



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 17, 2023 – 11:34 AM EDT

PDB ID : 4Y4F  
Title : Crystal structure of the mCD1d/GCK127/iNKTCR ternary complex  
Authors : Zajonc, D.M.; Yu, E.D.  
Deposited on : 2015-02-10  
Resolution : 3.19 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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.35.1  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.35.1

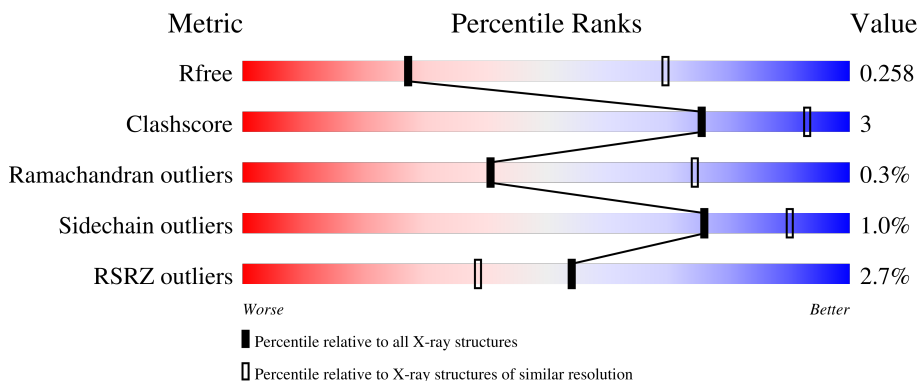
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.19 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	285	<div style="display: flex; align-items: center;"> <div style="width: 4%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 85%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 9%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 6%; height: 10px; background-color: grey;"></div> </div> <p style="text-align: center;">85%      9%      6%</p>
1	E	285	<div style="display: flex; align-items: center;"> <div style="width: 8%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 84%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 10%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 6%; height: 10px; background-color: grey;"></div> </div> <p style="text-align: center;">84%      10%      6%</p>
2	B	99	<div style="display: flex; align-items: center;"> <div style="width: 4%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 92%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 5%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 1%; height: 10px; background-color: grey;"></div> </div> <p style="text-align: center;">92%      5%      .</p>
2	F	99	<div style="display: flex; align-items: center;"> <div style="width: 89%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 8%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 3%; height: 10px; background-color: grey;"></div> </div> <p style="text-align: center;">89%      8%      .</p>
3	C	209	<div style="display: flex; align-items: center;"> <div style="width: 89%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 8%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 3%; height: 10px; background-color: grey;"></div> </div> <p style="text-align: center;">89%      8%      .</p>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	G	209	 88% 9% •
4	D	241	 % 88% 11% •
4	H	241	 89% 10% ••
5	I	2	 100%
5	J	2	 100%

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 12597 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 Antigen-presenting glycoprotein CD1d1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	267	Total	C	N	O	S	0	0	0
			2076	1331	345	387	13			
1	E	267	Total	C	N	O	S	0	0	0
			2025	1292	337	383	13			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	201	HIS	ASP	variant	UNP P11609
A	280	HIS	-	expression tag	UNP P11609
A	281	HIS	-	expression tag	UNP P11609
A	282	HIS	-	expression tag	UNP P11609
A	283	HIS	-	expression tag	UNP P11609
A	284	HIS	-	expression tag	UNP P11609
A	285	HIS	-	expression tag	UNP P11609
E	201	HIS	ASP	variant	UNP P11609
E	280	HIS	-	expression tag	UNP P11609
E	281	HIS	-	expression tag	UNP P11609
E	282	HIS	-	expression tag	UNP P11609
E	283	HIS	-	expression tag	UNP P11609
E	284	HIS	-	expression tag	UNP P11609
E	285	HIS	-	expression tag	UNP P11609

- Molecule 2 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	96	Total	C	N	O	S	0	0	0
			755	483	125	141	6			
2	F	96	Total	C	N	O	S	0	0	0
			753	481	127	139	6			

- Molecule 3 is a protein called Chimeric TCR Valpha14/Jalpha18 chain (mouse variable do-

main/ human constant domain).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	202	Total 1537	C 951	N 262	O 316	S 8	0	1	0
3	G	202	Total 1521	C 941	N 259	O 313	S 8	0	0	0

- Molecule 4 is a protein called Chimeric TCR Vbeta8.2 chain (mouse variable domain, human constant domain).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	239	Total 1858	C 1165	N 331	O 356	S 6	0	0	0
4	H	239	Total 1862	C 1168	N 332	O 356	S 6	0	0	0

- Molecule 5 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



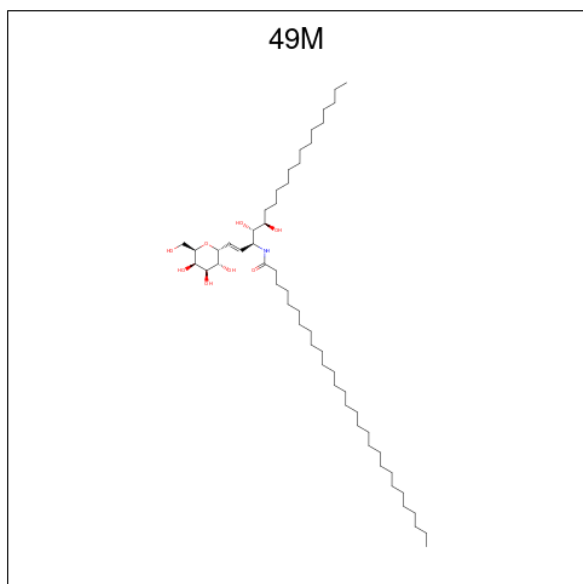
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
5	I	2	Total 28	C 16	N 2	O 10	0	0	0
5	J	2	Total 28	C 16	N 2	O 10	0	0	0

- Molecule 6 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
			Total	C	N	O		
6	A	1	14	8	1	5	0	0
6	E	1	14	8	1	5	0	0

- Molecule 7 is (1R)-1,5-anhydro-1-[(1E,3S,4S,5R)-4,5-dihydroxy-3-(nonacosanoylamino)nona dec-1-en-1-yl]-D-galactitol (three-letter code: 49M) (formula: C<sub>54</sub>H<sub>105</sub>NO<sub>8</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
7	A	1	63	54	1	8	0	0

*Continued on next page...*

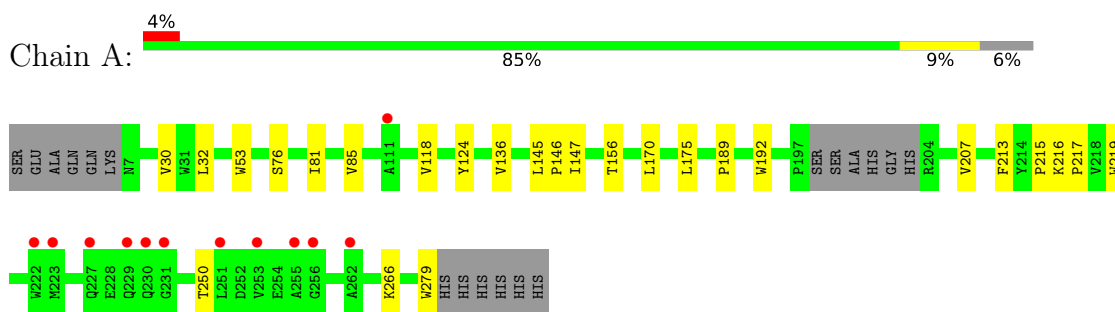
*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
7	E	1	63	54	1	8	0	0

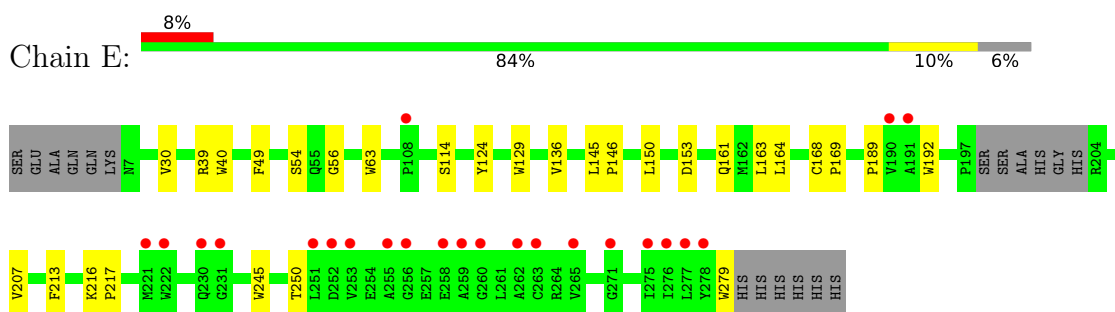
### 3 Residue-property plots [i](#)

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.

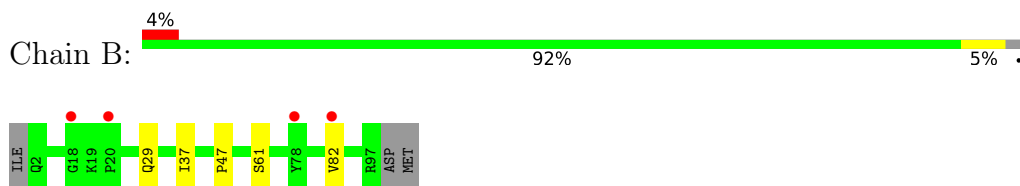
- Molecule 1: Antigen-presenting glycoprotein CD1d1



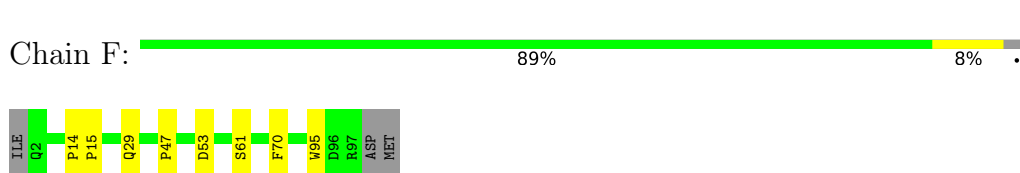
- Molecule 1: Antigen-presenting glycoprotein CD1d1



- Molecule 2: Beta-2-microglobulin




- Molecule 2: Beta-2-microglobulin




- Molecule 3: Chimeric TCR Valpha14/Jalpha18 chain (mouse variable domain/ human constant domain)

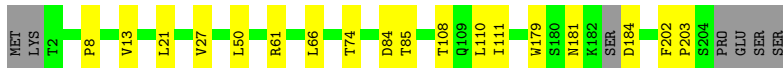


Chain C:  89% 8%




- Molecule 3: Chimeric TCR Valpha14/Jalpha18 chain (mouse variable domain/ human constant domain)

Chain G:  88% 9%




- Molecule 4: Chimeric TCR Vbeta8.2 chain (mouse variable domain, human constant domain)

Chain D:  88% 11%



- Molecule 4: Chimeric TCR Vbeta8.2 chain (mouse variable domain, human constant domain)

Chain H:  89% 10%

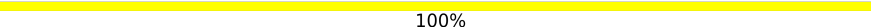


- Molecule 5: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain I:  100%



- Molecule 5: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain J:  100%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	78.59Å 149.73Å 101.39Å 90.00° 96.46° 90.00°	Depositor
Resolution (Å)	35.09 – 3.19 35.09 – 3.19	Depositor EDS
% Data completeness (in resolution range)	98.5 (35.09-3.19) 98.6 (35.09-3.19)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.90 (at 3.18Å)	Xtrriage
Refinement program	REFMAC 5.6.0104	Depositor
R, $R_{free}$	0.218 , 0.257 0.217 , 0.258	Depositor DCC
$R_{free}$ test set	1190 reflections (3.11%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	78.6	Xtrriage
Anisotropy	0.027	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 34.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	12597	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	90.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.36% 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: NAG, 49M

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	0.61	3/2138 (0.1%)	0.50	0/2919
1	E	0.62	6/2085 (0.3%)	0.49	0/2854
2	B	0.53	0/781	0.48	0/1070
2	F	0.53	1/779 (0.1%)	0.50	0/1068
3	C	0.43	1/1567 (0.1%)	0.48	0/2134
3	G	0.43	1/1548 (0.1%)	0.50	0/2111
4	D	0.51	3/1909 (0.2%)	0.49	0/2606
4	H	0.51	1/1913 (0.1%)	0.49	0/2610
All	All	0.53	16/12720 (0.1%)	0.49	0/17372

All (16) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	35	TRP	CD2-CE2	5.20	1.47	1.41
1	E	40	TRP	CD2-CE2	5.20	1.47	1.41
1	A	192	TRP	CD2-CE2	5.17	1.47	1.41
3	G	179	TRP	CD2-CE2	5.15	1.47	1.41
4	D	237	TRP	CD2-CE2	5.10	1.47	1.41
1	E	245	TRP	CD2-CE2	5.07	1.47	1.41
1	A	279	TRP	CD2-CE2	5.07	1.47	1.41
1	E	129	TRP	CD2-CE2	5.06	1.47	1.41
1	E	63	TRP	CD2-CE2	5.05	1.47	1.41
4	D	220	TRP	CD2-CE2	5.05	1.47	1.41
4	H	237	TRP	CD2-CE2	5.04	1.47	1.41
4	D	34	TRP	CD2-CE2	5.04	1.47	1.41
1	E	192	TRP	CD2-CE2	5.02	1.47	1.41
2	F	95	TRP	CD2-CE2	5.02	1.47	1.41
1	A	53	TRP	CD2-CE2	5.01	1.47	1.41
1	E	279	TRP	CD2-CE2	5.01	1.47	1.41

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2076	0	1924	14	0
1	E	2025	0	1810	17	0
2	B	755	0	684	2	0
2	F	753	0	680	3	0
3	C	1537	0	1440	7	0
3	G	1521	0	1410	9	0
4	D	1858	0	1741	12	0
4	H	1862	0	1752	12	0
5	I	28	0	25	0	0
5	J	28	0	25	1	0
6	A	14	0	13	0	0
6	E	14	0	13	0	0
7	A	63	0	105	4	0
7	E	63	0	105	4	0
All	All	12597	0	11727	75	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 3.

All (75) 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)
4:H:115:ARG:HB2	4:H:222:GLN:HE22	1.50	0.76
4:D:204:HIS:NE2	4:D:235:GLU:HG3	2.01	0.74
1:E:49:PHE:CD1	1:E:54:SER:HB2	2.23	0.72
1:E:30:VAL:HG21	7:E:504:49M:H80	1.72	0.70
1:E:49:PHE:HD1	1:E:54:SER:HB2	1.59	0.67
1:A:219:TRP:HB3	1:A:266:LYS:HB2	1.78	0.66
1:E:189:PRO:HB3	1:E:213:PHE:HB3	1.79	0.64
1:A:30:VAL:HG21	7:A:504:49M:H80	1.82	0.61
2:F:29:GLN:HA	2:F:61:SER:HB2	1.84	0.60

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:145:LEU:HB3	1:E:146:PRO:HD3	1.82	0.60
3:C:85:THR:HG23	3:C:111:ILE:HA	1.84	0.59
3:G:85:THR:HG23	3:G:111:ILE:HA	1.85	0.57
1:A:216:LYS:N	1:A:217:PRO:HD2	2.21	0.56
1:E:207:VAL:HG22	1:E:250:THR:HG22	1.88	0.56
1:A:124:TYR:CZ	1:A:136:VAL:HG11	2.41	0.55
1:A:189:PRO:HB3	1:A:213:PHE:HB3	1.88	0.54
1:A:215:PRO:HB2	1:A:217:PRO:HD2	1.89	0.54
1:E:150:LEU:HD11	7:E:504:49M:H13	1.90	0.54
4:H:127:PRO:HD3	4:H:140:LEU:HG	1.89	0.54
4:H:150:ASP:HB2	4:H:173:PRO:HG2	1.91	0.53
1:A:81:ILE:O	1:A:85:VAL:HG23	2.10	0.52
4:D:199:GLN:HA	4:D:239:ARG:O	2.10	0.52
1:E:161:GLN:HB3	5:J:1:NAG:H82	1.93	0.51
2:B:29:GLN:HA	2:B:61:SER:HB2	1.91	0.51
4:D:124:LEU:HD11	4:D:140:LEU:HD23	1.92	0.51
1:A:156:THR:HG23	7:A:504:49M:H52	1.92	0.50
1:E:49:PHE:HD1	1:E:54:SER:CB	2.24	0.50
4:D:150:ASP:HB2	4:D:173:PRO:HG2	1.94	0.49
7:A:504:49M:H38	7:A:504:49M:H43	1.93	0.49
3:C:56:LYS:HG2	3:C:65:THR:HG22	1.94	0.49
1:E:153:ASP:OD2	7:E:504:49M:O2A	2.29	0.48
4:H:214:LEU:HD13	4:H:227:PRO:HG2	1.95	0.48
3:C:61:ARG:NH2	3:C:84:ASP:OD2	2.46	0.48
4:H:21:LEU:HD12	4:H:76:LEU:HD23	1.95	0.48
1:A:118:VAL:HG11	7:A:504:49M:H11	1.95	0.48
1:E:163:LEU:HB3	7:E:504:49M:H101	1.95	0.47
1:A:145:LEU:HB3	1:A:146:PRO:HD3	1.95	0.47
1:E:168:CYS:HB3	1:E:169:PRO:HD3	1.96	0.47
3:C:50:LEU:HD13	3:C:66:LEU:HB2	1.96	0.47
4:H:204:HIS:NE2	4:H:235:GLU:HG3	2.29	0.47
3:G:21:LEU:O	3:G:74:THR:HA	2.15	0.47
2:B:37:ILE:HG12	2:B:82:VAL:HG22	1.97	0.46
4:D:214:LEU:HD13	4:D:227:PRO:HG2	1.96	0.46
4:D:152:VAL:HG12	4:D:211:PHE:HA	1.96	0.46
1:E:216:LYS:N	1:E:217:PRO:HD2	2.31	0.45
3:G:8:PRO:O	3:G:108:THR:HG23	2.17	0.44
4:D:46:ILE:HG22	4:D:47:HIS:HD2	1.83	0.44
1:E:168:CYS:HB3	1:E:169:PRO:CD	2.47	0.44
4:D:46:ILE:HG22	4:D:47:HIS:CD2	2.52	0.44
4:H:36:ARG:HH21	4:H:87:SER:HB2	1.81	0.44

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:D:115:ARG:HA	4:D:222:GLN:HE22	1.83	0.44
1:E:39:ARG:HD3	2:F:53:ASP:OD2	2.18	0.44
1:E:114:SER:HB3	1:E:164:LEU:HD22	2.00	0.43
3:G:13:VAL:HG22	3:G:110:LEU:HD11	2.01	0.43
4:H:78:LEU:HD23	4:H:85:GLN:OE1	2.18	0.43
4:H:25:GLN:HE21	4:H:27:ASN:H	1.67	0.43
4:D:21:LEU:HD22	4:D:106:THR:HG21	2.00	0.43
4:D:144:ALA:HB2	4:D:209:VAL:HG21	2.00	0.43
3:G:50:LEU:HD13	3:G:66:LEU:HB2	2.01	0.43
3:G:202:PHE:HA	3:G:203:PRO:HD3	1.91	0.42
1:A:76:SER:HA	3:C:95:ARG:NH1	2.35	0.42
3:C:156:VAL:HG13	3:C:180:SER:HB2	2.01	0.42
4:H:132:ILE:HG23	4:H:195:ALA:HB1	2.02	0.42
4:H:127:PRO:HD2	4:H:198:TRP:CZ2	2.55	0.41
3:G:184:ASP:N	3:G:184:ASP:OD1	2.53	0.41
1:A:147:ILE:HD13	1:A:147:ILE:HA	1.92	0.41
1:A:207:VAL:HG22	1:A:250:THR:HG22	2.01	0.41
3:G:61:ARG:NH2	3:G:84:ASP:OD2	2.43	0.41
1:E:124:TYR:CZ	1:E:136:VAL:HG11	2.56	0.41
4:H:36:ARG:HB3	4:H:46:ILE:HD11	2.02	0.41
4:D:117:VAL:HG12	4:D:227:PRO:HB2	2.03	0.41
1:A:32:LEU:HD22	1:A:175:LEU:HD13	2.03	0.40
2:F:14:PRO:HA	2:F:15:PRO:HD3	1.85	0.40
3:C:128:ASP:HB3	3:C:131:SER:O	2.20	0.40
3:G:21:LEU:HD23	3:G:108:THR:HB	2.02	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	263/285 (92%)	257 (98%)	6 (2%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	E	263/285 (92%)	252 (96%)	10 (4%)	1 (0%)	34	69
2	B	94/99 (95%)	92 (98%)	1 (1%)	1 (1%)	14	51
2	F	94/99 (95%)	91 (97%)	2 (2%)	1 (1%)	14	51
3	C	199/209 (95%)	188 (94%)	11 (6%)	0	100	100
3	G	198/209 (95%)	188 (95%)	9 (4%)	1 (0%)	29	67
4	D	237/241 (98%)	235 (99%)	2 (1%)	0	100	100
4	H	237/241 (98%)	233 (98%)	4 (2%)	0	100	100
All	All	1585/1668 (95%)	1536 (97%)	45 (3%)	4 (0%)	41	74

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	F	47	PRO
1	E	56	GLY
2	B	47	PRO
3	G	27	VAL

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

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	218/249 (88%)	217 (100%)	1 (0%)	88	95
1	E	204/249 (82%)	204 (100%)	0	100	100
2	B	81/93 (87%)	81 (100%)	0	100	100
2	F	80/93 (86%)	79 (99%)	1 (1%)	69	87
3	C	174/188 (93%)	171 (98%)	3 (2%)	60	83
3	G	170/188 (90%)	169 (99%)	1 (1%)	86	94
4	D	198/208 (95%)	194 (98%)	4 (2%)	55	80
4	H	199/208 (96%)	195 (98%)	4 (2%)	55	80
All	All	1324/1476 (90%)	1310 (99%)	14 (1%)	76	88

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	170	LEU
3	C	141	ASP
3	C	181[A]	ASN
3	C	181[B]	ASN
4	D	25	GLN
4	D	55	THR
4	D	190	ARG
4	D	235	GLU
2	F	70	PHE
3	G	181	ASN
4	H	25	GLN
4	H	168	CYS
4	H	190	ARG
4	H	235	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
3	C	31	ASN
3	C	125	GLN
3	C	193	ASN
4	D	47	HIS
4	D	116	ASN
4	D	217	ASN
4	D	222	GLN
1	E	36	GLN
1	E	117	HIS
3	G	31	ASN
3	G	117	GLN
4	H	25	GLN
4	H	222	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

4 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
5	NAG	I	1	5,1	14,14,15	0.48	0	17,19,21	0.97	1 (5%)
5	NAG	I	2	5	14,14,15	0.49	0	17,19,21	0.86	1 (5%)
5	NAG	J	1	5,1	14,14,15	0.52	0	17,19,21	0.75	0
5	NAG	J	2	5	14,14,15	0.56	0	17,19,21	1.10	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
5	NAG	I	1	5,1	-	0/6/23/26	0/1/1/1
5	NAG	I	2	5	-	2/6/23/26	0/1/1/1
5	NAG	J	1	5,1	-	0/6/23/26	0/1/1/1
5	NAG	J	2	5	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	J	2	NAG	C4-C3-C2	2.75	115.04	111.02
5	I	1	NAG	C1-O5-C5	2.16	115.12	112.19
5	I	2	NAG	O5-C5-C6	2.13	110.55	107.20

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	I	2	NAG	O5-C5-C6-O6

*Continued on next page...*

Continued from previous page...

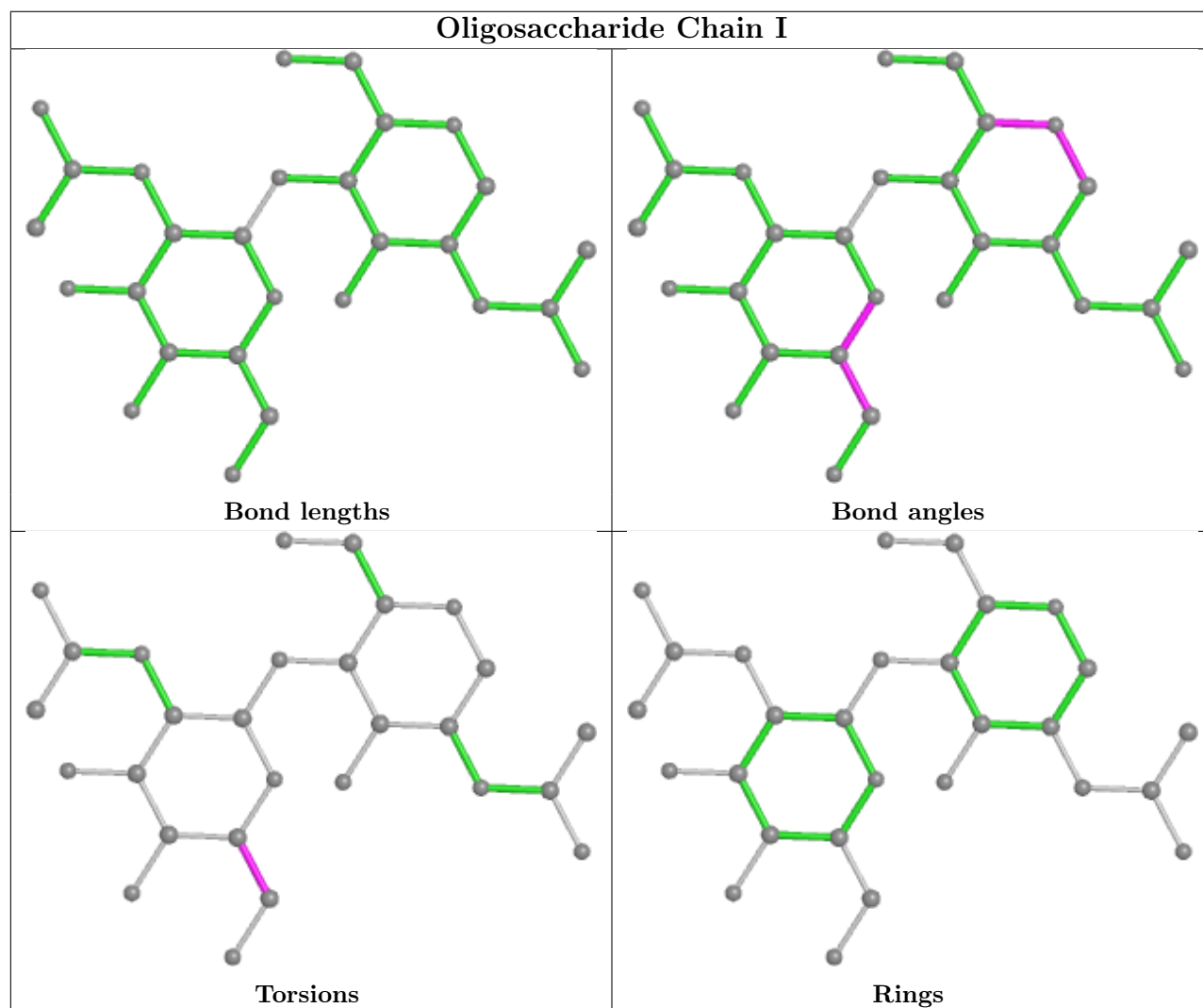
Mol	Chain	Res	Type	Atoms
5	J	2	NAG	O5-C5-C6-O6
5	I	2	NAG	C4-C5-C6-O6
5	J	2	NAG	C4-C5-C6-O6

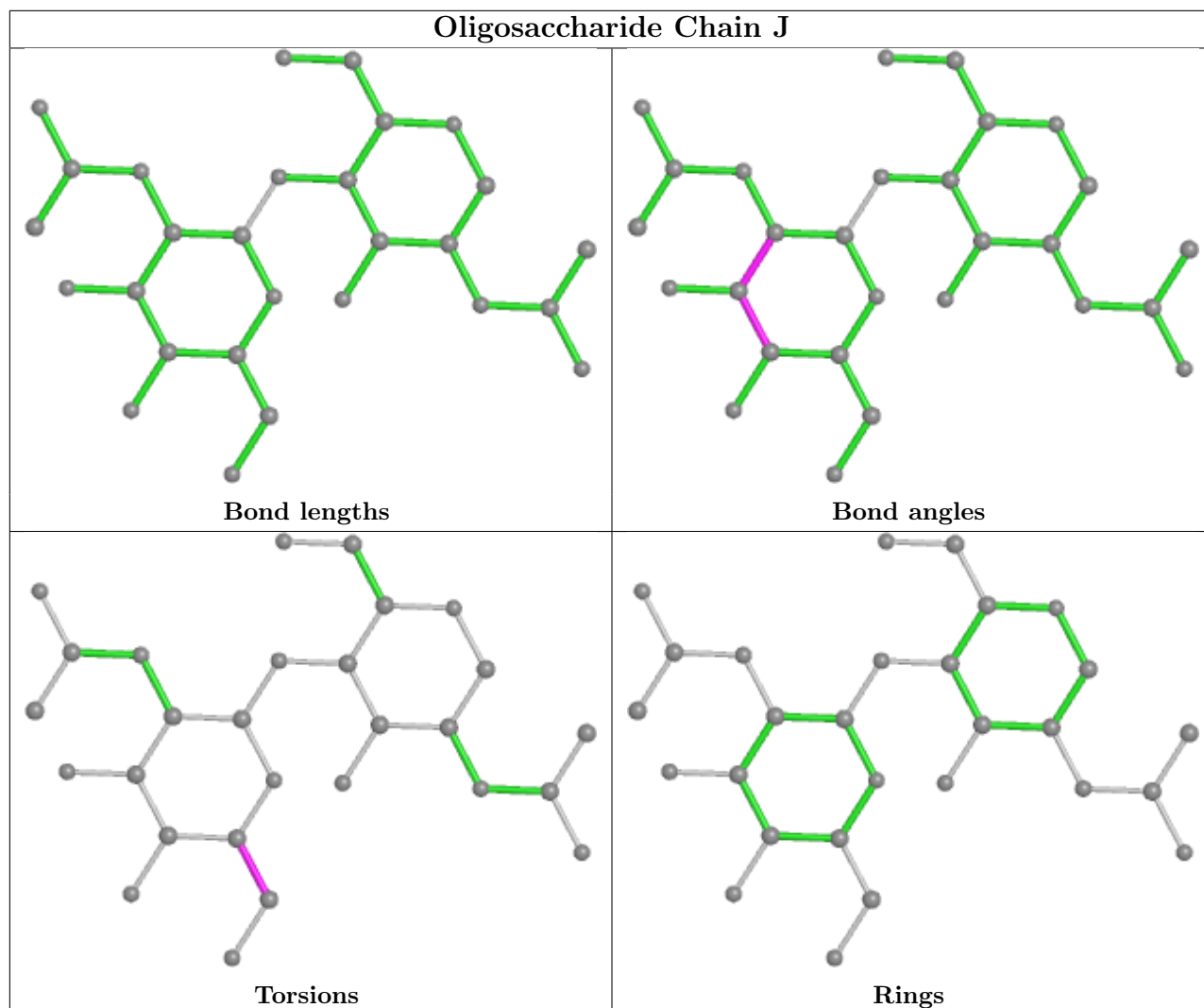
There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	J	1	NAG	1	0

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





## 5.6 Ligand geometry [i](#)

4 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$
6	NAG	E	501	1	14,14,15	0.43	0	17,19,21	1.45	1 (5%)
7	49M	E	504	-	63,63,63	0.35	0	64,72,72	0.80	2 (3%)
7	49M	A	504	-	63,63,63	0.36	0	64,72,72	0.68	2 (3%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	NAG	A	501	1	14,14,15	0.49	0	17,19,21	1.06	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	E	501	1	-	2/6/23/26	0/1/1/1
7	49M	E	504	-	-	42/61/81/81	0/1/1/1
7	49M	A	504	-	-	41/61/81/81	0/1/1/1
6	NAG	A	501	1	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	E	501	NAG	C1-O5-C5	5.03	119.01	112.19
6	A	501	NAG	C1-O5-C5	3.43	116.84	112.19
7	E	504	49M	C3A-C2A-C1A	3.09	114.02	110.22
7	E	504	49M	C5M-O6A-C1A	2.56	117.37	112.81
7	A	504	49M	C3A-C2A-C1A	2.35	113.11	110.22
7	A	504	49M	C5M-O6A-C1A	2.34	116.97	112.81

There are no chirality outliers.

All (87) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	A	504	49M	N2-C2-C3-C4
7	A	504	49M	N2-C2-C3-O3
7	A	504	49M	O4-C4-C5-C6
7	A	504	49M	O3-C3-C4-C5
7	A	504	49M	C2A-C1A-C1B-C1
7	E	504	49M	N2-C2-C3-C4
7	E	504	49M	N2-C2-C3-O3
7	E	504	49M	C1B-C1-C2-N2
7	E	504	49M	C1B-C1-C2-C3
7	E	504	49M	O4-C4-C5-C6
7	E	504	49M	C2A-C1A-C1B-C1
7	E	504	49M	O6A-C1A-C1B-C1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
7	A	504	49M	CAA-CAB-CAC-CAD
7	E	504	49M	CAA-CAB-CAC-CAD
7	E	504	49M	C1-C2-C3-O3
7	A	504	49M	CAT-CAU-CAV-CAW
7	E	504	49M	CAN-CAO-CAP-CAQ
7	A	504	49M	CAV-CAW-CAX-CAY
7	E	504	49M	C14-C15-C16-C17
7	E	504	49M	CAV-CAW-CAX-CAY
7	E	504	49M	CAW-CAX-CAY-CAZ
7	A	504	49M	CAX-CAY-CAZ-C19
7	E	504	49M	CAC-CAD-CAE-CAF
7	E	504	49M	C12-C13-C14-C15
7	E	504	49M	C7-C8-C9-C10
7	E	504	49M	C20-C19-CAZ-CAY
7	A	504	49M	C14-C15-C16-C17
7	A	504	49M	C6-C7-C8-C9
7	A	504	49M	C11-C10-C9-C8
7	E	504	49M	C6-C7-C8-C9
7	A	504	49M	CAG-CAH-CAI-CAJ
7	A	504	49M	C7-C8-C9-C10
7	A	504	49M	CAH-CAI-CAJ-CAK
7	E	504	49M	C11-C12-C13-C14
7	E	504	49M	CAP-CAQ-CAR-CAS
7	A	504	49M	C10-C11-C12-C13
7	E	504	49M	CAT-CAU-CAV-CAW
7	A	504	49M	CAW-CAX-CAY-CAZ
7	A	504	49M	C20-C19-CAZ-CAY
7	E	504	49M	C13-C14-C15-C16
7	E	504	49M	CAD-CAE-CAF-CAG
7	E	504	49M	C10-C11-C12-C13
7	A	504	49M	C11-C12-C13-C14
7	A	504	49M	CAC-CAD-CAE-CAF
7	E	504	49M	C11-C10-C9-C8
7	A	504	49M	CAS-CAT-CAU-CAV
7	A	504	49M	C1-C2-C3-O3
7	E	504	49M	CAM-CAN-CAO-CAP
7	A	504	49M	C2-C3-C4-O4
7	A	504	49M	CAL-CAM-CAN-CAO
7	A	504	49M	CAF-CAG-CAH-CAI
7	E	504	49M	CAI-CAJ-CAK-CAL
7	A	504	49M	CAM-CAN-CAO-CAP
7	E	504	49M	CAX-CAY-CAZ-C19

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
7	A	504	49M	CAP-CAQ-CAR-CAS
7	A	504	49M	CAJ-CAK-CAL-CAM
7	E	504	49M	CAS-CAT-CAU-CAV
7	E	504	49M	CAK-CAL-CAM-CAN
7	E	504	49M	CAH-CAI-CAJ-CAK
7	A	504	49M	CAZ-C19-C20-C21
7	E	504	49M	CAZ-C19-C20-C21
7	E	504	49M	CAR-CAS-CAT-CAU
7	A	504	49M	C9-C10-C11-C12
7	A	504	49M	C5-C6-C7-C8
6	A	501	NAG	C4-C5-C6-O6
6	E	501	NAG	C4-C5-C6-O6
7	A	504	49M	C4A-C5M-C6A-O5A
7	A	504	49M	CAK-CAL-CAM-CAN
7	E	504	49M	CAB-CAC-CAD-CAE
7	A	504	49M	C1-C2-C3-C4
7	E	504	49M	C1-C2-C3-C4
7	A	504	49M	C12-C13-C14-C15
7	E	504	49M	C5-C6-C7-C8
7	E	504	49M	C15-C16-C17-C18
7	A	504	49M	CAI-CAJ-CAK-CAL
7	A	504	49M	CAD-CAE-CAF-CAG
6	A	501	NAG	O5-C5-C6-O6
7	E	504	49M	C9-C10-C11-C12
7	A	504	49M	O3-C3-C4-O4
7	A	504	49M	CAE-CAF-CAG-CAH
6	E	501	NAG	O5-C5-C6-O6
7	E	504	49M	CAE-CAF-CAG-CAH
7	E	504	49M	C2-C3-C4-O4
7	E	504	49M	C3-C4-C5-C6
7	A	504	49M	C2-C3-C4-C5
7	E	504	49M	CAQ-CAR-CAS-CAT
7	A	504	49M	C1B-C1-C2-N2

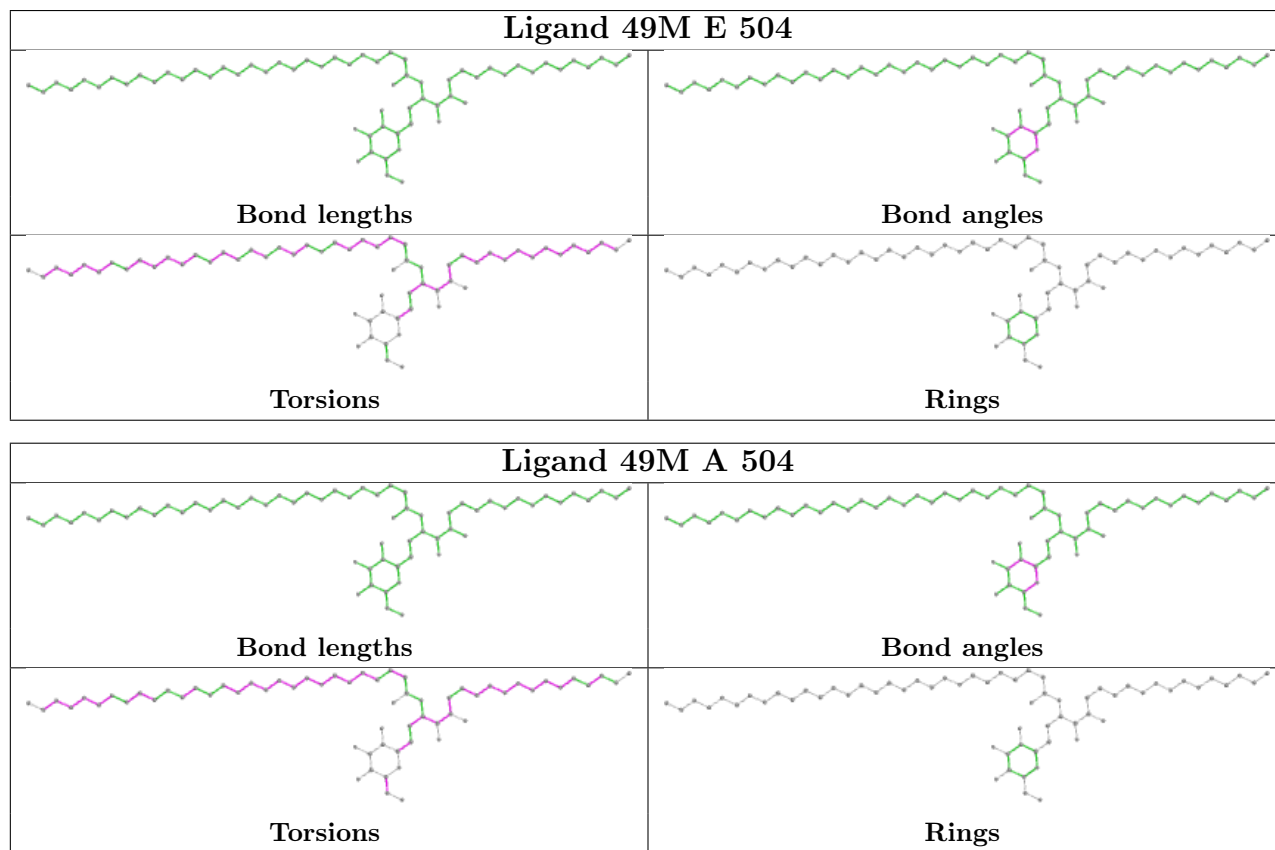
There are no ring outliers.

2 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	E	504	49M	4	0
7	A	504	49M	4	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths,

bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	267/285 (93%)	-0.00	12 (4%) 33 21	64, 89, 139, 166	0
1	E	267/285 (93%)	0.14	23 (8%) 10 5	62, 97, 172, 189	0
2	B	96/99 (96%)	0.26	4 (4%) 36 23	79, 109, 124, 133	1 (1%)
2	F	96/99 (96%)	0.01	0 100 100	68, 98, 125, 133	1 (1%)
3	C	202/209 (96%)	-0.15	1 (0%) 91 86	56, 77, 128, 141	0
3	G	202/209 (96%)	-0.26	0 100 100	59, 75, 122, 134	0
4	D	239/241 (99%)	-0.19	2 (0%) 86 78	61, 78, 108, 118	0
4	H	239/241 (99%)	-0.21	1 (0%) 92 89	65, 79, 110, 131	0
All	All	1608/1668 (96%)	-0.07	43 (2%) 54 39	56, 85, 135, 189	2 (0%)

All (43) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	256	GLY	5.4
1	A	262	ALA	5.4
1	E	259	ALA	4.0
1	E	253	VAL	3.9
1	E	222	TRP	3.8
1	E	260	GLY	3.7
4	D	179	ALA	3.5
1	A	222	TRP	3.5
4	H	240	ALA	3.4
1	E	255	ALA	3.3
1	E	108	PRO	3.3
1	E	276	ILE	3.2
1	E	231	GLY	3.1
1	E	275	ILE	3.0
1	E	191	ALA	3.0
1	E	252	ASP	3.0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	E	251	LEU	3.0
1	E	277	LEU	2.9
1	A	231	GLY	2.9
1	A	253	VAL	2.8
1	A	255	ALA	2.8
1	E	263	CYS	2.8
1	A	256	GLY	2.8
1	E	271	GLY	2.5
2	B	18	GLY	2.4
1	A	223	MET	2.4
1	E	190	VAL	2.4
1	E	262	ALA	2.3
1	E	278	TYR	2.3
1	E	221	MET	2.3
1	A	230	GLN	2.3
1	A	251	LEU	2.3
1	A	227	GLN	2.3
2	B	20	PRO	2.3
1	E	265	VAL	2.3
1	E	258	GLU	2.2
4	D	221	THR	2.2
1	A	111	ALA	2.2
2	B	82	VAL	2.2
1	E	230	GLN	2.2
1	A	229	GLN	2.1
3	C	191	PHE	2.1
2	B	78	TYR	2.0

## 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.

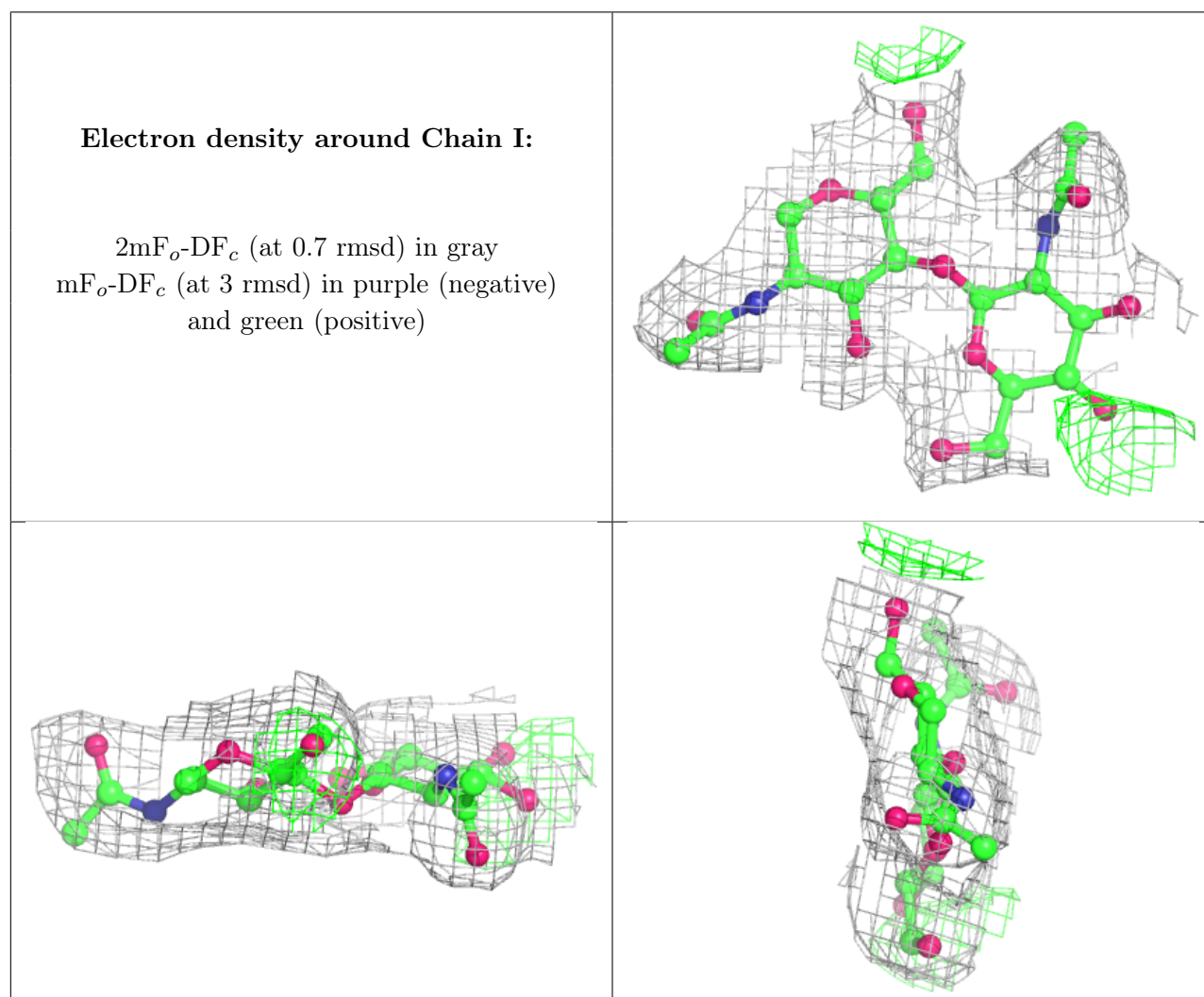
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q < 0.9
5	NAG	J	2	14/15	0.90	0.29	94,100,103,105	0

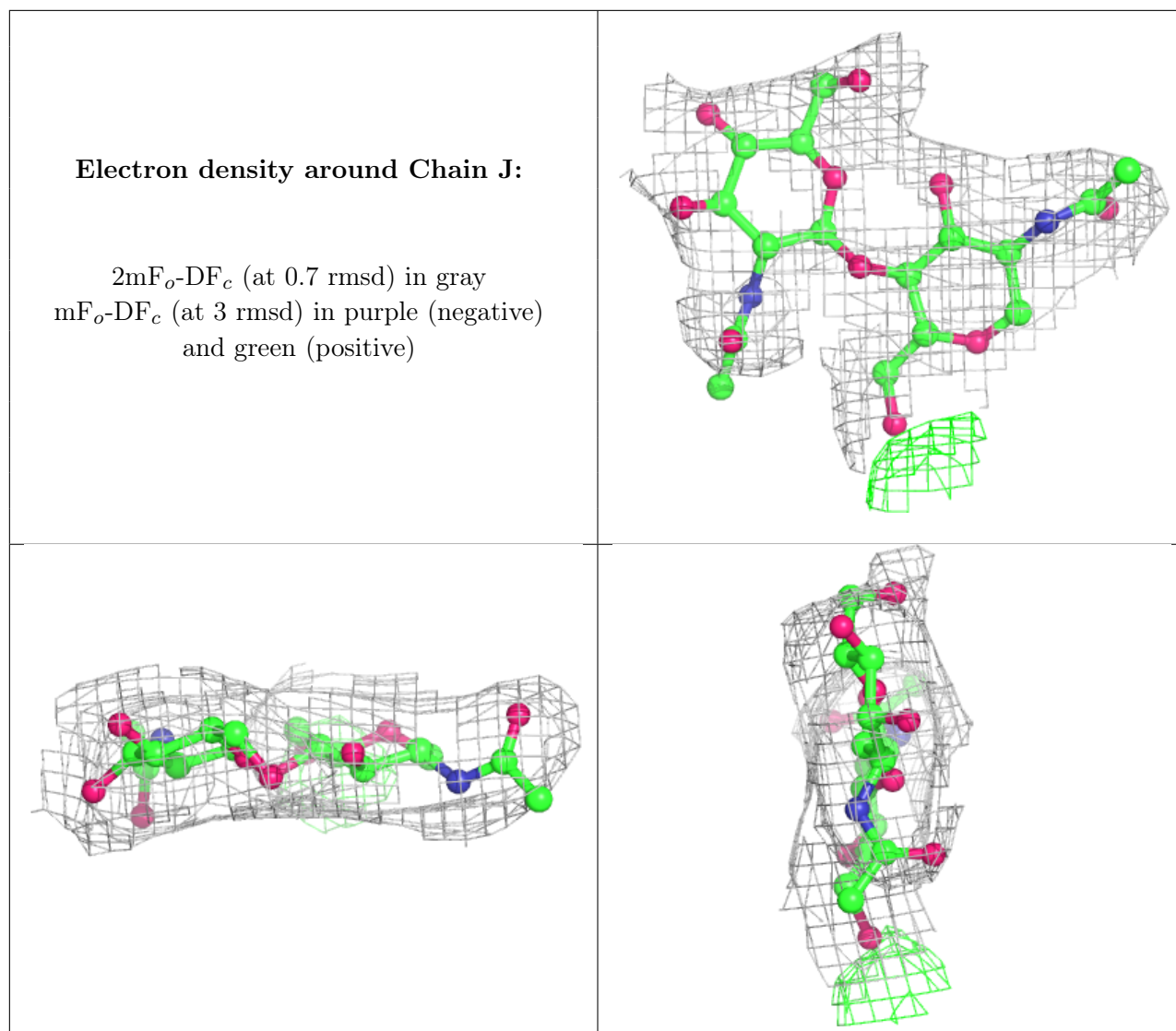
*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	NAG	I	1	14/15	0.91	0.21	74,76,78,79	0
5	NAG	J	1	14/15	0.92	0.15	81,83,86,89	0
5	NAG	I	2	14/15	0.94	0.20	82,85,87,87	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.





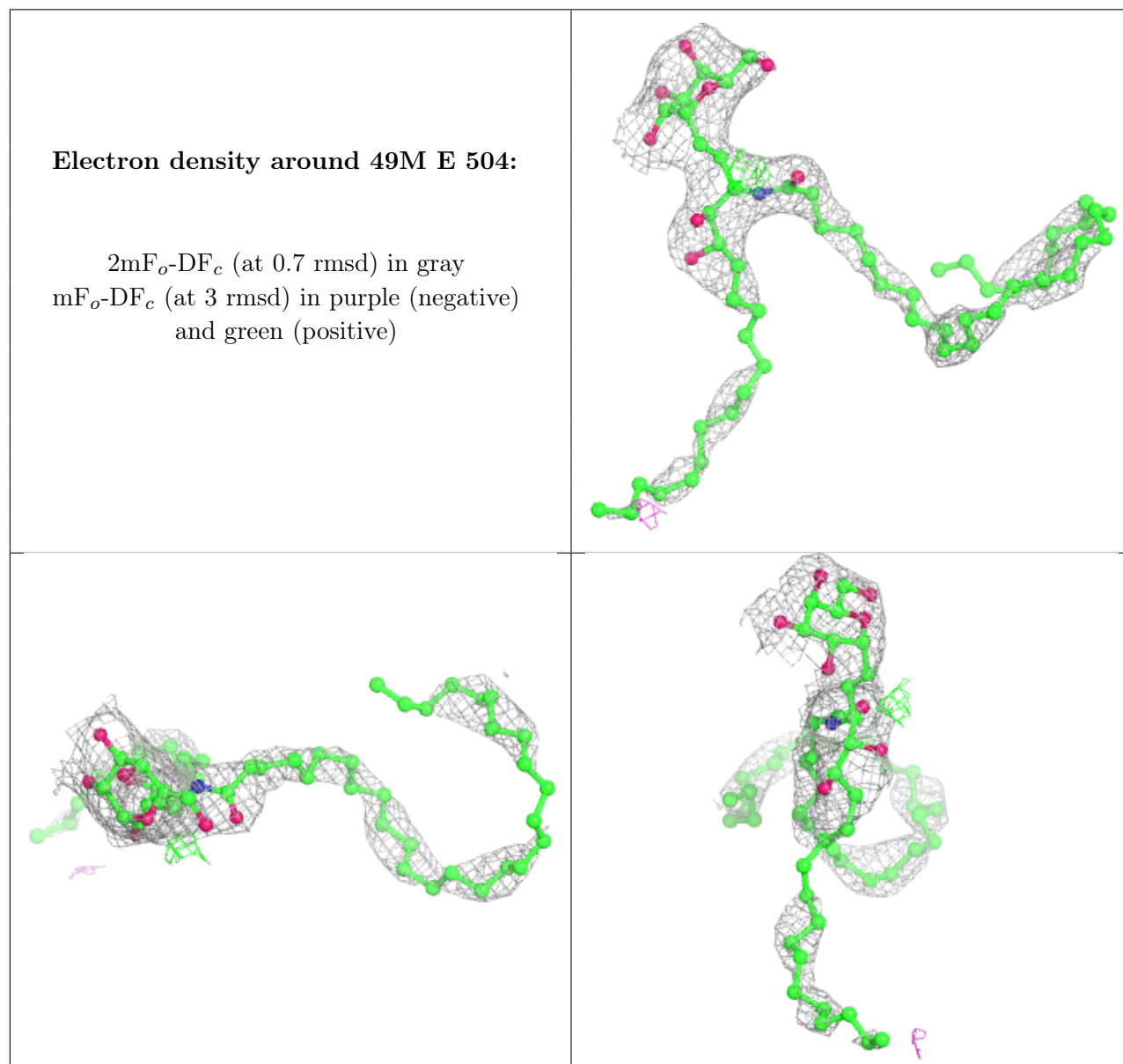
## 6.4 Ligands [i](#)

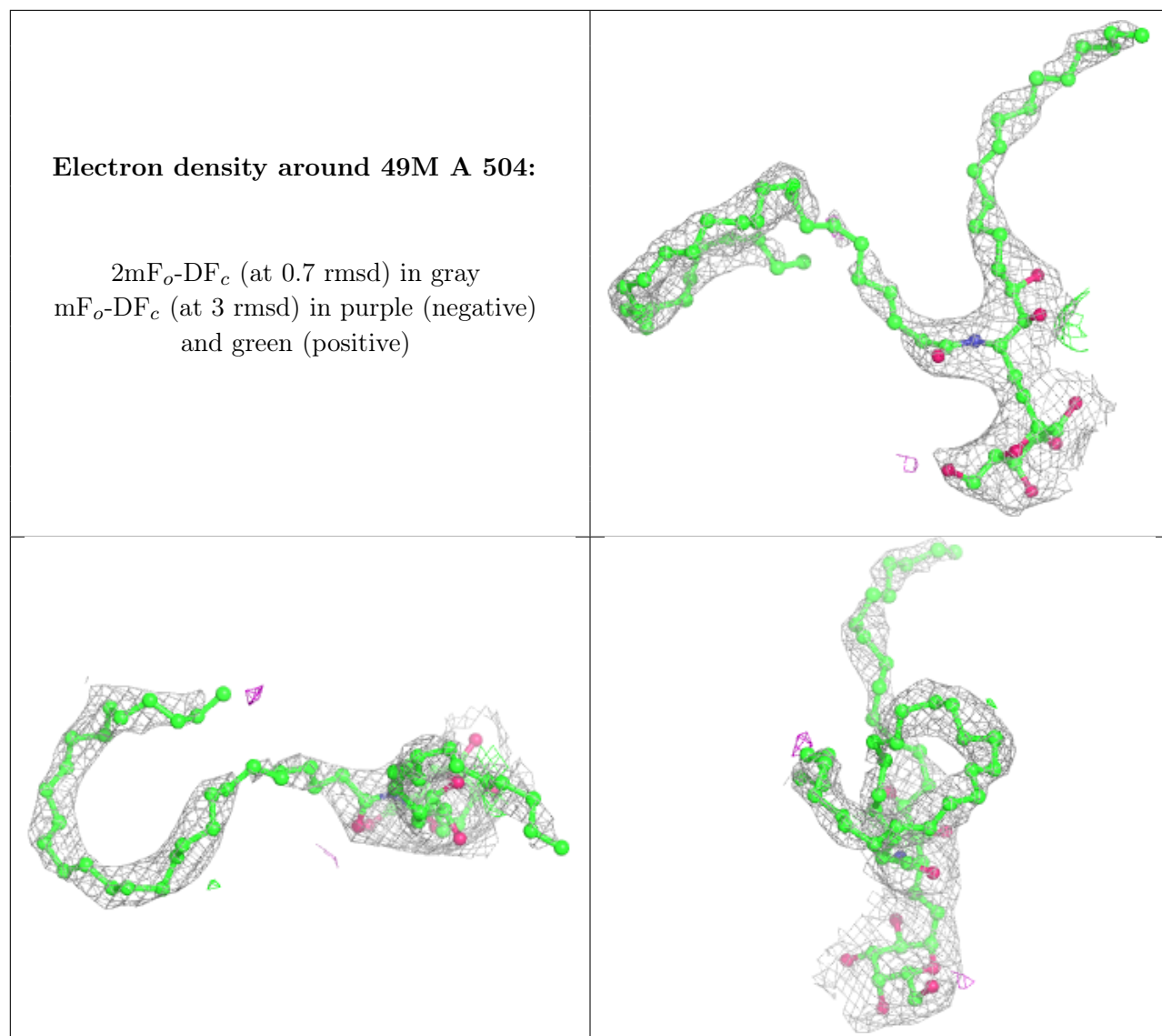
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
6	NAG	A	501	14/15	0.89	0.23	87,89,91,92	0
6	NAG	E	501	14/15	0.90	0.25	89,91,93,94	0
7	49M	E	504	63/63	0.91	0.36	71,79,94,98	0
7	49M	A	504	63/63	0.93	0.30	62,67,72,73	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.