



# Full wwPDB X-ray Structure Validation Report ⓘ

Dec 15, 2024 – 06:36 PM EST

PDB ID : 1CXP  
Title : CRYOGENIC CRYSTAL STRUCTURE OF HUMAN MYELOPEROXI-  
DASE ISOFORM C  
Authors : Fiedler, T.J.; Fenna, R.E.  
Deposited on : 1999-08-30  
Resolution : 1.80 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](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 : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 1.21  
EDS : 3.0  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.004 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.40

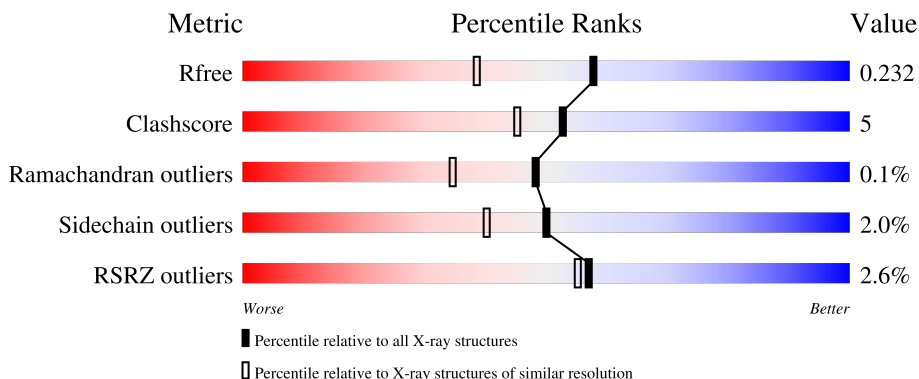
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.80 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))
$R_{free}$	164625	7108 (1.80-1.80)
Clashscore	180529	8162 (1.80-1.80)
Ramachandran outliers	177936	8077 (1.80-1.80)
Sidechain outliers	177891	8076 (1.80-1.80)
RSRZ outliers	164620	7108 (1.80-1.80)

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	104	
1	B	104	
2	C	466	
2	D	466	
3	E	6	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	F	6	 83% 17%

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	ACT	C	606	-	-	X	-

## 2 Entry composition i

There are 10 unique types of molecules in this entry. The entry contains 10307 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 MYELOPEROXIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	104	Total 838	C 529	N 148	O 156	S 5	0	0	0
1	B	104	Total 838	C 529	N 148	O 156	S 5	0	0	0

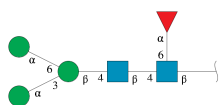
- Molecule 2 is a protein called MYELOPEROXIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	466	Total 3733	C 2351	N 687	O 668	S 27	0	0	0
2	D	466	Total 3733	C 2351	N 687	O 668	S 27	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	150	CSO	CYS	modified residue	UNP P05164
D	150	CSO	CYS	modified residue	UNP P05164

- Molecule 3 is an oligosaccharide called alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose.

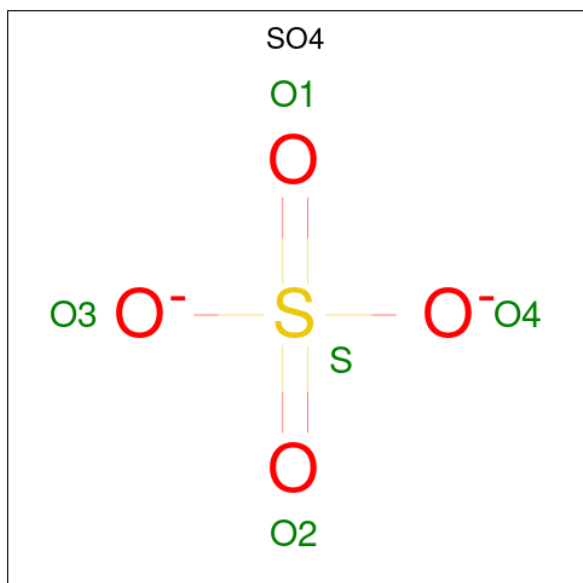


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	E	6	Total 71	C 40	N 2	O 29	0	0	0
3	F	6	Total 71	C 40	N 2	O 29	0	0	0

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Cl 1 1	0	0
4	B	1	Total Cl 1 1	0	0

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total O S 5 4 1	0	0
5	C	1	Total O S 5 4 1	0	0
5	B	1	Total O S 5 4 1	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	
6	C	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
6	D	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		

- Molecule 7 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

*Continued on next page...*

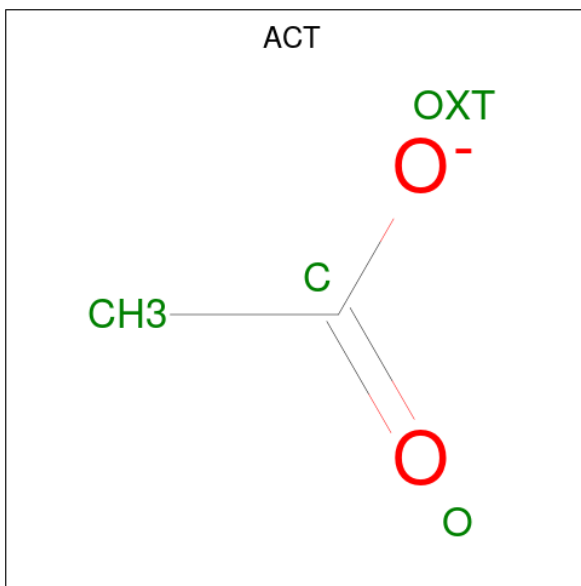
Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	C	1	Total	C	N	O	0	0
			14	8	1	5		
7	D	1	Total	C	N	O	0	0
			14	8	1	5		
7	D	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 8 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	C	1	Total	Ca	0	0
			1	1		
8	D	1	Total	Ca	0	0
			1	1		

- Molecule 9 is ACETATE ION (three-letter code: ACT) (formula: C<sub>2</sub>H<sub>3</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	C	1	Total	C	O	0	0
			4	2	2		
9	C	1	Total	C	O	0	0
			4	2	2		
9	C	1	Total	C	O	0	0
			4	2	2		
9	D	1	Total	C	O	0	0
			4	2	2		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	D	1	Total	C	O	0	0
			4	2	2		
9	D	1	Total	C	O	0	0
			4	2	2		

- Molecule 10 is water.

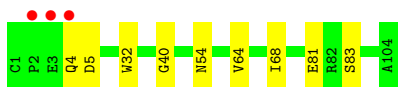
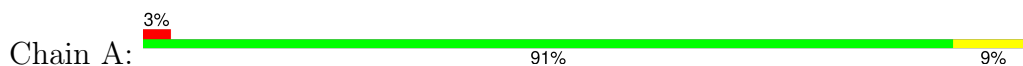
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	101	Total	O	0	0
			101	101		
10	C	319	Total	O	0	0
			319	319		
10	B	105	Total	O	0	0
			105	105		
10	D	313	Total	O	0	0
			313	313		



### 3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

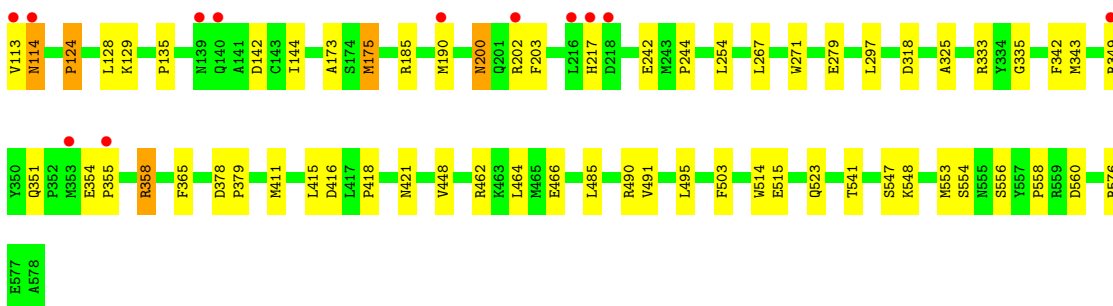
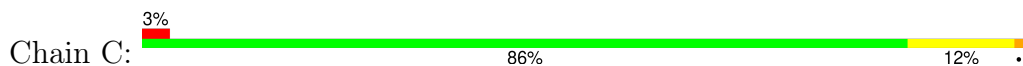
- Molecule 1: MYELOPEROXIDASE



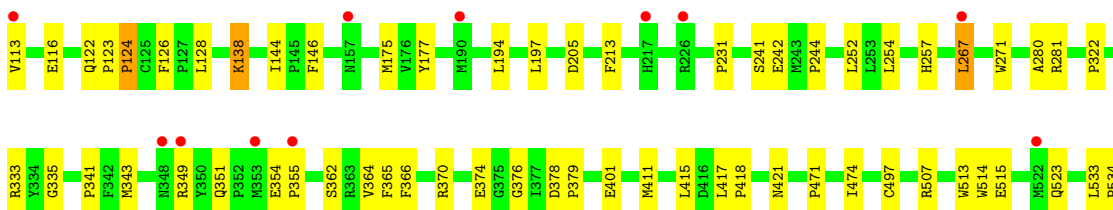
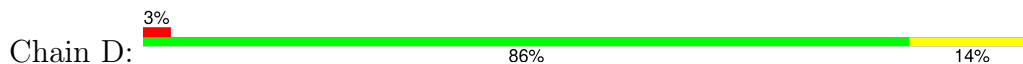
- Molecule 1: MYELOPEROXIDASE

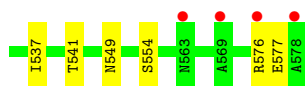


- Molecule 2: MYELOPEROXIDASE



- Molecule 2: MYELOPEROXIDASE





- Molecule 3: alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain E: 67% 33%



- Molecule 3: alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain F: 83% 17%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	111.03Å 63.38Å 92.17Å 90.00° 97.36° 90.00°	Depositor
Resolution (Å)	30.00 – 1.80 30.00 – 1.80	Depositor EDS
% Data completeness (in resolution range)	96.1 (30.00-1.80) 99.8 (30.00-1.80)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.55 (at 1.75Å)	Xtrriage
Refinement program	X-PLOR 3.851	Depositor
R, $R_{free}$	0.197 , 0.239 0.192 , 0.232	Depositor DCC
$R_{free}$ test set	5940 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	12.8	Xtrriage
Anisotropy	0.406	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 53.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	10307	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	12.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.78% 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: MAN, SO4, CL, CSO, FUC, HEM, CA, NAG, BMA, ACT

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	0.67	0/863	0.71	0/1174
1	B	0.68	0/863	0.72	0/1174
2	C	0.62	0/3811	0.64	0/5168
2	D	0.61	0/3811	0.61	0/5168
All	All	0.63	0/9348	0.64	0/12684

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	838	0	798	8	0
1	B	838	0	798	2	0
2	C	3733	0	3725	42	0
2	D	3733	0	3725	43	0
3	E	71	0	61	0	0
3	F	71	0	61	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
5	A	5	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	B	5	0	0	0	0
5	C	5	0	0	0	0
6	C	43	0	30	4	0
6	D	43	0	30	3	0
7	C	28	0	26	0	0
7	D	28	0	26	0	0
8	C	1	0	0	0	0
8	D	1	0	0	0	0
9	C	12	0	9	2	0
9	D	12	0	9	0	0
10	A	101	0	0	1	0
10	B	105	0	0	0	0
10	C	319	0	0	6	0
10	D	313	0	0	1	0
All	All	10307	0	9298	92	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 5.

All (92) 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)
10:A:1702:HOH:O	2:C:129:LYS:HD3	1.83	0.78
2:D:354:GLU:HB3	2:D:355:PRO:HA	1.77	0.67
2:D:349:ARG:HG3	2:D:351:GLN:HG2	1.80	0.64
2:C:333:ARG:HH11	2:C:421:ASN:HD22	1.47	0.63
2:C:200:ASN:ND2	2:C:202:ARG:H	1.97	0.62
2:C:354:GLU:HB3	2:C:355:PRO:HA	1.82	0.62
2:C:200:ASN:HD22	2:C:202:ARG:H	1.48	0.60
2:C:142:ASP:HB3	10:C:897:HOH:O	2.02	0.59
2:C:135:PRO:HG2	10:C:735:HOH:O	2.03	0.58
2:C:485:LEU:HD13	2:C:490:ARG:HA	1.84	0.58
1:A:4:GLN:HG2	1:A:5:ASP:N	2.18	0.58
2:D:333:ARG:HH11	2:D:421:ASN:ND2	2.01	0.57
2:D:333:ARG:HH11	2:D:421:ASN:HD22	1.50	0.57
6:C:601:HEM:HBC2	6:C:601:HEM:HMC2	1.87	0.56
2:D:411:MET:HE2	2:D:415:LEU:HD21	1.87	0.56
2:D:349:ARG:HE	2:D:351:GLN:HG3	1.71	0.56
2:D:411:MET:CE	2:D:415:LEU:HD21	2.35	0.56
1:A:4:GLN:HG2	1:A:5:ASP:H	1.71	0.55
2:C:349:ARG:HG3	2:C:351:GLN:HG2	1.87	0.55

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:113:VAL:HG21	2:D:122:GLN:HB2	1.88	0.54
2:D:378:ASP:HB2	2:D:379:PRO:HD3	1.90	0.54
2:C:333:ARG:HH11	2:C:421:ASN:ND2	2.06	0.53
1:A:68:ILE:HD13	2:C:464:LEU:HD23	1.91	0.53
2:C:128:LEU:HB2	2:C:144:ILE:HB	1.91	0.53
1:B:83:SER:HB3	2:D:554:SER:O	2.09	0.52
2:D:138:LYS:NZ	2:D:138:LYS:HA	2.24	0.52
2:D:417:LEU:HB3	2:D:418:PRO:HD3	1.92	0.52
2:D:197:LEU:HG	2:D:257:HIS:CD2	2.45	0.52
2:D:194:LEU:HD13	10:D:965:HOH:O	2.09	0.52
2:C:548:LYS:HD2	2:C:560:ASP:HA	1.90	0.52
2:D:122:GLN:HE21	2:D:122:GLN:HA	1.76	0.51
2:D:244:PRO:HD3	2:D:364:VAL:O	2.10	0.51
2:D:335:GLY:HA3	6:D:601:HEM:CBC	2.41	0.51
2:C:113:VAL:HG12	2:C:114:ASN:N	2.26	0.51
2:C:411:MET:HE2	2:C:415:LEU:HD21	1.93	0.51
2:D:241:SER:O	2:D:366:PHE:HA	2.11	0.51
2:C:242:GLU:OE1	6:C:601:HEM:HBB2	2.10	0.50
2:D:514:TRP:CE2	2:D:515:GLU:HG3	2.47	0.50
2:C:556:SER:HA	10:C:817:HOH:O	2.12	0.50
2:C:491:VAL:HB	2:C:495:LEU:HB2	1.94	0.49
2:C:271:TRP:CZ3	2:C:279:GLU:HG3	2.48	0.48
2:C:514:TRP:CE2	2:C:515:GLU:HG3	2.49	0.48
2:D:244:PRO:HB2	2:D:343:MET:SD	2.54	0.47
1:A:83:SER:HB3	2:C:554:SER:O	2.14	0.47
2:D:242:GLU:OE1	6:D:601:HEM:HBB2	2.14	0.47
2:C:378:ASP:HB2	2:C:379:PRO:HD3	1.96	0.47
2:C:462:ARG:O	2:C:466:GLU:HG2	2.16	0.47
2:C:378:ASP:OD1	2:C:541:THR:HB	2.14	0.46
2:D:257:HIS:CE1	2:D:280:ALA:HB3	2.50	0.46
2:D:242:GLU:O	2:D:365:PHE:HA	2.16	0.46
2:D:335:GLY:HA3	6:D:601:HEM:HBC2	1.98	0.46
6:C:601:HEM:HBC2	6:C:601:HEM:CMC	2.46	0.45
2:D:177:TYR:OH	2:D:281:ARG:HA	2.17	0.45
9:C:606:ACT:H2	10:C:939:HOH:O	2.16	0.45
2:C:342:PHE:CD1	2:C:358:ARG:NH1	2.85	0.45
2:C:135:PRO:CG	10:C:735:HOH:O	2.64	0.44
2:D:267:LEU:HD22	2:D:576:ARG:CZ	2.48	0.44
1:A:40:GLY:HA2	1:B:20:PRO:HD2	1.99	0.44
2:D:126:PHE:HB3	2:D:146:PHE:CD2	2.52	0.44
2:C:448:VAL:HG21	2:C:462:ARG:HG2	1.99	0.44

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:81:GLU:O	2:C:553:MET:HG3	2.17	0.43
2:C:124:PRO:HA	9:C:606:ACT:CH3	2.49	0.43
2:D:471:PRO:HA	2:D:474:ILE:HG13	2.01	0.43
2:D:123:PRO:HA	2:D:124:PRO:HA	1.85	0.43
2:D:370:ARG:O	2:D:374:GLU:HB2	2.17	0.43
2:C:267:LEU:HD12	2:C:576:ARG:HB2	2.01	0.43
1:A:64:VAL:HG13	1:A:68:ILE:HD12	2.00	0.43
2:C:342:PHE:CE1	2:C:358:ARG:NH1	2.87	0.42
2:D:213:PHE:CD2	2:D:231:PRO:HG2	2.55	0.42
2:D:128:LEU:HB2	2:D:144:ILE:HB	2.02	0.42
2:C:242:GLU:O	2:C:365:PHE:HA	2.20	0.42
2:D:205:ASP:OD2	2:D:376:GLY:HA3	2.18	0.42
2:D:122:GLN:HA	2:D:122:GLN:NE2	2.34	0.42
1:A:32:TRP:CE2	2:C:325:ALA:HB2	2.55	0.41
2:C:200:ASN:HD22	2:C:203:PHE:H	1.66	0.41
2:C:244:PRO:HB2	2:C:343:MET:SD	2.60	0.41
2:D:341:PRO:O	2:D:401:GLU:HG3	2.20	0.41
2:D:362:SER:HA	2:D:365:PHE:CE1	2.55	0.41
2:C:114:ASN:HA	10:C:821:HOH:O	2.21	0.41
2:D:116:GLU:OE1	2:D:411:MET:HE3	2.19	0.41
2:D:534:PRO:HG2	2:D:549:ASN:ND2	2.36	0.41
2:D:271:TRP:HE1	2:D:577:GLU:CD	2.24	0.41
2:C:333:ARG:HD3	2:C:421:ASN:ND2	2.35	0.41
2:C:335:GLY:HA3	6:C:601:HEM:CBC	2.51	0.41
2:C:173:ALA:HA	2:C:175:MET:SD	2.60	0.41
2:C:297:LEU:HD11	2:C:503:PHE:CD1	2.55	0.40
2:D:507:ARG:HG3	2:D:513:TRP:CE2	2.56	0.40
2:C:185:ARG:HG3	2:C:190:MET:CE	2.51	0.40
2:D:252:LEU:HD11	2:D:537:ILE:HA	2.03	0.40
2:D:378:ASP:OD1	2:D:541:THR:HB	2.22	0.40
2:D:533:LEU:HB3	2:D:534:PRO:HD3	2.02	0.40
2:C:416:ASP:OD2	2:C:418:PRO:HD2	2.21	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries

of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	102/104 (98%)	100 (98%)	2 (2%)	0	100	100
1	B	102/104 (98%)	100 (98%)	2 (2%)	0	100	100
2	C	463/466 (99%)	449 (97%)	13 (3%)	1 (0%)	44	31
2	D	463/466 (99%)	450 (97%)	13 (3%)	0	100	100
All	All	1130/1140 (99%)	1099 (97%)	30 (3%)	1 (0%)	48	34

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	114	ASN

### 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	90/90 (100%)	89 (99%)	1 (1%)	70	65
1	B	90/90 (100%)	89 (99%)	1 (1%)	70	65
2	C	410/410 (100%)	400 (98%)	10 (2%)	44	32
2	D	410/410 (100%)	402 (98%)	8 (2%)	50	40
All	All	1000/1000 (100%)	980 (98%)	20 (2%)	50	40

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	54	ASN
2	C	124	PRO
2	C	175	MET
2	C	200	ASN
2	C	217	HIS

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
2	C	254	LEU
2	C	318	ASP
2	C	358	ARG
2	C	523	GLN
2	C	547	SER
2	C	558	PRO
1	B	54	ASN
2	D	124	PRO
2	D	138	LYS
2	D	175	MET
2	D	254	LEU
2	D	267	LEU
2	D	322	PRO
2	D	497	CYS
2	D	523	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
2	C	200	ASN
2	C	421	ASN
2	C	530	GLN
1	B	54	ASN
2	D	122	GLN
2	D	133	ASN
2	D	421	ASN
2	D	549	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](#)

2 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the

expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
2	CSO	C	150	2	3,6,7	0.96	0	1,6,8	1.16	0
2	CSO	D	150	2	3,6,7	1.05	0	1,6,8	1.21	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
2	CSO	C	150	2	-	0/1/5/7	-
2	CSO	D	150	2	-	0/1/5/7	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

12 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	$\# Z  > 2$
3	NAG	E	1	3,2	14,14,15	0.83	0	17,19,21	0.89	0
3	NAG	E	2	3	14,14,15	0.69	0	17,19,21	0.89	0
3	BMA	E	3	3	11,11,12	0.65	0	15,15,17	0.48	0
3	MAN	E	4	3	11,11,12	0.75	0	15,15,17	0.77	1 (6%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	MAN	E	5	3	11,11,12	0.57	0	15,15,17	0.68	1 (6%)
3	FUC	E	6	3	10,10,11	0.79	0	14,14,16	0.42	0
3	NAG	F	1	3,2	14,14,15	0.63	0	17,19,21	0.89	0
3	NAG	F	2	3	14,14,15	0.75	0	17,19,21	0.99	2 (11%)
3	BMA	F	3	3	11,11,12	0.50	0	15,15,17	0.52	0
3	MAN	F	4	3	11,11,12	0.64	0	15,15,17	0.57	0
3	MAN	F	5	3	11,11,12	0.90	0	15,15,17	0.51	0
3	FUC	F	6	3	10,10,11	0.79	0	14,14,16	0.49	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	E	1	3,2	-	0/6/23/26	0/1/1/1
3	NAG	E	2	3	-	0/6/23/26	0/1/1/1
3	BMA	E	3	3	-	0/2/19/22	0/1/1/1
3	MAN	E	4	3	-	0/2/19/22	0/1/1/1
3	MAN	E	5	3	-	0/2/19/22	0/1/1/1
3	FUC	E	6	3	-	-	0/1/1/1
3	NAG	F	1	3,2	-	0/6/23/26	0/1/1/1
3	NAG	F	2	3	-	0/6/23/26	0/1/1/1
3	BMA	F	3	3	-	0/2/19/22	0/1/1/1
3	MAN	F	4	3	-	0/2/19/22	0/1/1/1
3	MAN	F	5	3	-	0/2/19/22	0/1/1/1
3	FUC	F	6	3	-	-	0/1/1/1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	4	MAN	C1-O5-C5	2.61	115.68	112.19
3	F	2	NAG	C2-N2-C7	-2.44	119.63	122.90
3	F	2	NAG	C4-C3-C2	-2.29	107.67	111.02
3	E	5	MAN	C1-O5-C5	2.15	115.07	112.19

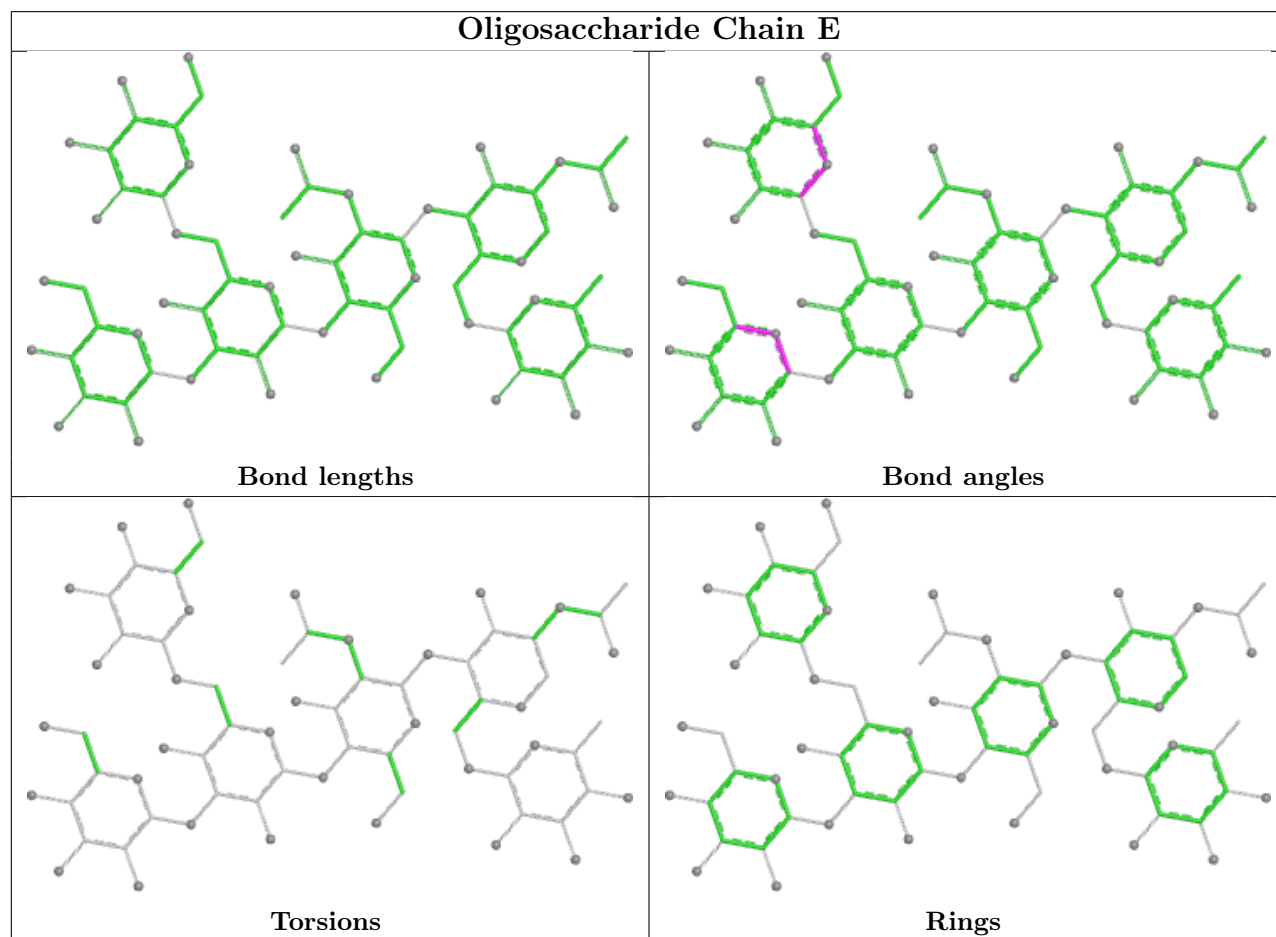
There are no chirality outliers.

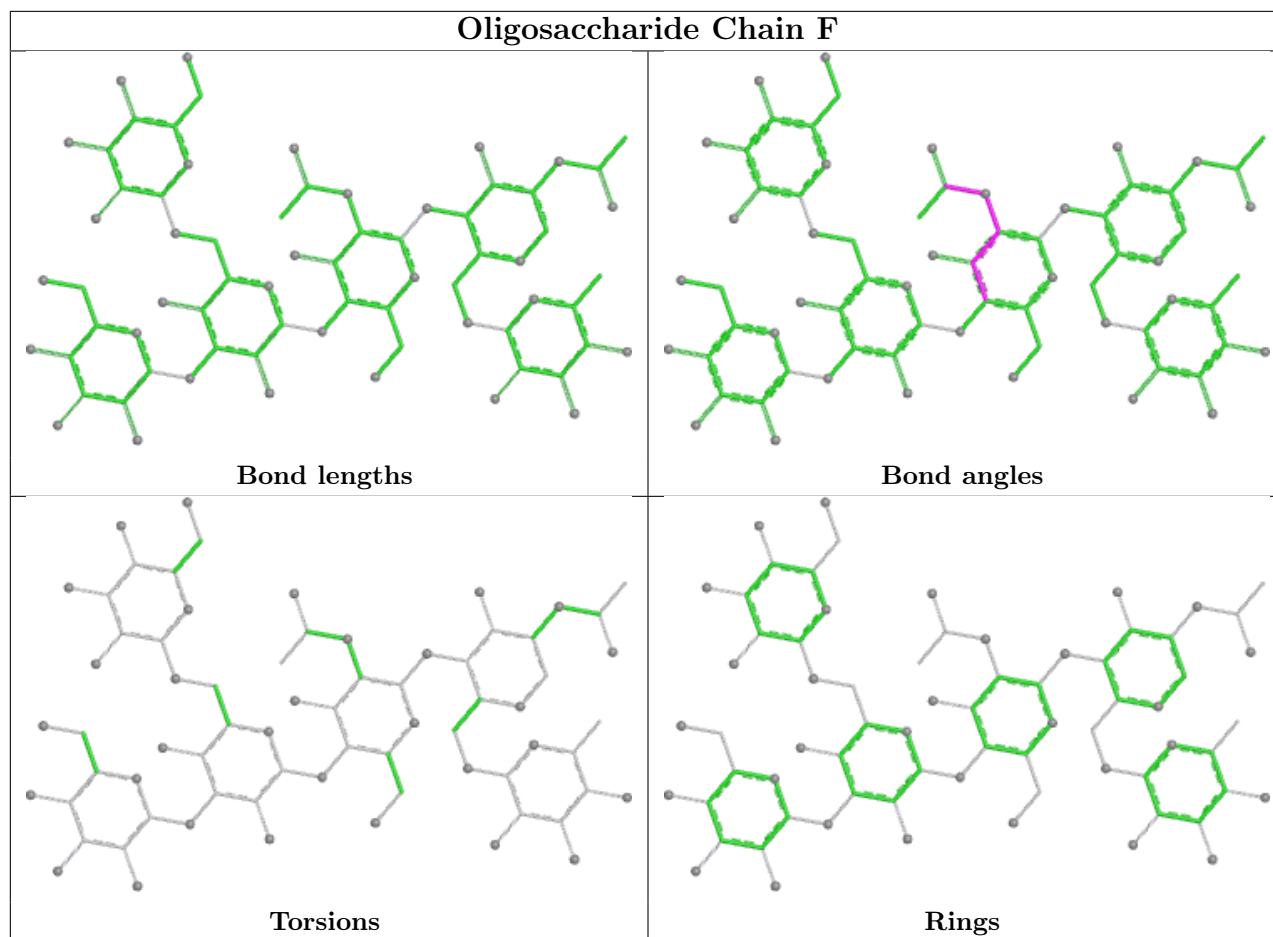
There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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





## 5.6 Ligand geometry [i](#)

Of 19 ligands modelled in this entry, 4 are monoatomic - leaving 15 for Mogul analysis.

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$
5	SO4	A	1602	-	4,4,4	0.35	0	6,6,6	0.15	0
9	ACT	D	605	-	3,3,3	1.22	1 (33%)	3,3,3	0.35	0
7	NAG	D	603	2	14,14,15	0.49	0	17,19,21	0.78	0
7	NAG	C	602	2	14,14,15	0.51	0	17,19,21	0.96	2 (11%)
9	ACT	C	606	-	3,3,3	1.21	1 (33%)	3,3,3	0.33	0
7	NAG	C	603	2	14,14,15	0.49	0	17,19,21	0.64	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	HEM	C	601	10,2	42,50,50	1.41	6 (14%)	46,82,82	1.17	2 (4%)
9	ACT	C	607	-	3,3,3	1.31	1 (33%)	3,3,3	0.45	0
9	ACT	C	608	-	3,3,3	1.03	0	3,3,3	0.41	0
7	NAG	D	602	2	14,14,15	0.80	1 (7%)	17,19,21	0.81	1 (5%)
5	SO4	C	605	-	4,4,4	0.38	0	6,6,6	0.12	0
6	HEM	D	601	10,2	42,50,50	1.34	6 (14%)	46,82,82	1.24	2 (4%)
9	ACT	D	607	-	3,3,3	1.19	1 (33%)	3,3,3	0.45	0
5	SO4	B	202	-	4,4,4	0.40	0	6,6,6	0.07	0
9	ACT	D	606	-	3,3,3	1.12	0	3,3,3	0.43	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
7	NAG	D	603	2	-	0/6/23/26	0/1/1/1
7	NAG	C	602	2	-	0/6/23/26	0/1/1/1
7	NAG	C	603	2	-	0/6/23/26	0/1/1/1
6	HEM	C	601	10,2	-	4/12/54/54	-
7	NAG	D	602	2	-	0/6/23/26	0/1/1/1
6	HEM	D	601	10,2	-	4/12/54/54	-

All (17) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	C	601	HEM	CBB-CAB	3.37	1.46	1.30
6	D	601	HEM	CBB-CAB	3.32	1.46	1.30
6	D	601	HEM	C3C-C4C	3.06	1.45	1.41
6	C	601	HEM	C3C-C4C	2.96	1.45	1.41
6	D	601	HEM	CBC-CAC	2.62	1.45	1.29
6	C	601	HEM	C3C-CAC	-2.56	1.41	1.47
6	C	601	HEM	CBC-CAC	2.53	1.45	1.29
6	C	601	HEM	CHB-C1B	2.48	1.40	1.34
9	C	607	ACT	O-C	-2.26	1.12	1.22
6	C	601	HEM	CAB-C3B	-2.20	1.41	1.47
7	D	602	NAG	C1-C2	2.18	1.55	1.52
6	D	601	HEM	C3C-CAC	-2.16	1.42	1.47
6	D	601	HEM	CAB-C3B	-2.12	1.41	1.47
9	D	605	ACT	O-C	-2.08	1.13	1.22
9	C	606	ACT	O-C	-2.06	1.13	1.22

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	D	607	ACT	O-C	-2.05	1.13	1.22
6	D	601	HEM	CHB-C1B	2.01	1.39	1.34

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	D	601	HEM	C3B-C4B-NB	4.22	112.50	109.47
6	C	601	HEM	C3B-C4B-NB	3.77	112.18	109.47
6	D	601	HEM	C4C-CHD-C1D	2.19	125.45	122.56
7	D	602	NAG	C1-O5-C5	2.16	115.09	112.19
6	C	601	HEM	C4C-CHD-C1D	2.14	125.38	122.56
7	C	602	NAG	C1-O5-C5	2.09	114.99	112.19
7	C	602	NAG	C2-N2-C7	-2.03	120.17	122.90

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	C	601	HEM	CAD-CBD-CGD-O1D
6	C	601	HEM	CAA-CBA-CGA-O2A
6	D	601	HEM	CAD-CBD-CGD-O1D
6	C	601	HEM	CAD-CBD-CGD-O2D
6	D	601	HEM	CAD-CBD-CGD-O2D
6	C	601	HEM	CAA-CBA-CGA-O1A
6	D	601	HEM	CAA-CBA-CGA-O1A
6	D	601	HEM	CAA-CBA-CGA-O2A

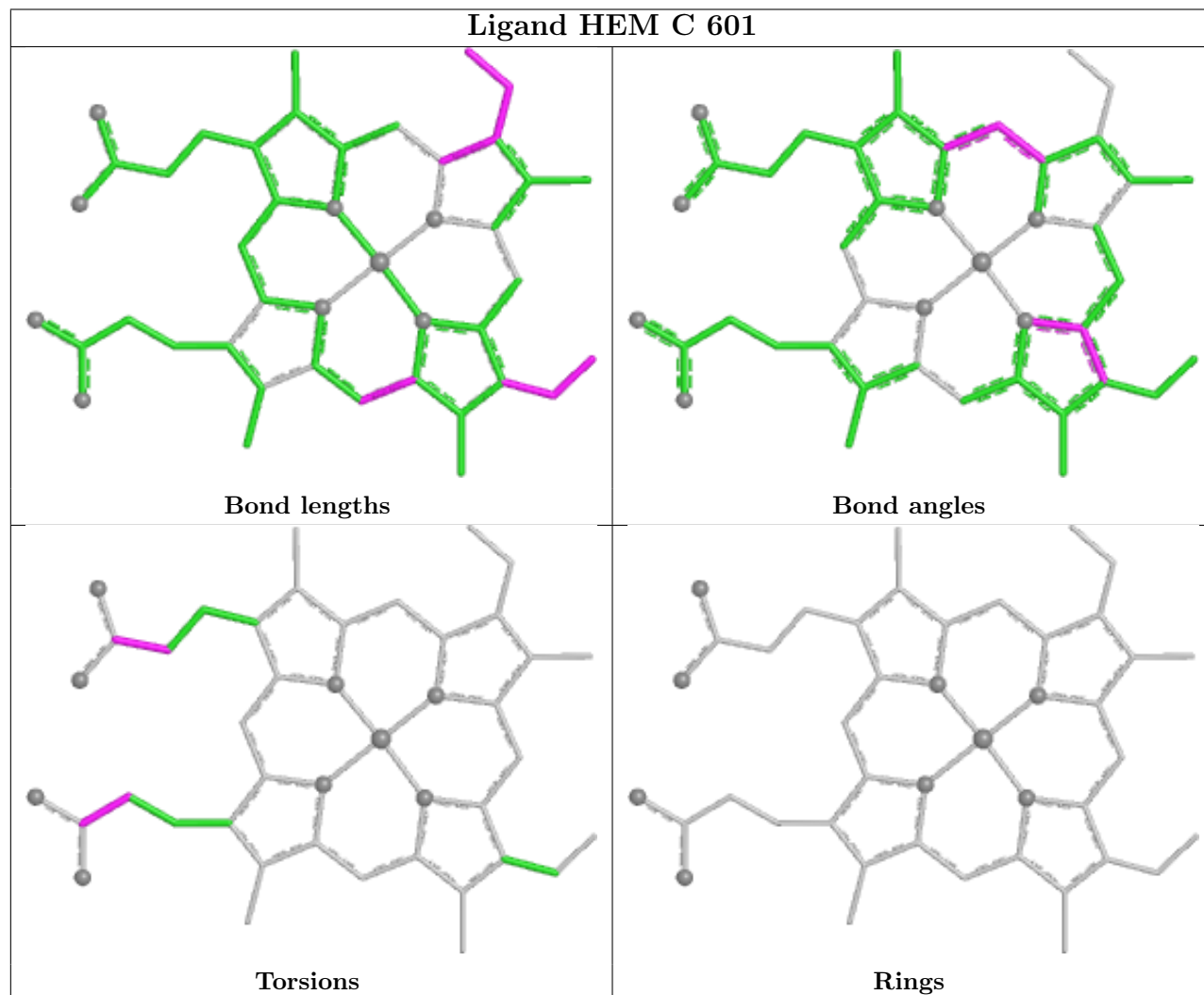
There are no ring outliers.

3 monomers are involved in 9 short contacts:

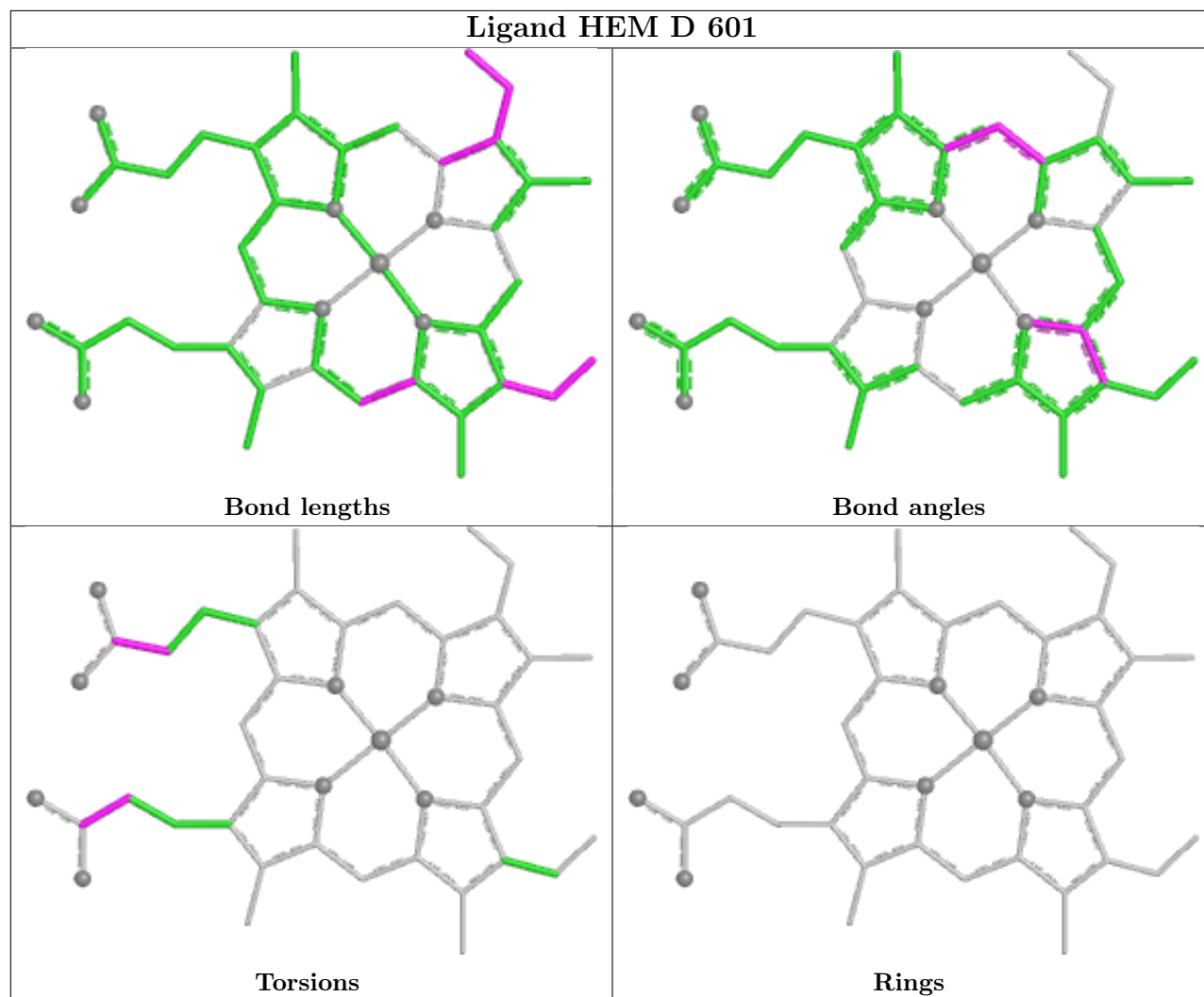
Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	C	606	ACT	2	0
6	C	601	HEM	4	0
6	D	601	HEM	3	0

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

### 6.1 Protein, DNA and RNA chains

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<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	104/104 (100%)	-0.15	3 (2%) 54 52	5, 8, 22, 44	0
1	B	104/104 (100%)	-0.18	0 100 100	5, 9, 18, 23	0
2	C	465/466 (99%)	-0.09	12 (2%) 57 55	4, 9, 22, 38	0
2	D	465/466 (99%)	0.14	15 (3%) 50 48	4, 12, 27, 42	0
All	All	1138/1140 (99%)	-0.01	30 (2%) 57 55	4, 10, 24, 44	0

All (30) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	113	VAL	8.0
2	D	578	ALA	4.5
2	C	217	HIS	3.6
1	A	2	PRO	3.6
2	D	576	ARG	3.6
2	D	217	HIS	3.2
2	D	267	LEU	3.2
2	C	353	MET	3.1
2	D	190	MET	3.0
2	C	218	ASP	3.0
2	D	349	ARG	3.0
2	D	355	PRO	2.9
2	D	348	ASN	2.7
2	C	355	PRO	2.7
2	C	140	GLN	2.7
2	C	216	LEU	2.5
2	D	353	MET	2.5
1	A	3	GLU	2.5
1	A	4	GLN	2.4
2	C	139	ASN	2.4
2	D	563	ASN	2.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	D	522	MET	2.3
2	D	226	ARG	2.3
2	D	157	ASN	2.3
2	C	202	ARG	2.2
2	C	114	ASN	2.2
2	D	569	ALA	2.1
2	D	113	VAL	2.1
2	C	190	MET	2.0
2	C	349	ARG	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [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.

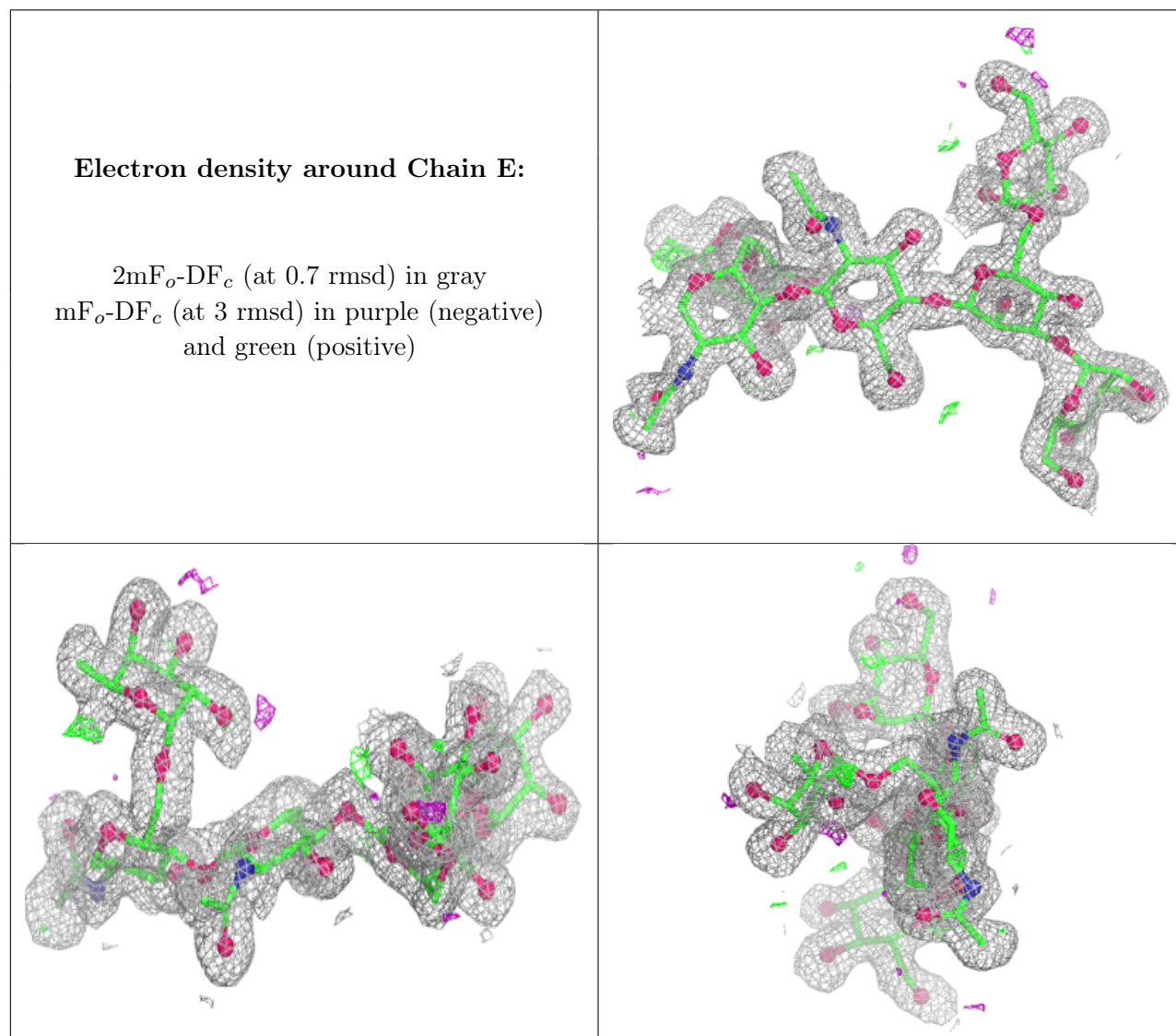
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	CSO	C	150	7/8	0.94	0.07	6,8,10,13	0
2	CSO	D	150	7/8	0.94	0.08	5,7,8,12	0

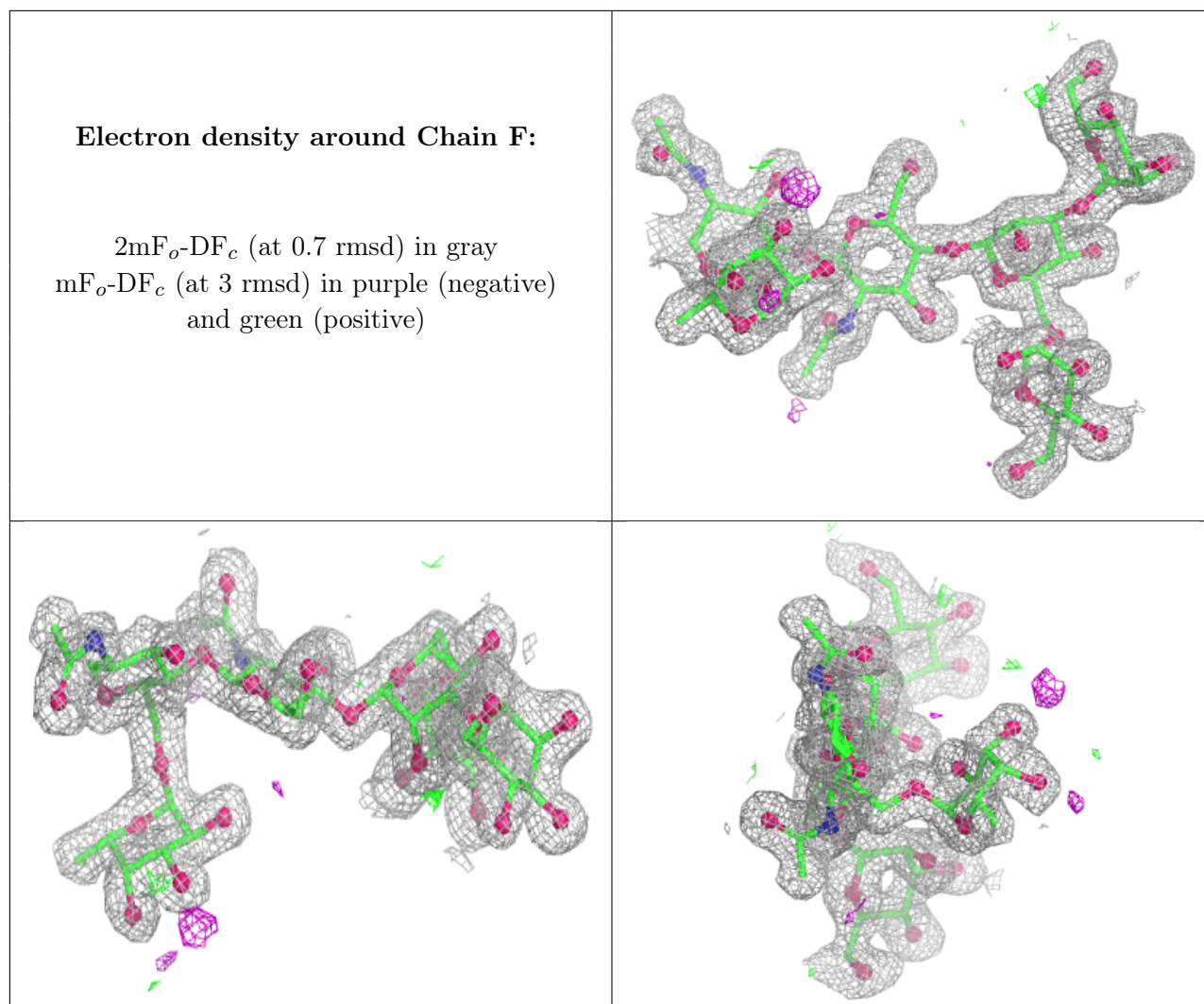
## 6.3 Carbohydrates [i](#)

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
3	MAN	F	4	11/12	0.84	0.12	20,22,26,28	0
3	MAN	E	4	11/12	0.86	0.11	22,26,27,27	0
3	FUC	E	6	10/11	0.92	0.08	14,16,17,17	0
3	NAG	F	1	14/15	0.93	0.08	8,10,17,18	0
3	NAG	E	2	14/15	0.93	0.08	7,9,10,10	0
3	NAG	E	1	14/15	0.94	0.08	8,9,12,15	0
3	FUC	F	6	10/11	0.94	0.08	11,13,16,17	0
3	BMA	F	3	11/12	0.95	0.07	10,11,14,17	0
3	BMA	E	3	11/12	0.95	0.07	9,12,14,19	0
3	MAN	F	5	11/12	0.95	0.07	11,13,16,16	0
3	NAG	F	2	14/15	0.95	0.06	6,9,11,11	0
3	MAN	E	5	11/12	0.96	0.06	13,15,16,16	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.





## 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.

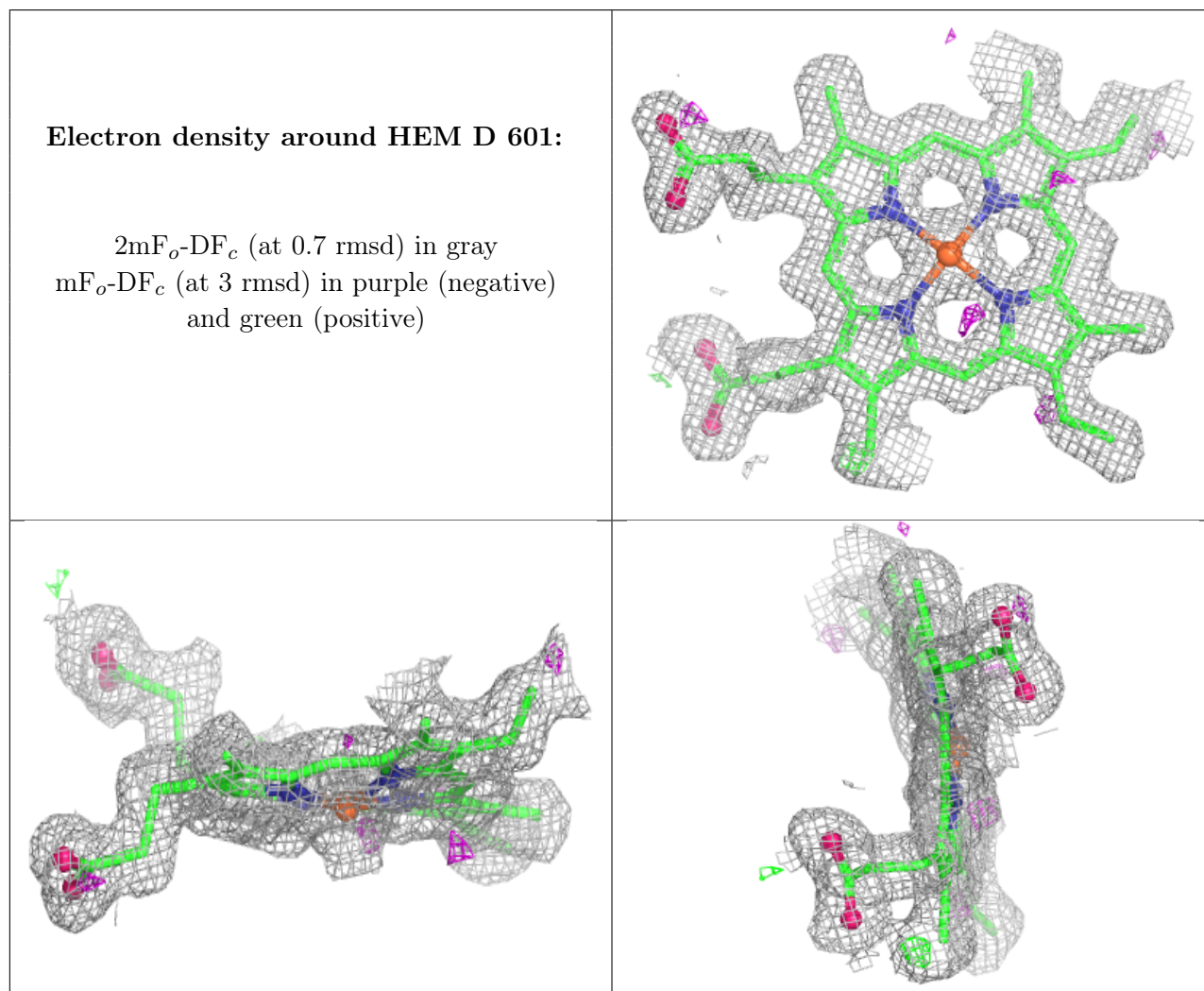
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	SO4	B	202	5/5	0.64	0.27	46,46,47,47	5
9	ACT	D	606	4/4	0.70	0.16	35,35,35,36	0
9	ACT	C	606	4/4	0.75	0.14	23,25,25,26	0
9	ACT	C	608	4/4	0.78	0.16	28,28,29,30	0
7	NAG	D	603	14/15	0.78	0.15	30,33,36,37	0
7	NAG	C	603	14/15	0.79	0.15	20,27,35,37	0
9	ACT	C	607	4/4	0.82	0.12	17,19,19,22	0
9	ACT	D	605	4/4	0.83	0.13	20,20,22,23	0

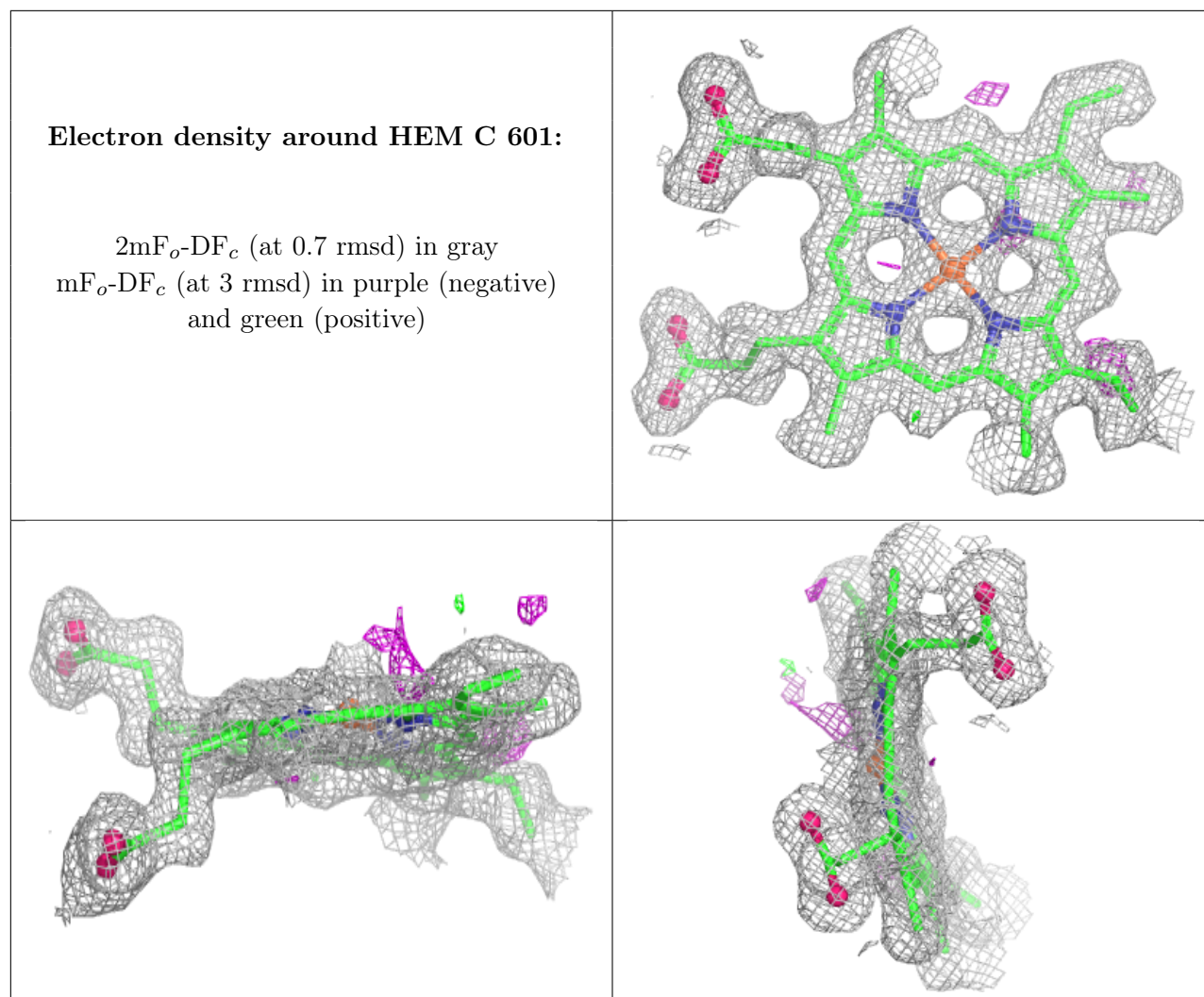
*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
9	ACT	D	607	4/4	0.86	0.12	28,29,29,30	0
7	NAG	C	602	14/15	0.87	0.11	13,17,21,22	0
7	NAG	D	602	14/15	0.88	0.11	14,23,28,31	0
5	SO4	C	605	5/5	0.92	0.12	24,24,28,29	0
5	SO4	A	1602	5/5	0.94	0.09	24,24,26,29	0
6	HEM	D	601	43/43	0.96	0.07	2,7,9,9	0
6	HEM	C	601	43/43	0.97	0.06	4,6,8,9	0
8	CA	C	604	1/1	1.00	0.02	8,8,8,8	0
8	CA	D	604	1/1	1.00	0.02	7,7,7,7	0
4	CL	B	201	1/1	1.00	0.01	7,7,7,7	0
4	CL	A	1601	1/1	1.00	0.02	6,6,6,6	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.





## 6.5 Other polymers [i](#)

There are no such residues in this entry.