



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 4, 2024 – 05:08 AM EST

PDB ID : 4F4O  
Title : Structure of the Haptoglobin-Haemoglobin Complex  
Authors : Andersen, C.B.F.; Torvund-Jensen, M.; Nielsen, M.J.; Oliveira, C.L.P.; Hersleth, H.P.; Andersen, N.H.; Pedersen, J.S.; Andersen, G.R.; Moestrup, S.K.  
Deposited on : 2012-05-11  
Resolution : 2.90 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](mailto: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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtrriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

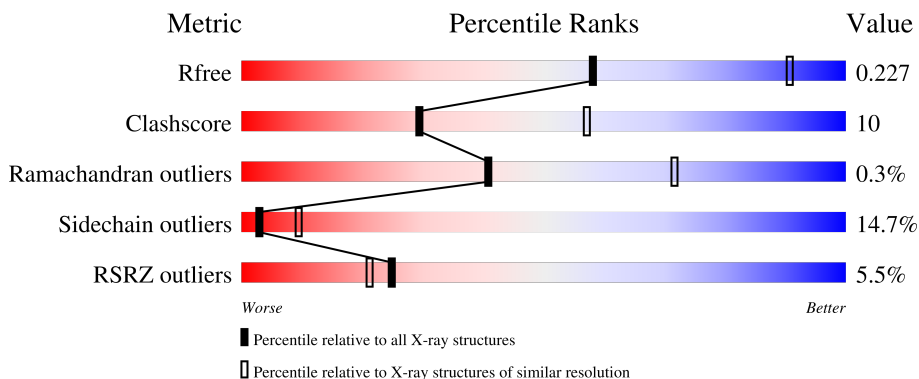
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.90 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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  $\geq 3$ , 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  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	141	
1	D	141	
1	G	141	
1	J	141	
2	B	146	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	E	146	
2	H	146	
2	K	146	
3	C	347	
3	F	347	
3	I	347	
3	L	347	
4	M	2	
4	O	2	
4	Q	2	
5	N	2	
5	P	2	
5	R	2	
5	S	2	

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
5	NAG	P	1	-	-	-	X
5	FUC	P	2	-	-	-	X
8	NAG	C	1003	-	-	-	X
8	NAG	F	1003	-	-	-	X
8	NAG	I	1003	-	-	-	X

## 2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 19196 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 Hemoglobin subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	141	1064	677	192	193	2	0	0	0
1	D	141	1064	677	192	193	2	0	0	0
1	G	141	1064	677	192	193	2	0	0	0
1	J	141	1064	677	192	193	2	0	0	0

- Molecule 2 is a protein called Hemoglobin subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	146	1144	731	205	206	2	0	1	0
2	E	146	1144	731	205	206	2	0	1	0
2	H	146	1144	731	205	206	2	0	1	0
2	K	146	1144	731	205	206	2	0	1	0

- Molecule 3 is a protein called Haptoglobin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	309	2428	1546	411	457	14	0	0	0
3	F	309	2428	1546	411	457	14	0	0	0
3	I	309	2428	1546	411	457	14	0	0	0
3	L	309	2428	1546	411	457	14	0	0	0

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



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
4	M	2	28	16	2	10	0	0	0
4	O	2	28	16	2	10	0	0	0
4	Q	2	28	16	2	10	0	0	0

- Molecule 5 is an oligosaccharide called alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose.



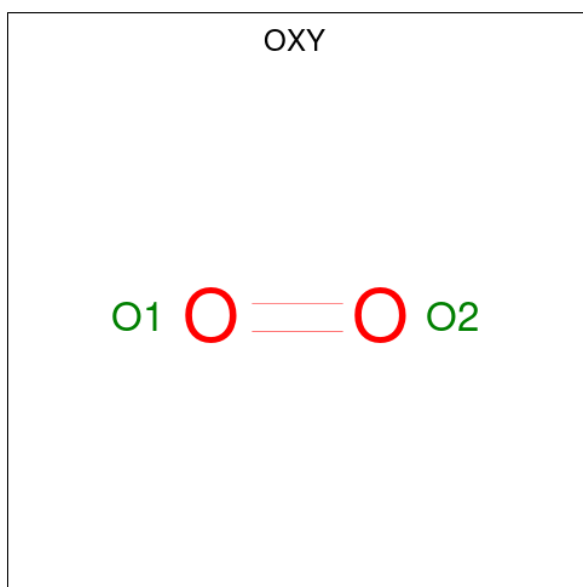
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
5	N	2	24	14	1	9	0	0	0
5	P	2	24	14	1	9	0	0	0
5	R	2	24	14	1	9	0	0	0
5	S	2	24	14	1	9	0	0	0

- Molecule 6 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: C<sub>34</sub>H<sub>32</sub>FeN<sub>4</sub>O<sub>4</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
			Total	C	Fe	N			O
6	A	1	43	34	1	4	4	0	0
6	B	1	43	34	1	4	4	0	0
6	D	1	43	34	1	4	4	0	0
6	E	1	43	34	1	4	4	0	0
6	G	1	43	34	1	4	4	0	0
6	H	1	43	34	1	4	4	0	0
6	J	1	43	34	1	4	4	0	0
6	K	1	43	34	1	4	4	0	0

- Molecule 7 is OXYGEN MOLECULE (three-letter code: OXY) (formula: O<sub>2</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total O 2 2	0	0
7	B	1	Total O 2 2	0	0
7	D	1	Total O 2 2	0	0
7	E	1	Total O 2 2	0	0
7	G	1	Total O 2 2	0	0
7	H	1	Total O 2 2	0	0
7	J	1	Total O 2 2	0	0
7	K	1	Total O 2 2	0	0

- Molecule 8 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).




Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
8	C	1	Total 14	C 8	N 1	O 5	0	0
8	C	1	Total 14	C 8	N 1	O 5	0	0
8	F	1	Total 14	C 8	N 1	O 5	0	0
8	F	1	Total 14	C 8	N 1	O 5	0	0
8	I	1	Total 14	C 8	N 1	O 5	0	0
8	I	1	Total 14	C 8	N 1	O 5	0	0
8	L	1	Total 14	C 8	N 1	O 5	0	0
8	L	1	Total 14	C 8	N 1	O 5	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.

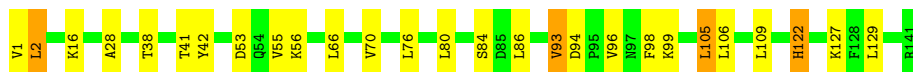
- Molecule 1: Hemoglobin subunit alpha

Chain A:  82% 15%




- Molecule 1: Hemoglobin subunit alpha

Chain D:  81% 16%




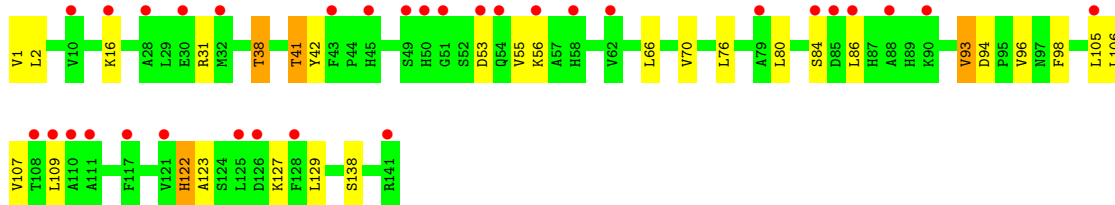
- Molecule 1: Hemoglobin subunit alpha

Chain G:  84% 13%



- Molecule 1: Hemoglobin subunit alpha

Chain J:  23% 79% 18%



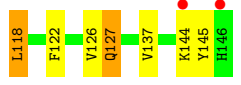
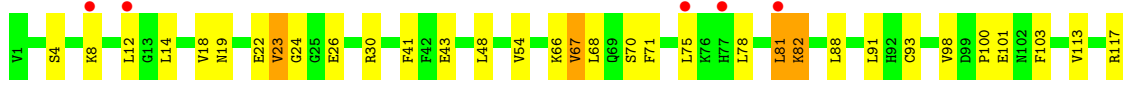
- Molecule 2: Hemoglobin subunit beta

Chain B:  75% 22%

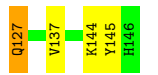
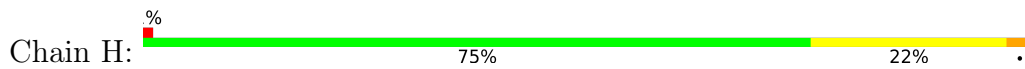




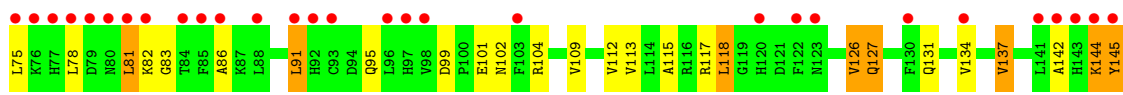
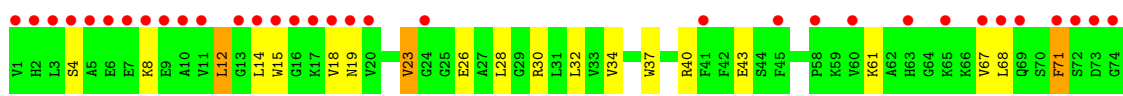
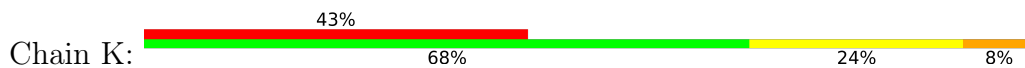
• Molecule 2: Hemoglobin subunit beta



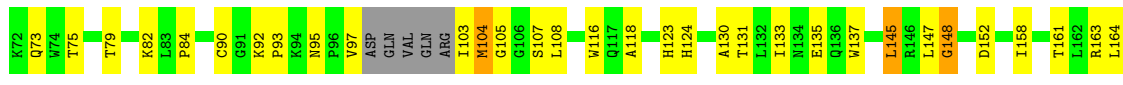
• Molecule 2: Hemoglobin subunit beta

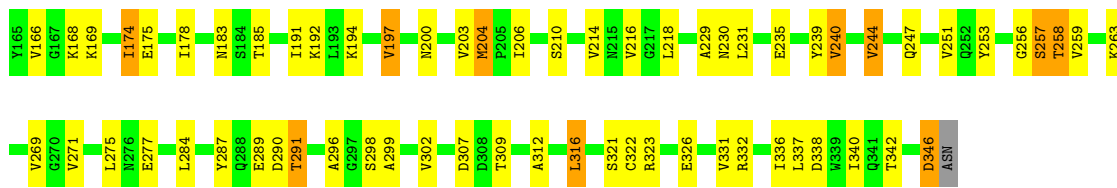


• Molecule 2: Hemoglobin subunit beta

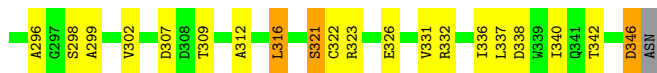
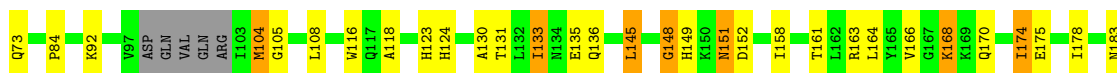
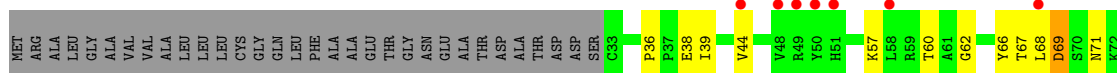


• Molecule 3: Haptoglobin

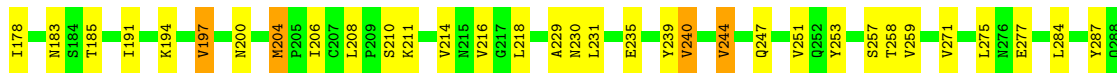
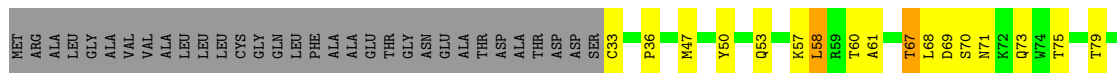




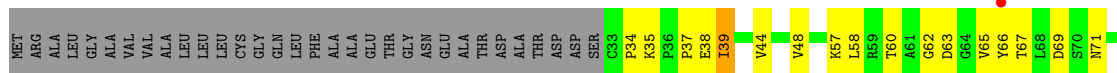
• Molecule 3: Haptoglobin

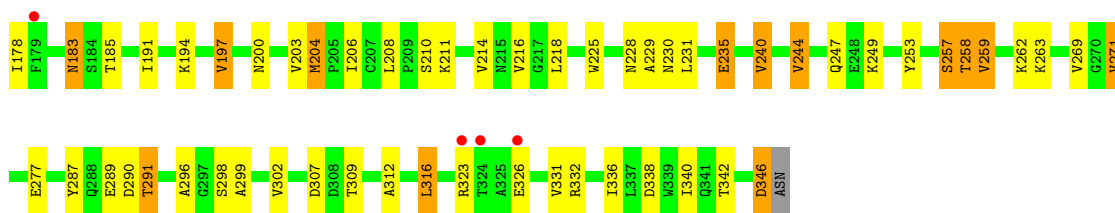


• Molecule 3: Haptoglobin



• Molecule 3: Haptoglobin





- Molecule 4: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain M: 100%

MAG1  
MAG2

- Molecule 4: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain O: 100%

MAG1  
MAG2

- Molecule 4: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain Q: 100%

MAG1  
MAG2

- Molecule 5: alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain N: 100%

MAG1  
FUC2

- Molecule 5: alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain P: 50% 50%

MAG1  
FUC2

- Molecule 5: alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain R: 100%

MAG1  
FUC2

- Molecule 5: alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain S:

100%

MGI  
FUC2

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	72.88Å 197.78Å 322.07Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.90 49.44 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.0 (20.00-2.90) 99.0 (49.44-2.90)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.09 (at 2.91Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.7.2_869)	Depositor
R, $R_{free}$	0.211 , 0.229 0.209 , 0.227	Depositor DCC
$R_{free}$ test set	2251 reflections (2.18%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	75.0	Xtrriage
Anisotropy	0.173	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 50.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	19196	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	93.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.51% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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: HEM, FUC, OXY, 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	0.59	0/1091	0.70	0/1480
1	D	0.61	0/1091	0.68	0/1480
1	G	0.56	0/1091	0.67	0/1480
1	J	0.35	0/1091	0.58	0/1480
2	B	0.57	0/1169	0.68	0/1580
2	E	0.48	0/1169	0.65	0/1580
2	H	0.63	0/1169	0.69	0/1580
2	K	0.44	0/1169	0.55	0/1580
3	C	0.57	0/2487	0.79	1/3377 (0.0%)
3	F	0.61	1/2487 (0.0%)	0.81	3/3377 (0.1%)
3	I	0.62	0/2487	0.81	3/3377 (0.1%)
3	L	0.51	0/2487	0.75	1/3377 (0.0%)
All	All	0.56	1/18988 (0.0%)	0.73	8/25748 (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
3	C	0	2
3	F	0	2
3	I	0	2
3	L	0	3
All	All	0	9

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	F	151	ASN	C-N	-5.51	1.21	1.34

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	I	346	ASP	CB-CG-OD2	5.56	123.30	118.30
3	F	322	CYS	CA-CB-SG	-5.23	104.59	114.00
3	I	136	GLN	N-CA-C	5.18	125.00	111.00
3	F	322	CYS	C-N-CA	-5.15	108.83	121.70
3	I	322	CYS	CA-CB-SG	-5.09	104.83	114.00

There are no chirality outliers.

5 of 9 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	C	148	GLY	Peptide
3	C	257	SER	Peptide
3	F	148	GLY	Peptide
3	F	257	SER	Peptide
3	I	148	GLY	Peptide

## 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1064	0	1060	12	0
1	D	1064	0	1060	12	0
1	G	1064	0	1060	11	0
1	J	1064	0	1060	14	0
2	B	1144	0	1141	21	0
2	E	1144	0	1141	22	0
2	H	1144	0	1141	21	0
2	K	1144	0	1141	35	0
3	C	2428	0	2396	61	0
3	F	2428	0	2394	54	0
3	I	2428	0	2395	57	0
3	L	2428	0	2397	66	0
4	M	28	0	25	1	0
4	O	28	0	25	1	0
4	Q	28	0	25	1	0
5	N	24	0	22	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	P	24	0	22	2	0
5	R	24	0	22	0	0
5	S	24	0	22	5	0
6	A	43	0	30	1	0
6	B	43	0	30	4	0
6	D	43	0	30	0	0
6	E	43	0	30	5	0
6	G	43	0	30	0	0
6	H	43	0	30	4	0
6	J	43	0	30	0	0
6	K	43	0	30	2	0
7	A	2	0	0	0	0
7	B	2	0	0	0	0
7	D	2	0	0	0	0
7	E	2	0	0	0	0
7	G	2	0	0	0	0
7	H	2	0	0	0	0
7	J	2	0	0	0	0
7	K	2	0	0	0	0
8	C	28	0	26	1	0
8	F	28	0	26	0	0
8	I	28	0	26	0	0
8	L	28	0	26	1	0
All	All	19196	0	18893	362	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 10.

The worst 5 of 362 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:K:101[B]:GLU:HG3	3:L:230:ASN:HD22	1.12	1.12
2:E:101[B]:GLU:HG3	3:F:230:ASN:HD22	1.08	1.11
2:H:101[B]:GLU:HG3	3:I:230:ASN:HD22	1.10	1.08
2:B:101[B]:GLU:HG3	3:C:230:ASN:HD22	1.12	1.07
3:F:92:LYS:H	3:F:204:MET:HE2	1.32	0.95

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	139/141 (99%)	134 (96%)	5 (4%)	0	100	100
1	D	139/141 (99%)	135 (97%)	4 (3%)	0	100	100
1	G	139/141 (99%)	133 (96%)	6 (4%)	0	100	100
1	J	139/141 (99%)	134 (96%)	5 (4%)	0	100	100
2	B	145/146 (99%)	143 (99%)	2 (1%)	0	100	100
2	E	145/146 (99%)	142 (98%)	3 (2%)	0	100	100
2	H	145/146 (99%)	143 (99%)	2 (1%)	0	100	100
2	K	145/146 (99%)	142 (98%)	3 (2%)	0	100	100
3	C	305/347 (88%)	288 (94%)	15 (5%)	2 (1%)	22	54
3	F	305/347 (88%)	284 (93%)	20 (7%)	1 (0%)	41	71
3	I	305/347 (88%)	287 (94%)	15 (5%)	3 (1%)	15	45
3	L	305/347 (88%)	285 (93%)	18 (6%)	2 (1%)	22	54
All	All	2356/2536 (93%)	2250 (96%)	98 (4%)	8 (0%)	41	71

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	F	258	THR
3	L	169	LYS
3	L	258	THR
3	C	169	LYS
3	C	258	THR

### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all 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	111/111 (100%)	97 (87%)	14 (13%)	4	13
1	D	111/111 (100%)	97 (87%)	14 (13%)	4	13
1	G	111/111 (100%)	97 (87%)	14 (13%)	4	13
1	J	111/111 (100%)	97 (87%)	14 (13%)	4	13
2	B	120/119 (101%)	101 (84%)	19 (16%)	2	8
2	E	120/119 (101%)	102 (85%)	18 (15%)	3	9
2	H	120/119 (101%)	103 (86%)	17 (14%)	3	10
2	K	120/119 (101%)	100 (83%)	20 (17%)	2	6
3	C	268/296 (90%)	225 (84%)	43 (16%)	2	7
3	F	268/296 (90%)	230 (86%)	38 (14%)	3	10
3	I	268/296 (90%)	225 (84%)	43 (16%)	2	7
3	L	268/296 (90%)	229 (85%)	39 (15%)	3	9
All	All	1996/2104 (95%)	1703 (85%)	293 (15%)	3	9

5 of 293 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	J	109	LEU
3	L	302	VAL
2	K	61	LYS
3	L	104	MET
2	E	75	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
3	L	230	ASN
1	J	122	HIS
1	G	122	HIS
3	F	230	ASN
3	I	230	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](#)

14 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
4	NAG	M	1	3,4	14,14,15	2.42	6 (42%)	17,19,21	3.15	10 (58%)
4	NAG	M	2	4	14,14,15	2.44	7 (50%)	17,19,21	1.63	3 (17%)
5	NAG	N	1	5,3	14,14,15	2.11	5 (35%)	17,19,21	2.35	6 (35%)
5	FUC	N	2	5	10,10,11	2.70	4 (40%)	14,14,16	2.60	4 (28%)
4	NAG	O	1	3,4	14,14,15	2.50	6 (42%)	17,19,21	3.64	13 (76%)
4	NAG	O	2	4	14,14,15	2.10	6 (42%)	17,19,21	1.49	3 (17%)
5	NAG	P	1	5,3	14,14,15	2.06	6 (42%)	17,19,21	2.30	6 (35%)
5	FUC	P	2	5	10,10,11	2.41	5 (50%)	14,14,16	2.27	4 (28%)
4	NAG	Q	1	3,4	14,14,15	2.19	6 (42%)	17,19,21	3.46	9 (52%)
4	NAG	Q	2	4	14,14,15	2.30	7 (50%)	17,19,21	1.60	2 (11%)
5	NAG	R	1	5,3	14,14,15	2.08	5 (35%)	17,19,21	2.14	5 (29%)
5	FUC	R	2	5	10,10,11	2.73	5 (50%)	14,14,16	2.86	5 (35%)
5	NAG	S	1	5,3	14,14,15	0.41	0	17,19,21	1.40	2 (11%)
5	FUC	S	2	5	10,10,11	0.47	0	14,14,16	0.96	1 (7%)

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
4	NAG	M	1	3,4	-	3/6/23/26	0/1/1/1
4	NAG	M	2	4	-	0/6/23/26	0/1/1/1

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	NAG	N	1	5,3	-	0/6/23/26	0/1/1/1
5	FUC	N	2	5	-	-	0/1/1/1
4	NAG	O	1	3,4	-	3/6/23/26	0/1/1/1
4	NAG	O	2	4	-	0/6/23/26	0/1/1/1
5	NAG	P	1	5,3	-	1/6/23/26	0/1/1/1
5	FUC	P	2	5	-	-	0/1/1/1
4	NAG	Q	1	3,4	-	3/6/23/26	0/1/1/1
4	NAG	Q	2	4	-	0/6/23/26	0/1/1/1
5	NAG	R	1	5,3	-	1/6/23/26	0/1/1/1
5	FUC	R	2	5	-	-	0/1/1/1
5	NAG	S	1	5,3	-	0/6/23/26	0/1/1/1
5	FUC	S	2	5	-	-	0/1/1/1

The worst 5 of 68 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	R	2	FUC	O5-C5	5.83	1.56	1.43
5	N	2	FUC	O5-C5	5.68	1.55	1.43
4	O	1	NAG	C2-N2	4.87	1.54	1.46
5	N	1	NAG	O5-C5	4.68	1.52	1.43
4	O	1	NAG	C7-N2	4.67	1.50	1.34

The worst 5 of 73 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	Q	1	NAG	O5-C1-C2	8.53	124.76	111.29
4	O	1	NAG	O5-C1-C2	8.04	123.98	111.29
5	R	2	FUC	O5-C5-C6	6.71	121.78	107.33
5	N	2	FUC	O5-C5-C6	6.37	121.05	107.33
4	M	1	NAG	O5-C1-C2	6.12	120.95	111.29

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

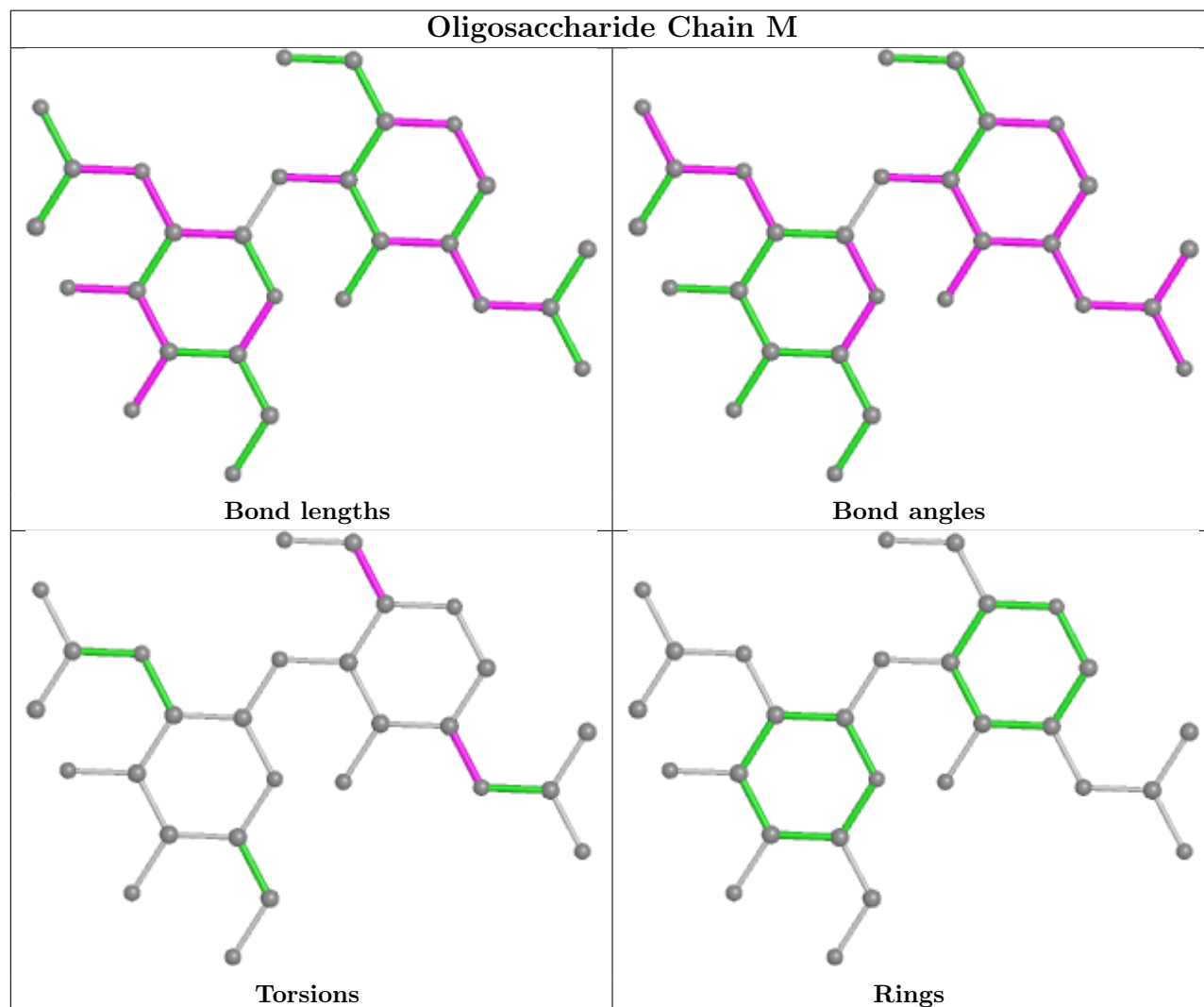
Mol	Chain	Res	Type	Atoms
4	M	1	NAG	C3-C2-N2-C7
4	O	1	NAG	C3-C2-N2-C7
4	Q	1	NAG	C3-C2-N2-C7
4	O	1	NAG	O5-C5-C6-O6
4	Q	1	NAG	O5-C5-C6-O6

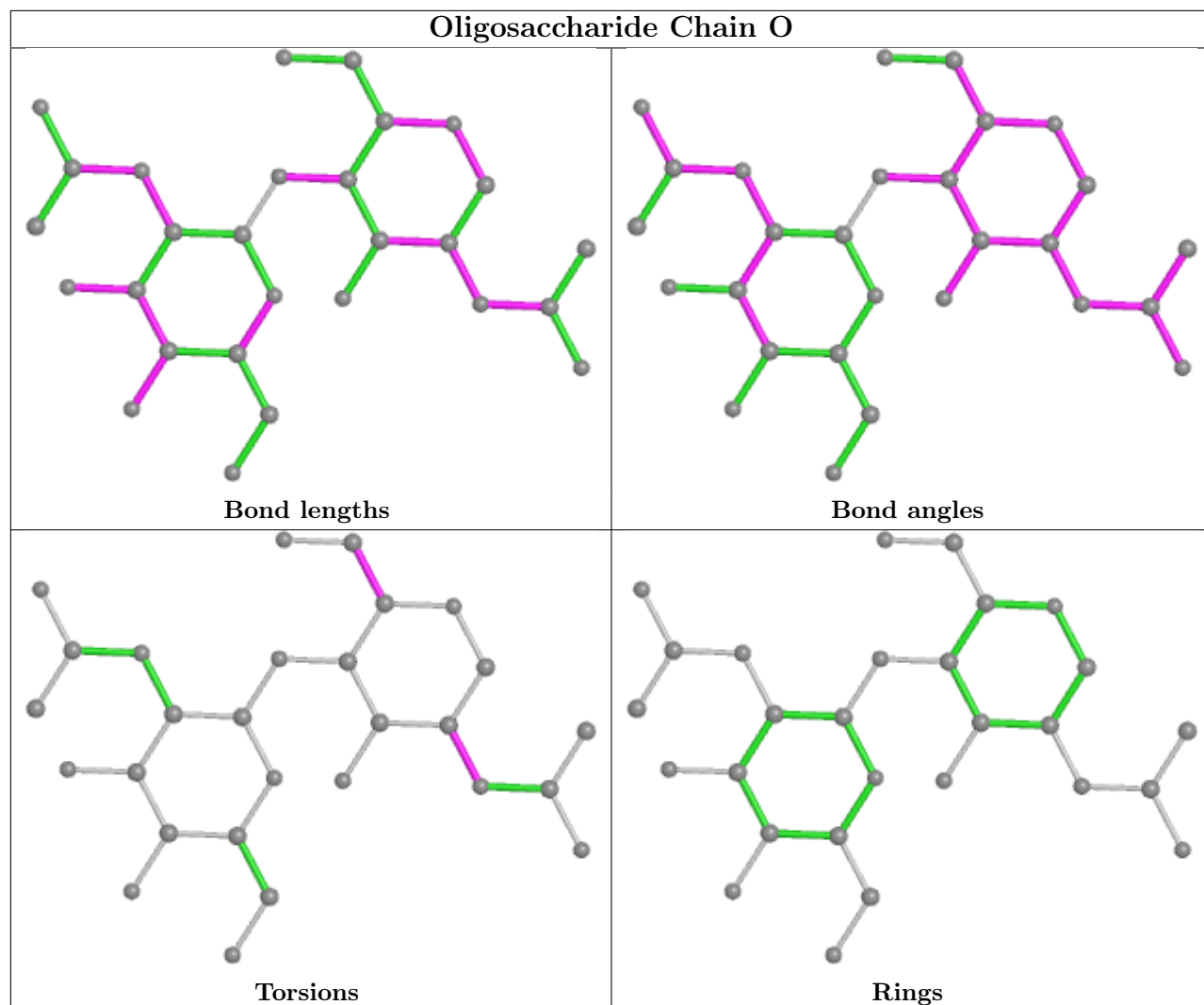
There are no ring outliers.

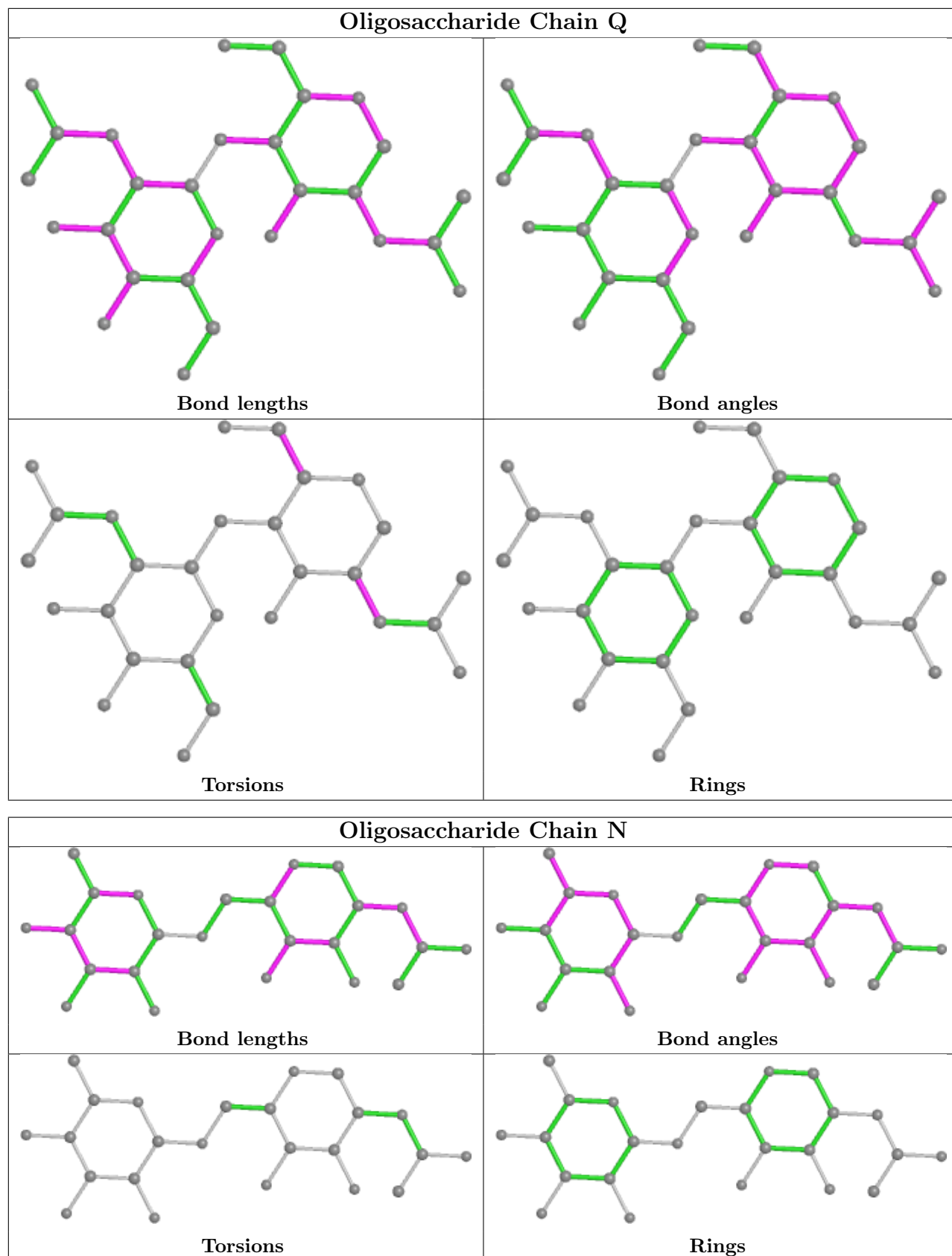
9 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	M	1	NAG	1	0
4	O	1	NAG	1	0
4	Q	1	NAG	1	0
4	M	2	NAG	1	0
4	O	2	NAG	1	0
5	S	2	FUC	4	0
5	S	1	NAG	1	0
5	P	2	FUC	2	0
4	Q	2	NAG	1	0

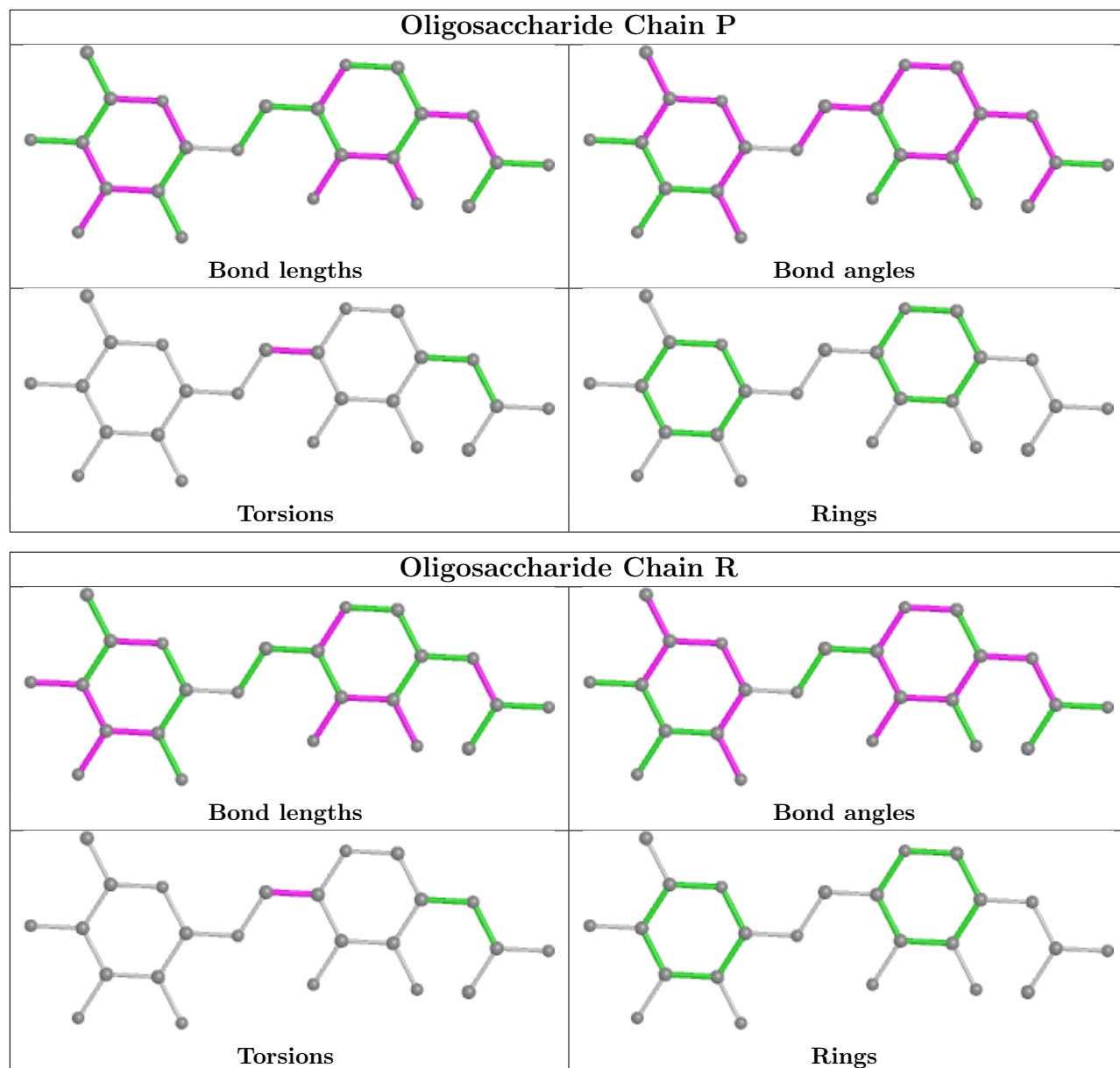
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

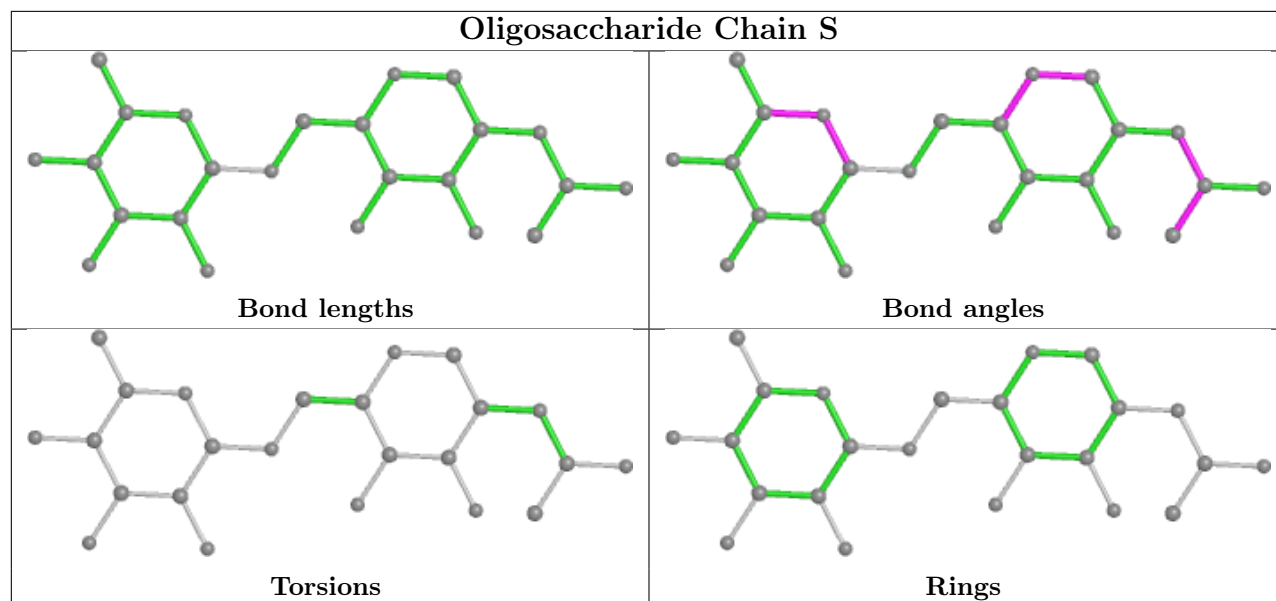












## 5.6 Ligand geometry [i](#)

24 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$
8	NAG	L	1002	3	14,14,15	0.42	0	17,19,21	1.38	2 (11%)
8	NAG	C	1003	3	14,14,15	2.19	5 (35%)	17,19,21	2.13	7 (41%)
7	OXY	J	202	6	1,1,1	0.15	0	-		
7	OXY	G	202	6	1,1,1	0.17	0	-		
6	HEM	D	201	7,1	41,50,50	1.86	8 (19%)	45,82,82	1.68	9 (20%)
6	HEM	G	201	7,1	41,50,50	1.84	9 (21%)	45,82,82	1.77	7 (15%)
7	OXY	E	202	6	1,1,1	0.15	0	-		
8	NAG	F	1004	3	14,14,15	1.80	5 (35%)	17,19,21	1.62	2 (11%)
7	OXY	H	202	6	1,1,1	0.12	0	-		
7	OXY	D	202	6	1,1,1	0.19	0	-		
8	NAG	L	1001	3	14,14,15	2.47	7 (50%)	17,19,21	2.12	3 (17%)
7	OXY	K	202	6	1,1,1	0.11	0	-		
6	HEM	A	201	7,1	41,50,50	1.92	8 (19%)	45,82,82	1.79	9 (20%)
8	NAG	I	1003	3	14,14,15	2.09	5 (35%)	17,19,21	1.58	3 (17%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
7	OXY	B	202	6	1,1,1	0.15	0	-		
6	HEM	H	201	2,7	41,50,50	1.95	6 (14%)	45,82,82	1.96	8 (17%)
6	HEM	K	201	2,7	41,50,50	2.01	7 (17%)	45,82,82	1.94	10 (22%)
8	NAG	C	1004	3	14,14,15	2.48	6 (42%)	17,19,21	2.64	9 (52%)
6	HEM	J	201	7,1	41,50,50	1.89	7 (17%)	45,82,82	1.74	6 (13%)
7	OXY	A	202	6	1,1,1	0.16	0	-		
6	HEM	E	201	2,7	41,50,50	1.92	7 (17%)	45,82,82	1.88	10 (22%)
6	HEM	B	201	2,7	41,50,50	2.01	8 (19%)	45,82,82	1.91	10 (22%)
8	NAG	F	1003	3	14,14,15	2.28	6 (42%)	17,19,21	2.23	4 (23%)
8	NAG	I	1004	3	14,14,15	1.87	5 (35%)	17,19,21	1.94	4 (23%)

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
8	NAG	L	1002	3	-	0/6/23/26	0/1/1/1
8	NAG	C	1003	3	-	0/6/23/26	0/1/1/1
6	HEM	A	201	7,1	-	2/12/54/54	-
6	HEM	E	201	2,7	-	6/12/54/54	-
8	NAG	I	1003	3	-	0/6/23/26	0/1/1/1
6	HEM	D	201	7,1	-	2/12/54/54	-
6	HEM	B	201	2,7	-	6/12/54/54	-
8	NAG	L	1001	3	-	3/6/23/26	0/1/1/1
6	HEM	H	201	2,7	-	6/12/54/54	-
6	HEM	K	201	2,7	-	5/12/54/54	-
8	NAG	C	1004	3	-	3/6/23/26	0/1/1/1
8	NAG	F	1003	3	-	0/6/23/26	0/1/1/1
6	HEM	G	201	7,1	-	2/12/54/54	-
8	NAG	F	1004	3	-	0/6/23/26	0/1/1/1
6	HEM	J	201	7,1	-	2/12/54/54	-
8	NAG	I	1004	3	-	0/6/23/26	0/1/1/1

The worst 5 of 99 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	K	201	HEM	C3D-C2D	8.42	1.54	1.36

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	B	201	HEM	C3D-C2D	8.00	1.53	1.36
6	A	201	HEM	C3D-C2D	7.93	1.53	1.36
6	H	201	HEM	C3D-C2D	7.79	1.53	1.36
6	J	201	HEM	C3D-C2D	7.75	1.53	1.36

The worst 5 of 103 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	K	201	HEM	C4D-ND-C1D	7.50	112.82	105.07
8	F	1003	NAG	O5-C1-C2	-7.04	100.17	111.29
6	H	201	HEM	C4D-ND-C1D	6.69	111.99	105.07
6	B	201	HEM	C4D-ND-C1D	6.44	111.73	105.07
6	E	201	HEM	C4D-ND-C1D	6.37	111.66	105.07

There are no chirality outliers.

5 of 37 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	B	201	HEM	C1A-C2A-CAA-CBA
6	B	201	HEM	C3A-C2A-CAA-CBA
6	E	201	HEM	C1A-C2A-CAA-CBA
6	E	201	HEM	C3A-C2A-CAA-CBA
6	H	201	HEM	C1A-C2A-CAA-CBA

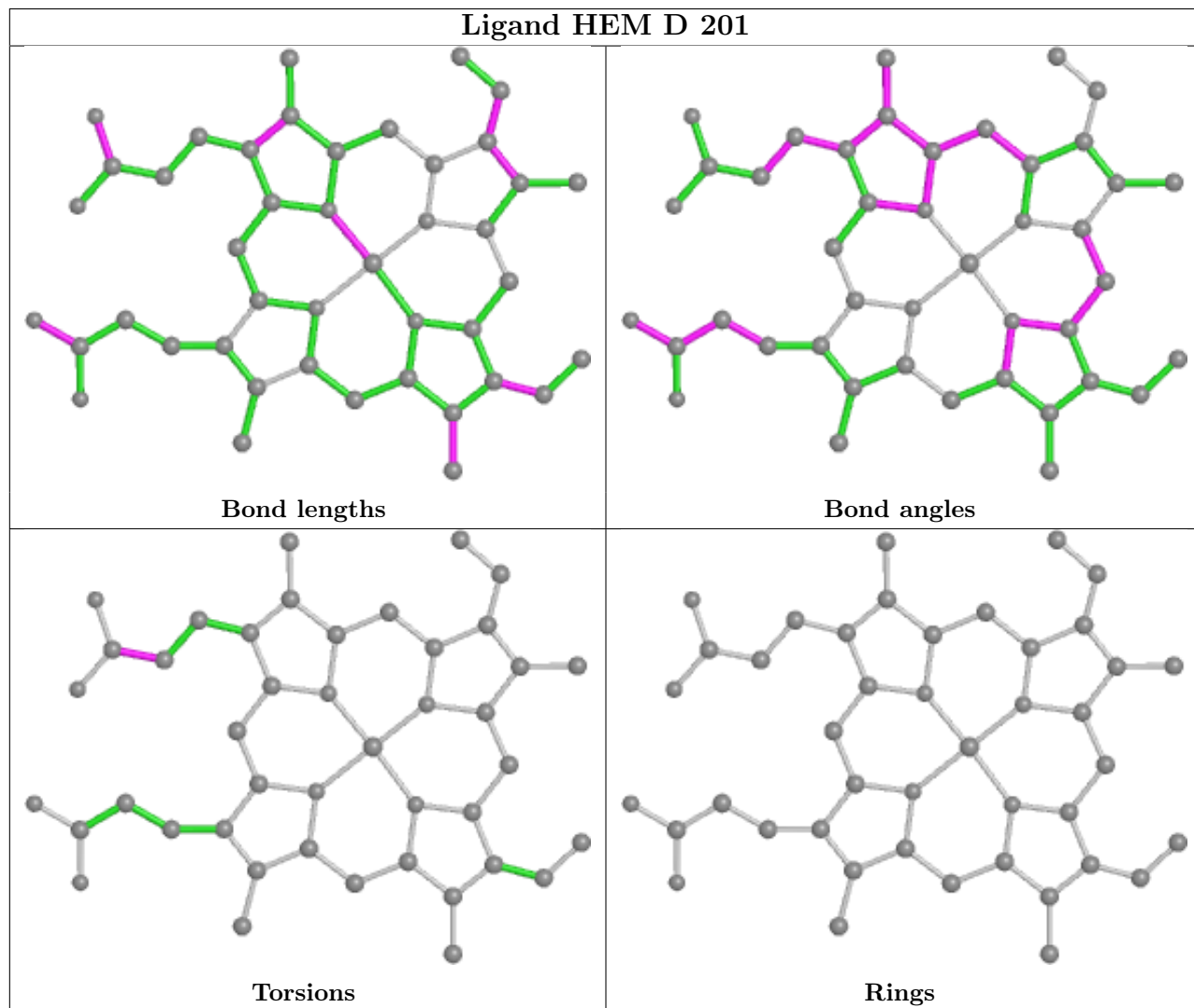
There are no ring outliers.

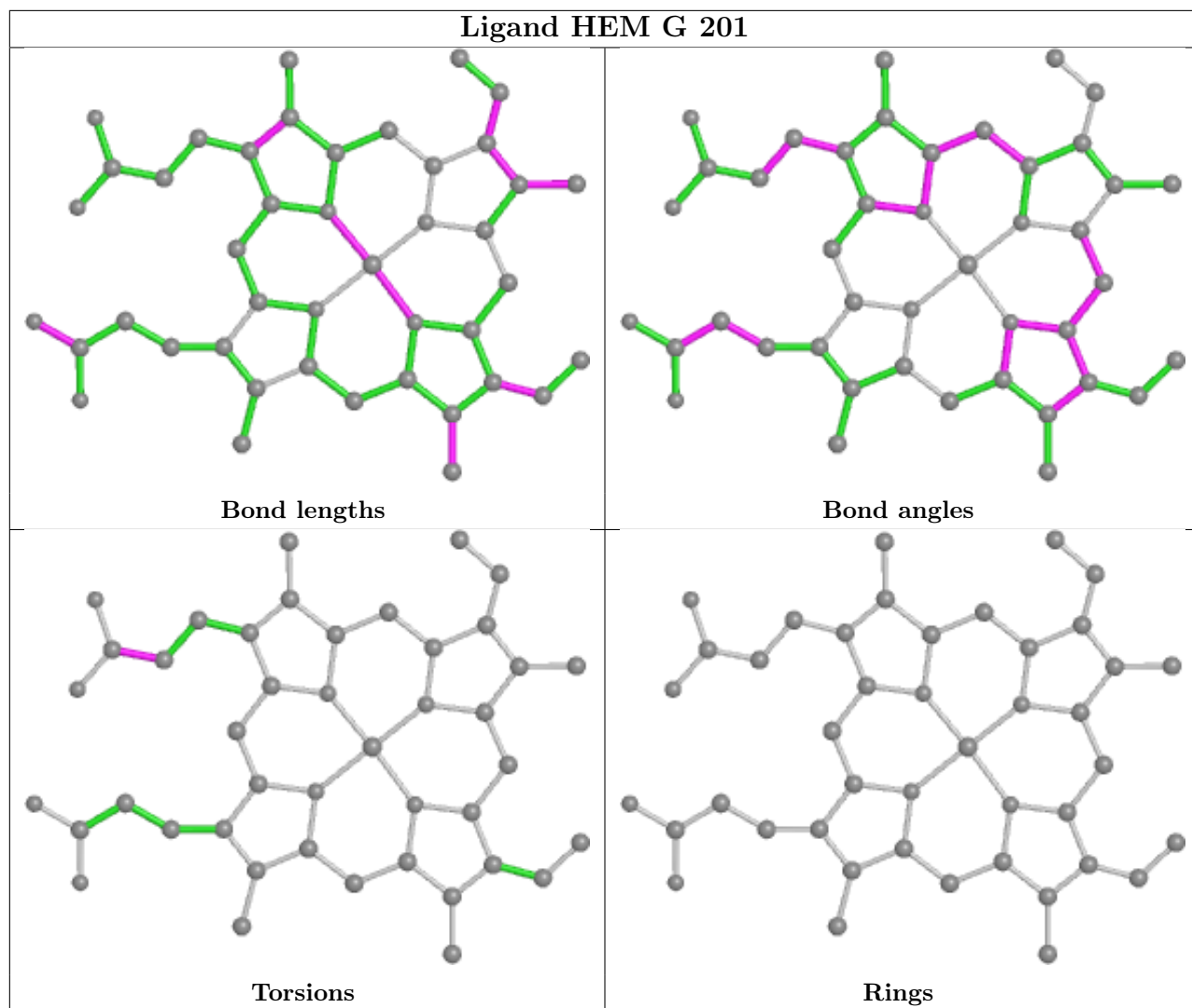
7 monomers are involved in 18 short contacts:

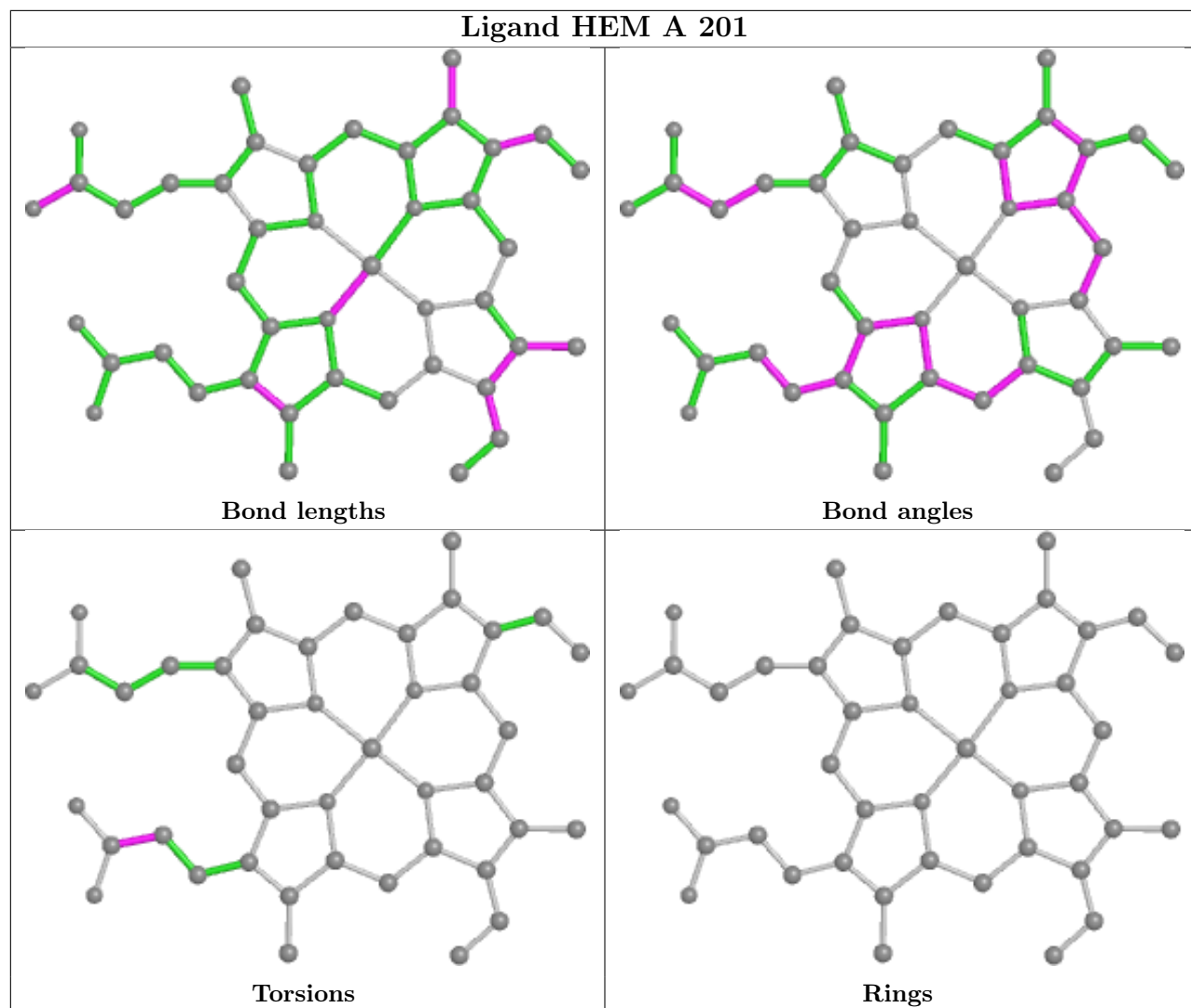
Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	L	1002	NAG	1	0
8	C	1003	NAG	1	0
6	A	201	HEM	1	0
6	H	201	HEM	4	0
6	K	201	HEM	2	0
6	E	201	HEM	5	0
6	B	201	HEM	4	0

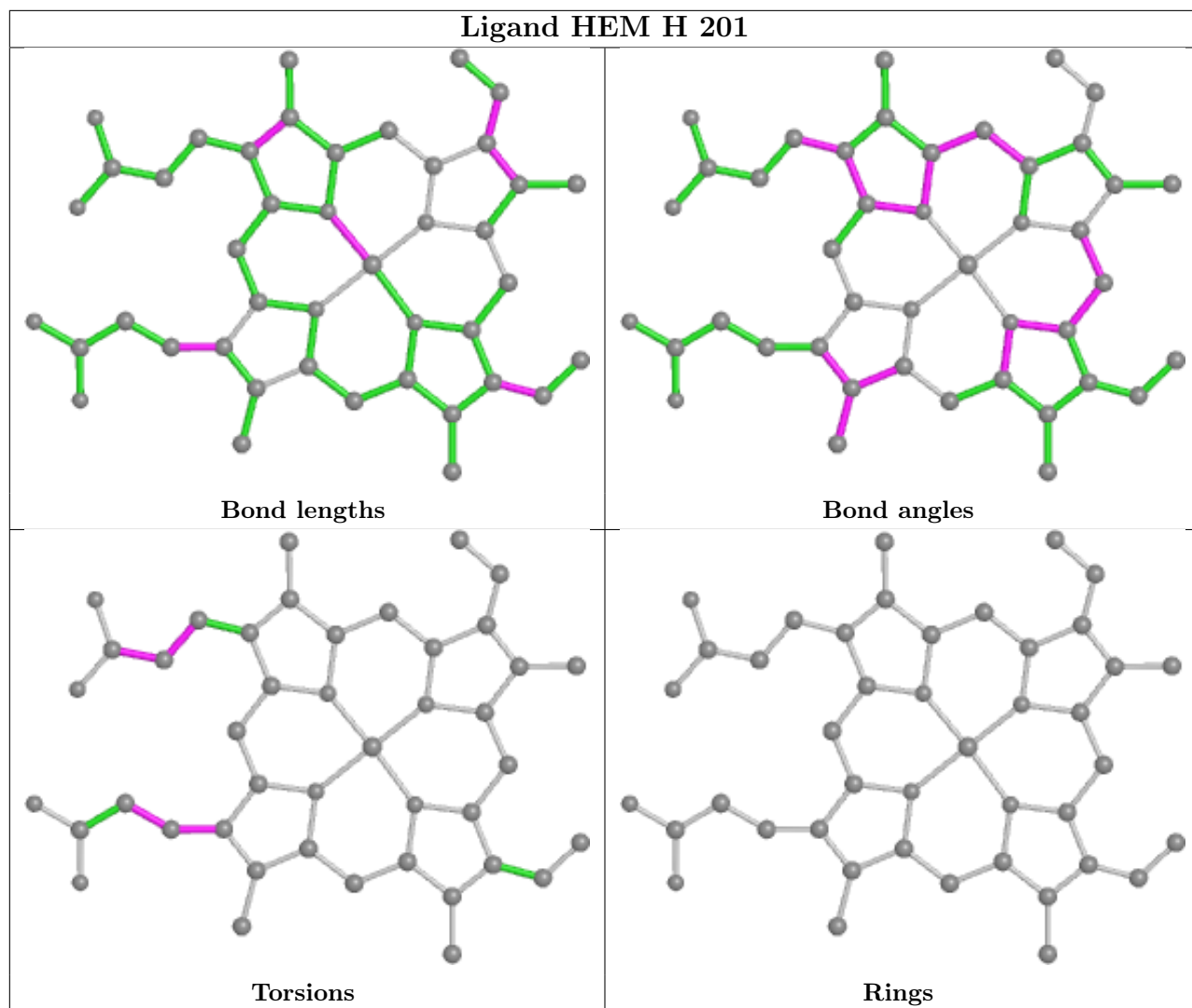
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 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.

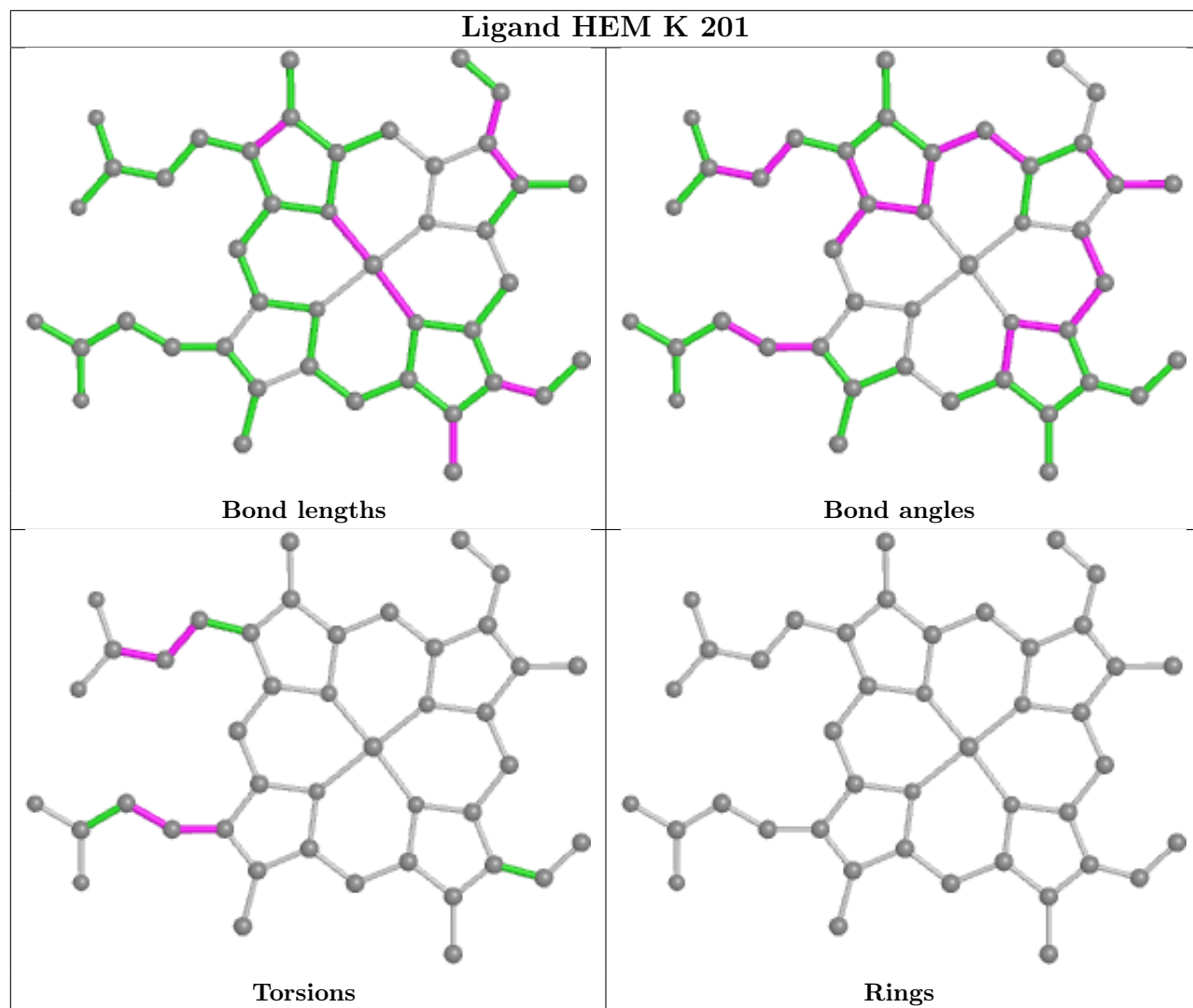


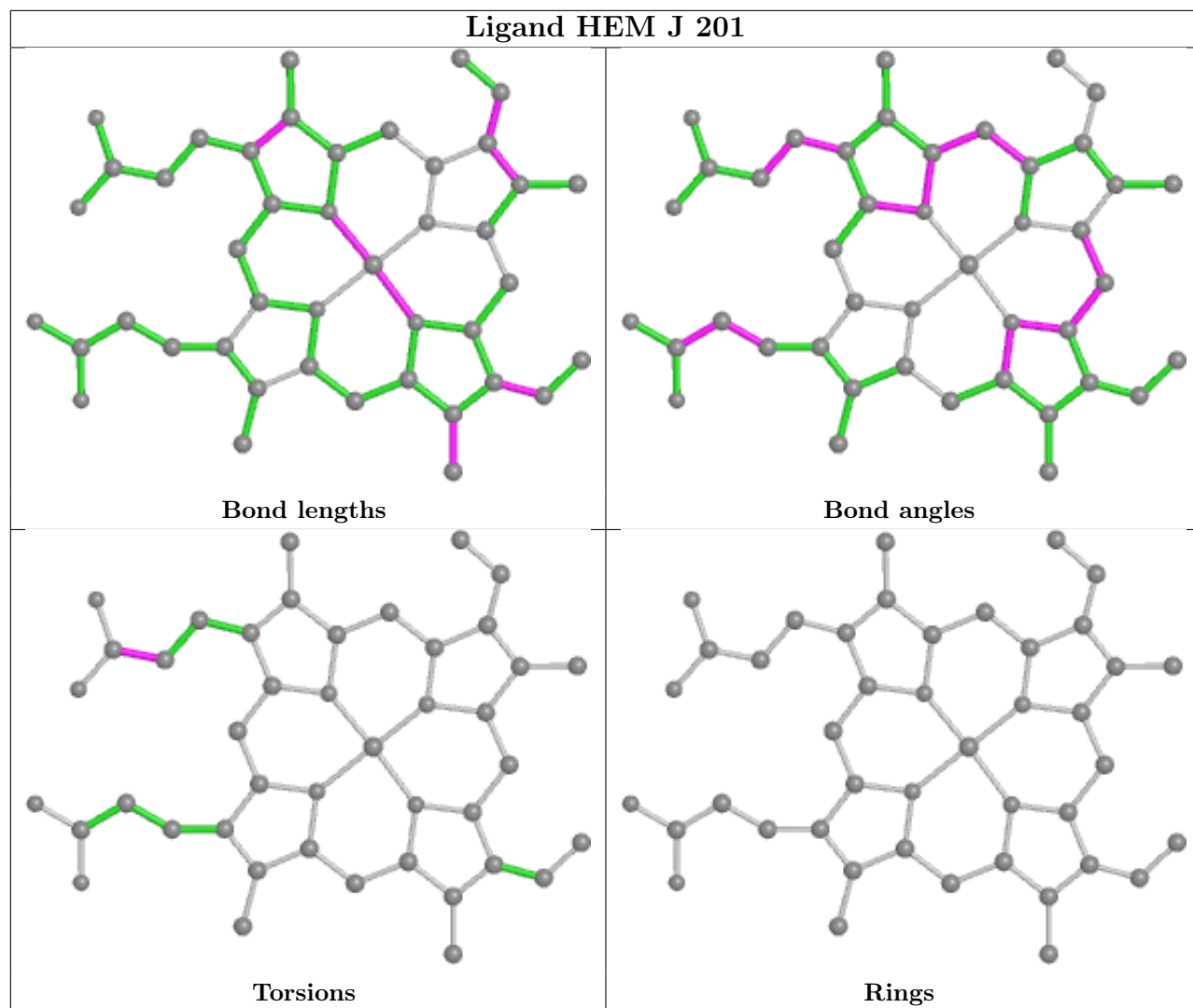


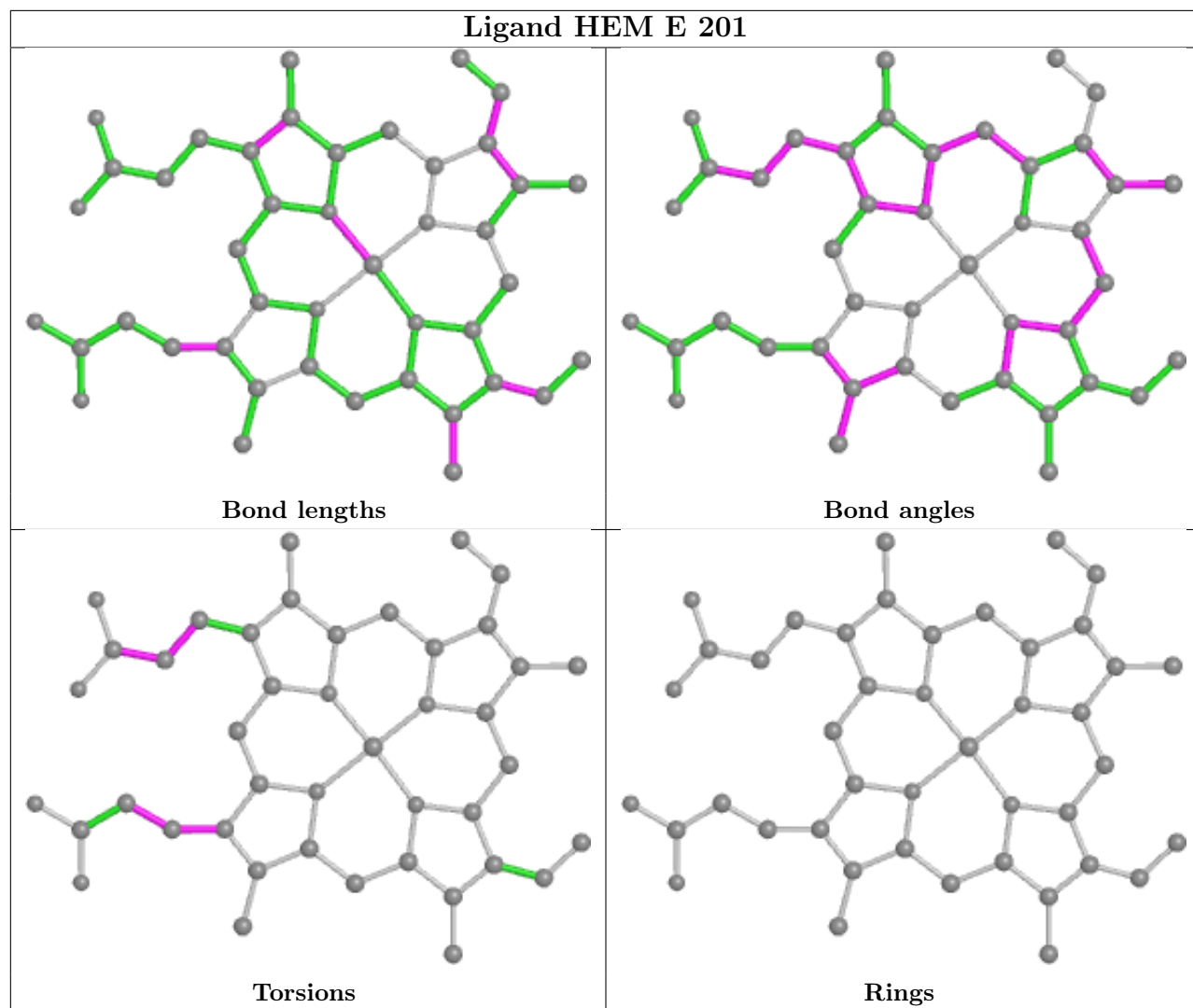


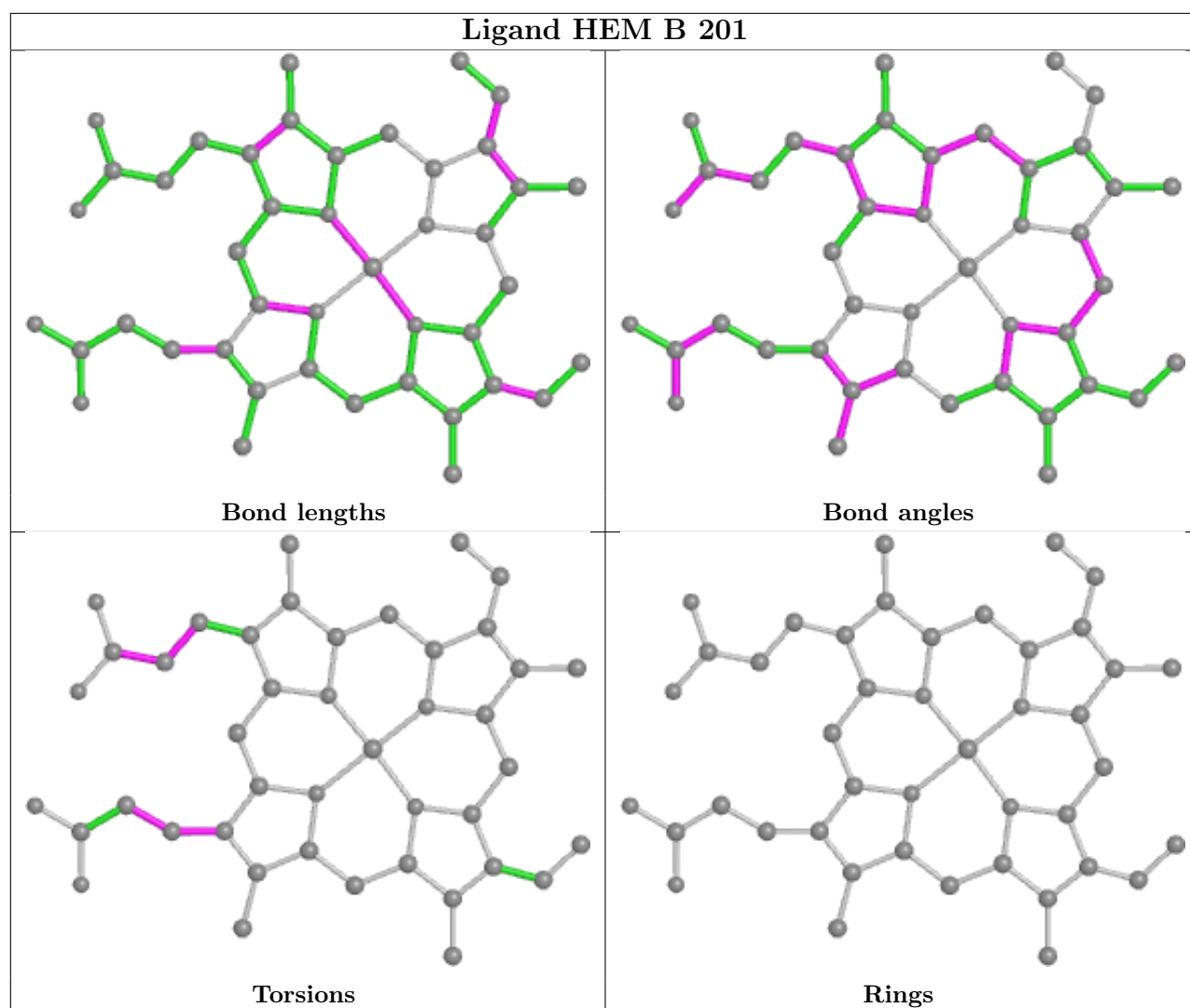












## 5.7 Other polymers [i](#)

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	141/141 (100%)	-0.14	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	37, 56, 86, 124	0
1	D	141/141 (100%)	-0.10	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	40, 57, 84, 103	0
1	G	141/141 (100%)	-0.07	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	41, 60, 96, 132	0
1	J	141/141 (100%)	1.02	32 (22%) <span style="border: 1px solid red; padding: 2px;">0</span> <span style="border: 1px solid red; padding: 2px;">0</span>	110, 156, 201, 300	0
2	B	146/146 (100%)	-0.16	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	40, 84, 129, 196	0
2	E	146/146 (100%)	0.19	7 (4%) <span style="border: 1px solid red; padding: 2px;">30</span> <span style="border: 1px solid red; padding: 2px;">27</span>	47, 93, 138, 233	0
2	H	146/146 (100%)	-0.10	2 (1%) <span style="border: 1px solid blue; padding: 2px;">75</span> <span style="border: 1px solid blue; padding: 2px;">75</span>	48, 85, 131, 204	0
2	K	146/146 (100%)	2.15	63 (43%) <span style="border: 1px solid red; padding: 2px;">0</span> <span style="border: 1px solid red; padding: 2px;">0</span>	127, 182, 250, 268	0
3	C	309/347 (89%)	0.11	8 (2%) <span style="border: 1px solid blue; padding: 2px;">56</span> <span style="border: 1px solid blue; padding: 2px;">52</span>	45, 81, 121, 186	0
3	F	309/347 (89%)	0.13	7 (2%) <span style="border: 1px solid blue; padding: 2px;">60</span> <span style="border: 1px solid blue; padding: 2px;">58</span>	42, 72, 118, 165	1 (0%)
3	I	309/347 (89%)	-0.03	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	45, 67, 106, 133	0
3	L	309/347 (89%)	0.18	11 (3%) <span style="border: 1px solid red; padding: 2px;">42</span> <span style="border: 1px solid red; padding: 2px;">37</span>	66, 103, 141, 187	1 (0%)
All	All	2384/2536 (94%)	0.22	130 (5%) <span style="border: 1px solid red; padding: 2px;">25</span> <span style="border: 1px solid red; padding: 2px;">21</span>	37, 81, 176, 300	2 (0%)

The worst 5 of 130 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	K	75	LEU	8.3
2	K	11	VAL	8.0
2	K	45	PHE	7.3
2	K	146	HIS	6.9
2	K	1	VAL	6.8

### 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

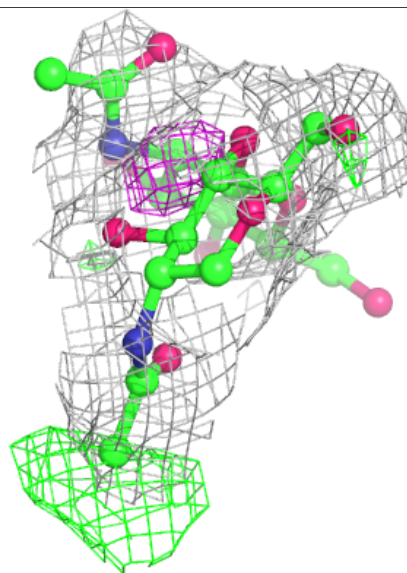
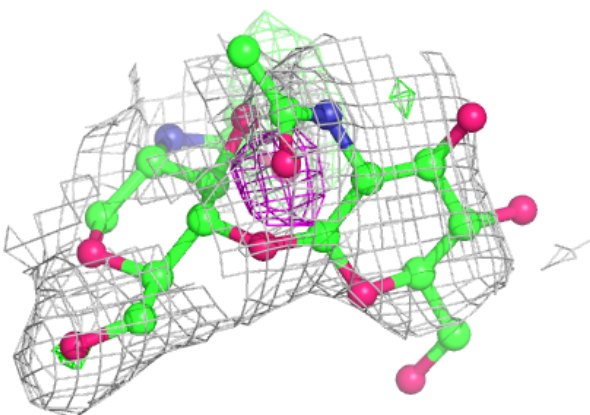
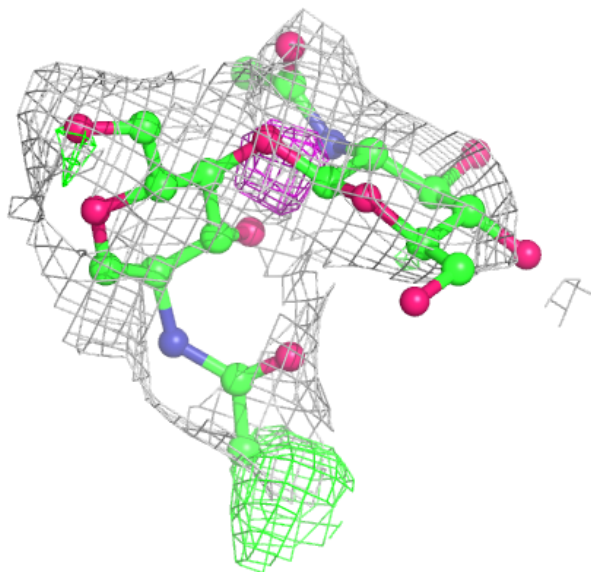
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	NAG	Q	1	14/15	0.72	0.19	126,126,126,126	0
5	FUC	P	2	10/11	0.72	0.46	149,149,149,149	0
4	NAG	O	1	14/15	0.74	0.20	118,118,118,118	0
5	NAG	N	1	14/15	0.77	0.37	138,138,138,138	0
5	NAG	S	1	14/15	0.77	0.28	212,212,212,212	0
5	NAG	P	1	14/15	0.79	0.42	135,135,135,135	0
4	NAG	M	1	14/15	0.79	0.16	122,122,122,122	0
4	NAG	O	2	14/15	0.79	0.25	161,161,161,161	0
4	NAG	M	2	14/15	0.80	0.34	180,180,180,180	0
5	NAG	R	1	14/15	0.82	0.35	141,141,141,141	0
5	FUC	S	2	10/11	0.82	0.28	188,188,188,188	0
4	NAG	Q	2	14/15	0.84	0.32	168,168,168,168	0
5	FUC	R	2	10/11	0.87	0.25	148,148,148,148	0
5	FUC	N	2	10/11	0.92	0.21	138,138,138,138	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.

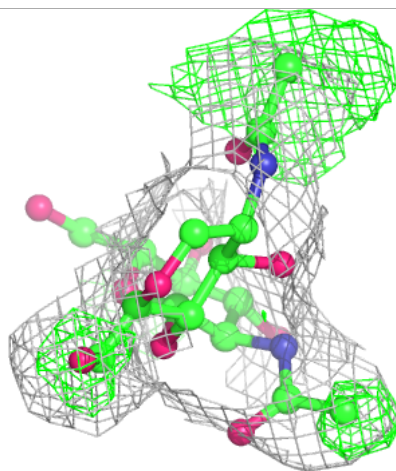
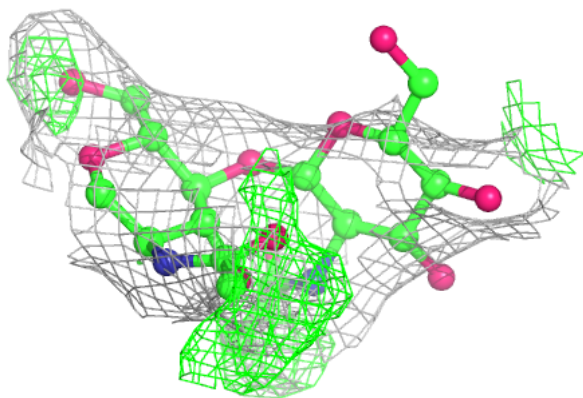
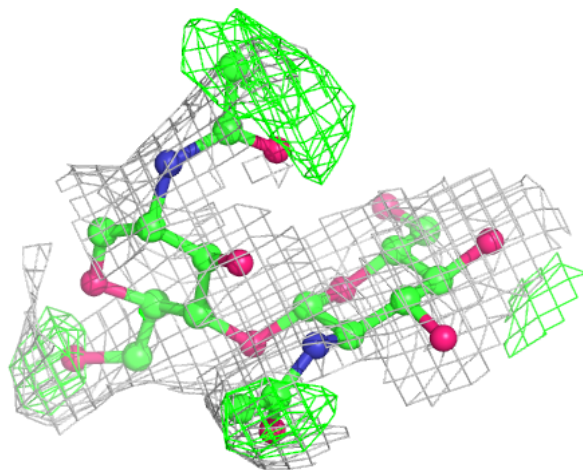
**Electron density around Chain M:**

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



**Electron density around Chain O:**

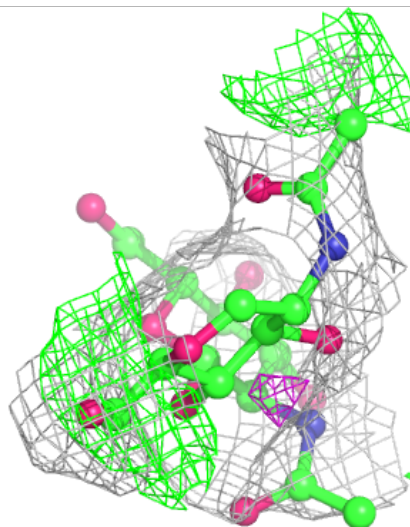
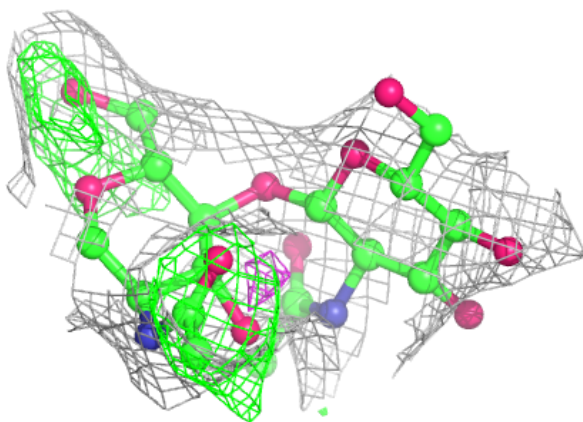
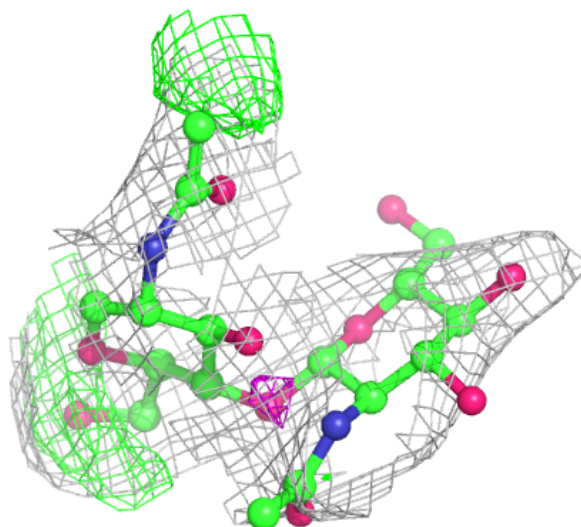
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





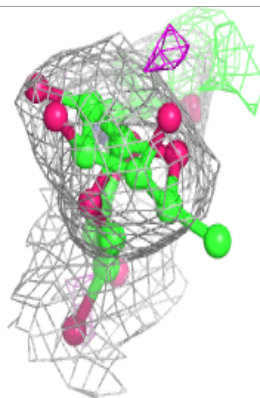
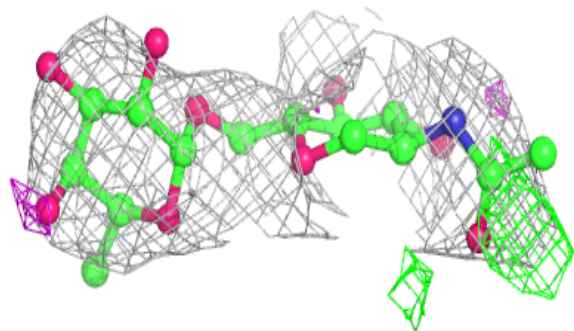
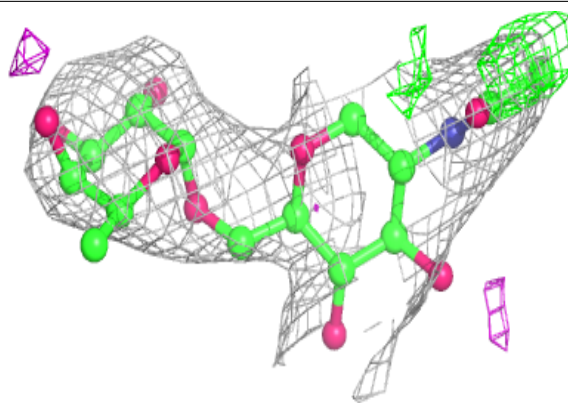
**Electron density around Chain Q:**

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

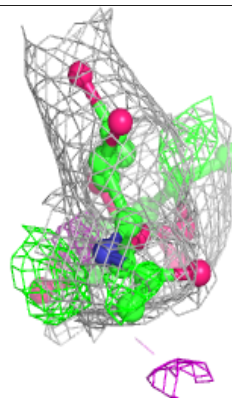
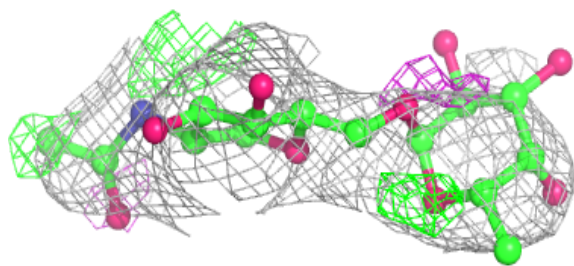
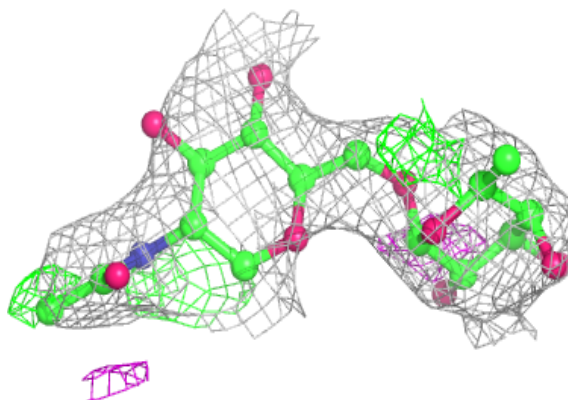


**Electron density around Chain N:**

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

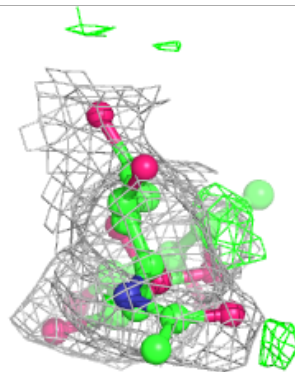
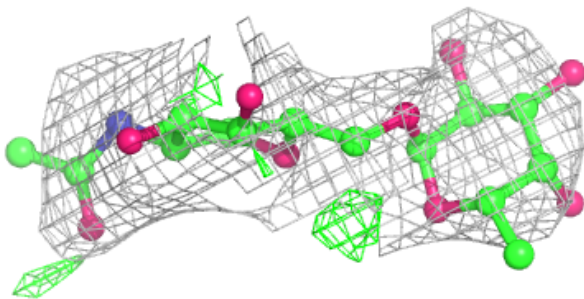
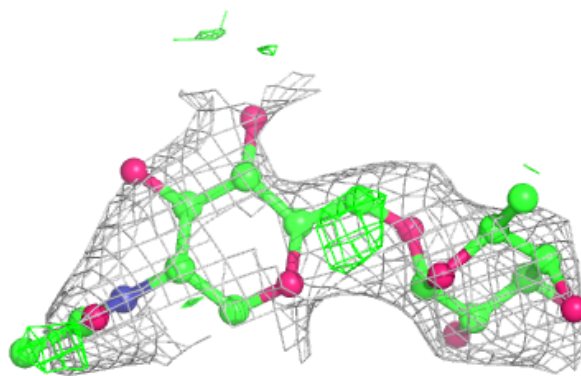
**Electron density around Chain P:**

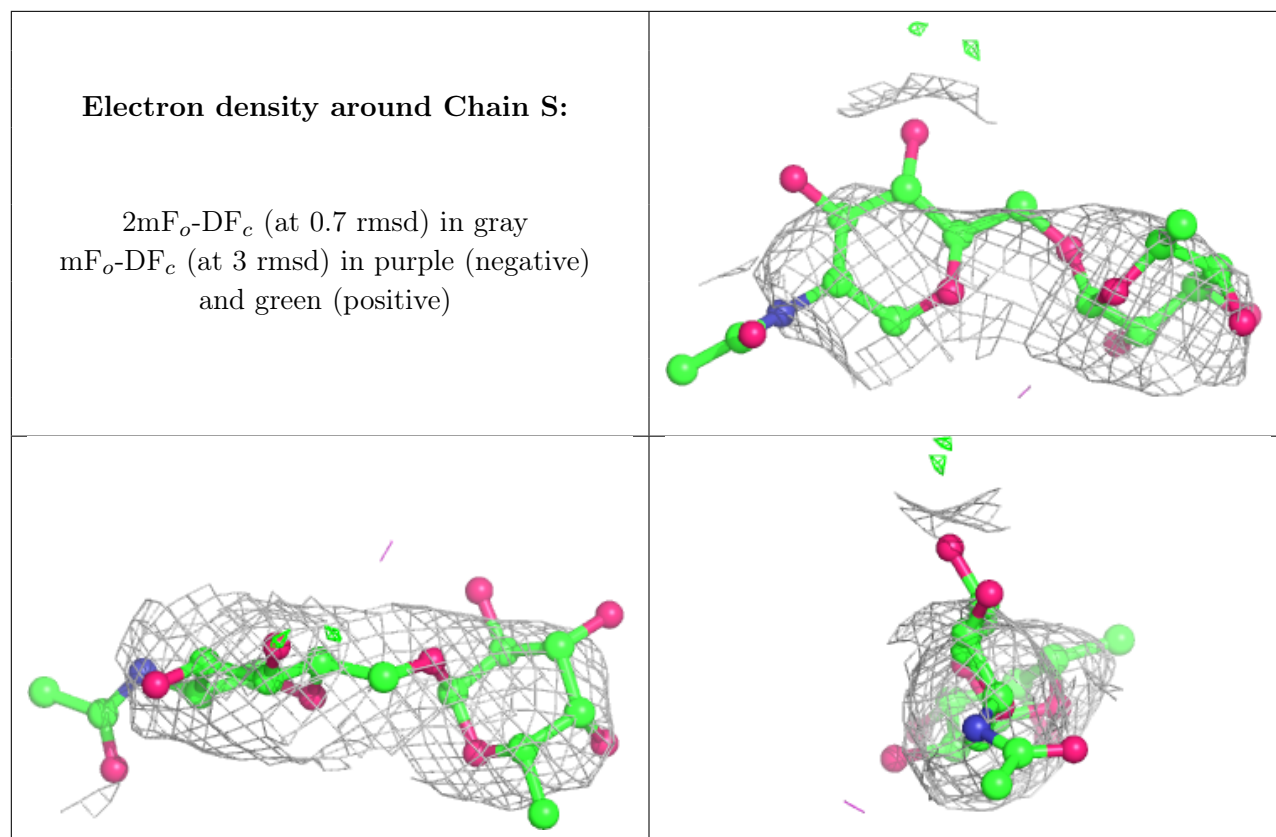
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around Chain R:**

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





## 6.4 Ligands [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, 95<sup>th</sup> 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.

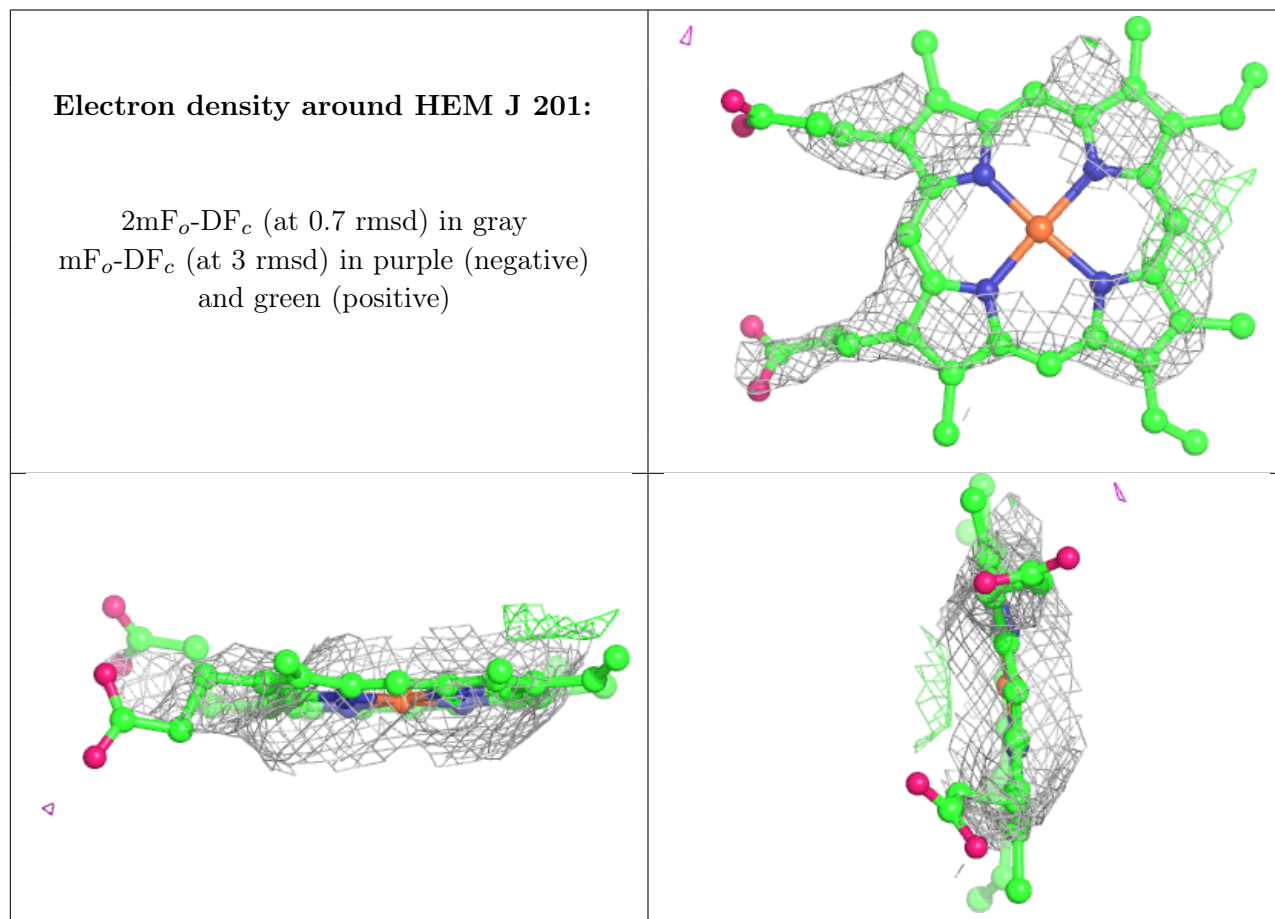
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
8	NAG	C	1003	14/15	0.49	0.46	172,172,172,172	0
8	NAG	I	1004	14/15	0.64	0.29	150,150,150,150	0
8	NAG	I	1003	14/15	0.67	0.41	164,164,164,164	0
8	NAG	L	1001	14/15	0.71	0.17	138,138,138,138	0
8	NAG	F	1004	14/15	0.72	0.30	144,144,144,144	0
8	NAG	F	1003	14/15	0.72	0.53	174,174,174,174	0
8	NAG	L	1002	14/15	0.81	0.29	171,171,171,171	0
8	NAG	C	1004	14/15	0.84	0.25	159,159,159,159	0
6	HEM	J	201	43/43	0.93	0.34	183,186,190,191	0
6	HEM	K	201	43/43	0.93	0.32	167,186,195,199	0
6	HEM	E	201	43/43	0.96	0.23	95,101,114,121	0
7	OXY	J	202	2/2	0.97	0.21	225,225,225,229	0
6	HEM	B	201	43/43	0.97	0.21	76,80,95,103	0
6	HEM	H	201	43/43	0.97	0.25	83,87,101,109	0

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
7	OXY	K	202	2/2	0.98	0.47	292,292,292,295	0
6	HEM	D	201	43/43	0.98	0.19	45,51,57,59	0
6	HEM	G	201	43/43	0.98	0.19	41,58,62,64	0
6	HEM	A	201	43/43	0.99	0.18	45,54,62,63	0
7	OXY	A	202	2/2	0.99	0.38	92,92,92,96	0
7	OXY	E	202	2/2	0.99	0.16	146,146,146,148	0
7	OXY	G	202	2/2	0.99	0.49	101,101,101,105	0
7	OXY	H	202	2/2	0.99	0.24	174,174,174,176	0
7	OXY	D	202	2/2	1.00	0.22	84,84,84,88	0
7	OXY	B	202	2/2	1.00	0.18	121,121,121,123	0

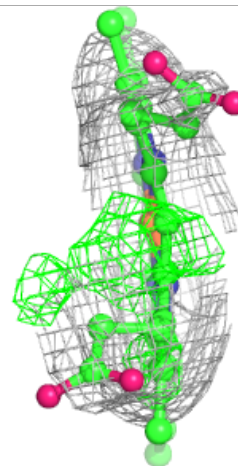
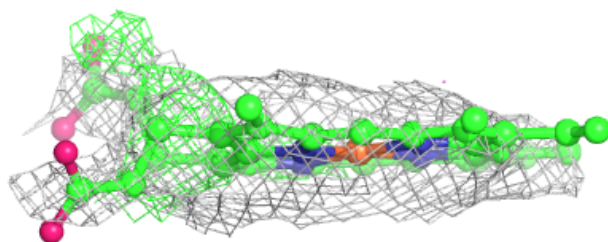
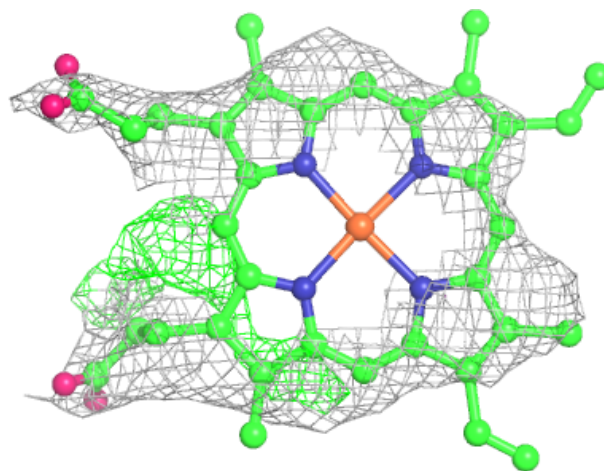
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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





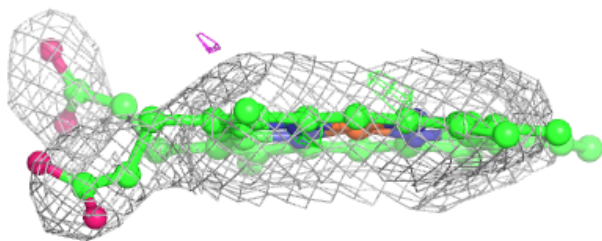
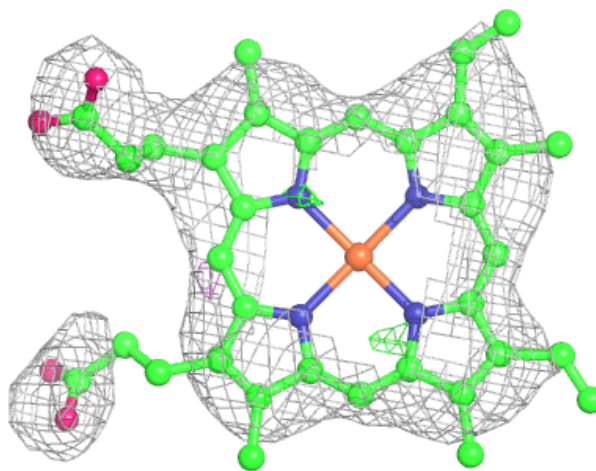
**Electron density around HEM K 201:**

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



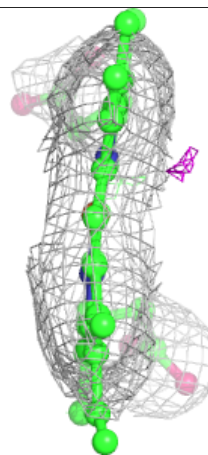
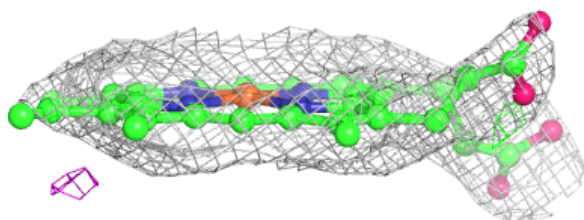
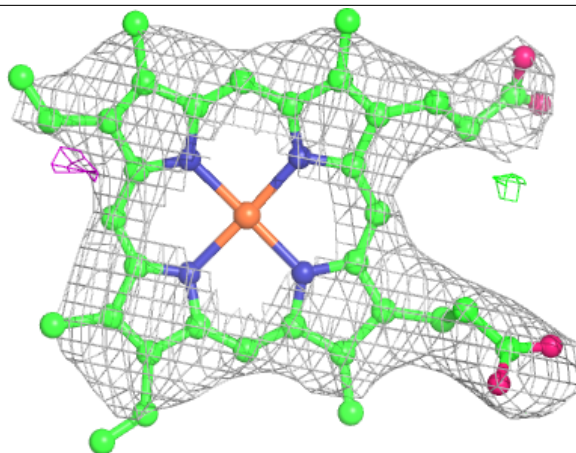
**Electron density around HEM E 201:**

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



**Electron density around HEM B 201:**

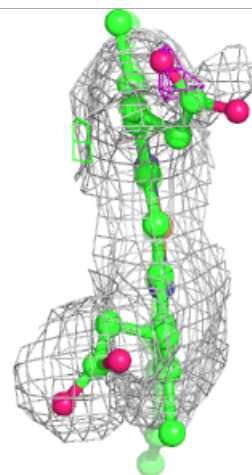
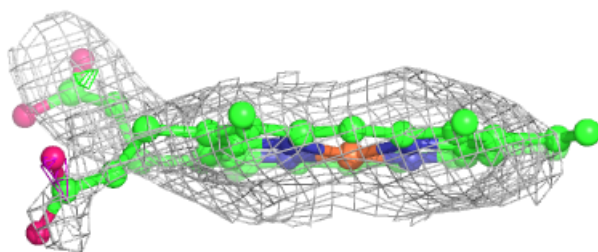
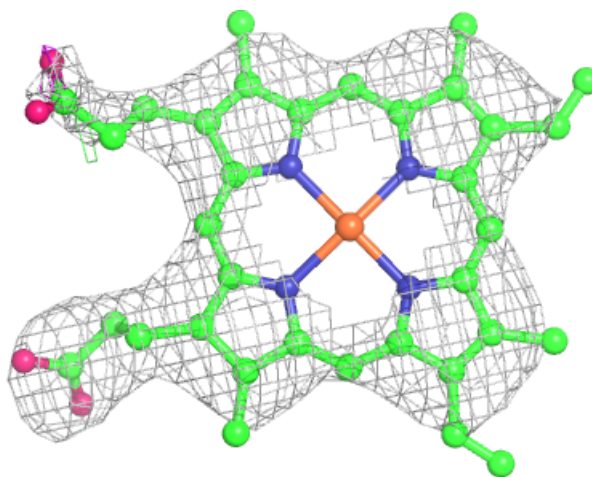
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





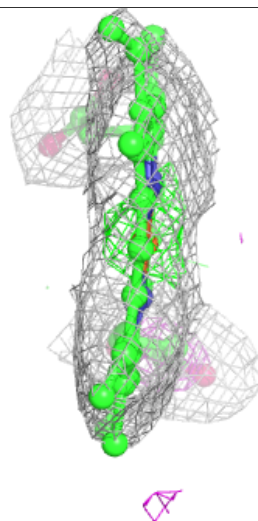
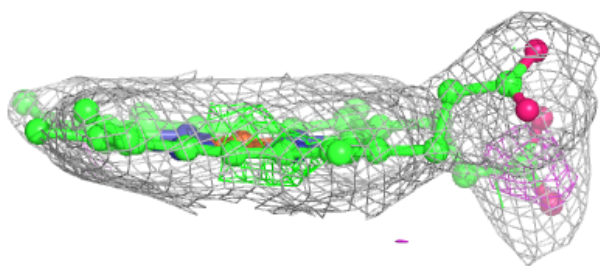
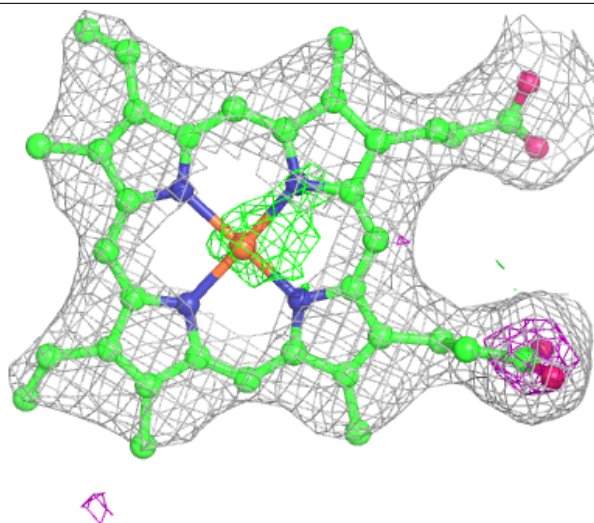
**Electron density around HEM H 201:**

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



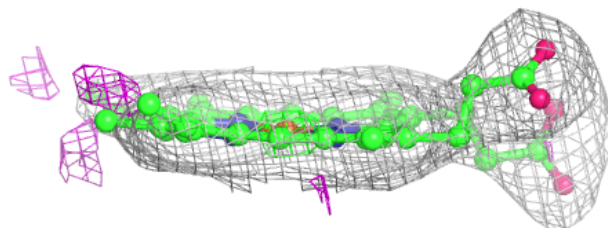
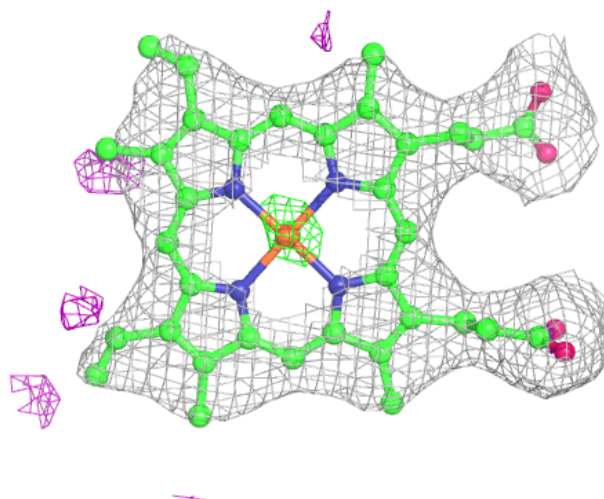
**Electron density around HEM D 201:**

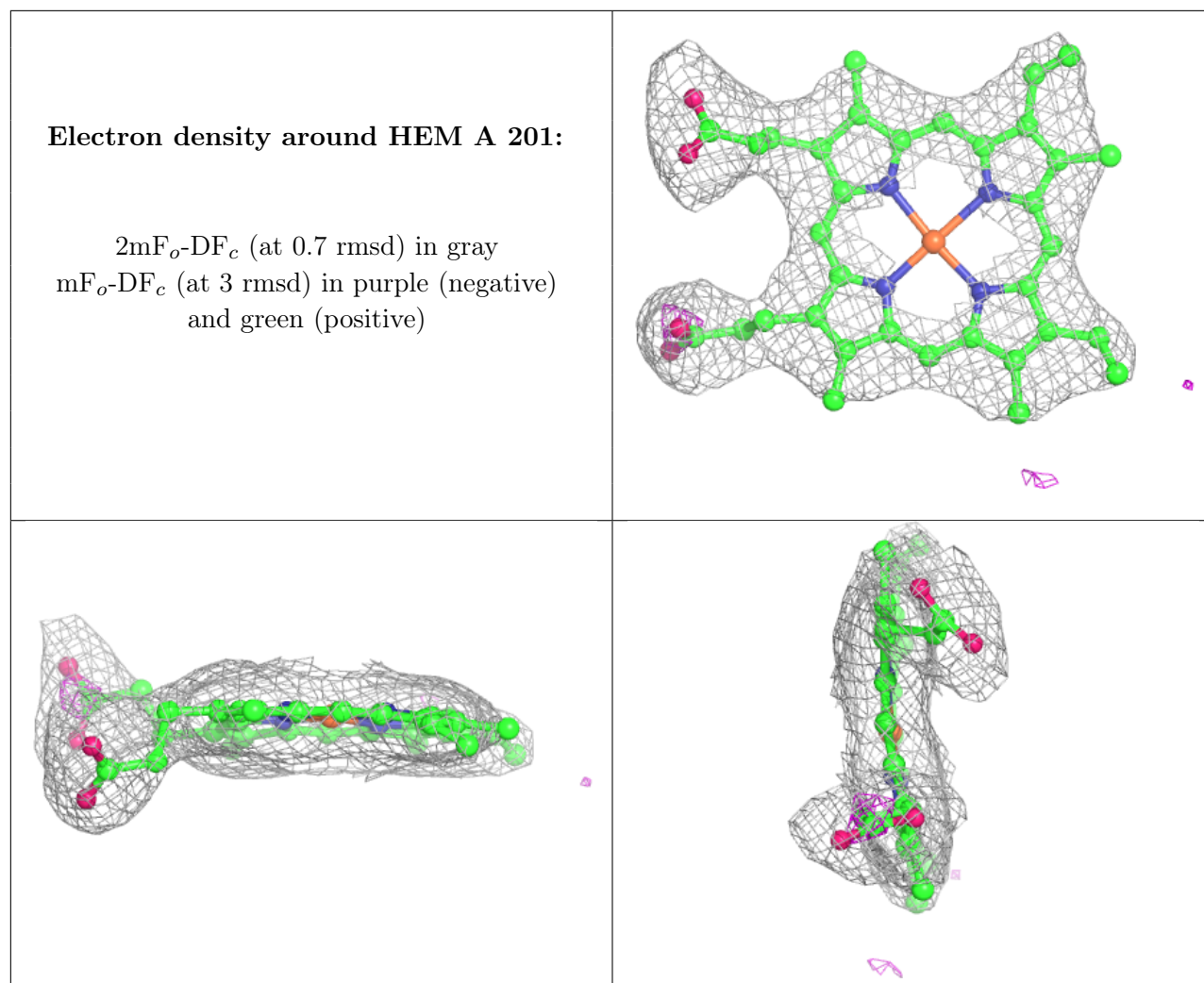
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around HEM G 201:**

$2mF_o-DF_c$  (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.