



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 6, 2020 – 01:45 PM BST

PDB ID : 2ID4  
Title : The 1.9 Å structure of Kex2 in complex with an Ac-R-E-R-K-chloromethyl ketone inhibitor.  
Authors : Wheatley, J.L.; Holyoak, T.  
Deposited on : 2006-09-14  
Resolution : 1.90 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.13.1  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.13.1

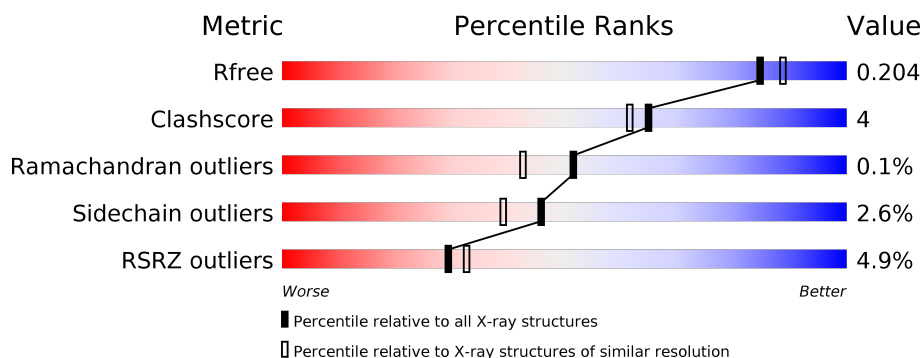
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 Å.

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



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	503	<div> <div>6%</div> <div> <div></div> <div>87%</div> <div>8%</div> <div>5%</div> </div> </div>
1	B	503	<div> <div>3%</div> <div> <div></div> <div>88%</div> <div>6%</div> <div>5%</div> </div> </div>
2	C	6	<div> <div>33%</div> <div>50%</div> <div>17%</div> </div>
2	D	6	<div> <div>67%</div> <div>33%</div> </div>
3	E	2	<div> <div>100%</div> </div>
3	F	2	<div> <div>100%</div> </div>

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
3	NDG	E	2	X	-	-	X
3	NDG	F	2	X	-	-	X
4	NAG	A	903	X	-	-	X
4	NAG	B	903	X	-	-	-

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 8487 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 Kexin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	480	Total	C	N	O	S	0	20	0
			3814	2395	646	762	11			
1	B	479	Total	C	N	O	S	0	12	0
			3771	2361	646	753	11			

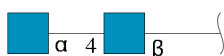
There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	614	ARG	-	expression tag	UNP P13134
A	615	GLN	-	expression tag	UNP P13134
A	616	ARG	-	expression tag	UNP P13134
B	614	ARG	-	expression tag	UNP P13134
B	615	GLN	-	expression tag	UNP P13134
B	616	ARG	-	expression tag	UNP P13134

- Molecule 2 is a protein called Ac-RERK-CMK inhibitor.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	6	Total	C	N	O	0	1	1
			50	30	11	9			
2	D	6	Total	C	N	O	0	0	1
			44	26	11	7			

- Molecule 3 is an oligosaccharide called 2-acetamido-2-deoxy-alpha-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	E	2	Total	C	N	O	0	0	0
			28	16	2	10			
3	F	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	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 CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	2	Total	Ca	0	0
			2	2		
5	A	2	Total	Ca	0	0
			2	2		

- Molecule 6 is SODIUM ION (three-letter code: NA) (formula: Na).

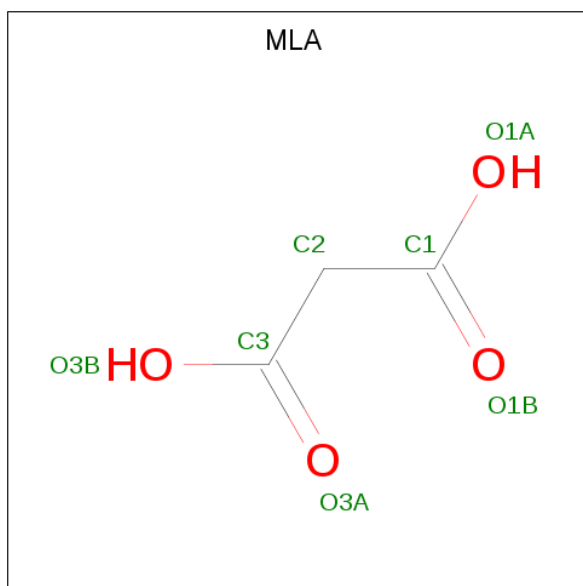
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	1	Total	Na	0	0
			1	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	Na	0	0
			1	1		

- Molecule 7 is MALONIC ACID (three-letter code: MLA) (formula:  $C_3H_4O_4$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	C	O	0	0
			7	3	4		

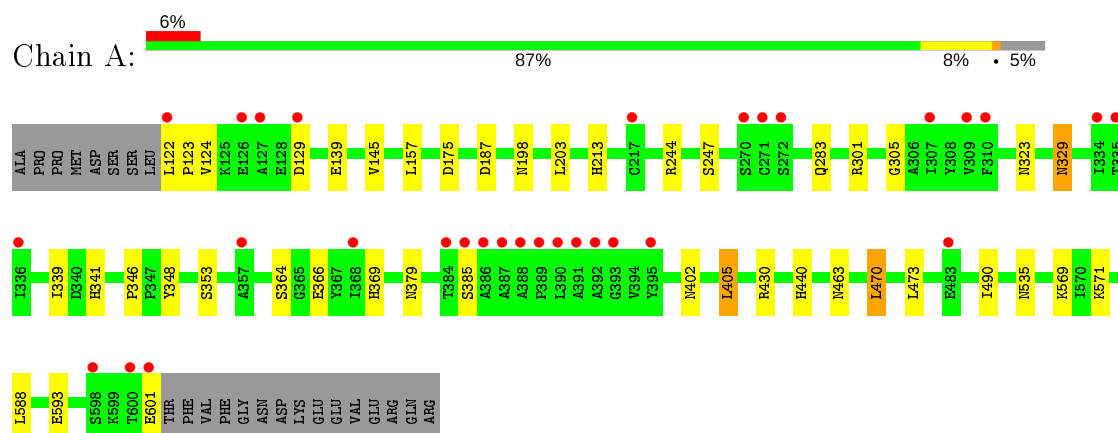
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	347	Total	O	0	0
			347	347		
8	B	354	Total	O	0	0
			354	354		
8	C	5	Total	O	0	0
			5	5		
8	D	5	Total	O	0	0
			5	5		

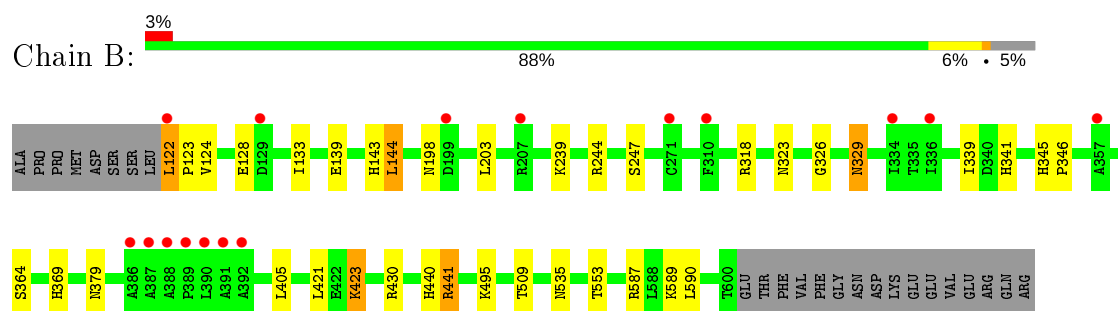
### 3 Residue-property plots

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: Kexin



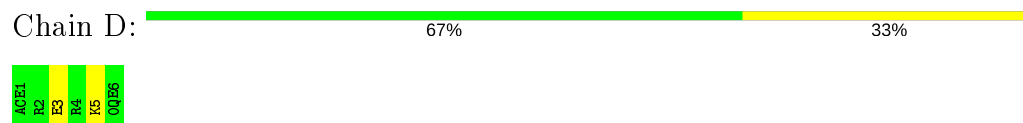
#### • Molecule 1: Kexin



#### • Molecule 2: Ac-RERK-CMK inhibitor



#### • Molecule 2: Ac-RERK-CMK inhibitor



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

Chain E:  100%

MDG1  
MDG2

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

Chain F:  100%

MDG1  
MDG2



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	112.85Å 112.85Å 370.17Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	40.79 – 1.90 40.78 – 1.90	Depositor EDS
% Data completeness (in resolution range)	96.6 (40.79-1.90) 96.6 (40.78-1.90)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.95 (at 1.89Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.177 , 0.206 0.176 , 0.204	Depositor DCC
$R_{free}$ test set	5345 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	35.6	Xtriage
Anisotropy	0.087	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 42.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	8487	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.29% 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: CSO, MLA, NAG, ACE, NA, CA, NDG, OQE, LYK

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	0.41	0/3946	0.55	0/5356
1	B	0.41	0/3876	0.57	0/5259
2	C	0.41	0/40	0.78	0/51
2	D	0.25	0/31	0.76	0/39
All	All	0.41	0/7893	0.56	0/10705

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	3814	0	3688	28	0
1	B	3771	0	3621	27	0
2	C	50	0	52	4	0
2	D	44	0	46	0	0
3	E	28	0	24	0	0
3	F	28	0	24	1	0
4	A	14	0	13	0	0
4	B	14	0	13	4	0
5	A	2	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	2	0	0	0	0
6	A	1	0	0	0	0
6	B	1	0	0	0	0
7	A	7	0	2	1	0
8	A	347	0	0	2	0
8	B	354	0	0	3	0
8	C	5	0	0	0	0
8	D	5	0	0	0	0
All	All	8487	0	7483	61	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 4.

All (61) 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:B:903:NAG:H3	4:B:903:NAG:H82	1.07	1.03
4:B:903:NAG:H3	4:B:903:NAG:C8	1.86	1.03
4:B:903:NAG:H82	4:B:903:NAG:C3	1.99	0.91
1:B:441[B]:ARG:HH11	1:B:441[B]:ARG:HG3	1.40	0.87
1:B:318:ARG:NH2	8:B:1163:HOH:O	2.17	0.76
1:B:441[B]:ARG:NH1	8:B:998:HOH:O	2.18	0.76
1:A:470[A]:LEU:HD22	1:A:490[A]:ILE:HG21	1.68	0.73
1:B:587[B]:ARG:HE	1:B:589:LYS:HE3	1.54	0.71
1:B:440:HIS:HE1	1:B:535:ASN:H	1.38	0.71
4:B:903:NAG:C8	4:B:903:NAG:C3	2.63	0.70
1:A:385:SER:CB	2:C:5:LYK:C	2.70	0.69
1:B:341:HIS:HE1	1:B:364:SER:O	1.74	0.69
1:A:440:HIS:HE1	1:A:535:ASN:H	1.41	0.68
1:A:369:HIS:HD2	1:A:379:ASN:OD1	1.75	0.67
1:A:440:HIS:CE1	1:A:535:ASN:H	2.13	0.67
1:B:440:HIS:CE1	1:B:535:ASN:H	2.13	0.66
3:F:1:NAG:H62	3:F:2:NDG:H4	1.84	0.60
1:A:470[A]:LEU:HD13	1:A:588:LEU:HD23	1.84	0.60
8:A:1006:HOH:O	1:B:430[B]:ARG:HD2	2.01	0.59
1:A:198:ASN:HD22	1:A:244:ARG:HG2	1.67	0.59
1:A:283:GLN:NE2	2:C:1:ACE:H3	2.18	0.58
1:B:369:HIS:HD2	1:B:379:ASN:OD1	1.86	0.58
1:A:329:ASN:HD22	1:A:329:ASN:C	2.08	0.56
1:B:198:ASN:HD22	1:B:244:ARG:HG2	1.70	0.56
1:B:143:HIS:CD2	1:B:144:LEU:HD13	2.41	0.56

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:329:ASN:C	1:B:329:ASN:HD22	2.11	0.55
1:A:341:HIS:HE1	1:A:364:SER:O	1.90	0.54
1:B:345:HIS:CE1	1:B:441[A]:ARG:HD2	2.43	0.54
1:A:122:LEU:N	1:A:123:PRO:HD2	2.24	0.53
1:A:470[A]:LEU:HD22	1:A:490[A]:ILE:CG2	2.39	0.53
2:C:1:ACE:H1	2:C:2:ARG:HG3	1.91	0.53
1:A:366:GLU:OE1	1:A:369:HIS:HE1	1.93	0.50
1:A:145[A]:VAL:HG11	1:A:157:LEU:HD11	1.94	0.49
1:B:587[B]:ARG:NE	1:B:589:LYS:HE3	2.24	0.48
1:B:128:GLU:HA	1:B:133:ILE:HB	1.96	0.48
1:B:421:LEU:HA	1:B:423[A]:LYS:HE2	1.97	0.47
1:A:341:HIS:HD2	8:A:953:HOH:O	1.96	0.47
1:B:441[B]:ARG:HH11	1:B:441[B]:ARG:CG	2.20	0.47
1:B:122:LEU:N	1:B:123:PRO:CD	2.78	0.47
1:A:463:ASN:HB2	1:A:593:GLU:HG2	1.98	0.46
1:B:326:GLY:HA2	1:B:329:ASN:ND2	2.31	0.46
1:A:198:ASN:HD21	1:A:247:SER:H	1.64	0.46
1:A:198:ASN:ND2	1:A:247:SER:H	2.15	0.44
1:B:339:ILE:HA	1:B:346:PRO:HD3	1.98	0.44
1:B:198:ASN:ND2	1:B:247:SER:H	2.16	0.44
1:A:348:TYR:CD2	7:A:908:MLA:HC21	2.53	0.43
1:A:402:ASN:HB3	1:A:405:LEU:HD22	2.00	0.43
1:A:301[A]:ARG:HD3	1:A:305:GLY:O	2.18	0.43
1:B:441[B]:ARG:HG3	1:B:441[B]:ARG:NH1	2.17	0.43
1:B:239:LYS:HA	1:B:239:LYS:HD3	1.88	0.42
1:A:569[A]:LYS:HE3	1:A:571:LYS:HD2	2.02	0.41
1:A:329:ASN:ND2	1:A:329:ASN:C	2.74	0.41
1:A:213:HIS:CD2	2:C:6:OQE:C1	2.97	0.41
1:A:124:VAL:HG11	1:A:139[B]:GLU:HG2	2.02	0.41
1:A:301[B]:ARG:NH1	1:A:405:LEU:O	2.52	0.41
1:B:124:VAL:HG11	1:B:139:GLU:HG2	2.02	0.41
1:A:339:ILE:HA	1:A:346:PRO:HD3	2.03	0.41
1:B:509:THR:HB	1:B:553:THR:CG2	2.51	0.41
1:A:353:SER:HB3	1:A:440:HIS:CE1	2.57	0.40
1:B:198:ASN:HD21	1:B:247:SER:H	1.69	0.40
1:B:341:HIS:HD2	8:B:1029:HOH:O	2.03	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	497/503 (99%)	485 (98%)	11 (2%)	1 (0%)	47	38
1	B	488/503 (97%)	475 (97%)	13 (3%)	0	100	100
2	C	4/6 (67%)	4 (100%)	0	0	100	100
2	D	3/6 (50%)	3 (100%)	0	0	100	100
All	All	992/1018 (97%)	967 (98%)	24 (2%)	1 (0%)	51	43

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	175	ASP

### 5.3.2 Protein sidechains

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	420/421 (100%)	408 (97%)	12 (3%)	42	35
1	B	411/421 (98%)	399 (97%)	12 (3%)	42	35
2	C	4/3 (133%)	4 (100%)	0	100	100
2	D	3/3 (100%)	2 (67%)	1 (33%)	0	0
All	All	838/848 (99%)	813 (97%)	25 (3%)	46	33

All (25) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	129	ASP
1	A	187	ASP
1	A	203	LEU
1	A	323	ASN
1	A	329	ASN
1	A	405	LEU
1	A	430	ARG
1	A	470[A]	LEU
1	A	470[B]	LEU
1	A	473[A]	LEU
1	A	473[B]	LEU
1	A	601	GLU
1	B	122	LEU
1	B	144	LEU
1	B	203	LEU
1	B	323	ASN
1	B	329	ASN
1	B	405	LEU
1	B	423[A]	LYS
1	B	423[B]	LYS
1	B	441[A]	ARG
1	B	441[B]	ARG
1	B	495	LYS
1	B	590	LEU
2	D	3	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	198	ASN
1	A	329	ASN
1	A	341	HIS
1	A	369	HIS
1	A	440	HIS
1	A	477	GLN
1	B	198	ASN
1	B	237	ASN
1	B	329	ASN
1	B	341	HIS
1	B	369	HIS
1	B	440	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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	CSO	B	190	1	3,6,7	0.77	0	0,6,8	0.00	-
1	CSO	A	190	1	3,6,7	0.70	0	0,6,8	0.00	-
2	LYK	C	5	1,2	8,8,9	1.54	1 (12%)	7,8,10	1.26	1 (14%)
2	LYK	D	5	1,2	8,8,9	1.54	1 (12%)	7,8,10	1.04	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
1	CSO	B	190	1	-	0/1/5/7	-
1	CSO	A	190	1	-	0/1/5/7	-
2	LYK	C	5	1,2	-	1/7/7/9	-
2	LYK	D	5	1,2	-	1/7/7/9	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	5	LYK	O-C	-4.24	1.24	1.42
2	D	5	LYK	O-C	-4.12	1.25	1.42

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	5	LYK	O-C-CA	2.45	121.22	111.52

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	5	LYK	O-C-CA-N
2	D	5	LYK	O-C-CA-N

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	5	LYK	1	0

## 5.5 Carbohydrates

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	$\# Z  > 2$
3	NAG	E	1	1,3	14,14,15	0.55	0	17,19,21	1.52	4 (23%)
3	NDG	E	2	3	14,14,15	0.49	0	17,19,21	1.68	3 (17%)
3	NAG	F	1	1,3	14,14,15	0.53	0	17,19,21	1.45	2 (11%)
3	NDG	F	2	3	14,14,15	0.56	0	17,19,21	1.49	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	-	0/6/23/26	0/1/1/1
3	NDG	E	2	3	1/1/5/7	5/6/23/26	0/1/1/1
3	NAG	F	1	1,3	-	2/6/23/26	0/1/1/1
3	NDG	F	2	3	1/1/5/7	4/6/23/26	0/1/1/1



There are no bond length outliers.

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	F	2	NDG	C1-O5-C5	5.25	119.31	112.19
3	E	2	NDG	C1-O5-C5	5.25	119.30	112.19
3	E	1	NAG	C1-O5-C5	3.82	117.37	112.19
3	F	1	NAG	O5-C1-C2	-3.66	105.50	111.29
3	E	2	NDG	C3-C4-C5	2.46	114.64	110.24
3	F	1	NAG	C4-C3-C2	2.26	114.33	111.02
3	E	1	NAG	O4-C4-C5	2.24	114.86	109.30
3	E	1	NAG	O4-C4-C3	2.21	115.46	110.35
3	E	2	NDG	C4-C3-C2	2.17	114.19	111.02
3	E	1	NAG	C4-C3-C2	-2.00	108.08	111.02

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	E	2	NDG	C1
3	F	2	NDG	C1

All (11) torsion outliers are listed below:

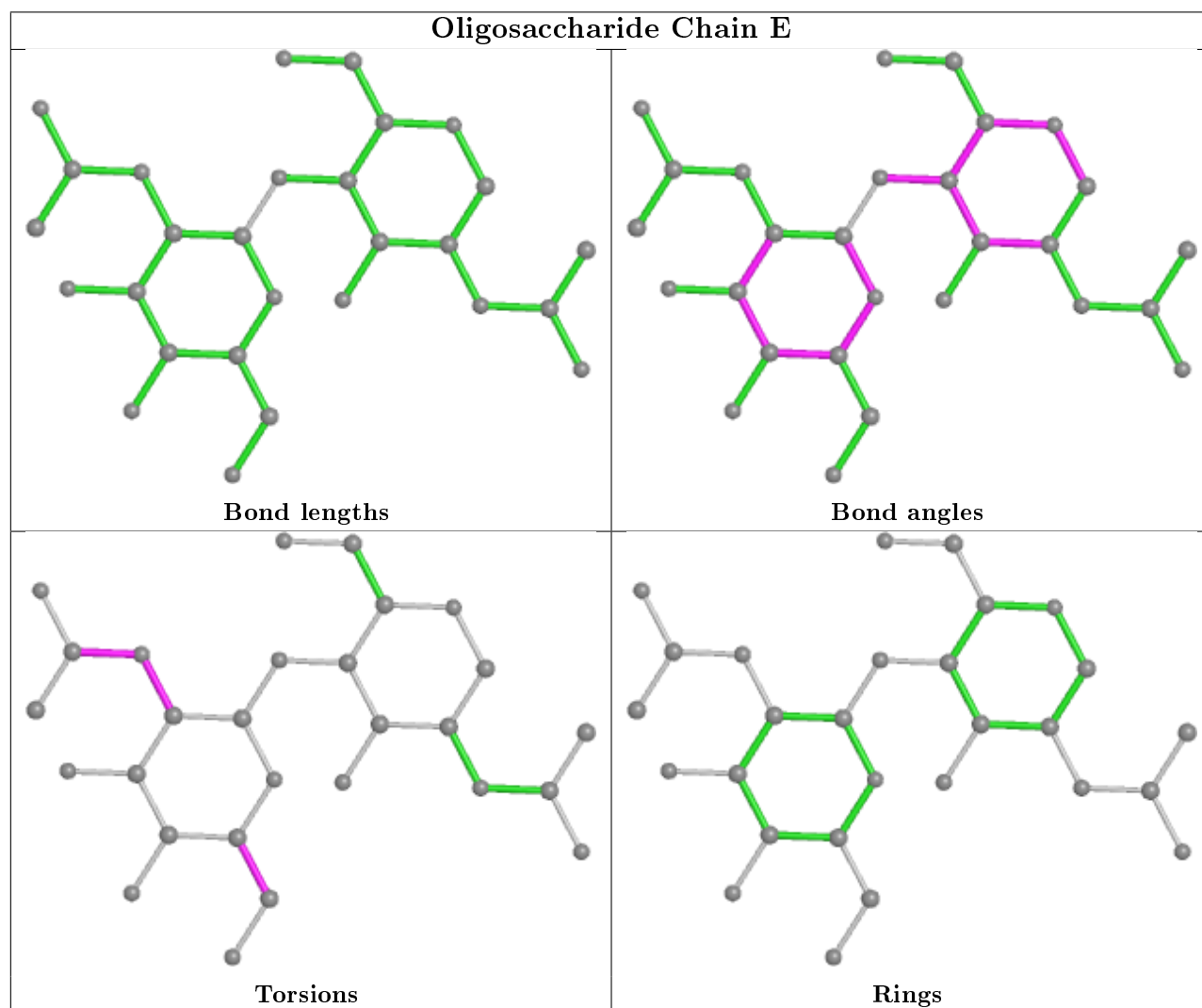
Mol	Chain	Res	Type	Atoms
3	E	2	NDG	C8-C7-N2-C2
3	E	2	NDG	O7-C7-N2-C2
3	F	2	NDG	C8-C7-N2-C2
3	F	2	NDG	O7-C7-N2-C2
3	F	1	NAG	C4-C5-C6-O6
3	F	1	NAG	O5-C5-C6-O6
3	F	2	NDG	O5-C5-C6-O6
3	E	2	NDG	C1-C2-N2-C7
3	E	2	NDG	O5-C5-C6-O6
3	E	2	NDG	C3-C2-N2-C7
3	F	2	NDG	C4-C5-C6-O6

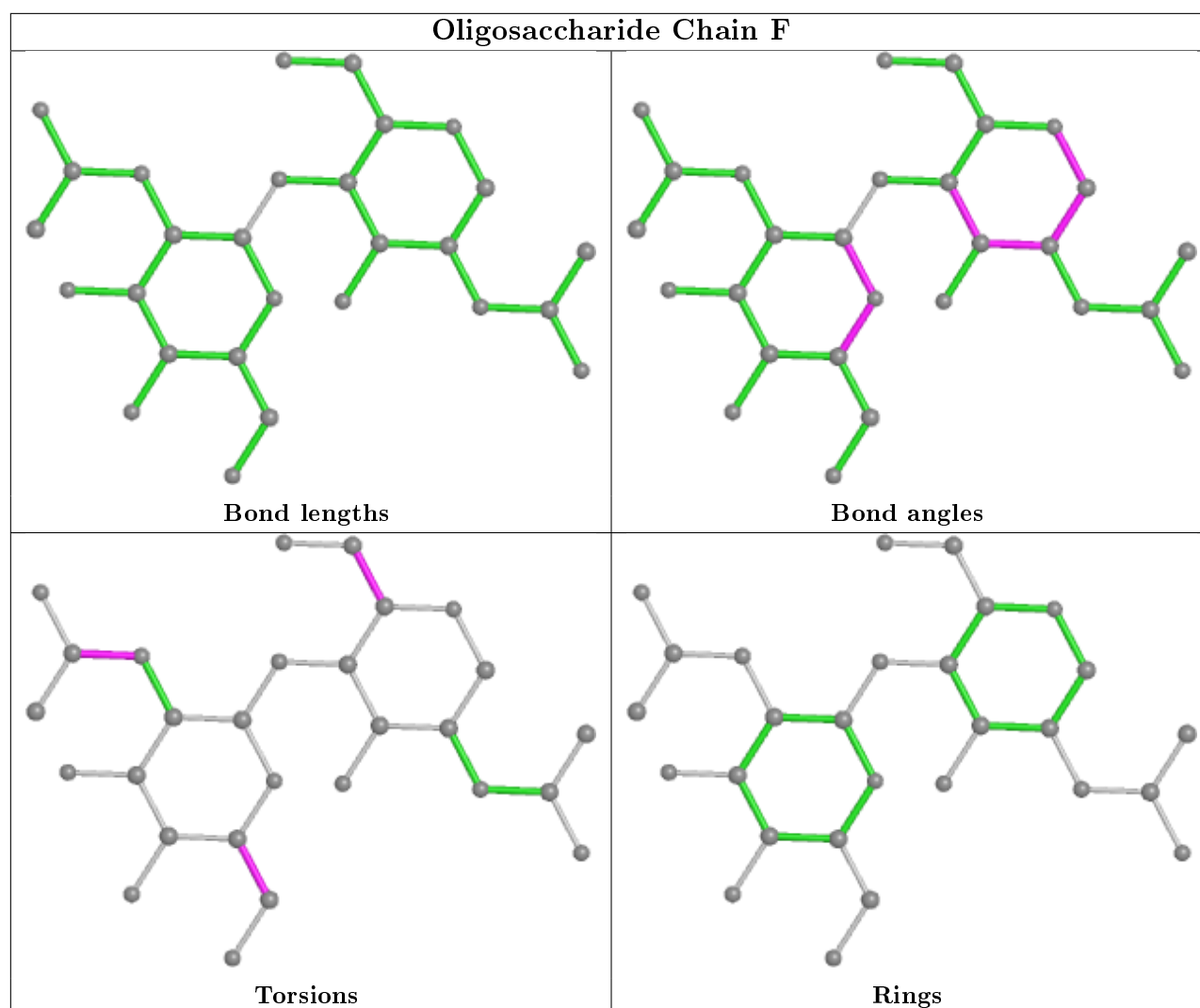
There are no ring outliers.

2 monomers are involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	F	2	NDG	1	0
3	F	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 ⓘ

Of 9 ligands modelled in this entry, 6 are monoatomic - leaving 3 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	$\# Z  > 2$
4	NAG	A	903	1	14,14,15	0.63	0	17,19,21	1.14	3 (17%)
7	MLA	A	908	-	0,6,6	0.00	-	0,7,7	0.00	-
4	NAG	B	903	1	14,14,15	1.08	1 (7%)	17,19,21	1.87	5 (29%)

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	903	1	1/1/5/7	3/6/23/26	0/1/1/1
7	MLA	A	908	-	-	0/0/4/4	-
4	NAG	B	903	1	1/1/5/7	4/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	903	NAG	O5-C1	-2.95	1.39	1.43

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	903	NAG	C2-N2-C7	-4.21	116.91	122.90
4	B	903	NAG	C3-C4-C5	-3.47	104.05	110.24
4	B	903	NAG	C1-C2-N2	-3.09	105.21	110.49
4	A	903	NAG	C4-C3-C2	2.74	115.04	111.02
4	B	903	NAG	C4-C3-C2	-2.34	107.58	111.02
4	B	903	NAG	O5-C5-C6	2.16	110.59	107.20
4	A	903	NAG	C2-N2-C7	2.10	125.90	122.90
4	A	903	NAG	O5-C5-C6	2.08	110.46	107.20

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
4	A	903	NAG	C1
4	B	903	NAG	C1

All (7) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	903	NAG	C3-C2-N2-C7
4	A	903	NAG	O7-C7-N2-C2
4	B	903	NAG	C8-C7-N2-C2
4	B	903	NAG	O7-C7-N2-C2
4	A	903	NAG	C8-C7-N2-C2
4	B	903	NAG	C1-C2-N2-C7
4	B	903	NAG	C3-C2-N2-C7

There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	A	908	MLA	1	0
4	B	903	NAG	4	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	479/503 (95%)	0.21	31 (6%) 18 21	25, 28, 32, 61	0
1	B	478/503 (95%)	0.10	16 (3%) 46 49	25, 28, 32, 38	0
2	C	3/6 (50%)	0.09	0 100 100	37, 37, 41, 48	0
2	D	3/6 (50%)	0.05	0 100 100	44, 44, 49, 53	0
All	All	963/1018 (94%)	0.16	47 (4%) 29 33	25, 28, 32, 61	0

All (47) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	601	GLU	6.3
1	A	390	LEU	5.4
1	B	122	LEU	4.7
1	A	336	ILE	4.6
1	A	391	ALA	4.2
1	A	309	VAL	4.0
1	A	600	THR	4.0
1	A	334	ILE	3.9
1	A	392	ALA	3.9
1	A	387	ALA	3.8
1	A	388	ALA	3.8
1	A	122	LEU	3.8
1	B	388	ALA	3.7
1	B	129	ASP	3.5
1	B	387	ALA	3.1
1	A	598	SER	3.1
1	A	271	CYS	2.9
1	A	368	ILE	2.9
1	A	384	THR	2.8
1	A	389	PRO	2.8
1	B	391	ALA	2.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	334	ILE	2.8
1	B	389	PRO	2.7
1	B	336	ILE	2.6
1	A	386	ALA	2.6
1	A	335	THR	2.6
1	B	392	ALA	2.5
1	A	270	SER	2.5
1	A	129	ASP	2.5
1	A	385	SER	2.5
1	A	217	CYS	2.5
1	B	207	ARG	2.4
1	A	395	TYR	2.4
1	A	272	SER	2.4
1	B	199	ASP	2.4
1	B	390	LEU	2.3
1	B	271	CYS	2.3
1	A	483	GLU	2.3
1	A	127	ALA	2.3
1	A	310	PHE	2.3
1	A	357	ALA	2.2
1	B	386	ALA	2.2
1	A	393	GLY	2.2
1	A	307	ILE	2.1
1	B	310	PHE	2.1
1	A	126	GLU	2.1
1	B	357	ALA	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.

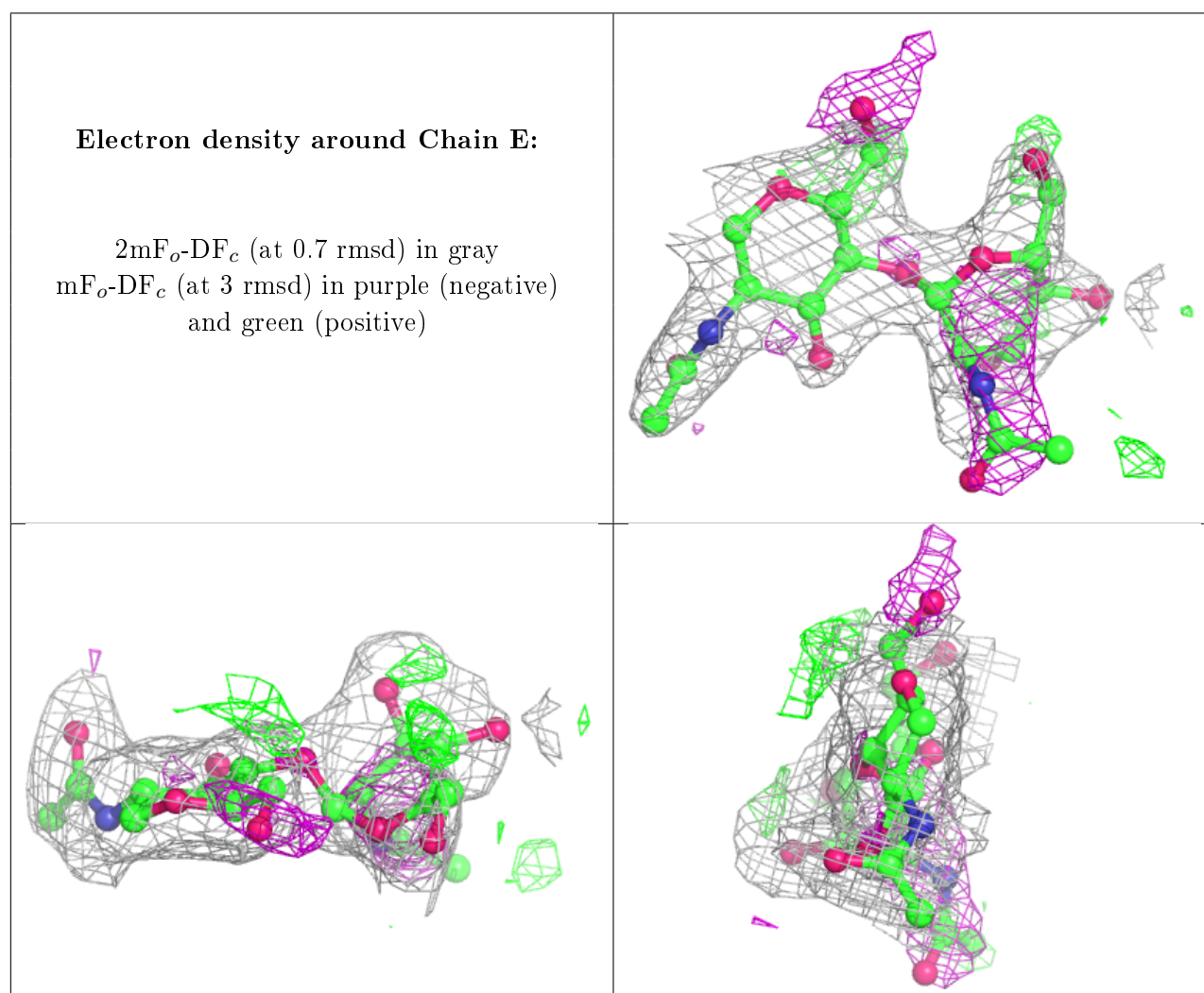
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
1	CSO	B	190	7/8	0.93	0.08	27,27,27,27	0
1	CSO	A	190	7/8	0.94	0.07	30,30,30,31	0
2	LYK	C	5	9/10	0.96	0.14	28,32,33,34	0
2	LYK	D	5	9/10	0.97	0.11	34,35,36,38	0

## 6.3 Carbohydrates ⓘ

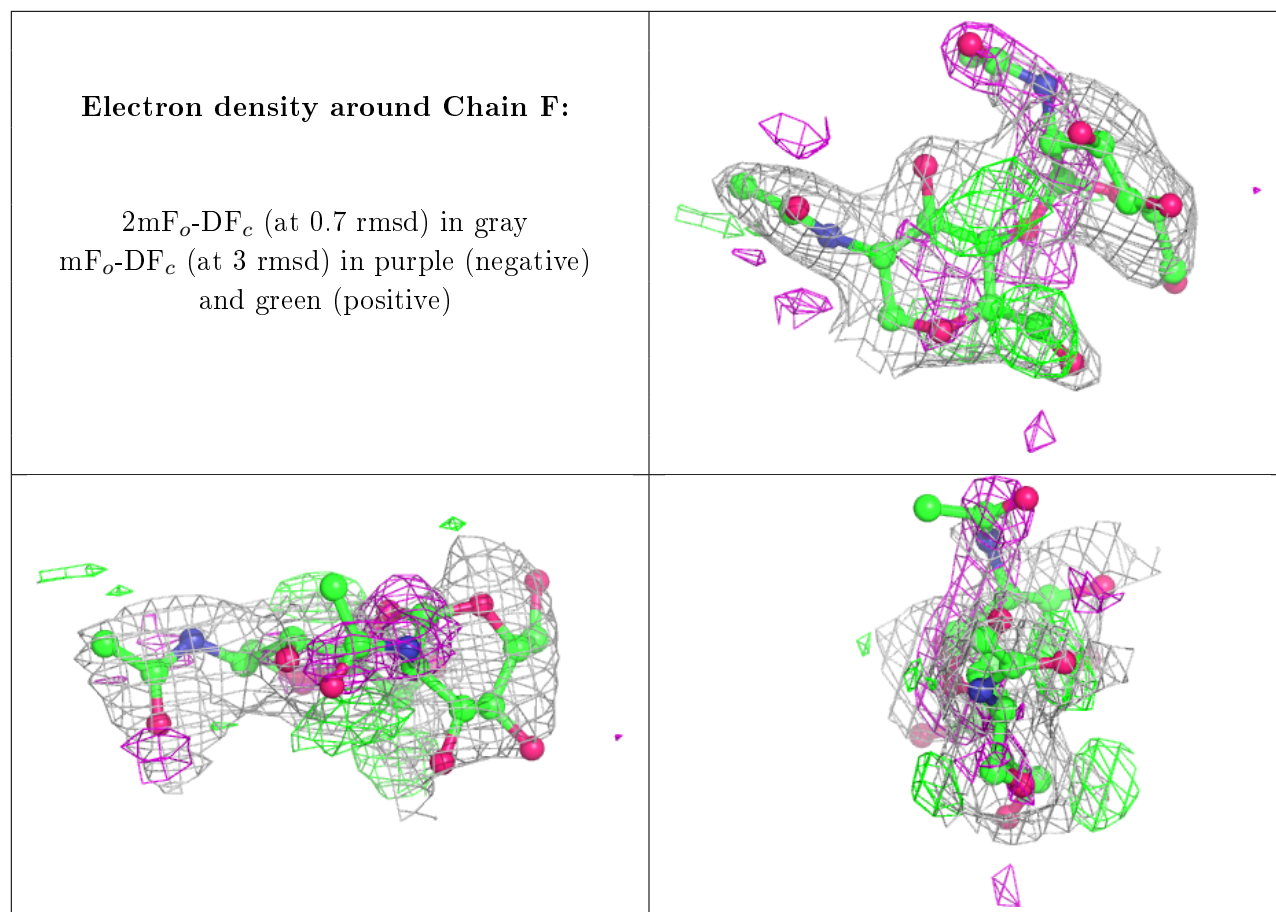
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
3	NDG	E	2	14/15	0.59	0.51	74,77,81,81	0
3	NDG	F	2	14/15	0.75	0.44	74,76,79,79	0
3	NAG	F	1	14/15	0.84	0.33	53,58,63,69	0
3	NAG	E	1	14/15	0.89	0.30	52,57,64,69	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 ⓘ

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	903	14/15	0.76	0.50	61,69,71,72	0
4	NAG	B	903	14/15	0.81	0.43	61,68,70,71	0
7	MLA	A	908	7/7	0.94	0.29	23,25,26,27	7
5	CA	B	906	1/1	0.97	0.06	45,45,45,45	0
5	CA	A	906	1/1	0.97	0.03	39,39,39,39	0
6	NA	B	907	1/1	0.99	0.07	29,29,29,29	0
6	NA	A	907	1/1	0.99	0.11	34,34,34,34	0
5	CA	A	901	1/1	0.99	0.02	28,28,28,28	0
5	CA	B	902	1/1	1.00	0.02	30,30,30,30	0

## 6.5 Other polymers

There are no such residues in this entry.