



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 11, 2023 – 08:26 am GMT

PDB ID : 8AE5  
Title : Crystal structure of human legumain in complex with macrocypin 1a  
Authors : Elamin, T.; Brandstetter, H.; Dall, E.  
Deposited on : 2022-07-12  
Resolution : 2.29 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 : 1.8.4, CSD as541be (2020)  
Xtrriage (Phenix) : 1.13  
EDS : 2.36  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

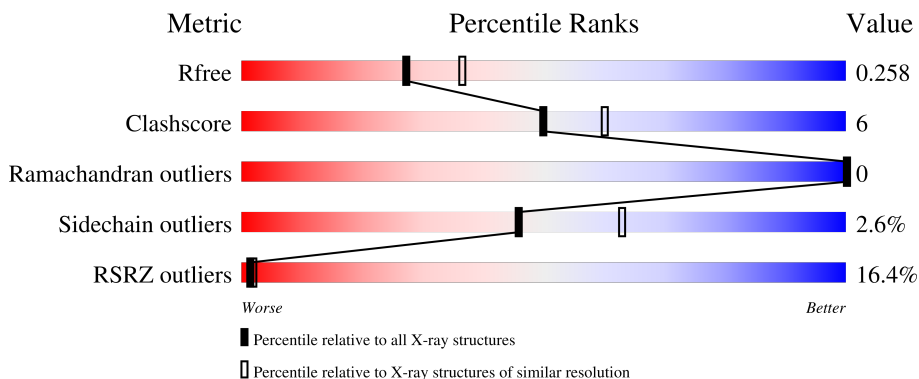
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.29 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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	263	 22% 81% 19%
1	B	263	 23% 82% 18%
2	C	177	 6% 79% 16% 5%
2	D	177	 6% 81% 9% 10%
3	E	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:

<b>Mol</b>	<b>Type</b>	<b>Chain</b>	<b>Res</b>	<b>Chirality</b>	<b>Geometry</b>	<b>Clashes</b>	<b>Electron density</b>
3	NAG	E	1	-	-	-	X

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 7053 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 Legumain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	263	Total	C	N	O	S	7	0	0
			2115	1335	361	403	16			
1	B	263	Total	C	N	O	S	41	0	0
			2115	1335	361	403	16			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	147	SNN	ASP	modified residue	UNP Q99538
B	147	SNN	ASP	modified residue	UNP Q99538

- Molecule 2 is a protein called Macrocypin-1a.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	168	Total	C	N	O	S	7	0	0
			1349	865	225	255	4			
2	D	160	Total	C	N	O	S	3	0	0
			1285	825	212	244	4			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	170	LEU	-	expression tag	UNP B9V973
C	171	GLU	-	expression tag	UNP B9V973
C	172	HIS	-	expression tag	UNP B9V973
C	173	HIS	-	expression tag	UNP B9V973
C	174	HIS	-	expression tag	UNP B9V973
C	175	HIS	-	expression tag	UNP B9V973
C	176	HIS	-	expression tag	UNP B9V973
C	177	HIS	-	expression tag	UNP B9V973
D	170	LEU	-	expression tag	UNP B9V973
D	171	GLU	-	expression tag	UNP B9V973

*Continued on next page...*

Continued from previous page...

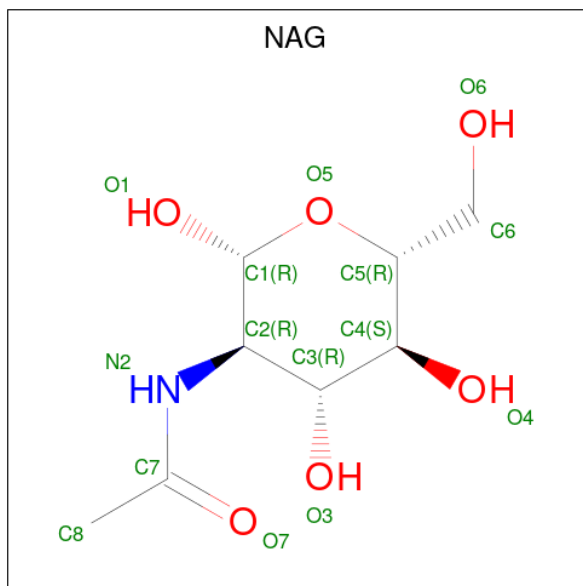
Chain	Residue	Modelled	Actual	Comment	Reference
D	172	HIS	-	expression tag	UNP B9V973
D	173	HIS	-	expression tag	UNP B9V973
D	174	HIS	-	expression tag	UNP B9V973
D	175	HIS	-	expression tag	UNP B9V973
D	176	HIS	-	expression tag	UNP B9V973
D	177	HIS	-	expression tag	UNP B9V973

- Molecule 3 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			
3	E	2	28	16	2	10	0	0	0

- Molecule 4 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>).



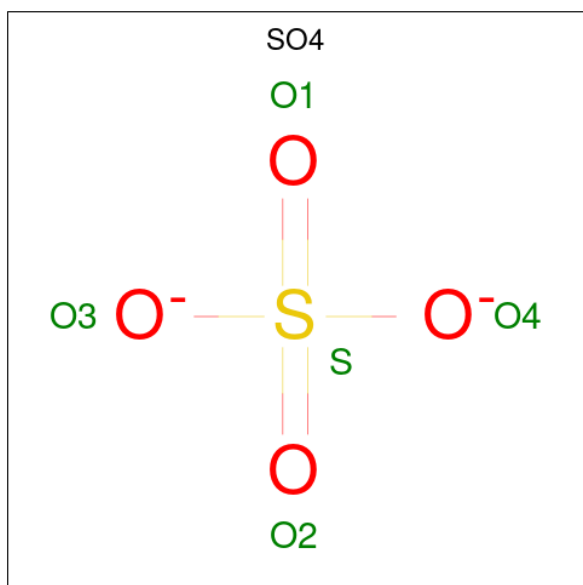
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
4	A	1	14	8	1	5	0	0
4	A	1	14	8	1	5	0	0

Continued on next page...

Continued from previous page...

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

- 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	0	0
			5	4	1		
5	C	1	Total	O	S	0	0
			5	4	1		
5	B	1	Total	O	S	0	0
			5	4	1		
5	D	1	Total	O	S	0	0
			5	4	1		
5	D	1	Total	O	S	0	0
			5	4	1		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	17	Total	O	0	0
			17	17		
6	C	32	Total	O	0	0
			32	32		

Continued on next page...

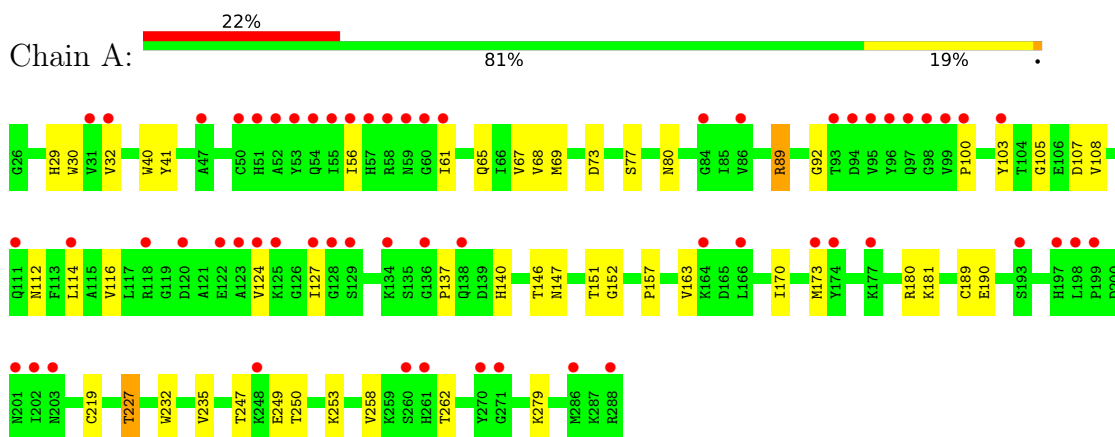
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
6	B	11	Total	O	0	0
			11	11		
6	D	20	Total	O	0	0
			20	20		

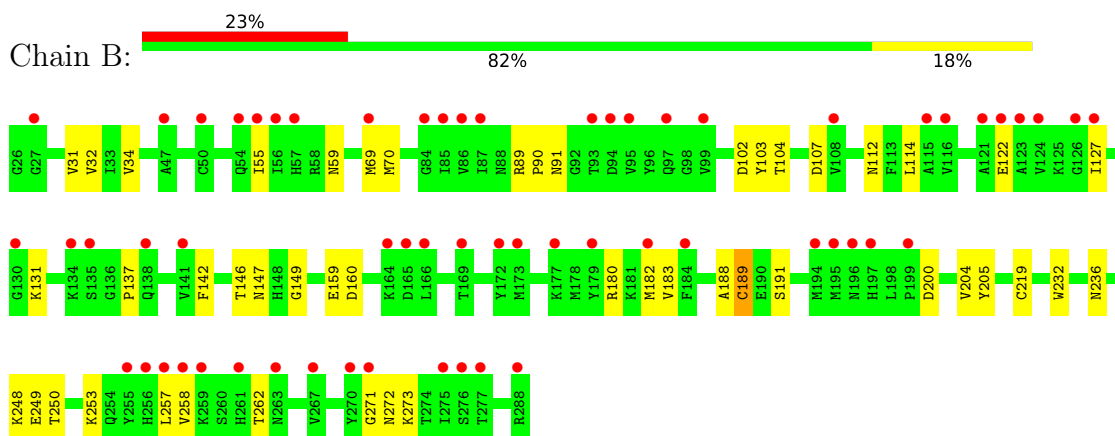
### 3 Residue-property plots [i](#)

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.

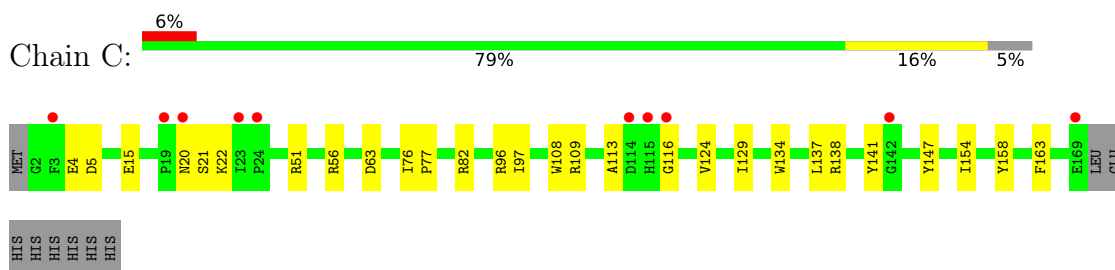
- Molecule 1: Legumain



- Molecule 1: Legumain




- Molecule 2: Macrocyprin-1a



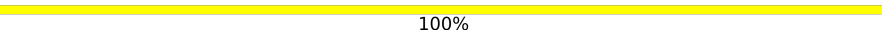


## ● Molecule 2: Macrocypin-1a

Chain D: 



## ● Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain E: 

MAG1  
MAG2

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	159.18Å 174.12Å 113.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.39 – 2.29 49.11 – 2.29	Depositor EDS
% Data completeness (in resolution range)	99.2 (47.39-2.29) 99.4 (49.11-2.29)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.95 (at 2.29Å)	Xtrriage
Refinement program	PHENIX 1.18.2_3874, PHENIX 1.18.2_3874	Depositor
R, $R_{free}$	0.221 , 0.258 0.221 , 0.258	Depositor DCC
$R_{free}$ test set	3674 reflections (5.23%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	73.8	Xtrriage
Anisotropy	0.381	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 58.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	7053	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	93.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.55% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: SCH, NAG, SNN, SO4

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	0.42	0/2154	0.58	0/2918
1	B	0.40	0/2154	0.61	0/2918
2	C	0.52	0/1394	0.69	0/1905
2	D	0.47	0/1326	0.63	0/1810
All	All	0.44	0/7028	0.62	0/9551

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
1	A	0	1
1	B	0	1
All	All	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	146	THR	Mainchain
1	B	146	THR	Mainchain

## 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	2115	0	1996	30	0
1	B	2115	0	1995	24	0
2	C	1349	0	1274	18	0
2	D	1285	0	1209	7	0
3	E	28	0	25	0	0
4	A	28	0	26	1	0
4	B	28	0	26	0	0
5	A	5	0	0	0	0
5	B	5	0	0	0	0
5	C	5	0	0	0	0
5	D	10	0	0	0	0
6	A	17	0	0	0	0
6	B	11	0	0	0	0
6	C	32	0	0	2	0
6	D	20	0	0	0	0
All	All	7053	0	6551	77	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 6.

All (77) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:250:THR:HG23	1:B:253:LYS:H	1.48	0.77
2:C:56:ARG:NH2	2:C:63:ASP:H	1.83	0.76
1:B:253:LYS:HE2	1:B:257:LEU:HD11	1.71	0.72
1:A:250:THR:HG23	1:A:253:LYS:H	1.58	0.69
1:A:170:ILE:HA	1:A:173:MET:HE2	1.78	0.65
1:A:105:GLY:O	1:A:108:VAL:HG22	1.98	0.63
1:A:247:THR:O	1:A:279:LYS:HG2	1.98	0.62
1:A:124:VAL:O	1:A:127:ILE:HG12	2.00	0.61
1:B:107:ASP:O	1:B:112:ASN:ND2	2.32	0.61
1:A:107:ASP:O	1:A:112:ASN:ND2	2.36	0.59
2:D:109:ARG:HB2	2:D:124:VAL:HB	1.83	0.59
2:C:21:SER:HB2	2:C:163:PHE:HZ	1.70	0.57

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:56:ILE:HG22	1:A:61:ILE:HB	1.87	0.56
1:A:69:MET:HG2	1:A:103:TYR:HB2	1.88	0.56
2:C:4:GLU:HG2	2:C:5:ASP:N	2.21	0.55
2:C:4:GLU:HG2	2:C:5:ASP:H	1.71	0.55
2:C:141:TYR:HB2	2:D:139:GLU:HB2	1.90	0.53
1:B:137:PRO:O	1:B:180:ARG:HB2	2.08	0.52
1:B:200:ASP:HB2	1:B:272:ASN:HA	1.91	0.52
1:A:40:TRP:O	1:A:227:THR:HG21	2.10	0.51
1:A:137:PRO:O	1:A:180:ARG:HB2	2.10	0.51
2:C:113:ALA:HB3	2:C:116:GLY:HA3	1.92	0.50
1:B:236:ASN:HD22	1:B:262:THR:HG22	1.77	0.49
1:A:152:GLY:HA2	1:A:163:VAL:CG2	2.43	0.49
2:C:56:ARG:HH22	2:C:63:ASP:H	1.61	0.49
1:A:89:ARG:HG3	1:A:92:GLY:HA3	1.96	0.48
2:C:21:SER:HB2	2:C:163:PHE:CZ	2.48	0.48
1:B:89:ARG:HG2	1:B:90:PRO:HD2	1.95	0.48
1:B:149:GLY:O	1:B:189:SCH:HB2	2.12	0.48
1:B:182:MET:HB3	1:B:204:VAL:HG22	1.95	0.48
1:A:258:VAL:O	1:A:262:THR:HG23	2.14	0.48
2:C:109:ARG:HB2	2:C:124:VAL:HB	1.96	0.47
2:C:82:ARG:HD3	2:C:108:TRP:CH2	2.50	0.46
1:A:127:ILE:HG13	1:A:127:ILE:O	2.13	0.46
1:A:73:ASP:O	1:A:77:SER:HB2	2.15	0.46
1:A:219:CYS:HB3	1:A:232:TRP:CE2	2.51	0.46
1:A:249:GLU:HG3	1:A:253:LYS:HD3	1.97	0.46
1:B:55:ILE:O	1:B:59:ASN:ND2	2.44	0.46
1:A:181:LYS:NZ	4:A:301:NAG:H61	2.32	0.45
1:A:151:THR:HG23	1:A:190:GLU:CD	2.36	0.45
1:B:249:GLU:HG3	1:B:253:LYS:HD3	1.98	0.45
1:A:29:HIS:NE2	1:A:140:HIS:HD2	2.15	0.45
1:B:219:CYS:HB3	1:B:232:TRP:CE2	2.51	0.45
1:B:258:VAL:O	1:B:262:THR:HG23	2.17	0.45
2:D:124:VAL:HG22	2:D:134:TRP:CZ3	2.52	0.45
2:D:71:ARG:HD2	2:D:77:PRO:HA	1.99	0.44
1:B:102:ASP:OD2	1:B:131:LYS:HE2	2.18	0.44
2:C:97:ILE:H	2:C:97:ILE:HD12	1.83	0.44
2:D:122:HIS:HB3	2:D:134:TRP:CE3	2.53	0.44
1:A:56:ILE:HG23	1:A:61:ILE:HD12	2.00	0.43
2:C:15:GLU:HG3	2:C:158:TYR:CD1	2.53	0.43
1:A:67:VAL:HG11	1:A:116:VAL:HG11	1.99	0.43
1:B:183:VAL:HA	1:B:205:TYR:O	2.18	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:31:VAL:HA	1:B:142:PHE:O	2.19	0.43
2:D:30:SER:HB2	2:D:42:GLU:HG2	1.99	0.43
1:B:114:LEU:HD13	1:B:114:LEU:HA	1.74	0.42
2:C:138:ARG:HB3	2:C:147:TYR:CE1	2.54	0.42
1:B:34:VAL:HG22	1:B:69:MET:HB2	2.00	0.42
1:B:69:MET:HG2	1:B:103:TYR:HB2	2.02	0.42
1:B:160:ASP:OD2	2:D:96:ARG:NH2	2.52	0.42
1:A:32:VAL:HG13	1:A:67:VAL:HG13	2.02	0.41
1:B:127:ILE:O	1:B:127:ILE:HG13	2.20	0.41
1:A:41:TYR:HA	1:A:227:THR:HG23	2.03	0.41
1:A:56:ILE:CG2	1:A:61:ILE:HB	2.51	0.41
1:B:188:ALA:HB3	1:B:191:SER:HB3	2.02	0.41
1:B:271:GLY:O	1:B:273:LYS:N	2.54	0.41
1:A:40:TRP:CH2	1:A:80:ASN:HA	2.56	0.41
1:A:68:VAL:HG21	1:A:100:PRO:HD2	2.02	0.41
2:C:76:ILE:HA	2:C:77:PRO:HD3	1.90	0.41
1:A:152:GLY:HA2	1:A:163:VAL:HG23	2.02	0.41
1:A:30:TRP:CZ3	1:A:65:GLN:HA	2.56	0.40
2:C:20:ASN:O	2:C:22:LYS:N	2.44	0.40
2:C:124:VAL:HG22	2:C:134:TRP:CZ3	2.56	0.40
2:C:129:ILE:O	6:C:301:HOH:O	2.22	0.40
2:C:63:ASP:HB2	6:C:314:HOH:O	2.21	0.40
1:B:70:MET:O	1:B:104:THR:HA	2.21	0.40
1:A:108:VAL:HG23	1:A:157:PRO:HG2	2.03	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	258/263 (98%)	248 (96%)	10 (4%)	0	<b>100</b> <b>100</b>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	258/263 (98%)	247 (96%)	11 (4%)	0	100	100
2	C	166/177 (94%)	159 (96%)	7 (4%)	0	100	100
2	D	156/177 (88%)	148 (95%)	8 (5%)	0	100	100
All	All	838/880 (95%)	802 (96%)	36 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 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	227/228 (100%)	223 (98%)	4 (2%)	59	75
1	B	227/228 (100%)	222 (98%)	5 (2%)	52	69
2	C	142/151 (94%)	138 (97%)	4 (3%)	43	60
2	D	134/151 (89%)	128 (96%)	6 (4%)	27	39
All	All	730/758 (96%)	711 (97%)	19 (3%)	46	63

All (19) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	89	ARG
1	A	114	LEU
1	A	227	THR
1	A	235	VAL
2	C	51	ARG
2	C	96	ARG
2	C	137	LEU
2	C	154	ILE
1	B	32	VAL
1	B	91	ASN
1	B	122	GLU
1	B	159	GLU
1	B	248	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	D	3	PHE
2	D	37	VAL
2	D	59	GLN
2	D	105	THR
2	D	132	THR
2	D	169	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	140	HIS

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

4 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
1	SNN	B	147	1	7,8,8	1.34	1 (14%)	7,11,11	2.65	4 (57%)
1	SNN	A	147	1	7,8,8	1.23	2 (28%)	7,11,11	3.24	3 (42%)
1	SCH	B	189	1	6,7,8	1.19	0	3,7,9	2.19	1 (33%)
1	SCH	A	189	1	6,7,8	1.24	0	3,7,9	2.44	1 (33%)

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. '2' means no outliers of that kind were identified.



Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	SNN	B	147	1	-	-	0/1/1/1
1	SNN	A	147	1	-	-	0/1/1/1
1	SCH	B	189	1	-	0/2/6/8	-
1	SCH	A	189	1	-	1/2/6/8	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	147	SNN	C-N1	-2.69	1.34	1.37
1	A	147	SNN	C-N1	-2.05	1.34	1.37
1	A	147	SNN	C5-N1	-2.05	1.34	1.37

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	147	SNN	CA-C-N1	5.40	111.47	107.30
1	A	147	SNN	O-C-CA	-5.13	122.47	126.18
1	B	147	SNN	CA-C-N1	4.80	111.01	107.30
1	A	189	SCH	CB-SG-SD	3.97	114.10	103.82
1	A	147	SNN	O5-C5-C4	-3.59	121.66	126.39
1	B	189	SCH	CB-SG-SD	3.50	112.88	103.82
1	B	147	SNN	O-C-CA	-3.37	123.75	126.18
1	B	147	SNN	O5-C5-C4	-3.00	122.44	126.39
1	B	147	SNN	O5-C5-N1	2.06	127.77	125.00

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	189	SCH	CA-CB-SG-SD

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	B	189	SCH	1	0

## 5.5 Carbohydrates [i](#)

2 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	1,3	14,14,15	1.02	2 (14%)	17,19,21	1.27	2 (11%)
3	NAG	E	2	3	14,14,15	0.55	0	17,19,21	1.04	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
3	NAG	E	1	1,3	-	3/6/23/26	0/1/1/1
3	NAG	E	2	3	-	0/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	1	NAG	C1-C2	2.87	1.56	1.52
3	E	1	NAG	O5-C1	2.22	1.47	1.43

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	1	NAG	O4-C4-C5	3.82	118.79	109.30
3	E	2	NAG	C1-O5-C5	3.81	117.36	112.19
3	E	1	NAG	C1-O5-C5	2.47	115.54	112.19

There are no chirality outliers.

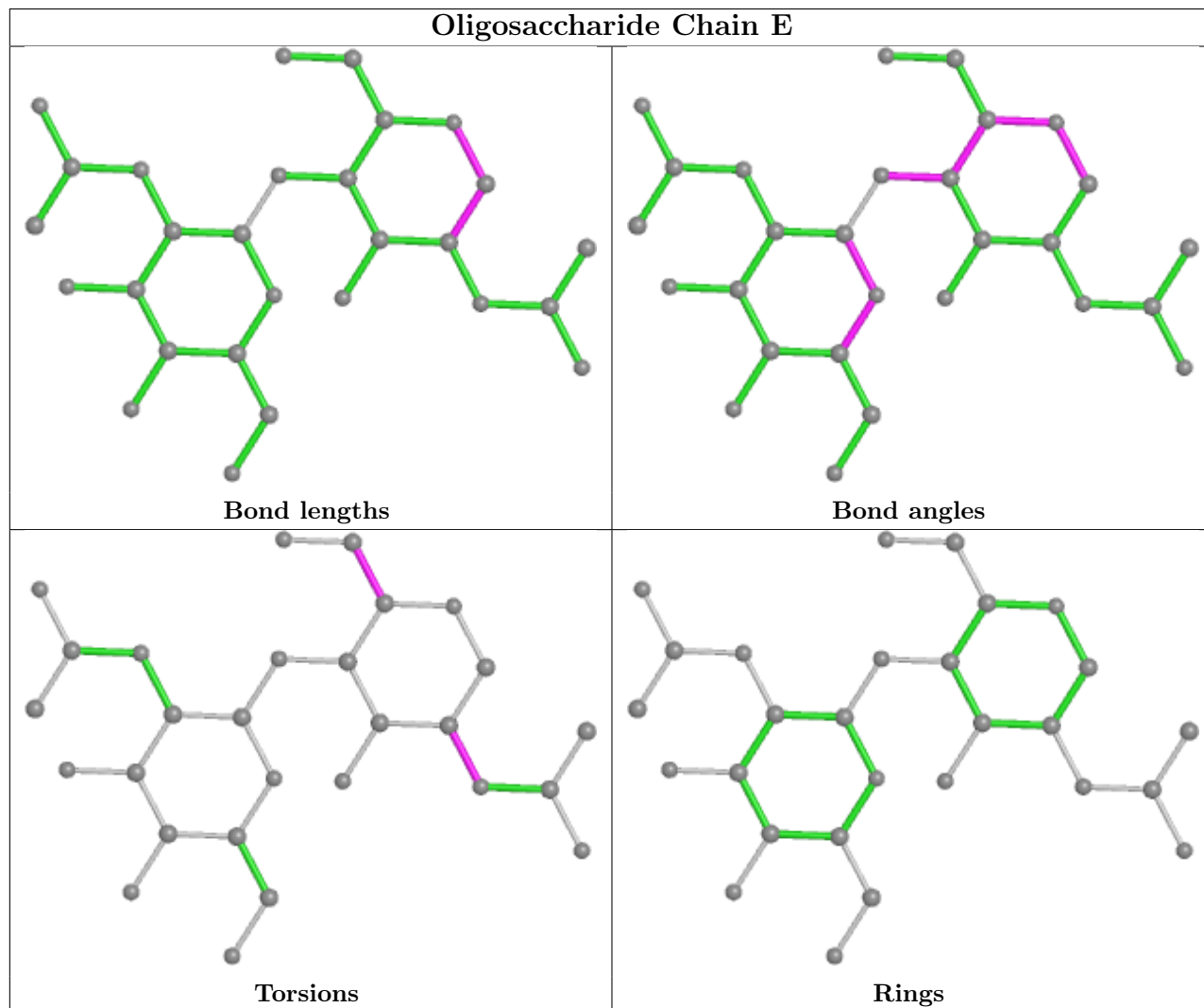
All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	E	1	NAG	C4-C5-C6-O6
3	E	1	NAG	C3-C2-N2-C7
3	E	1	NAG	O5-C5-C6-O6

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](#)

9 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	NAG	A	302	1	14,14,15	0.68	0	17,19,21	0.61	0
4	NAG	B	301	1	14,14,15	0.36	0	17,19,21	0.77	1 (5%)
5	SO4	D	201	-	4,4,4	0.20	0	6,6,6	0.28	0
5	SO4	D	202	-	4,4,4	0.19	0	6,6,6	0.53	0
4	NAG	B	302	1	14,14,15	0.59	0	17,19,21	0.44	0
5	SO4	C	201	-	4,4,4	0.15	0	6,6,6	0.35	0
5	SO4	B	303	-	4,4,4	0.11	0	6,6,6	0.37	0
5	SO4	A	303	-	4,4,4	0.23	0	6,6,6	0.31	0
4	NAG	A	301	1	14,14,15	1.78	1 (7%)	17,19,21	0.83	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
4	NAG	A	302	1	-	0/6/23/26	0/1/1/1
4	NAG	B	302	1	-	0/6/23/26	0/1/1/1
4	NAG	A	301	1	-	1/6/23/26	0/1/1/1
4	NAG	B	301	1	-	2/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	301	NAG	O5-C1	6.34	1.53	1.43

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	301	NAG	C1-O5-C5	2.58	115.69	112.19
4	A	301	NAG	C1-O5-C5	2.32	115.33	112.19

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	301	NAG	O5-C5-C6-O6
4	A	301	NAG	O5-C5-C6-O6
4	B	301	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
4	A	301	NAG	1	0

## 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	261/263 (99%)	1.19	59 (22%) 0 1	66, 100, 139, 168	3 (1%)
1	B	260/263 (98%)	1.17	60 (23%) 0 1	66, 105, 137, 162	11 (4%)
2	C	168/177 (94%)	0.50	10 (5%) 21 28	52, 67, 126, 173	2 (1%)
2	D	160/177 (90%)	0.49	10 (6%) 20 25	55, 71, 118, 178	1 (0%)
All	All	849/880 (96%)	0.92	139 (16%) 1 2	52, 90, 137, 178	17 (2%)

All (139) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	248	LYS	4.9
1	A	53	TYR	4.9
1	B	123	ALA	4.7
1	A	127	ILE	4.6
1	B	261	HIS	4.5
1	B	255	TYR	4.4
1	B	195	MET	4.3
1	B	127	ILE	4.2
1	A	54	GLN	4.2
1	A	98	GLY	4.2
1	B	270	TYR	4.1
1	A	95	VAL	4.1
1	B	115	ALA	4.1
1	A	114	LEU	4.1
1	A	55	ILE	4.0
1	A	202	ILE	4.0
1	B	126	GLY	3.9
1	B	166	LEU	3.9
1	B	27	GLY	3.9
1	A	99	VAL	3.8
1	B	197	HIS	3.8

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	B	256	HIS	3.8
1	A	124	VAL	3.7
1	B	122	GLU	3.7
1	A	56	ILE	3.7
1	A	270	TYR	3.6
1	A	198	LEU	3.5
2	D	115	HIS	3.5
1	A	50	CYS	3.5
1	A	271	GLY	3.5
1	B	95	VAL	3.5
1	A	52	ALA	3.5
1	A	122	GLU	3.5
1	A	203	ASN	3.4
1	B	257	LEU	3.4
1	B	258	VAL	3.4
1	B	121	ALA	3.4
1	A	201	ASN	3.4
1	B	277	THR	3.3
1	B	182	MET	3.3
1	B	288	ARG	3.3
1	A	57	HIS	3.3
1	B	97	GLN	3.3
1	B	276	SER	3.2
1	B	55	ILE	3.2
2	D	169	GLU	3.2
1	B	50	CYS	3.2
2	C	142	GLY	3.2
2	D	2	GLY	3.2
1	B	130	GLY	3.1
1	A	86	VAL	3.1
1	B	173	MET	3.1
2	D	3	PHE	3.1
1	B	275	ILE	3.0
1	A	177	LYS	3.0
1	A	111	GLN	3.0
1	A	61	ILE	3.0
1	A	123	ALA	3.0
1	A	58	ARG	3.0
1	A	31	VAL	3.0
1	B	86	VAL	2.9
2	C	20	ASN	2.9
1	A	125	LYS	2.9

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	B	177	LYS	2.9
1	A	47	ALA	2.8
2	C	23	ILE	2.8
1	B	267	VAL	2.8
1	A	199	PRO	2.8
1	A	136	GLY	2.8
2	C	114	ASP	2.8
2	C	169	GLU	2.8
2	D	16	GLY	2.8
1	B	93	THR	2.8
1	A	96	TYR	2.7
1	A	261	HIS	2.7
1	B	99	VAL	2.6
1	B	116	VAL	2.6
2	C	115	HIS	2.6
1	A	32	VAL	2.6
1	A	174	TYR	2.6
1	A	260	SER	2.5
1	B	172	TYR	2.5
1	B	169	THR	2.5
2	D	63	ASP	2.5
1	A	103	TYR	2.5
1	B	56	ILE	2.5
1	A	166	LEU	2.5
2	D	116	GLY	2.5
1	B	271	GLY	2.5
1	A	173	MET	2.5
1	B	194	MET	2.5
2	C	3	PHE	2.5
1	A	120	ASP	2.5
1	B	124	VAL	2.5
1	B	47	ALA	2.5
1	A	93	THR	2.5
1	B	263	ASN	2.5
1	A	288	ARG	2.4
1	A	60	GLY	2.4
1	A	197	HIS	2.4
1	B	164	LYS	2.4
1	B	179	TYR	2.4
2	D	114	ASP	2.4
1	B	108	VAL	2.4
2	C	24	PRO	2.4

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	51	HIS	2.3
1	B	259	LYS	2.3
1	A	97	GLN	2.3
1	A	100	PRO	2.3
1	B	199	PRO	2.3
1	A	94	ASP	2.3
1	B	57	HIS	2.3
1	A	128	GLY	2.3
1	A	129	SER	2.2
1	A	134	LYS	2.2
1	B	94	ASP	2.2
2	C	19	PRO	2.2
2	D	109	ARG	2.2
1	B	84	GLY	2.2
1	A	164	LYS	2.2
1	B	135	SER	2.2
1	B	141	VAL	2.2
1	B	85	ILE	2.2
1	A	286	MET	2.2
1	B	134	LYS	2.1
1	A	138	GLN	2.1
1	B	138	GLN	2.1
1	A	118	ARG	2.1
2	C	116	GLY	2.1
1	A	59	ASN	2.1
1	B	184	PHE	2.1
1	B	54	GLN	2.1
2	D	62	ASP	2.1
1	B	196	ASN	2.1
1	B	87	ILE	2.0
1	B	69	MET	2.0
1	A	84	GLY	2.0
1	A	193	SER	2.0
1	B	165	ASP	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.

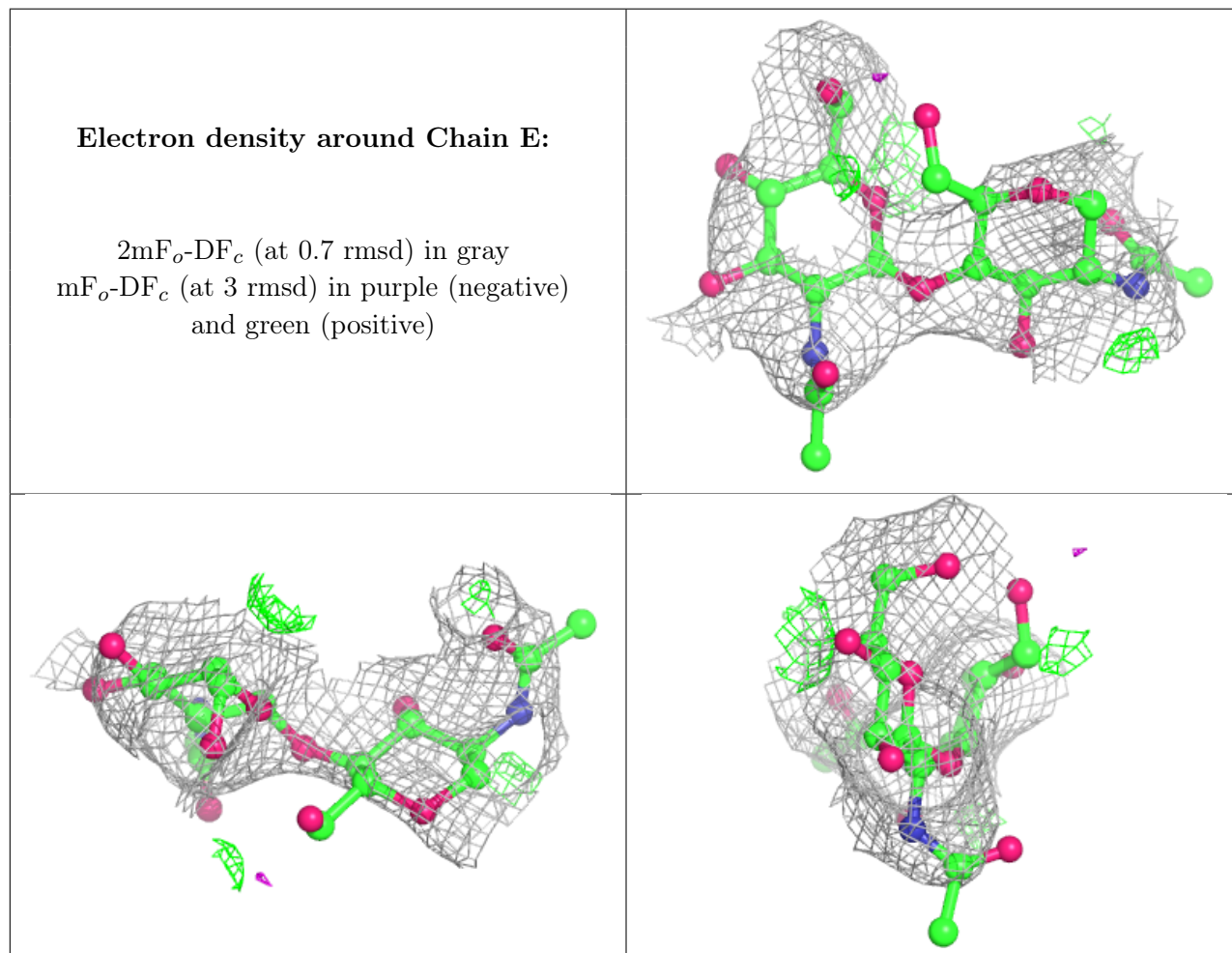
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
1	SCH	A	189	8/9	0.88	0.21	78,86,110,134	0
1	SCH	B	189	8/9	0.88	0.13	91,94,109,137	0
1	SNN	B	147	8/8	0.90	0.14	79,87,93,93	0
1	SNN	A	147	8/8	0.95	0.13	69,75,81,83	0

### 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( $\text{\AA}^2$ )	Q<0.9
3	NAG	E	1	14/15	0.77	0.40	141,152,159,166	0
3	NAG	E	2	14/15	0.80	0.30	140,161,174,183	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
4	NAG	A	301	14/15	0.23	0.33	121,137,150,151	0
4	NAG	B	302	14/15	0.35	0.40	146,168,179,184	0
4	NAG	B	301	14/15	0.78	0.20	120,152,167,167	0
4	NAG	A	302	14/15	0.80	0.30	150,166,176,183	0
5	SO4	D	202	5/5	0.88	0.22	107,116,129,130	0
5	SO4	D	201	5/5	0.95	0.14	84,90,92,111	0
5	SO4	C	201	5/5	0.95	0.14	77,85,92,94	0
5	SO4	A	303	5/5	0.97	0.15	74,95,98,107	0
5	SO4	B	303	5/5	0.99	0.18	75,79,88,94	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.