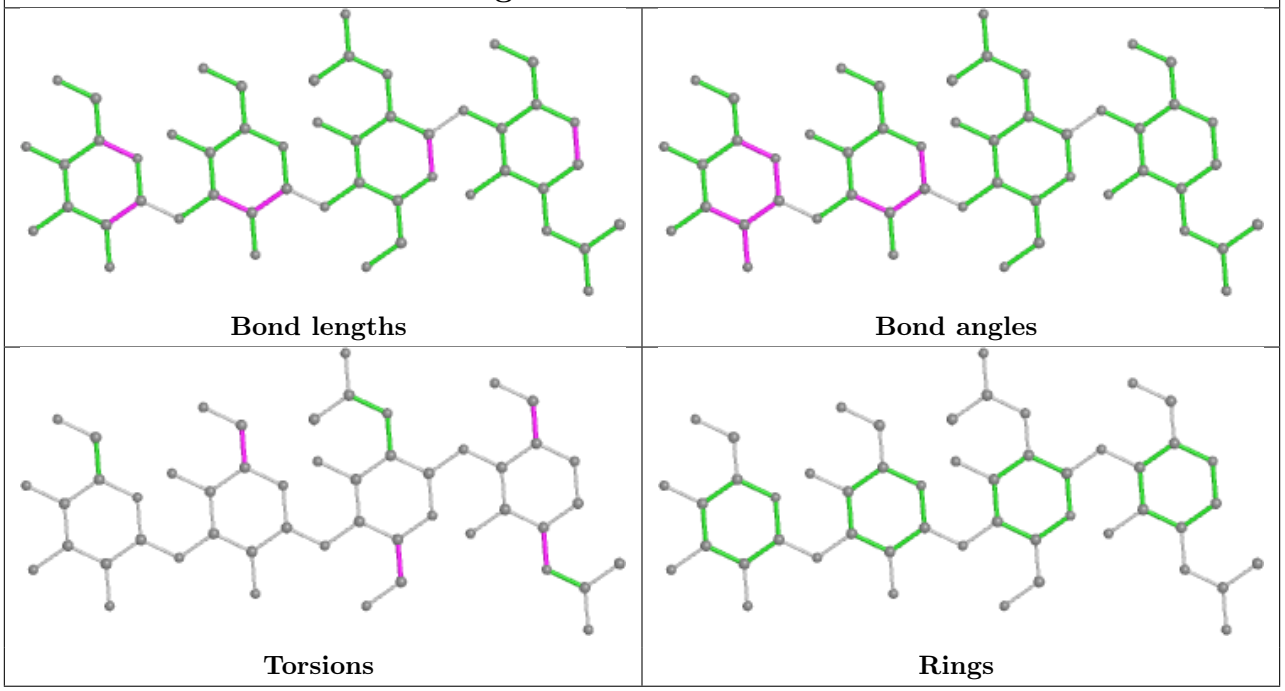
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	1	NAG	1	0
2	C	1	NAG	7	0
2	D	3	BMA	3	0
2	D	2	NAG	1	0
2	C	2	NAG	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

Oligosaccharide Chain C



Oligosaccharide Chain D



5.6 Ligand geometry (i)

3 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	$\#\|Z\| > 2$
4	PO4	B	1202	-	4,4,4	0.82	0	6,6,6	0.42	0
4	PO4	B	1203	-	4,4,4	0.84	0	6,6,6	0.62	0
3	NAG	B	1201	-	14,14,15	0.34	0	17,19,21	0.40	0

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	B	1201	-	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

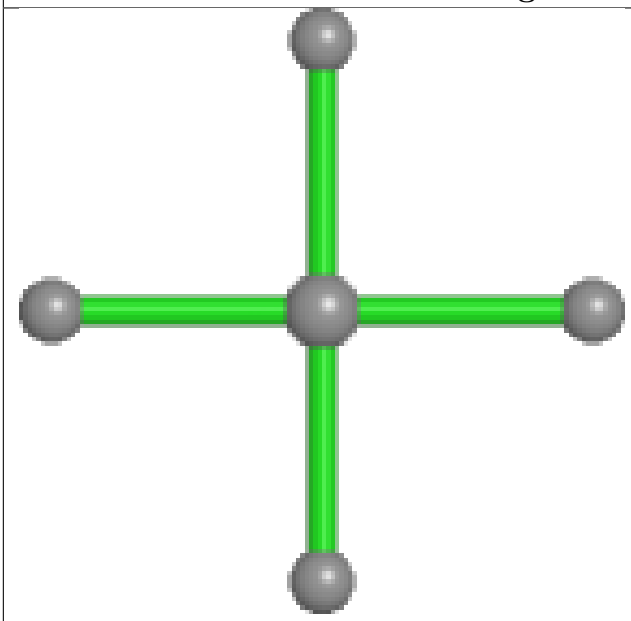
Mol	Chain	Res	Type	Atoms
3	B	1201	NAG	C4-C5-C6-O6
3	B	1201	NAG	O5-C5-C6-O6

There are no ring outliers.

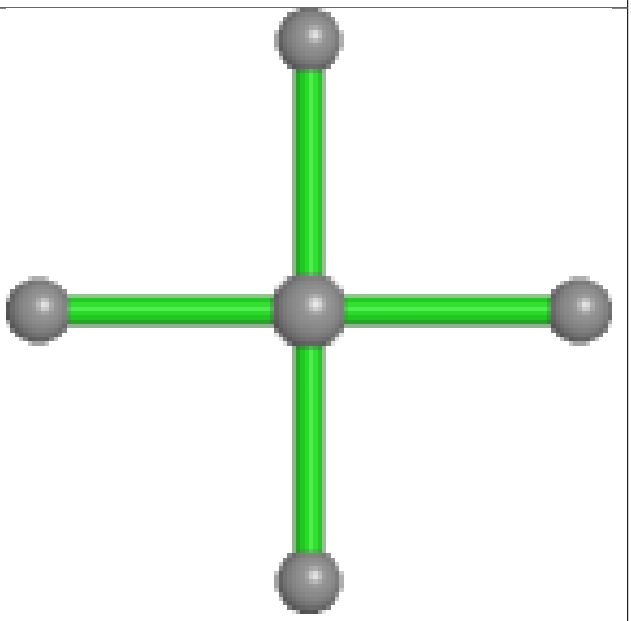
No monomer is involved in short contacts.

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.

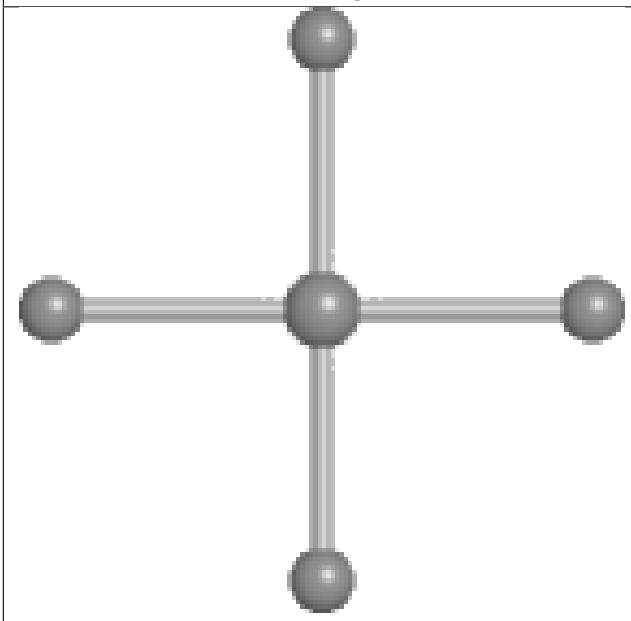
Ligand PO4 B 1202



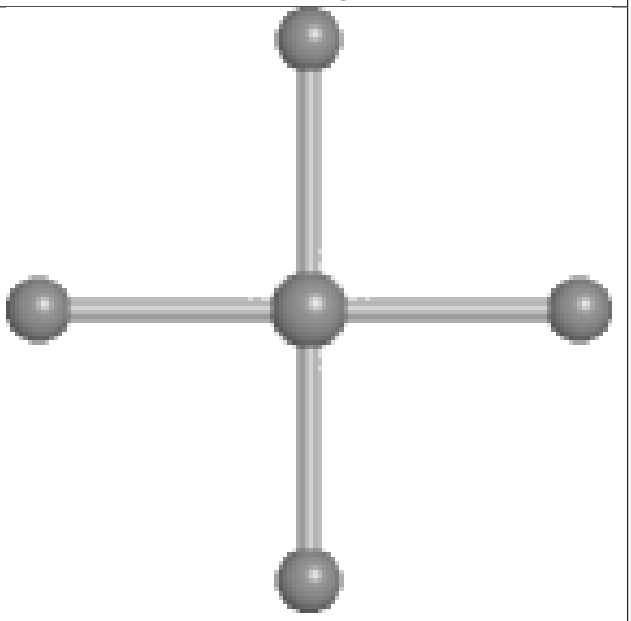
Bond lengths



Bond angles

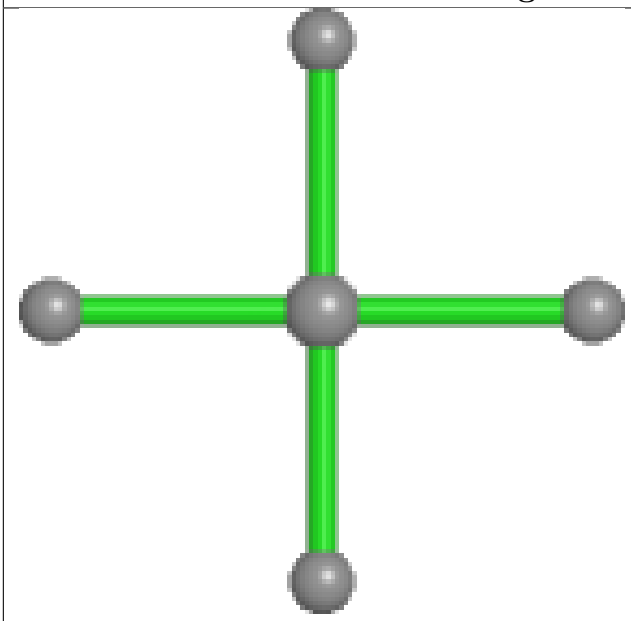


Torsions

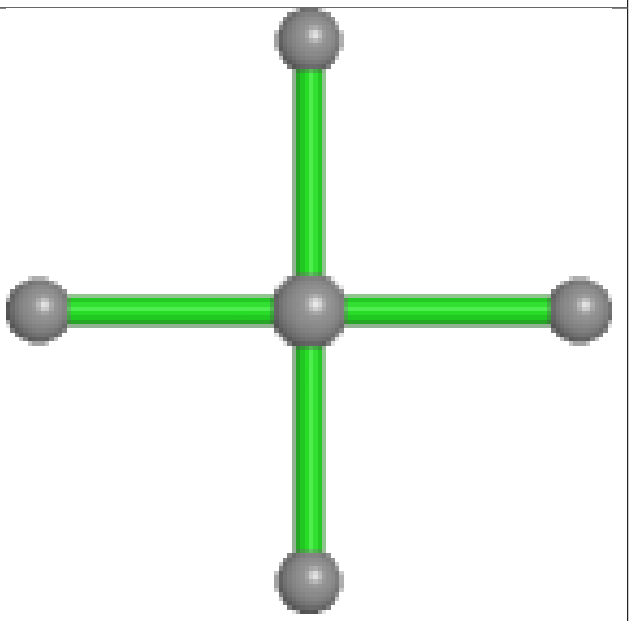


Rings

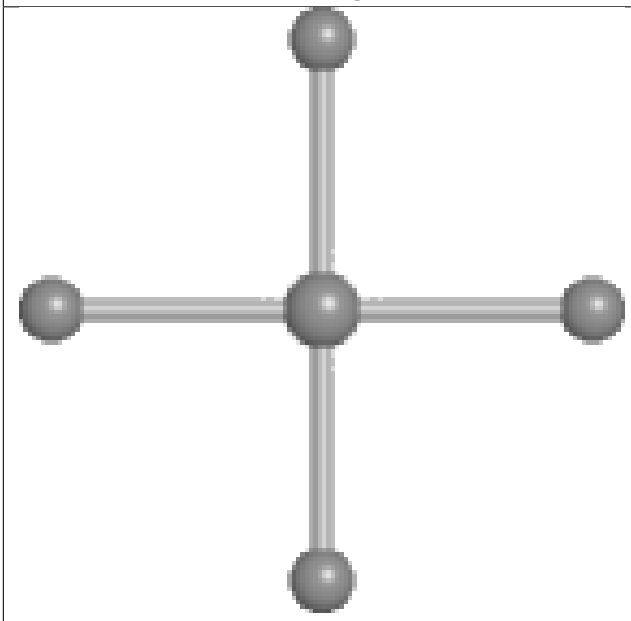
Ligand PO4 B 1203



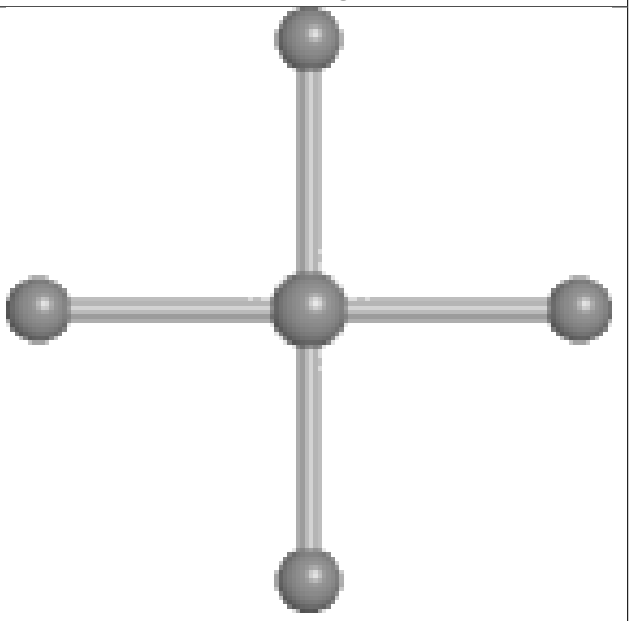
Bond lengths



Bond angles

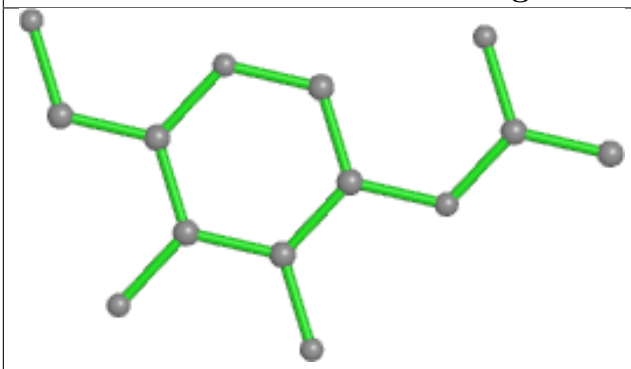


Torsions

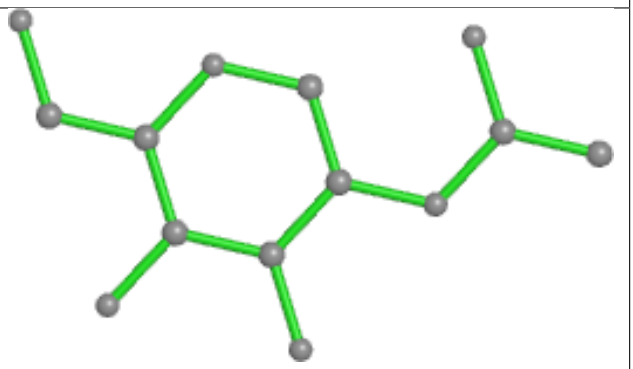


Rings

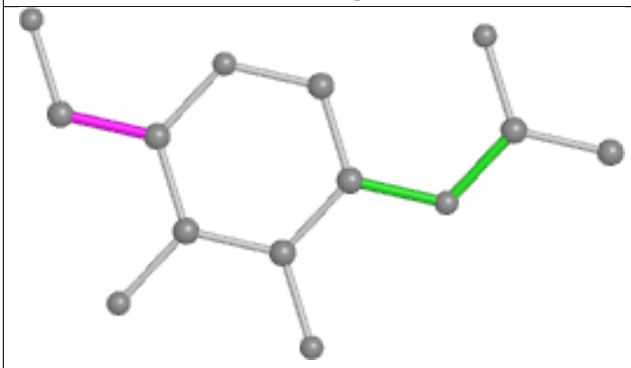
Ligand NAG B 1201



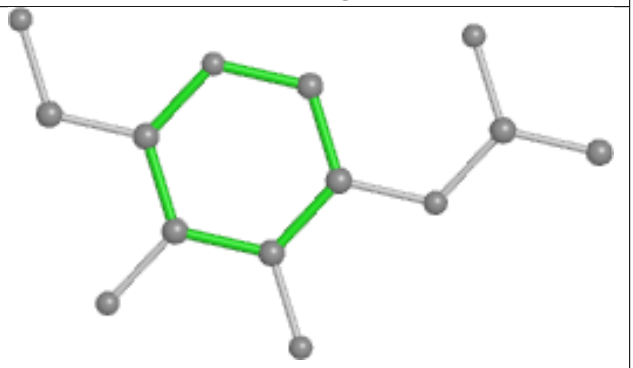
Bond lengths



Bond angles



Torsions



Rings

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^{th} 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	353/893 (39%)	-0.04	1 (0%)	94	92	77, 101, 135, 164	0
1	B	417/893 (46%)	0.14	9 (2%)	62	48	67, 103, 133, 171	0
All	All	770/1786 (43%)	0.06	10 (1%)	77	65	67, 102, 134, 171	0

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	203	HIS	4.2
1	B	978	THR	3.1
1	B	1042	GLY	2.7
1	B	1009	ILE	2.7
1	B	969	ASP	2.6

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

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, 95^{th} 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	MAN	C	4	11/12	0.75	0.28	113,130,139,142	0
2	NAG	C	1	14/15	0.85	0.27	90,104,129,132	0
2	BMA	D	3	11/12	0.88	0.12	127,139,147,150	0
2	MAN	D	4	11/12	0.90	0.24	92,125,141,143	0
2	BMA	C	3	11/12	0.93	0.12	120,125,131,134	0

Continued on next page...

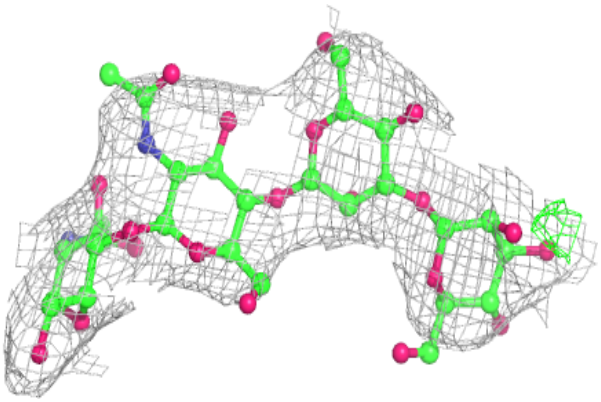
Continued from previous page...

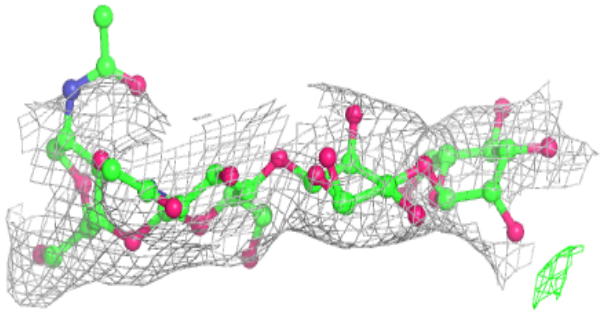
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	NAG	D	2	14/15	0.94	0.16	85,98,107,115	0
2	NAG	D	1	14/15	0.95	0.20	72,100,108,115	0
2	NAG	C	2	14/15	0.96	0.19	89,105,120,139	0

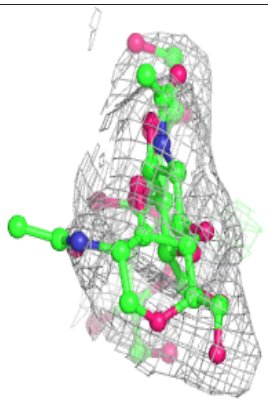
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around Chain C:

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

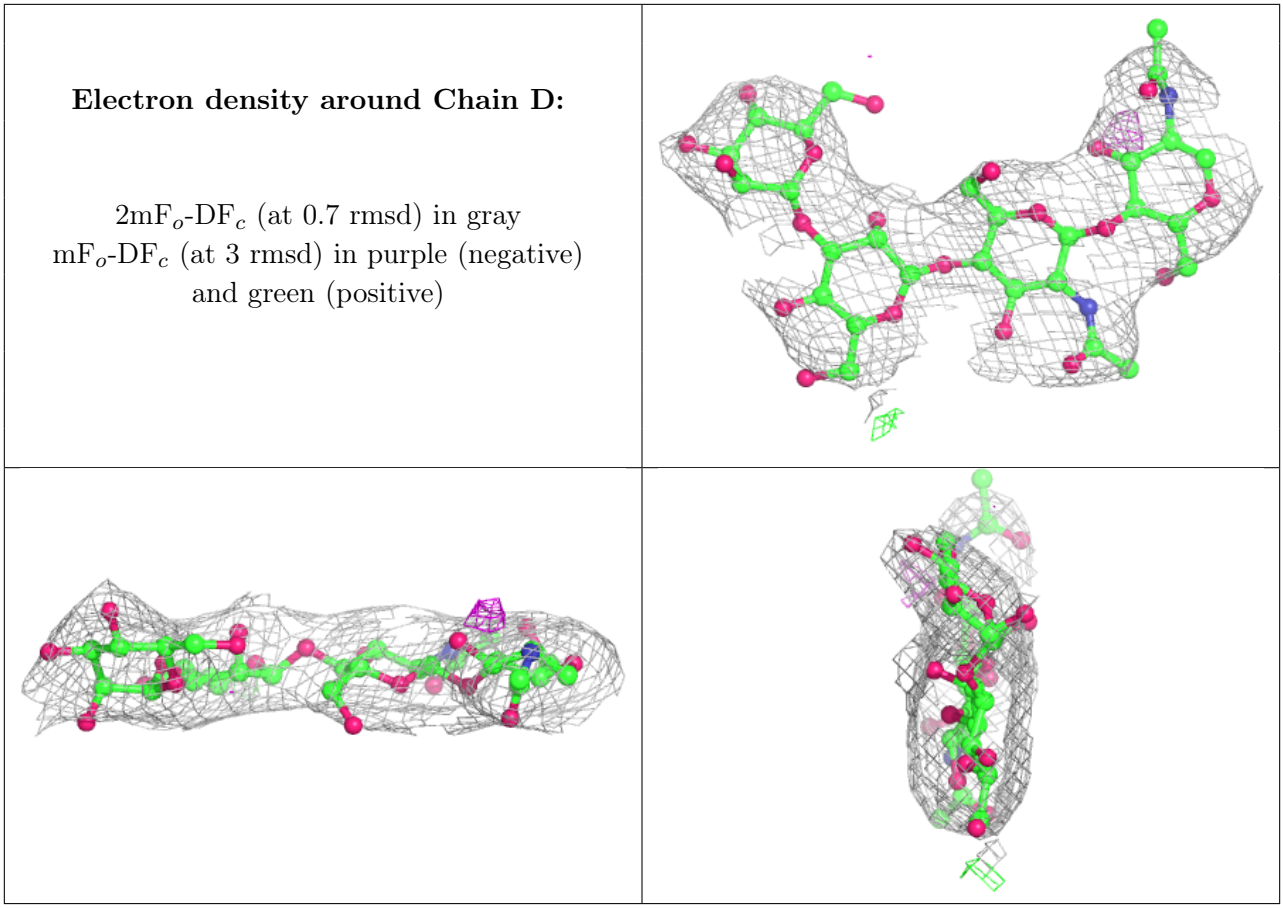






Electron density around Chain D:

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



6.4 Ligands

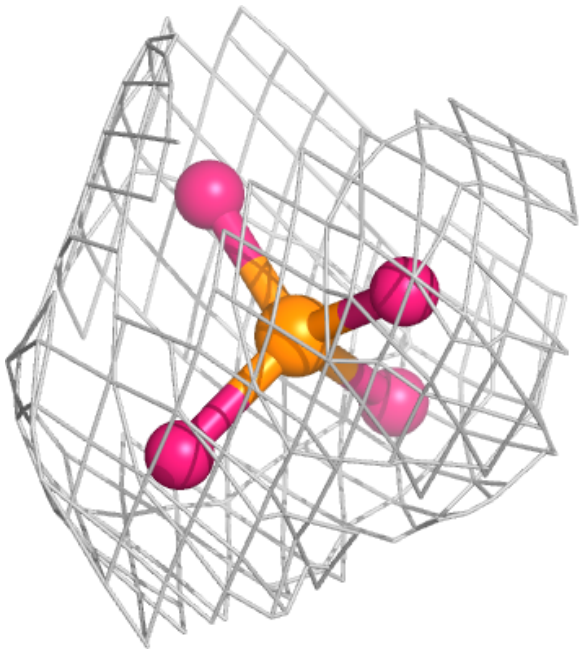
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

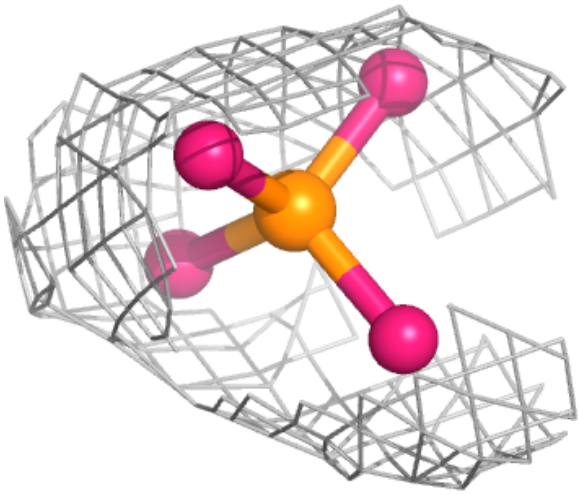
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	PO4	B	1203	5/5	0.75	0.20	127,136,176,178	0
4	PO4	B	1202	5/5	0.81	0.20	146,150,169,182	0
3	NAG	B	1201	14/15	0.83	0.19	85,126,138,138	0

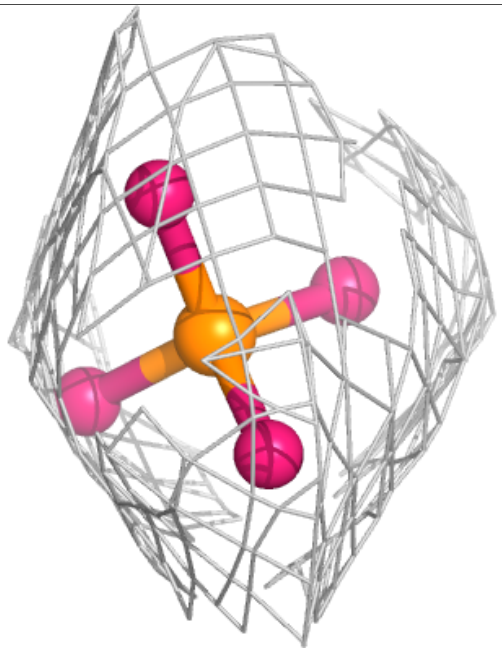
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around PO4 B 1203:

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

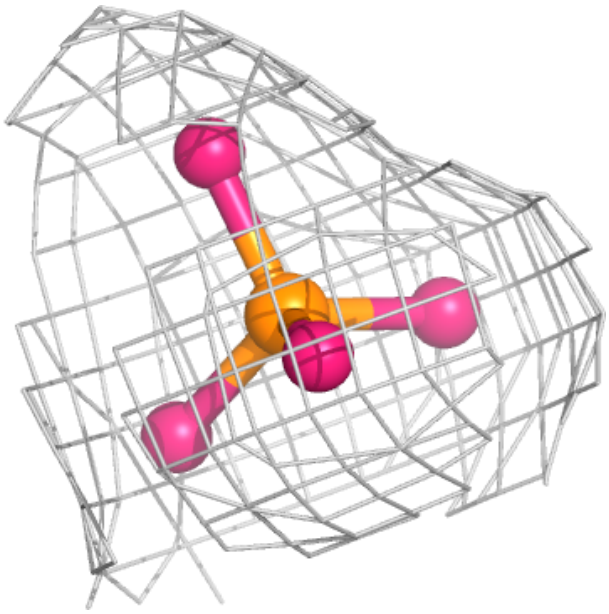


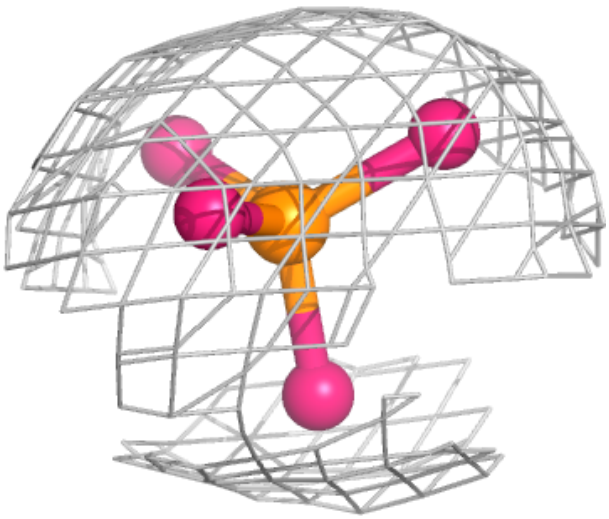


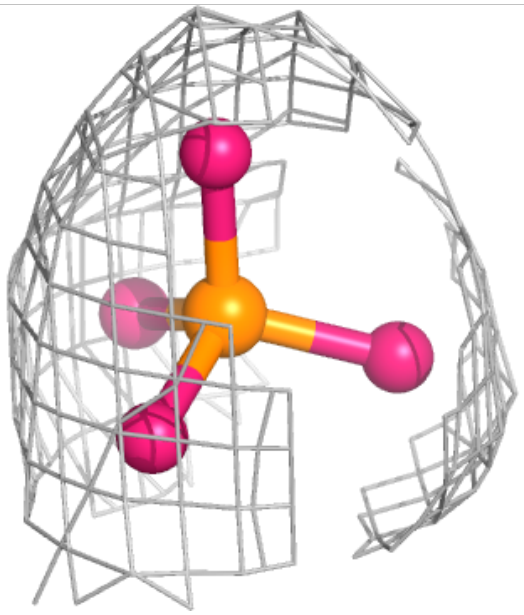


Electron density around PO4 B 1202:

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

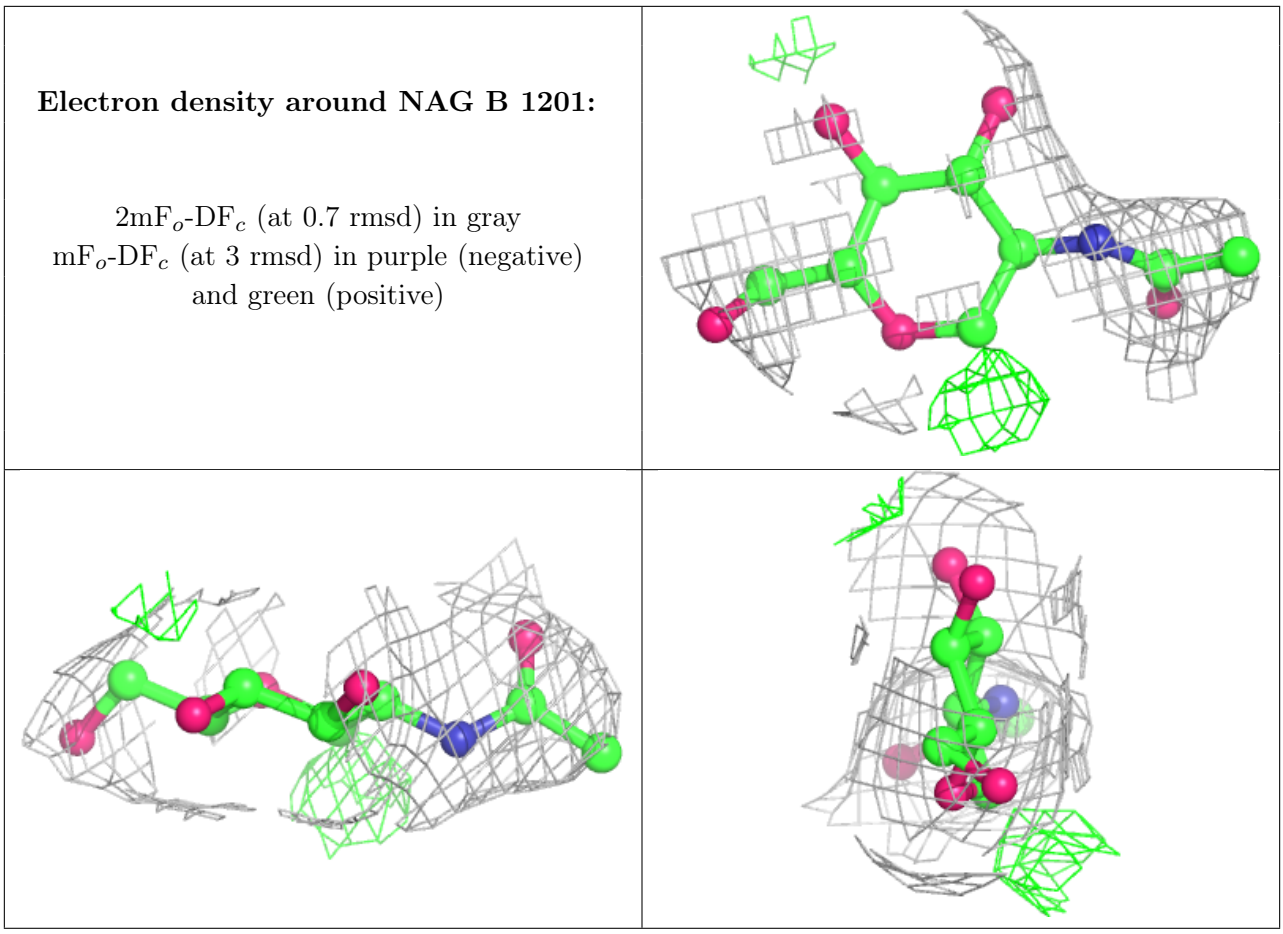






Electron density around NAG B 1201:

$2mF_o$-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.