



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 9, 2020 – 01:23 AM BST

PDB ID : 4WJG  
Title : Structure of *T. brucei* haptoglobin-hemoglobin receptor binding to human haptoglobin-hemoglobin  
Authors : Stoedkilde, K.; Torvund-Jensen, M.; Moestrup, S.K.; Andersen, C.B.F.  
Deposited on : 2014-09-30  
Resolution : 3.10 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 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)  
Xtriage (Phenix) : 1.13  
EDS : 2.13.1  
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) : 2.13.1

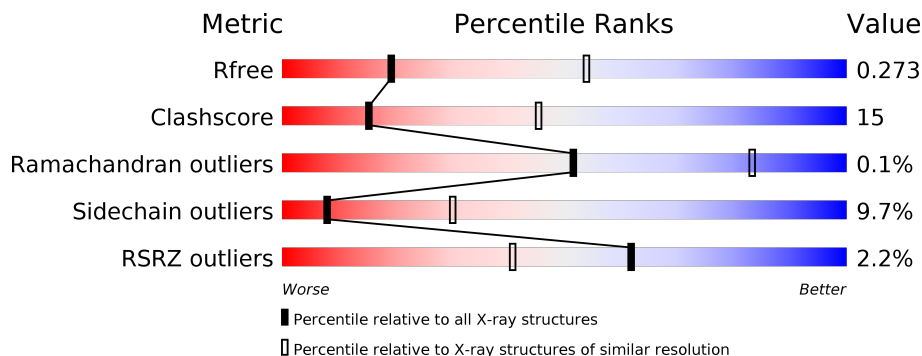
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.10 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.



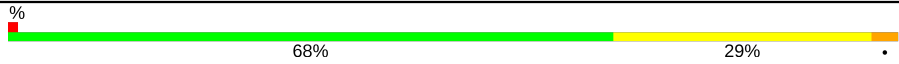
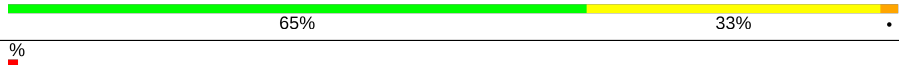
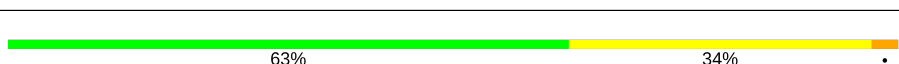
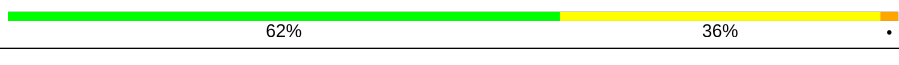
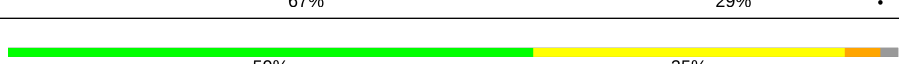




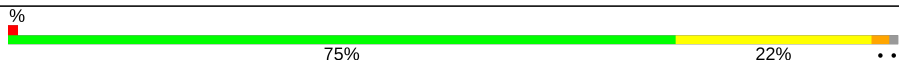

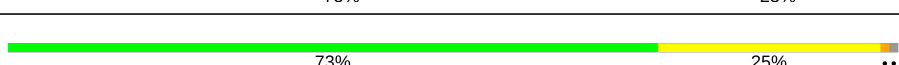


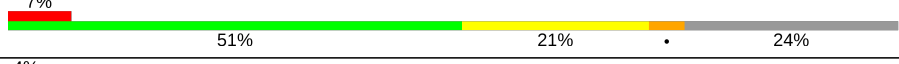


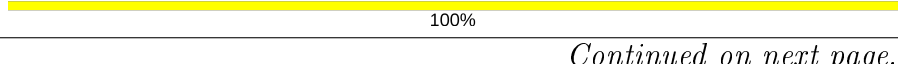


Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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 on 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	F	141	
1	K	141	
1	P	141	
1	U	141	
1	Z	141	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	1	146	 % 68% 29% .
2	B	146	 % 65% 33% .
2	G	146	 % 66% 31% .
2	L	146	 % 63% 34% .
2	Q	146	 % 62% 36% .
2	V	146	 % 67% 29% .
3	2	315	 % 59% 35% . .
3	C	315	 % 57% 36% 6% .
3	H	315	 % 57% 36% 5% .
3	M	315	 % 57% 37% 5% .
3	R	315	 % 57% 36% 5% .
3	W	315	 % 58% 34% 6% .
4	3	146	 % 75% 22% ..
4	D	146	 % 75% 22% ..
4	I	146	 % 74% 23% ..
4	N	146	 % 75% 23% ..
4	S	146	 % 73% 25% ..
4	X	146	 % 75% 22% ..
5	4	343	 5% 51% 20% . 24%
5	E	343	 4% 52% 19% 5% 24%
5	J	343	7% 51% 21% . 24%
5	O	343	4% 51% 20% . 24%
5	T	343	7% 50% 21% . 24%
5	Y	343	7% 51% 21% . 24%
6	a	2	100%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
6	b	2	 50% 50%
6	c	2	 100%
6	d	2	 100%
6	e	2	 100%
6	f	2	 100%
6	g	2	 100%
6	h	2	 100%
6	i	2	 100%
6	j	2	 100%

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
9	NAG	C	1001	-	-	-	X
9	NAG	C	1004	-	-	-	X
9	NAG	J	1001	-	-	-	X
9	NAG	M	1004	-	-	-	X
9	NAG	R	1004	-	-	-	X
9	NAG	T	1001	-	-	-	X

## 2 Entry composition [i](#)

There are 9 unique types of molecules in this entry. The entry contains 47792 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	1069	685	187	194	3	0	0	0
1	F	141	1069	685	187	194	3	0	0	0
1	K	141	1069	685	187	194	3	0	0	0
1	P	141	1069	685	187	194	3	0	0	0
1	U	141	1069	685	187	194	3	0	0	0
1	Z	141	1069	685	187	194	3	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	1123	724	195	201	3	0	0	0
2	G	146	1123	724	195	201	3	0	0	0
2	L	146	1123	724	195	201	3	0	0	0
2	Q	146	1123	724	195	201	3	0	0	0
2	V	146	1123	724	195	201	3	0	0	0
2	1	146	1123	724	195	201	3	0	0	0

- Molecule 3 is a protein called Haptoglobin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	310	Total	C	N	O	S	0	0	0
			2437	1552	413	459	13			
3	H	310	Total	C	N	O	S	0	0	0
			2437	1552	413	459	13			
3	M	310	Total	C	N	O	S	0	0	0
			2437	1552	413	459	13			
3	R	310	Total	C	N	O	S	0	0	0
			2437	1552	413	459	13			
3	W	310	Total	C	N	O	S	0	0	0
			2437	1552	413	459	13			
3	2	310	Total	C	N	O	S	0	0	0
			2437	1552	413	459	13			

- Molecule 4 is a protein called Iron-regulated surface determinant protein H.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	D	144	Total	C	N	O	0	0	0
			1183	754	189	240			
4	I	144	Total	C	N	O	0	0	0
			1183	754	189	240			
4	N	144	Total	C	N	O	0	0	0
			1183	754	189	240			
4	S	144	Total	C	N	O	0	0	0
			1183	754	189	240			
4	X	144	Total	C	N	O	0	0	0
			1183	754	189	240			
4	3	144	Total	C	N	O	0	0	0
			1183	754	189	240			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	84	GLY	-	expression tag	UNP Q99TD3
D	85	SER	-	expression tag	UNP Q99TD3
I	84	GLY	-	expression tag	UNP Q99TD3
I	85	SER	-	expression tag	UNP Q99TD3
N	84	GLY	-	expression tag	UNP Q99TD3
N	85	SER	-	expression tag	UNP Q99TD3
S	84	GLY	-	expression tag	UNP Q99TD3
S	85	SER	-	expression tag	UNP Q99TD3
X	84	GLY	-	expression tag	UNP Q99TD3
X	85	SER	-	expression tag	UNP Q99TD3
3	84	GLY	-	expression tag	UNP Q99TD3
3	85	SER	-	expression tag	UNP Q99TD3

- Molecule 5 is a protein called Haptoglobin-hemoglobin receptor.

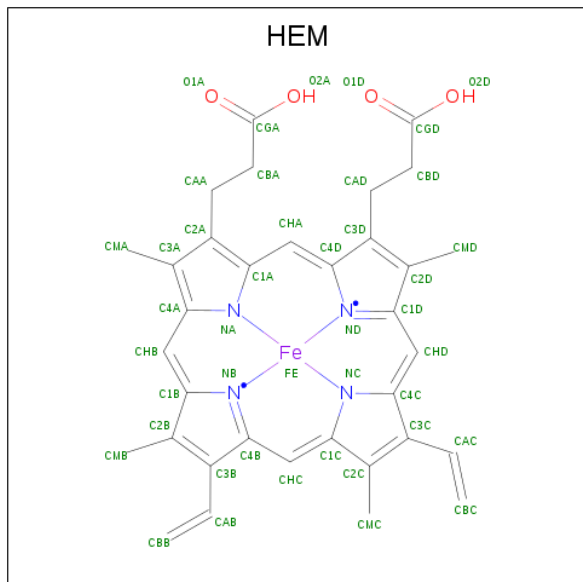
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	E	261	Total 1963	C 1202	N 365	O 393	S 3	0	0	0
5	J	261	Total 1963	C 1202	N 365	O 393	S 3	0	0	0
5	O	261	Total 1963	C 1202	N 365	O 393	S 3	0	0	0
5	T	261	Total 1963	C 1202	N 365	O 393	S 3	0	0	0
5	Y	261	Total 1963	C 1202	N 365	O 393	S 3	0	0	0
5	4	261	Total 1963	C 1202	N 365	O 393	S 3	0	0	0

- Molecule 6 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			
6	a	2	Total 28	C 16	N 2	O 10	0	0	0
6	b	2	Total 28	C 16	N 2	O 10	0	0	0
6	c	2	Total 28	C 16	N 2	O 10	0	0	0
6	d	2	Total 28	C 16	N 2	O 10	0	0	0
6	e	2	Total 28	C 16	N 2	O 10	0	0	0
6	f	2	Total 28	C 16	N 2	O 10	0	0	0
6	g	2	Total 28	C 16	N 2	O 10	0	0	0
6	h	2	Total 28	C 16	N 2	O 10	0	0	0
6	i	2	Total 28	C 16	N 2	O 10	0	0	0
6	j	2	Total 28	C 16	N 2	O 10	0	0	0

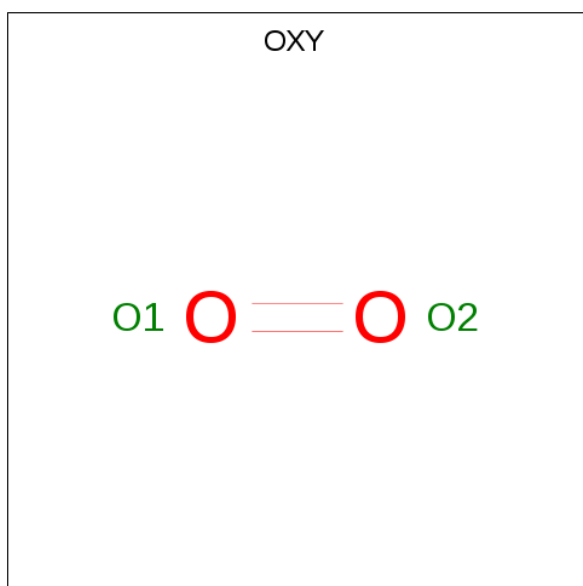
- Molecule 7 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula:  $C_{34}H_{32}FeN_4O_4$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
7	A	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
7	B	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
7	F	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
7	G	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
7	K	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
7	L	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
7	P	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
7	Q	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
7	U	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
7	V	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
7	Z	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
7	1	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		



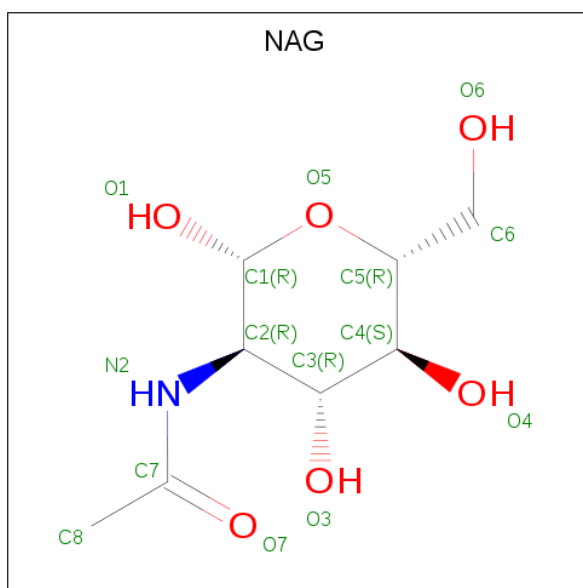
- Molecule 8 is OXYGEN MOLECULE (three-letter code: OXY) (formula: O<sub>2</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	1	Total O 2 2	0	0
8	B	1	Total O 2 2	0	0
8	F	1	Total O 2 2	0	0
8	G	1	Total O 2 2	0	0
8	K	1	Total O 2 2	0	0
8	L	1	Total O 2 2	0	0
8	P	1	Total O 2 2	0	0
8	Q	1	Total O 2 2	0	0
8	U	1	Total O 2 2	0	0
8	V	1	Total O 2 2	0	0
8	Z	1	Total O 2 2	0	0
8	1	1	Total O 2 2	0	0

- Molecule 9 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:

C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
9	C	1	Total	C	N	O	0	0
			14	8	1	5		
9	C	1	Total	C	N	O	0	0
			14	8	1	5		
9	E	1	Total	C	N	O	0	0
			14	8	1	5		
9	E	1	Total	C	N	O	0	0
			14	8	1	5		
9	H	1	Total	C	N	O	0	0
			14	8	1	5		
9	H	1	Total	C	N	O	0	0
			14	8	1	5		
9	J	1	Total	C	N	O	0	0
			14	8	1	5		
9	J	1	Total	C	N	O	0	0
			14	8	1	5		
9	M	1	Total	C	N	O	0	0
			14	8	1	5		
9	M	1	Total	C	N	O	0	0
			14	8	1	5		
9	O	1	Total	C	N	O	0	0
			14	8	1	5		
9	O	1	Total	C	N	O	0	0
			14	8	1	5		
9	R	1	Total	C	N	O	0	0
			14	8	1	5		

Continued on next page...

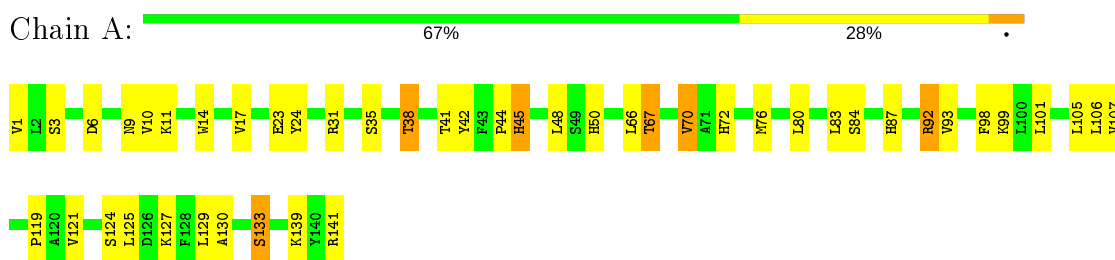
*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
9	R	1	Total 14	8	1	5	0	0
9	T	1	Total 14	8	1	5	0	0
9	T	1	Total 14	8	1	5	0	0
9	W	1	Total 14	8	1	5	0	0
9	W	1	Total 14	8	1	5	0	0
9	Y	1	Total 14	8	1	5	0	0
9	Y	1	Total 14	8	1	5	0	0
9	2	1	Total 14	8	1	5	0	0
9	4	1	Total 14	8	1	5	0	0
9	4	1	Total 14	8	1	5	0	0

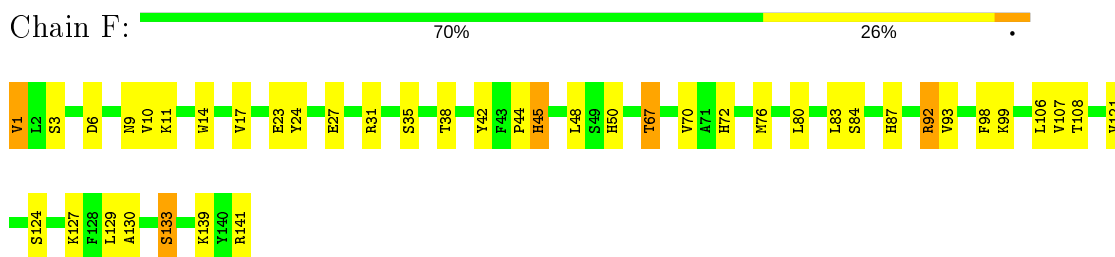
### 3 Residue-property plots [i](#)

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.

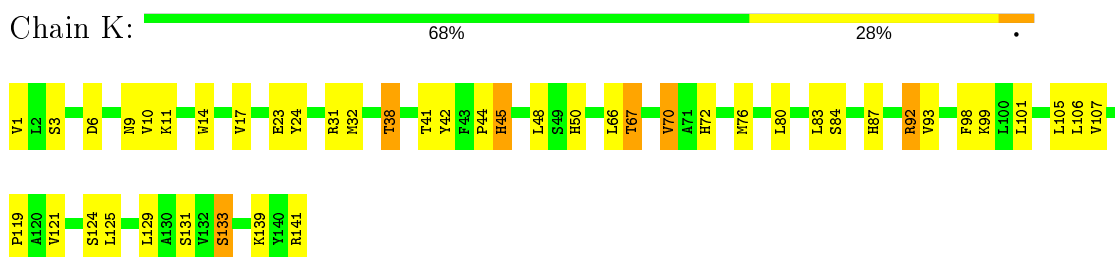
- Molecule 1: Hemoglobin subunit alpha



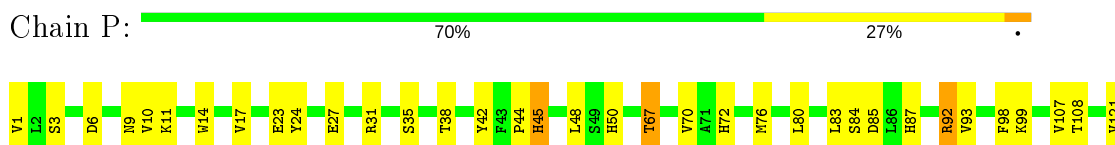
- Molecule 1: Hemoglobin subunit alpha



- Molecule 1: Hemoglobin subunit alpha



- Molecule 1: Hemoglobin subunit alpha

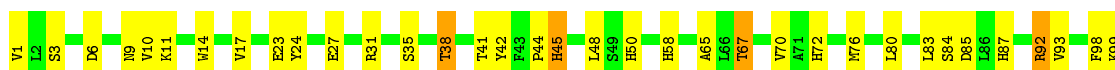




- Molecule 1: Hemoglobin subunit alpha



- Molecule 1: Hemoglobin subunit alpha



- Molecule 2: Hemoglobin subunit beta

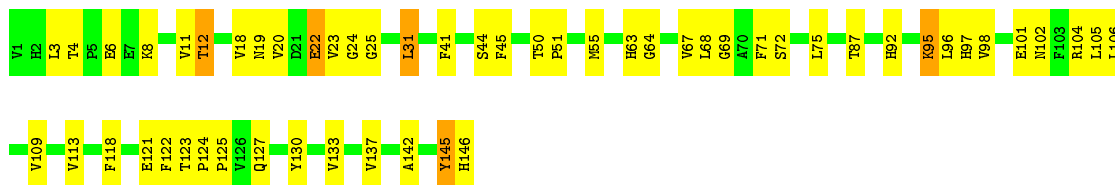


- Molecule 2: Hemoglobin subunit beta

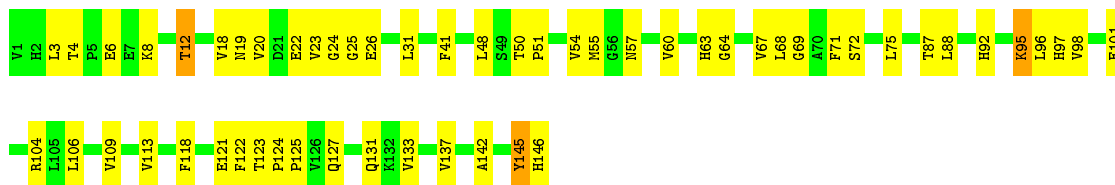


- Molecule 2: Hemoglobin subunit beta

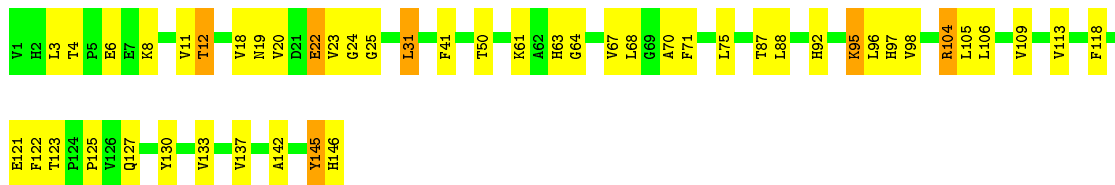




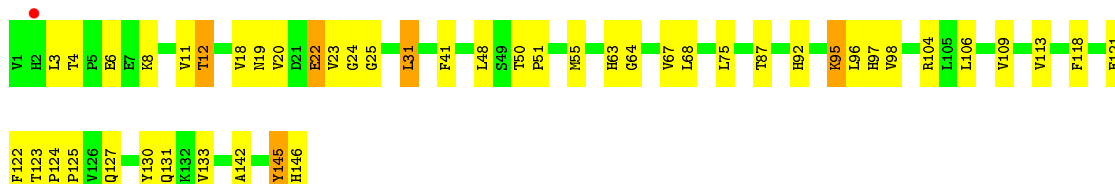
• Molecule 2: Hemoglobin subunit beta



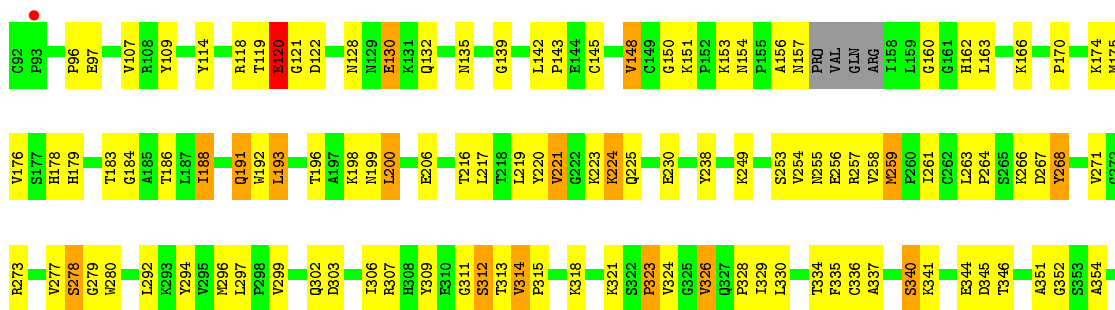
• Molecule 2: Hemoglobin subunit beta

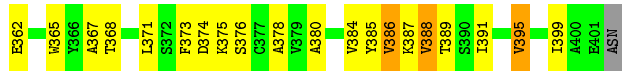


• Molecule 2: Hemoglobin subunit beta

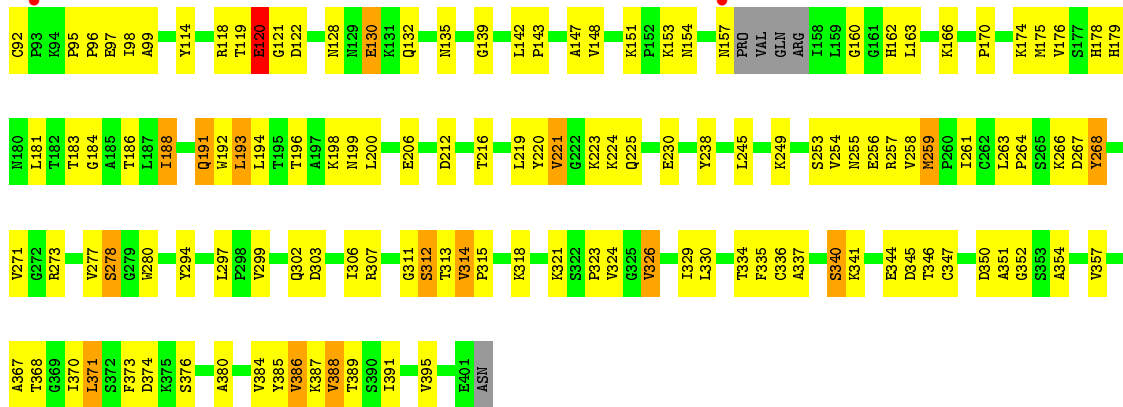


• Molecule 3: Haptoglobin

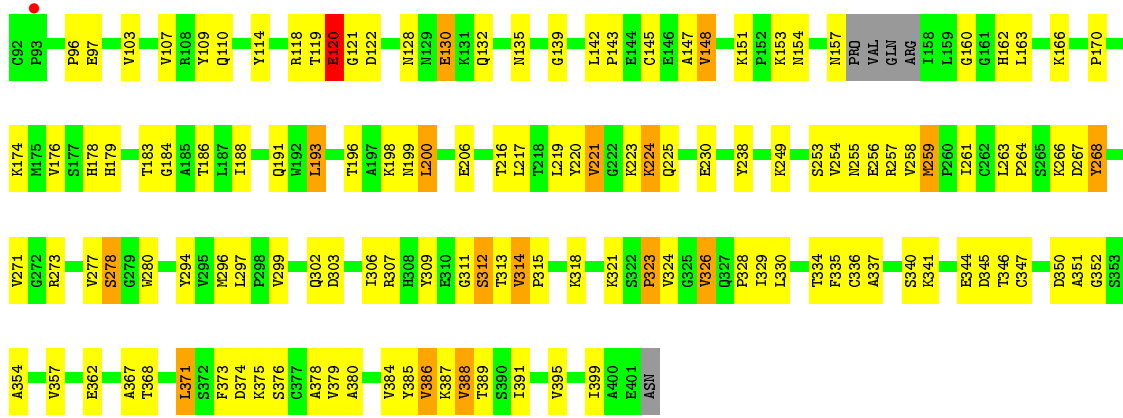




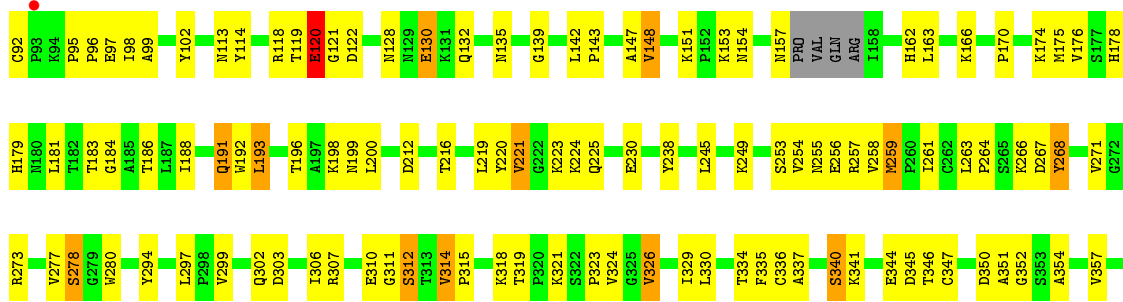
● Molecule 3: Haptoglobin

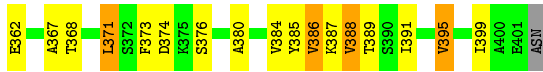


● Molecule 3: Haptoglobin

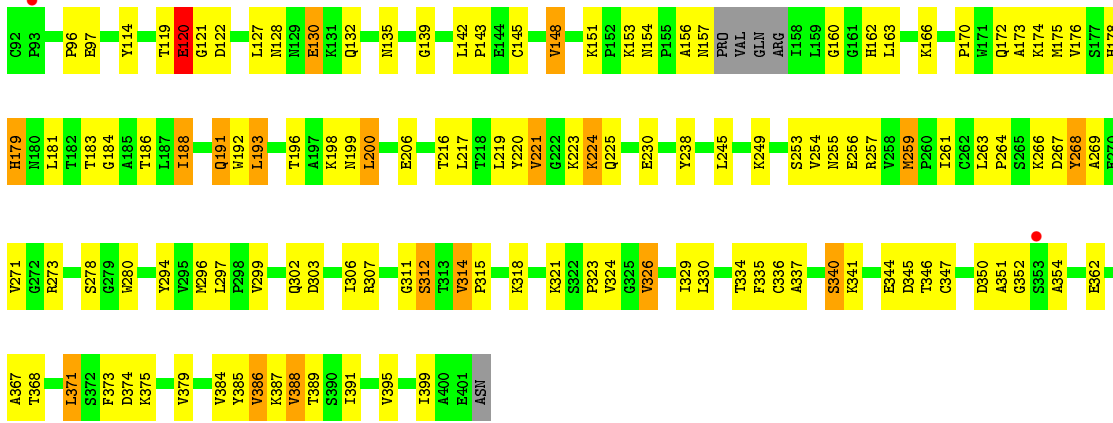


● Molecule 3: Haptoglobin

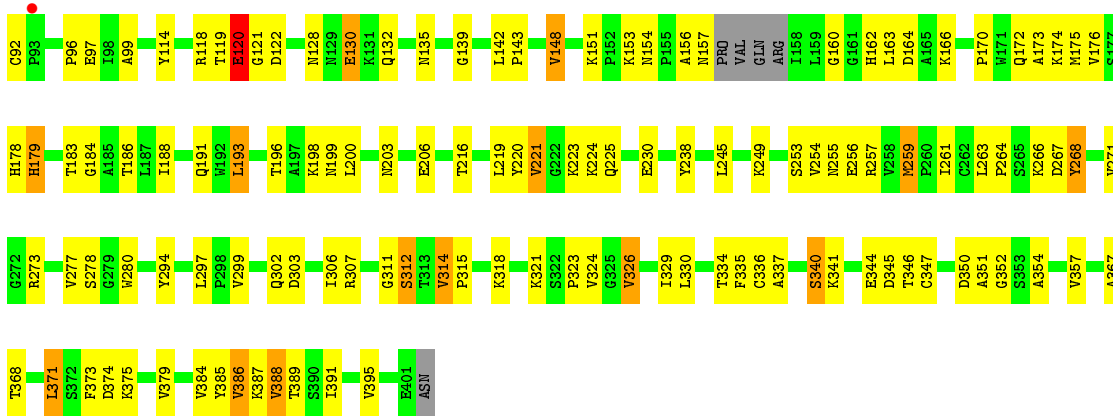
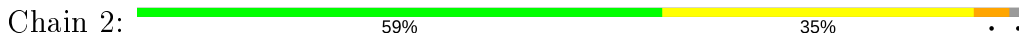




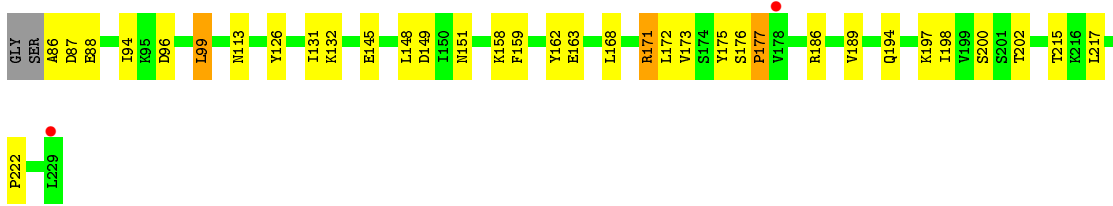
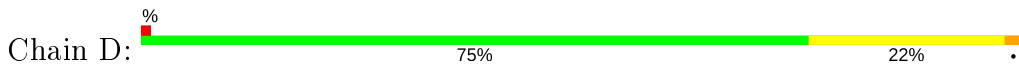
• Molecule 3: Haptoglobin



• Molecule 3: Haptoglobin



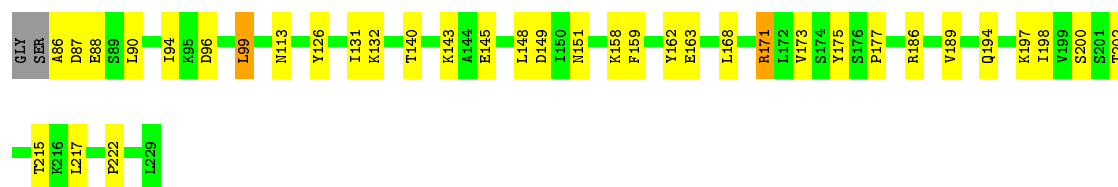
• Molecule 4: Iron-regulated surface determinant protein H




• Molecule 4: Iron-regulated surface determinant protein H

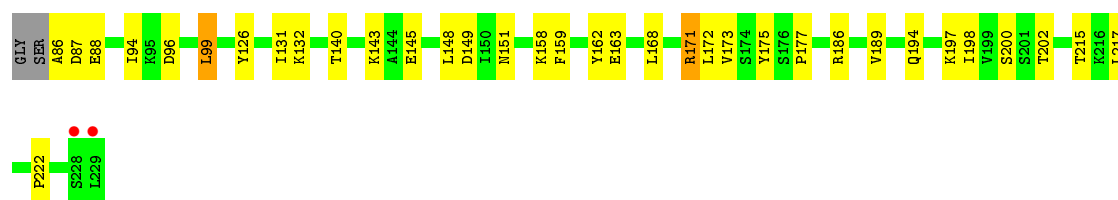


Chain I:  74% 23% ..



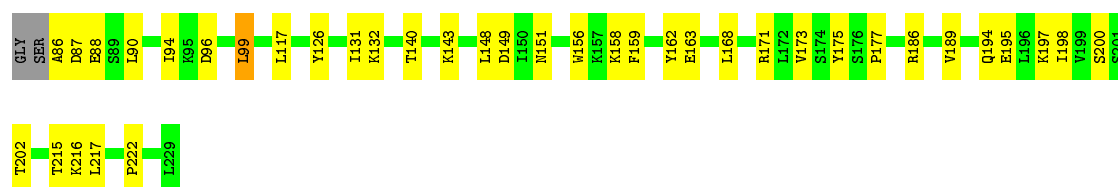
• Molecule 4: Iron-regulated surface determinant protein H

Chain N:  75% 23% ..




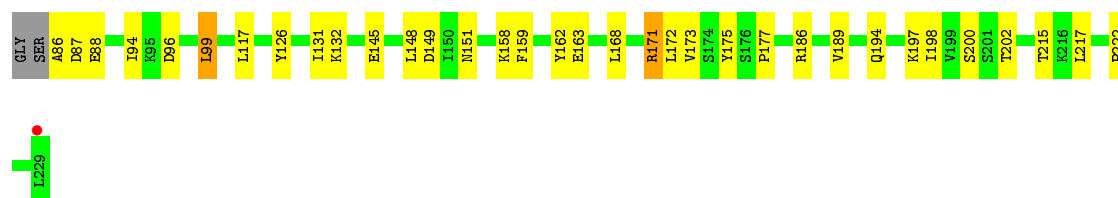
• Molecule 4: Iron-regulated surface determinant protein H

Chain S:  73% 25% ..



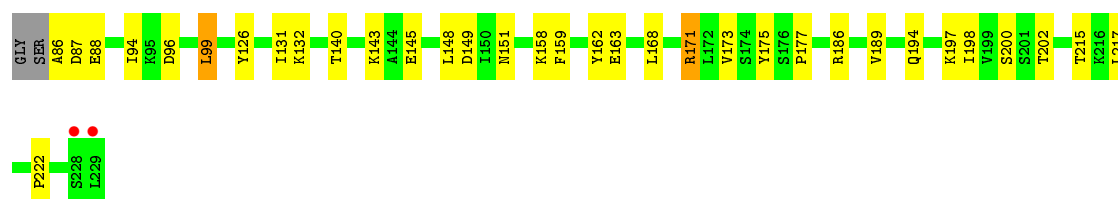
• Molecule 4: Iron-regulated surface determinant protein H

Chain X:  75% 22% ..

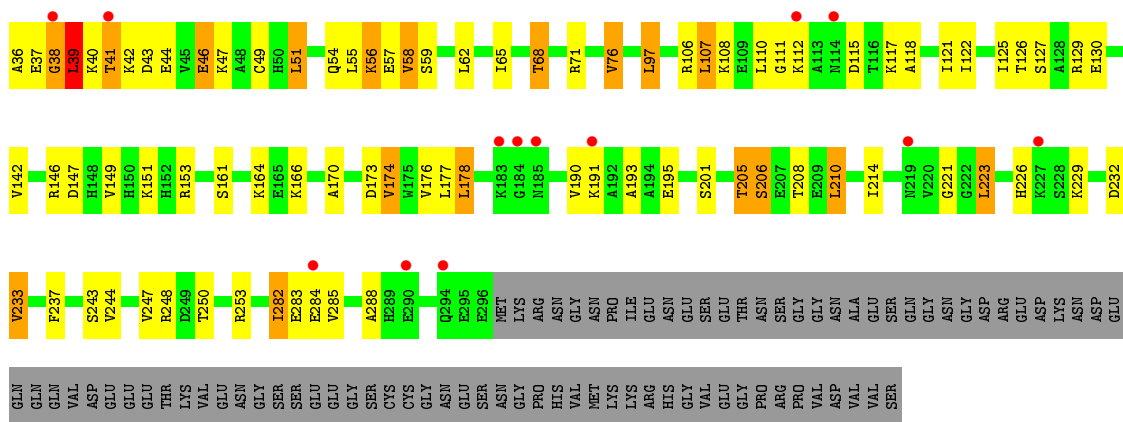


• Molecule 4: Iron-regulated surface determinant protein H

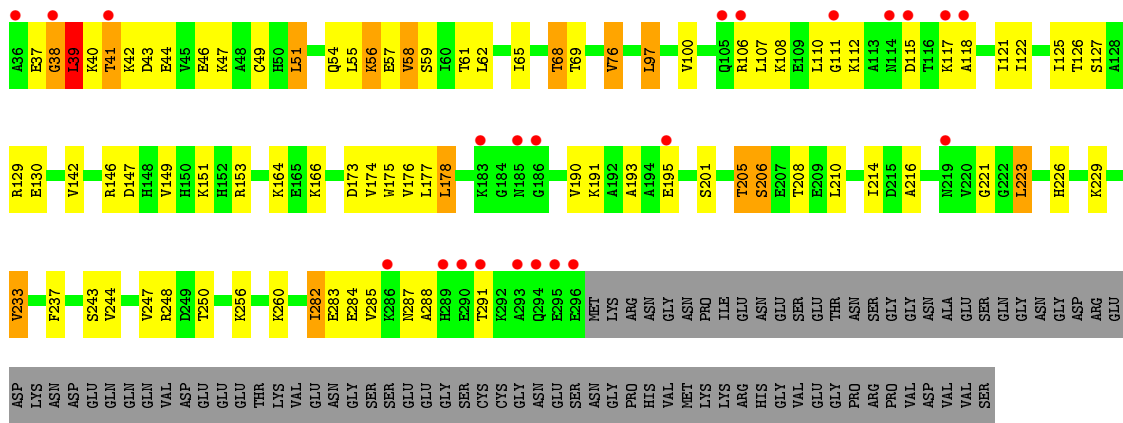
Chain 3:  75% 22% ..



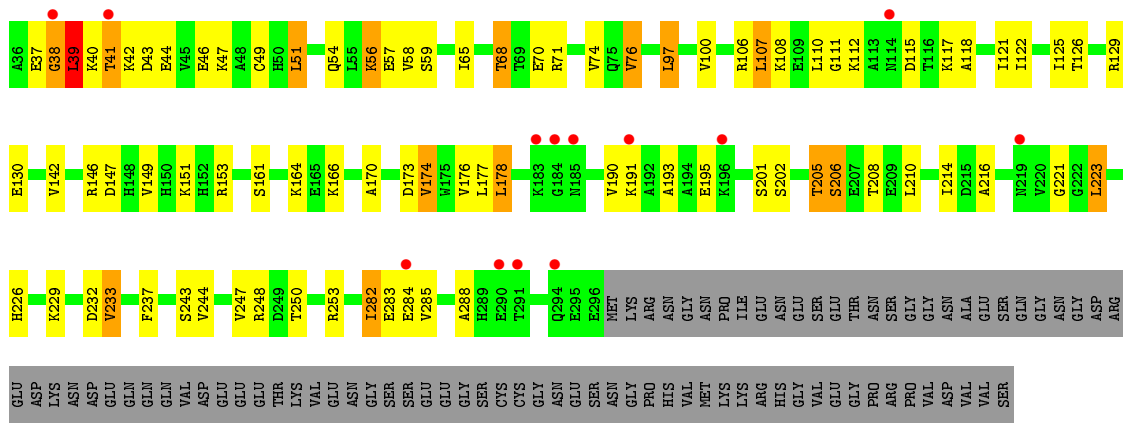
● Molecule 5: Haptoglobin-hemoglobin receptor



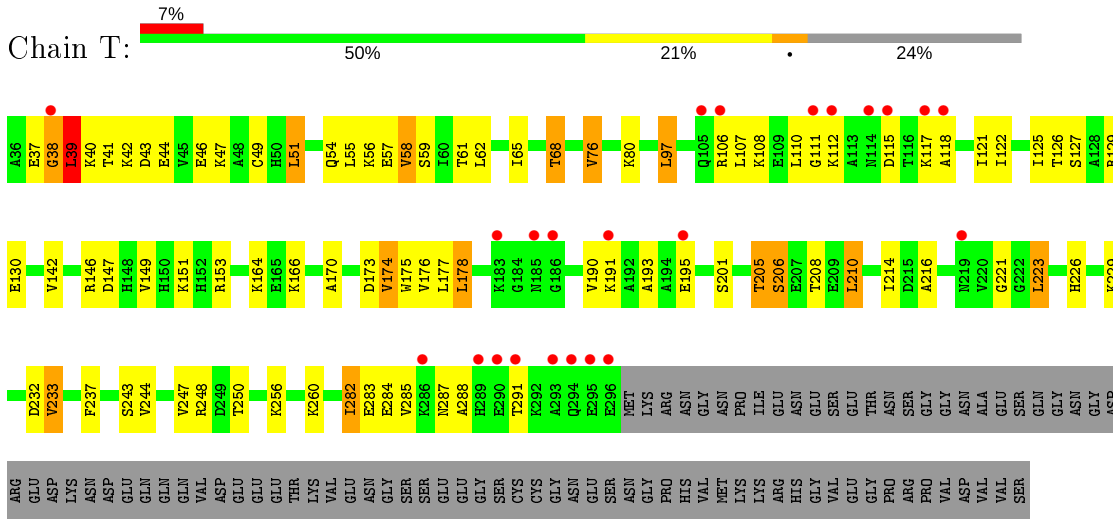
● Molecule 5: Haptoglobin-hemoglobin receptor



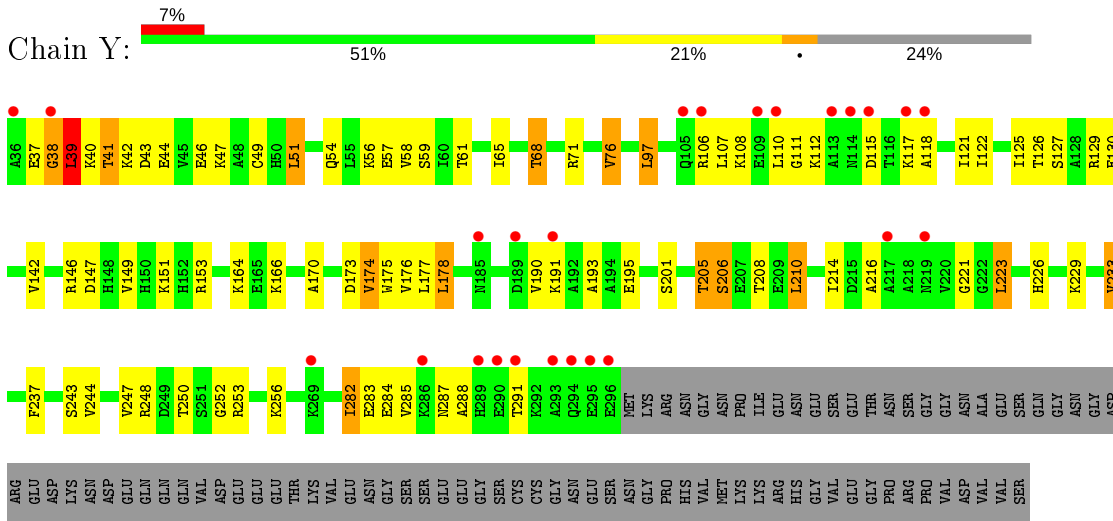
● Molecule 5: Haptoglobin-hemoglobin receptor



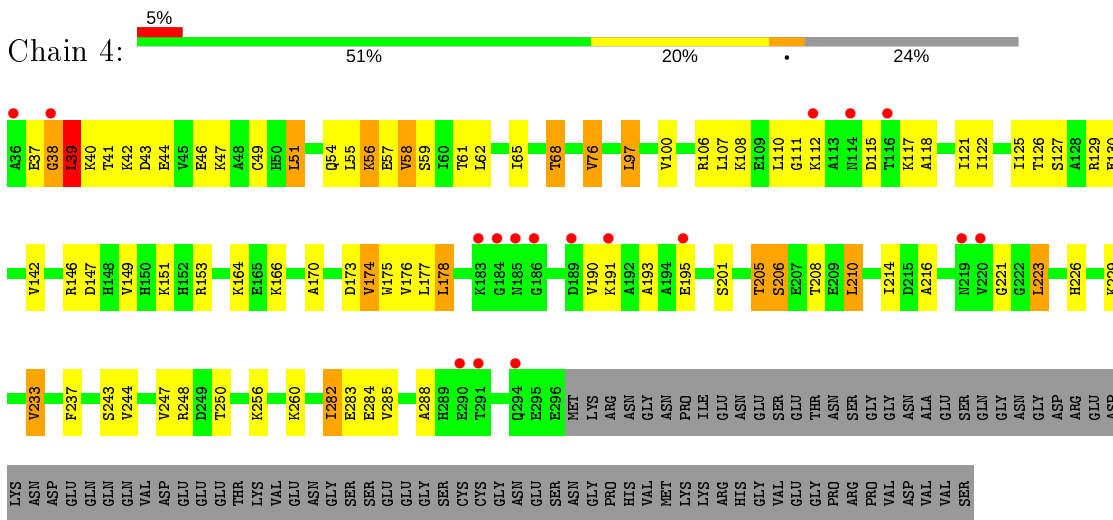
• Molecule 5: Haptoglobin-hemoglobin receptor




• Molecule 5: Haptoglobin-hemoglobin receptor



• Molecule 5: Haptoglobin-hemoglobin receptor



- Molecule 6: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain a:  100%

■ NAG1  
■ NAG2

- Molecule 6: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain b:  50% 50%


■ NAG1  
■ NAG2

- Molecule 6: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain c:  100%

■ NAG1  
■ NAG2

- Molecule 6: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain d:  100%

■ NAG1  
■ NAG2

- Molecule 6: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain e:  100%

■ NAG1  
■ NAG2

- Molecule 6: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain f:  100%

■ NAG1  
■ NAG2

- Molecule 6: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain g:  100%

MAG1  
MAG2

- Molecule 6: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain h:  100%MAG1  
MAG2

- Molecule 6: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain i:  100%MAG1  
MAG2

- Molecule 6: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain j:  100%MAG1  
MAG2

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	143.23Å 140.95Å 267.18Å 90.00° 98.54° 90.00°	Depositor
Resolution (Å)	29.00 – 3.10 28.98 – 3.10	Depositor EDS
% Data completeness (in resolution range)	96.1 (29.00-3.10) 96.7 (28.98-3.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.80 (at 3.11Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: dev_1702)	Depositor
R, $R_{free}$	0.255 , 0.271 0.260 , 0.273	Depositor DCC
$R_{free}$ test set	1758 reflections (0.96%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	65.7	Xtrriage
Anisotropy	0.870	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 38.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.27$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	47792	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	81.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 53.03 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 4.4672e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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, NAG, OXY

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	0.54	0/1097	0.66	0/1491
1	F	0.52	0/1097	0.69	2/1491 (0.1%)
1	K	0.54	0/1097	0.68	0/1491
1	P	0.51	0/1097	0.64	0/1491
1	U	0.46	0/1097	0.61	0/1491
1	Z	0.47	0/1097	0.61	0/1491
2	1	0.49	0/1153	0.61	0/1566
2	B	0.59	0/1153	0.67	0/1566
2	G	0.57	0/1153	0.65	0/1566
2	L	0.59	0/1153	0.68	0/1566
2	Q	0.56	0/1153	0.65	0/1566
2	V	0.51	0/1153	0.61	0/1566
3	2	0.52	0/2497	0.71	0/3391
3	C	0.64	0/2497	0.77	0/3391
3	H	0.57	0/2497	0.74	0/3391
3	M	0.64	0/2497	0.77	0/3391
3	R	0.57	0/2497	0.74	0/3391
3	W	0.49	0/2497	0.71	0/3391
4	3	0.44	0/1212	0.62	0/1647
4	D	0.51	0/1212	0.66	0/1647
4	I	0.49	0/1212	0.65	0/1647
4	N	0.52	0/1212	0.66	0/1647
4	S	0.49	0/1212	0.65	0/1647
4	X	0.43	0/1212	0.63	0/1647
5	4	0.41	0/1981	0.62	1/2666 (0.0%)
5	E	0.45	0/1981	0.63	1/2666 (0.0%)
5	J	0.42	0/1981	0.62	1/2666 (0.0%)
5	O	0.48	0/1981	0.64	1/2666 (0.0%)
5	T	0.42	0/1981	0.62	1/2666 (0.0%)
5	Y	0.39	0/1981	0.61	1/2666 (0.0%)
All	All	0.52	0/47640	0.67	8/64566 (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	2	0	3
3	C	0	4
3	H	0	3
3	M	0	4
3	R	0	3
3	W	0	3
4	3	0	1
4	D	0	1
4	I	0	1
4	N	0	1
4	S	0	1
4	X	0	1
5	4	0	1
5	E	0	1
5	J	0	1
5	O	0	1
5	T	0	1
5	Y	0	1
All	All	0	32

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	1	VAL	CG1-CB-CG2	6.10	120.66	110.90
1	F	1	VAL	CB-CA-C	-5.41	101.12	111.40
5	E	223	LEU	CA-CB-CG	5.25	127.38	115.30
5	T	223	LEU	CA-CB-CG	5.12	127.07	115.30
5	4	223	LEU	CA-CB-CG	5.09	127.01	115.30

There are no chirality outliers.

5 of 32 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	C	120	GLU	Peptide
3	C	311	GLY	Peptide
3	C	312	SER	Peptide

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Group
3	C	378	ALA	Peptide
4	D	177	PRO	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	1069	0	1073	30	0
1	F	1069	0	1073	27	0
1	K	1069	0	1073	30	0
1	P	1069	0	1073	24	0
1	U	1069	0	1073	33	0
1	Z	1069	0	1073	31	0
2	1	1123	0	1118	32	0
2	B	1123	0	1118	37	0
2	G	1123	0	1118	30	0
2	L	1123	0	1118	39	0
2	Q	1123	0	1118	34	0
2	V	1123	0	1118	33	0
3	2	2437	0	2388	93	0
3	C	2437	0	2387	103	0
3	H	2437	0	2388	99	0
3	M	2437	0	2387	102	0
3	R	2437	0	2387	102	0
3	W	2437	0	2388	94	0
4	3	1183	0	1132	16	0
4	D	1183	0	1132	19	0
4	I	1183	0	1132	17	0
4	N	1183	0	1132	18	0
4	S	1183	0	1132	17	0
4	X	1183	0	1132	17	0
5	4	1963	0	1978	57	0
5	E	1963	0	1978	65	0
5	J	1963	0	1978	60	0
5	O	1963	0	1978	58	0
5	T	1963	0	1978	58	0
5	Y	1963	0	1978	62	0
6	a	28	0	25	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	b	28	0	25	0	0
6	c	28	0	25	0	0
6	d	28	0	25	0	0
6	e	28	0	25	0	0
6	f	28	0	25	0	0
6	g	28	0	25	0	0
6	h	28	0	25	0	0
6	i	28	0	25	0	0
6	j	28	0	25	0	0
7	1	43	0	30	6	0
7	A	43	0	30	1	0
7	B	43	0	30	6	0
7	F	43	0	30	2	0
7	G	43	0	30	3	0
7	K	43	0	30	1	0
7	L	43	0	30	6	0
7	P	43	0	30	2	0
7	Q	43	0	30	4	0
7	U	43	0	30	4	0
7	V	43	0	30	3	0
7	Z	43	0	30	2	0
8	1	2	0	0	0	0
8	A	2	0	0	0	0
8	B	2	0	0	0	0
8	F	2	0	0	0	0
8	G	2	0	0	0	0
8	K	2	0	0	0	0
8	L	2	0	0	0	0
8	P	2	0	0	0	0
8	Q	2	0	0	0	0
8	U	2	0	0	1	0
8	V	2	0	0	1	0
8	Z	2	0	0	1	0
9	2	14	0	13	0	0
9	4	28	0	26	2	0
9	C	28	0	26	0	0
9	E	28	0	26	1	0
9	H	28	0	26	0	0
9	J	28	0	26	2	0
9	M	28	0	26	1	0
9	O	28	0	26	1	0
9	R	28	0	26	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
9	T	28	0	26	1	0
9	W	28	0	26	1	0
9	Y	28	0	26	1	0
All	All	47792	0	47040	1392	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 15.

The worst 5 of 1392 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)
5:E:43:ASP:OD2	5:E:47:LYS:HE2	1.30	1.29
5:E:43:ASP:OD2	5:E:47:LYS:CE	2.15	0.94
3:2:183:THR:HB	3:2:199:ASN:HD22	1.36	0.89
3:M:183:THR:HB	3:M:199:ASN:HD22	1.35	0.89
5:O:57:GLU:HG3	5:O:201:SER:HB3	1.55	0.88

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%)	130 (94%)	9 (6%)	0	100	100
1	F	139/141 (99%)	129 (93%)	10 (7%)	0	100	100
1	K	139/141 (99%)	129 (93%)	10 (7%)	0	100	100
1	P	139/141 (99%)	128 (92%)	11 (8%)	0	100	100
1	U	139/141 (99%)	130 (94%)	9 (6%)	0	100	100
1	Z	139/141 (99%)	128 (92%)	11 (8%)	0	100	100
2	1	144/146 (99%)	141 (98%)	3 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	144/146 (99%)	142 (99%)	2 (1%)	0	100	100
2	G	144/146 (99%)	140 (97%)	4 (3%)	0	100	100
2	L	144/146 (99%)	142 (99%)	2 (1%)	0	100	100
2	Q	144/146 (99%)	141 (98%)	3 (2%)	0	100	100
2	V	144/146 (99%)	142 (99%)	2 (1%)	0	100	100
3	2	306/315 (97%)	283 (92%)	23 (8%)	0	100	100
3	C	306/315 (97%)	285 (93%)	21 (7%)	0	100	100
3	H	306/315 (97%)	283 (92%)	23 (8%)	0	100	100
3	M	306/315 (97%)	282 (92%)	24 (8%)	0	100	100
3	R	306/315 (97%)	283 (92%)	23 (8%)	0	100	100
3	W	306/315 (97%)	283 (92%)	23 (8%)	0	100	100
4	3	142/146 (97%)	135 (95%)	7 (5%)	0	100	100
4	D	142/146 (97%)	134 (94%)	8 (6%)	0	100	100
4	I	142/146 (97%)	135 (95%)	7 (5%)	0	100	100
4	N	142/146 (97%)	134 (94%)	8 (6%)	0	100	100
4	S	142/146 (97%)	134 (94%)	8 (6%)	0	100	100
4	X	142/146 (97%)	135 (95%)	7 (5%)	0	100	100
5	4	259/343 (76%)	232 (90%)	26 (10%)	1 (0%)	34	69
5	E	259/343 (76%)	231 (89%)	27 (10%)	1 (0%)	34	69
5	J	259/343 (76%)	234 (90%)	24 (9%)	1 (0%)	34	69
5	O	259/343 (76%)	232 (90%)	26 (10%)	1 (0%)	34	69
5	T	259/343 (76%)	233 (90%)	25 (10%)	1 (0%)	34	69
5	Y	259/343 (76%)	232 (90%)	26 (10%)	1 (0%)	34	69
All	All	5940/6546 (91%)	5522 (93%)	412 (7%)	6 (0%)	51	83

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	E	39	LEU
5	J	39	LEU
5	O	39	LEU
5	T	39	LEU
5	Y	39	LEU

### 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	113/113 (100%)	100 (88%)	13 (12%)	5	22
1	F	113/113 (100%)	100 (88%)	13 (12%)	5	22
1	K	113/113 (100%)	100 (88%)	13 (12%)	5	22
1	P	113/113 (100%)	100 (88%)	13 (12%)	5	22
1	U	113/113 (100%)	101 (89%)	12 (11%)	6	26
1	Z	113/113 (100%)	100 (88%)	13 (12%)	5	22
2	1	118/118 (100%)	106 (90%)	12 (10%)	7	27
2	B	118/118 (100%)	107 (91%)	11 (9%)	9	32
2	G	118/118 (100%)	107 (91%)	11 (9%)	9	32
2	L	118/118 (100%)	106 (90%)	12 (10%)	7	27
2	Q	118/118 (100%)	108 (92%)	10 (8%)	10	37
2	V	118/118 (100%)	106 (90%)	12 (10%)	7	27
3	2	266/271 (98%)	238 (90%)	28 (10%)	7	26
3	C	266/271 (98%)	237 (89%)	29 (11%)	6	25
3	H	266/271 (98%)	237 (89%)	29 (11%)	6	25
3	M	266/271 (98%)	235 (88%)	31 (12%)	5	22
3	R	266/271 (98%)	235 (88%)	31 (12%)	5	22
3	W	266/271 (98%)	236 (89%)	30 (11%)	6	23
4	3	133/134 (99%)	129 (97%)	4 (3%)	41	71
4	D	133/134 (99%)	129 (97%)	4 (3%)	41	71
4	I	133/134 (99%)	127 (96%)	6 (4%)	27	60
4	N	133/134 (99%)	129 (97%)	4 (3%)	41	71
4	S	133/134 (99%)	128 (96%)	5 (4%)	33	66
4	X	133/134 (99%)	128 (96%)	5 (4%)	33	66
5	4	202/272 (74%)	180 (89%)	22 (11%)	6	25
5	E	202/272 (74%)	178 (88%)	24 (12%)	5	20

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
5	J	202/272 (74%)	180 (89%)	22 (11%)	6	25
5	O	202/272 (74%)	179 (89%)	23 (11%)	5	23
5	T	202/272 (74%)	180 (89%)	22 (11%)	6	25
5	Y	202/272 (74%)	180 (89%)	22 (11%)	6	25
All	All	4992/5448 (92%)	4506 (90%)	486 (10%)	8	30

5 of 486 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	N	171	ARG
3	R	148	VAL
3	2	278	SER
5	O	56	LYS
1	P	23	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
3	M	327	GLN
2	Q	131	GLN
3	2	100	HIS
4	N	113	ASN
5	O	54	GLN

### 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](#)

20 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
6	NAG	a	1	3,6	14,14,15	1.14	1 (7%)	17,19,21	1.74	2 (11%)
6	NAG	a	2	6	14,14,15	1.14	1 (7%)	17,19,21	1.20	1 (5%)
6	NAG	b	1	3,6	14,14,15	0.59	0	17,19,21	2.02	5 (29%)
6	NAG	b	2	6	14,14,15	0.24	0	17,19,21	0.63	0
6	NAG	c	1	3,6	14,14,15	1.01	1 (7%)	17,19,21	1.99	5 (29%)
6	NAG	c	2	6	14,14,15	0.58	0	17,19,21	0.81	1 (5%)
6	NAG	d	1	3,6	14,14,15	0.49	0	17,19,21	2.08	2 (11%)
6	NAG	d	2	6	14,14,15	1.19	2 (14%)	17,19,21	1.13	1 (5%)
6	NAG	e	1	3,6	14,14,15	0.93	1 (7%)	17,19,21	1.66	5 (29%)
6	NAG	e	2	6	14,14,15	0.33	0	17,19,21	0.77	1 (5%)
6	NAG	f	1	3,6	14,14,15	0.53	0	17,19,21	1.59	2 (11%)
6	NAG	f	2	6	14,14,15	1.28	2 (14%)	17,19,21	1.14	1 (5%)
6	NAG	g	1	3,6	14,14,15	0.83	1 (7%)	17,19,21	1.67	4 (23%)
6	NAG	g	2	6	14,14,15	0.54	0	17,19,21	0.82	1 (5%)
6	NAG	h	1	3,6	14,14,15	0.78	1 (7%)	17,19,21	1.92	5 (29%)
6	NAG	h	2	6	14,14,15	0.34	0	17,19,21	0.77	1 (5%)
6	NAG	i	1	3,6	14,14,15	0.36	0	17,19,21	1.32	1 (5%)
6	NAG	i	2	6	14,14,15	0.78	1 (7%)	17,19,21	0.87	1 (5%)
6	NAG	j	1	3,6	14,14,15	0.83	1 (7%)	17,19,21	1.77	5 (29%)
6	NAG	j	2	6	14,14,15	0.51	0	17,19,21	0.86	1 (5%)

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
6	NAG	a	1	3,6	-	2/6/23/26	0/1/1/1
6	NAG	a	2	6	-	1/6/23/26	0/1/1/1
6	NAG	b	1	3,6	-	3/6/23/26	0/1/1/1
6	NAG	b	2	6	-	3/6/23/26	0/1/1/1
6	NAG	c	1	3,6	-	3/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	NAG	c	2	6	-	3/6/23/26	0/1/1/1
6	NAG	d	1	3,6	-	3/6/23/26	0/1/1/1
6	NAG	d	2	6	-	2/6/23/26	0/1/1/1
6	NAG	e	1	3,6	-	3/6/23/26	0/1/1/1
6	NAG	e	2	6	-	2/6/23/26	0/1/1/1
6	NAG	f	1	3,6	-	3/6/23/26	0/1/1/1
6	NAG	f	2	6	-	2/6/23/26	0/1/1/1
6	NAG	g	1	3,6	-	4/6/23/26	0/1/1/1
6	NAG	g	2	6	-	3/6/23/26	0/1/1/1
6	NAG	h	1	3,6	-	3/6/23/26	0/1/1/1
6	NAG	h	2	6	-	4/6/23/26	0/1/1/1
6	NAG	i	1	3,6	-	3/6/23/26	0/1/1/1
6	NAG	i	2	6	-	0/6/23/26	0/1/1/1
6	NAG	j	1	3,6	-	3/6/23/26	0/1/1/1
6	NAG	j	2	6	-	2/6/23/26	0/1/1/1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	f	2	NAG	O5-C1	4.21	1.50	1.43
6	a	1	NAG	O5-C1	3.90	1.50	1.43
6	a	2	NAG	O5-C1	3.72	1.49	1.43
6	d	2	NAG	O5-C1	3.48	1.49	1.43
6	c	1	NAG	O5-C1	3.27	1.48	1.43

The worst 5 of 45 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	d	1	NAG	C1-O5-C5	7.29	122.06	112.19
6	a	1	NAG	C1-O5-C5	5.81	120.06	112.19
6	b	1	NAG	C1-O5-C5	5.15	119.17	112.19
6	f	1	NAG	C1-O5-C5	5.09	119.09	112.19
6	c	1	NAG	C1-O5-C5	4.61	118.43	112.19

There are no chirality outliers.

5 of 52 torsion outliers are listed below:



Mol	Chain	Res	Type	Atoms
6	g	2	NAG	C3-C2-N2-C7
6	b	2	NAG	C1-C2-N2-C7
6	a	1	NAG	C3-C2-N2-C7
6	d	1	NAG	C3-C2-N2-C7
6	i	1	NAG	C3-C2-N2-C7

There are no ring outliers.

No monomer is involved in short contacts.

## 5.6 Ligand geometry [i](#)

47 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
9	NAG	Y	1002	5	14,14,15	0.82	1 (7%)	17,19,21	0.80	1 (5%)
8	OXY	1	202	7	1,1,1	0.15	0	-		
7	HEM	A	201	1,8	27,50,50	2.04	5 (18%)	17,82,82	1.85	5 (29%)
9	NAG	4	1002	5	14,14,15	0.84	1 (7%)	17,19,21	0.69	0
8	OXY	F	202	7	1,1,1	0.08	0	-		
9	NAG	2	1001	3	14,14,15	0.26	0	17,19,21	0.53	0
8	OXY	A	202	7	1,1,1	0.10	0	-		
7	HEM	1	201	8,2	27,50,50	2.45	6 (22%)	17,82,82	1.92	6 (35%)
8	OXY	P	202	7	1,1,1	0.12	0	-		
7	HEM	K	201	1,8	27,50,50	2.14	6 (22%)	17,82,82	2.41	7 (41%)
9	NAG	C	1001	3	14,14,15	0.44	0	17,19,21	0.57	0
9	NAG	E	1002	5	14,14,15	0.75	1 (7%)	17,19,21	1.17	1 (5%)
8	OXY	U	202	7	1,1,1	0.13	0	-		
8	OXY	Q	202	7	1,1,1	0.10	0	-		
9	NAG	O	1001	5	14,14,15	0.21	0	17,19,21	0.55	0
8	OXY	L	202	7	1,1,1	0.13	0	-		
9	NAG	W	1002	3	14,14,15	0.52	0	17,19,21	1.42	1 (5%)
9	NAG	C	1004	3	14,14,15	1.10	1 (7%)	17,19,21	0.99	1 (5%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
8	OXY	G	202	7	1,1,1	0.11	0	-		
9	NAG	T	1001	5	14,14,15	0.25	0	17,19,21	0.42	0
7	HEM	Q	201	8,2	27,50,50	2.37	6 (22%)	17,82,82	2.25	5 (29%)
9	NAG	E	1001	5	14,14,15	0.32	0	17,19,21	0.53	0
9	NAG	H	1001	3	14,14,15	0.68	1 (7%)	17,19,21	0.63	0
9	NAG	J	1001	5	14,14,15	0.39	0	17,19,21	0.53	0
9	NAG	H	1002	3	14,14,15	0.60	0	17,19,21	0.64	0
8	OXY	K	202	7	1,1,1	0.08	0	-		
9	NAG	W	1001	3	14,14,15	0.62	0	17,19,21	0.42	0
7	HEM	V	201	8,2	27,50,50	2.32	5 (18%)	17,82,82	1.78	3 (17%)
7	HEM	L	201	8,2	27,50,50	2.33	7 (25%)	17,82,82	2.48	7 (41%)
9	NAG	M	1001	3	14,14,15	0.94	1 (7%)	17,19,21	0.83	1 (5%)
7	HEM	B	201	8,2	27,50,50	2.30	6 (22%)	17,82,82	2.30	5 (29%)
9	NAG	4	1001	5	14,14,15	0.19	0	17,19,21	0.49	0
8	OXY	Z	202	7	1,1,1	0.11	0	-		
9	NAG	Y	1001	5	14,14,15	0.38	0	17,19,21	0.49	0
8	OXY	V	202	7	1,1,1	0.14	0	-		
9	NAG	M	1004	3	14,14,15	1.16	1 (7%)	17,19,21	0.93	1 (5%)
7	HEM	P	201	1,8	27,50,50	2.08	6 (22%)	17,82,82	1.85	5 (29%)
7	HEM	G	201	8,2	27,50,50	2.42	6 (22%)	17,82,82	1.96	6 (35%)
7	HEM	U	201	1,8	27,50,50	2.18	6 (22%)	17,82,82	1.53	3 (17%)
9	NAG	T	1002	5	14,14,15	0.62	1 (7%)	17,19,21	1.05	1 (5%)
7	HEM	F	201	1,8	27,50,50	2.16	6 (22%)	17,82,82	1.95	4 (23%)
9	NAG	R	1001	3	14,14,15	0.40	0	17,19,21	0.45	0
9	NAG	J	1002	5	14,14,15	0.69	1 (7%)	17,19,21	0.45	0
9	NAG	O	1002	5	14,14,15	0.97	1 (7%)	17,19,21	0.72	0
7	HEM	Z	201	1,8	27,50,50	2.17	6 (22%)	17,82,82	1.81	7 (41%)
9	NAG	R	1004	3	14,14,15	0.37	0	17,19,21	0.46	0
8	OXY	B	202	7	1,1,1	0.11	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
9	NAG	Y	1002	5	-	1/6/23/26	0/1/1/1
7	HEM	A	201	1,8	-	2/6/54/54	-
9	NAG	4	1002	5	-	2/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
9	NAG	2	1001	3	-	0/6/23/26	0/1/1/1
7	HEM	F	201	1,8	-	3/6/54/54	-
7	HEM	1	201	8,2	-	3/6/54/54	-
7	HEM	K	201	1,8	-	1/6/54/54	-
9	NAG	C	1001	3	-	2/6/23/26	0/1/1/1
9	NAG	E	1002	5	-	1/6/23/26	0/1/1/1
9	NAG	O	1001	5	-	0/6/23/26	0/1/1/1
9	NAG	W	1002	3	-	3/6/23/26	0/1/1/1
9	NAG	C	1004	3	-	0/6/23/26	0/1/1/1
9	NAG	T	1001	5	-	2/6/23/26	0/1/1/1
7	HEM	Q	201	8,2	-	1/6/54/54	-
9	NAG	E	1001	5	-	0/6/23/26	0/1/1/1
9	NAG	H	1001	3	-	1/6/23/26	0/1/1/1
9	NAG	J	1001	5	-	2/6/23/26	0/1/1/1
9	NAG	H	1002	3	-	1/6/23/26	0/1/1/1
9	NAG	W	1001	3	-	2/6/23/26	0/1/1/1
7	HEM	V	201	8,2	-	0/6/54/54	-
7	HEM	L	201	8,2	-	5/6/54/54	-
9	NAG	M	1001	3	-	2/6/23/26	0/1/1/1
7	HEM	B	201	8,2	-	3/6/54/54	-
9	NAG	4	1001	5	-	0/6/23/26	0/1/1/1
9	NAG	Y	1001	5	-	2/6/23/26	0/1/1/1
9	NAG	M	1004	3	-	0/6/23/26	0/1/1/1
7	HEM	P	201	1,8	-	1/6/54/54	-
7	HEM	G	201	8,2	-	0/6/54/54	-
7	HEM	U	201	1,8	-	0/6/54/54	-
9	NAG	T	1002	5	-	1/6/23/26	0/1/1/1
9	NAG	R	1001	3	-	1/6/23/26	0/1/1/1
9	NAG	J	1002	5	-	2/6/23/26	0/1/1/1
9	NAG	O	1002	5	-	2/6/23/26	0/1/1/1
7	HEM	Z	201	1,8	-	4/6/54/54	-
9	NAG	R	1004	3	-	0/6/23/26	0/1/1/1

The worst 5 of 81 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	G	201	HEM	C3C-C2C	-6.58	1.31	1.40
7	1	201	HEM	C3B-C2B	-6.15	1.31	1.40
7	Q	201	HEM	C3C-C2C	-6.15	1.31	1.40

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	Q	201	HEM	C3B-C2B	-6.11	1.31	1.40
7	V	201	HEM	C3D-C2D	6.01	1.55	1.37

The worst 5 of 70 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	Q	201	HEM	CAA-CBA-CGA	-6.09	102.45	112.67
7	L	201	HEM	CAD-CBD-CGD	-5.71	103.09	112.67
7	B	201	HEM	CAA-CBA-CGA	-5.59	103.30	112.67
7	B	201	HEM	CAD-CBD-CGD	-4.98	104.31	112.67
7	F	201	HEM	CMA-C3A-C4A	-4.63	121.36	128.46

There are no chirality outliers.

5 of 50 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	A	201	HEM	C2D-C3D-CAD-CBD
7	A	201	HEM	C4D-C3D-CAD-CBD
7	1	201	HEM	C1A-C2A-CAA-CBA
7	1	201	HEM	C3A-C2A-CAA-CBA
7	1	201	HEM	C2A-CAA-CBA-CGA

There are no ring outliers.

23 monomers are involved in 53 short contacts:

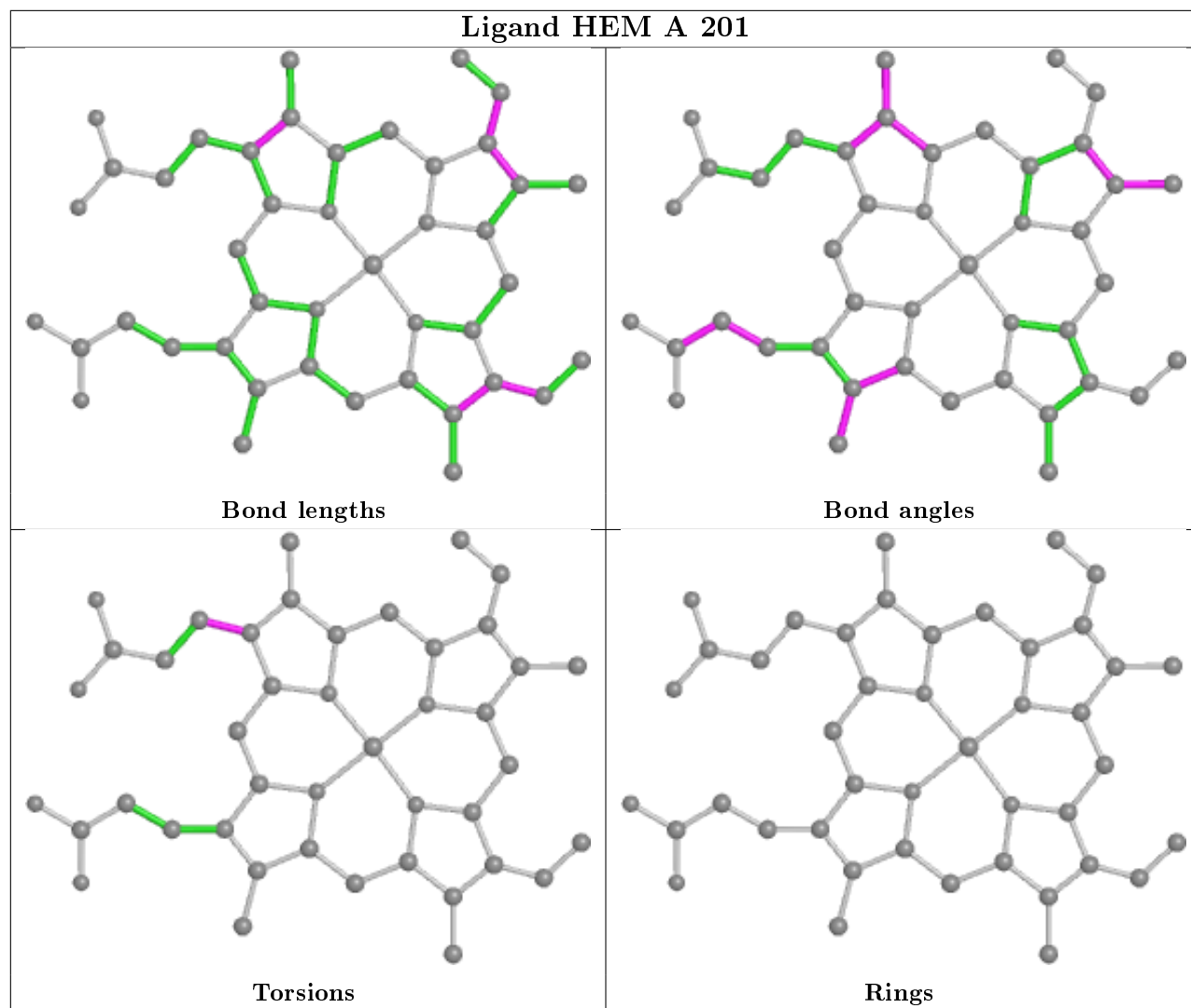
Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	Y	1002	NAG	1	0
7	A	201	HEM	1	0
9	4	1002	NAG	2	0
7	1	201	HEM	6	0
7	K	201	HEM	1	0
9	E	1002	NAG	1	0
8	U	202	OXY	1	0
9	W	1002	NAG	1	0
7	Q	201	HEM	4	0
7	V	201	HEM	3	0
7	L	201	HEM	6	0
9	M	1001	NAG	1	0
7	B	201	HEM	6	0
8	Z	202	OXY	1	0
8	V	202	OXY	1	0

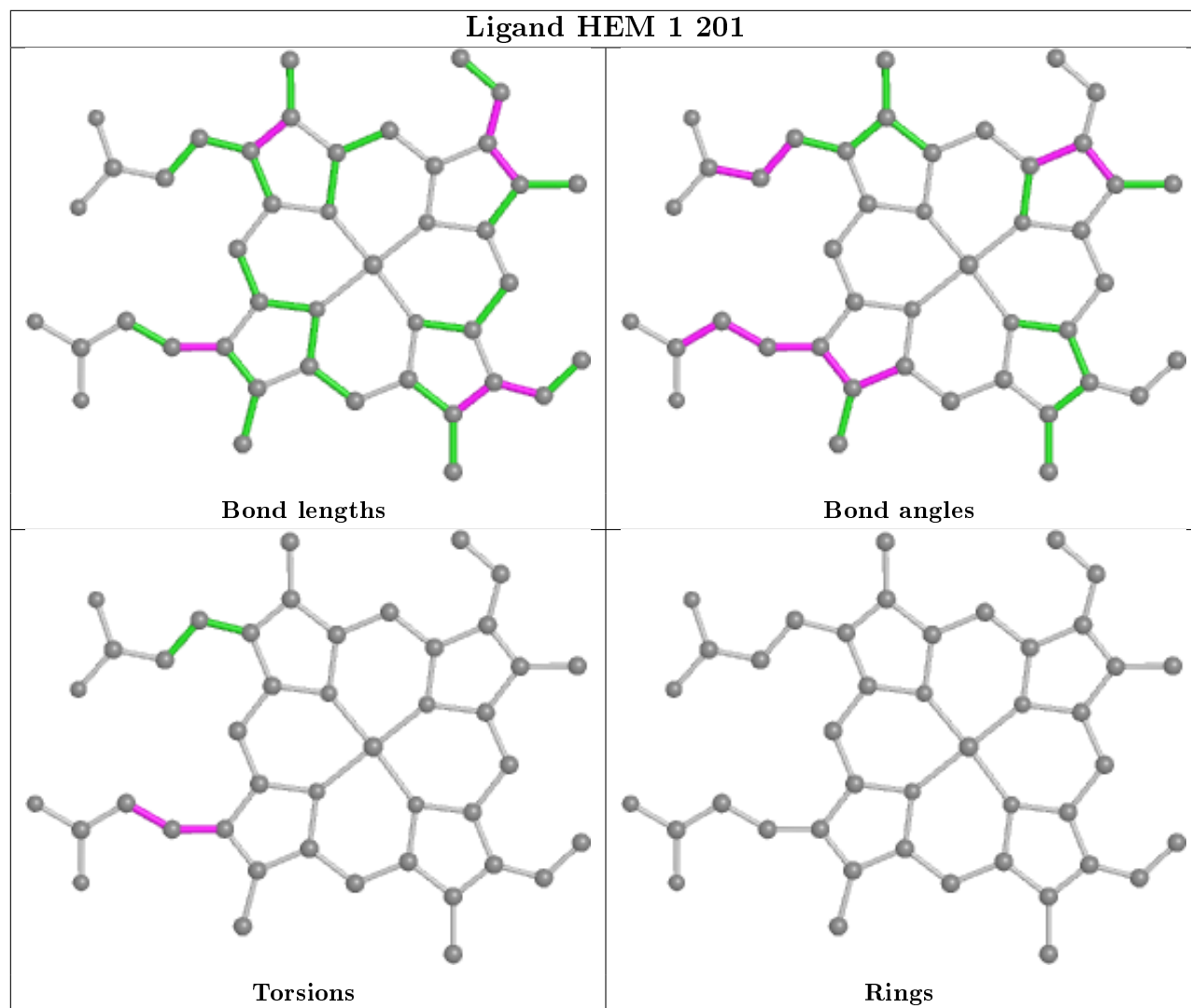
*Continued on next page...*

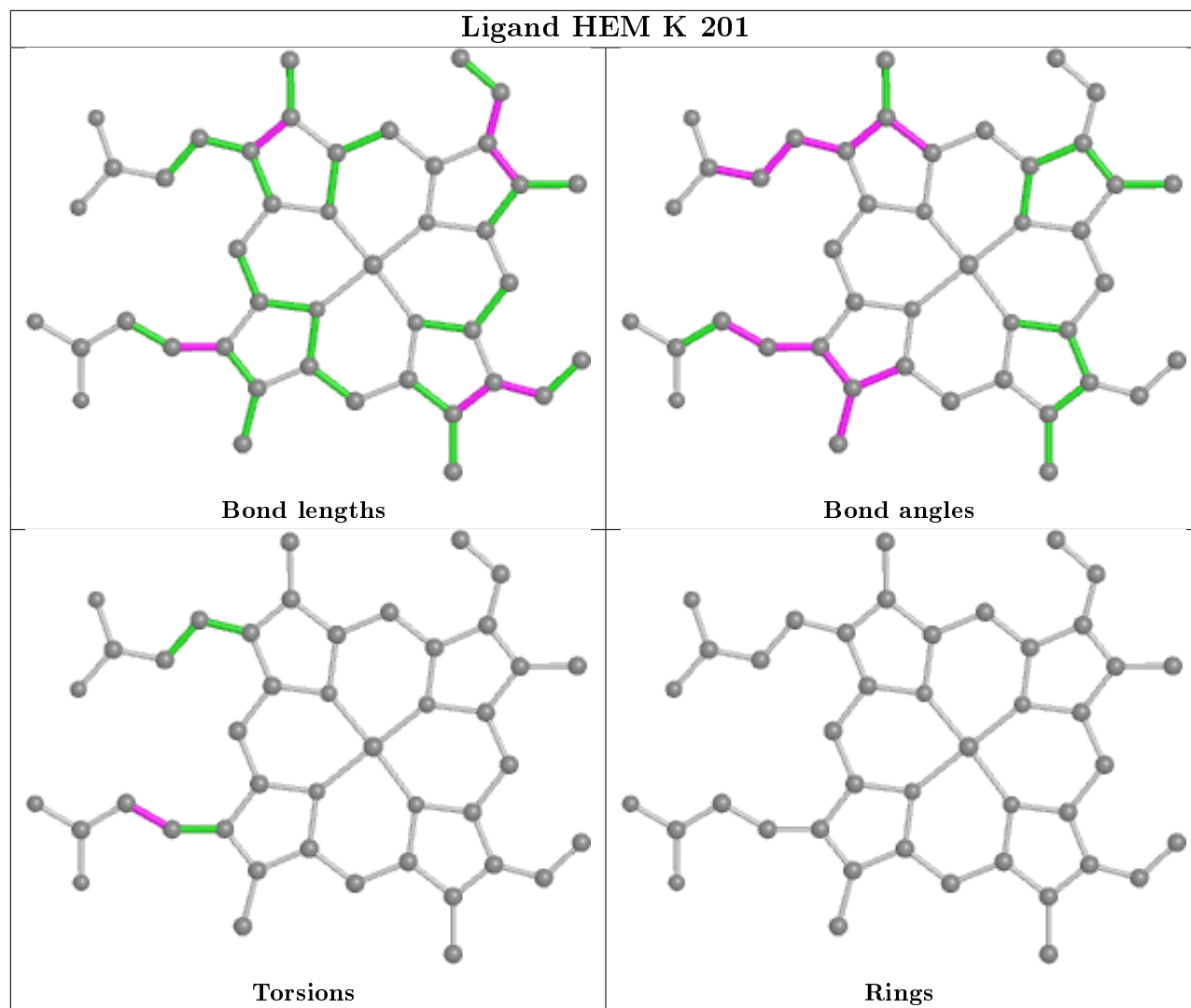
*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	P	201	HEM	2	0
7	G	201	HEM	3	0
7	U	201	HEM	4	0
9	T	1002	NAG	1	0
7	F	201	HEM	2	0
9	J	1002	NAG	2	0
9	O	1002	NAG	1	0
7	Z	201	HEM	2	0

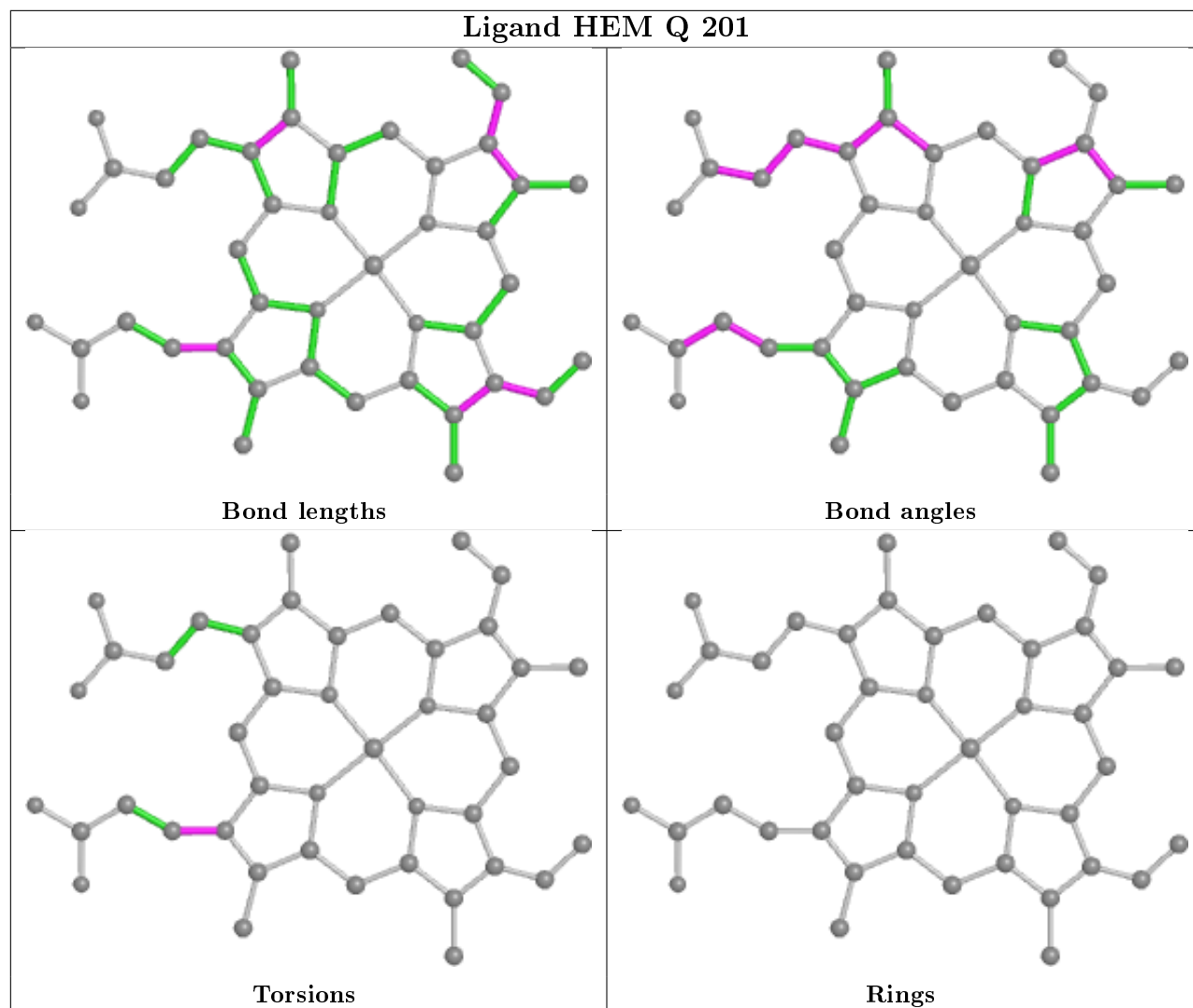
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

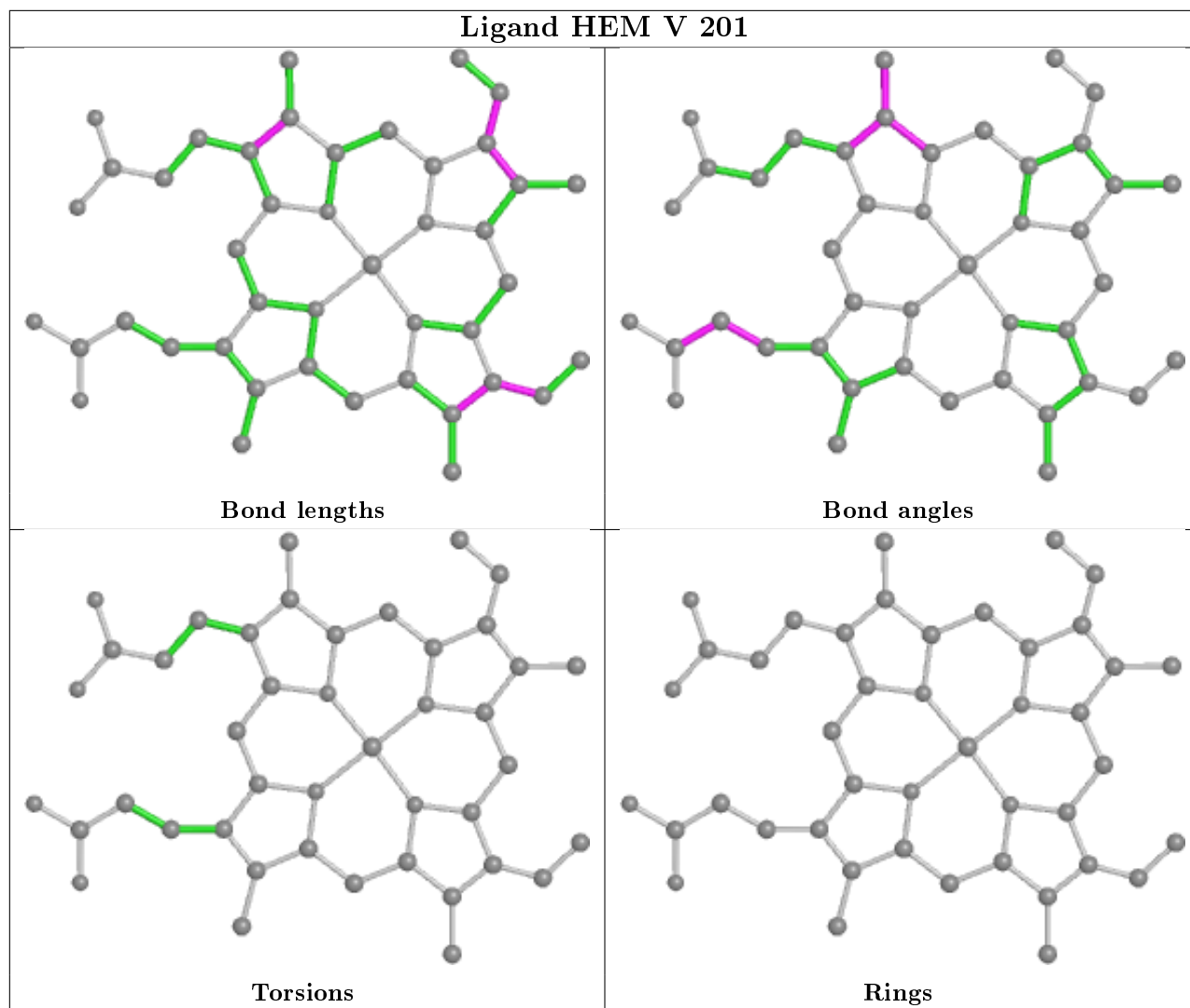


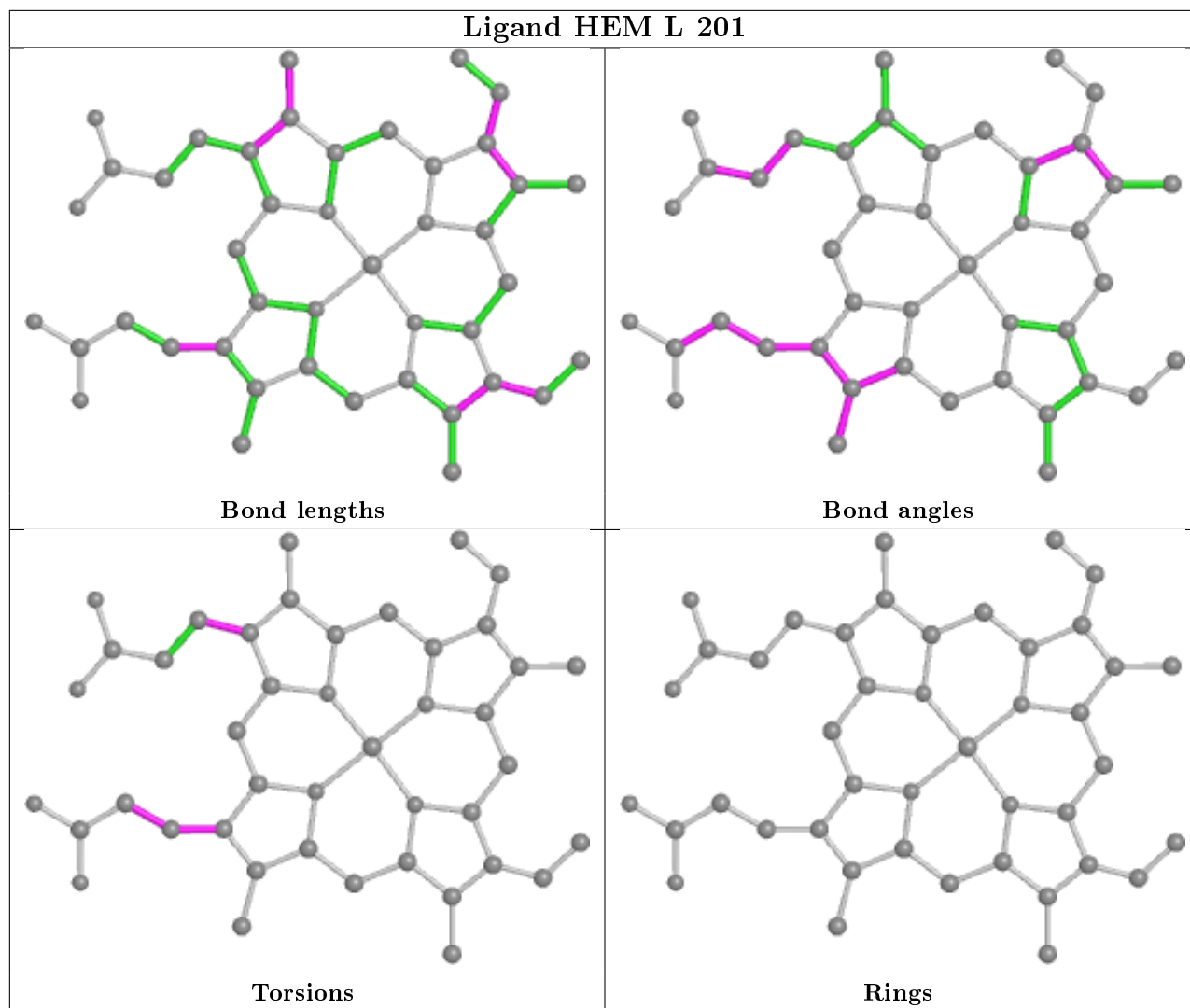


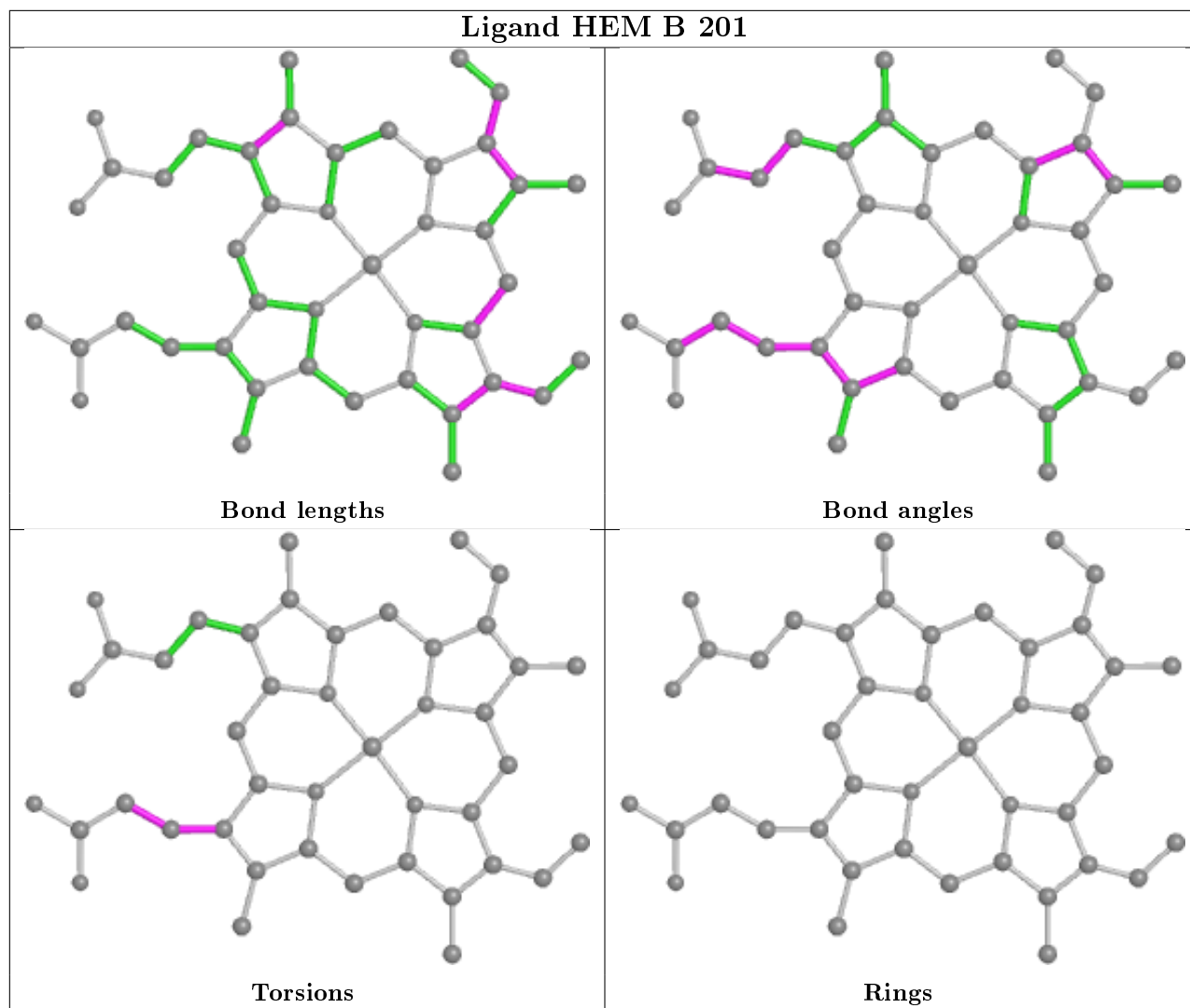


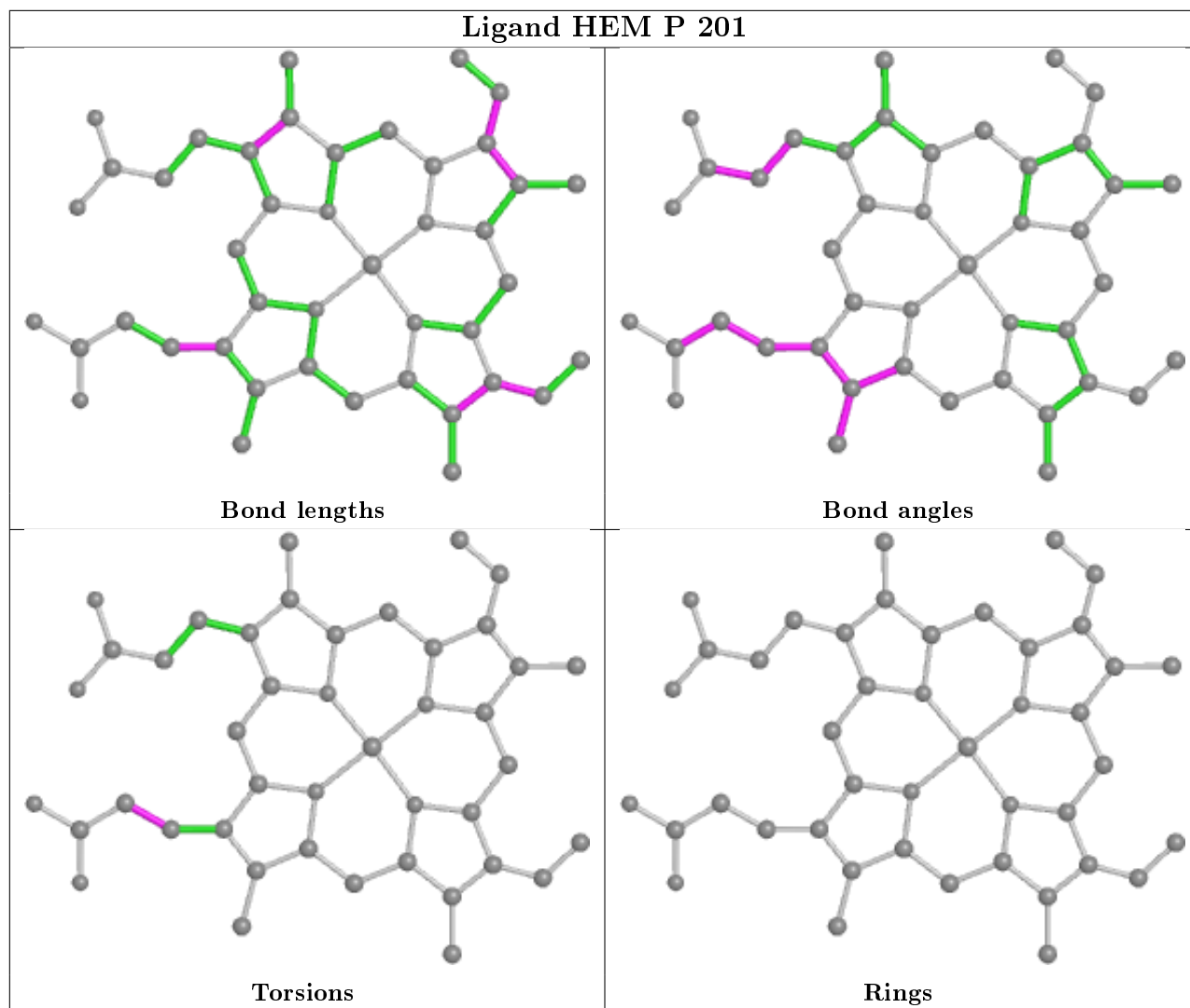


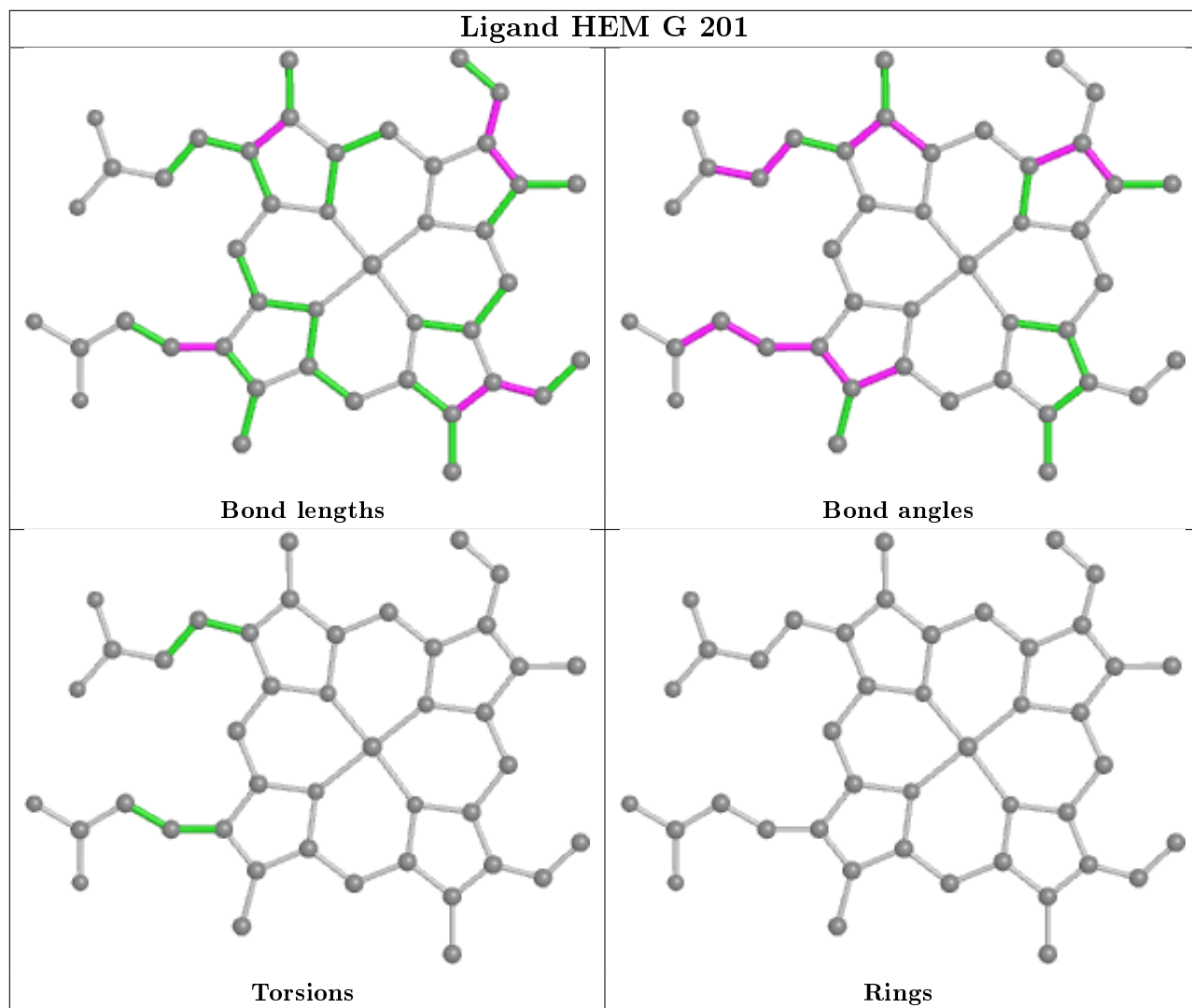


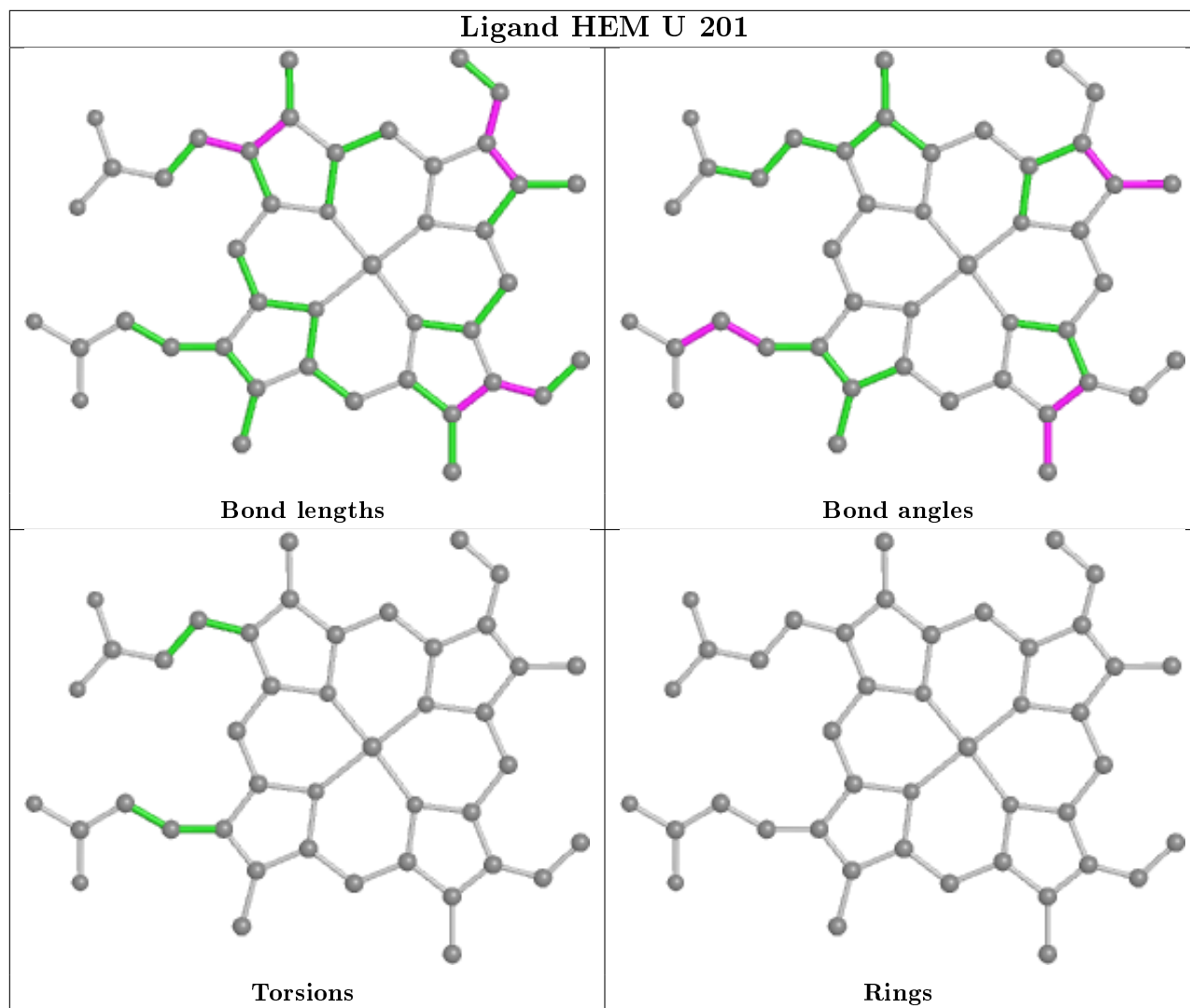


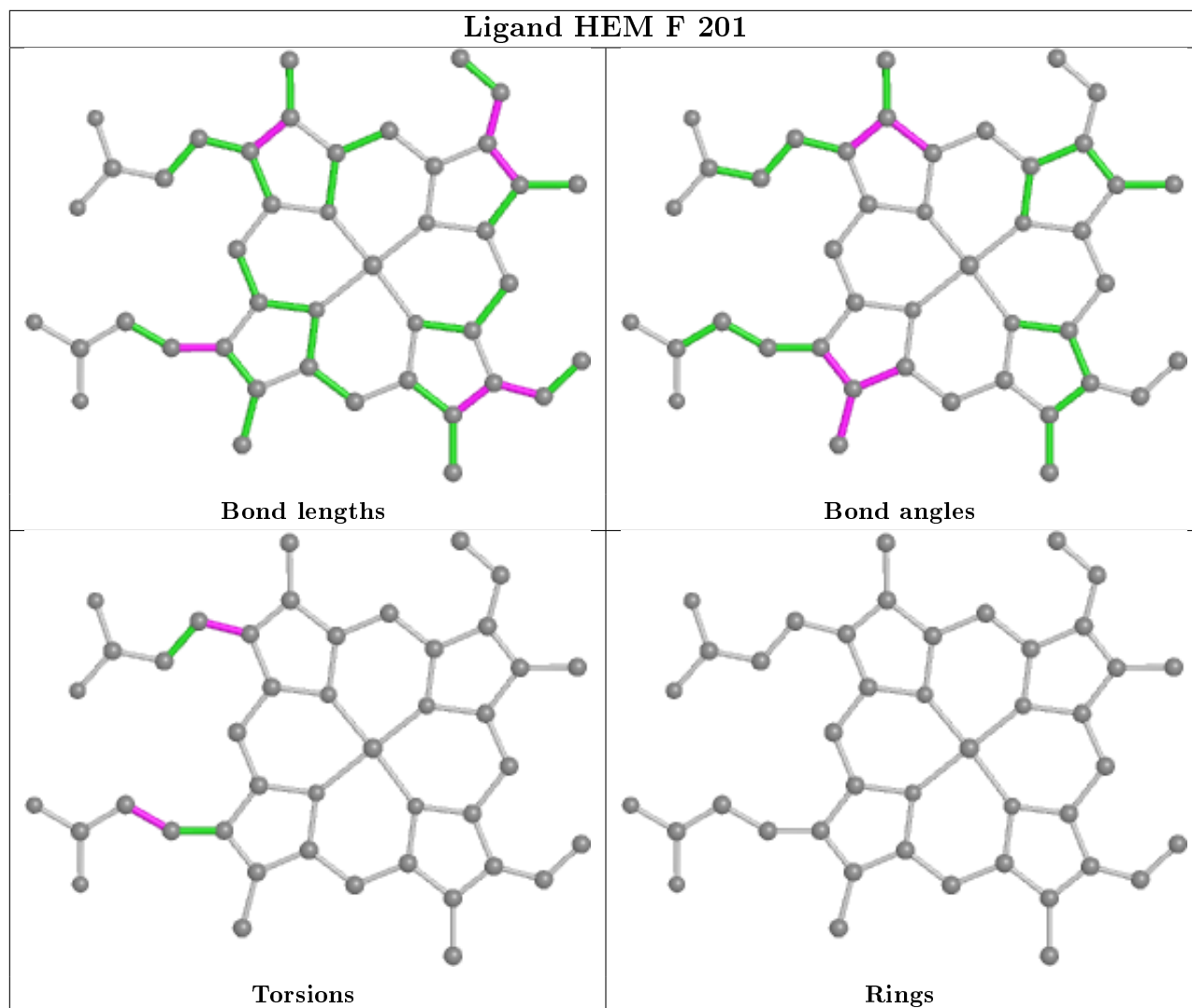




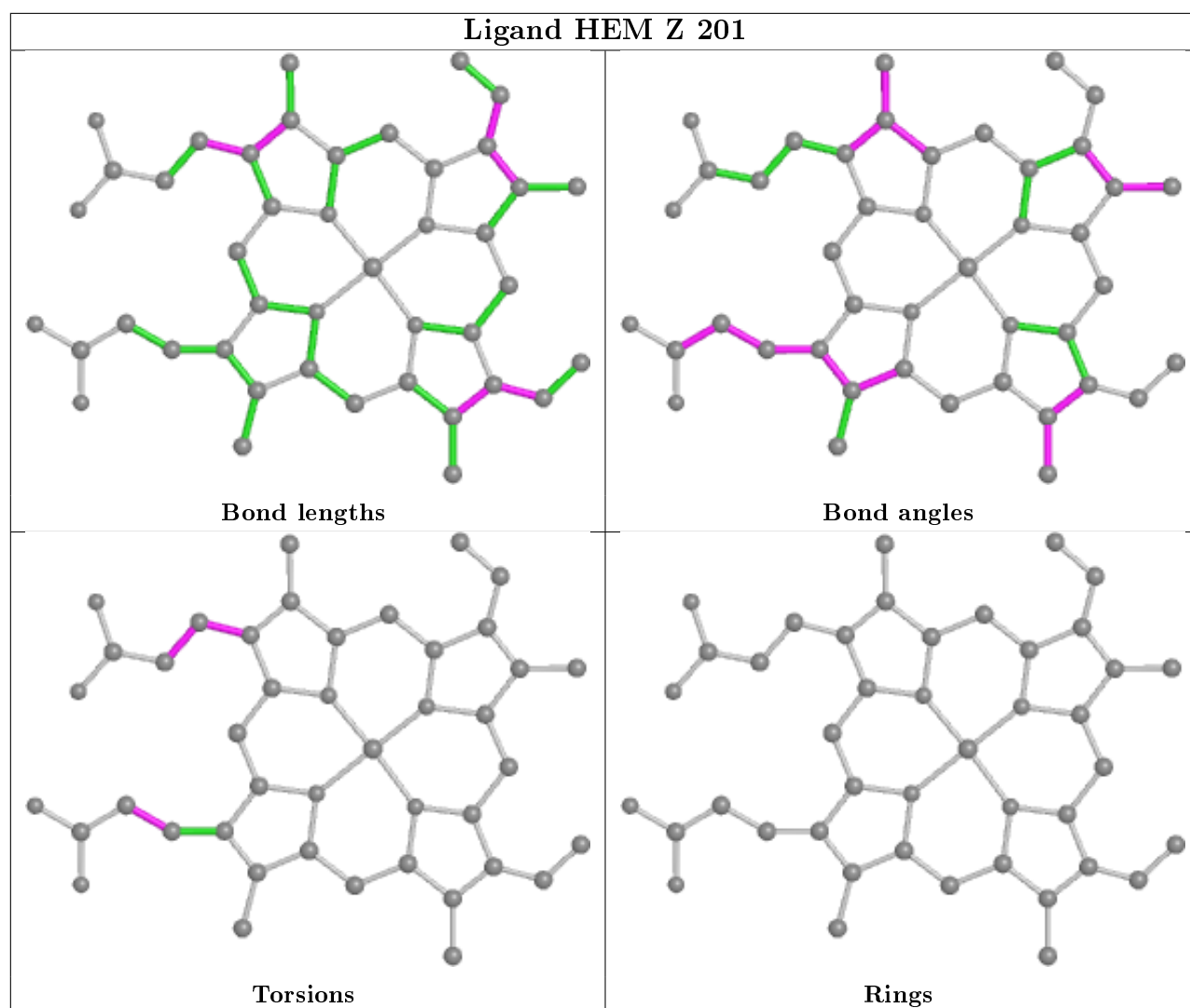












## 5.7 Other polymers [i](#)

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%)	-0.38	0 <a href="#">100</a> <a href="#">100</a>	40, 52, 72, 81	0
1	F	141/141 (100%)	-0.37	0 <a href="#">100</a> <a href="#">100</a>	49, 67, 83, 92	0
1	K	141/141 (100%)	-0.39	0 <a href="#">100</a> <a href="#">100</a>	37, 53, 76, 86	0
1	P	141/141 (100%)	-0.41	0 <a href="#">100</a> <a href="#">100</a>	53, 67, 91, 107	0
1	U	141/141 (100%)	-0.30	0 <a href="#">100</a> <a href="#">100</a>	59, 76, 98, 100	0
1	Z	141/141 (100%)	-0.38	0 <a href="#">100</a> <a href="#">100</a>	58, 72, 89, 93	0
2	1	146/146 (100%)	-0.41	1 (0%) <a href="#">87</a> <a href="#">75</a>	67, 79, 104, 138	0
2	B	146/146 (100%)	-0.37	0 <a href="#">100</a> <a href="#">100</a>	39, 60, 90, 117	0
2	G	146/146 (100%)	-0.32	2 (1%) <a href="#">75</a> <a href="#">56</a>	50, 66, 97, 118	0
2	L	146/146 (100%)	-0.38	0 <a href="#">100</a> <a href="#">100</a>	39, 62, 94, 120	0
2	Q	146/146 (100%)	-0.34	0 <a href="#">100</a> <a href="#">100</a>	53, 69, 102, 133	0
2	V	146/146 (100%)	-0.36	0 <a href="#">100</a> <a href="#">100</a>	66, 78, 105, 137	0
3	2	310/315 (98%)	-0.20	1 (0%) <a href="#">94</a> <a href="#">88</a>	63, 80, 96, 115	0
3	C	310/315 (98%)	-0.27	1 (0%) <a href="#">94</a> <a href="#">88</a>	38, 57, 75, 93	0
3	H	310/315 (98%)	-0.19	2 (0%) <a href="#">89</a> <a href="#">78</a>	56, 74, 94, 114	0
3	M	310/315 (98%)	-0.24	1 (0%) <a href="#">94</a> <a href="#">88</a>	38, 55, 79, 96	0
3	R	310/315 (98%)	-0.21	1 (0%) <a href="#">94</a> <a href="#">88</a>	65, 79, 100, 107	0
3	W	310/315 (98%)	-0.08	2 (0%) <a href="#">89</a> <a href="#">78</a>	76, 92, 112, 121	0
4	3	144/146 (98%)	-0.34	2 (1%) <a href="#">75</a> <a href="#">56</a>	60, 75, 95, 115	0
4	D	144/146 (98%)	-0.38	2 (1%) <a href="#">75</a> <a href="#">56</a>	51, 65, 84, 101	0
4	I	144/146 (98%)	-0.34	0 <a href="#">100</a> <a href="#">100</a>	49, 72, 95, 111	0
4	N	144/146 (98%)	-0.30	2 (1%) <a href="#">75</a> <a href="#">56</a>	54, 68, 88, 100	0
4	S	144/146 (98%)	-0.34	0 <a href="#">100</a> <a href="#">100</a>	53, 74, 94, 109	0
4	X	144/146 (98%)	-0.35	1 (0%) <a href="#">87</a> <a href="#">75</a>	61, 78, 100, 115	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
5	4	261/343 (76%)	-0.03	17 (6%) 18 8	74, 102, 153, 172	2 (0%)
5	E	261/343 (76%)	-0.06	13 (4%) 28 13	42, 82, 146, 155	2 (0%)
5	J	261/343 (76%)	0.23	23 (8%) 10 4	68, 110, 203, 212	2 (0%)
5	O	261/343 (76%)	-0.09	13 (4%) 28 13	43, 83, 140, 151	2 (0%)
5	T	261/343 (76%)	0.22	23 (8%) 10 4	68, 111, 197, 212	2 (0%)
5	Y	261/343 (76%)	0.26	25 (9%) 8 2	82, 111, 205, 213	2 (0%)
All	All	6012/6546 (91%)	-0.19	132 (2%) 62 41	37, 76, 138, 213	12 (0%)

The worst 5 of 132 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	T	293	ALA	8.7
5	T	294	GLN	8.0
5	J	118	ALA	7.5
5	J	296	GLU	7.5
5	J	219	ASN	7.4

## 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](#)

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
6	NAG	f	1	14/15	0.69	0.33	105,105,105,105	0
6	NAG	i	1	14/15	0.70	0.37	118,118,118,118	0
6	NAG	d	1	14/15	0.71	0.32	90,90,90,90	0
6	NAG	h	2	14/15	0.74	0.23	116,116,116,116	0
6	NAG	a	1	14/15	0.78	0.30	86,86,86,86	0
6	NAG	g	2	14/15	0.78	0.22	94,94,94,94	0
6	NAG	c	2	14/15	0.78	0.22	98,98,98,98	0
6	NAG	d	2	14/15	0.80	0.40	97,97,97,97	0
6	NAG	e	2	14/15	0.81	0.22	86,86,86,86	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
6	NAG	j	2	14/15	0.81	0.18	106,106,106,106	0
6	NAG	i	2	14/15	0.82	0.36	123,123,123,123	0
6	NAG	b	2	14/15	0.82	0.17	89,89,89,89	0
6	NAG	f	2	14/15	0.83	0.38	112,112,112,112	0
6	NAG	e	1	14/15	0.84	0.21	75,75,75,75	0
6	NAG	h	1	14/15	0.85	0.22	107,107,107,107	0
6	NAG	c	1	14/15	0.85	0.23	89,89,89,89	0
6	NAG	g	1	14/15	0.87	0.20	88,88,88,88	0
6	NAG	b	1	14/15	0.87	0.18	77,77,77,77	0
6	NAG	j	1	14/15	0.89	0.22	97,97,97,97	0
6	NAG	a	2	14/15	0.90	0.31	94,94,94,94	0

## 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
9	NAG	R	1004	14/15	0.48	0.50	123,123,123,123	0
9	NAG	W	1002	14/15	0.66	0.33	112,112,112,112	0
9	NAG	M	1004	14/15	0.69	0.44	116,116,116,116	0
9	NAG	C	1004	14/15	0.71	0.48	112,112,112,112	0
9	NAG	T	1001	14/15	0.73	0.50	143,148,153,154	0
9	NAG	J	1001	14/15	0.75	0.50	136,144,151,154	0
9	NAG	H	1002	14/15	0.76	0.26	85,85,85,85	0
9	NAG	C	1001	14/15	0.79	0.42	84,84,84,84	0
9	NAG	O	1001	14/15	0.80	0.39	101,107,112,116	0
9	NAG	4	1001	14/15	0.80	0.45	125,128,131,132	0
9	NAG	Y	1001	14/15	0.82	0.39	133,142,147,151	0
9	NAG	R	1001	14/15	0.83	0.39	94,94,94,94	0
9	NAG	M	1001	14/15	0.83	0.40	92,92,92,92	0
9	NAG	Y	1002	14/15	0.84	0.22	81,85,88,90	0
9	NAG	2	1001	14/15	0.86	0.37	101,101,101,101	0
9	NAG	E	1001	14/15	0.86	0.42	103,107,112,113	0
9	NAG	O	1002	14/15	0.88	0.17	60,65,68,70	0
9	NAG	W	1001	14/15	0.88	0.39	101,101,101,101	0
9	NAG	H	1001	14/15	0.90	0.36	81,81,81,81	0
9	NAG	J	1002	14/15	0.90	0.17	80,83,86,87	0
9	NAG	4	1002	14/15	0.91	0.19	81,83,86,87	0
9	NAG	E	1002	14/15	0.92	0.14	63,67,70,70	0

Continued on next page...

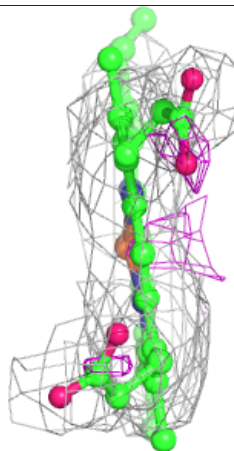
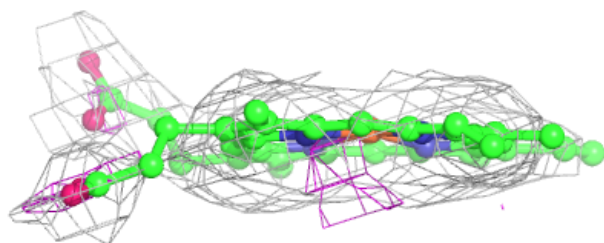
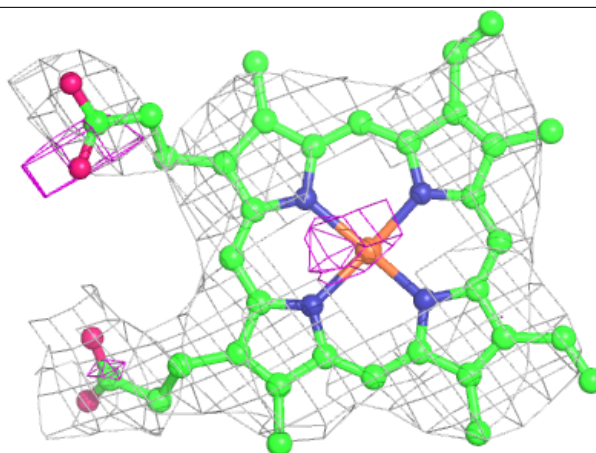
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
8	OXY	Q	202	2/2	0.93	0.31	62,62,62,63	0
9	NAG	T	1002	14/15	0.93	0.11	76,80,86,86	0
8	OXY	A	202	2/2	0.94	0.31	61,61,61,62	0
7	HEM	F	201	43/43	0.94	0.23	56,65,76,80	0
7	HEM	P	201	43/43	0.95	0.22	63,71,79,83	0
7	HEM	G	201	43/43	0.95	0.23	53,58,64,67	0
7	HEM	U	201	43/43	0.95	0.22	68,78,88,92	0
7	HEM	Q	201	43/43	0.95	0.23	58,60,62,63	0
7	HEM	L	201	43/43	0.95	0.21	45,50,57,59	0
7	HEM	A	201	43/43	0.95	0.21	51,55,60,62	0
7	HEM	1	201	43/43	0.95	0.24	63,68,73,76	0
8	OXY	U	202	2/2	0.95	0.25	80,80,80,81	0
7	HEM	Z	201	43/43	0.95	0.21	63,71,79,82	0
7	HEM	K	201	43/43	0.95	0.20	43,57,69,73	0
7	HEM	B	201	43/43	0.96	0.23	41,50,60,64	0
7	HEM	V	201	43/43	0.96	0.21	65,69,74,76	0
8	OXY	Z	202	2/2	0.96	0.19	75,75,75,76	0
8	OXY	K	202	2/2	0.96	0.24	62,62,62,64	0
8	OXY	G	202	2/2	0.96	0.28	59,59,59,60	0
8	OXY	F	202	2/2	0.97	0.29	72,72,72,72	0
8	OXY	B	202	2/2	0.97	0.22	49,49,49,51	0
8	OXY	P	202	2/2	0.98	0.23	70,70,70,72	0
8	OXY	L	202	2/2	0.98	0.19	47,47,47,47	0
8	OXY	V	202	2/2	0.98	0.17	69,69,69,69	0
8	OXY	1	202	2/2	0.99	0.21	69,69,69,69	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.

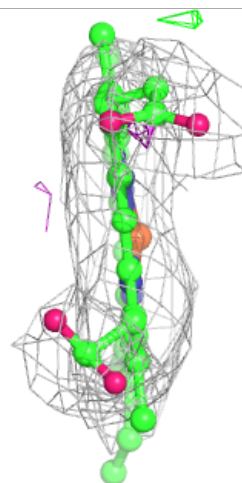
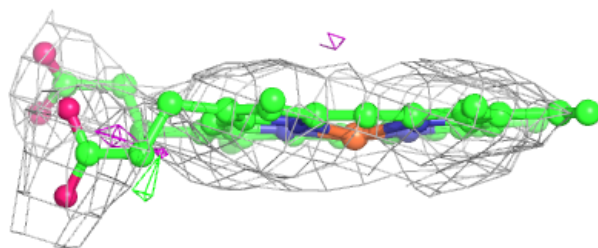
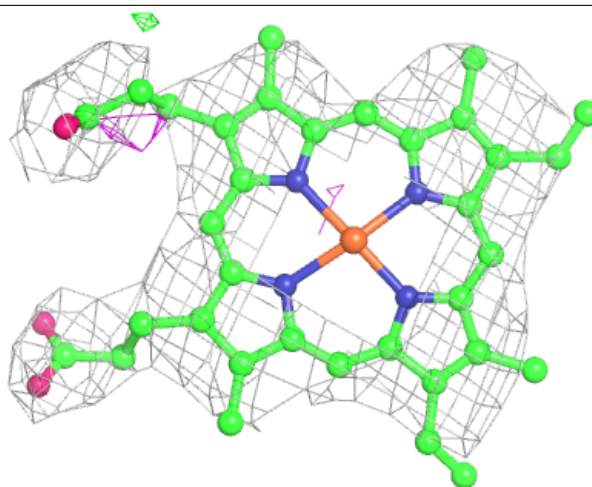
**Electron density around HEM F 201:**

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



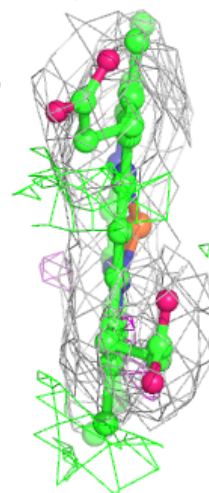
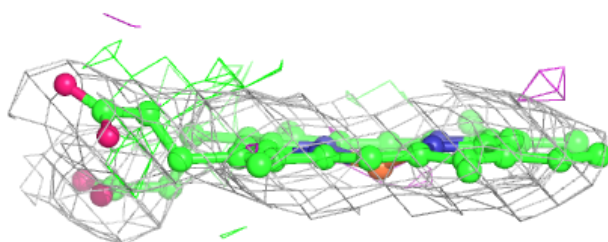
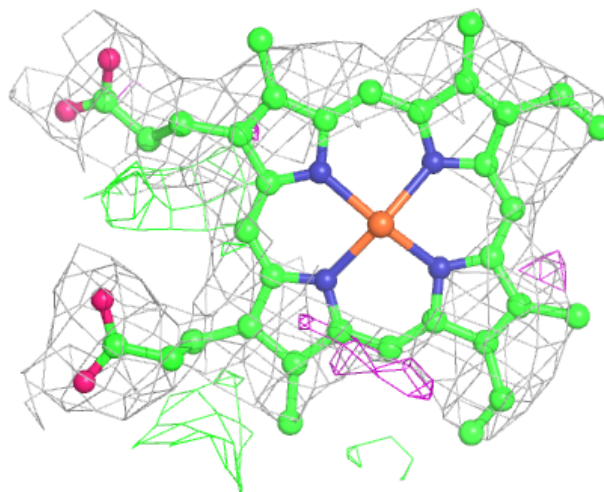
**Electron density around HEM P 201:**

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

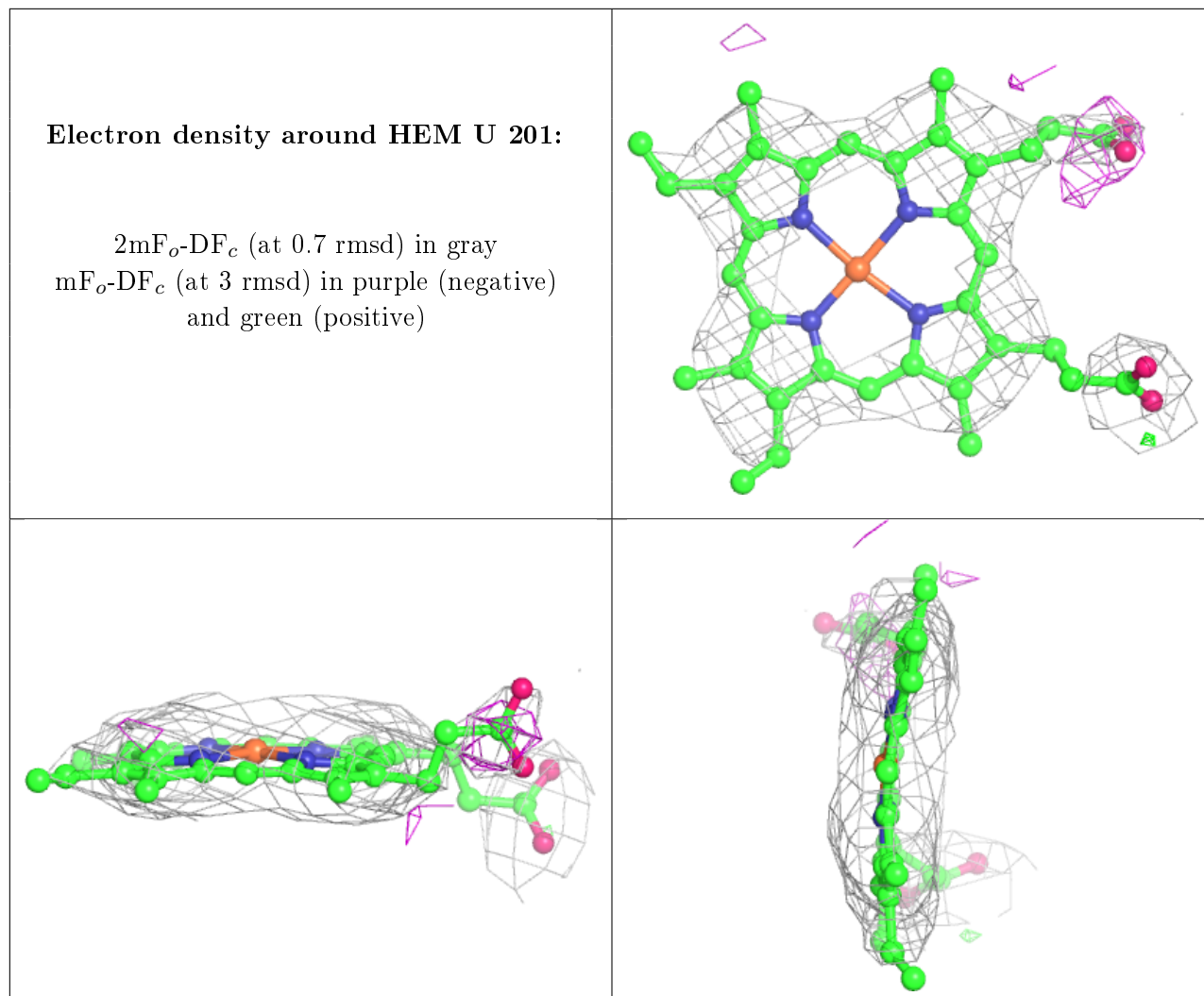


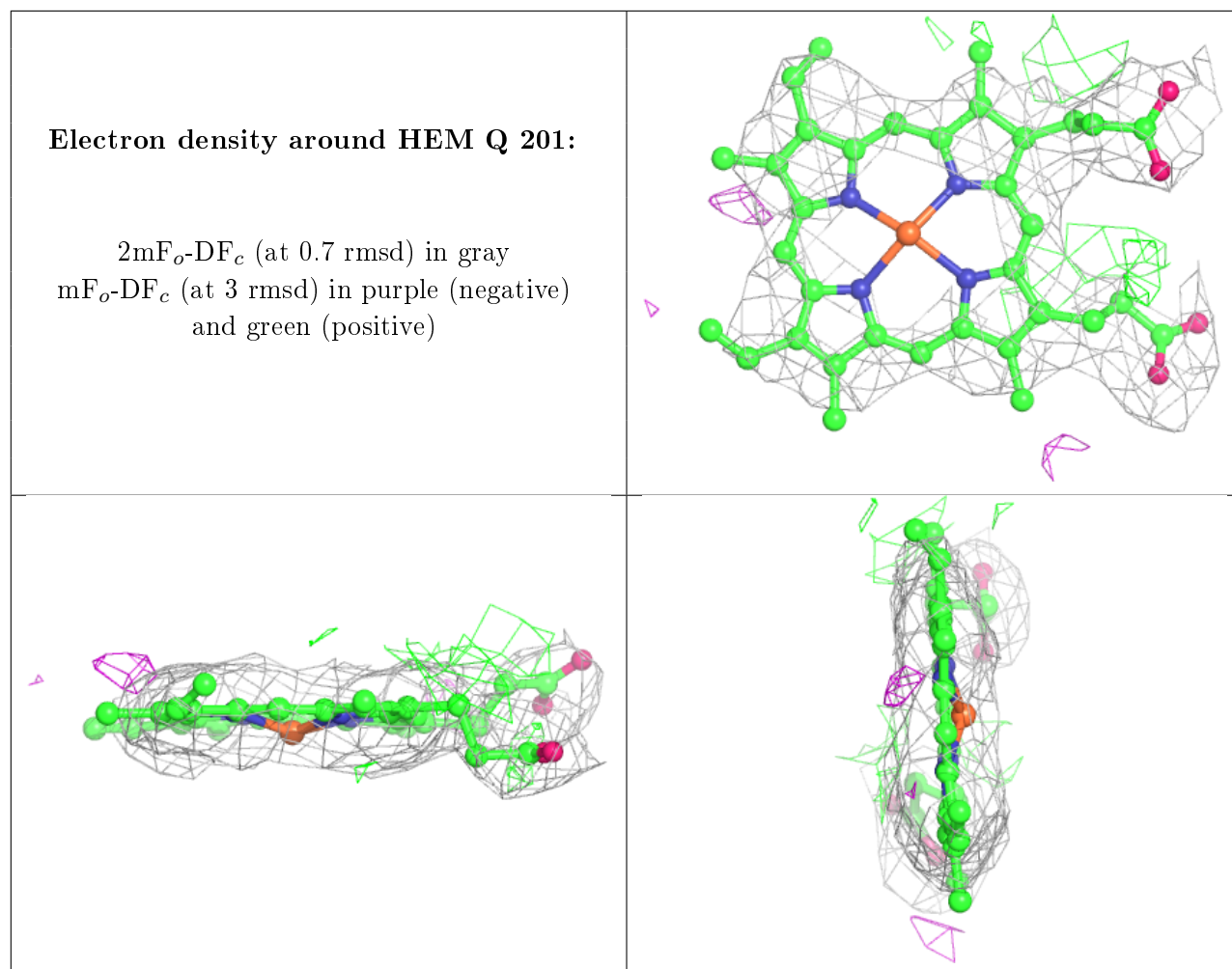
**Electron density around HEM G 201:**

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



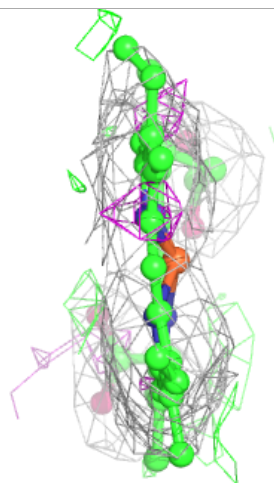
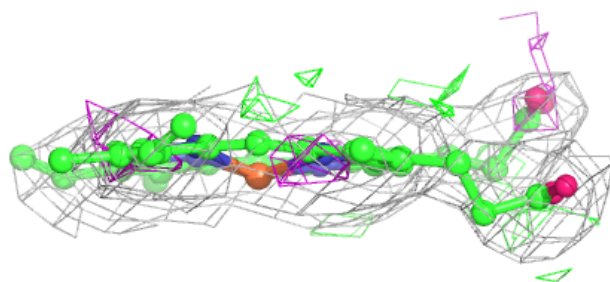
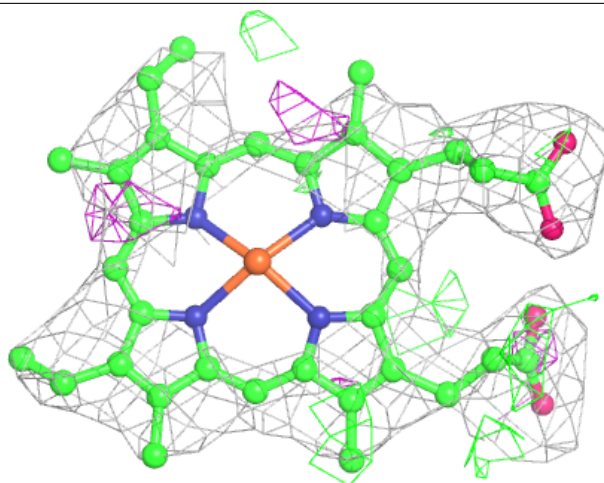






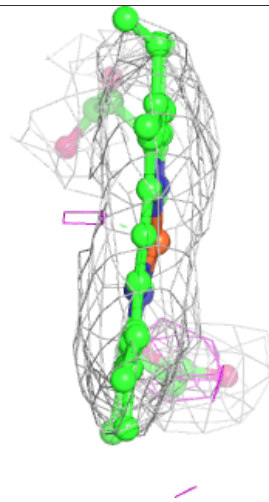
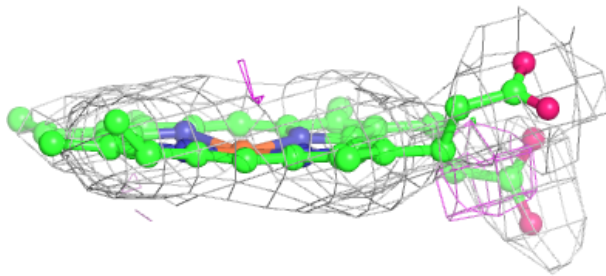
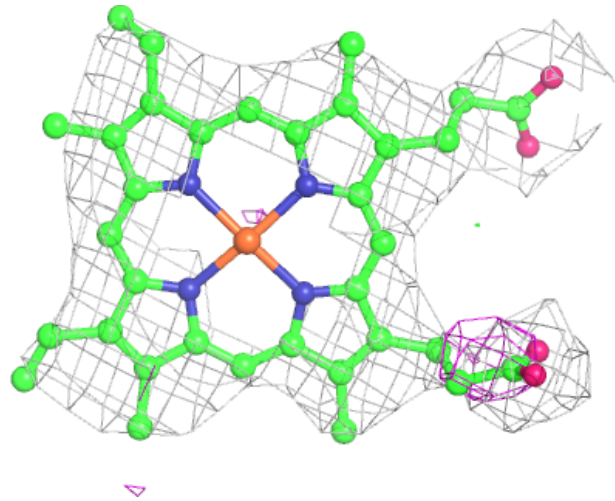
**Electron density around HEM L 201:**

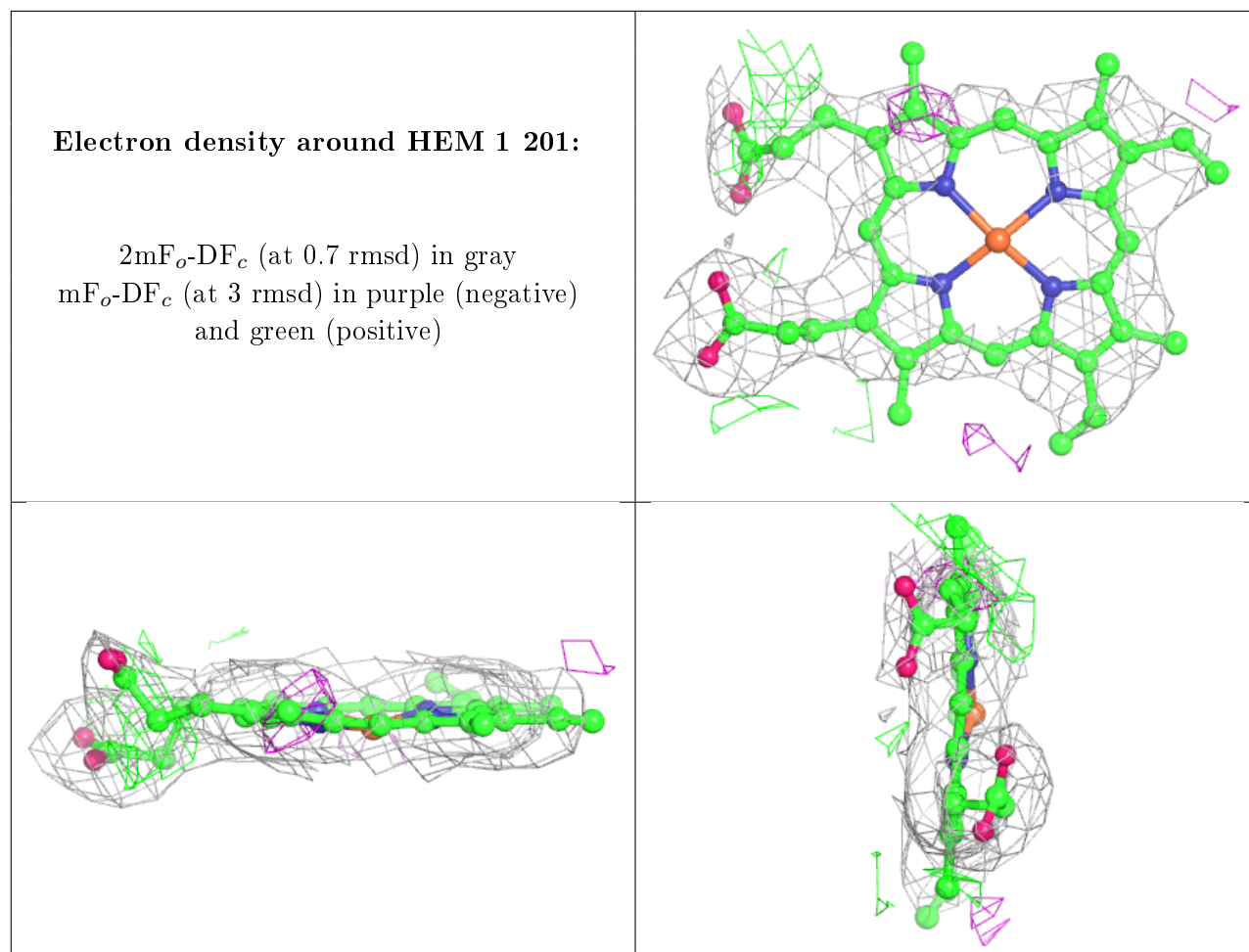
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

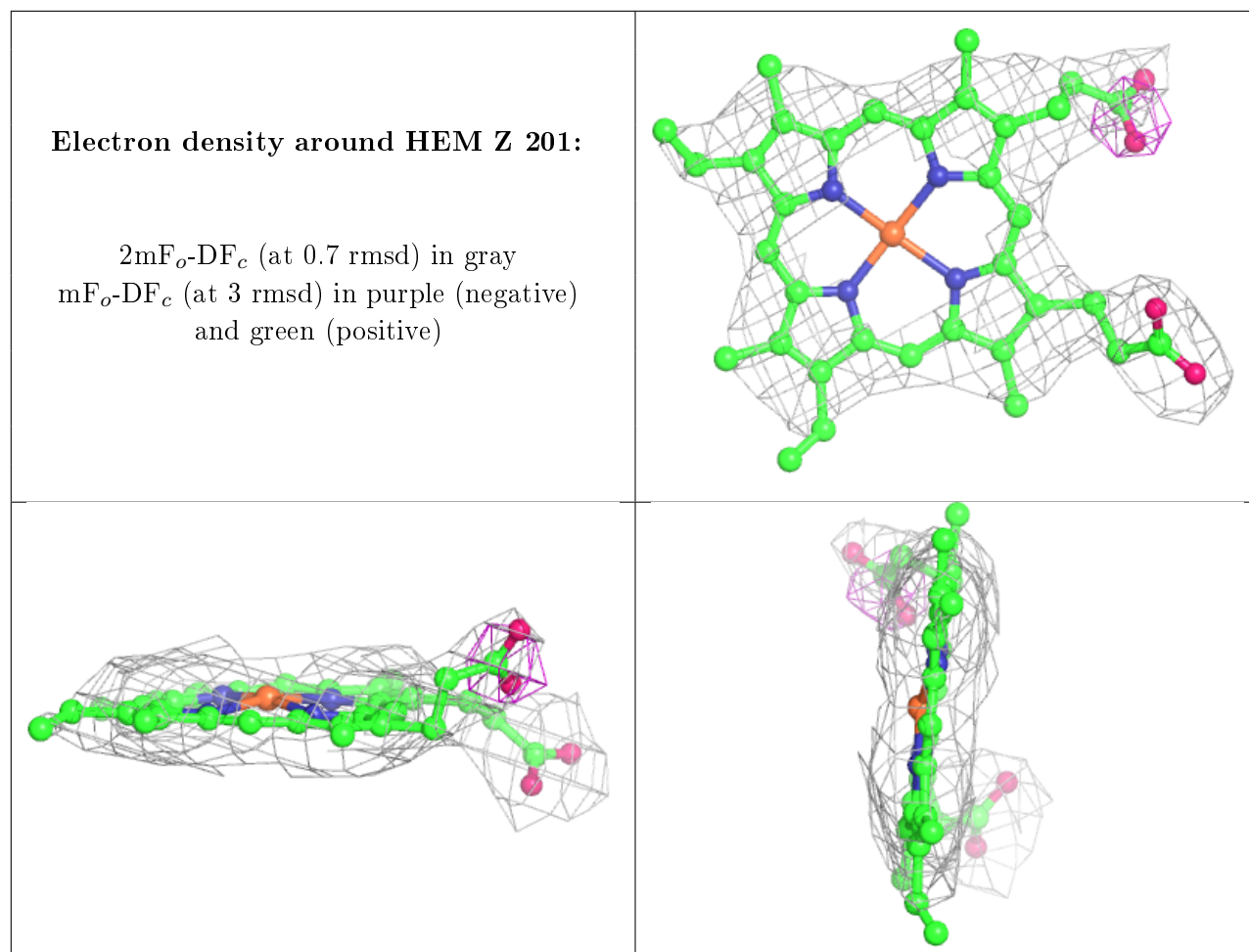


**Electron density around HEM A 201:**

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

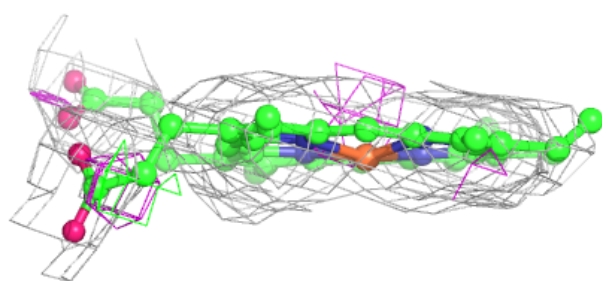
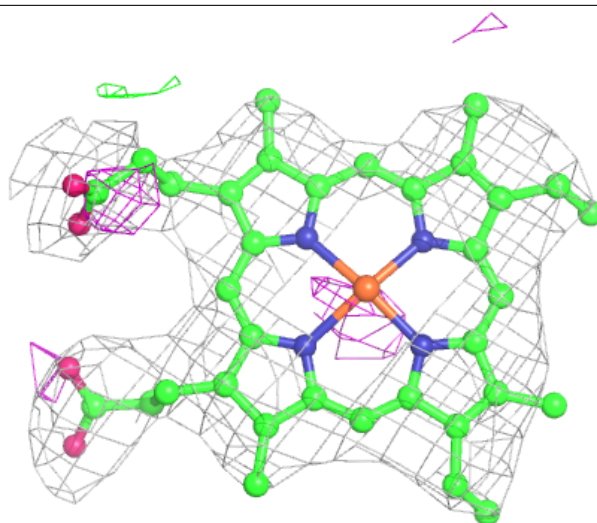






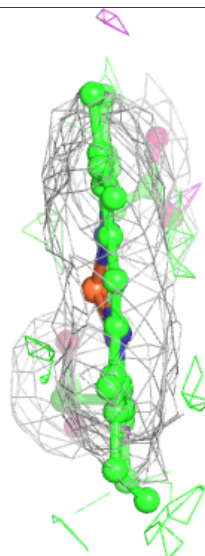
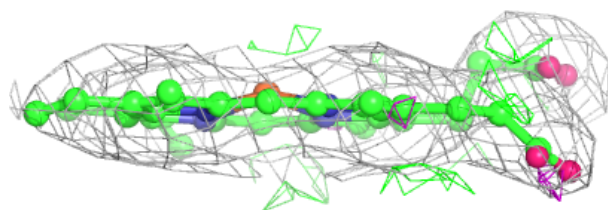
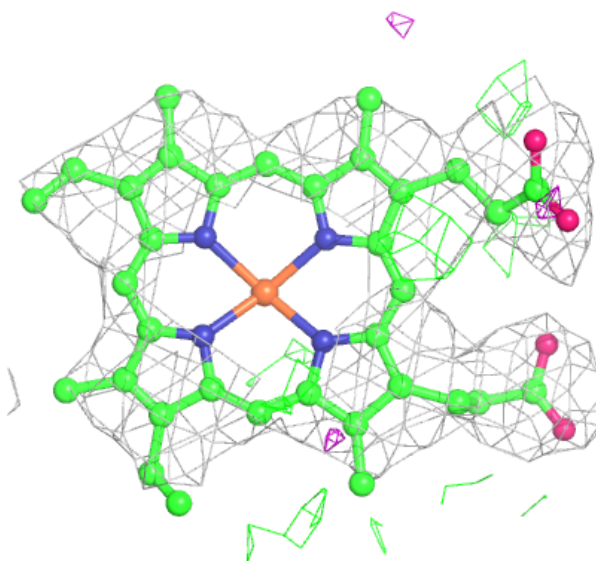
**Electron density around HEM K 201:**

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

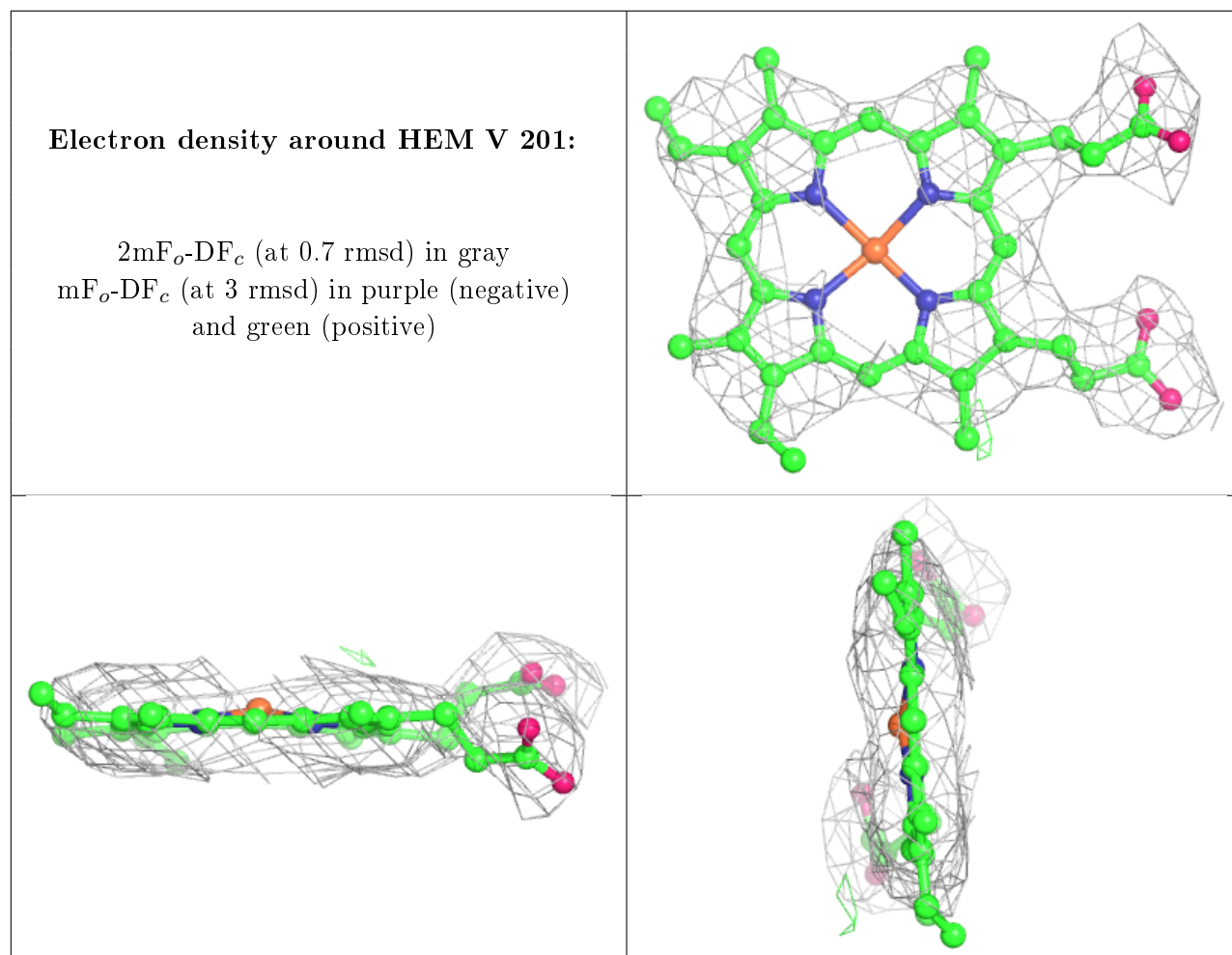


**Electron density around HEM B 201:**

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







## 6.5 Other polymers [\(i\)](#)

There are no such residues in this entry.