Full wwPDB X-ray Structure Validation Report (i)

Oct 4, 2023 – 07:23 AM EDT

PDB ID	:	6PKU
Title	:	Guinea pig N-acetylglucosamine-1-phosphodiester alpha-N-acetylglucosaminidase (NAGPA) catalytic domain (C51S C221S) in complex with N-acetyl-alpha-D-glucosamine (alpha-GlcNAc) and mannose 6-phosphate (M6P)
Authors	:	Gorelik, A.; Illes, K.; Nagar, B.
Deposited on	:	2019-06-29
Resolution	:	1.95 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

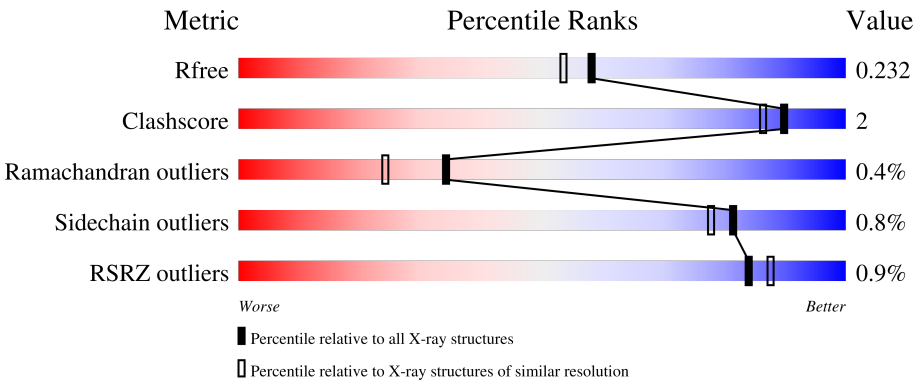
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 1.95 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	4310 (1.96-1.92)
Clashscore	141614	1023 (1.94-1.94)
Ramachandran outliers	138981	1007 (1.94-1.94)
Sidechain outliers	138945	1007 (1.94-1.94)
RSRZ outliers	127900	4250 (1.96-1.92)

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	300	86% • 12%
1	B	300	% 84% • 12%
1	C	300	% 83% 5% 12%
1	D	300	% 82% 5% 12%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 16604 atoms, of which 7802 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 N-acetylglucosamine-1-phosphodiester alpha-N-acetylglucosa minidase (NAGPA).

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	263	3919	1234	1919	371	385	10	0	0	0
1	B	264	3932	1239	1925	372	386	10	0	0	0
1	C	263	3920	1234	1920	371	385	10	0	0	0
1	D	263	3920	1234	1920	371	385	10	0	0	0

There are 52 discrepancies between the modelled and reference sequences:

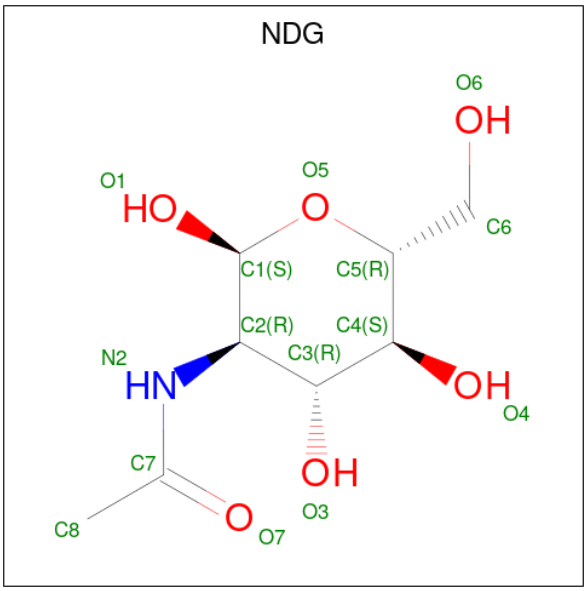
Chain	Residue	Modelled	Actual	Comment	Reference
A	28	ASP	-	expression tag	UNP H0VTT5
A	29	ARG	-	expression tag	UNP H0VTT5
A	30	HIS	-	expression tag	UNP H0VTT5
A	31	HIS	-	expression tag	UNP H0VTT5
A	32	HIS	-	expression tag	UNP H0VTT5
A	33	HIS	-	expression tag	UNP H0VTT5
A	34	HIS	-	expression tag	UNP H0VTT5
A	35	HIS	-	expression tag	UNP H0VTT5
A	36	GLY	-	expression tag	UNP H0VTT5
A	37	SER	-	expression tag	UNP H0VTT5
A	38	GLY	-	expression tag	UNP H0VTT5
A	51	SER	CYS	engineered mutation	UNP H0VTT5
A	221	SER	CYS	engineered mutation	UNP H0VTT5
B	28	ASP	-	expression tag	UNP H0VTT5
B	29	ARG	-	expression tag	UNP H0VTT5
B	30	HIS	-	expression tag	UNP H0VTT5
B	31	HIS	-	expression tag	UNP H0VTT5
B	32	HIS	-	expression tag	UNP H0VTT5
B	33	HIS	-	expression tag	UNP H0VTT5
B	34	HIS	-	expression tag	UNP H0VTT5

Continued on next page...

Continued from previous page...

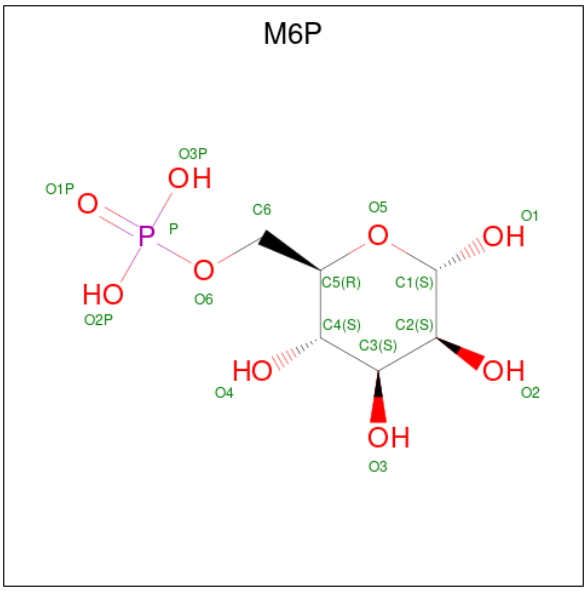
Chain	Residue	Modelled	Actual	Comment	Reference
B	35	HIS	-	expression tag	UNP H0VTT5
B	36	GLY	-	expression tag	UNP H0VTT5
B	37	SER	-	expression tag	UNP H0VTT5
B	38	GLY	-	expression tag	UNP H0VTT5
B	51	SER	CYS	engineered mutation	UNP H0VTT5
B	221	SER	CYS	engineered mutation	UNP H0VTT5
C	28	ASP	-	expression tag	UNP H0VTT5
C	29	ARG	-	expression tag	UNP H0VTT5
C	30	HIS	-	expression tag	UNP H0VTT5
C	31	HIS	-	expression tag	UNP H0VTT5
C	32	HIS	-	expression tag	UNP H0VTT5
C	33	HIS	-	expression tag	UNP H0VTT5
C	34	HIS	-	expression tag	UNP H0VTT5
C	35	HIS	-	expression tag	UNP H0VTT5
C	36	GLY	-	expression tag	UNP H0VTT5
C	37	SER	-	expression tag	UNP H0VTT5
C	38	GLY	-	expression tag	UNP H0VTT5
C	51	SER	CYS	engineered mutation	UNP H0VTT5
C	221	SER	CYS	engineered mutation	UNP H0VTT5
D	28	ASP	-	expression tag	UNP H0VTT5
D	29	ARG	-	expression tag	UNP H0VTT5
D	30	HIS	-	expression tag	UNP H0VTT5
D	31	HIS	-	expression tag	UNP H0VTT5
D	32	HIS	-	expression tag	UNP H0VTT5
D	33	HIS	-	expression tag	UNP H0VTT5
D	34	HIS	-	expression tag	UNP H0VTT5
D	35	HIS	-	expression tag	UNP H0VTT5
D	36	GLY	-	expression tag	UNP H0VTT5
D	37	SER	-	expression tag	UNP H0VTT5
D	38	GLY	-	expression tag	UNP H0VTT5
D	51	SER	CYS	engineered mutation	UNP H0VTT5
D	221	SER	CYS	engineered mutation	UNP H0VTT5

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



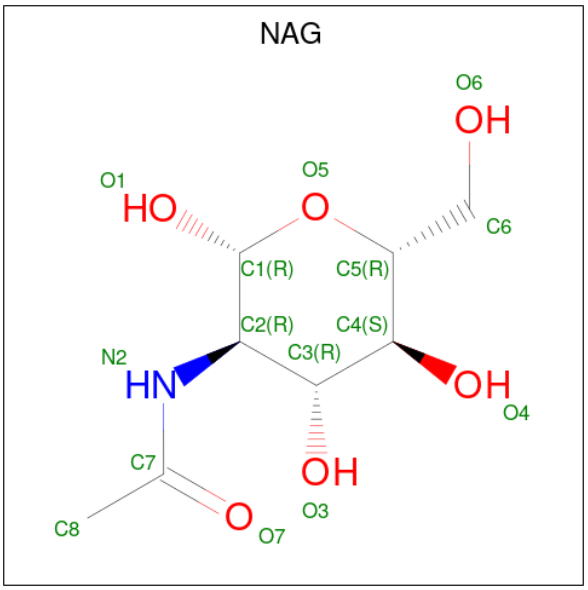
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	H	N	O	0	0
			30	8	15	1	6		
2	B	1	Total	C	H	N	O	0	0
			30	8	15	1	6		
2	C	1	Total	C	H	N	O	0	0
			30	8	15	1	6		
2	D	1	Total	C	H	N	O	0	0
			30	8	15	1	6		

- Molecule 3 is 6-O-phosphono-alpha-D-mannopyranose (three-letter code: M6P) (formula: $C_6H_{13}O_9P$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	H	O	P	0	0
			27	6	11	9	1		
3	B	1	Total	C	H	O	P	0	0
			27	6	11	9	1		
3	C	1	Total	C	H	O	P	0	0
			27	6	11	9	1		
3	D	1	Total	C	H	O	P	0	0
			27	6	11	9	1		

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	B	1	Total	C	H	N	O	0	0
			28	8	14	1	5		

- Molecule 5 is water.

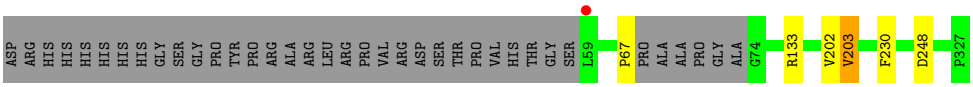
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	181	Total	O	0	0
			181	181		
5	B	165	Total	O	0	0
			165	165		
5	C	163	Total	O	0	0
			163	163		
5	D	148	Total	O	0	0
			148	148		

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.

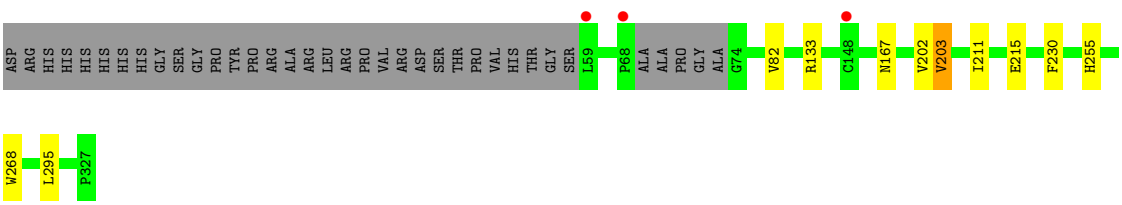
• Molecule 1: N-acetylglucosamine-1-phosphodiester alpha-N-acetylglucosaminidase (NAGPA)

Chain A:



• Molecule 1: N-acetylglucosamine-1-phosphodiester alpha-N-acetylglucosaminidase (NAGPA)

Chain B:

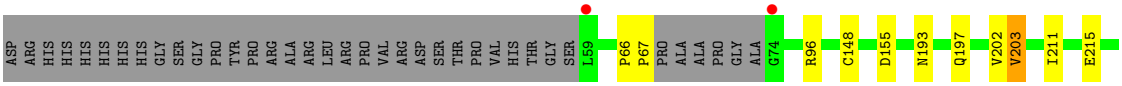


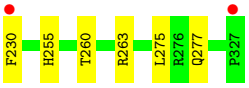
• Molecule 1: N-acetylglucosamine-1-phosphodiester alpha-N-acetylglucosaminidase (NAGPA)

Chain C:

• Molecule 1: N-acetylglucosamine-1-phosphodiester alpha-N-acetylglucosaminidase (NAGPA)

Chain D:


ASP ARG HIS HIS HIS HIS HIS GLY SER GLY PRO TYR PRO ARG ALA ARG LEU ARG PRO VAL ARG ASP SER THR PRO VAL HIS THR GLY SER
L59
P66 P67
PRO ALA ALA PRO GLY ALA
G74
R96
C148
D155
N193
Q197
V202 V203
I211
E215


F230
H255
T260
R263
L275
R276
Q277
P327

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	43.46Å 232.56Å 65.31Å 90.00° 109.15° 90.00°	Depositor
Resolution (Å)	30.78 – 1.95 30.78 – 1.95	Depositor EDS
% Data completeness (in resolution range)	85.0 (30.78-1.95) 79.5 (30.78-1.95)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.65 (at 1.95Å)	Xtriage
Refinement program	PHENIX (1.15_3459: ???)	Depositor
R, R_{free}	0.201 , 0.232 0.201 , 0.232	Depositor DCC
R_{free} test set	1778 reflections (2.34%)	wwPDB-VP
Wilson B-factor (Å^2)	21.1	Xtriage
Anisotropy	0.057	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 26.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.27$	Xtriage
Estimated twinning fraction	0.327 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	16604	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.18% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: NDG, NAG, M6P

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	0.36	0/2040	0.56	0/2764
1	B	0.35	0/2048	0.54	0/2776
1	C	0.34	0/2040	0.55	0/2764
1	D	0.35	0/2040	0.53	0/2764
All	All	0.35	0/8168	0.55	0/11068

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 (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	2000	1919	1919	3	0
1	B	2007	1925	1925	7	0
1	C	2000	1920	1919	8	0
1	D	2000	1920	1919	11	0
2	A	15	15	12	0	0
2	B	15	15	12	0	0
2	C	15	15	12	0	0
2	D	15	15	12	0	0
3	A	16	11	10	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	16	11	10	0	0
3	C	16	11	10	0	0
3	D	16	11	10	1	0
4	B	14	14	13	1	0
5	A	181	0	0	1	0
5	B	165	0	0	2	0
5	C	163	0	0	2	0
5	D	148	0	0	4	0
All	All	8802	7802	7783	28	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 2.

All (28) 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:295:LEU:HD21	4:B:401:NAG:H81	1.71	0.72
1:D:215:GLU:OE1	1:D:215:GLU:N	2.40	0.53
1:D:277:GLN:OE1	5:D:501:HOH:O	2.19	0.53
1:C:202:VAL:O	1:C:203:VAL:HB	2.10	0.51
1:C:157:ARG:NH1	5:C:512:HOH:O	2.44	0.50
1:A:202:VAL:O	1:A:203:VAL:HB	2.14	0.48
1:D:202:VAL:O	1:D:203:VAL:HB	2.15	0.46
1:B:133:ARG:NH1	5:B:514:HOH:O	2.49	0.46
1:D:96:ARG:NH2	5:D:501:HOH:O	2.48	0.45
1:C:88:ARG:NH1	5:C:519:HOH:O	2.48	0.45
1:D:211:ILE:HG13	1:D:275:LEU:HD11	1.99	0.45
1:D:148:CYS:O	5:D:502:HOH:O	2.20	0.44
1:B:82:VAL:HG11	1:D:263:ARG:CZ	2.48	0.44
1:D:155:ASP:OD1	1:D:197:GLN:NE2	2.42	0.44
1:A:67:PRO:HG3	1:A:248:ASP:O	2.18	0.43
1:C:133:ARG:NH1	1:C:326:GLU:OE2	2.51	0.43
1:C:237:MET:HG2	1:C:263:ARG:C	2.38	0.43
1:B:211:ILE:HD11	1:B:268:TRP:CE3	2.54	0.43
1:B:202:VAL:O	1:B:203:VAL:HB	2.18	0.43
1:C:167:ASN:HA	1:C:220:GLU:HG2	2.00	0.43
1:A:133:ARG:NH1	5:A:517:HOH:O	2.52	0.42
1:D:193:ASN:ND2	5:D:505:HOH:O	2.45	0.42
1:D:260:THR:OG1	3:D:402:M6P:O2P	2.32	0.42
1:C:84:HIS:HA	1:C:88:ARG:O	2.20	0.41
1:B:215:GLU:OE1	1:B:215:GLU:N	2.45	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:66:PRO:HA	1:D:67:PRO:HD3	1.97	0.41
1:C:66:PRO:HA	1:C:67:PRO:HD3	1.97	0.40
1:B:167:ASN:ND2	5:B:506:HOH:O	2.39	0.40

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	259/300 (86%)	254 (98%)	4 (2%)	1 (0%)	34	24
1	B	260/300 (87%)	254 (98%)	5 (2%)	1 (0%)	34	24
1	C	259/300 (86%)	254 (98%)	4 (2%)	1 (0%)	34	24
1	D	259/300 (86%)	253 (98%)	5 (2%)	1 (0%)	34	24
All	All	1037/1200 (86%)	1015 (98%)	18 (2%)	4 (0%)	34	24

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	203	VAL
1	B	203	VAL
1	C	203	VAL
1	D	203	VAL

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	213/242 (88%)	212 (100%)	1 (0%)	88	88
1	B	214/242 (88%)	212 (99%)	2 (1%)	78	75
1	C	213/242 (88%)	211 (99%)	2 (1%)	78	75
1	D	213/242 (88%)	211 (99%)	2 (1%)	78	75
All	All	853/968 (88%)	846 (99%)	7 (1%)	81	78

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	230	PHE
1	B	230	PHE
1	B	255	HIS
1	C	230	PHE
1	C	255	HIS
1	D	230	PHE
1	D	255	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	61	HIS
1	C	61	HIS
1	C	325	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 (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

9 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	NDG	D	401	-	15,15,15	0.86	0	21,21,21	1.26	3 (14%)
4	NAG	B	401	1	14,14,15	0.77	1 (7%)	17,19,21	0.91	1 (5%)
3	M6P	A	402	-	16,16,16	1.21	1 (6%)	24,24,24	1.28	4 (16%)
3	M6P	C	402	-	16,16,16	1.13	1 (6%)	24,24,24	1.76	5 (20%)
2	NDG	C	401	-	15,15,15	0.75	0	21,21,21	1.22	3 (14%)
2	NDG	B	402	-	15,15,15	0.83	1 (6%)	21,21,21	1.26	3 (14%)
3	M6P	D	402	-	16,16,16	1.18	1 (6%)	24,24,24	0.97	2 (8%)
3	M6P	B	403	-	16,16,16	1.20	1 (6%)	24,24,24	1.67	4 (16%)
2	NDG	A	401	-	15,15,15	1.03	1 (6%)	21,21,21	2.12	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	NDG	D	401	-	-	1/6/26/26	0/1/1/1
4	NAG	B	401	1	-	4/6/23/26	0/1/1/1
3	M6P	A	402	-	-	0/6/26/26	0/1/1/1
3	M6P	C	402	-	-	1/6/26/26	0/1/1/1
2	NDG	C	401	-	-	0/6/26/26	0/1/1/1
2	NDG	B	402	-	-	0/6/26/26	0/1/1/1
3	M6P	D	402	-	-	1/6/26/26	0/1/1/1
3	M6P	B	403	-	-	2/6/26/26	0/1/1/1
2	NDG	A	401	-	-	0/6/26/26	0/1/1/1

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	402	M6P	O5-C1	3.53	1.51	1.42
3	B	403	M6P	O5-C1	3.41	1.51	1.42
3	D	402	M6P	O5-C1	3.15	1.50	1.42
2	A	401	NDG	C1-C2	3.12	1.56	1.52
3	C	402	M6P	O5-C1	2.85	1.50	1.42
4	B	401	NAG	C1-C2	2.38	1.55	1.52
2	B	402	NDG	C3-C2	2.11	1.57	1.53

All (29) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	401	NDG	C1-C2-N2	7.28	119.16	110.73
3	B	403	M6P	O5-C1-C2	4.82	118.88	110.28
3	C	402	M6P	C4-C3-C2	4.78	119.16	110.82
3	C	402	M6P	C3-C4-C5	3.80	117.02	110.24
2	D	401	NDG	O5-C1-C2	-3.66	105.83	109.52
3	B	403	M6P	C1-C2-C3	3.45	117.48	110.31
2	A	401	NDG	C3-C2-N2	-3.43	104.13	110.62
3	B	403	M6P	C3-C4-C5	-3.24	104.46	110.24
2	C	401	NDG	O5-C1-C2	-3.23	106.27	109.52
3	A	402	M6P	O5-C1-C2	3.12	115.86	110.28
3	A	402	M6P	C3-C4-C5	-3.10	104.71	110.24
3	C	402	M6P	C1-O5-C5	-3.01	107.98	113.66
4	B	401	NAG	C1-O5-C5	2.84	116.04	112.19
2	D	401	NDG	C1-C2-C3	-2.70	106.86	110.54
3	C	402	M6P	O5-C5-C4	2.68	114.55	109.69
2	A	401	NDG	O4-C4-C3	-2.67	104.17	110.35
2	C	401	NDG	C1-C2-C3	-2.66	106.92	110.54
3	D	402	M6P	O5-C5-C4	2.59	114.39	109.69
2	B	402	NDG	C1-C2-C3	-2.57	107.04	110.54
3	A	402	M6P	C1-C2-C3	2.55	115.60	110.31
2	D	401	NDG	C3-C2-N2	2.36	115.07	110.62
3	B	403	M6P	O6-C6-C5	-2.36	100.88	108.99
2	C	401	NDG	C3-C2-N2	2.22	114.81	110.62
3	C	402	M6P	O3-C3-C4	-2.20	105.27	110.35
2	B	402	NDG	C1-O5-C5	2.15	117.72	113.66
2	B	402	NDG	O3-C3-C2	2.13	113.95	109.66
2	A	401	NDG	O3-C3-C2	-2.05	105.52	109.66
3	A	402	M6P	O6-C6-C5	-2.03	102.01	108.99
3	D	402	M6P	O6-P-O1P	2.02	112.15	106.47

There are no chirality outliers.

All (9) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	403	M6P	C4-C5-C6-O6
3	B	403	M6P	O5-C5-C6-O6
4	B	401	NAG	C3-C2-N2-C7
4	B	401	NAG	O5-C5-C6-O6
4	B	401	NAG	C4-C5-C6-O6
3	C	402	M6P	O5-C5-C6-O6
4	B	401	NAG	C1-C2-N2-C7
2	D	401	NDG	C4-C5-C6-O6
3	D	402	M6P	C4-C5-C6-O6

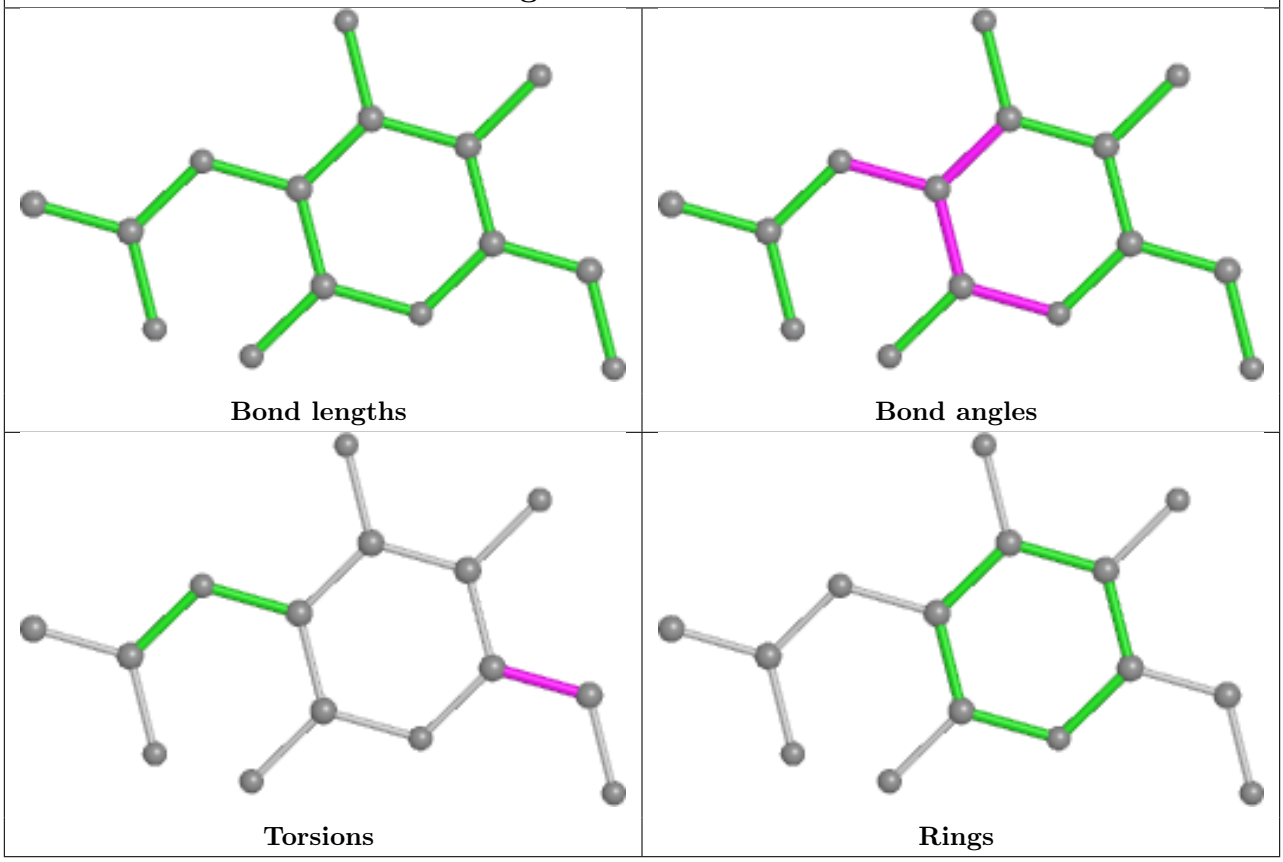
There are no ring outliers.

2 monomers are involved in 2 short contacts:

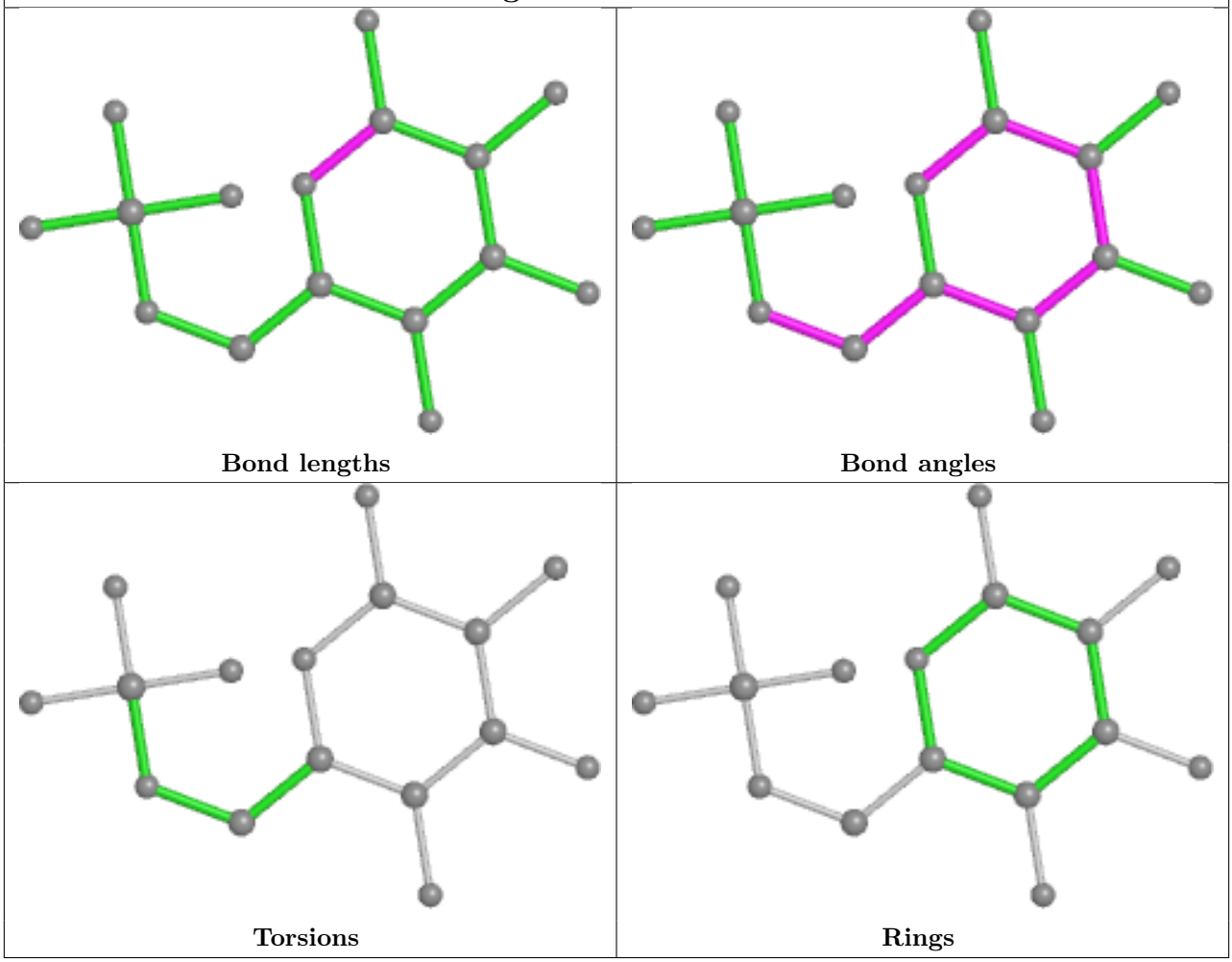
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	401	NAG	1	0
3	D	402	M6P	1	0

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

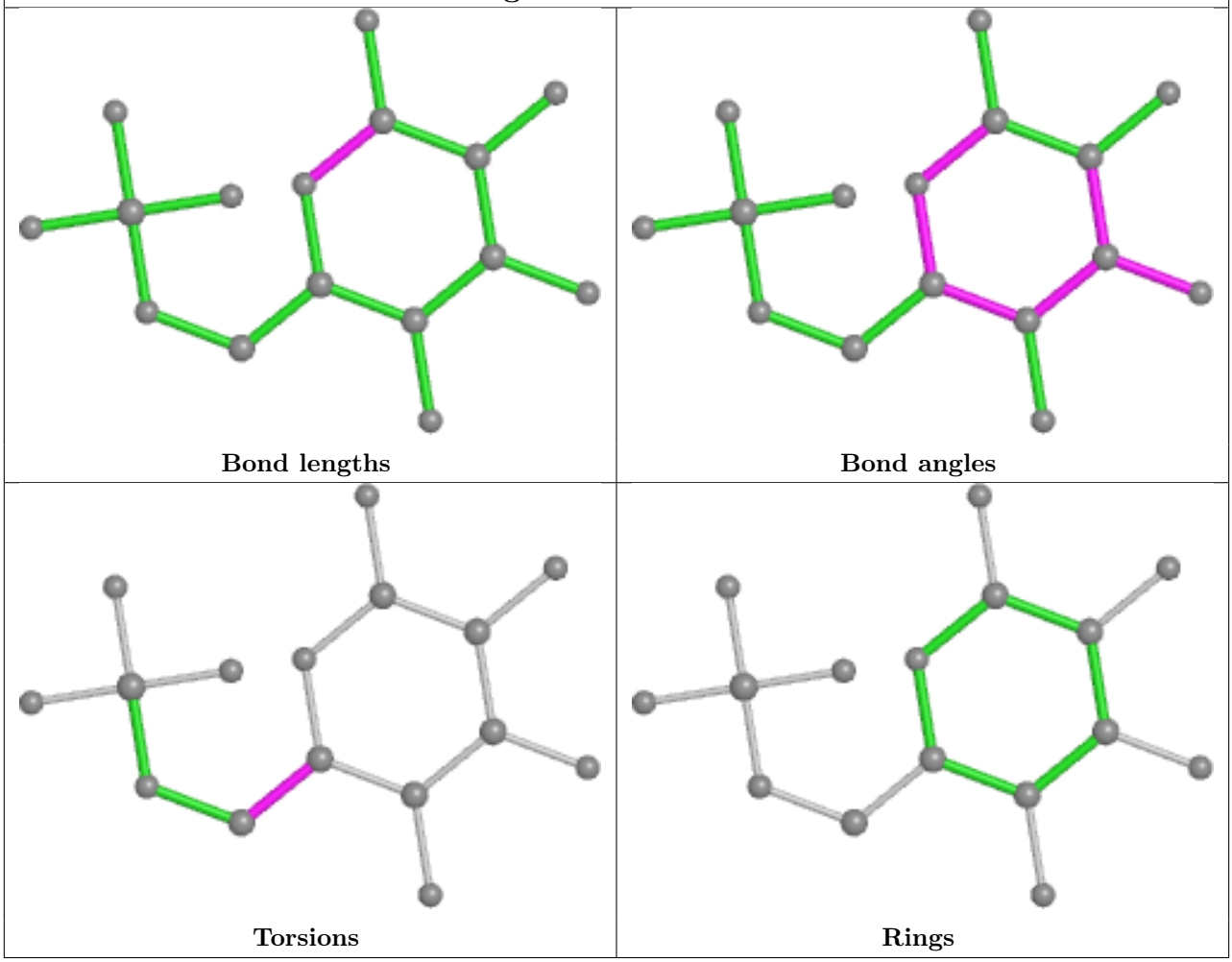
Ligand NDG D 401



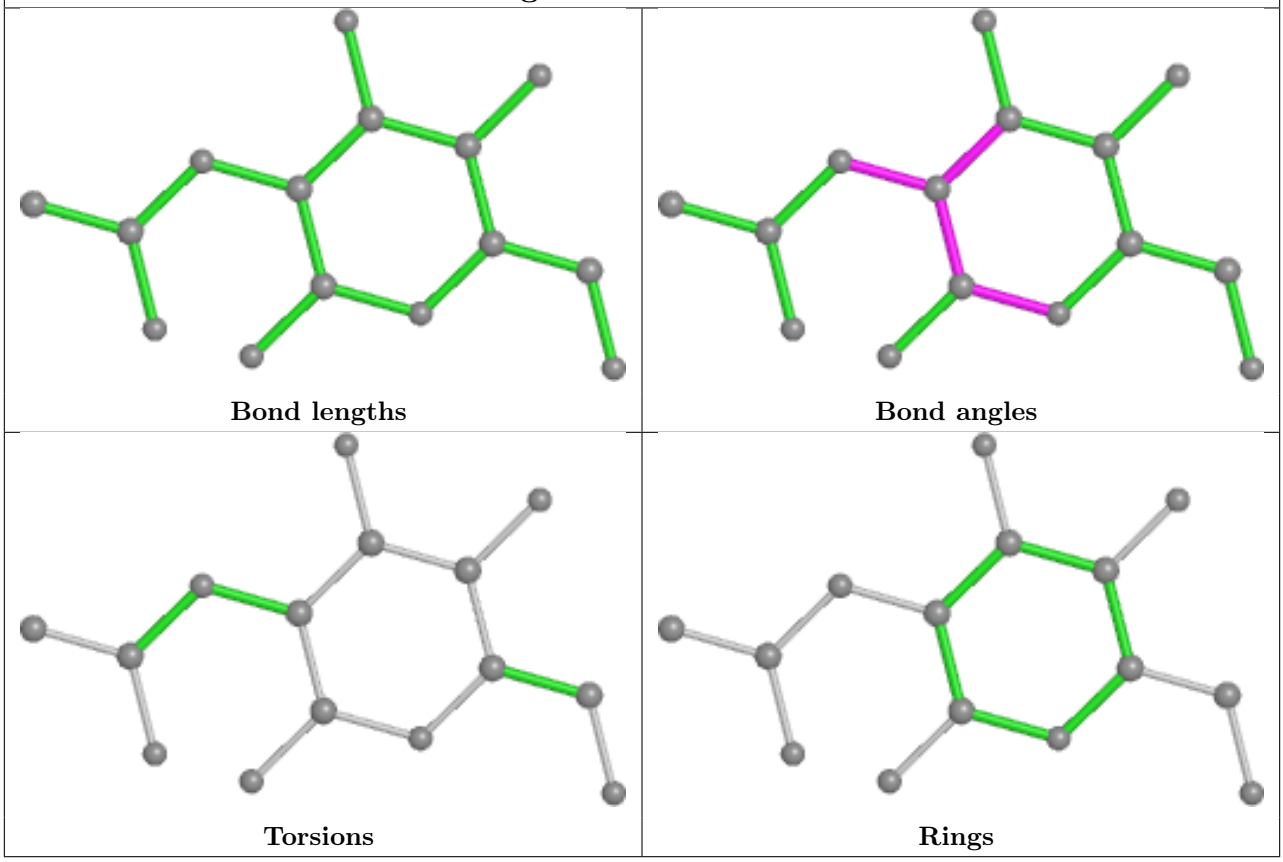
Ligand M6P A 402



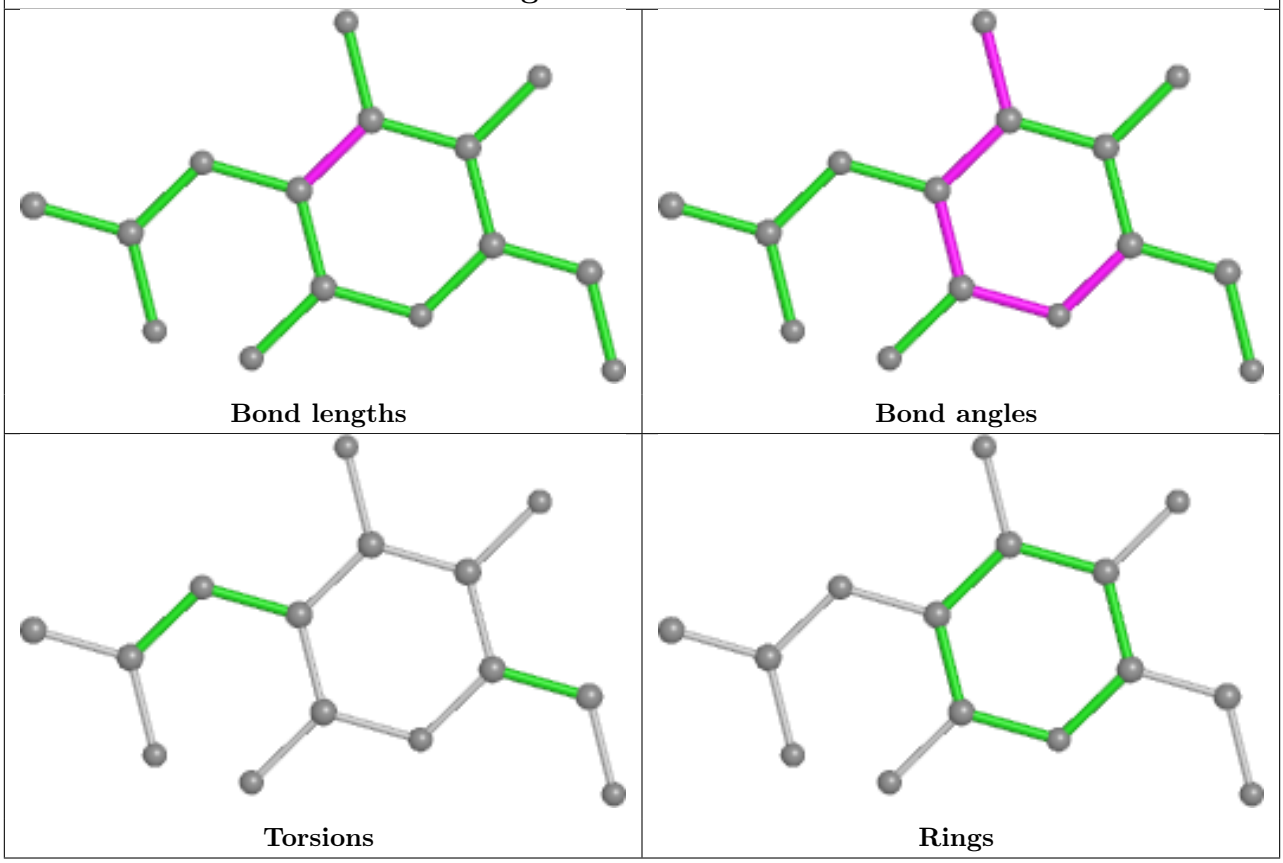
Ligand M6P C 402



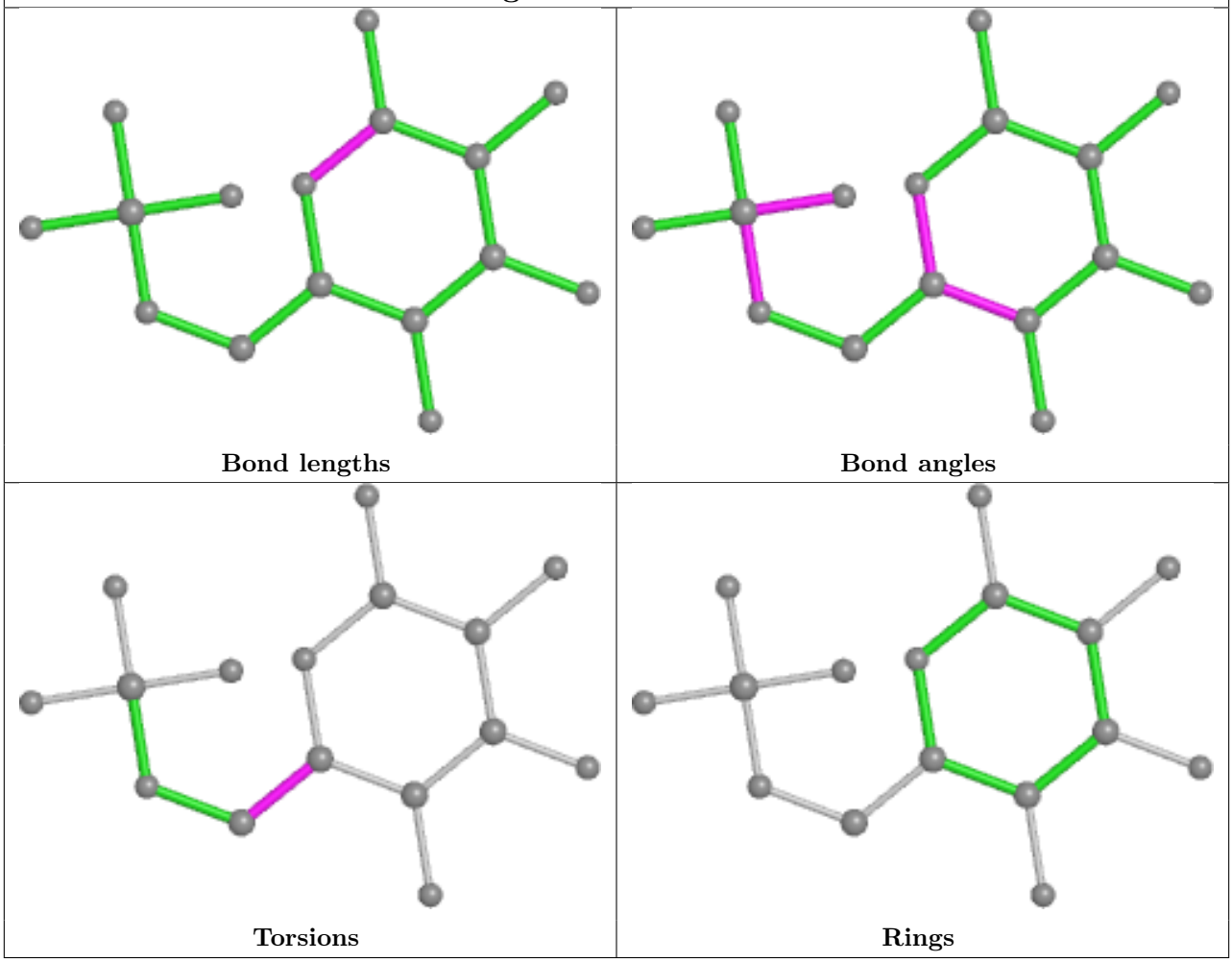
Ligand NDG C 401



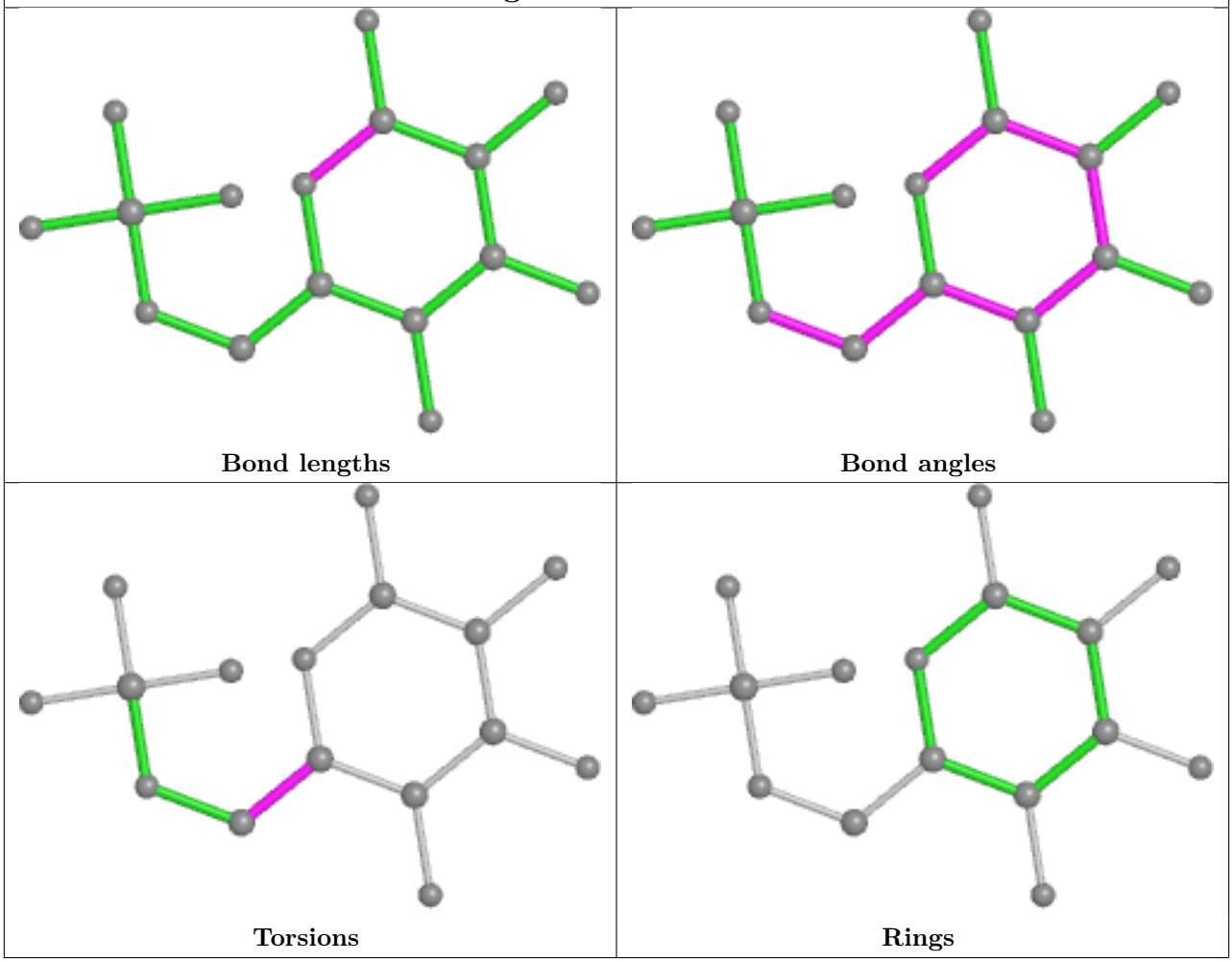
Ligand NDG B 402



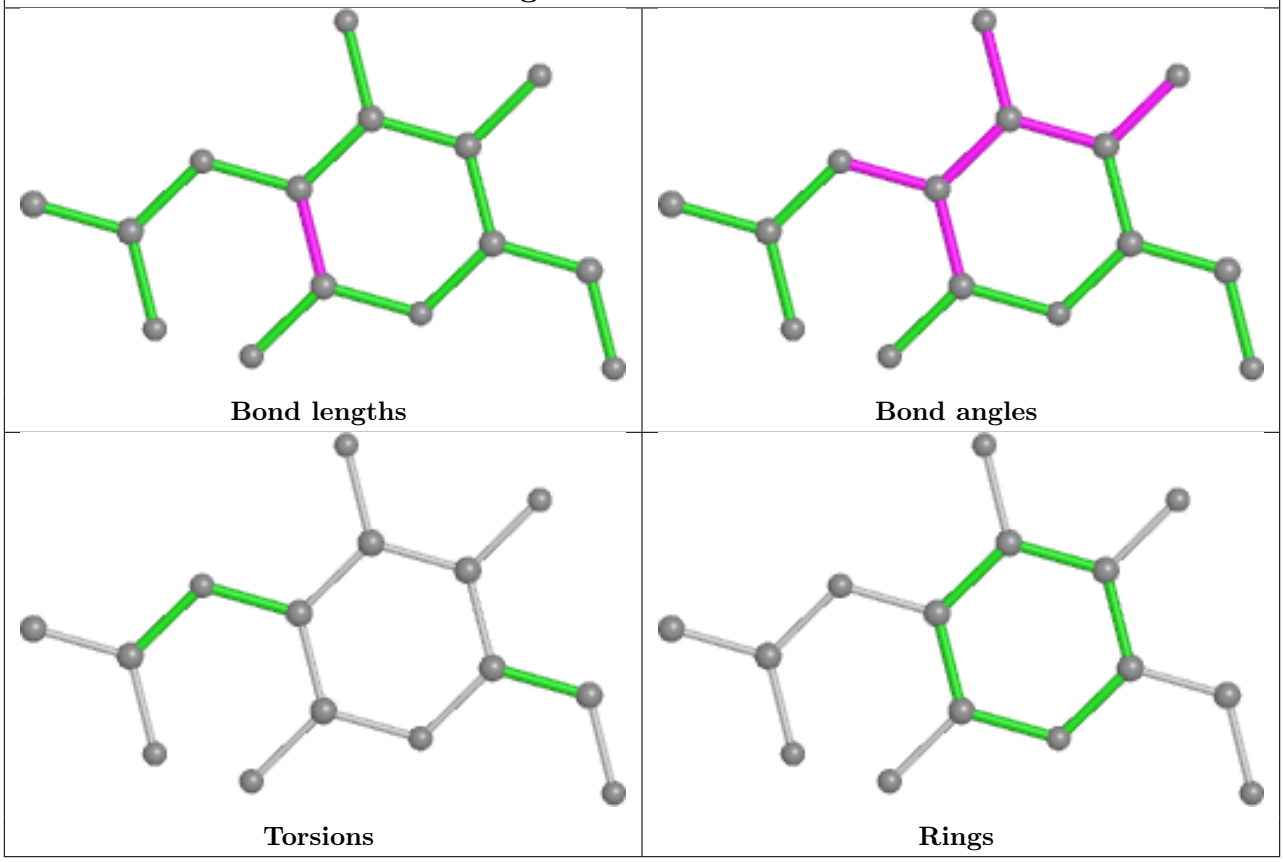
Ligand M6P D 402



Ligand M6P B 403



Ligand NDG A 401



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	263/300 (87%)	-0.09	1 (0%)	92	95	15, 24, 41, 55	0
1	B	264/300 (88%)	-0.05	3 (1%)	80	84	16, 24, 42, 64	0
1	C	263/300 (87%)	-0.06	2 (0%)	86	89	15, 26, 44, 57	0
1	D	263/300 (87%)	0.05	4 (1%)	73	79	16, 28, 47, 64	0
All	All	1053/1200 (87%)	-0.04	10 (0%)	84	87	15, 26, 44, 64	0

All (10) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	74	GLY	6.9
1	D	59	LEU	6.4
1	B	59	LEU	4.2
1	D	327	PRO	3.4
1	A	59	LEU	2.8
1	C	327	PRO	2.8
1	C	74	GLY	2.2
1	B	68	PRO	2.2
1	B	148	CYS	2.2
1	D	230	PHE	2.1

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

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

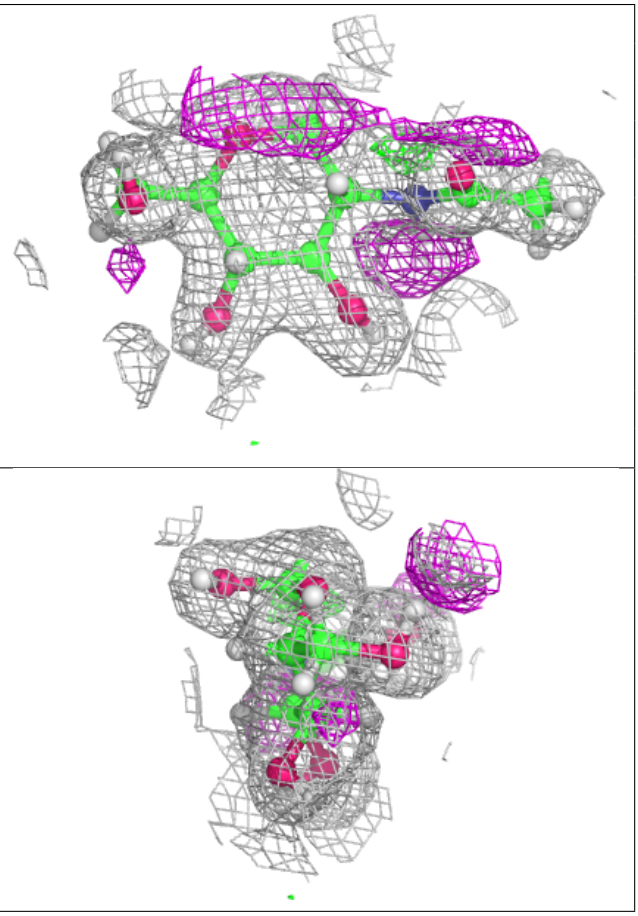
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	NAG	B	401	14/15	0.71	0.15	48,65,81,89	0
2	NDG	A	401	15/15	0.88	0.11	19,29,34,36	0
2	NDG	C	401	15/15	0.93	0.10	18,27,32,34	0
3	M6P	C	402	16/16	0.94	0.18	20,66,89,95	0
2	NDG	B	402	15/15	0.94	0.10	22,29,37,41	0
3	M6P	B	403	16/16	0.95	0.19	23,78,102,109	0
3	M6P	A	402	16/16	0.96	0.17	22,83,103,117	0
3	M6P	D	402	16/16	0.96	0.17	21,67,97,102	0
2	NDG	D	401	15/15	0.96	0.09	22,27,31,35	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.

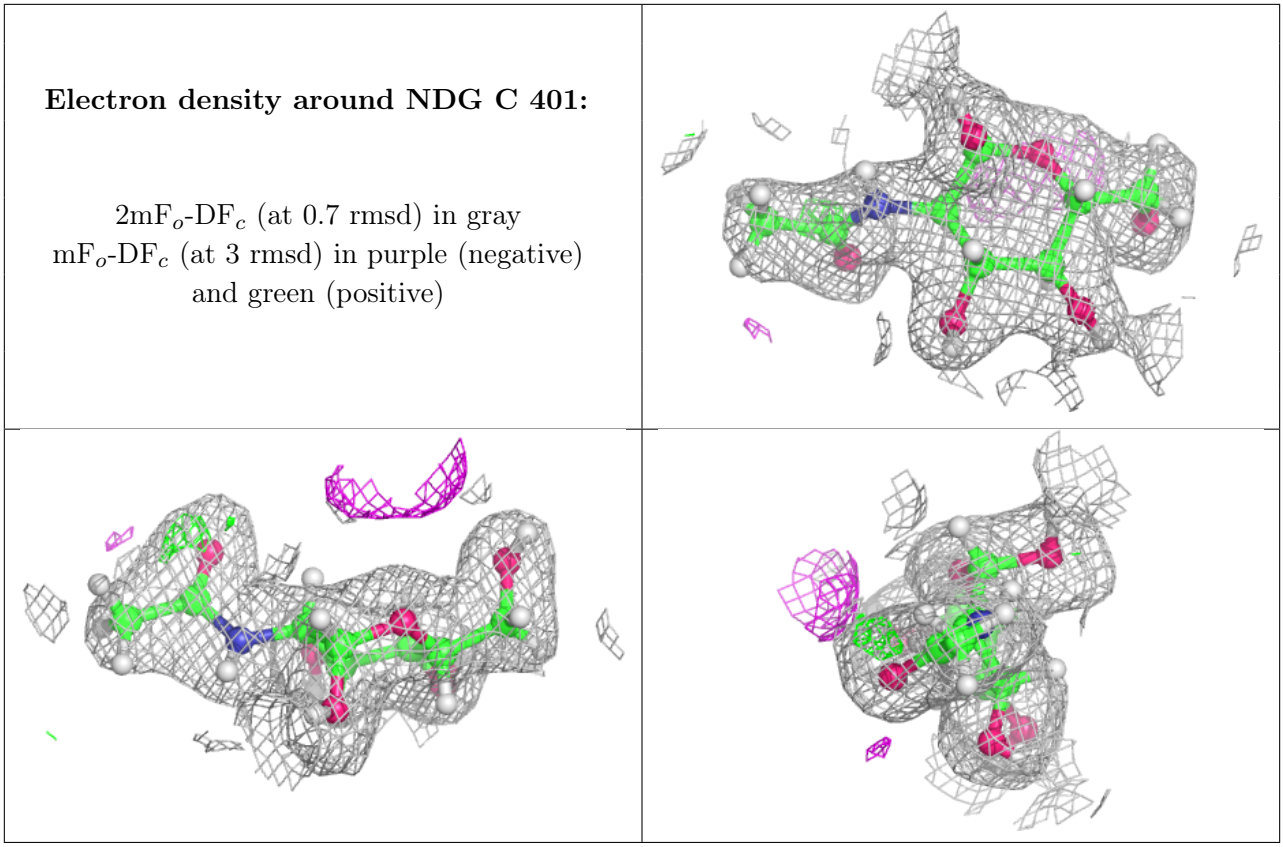
Electron density around NDG A 401:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)



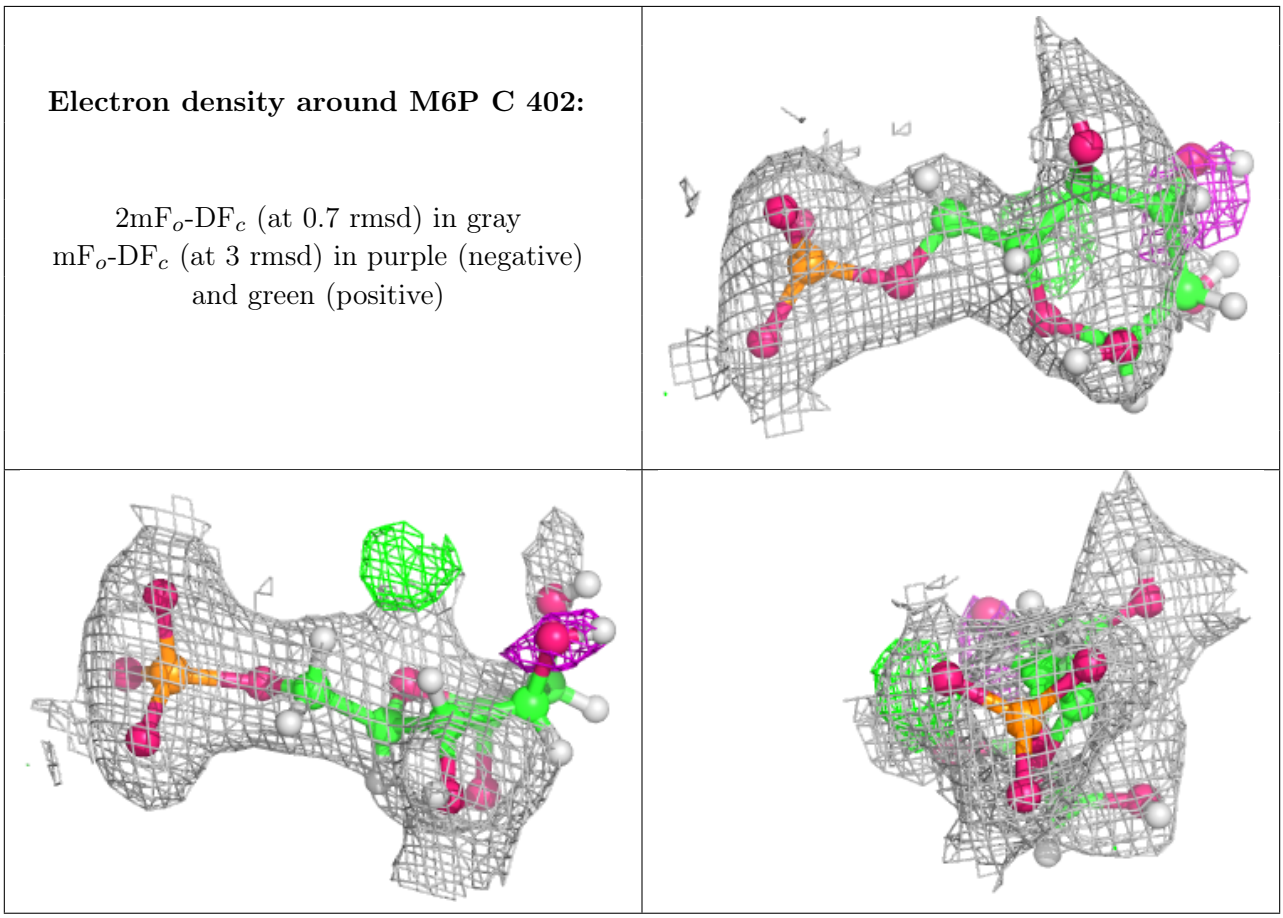
Electron density around NDG C 401:

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



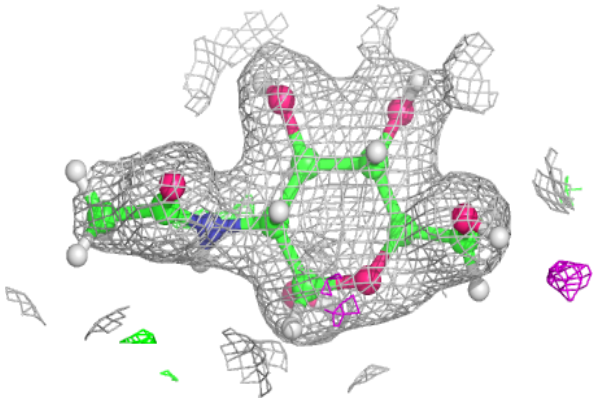
Electron density around M6P C 402:

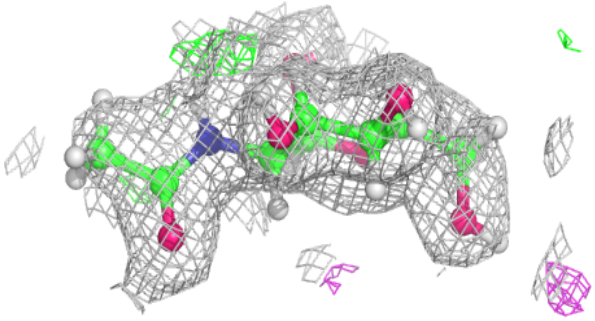
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

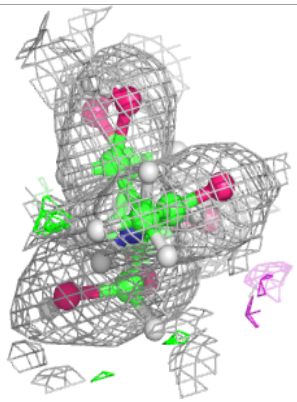


Electron density around NDG B 402:

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

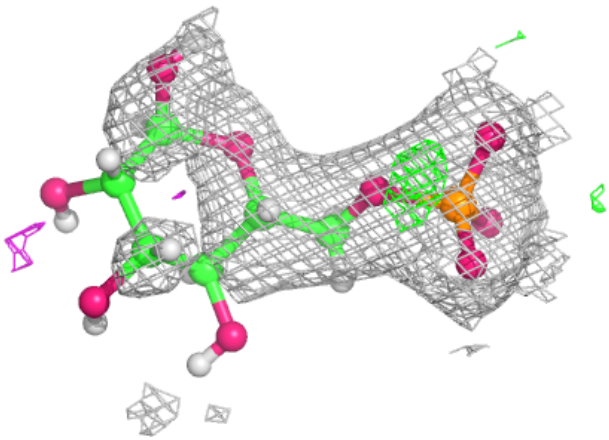


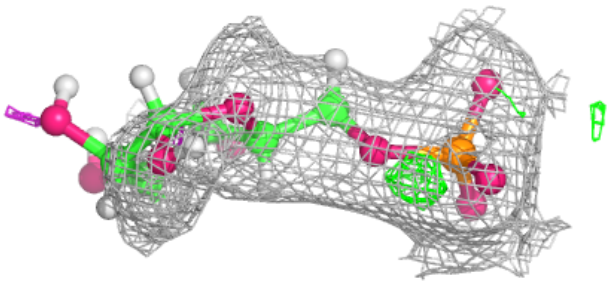


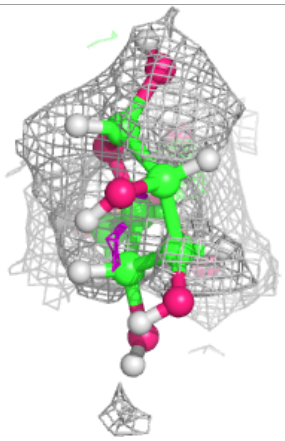


Electron density around M6P B 403:

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

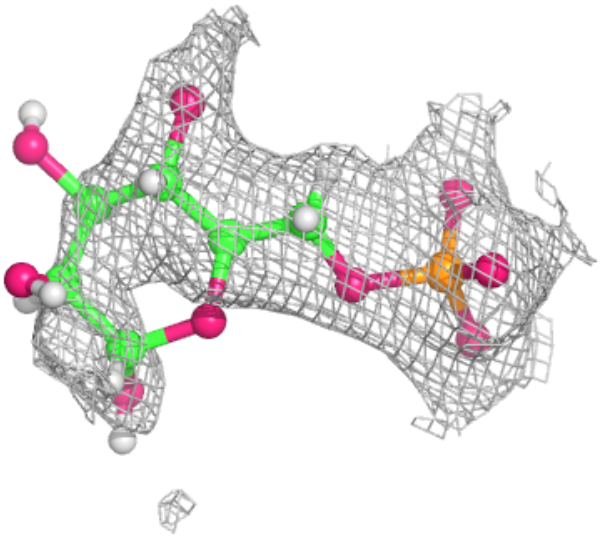


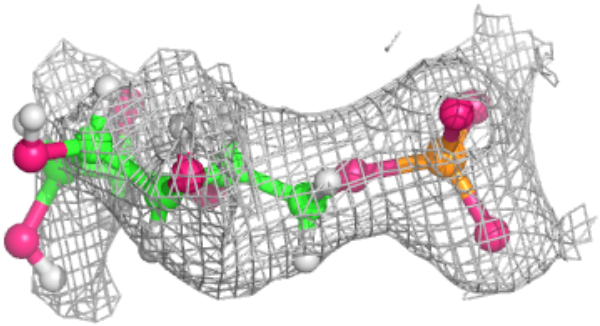


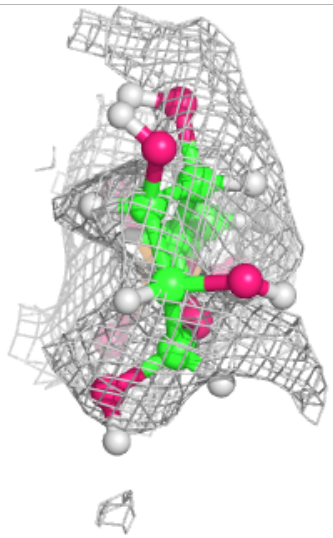


Electron density around M6P A 402:

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

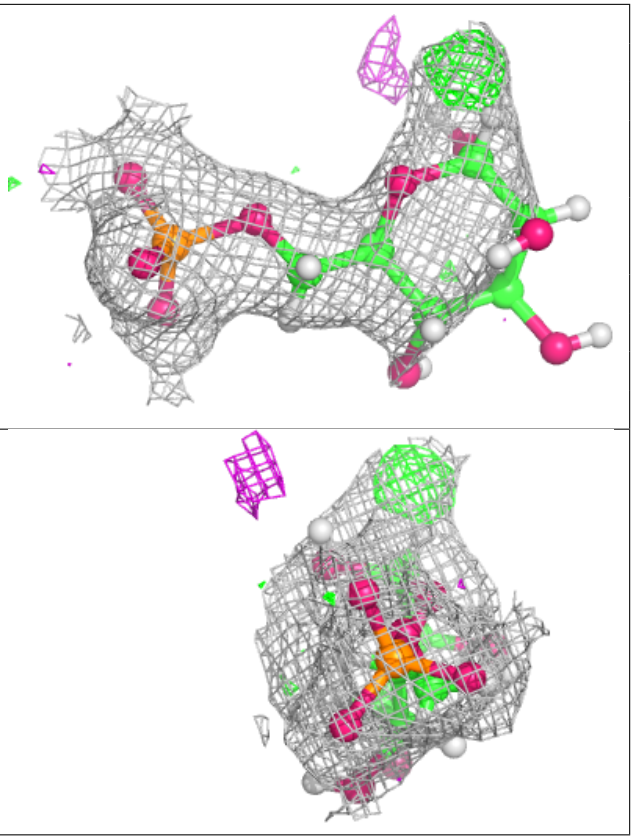






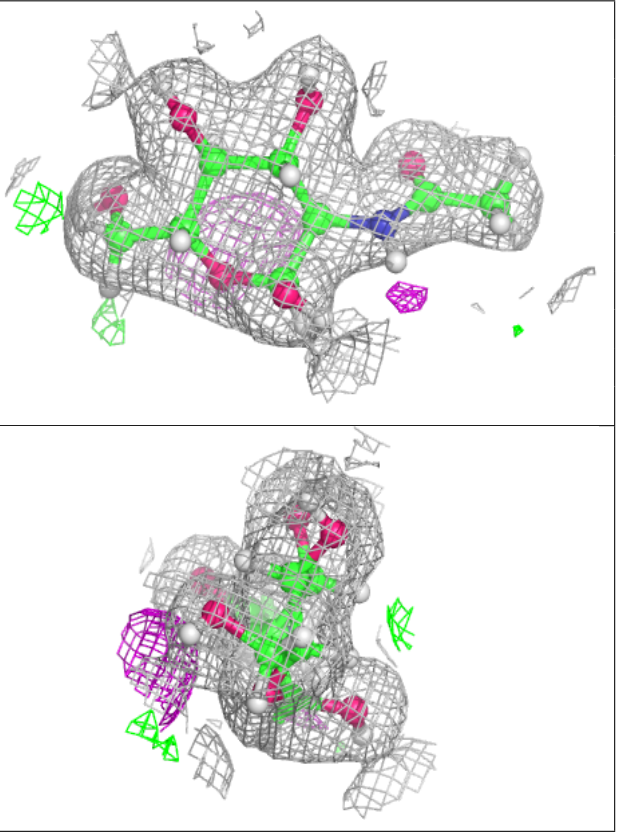
Electron density around M6P D 402:

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



Electron density around NDG 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 (i)

There are no such residues in this entry.