



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 24, 2023 – 12:33 AM EDT

PDB ID : 5TJ4  
Title : Gasdermin-B C-terminal domain containing the polymorphism residues Gly299:Pro306 fused to maltose binding protein  
Authors : Chao, L.K.; Herzberg, O.  
Deposited on : 2016-10-03  
Resolution : 3.50 Å(reported)

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

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The 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  
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

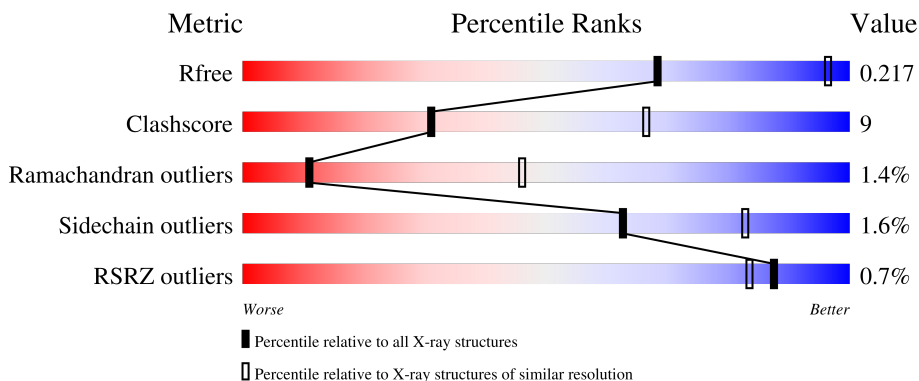
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.50 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	1659 (3.60-3.40)
Clashscore	141614	1036 (3.58-3.42)
Ramachandran outliers	138981	1005 (3.58-3.42)
Sidechain outliers	138945	1006 (3.58-3.42)
RSRZ outliers	127900	1559 (3.60-3.40)

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	561	 84% 13% ..
1	B	561	 84% 13% ..
1	C	561	 80% 16% ..
1	D	561	 81% 14% ..
1	E	561	 85% 13% ..

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	561	83% 11% . .
1	G	561	82% 13% . .
1	H	561	81% 13% . .
1	I	561	81% 13% . 5%
1	J	561	86% 11% . .
2	K	2	50% 50%
2	L	2	100%
2	M	2	50% 50%
2	N	2	50% 50%
2	O	2	100%
2	P	2	50% 50%
2	Q	2	100%
2	R	2	100%
2	S	2	50% 50%
2	T	2	100%

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 39838 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 Sugar ABC transporter substrate-binding protein, Gasdermin-B fusion protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	553	Total	C	N	O	S	0	0	0
			4066	2613	658	781	14			
1	B	554	Total	C	N	O	S	0	0	0
			4064	2603	664	784	13			
1	C	554	Total	C	N	O	S	0	0	0
			4070	2608	664	785	13			
1	D	541	Total	C	N	O	S	0	0	0
			3959	2542	643	762	12			
1	E	556	Total	C	N	O	S	0	0	0
			4077	2618	669	777	13			
1	F	541	Total	C	N	O	S	0	0	0
			3956	2546	633	765	12			
1	G	543	Total	C	N	O	S	0	0	0
			3885	2504	629	740	12			
1	H	537	Total	C	N	O	S	0	0	0
			3832	2459	623	739	11			
1	I	535	Total	C	N	O	S	0	0	0
			3795	2437	619	727	12			
1	J	548	Total	C	N	O	S	0	0	0
			3902	2500	638	754	10			

There are 120 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	82	ALA	ASP	see remark 999	UNP A0A178SBV6
A	83	ALA	LYS	see remark 999	UNP A0A178SBV6
A	172	ALA	GLU	see remark 999	UNP A0A178SBV6
A	173	ALA	ASN	see remark 999	UNP A0A178SBV6
A	239	ALA	LYS	see remark 999	UNP A0A178SBV6
A	362	ALA	LYS	see remark 999	UNP A0A178SBV6
A	363	ALA	ASP	see remark 999	UNP A0A178SBV6
A	367	ASN	-	linker	UNP A0A178SBV6

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	368	ALA	-	linker	UNP A0A178SBV6
A	369	ALA	-	linker	UNP A0A178SBV6
A	370	ALA	-	linker	UNP A0A178SBV6
A	1220	MET	MET	linker	UNP Q8TAX9
B	82	ALA	ASP	see remark 999	UNP A0A178SBV6
B	83	ALA	LYS	see remark 999	UNP A0A178SBV6
B	172	ALA	GLU	see remark 999	UNP A0A178SBV6
B	173	ALA	ASN	see remark 999	UNP A0A178SBV6
B	239	ALA	LYS	see remark 999	UNP A0A178SBV6
B	362	ALA	LYS	see remark 999	UNP A0A178SBV6
B	363	ALA	ASP	see remark 999	UNP A0A178SBV6
B	367	ASN	-	linker	UNP A0A178SBV6
B	368	ALA	-	linker	UNP A0A178SBV6
B	369	ALA	-	linker	UNP A0A178SBV6
B	370	ALA	-	linker	UNP A0A178SBV6
B	1220	MET	MET	linker	UNP Q8TAX9
C	82	ALA	ASP	see remark 999	UNP A0A178SBV6
C	83	ALA	LYS	see remark 999	UNP A0A178SBV6
C	172	ALA	GLU	see remark 999	UNP A0A178SBV6
C	173	ALA	ASN	see remark 999	UNP A0A178SBV6
C	239	ALA	LYS	see remark 999	UNP A0A178SBV6
C	362	ALA	LYS	see remark 999	UNP A0A178SBV6
C	363	ALA	ASP	see remark 999	UNP A0A178SBV6
C	367	ASN	-	linker	UNP A0A178SBV6
C	368	ALA	-	linker	UNP A0A178SBV6
C	369	ALA	-	linker	UNP A0A178SBV6
C	370	ALA	-	linker	UNP A0A178SBV6
C	1220	MET	MET	linker	UNP Q8TAX9
D	82	ALA	ASP	see remark 999	UNP A0A178SBV6
D	83	ALA	LYS	see remark 999	UNP A0A178SBV6
D	172	ALA	GLU	see remark 999	UNP A0A178SBV6
D	173	ALA	ASN	see remark 999	UNP A0A178SBV6
D	239	ALA	LYS	see remark 999	UNP A0A178SBV6
D	362	ALA	LYS	see remark 999	UNP A0A178SBV6
D	363	ALA	ASP	see remark 999	UNP A0A178SBV6
D	367	ASN	-	linker	UNP A0A178SBV6
D	368	ALA	-	linker	UNP A0A178SBV6
D	369	ALA	-	linker	UNP A0A178SBV6
D	370	ALA	-	linker	UNP A0A178SBV6
D	1220	MET	MET	linker	UNP Q8TAX9
E	82	ALA	ASP	see remark 999	UNP A0A178SBV6
E	83	ALA	LYS	see remark 999	UNP A0A178SBV6

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	172	ALA	GLU	see remark 999	UNP A0A178SBV6
E	173	ALA	ASN	see remark 999	UNP A0A178SBV6
E	239	ALA	LYS	see remark 999	UNP A0A178SBV6
E	362	ALA	LYS	see remark 999	UNP A0A178SBV6
E	363	ALA	ASP	see remark 999	UNP A0A178SBV6
E	367	ASN	-	linker	UNP A0A178SBV6
E	368	ALA	-	linker	UNP A0A178SBV6
E	369	ALA	-	linker	UNP A0A178SBV6
E	370	ALA	-	linker	UNP A0A178SBV6
E	1220	MET	MET	linker	UNP Q8TAX9
F	82	ALA	ASP	see remark 999	UNP A0A178SBV6
F	83	ALA	LYS	see remark 999	UNP A0A178SBV6
F	172	ALA	GLU	see remark 999	UNP A0A178SBV6
F	173	ALA	ASN	see remark 999	UNP A0A178SBV6
F	239	ALA	LYS	see remark 999	UNP A0A178SBV6
F	362	ALA	LYS	see remark 999	UNP A0A178SBV6
F	363	ALA	ASP	see remark 999	UNP A0A178SBV6
F	367	ASN	-	linker	UNP A0A178SBV6
F	368	ALA	-	linker	UNP A0A178SBV6
F	369	ALA	-	linker	UNP A0A178SBV6
F	370	ALA	-	linker	UNP A0A178SBV6
F	1220	MET	MET	linker	UNP Q8TAX9
G	82	ALA	ASP	see remark 999	UNP A0A178SBV6
G	83	ALA	LYS	see remark 999	UNP A0A178SBV6
G	172	ALA	GLU	see remark 999	UNP A0A178SBV6
G	173	ALA	ASN	see remark 999	UNP A0A178SBV6
G	239	ALA	LYS	see remark 999	UNP A0A178SBV6
G	362	ALA	LYS	see remark 999	UNP A0A178SBV6
G	363	ALA	ASP	see remark 999	UNP A0A178SBV6
G	367	ASN	-	linker	UNP A0A178SBV6
G	368	ALA	-	linker	UNP A0A178SBV6
G	369	ALA	-	linker	UNP A0A178SBV6
G	370	ALA	-	linker	UNP A0A178SBV6
G	1220	MET	MET	linker	UNP Q8TAX9
H	82	ALA	ASP	see remark 999	UNP A0A178SBV6
H	83	ALA	LYS	see remark 999	UNP A0A178SBV6
H	172	ALA	GLU	see remark 999	UNP A0A178SBV6
H	173	ALA	ASN	see remark 999	UNP A0A178SBV6
H	239	ALA	LYS	see remark 999	UNP A0A178SBV6
H	362	ALA	LYS	see remark 999	UNP A0A178SBV6
H	363	ALA	ASP	see remark 999	UNP A0A178SBV6
H	367	ASN	-	linker	UNP A0A178SBV6

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	368	ALA	-	linker	UNP A0A178SBV6
H	369	ALA	-	linker	UNP A0A178SBV6
H	370	ALA	-	linker	UNP A0A178SBV6
H	1220	MET	MET	linker	UNP Q8TAX9
I	82	ALA	ASP	see remark 999	UNP A0A178SBV6
I	83	ALA	LYS	see remark 999	UNP A0A178SBV6
I	172	ALA	GLU	see remark 999	UNP A0A178SBV6
I	173	ALA	ASN	see remark 999	UNP A0A178SBV6
I	239	ALA	LYS	see remark 999	UNP A0A178SBV6
I	362	ALA	LYS	see remark 999	UNP A0A178SBV6
I	363	ALA	ASP	see remark 999	UNP A0A178SBV6
I	367	ASN	-	linker	UNP A0A178SBV6
I	368	ALA	-	linker	UNP A0A178SBV6
I	369	ALA	-	linker	UNP A0A178SBV6
I	370	ALA	-	linker	UNP A0A178SBV6
I	1220	MET	MET	linker	UNP Q8TAX9
J	82	ALA	ASP	see remark 999	UNP A0A178SBV6
J	83	ALA	LYS	see remark 999	UNP A0A178SBV6
J	172	ALA	GLU	see remark 999	UNP A0A178SBV6
J	173	ALA	ASN	see remark 999	UNP A0A178SBV6
J	239	ALA	LYS	see remark 999	UNP A0A178SBV6
J	362	ALA	LYS	see remark 999	UNP A0A178SBV6
J	363	ALA	ASP	see remark 999	UNP A0A178SBV6
J	367	ASN	-	linker	UNP A0A178SBV6
J	368	ALA	-	linker	UNP A0A178SBV6
J	369	ALA	-	linker	UNP A0A178SBV6
J	370	ALA	-	linker	UNP A0A178SBV6
J	1220	MET	MET	linker	UNP Q8TAX9

- Molecule 2 is an oligosaccharide called alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose.



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
2	K	2	Total	C	O	0	0	0
			23	12	11			
2	L	2	Total	C	O	0	0	0
			23	12	11			
2	M	2	Total	C	O	0	0	0
			23	12	11			

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
2	N	2	Total C O 23 12 11	0	0	0
2	O	2	Total C O 23 12 11	0	0	0
2	P	2	Total C O 23 12 11	0	0	0
2	Q	2	Total C O 23 12 11	0	0	0
2	R	2	Total C O 23 12 11	0	0	0
2	S	2	Total C O 23 12 11	0	0	0
2	T	2	Total C O 23 12 11	0	0	0

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

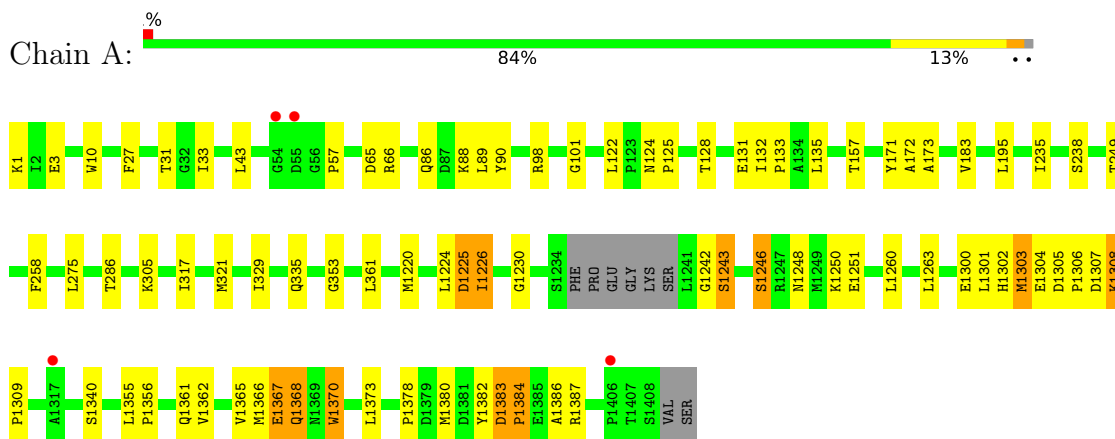
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total Na 1 1	0	0
3	E	1	Total Na 1 1	0	0



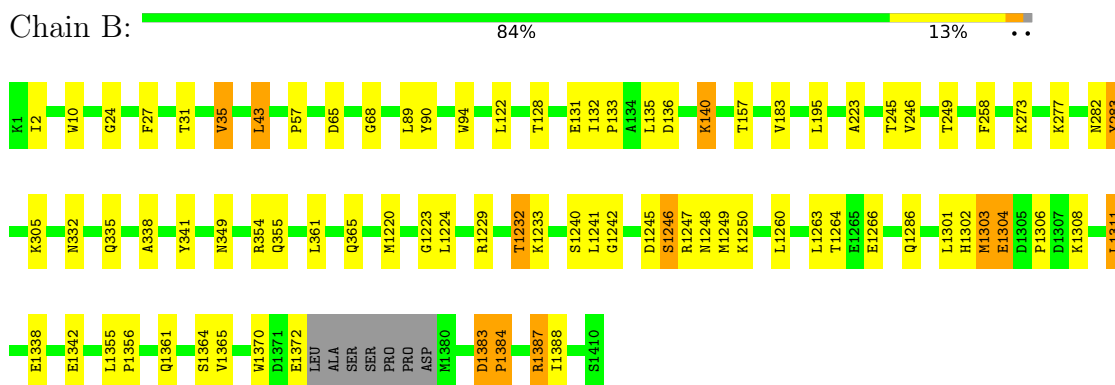
### 3 Residue-property plots

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.

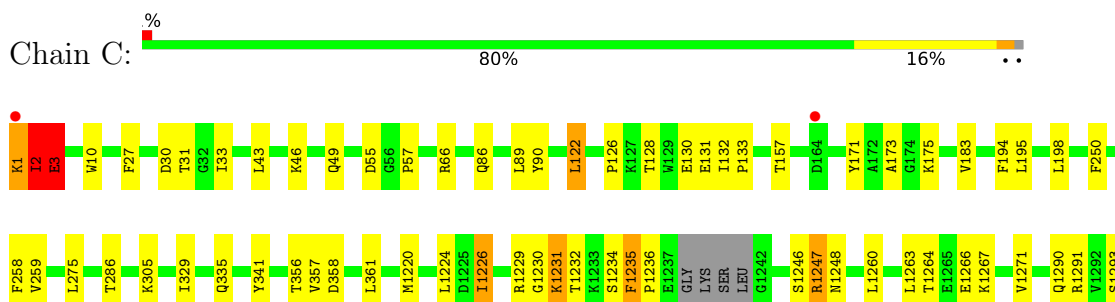
- Molecule 1: Sugar ABC transporter substrate-binding protein, Gasdermin-B fusion protein

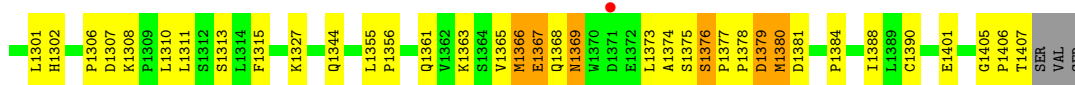


- Molecule 1: Sugar ABC transporter substrate-binding protein, Gasdermin-B fusion protein

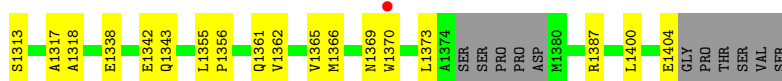
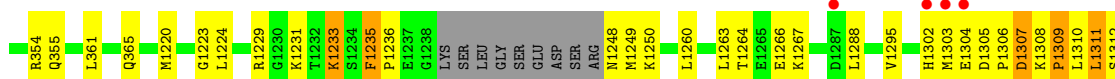
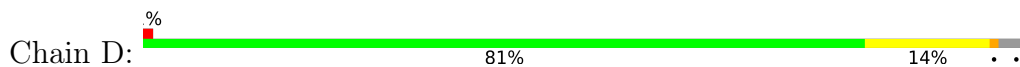


- Molecule 1: Sugar ABC transporter substrate-binding protein, Gasdermin-B fusion protein

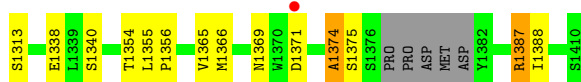
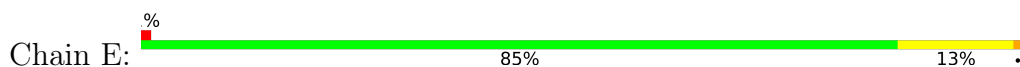




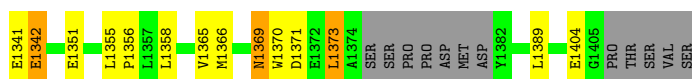
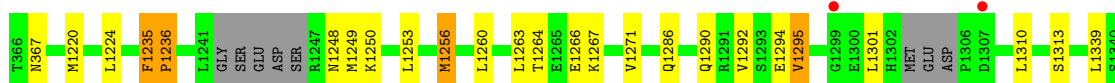
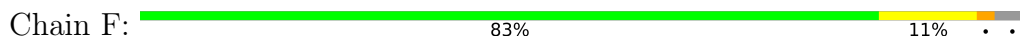
- Molecule 1: Sugar ABC transporter substrate-binding protein,Gasdermin-B fusion protein



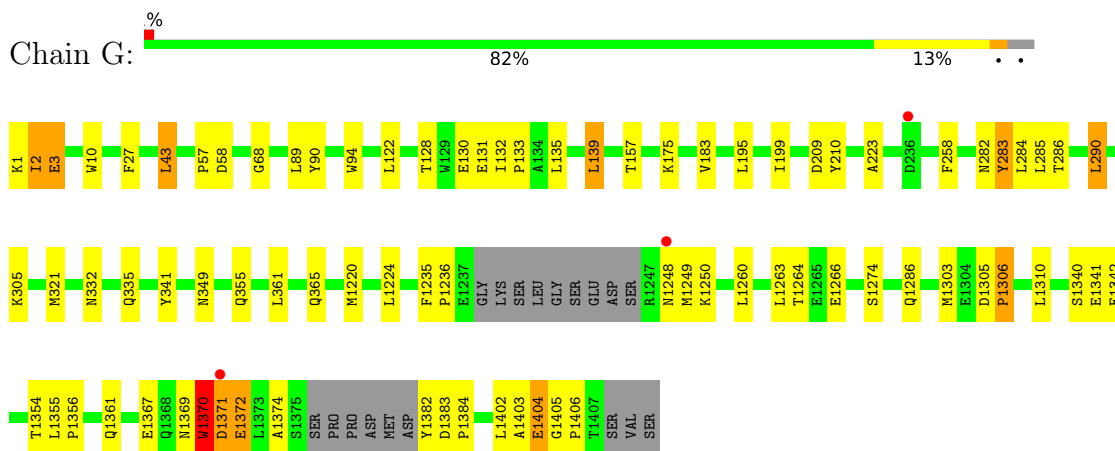
- Molecule 1: Sugar ABC transporter substrate-binding protein,Gasdermin-B fusion protein



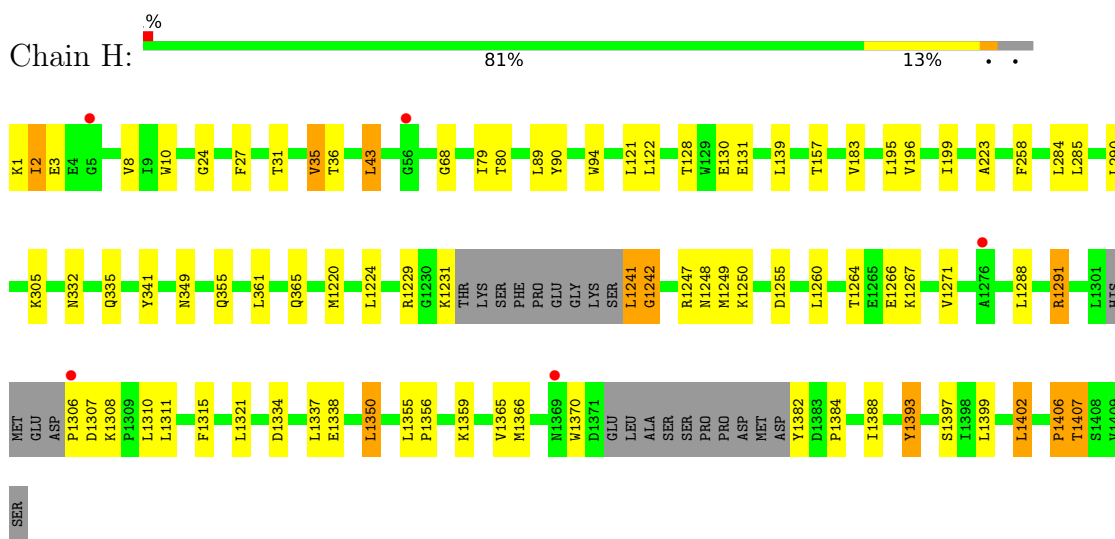
- Molecule 1: Sugar ABC transporter substrate-binding protein,Gasdermin-B fusion protein



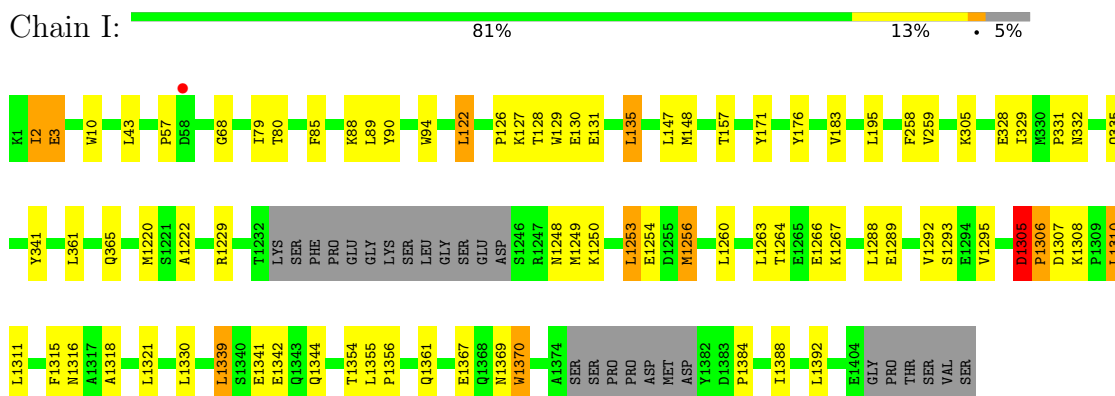
- Molecule 1: Sugar ABC transporter substrate-binding protein,Gasdermin-B fusion protein



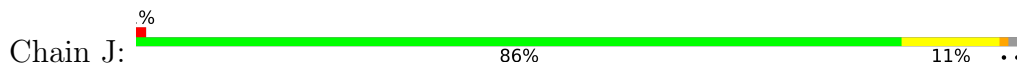
- Molecule 1: Sugar ABC transporter substrate-binding protein,Gasdermin-B fusion protein

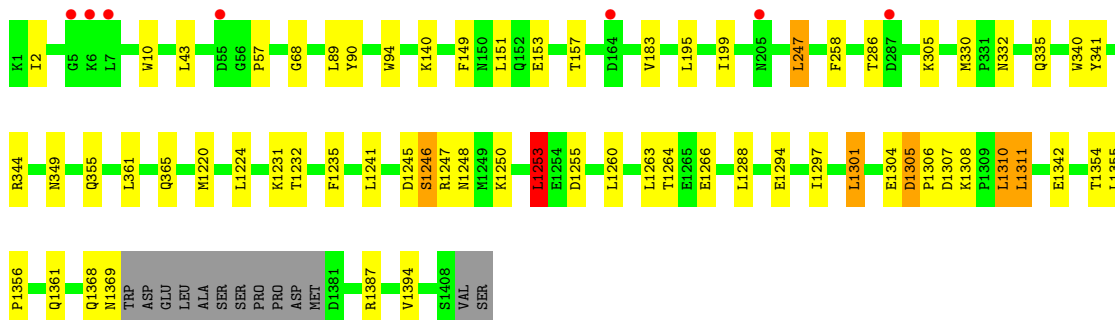


- Molecule 1: Sugar ABC transporter substrate-binding protein,Gasdermin-B fusion protein



- Molecule 1: Sugar ABC transporter substrate-binding protein,Gasdermin-B fusion protein





- Molecule 2: alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose

Chain K: 50% 50%

GLC1  
GLC2

- Molecule 2: alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose

Chain L: 100%

GLC1  
GLC2

- Molecule 2: alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose

Chain M: 50% 50%

GLC1  
GLC2

- Molecule 2: alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose

Chain N: 50% 50%

GLC1  
GLC2

- Molecule 2: alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose

Chain O: 100%

GLC1  
GLC2

- Molecule 2: alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose

Chain P: 50% 50%

GLC1  
GLC2

- Molecule 2: alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose

Chain Q:  100%

GLC1  
GLC2

- Molecule 2: alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose

Chain R:  100%

GLC1  
GLC2

- Molecule 2: alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose

Chain S:  50% 50%

GLC1  
GLC2

- Molecule 2: alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose

Chain T:  100%

GLC1  
GLC2

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	89.60Å 274.57Å 174.02Å 90.00° 96.09° 90.00°	Depositor
Resolution (Å)	173.04 – 3.50 37.39 – 3.50	Depositor EDS
% Data completeness (in resolution range)	99.7 (173.04-3.50) 99.8 (37.39-3.50)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.77 (at 3.48Å)	Xtrriage
Refinement program	REFMAC 5.8.0135	Depositor
R, $R_{free}$	0.213 , 0.248 0.214 , 0.217	Depositor DCC
$R_{free}$ test set	5232 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	89.4	Xtrriage
Anisotropy	0.011	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 80.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.27$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	39838	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	94.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.25% 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: NA, GLC

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	0.55	0/4149	0.79	5/5661 (0.1%)
1	B	0.53	0/4147	0.79	9/5659 (0.2%)
1	C	0.53	0/4157	0.80	5/5681 (0.1%)
1	D	0.50	0/4040	0.73	3/5519 (0.1%)
1	E	0.52	0/4161	0.78	7/5680 (0.1%)
1	F	0.52	0/4037	0.75	4/5518 (0.1%)
1	G	0.51	0/3968	0.76	11/5441 (0.2%)
1	H	0.52	0/3910	0.79	9/5356 (0.2%)
1	I	0.50	0/3874	0.72	3/5318 (0.1%)
1	J	0.52	0/3984	0.77	7/5461 (0.1%)
All	All	0.52	0/40427	0.77	63/55294 (0.1%)

There are no bond length outliers.

The worst 5 of 63 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	89	LEU	CB-CG-CD2	10.84	129.42	111.00
1	C	89	LEU	CB-CG-CD2	10.65	129.10	111.00
1	A	89	LEU	CB-CG-CD2	10.52	128.88	111.00
1	A	65	ASP	CB-CG-OD2	10.29	127.56	118.30
1	B	65	ASP	CB-CG-OD2	10.11	127.40	118.30

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	4066	0	3871	88	0
1	B	4064	0	3833	65	0
1	C	4070	0	3841	115	0
1	D	3959	0	3729	66	0
1	E	4077	0	3874	53	0
1	F	3956	0	3714	65	0
1	G	3885	0	3583	72	0
1	H	3832	0	3525	78	0
1	I	3795	0	3444	88	1
1	J	3902	0	3567	45	1
2	K	23	0	21	1	0
2	L	23	0	21	0	0
2	M	23	0	21	1	0
2	N	23	0	21	1	0
2	O	23	0	21	0	0
2	P	23	0	21	1	0
2	Q	23	0	21	0	0
2	R	23	0	21	0	0
2	S	23	0	21	0	0
2	T	23	0	21	0	0
3	B	1	0	0	0	0
3	E	1	0	0	0	0
All	All	39838	0	37191	704	1

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 9.

The worst 5 of 704 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:I:259:VAL:CG2	1:I:329:ILE:HA	1.73	1.18
1:E:184:ASP:CB	1:E:365:GLN:OE1	1.96	1.14
1:H:1241:LEU:HD23	1:H:1242:GLY:H	1.07	1.12
1:D:2:ILE:HG22	1:D:3:GLU:H	1.03	1.12
1:C:1247:ARG:HH22	1:C:1327:LYS:HE3	1.01	1.11

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.



Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:I:171:TYR:OH	1:J:1231:LYS:O[2_444]	2.12	0.08

### 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	547/561 (98%)	504 (92%)	30 (6%)	13 (2%)	6	35
1	B	550/561 (98%)	510 (93%)	29 (5%)	11 (2%)	7	39
1	C	550/561 (98%)	517 (94%)	21 (4%)	12 (2%)	6	37
1	D	535/561 (95%)	505 (94%)	22 (4%)	8 (2%)	10	45
1	E	552/561 (98%)	520 (94%)	27 (5%)	5 (1%)	17	56
1	F	533/561 (95%)	511 (96%)	18 (3%)	4 (1%)	19	58
1	G	537/561 (96%)	513 (96%)	17 (3%)	7 (1%)	12	48
1	H	529/561 (94%)	504 (95%)	21 (4%)	4 (1%)	19	58
1	I	529/561 (94%)	509 (96%)	14 (3%)	6 (1%)	14	52
1	J	544/561 (97%)	521 (96%)	19 (4%)	4 (1%)	22	61
All	All	5406/5610 (96%)	5114 (95%)	218 (4%)	74 (1%)	11	46

5 of 74 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1226	ILE
1	A	1303	MET
1	A	1368	GLN
1	A	1370	TRP
1	A	1383	ASP

### 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	388/454 (86%)	385 (99%)	3 (1%)	81	91
1	B	384/454 (85%)	377 (98%)	7 (2%)	59	81
1	C	389/454 (86%)	383 (98%)	6 (2%)	65	84
1	D	372/454 (82%)	368 (99%)	4 (1%)	73	88
1	E	385/454 (85%)	381 (99%)	4 (1%)	76	88
1	F	372/454 (82%)	361 (97%)	11 (3%)	41	71
1	G	350/454 (77%)	345 (99%)	5 (1%)	67	85
1	H	347/454 (76%)	340 (98%)	7 (2%)	55	79
1	I	335/454 (74%)	329 (98%)	6 (2%)	59	81
1	J	351/454 (77%)	345 (98%)	6 (2%)	60	82
All	All	3673/4540 (81%)	3614 (98%)	59 (2%)	62	83

5 of 59 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	365	GLN
1	J	1241	LEU
1	G	139	LEU
1	J	365	GLN
1	I	365	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	86	GLN
1	C	1227	HIS
1	G	124	ASN
1	D	325	GLN
1	D	1361	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates [i](#)

20 monosaccharides are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	GLC	K	1	2	12,12,12	1.02	0	17,17,17	1.81	3 (17%)
2	GLC	K	2	2	11,11,12	0.44	0	15,15,17	1.48	3 (20%)
2	GLC	L	1	2	12,12,12	0.67	0	17,17,17	1.22	2 (11%)
2	GLC	L	2	2	11,11,12	0.43	0	15,15,17	1.26	2 (13%)
2	GLC	M	1	2	12,12,12	0.48	0	17,17,17	1.13	1 (5%)
2	GLC	M	2	2	11,11,12	0.68	0	15,15,17	1.49	3 (20%)
2	GLC	N	1	2	12,12,12	0.67	0	17,17,17	1.14	2 (11%)
2	GLC	N	2	2	11,11,12	0.53	0	15,15,17	1.08	1 (6%)
2	GLC	O	1	2	12,12,12	0.72	0	17,17,17	1.51	4 (23%)
2	GLC	O	2	2	11,11,12	0.36	0	15,15,17	1.63	2 (13%)
2	GLC	P	1	2	12,12,12	0.65	0	17,17,17	1.17	2 (11%)
2	GLC	P	2	2	11,11,12	0.60	0	15,15,17	1.13	2 (13%)
2	GLC	Q	1	2	12,12,12	0.77	0	17,17,17	1.76	4 (23%)
2	GLC	Q	2	2	11,11,12	0.52	0	15,15,17	1.35	1 (6%)
2	GLC	R	1	2	12,12,12	0.63	0	17,17,17	1.41	3 (17%)
2	GLC	R	2	2	11,11,12	0.33	0	15,15,17	2.09	4 (26%)
2	GLC	S	1	2	12,12,12	0.60	0	17,17,17	1.00	0
2	GLC	S	2	2	11,11,12	0.84	0	15,15,17	1.51	2 (13%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	GLC	T	1	2	12,12,12	0.65	0	17,17,17	1.68	4 (23%)
2	GLC	T	2	2	11,11,12	0.39	0	15,15,17	0.83	1 (6%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GLC	K	1	2	-	0/2/22/22	0/1/1/1
2	GLC	K	2	2	-	0/2/19/22	0/1/1/1
2	GLC	L	1	2	-	0/2/22/22	0/1/1/1
2	GLC	L	2	2	-	2/2/19/22	0/1/1/1
2	GLC	M	1	2	-	2/2/22/22	0/1/1/1
2	GLC	M	2	2	-	1/2/19/22	0/1/1/1
2	GLC	N	1	2	-	0/2/22/22	0/1/1/1
2	GLC	N	2	2	-	0/2/19/22	0/1/1/1
2	GLC	O	1	2	-	2/2/22/22	0/1/1/1
2	GLC	O	2	2	-	0/2/19/22	0/1/1/1
2	GLC	P	1	2	-	2/2/22/22	0/1/1/1
2	GLC	P	2	2	-	0/2/19/22	0/1/1/1
2	GLC	Q	1	2	-	2/2/22/22	0/1/1/1
2	GLC	Q	2	2	-	1/2/19/22	0/1/1/1
2	GLC	R	1	2	-	0/2/22/22	0/1/1/1
2	GLC	R	2	2	-	0/2/19/22	0/1/1/1
2	GLC	S	1	2	-	2/2/22/22	0/1/1/1
2	GLC	S	2	2	-	1/2/19/22	0/1/1/1
2	GLC	T	1	2	-	2/2/22/22	0/1/1/1
2	GLC	T	2	2	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

The worst 5 of 46 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	R	2	GLC	C1-O5-C5	5.76	120.00	112.19
2	Q	2	GLC	C1-O5-C5	4.02	117.64	112.19
2	O	2	GLC	C1-O5-C5	3.82	117.37	112.19
2	K	1	GLC	C1-C2-C3	3.80	118.19	110.31
2	S	2	GLC	C1-O5-C5	3.79	117.33	112.19

There are no chirality outliers.

5 of 17 torsion outliers are listed below:

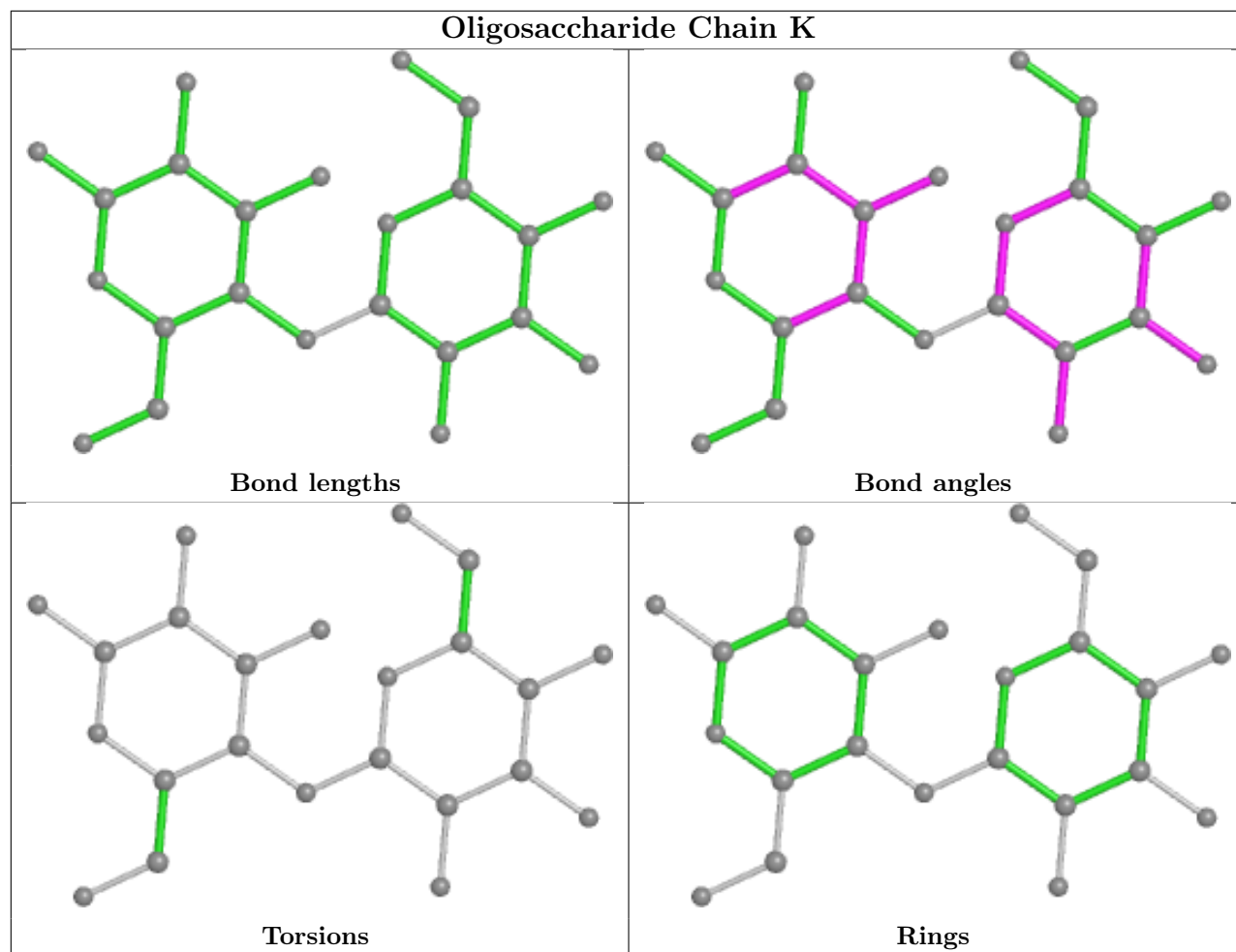
Mol	Chain	Res	Type	Atoms
2	O	1	GLC	O5-C5-C6-O6
2	S	1	GLC	C4-C5-C6-O6
2	O	1	GLC	C4-C5-C6-O6
2	S	1	GLC	O5-C5-C6-O6
2	P	1	GLC	O5-C5-C6-O6

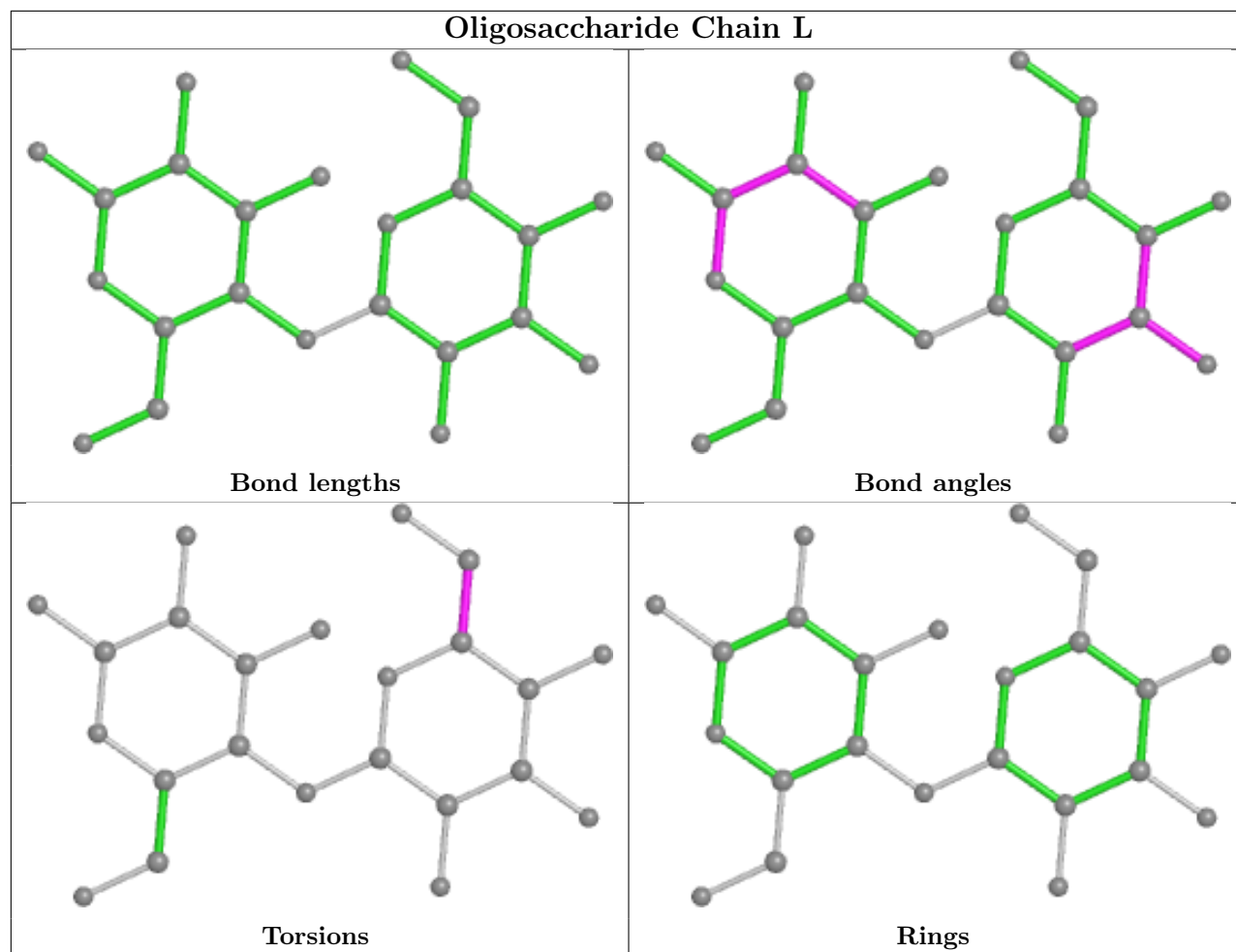
There are no ring outliers.

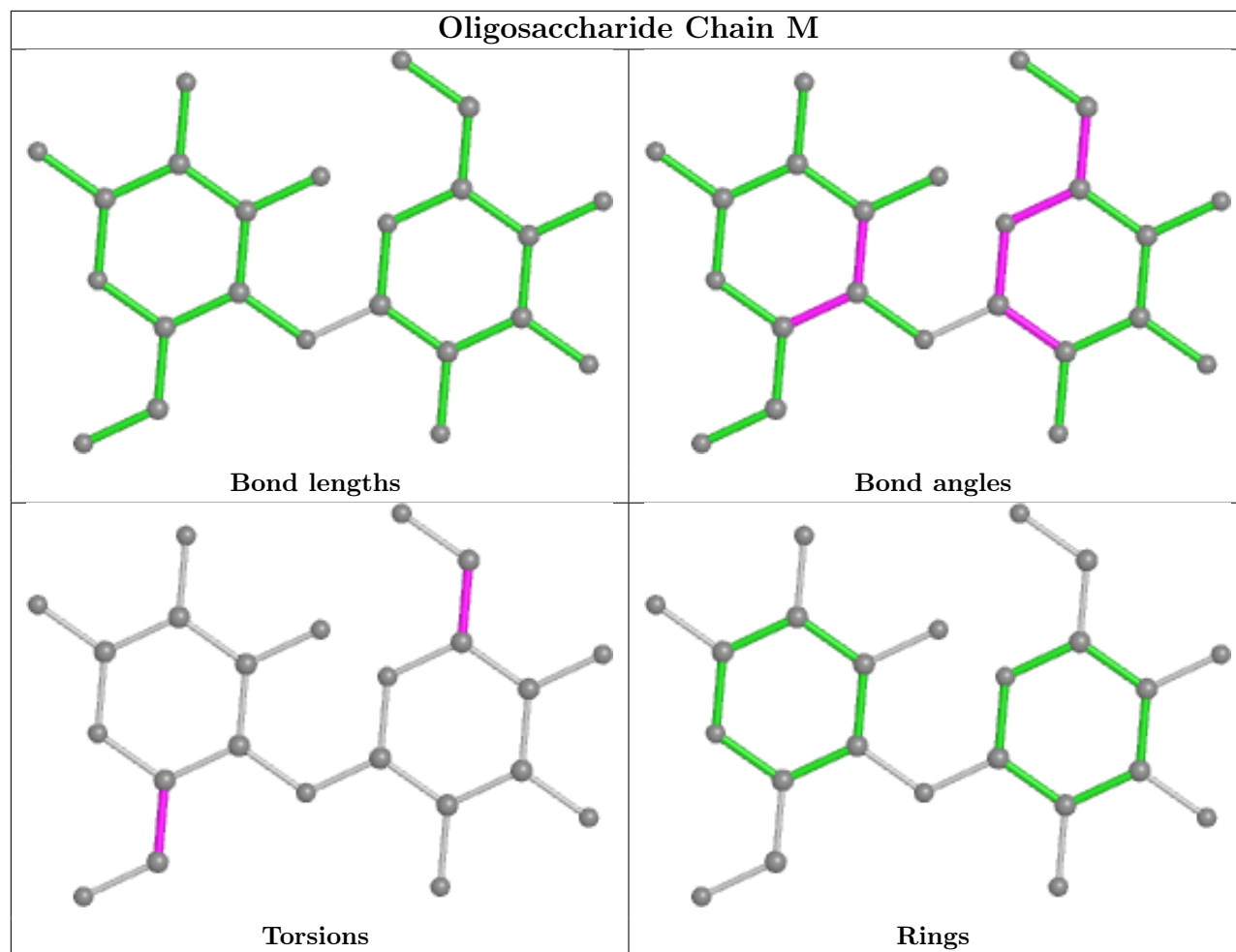
4 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	P	2	GLC	1	0
2	M	2	GLC	1	0
2	K	2	GLC	1	0
2	N	2	GLC	1	0

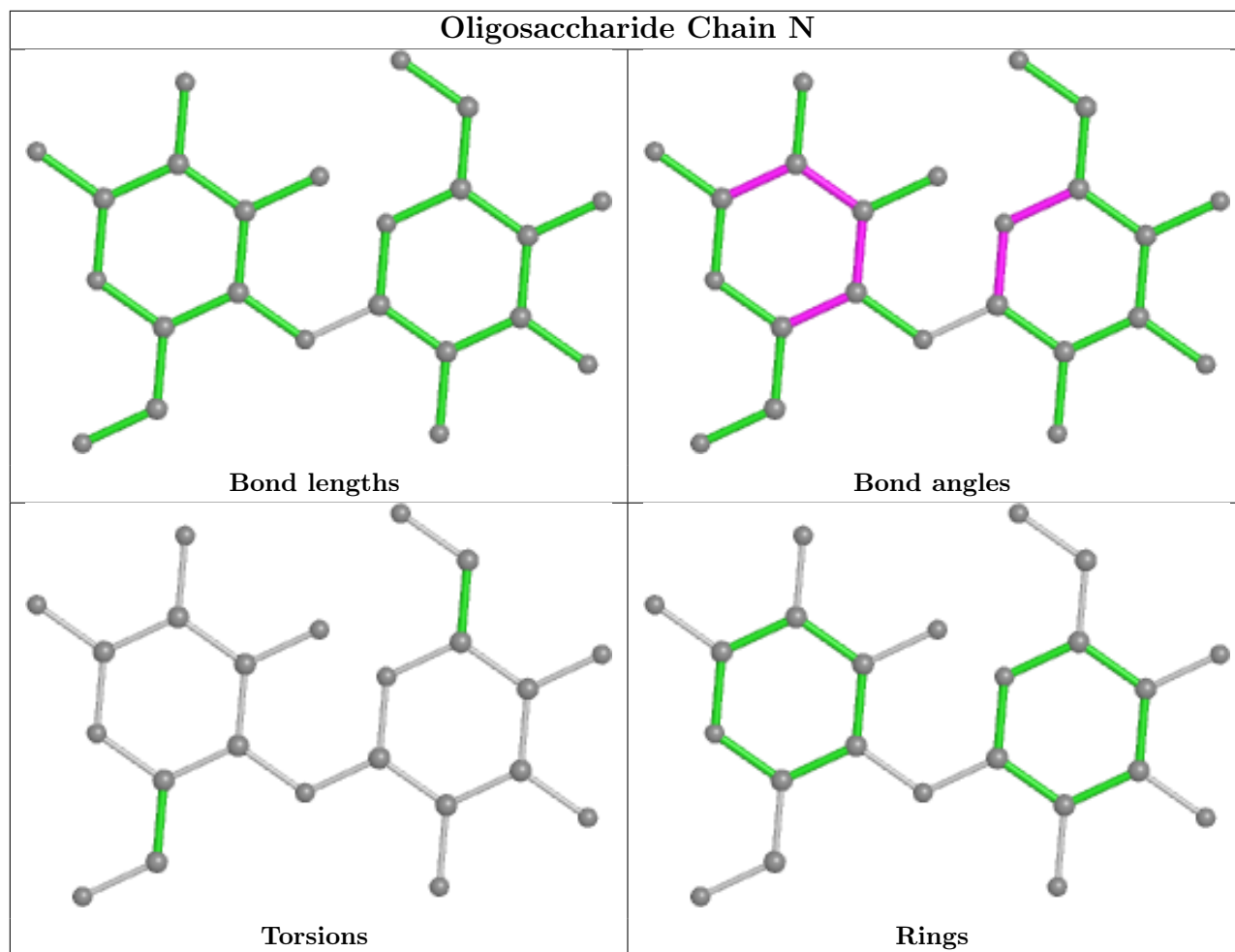
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

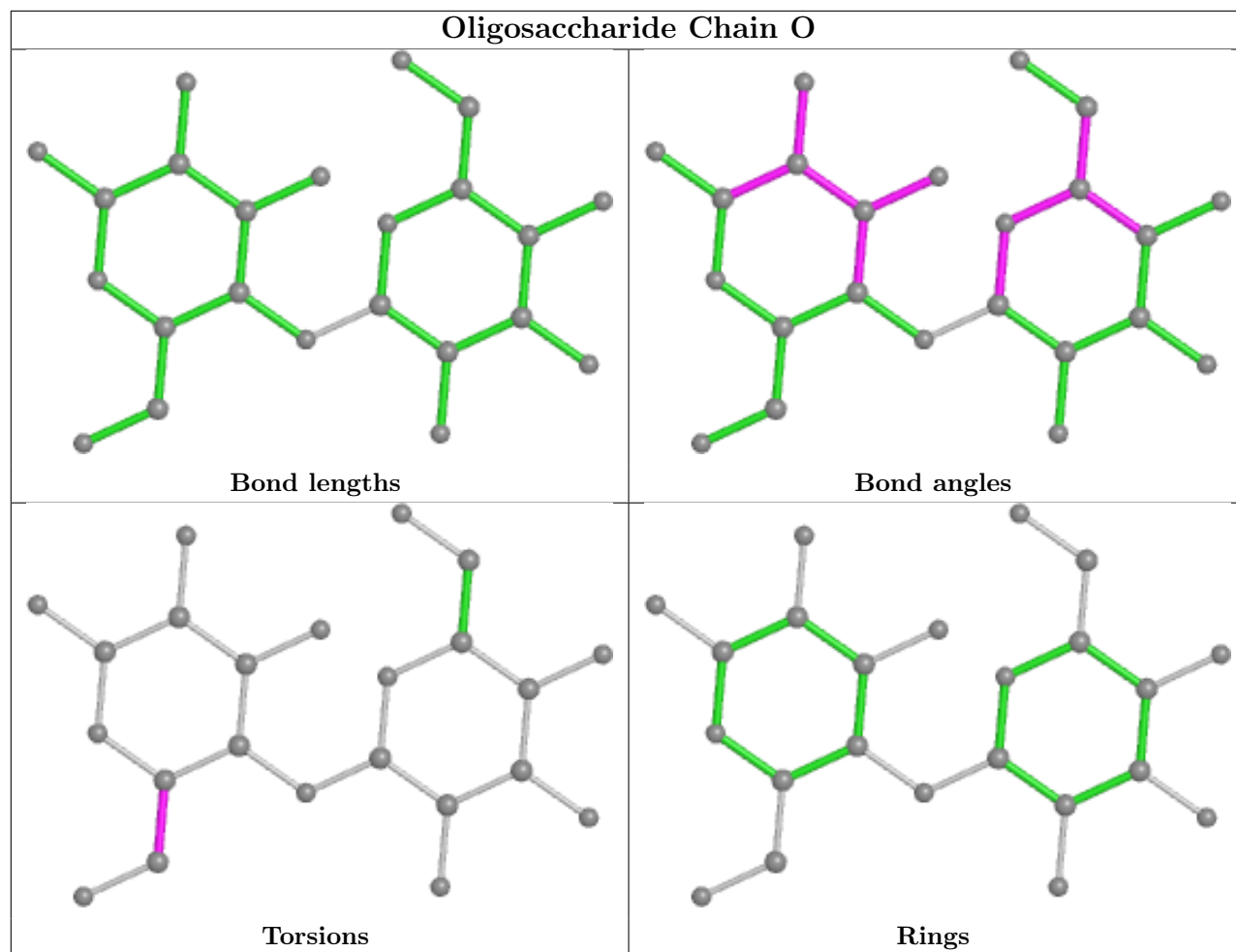


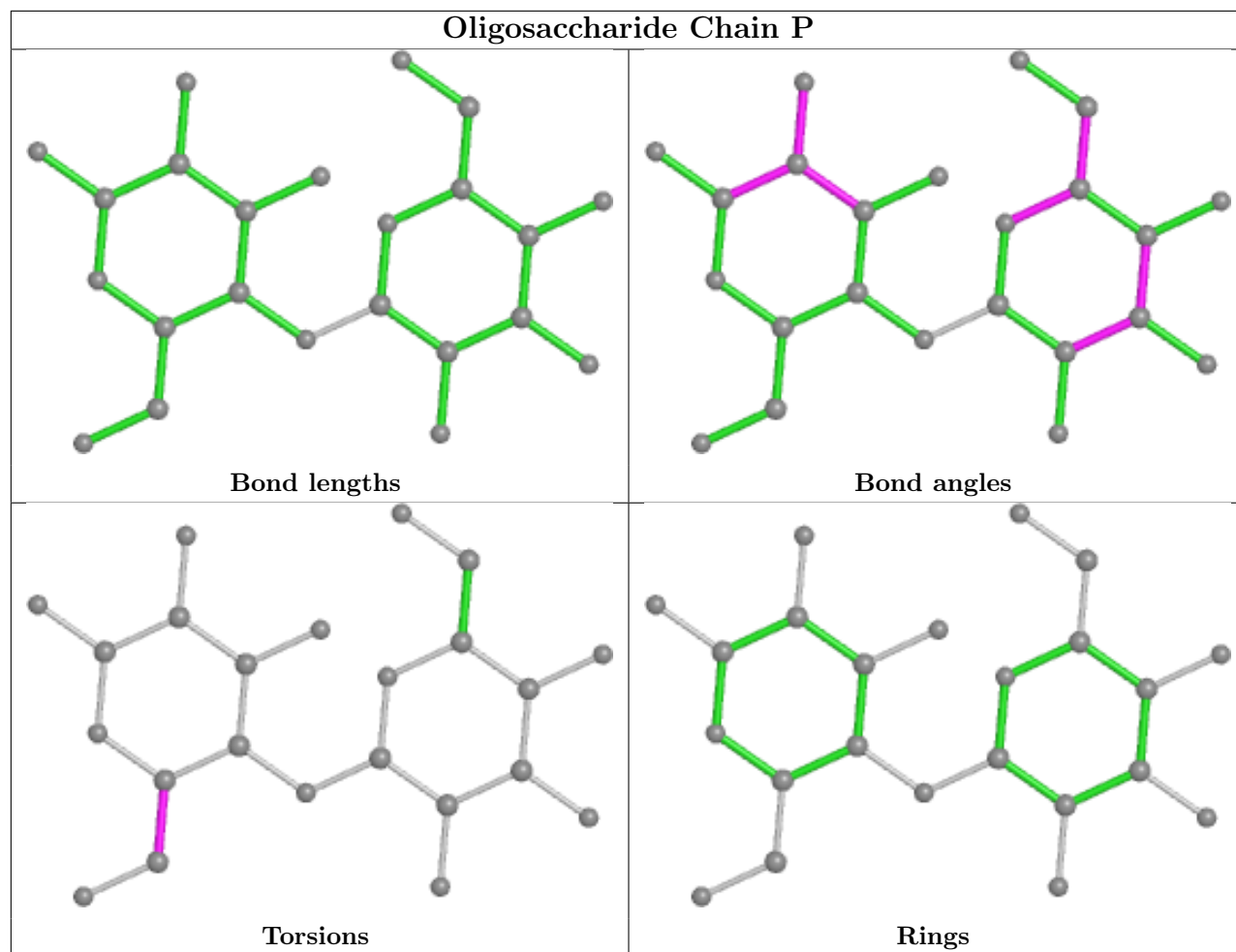


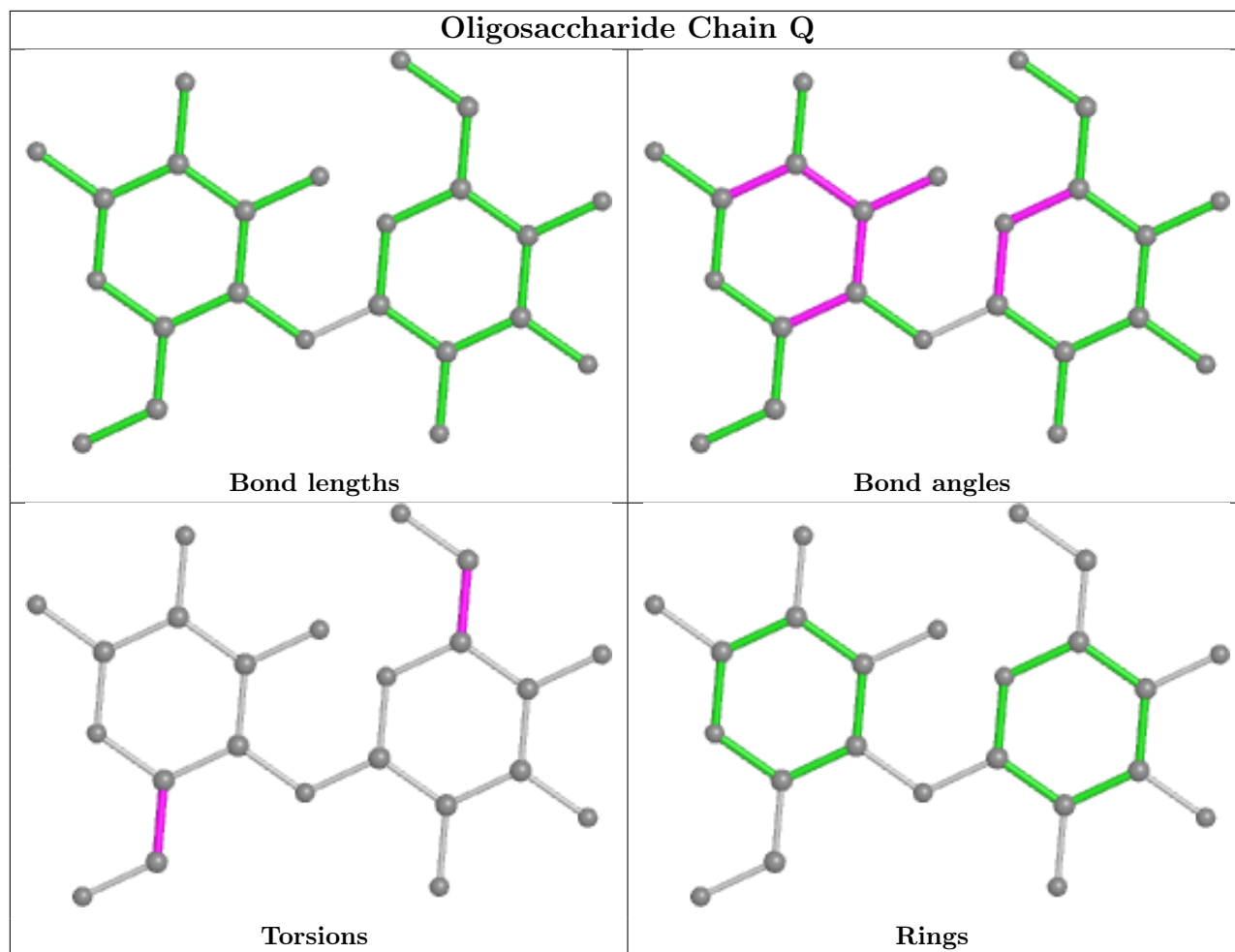


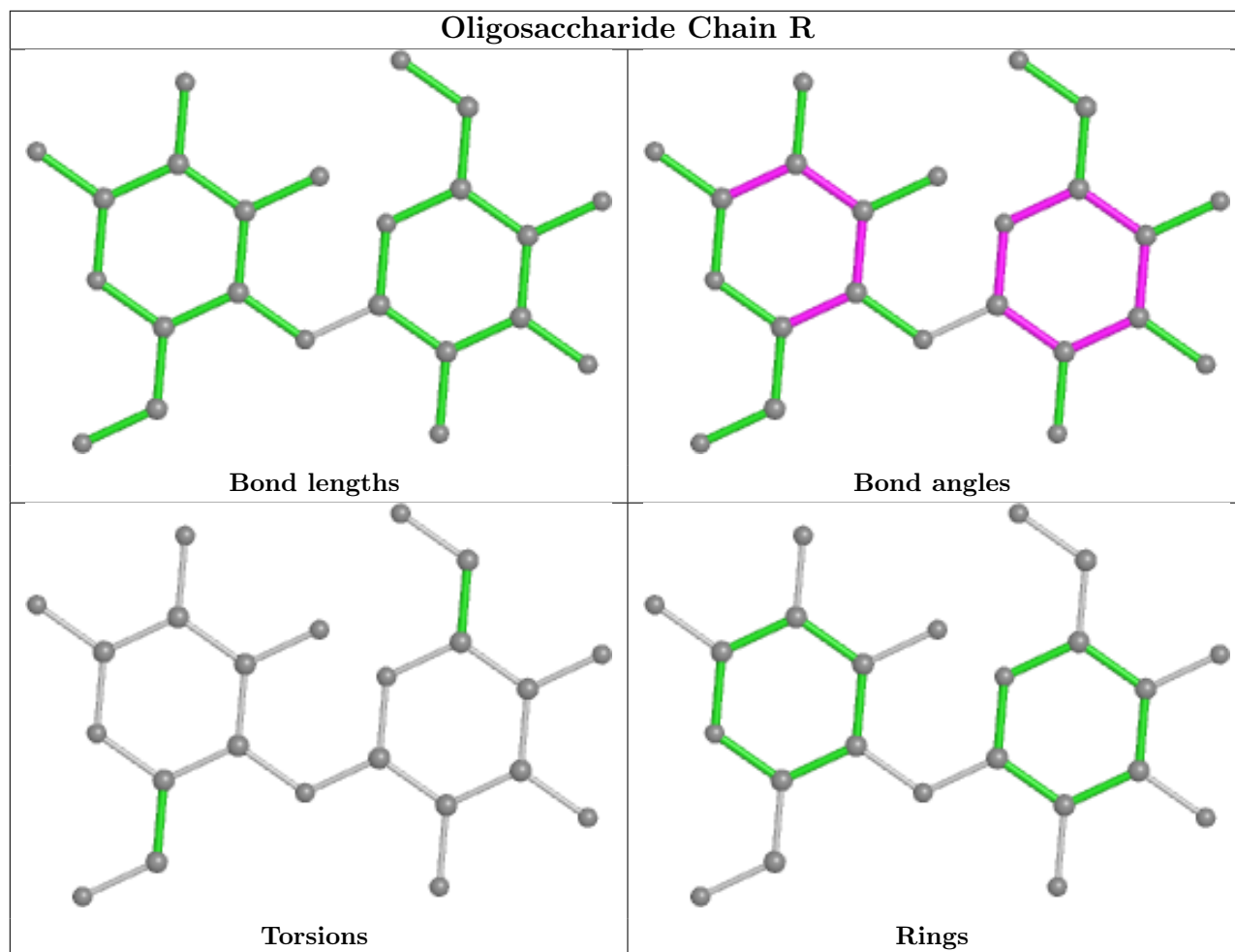


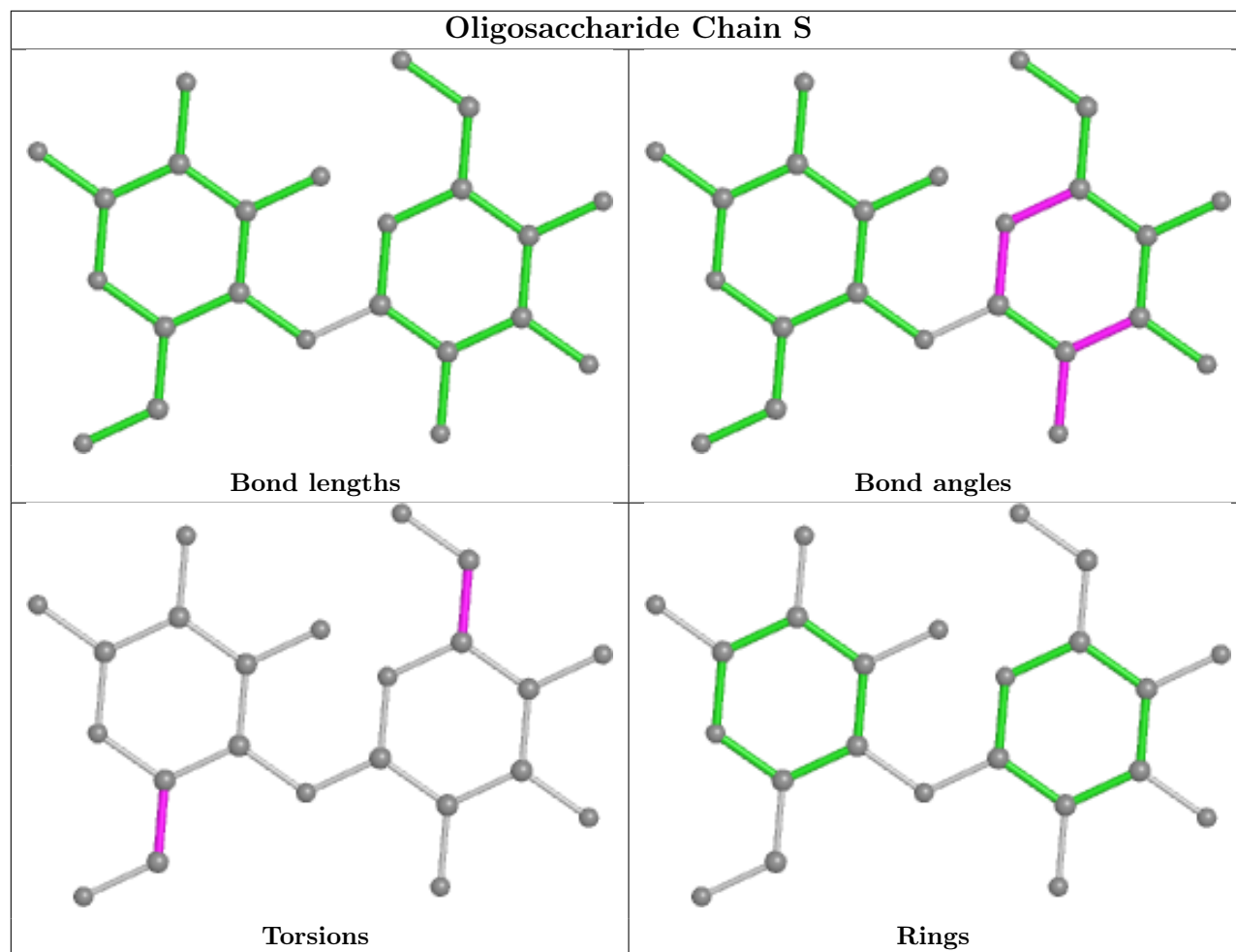


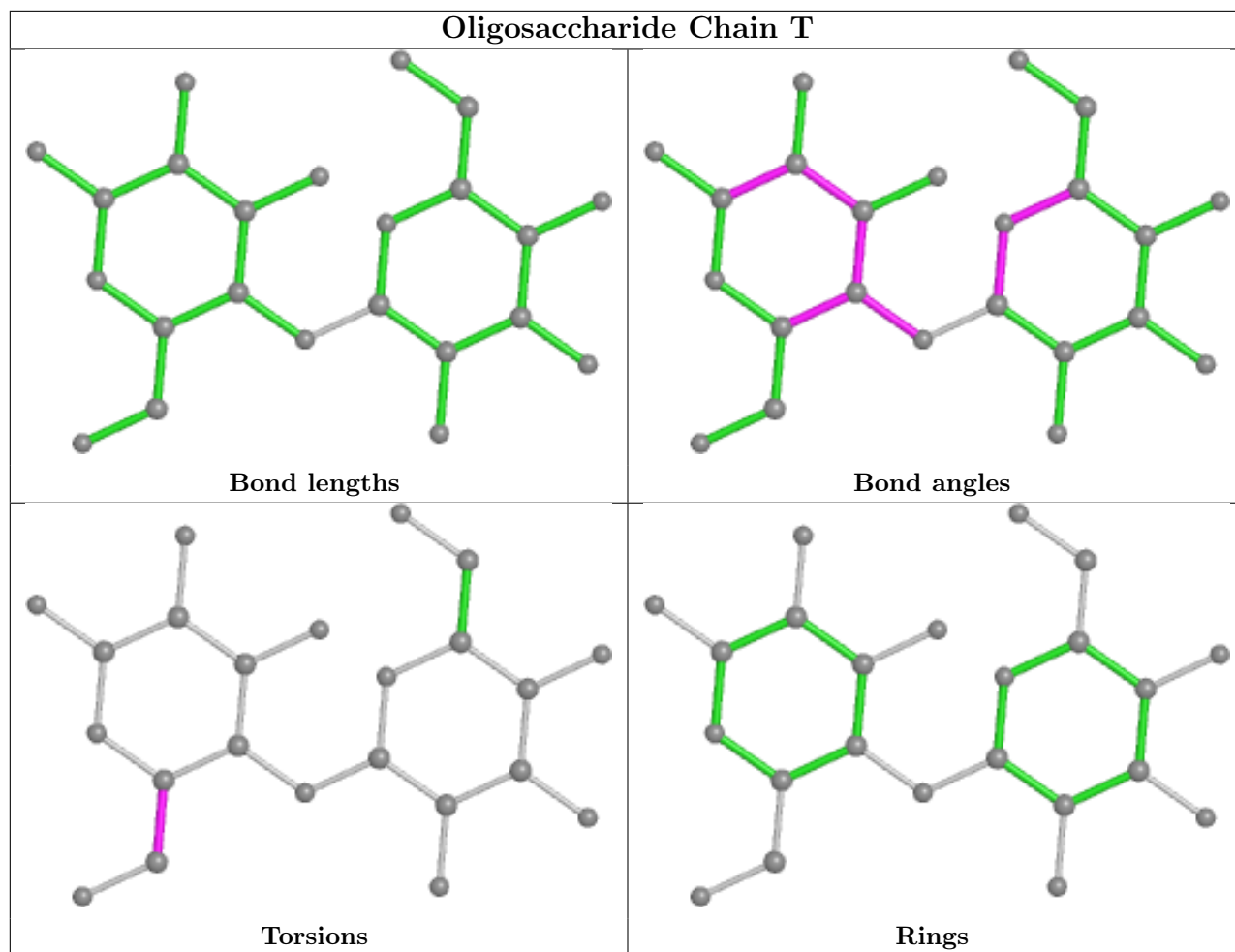












## 5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	1405:GLY	C	1406:PRO	N	3.42



## 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	553/561 (98%)	-0.43	4 (0%) 87 83	36, 77, 141, 196	0
1	B	554/561 (98%)	-0.50	0 100 100	38, 75, 128, 171	0
1	C	554/561 (98%)	-0.41	3 (0%) 91 88	39, 82, 130, 181	0
1	D	541/561 (96%)	-0.28	6 (1%) 80 75	40, 79, 144, 200	0
1	E	556/561 (99%)	-0.48	5 (0%) 84 79	36, 76, 125, 175	0
1	F	541/561 (96%)	-0.37	2 (0%) 92 90	47, 86, 129, 193	0
1	G	543/561 (96%)	-0.27	3 (0%) 89 86	55, 104, 146, 204	0
1	H	537/561 (95%)	-0.08	5 (0%) 84 79	64, 112, 148, 184	0
1	I	535/561 (95%)	-0.15	1 (0%) 95 93	68, 118, 161, 205	0
1	J	548/561 (97%)	-0.18	7 (1%) 77 71	54, 120, 157, 202	0
All	All	5462/5610 (97%)	-0.32	36 (0%) 87 83	36, 93, 147, 205	0

The worst 5 of 36 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	1371	ASP	4.0
1	D	1303	MET	3.3
1	J	205	ASN	3.2
1	D	1302	HIS	3.0
1	A	1317	ALA	3.0

### 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

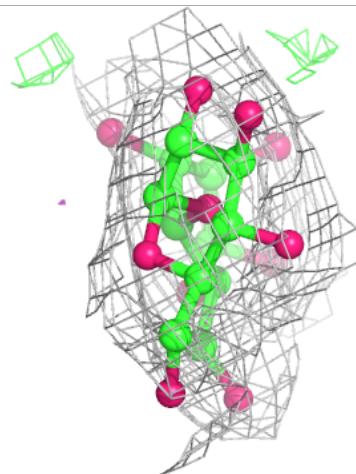
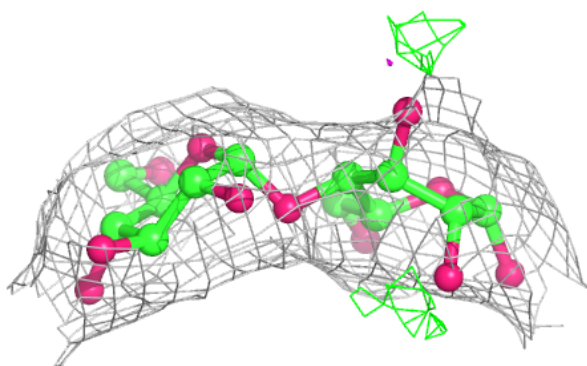
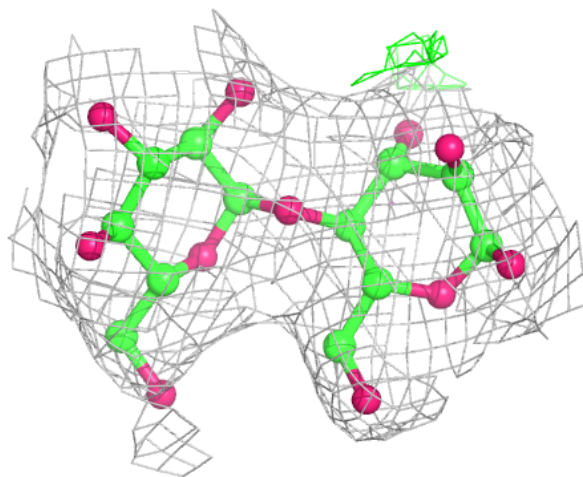
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	GLC	M	1	12/12	0.95	0.27	62,72,82,88	0
2	GLC	M	2	11/12	0.95	0.32	51,63,77,80	0
2	GLC	S	2	11/12	0.95	0.25	58,67,69,69	0
2	GLC	R	2	11/12	0.96	0.32	69,78,88,90	0
2	GLC	S	1	12/12	0.96	0.25	63,67,76,81	0
2	GLC	R	1	12/12	0.96	0.32	76,87,106,114	0
2	GLC	T	1	12/12	0.96	0.21	93,97,115,116	0
2	GLC	Q	2	11/12	0.97	0.24	53,59,62,63	0
2	GLC	K	1	12/12	0.97	0.27	42,46,54,67	0
2	GLC	L	1	12/12	0.97	0.23	50,56,64,66	0
2	GLC	N	1	12/12	0.97	0.29	48,59,70,76	0
2	GLC	N	2	11/12	0.97	0.24	39,44,48,48	0
2	GLC	Q	1	12/12	0.97	0.25	57,64,72,84	0
2	GLC	T	2	11/12	0.97	0.21	76,87,94,97	0
2	GLC	O	2	11/12	0.98	0.20	45,47,50,55	0
2	GLC	P	1	12/12	0.98	0.21	46,53,63,70	0
2	GLC	K	2	11/12	0.98	0.26	41,48,66,72	0
2	GLC	L	2	11/12	0.98	0.22	46,50,53,56	0
2	GLC	O	1	12/12	0.98	0.17	59,68,72,80	0
2	GLC	P	2	11/12	0.99	0.22	36,39,42,43	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.

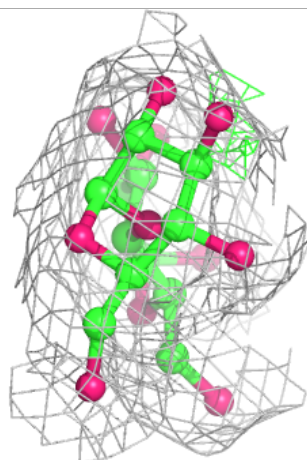
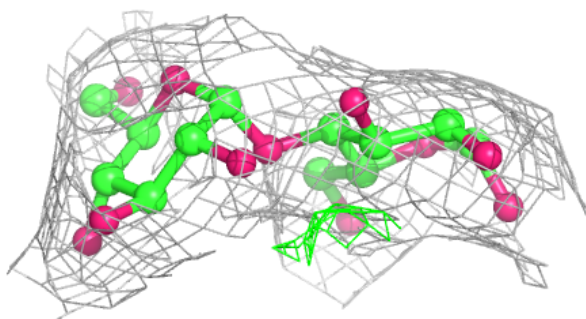
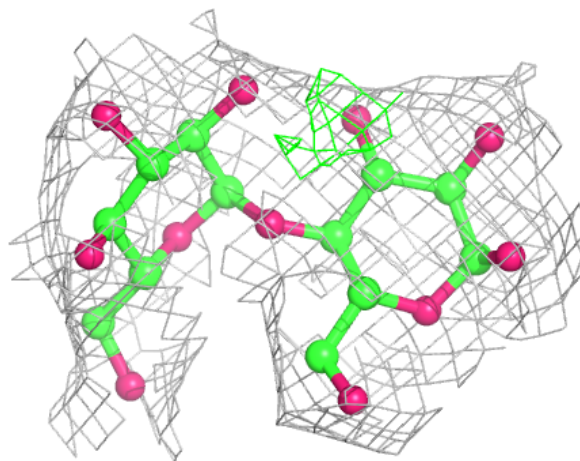
**Electron density around Chain K:**

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



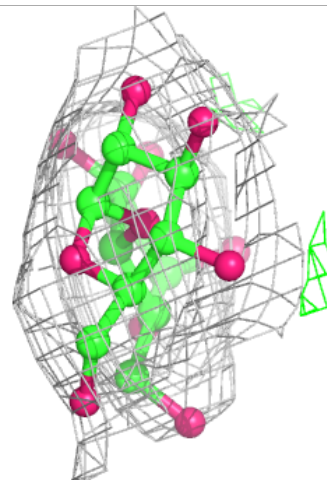
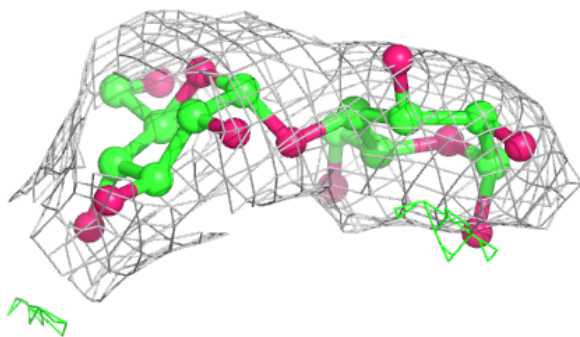
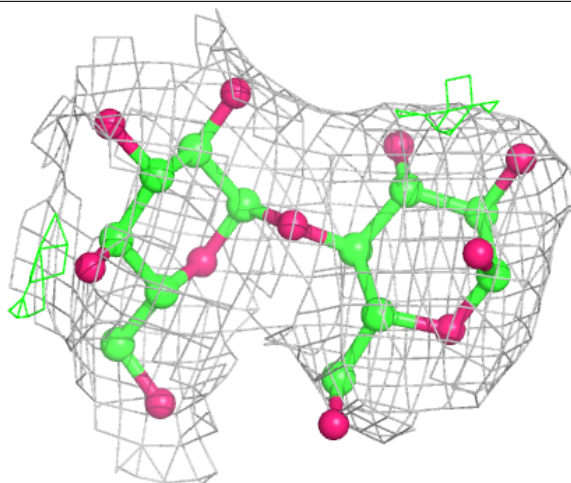
**Electron density around Chain L:**

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



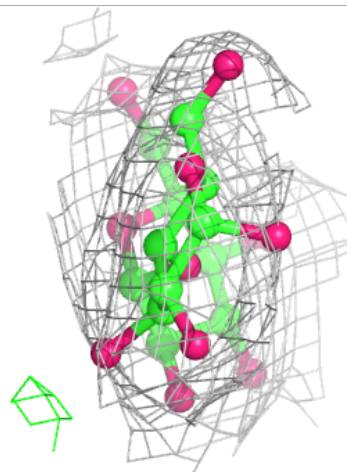
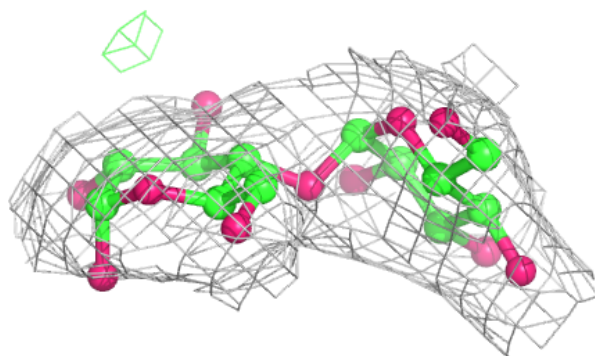
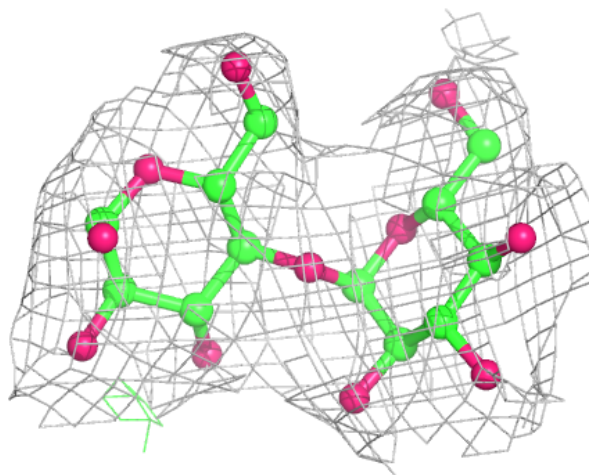
**Electron density around Chain M:**

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



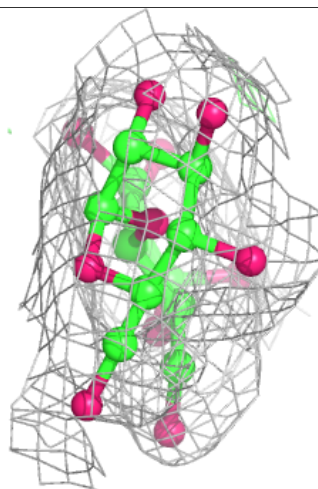
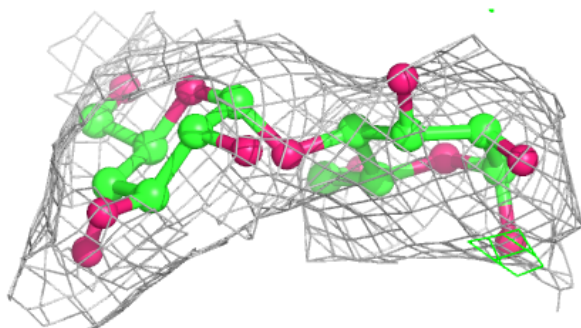
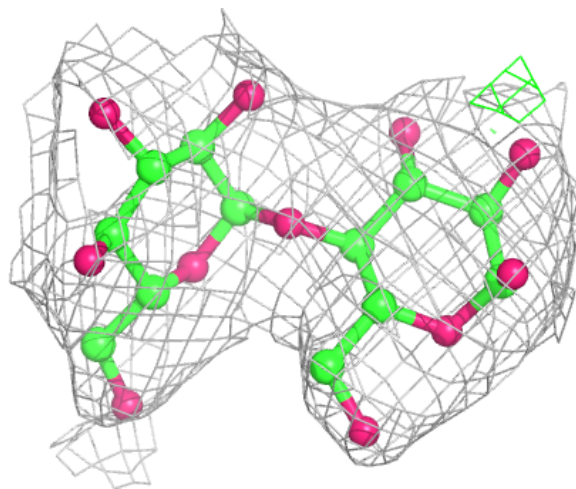
**Electron density around Chain N:**

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



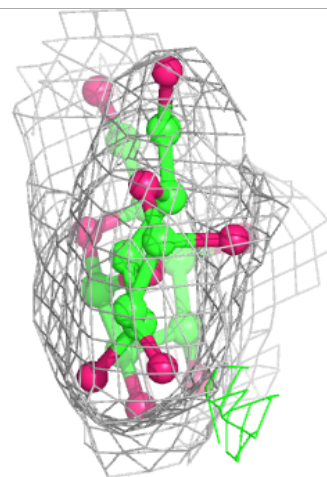
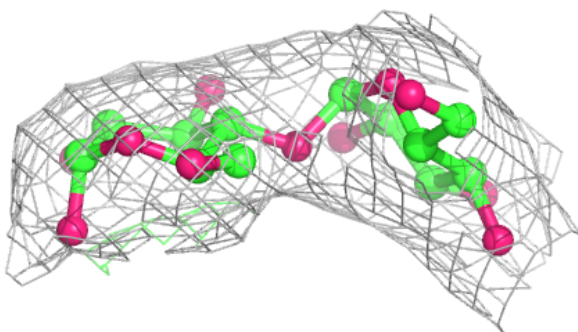
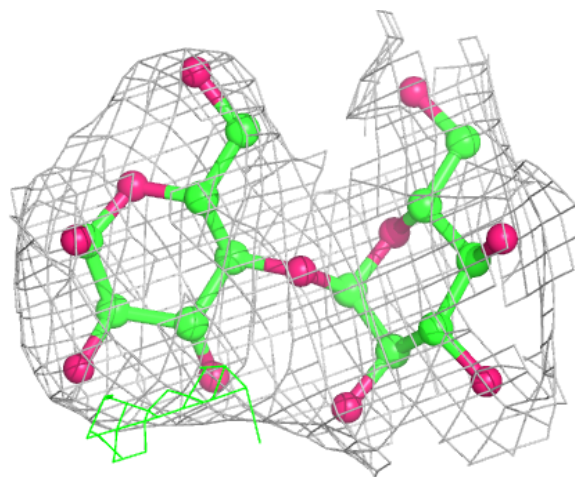
**Electron density around Chain O:**

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



**Electron density around Chain P:**

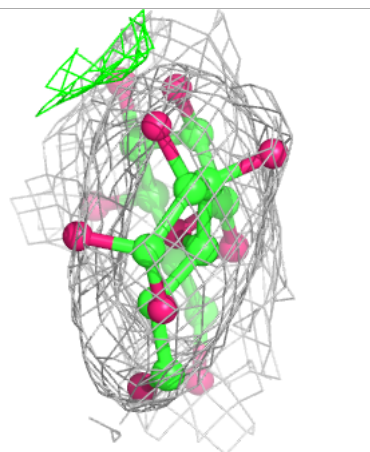
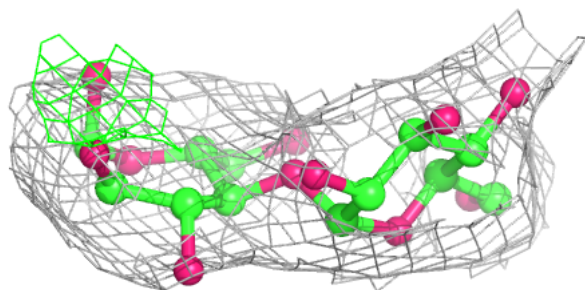
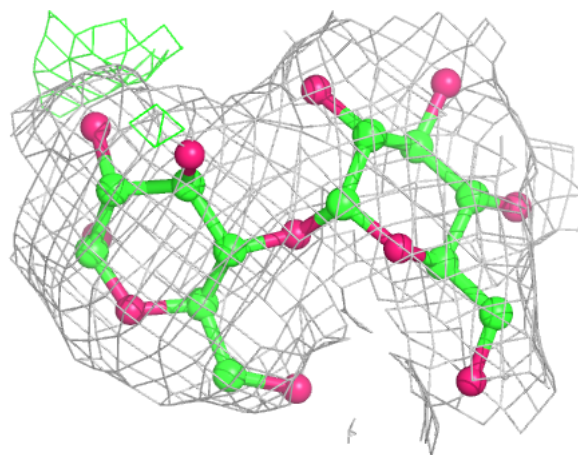
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





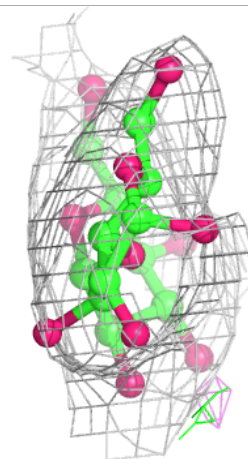
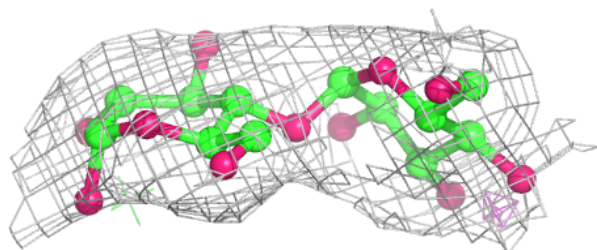
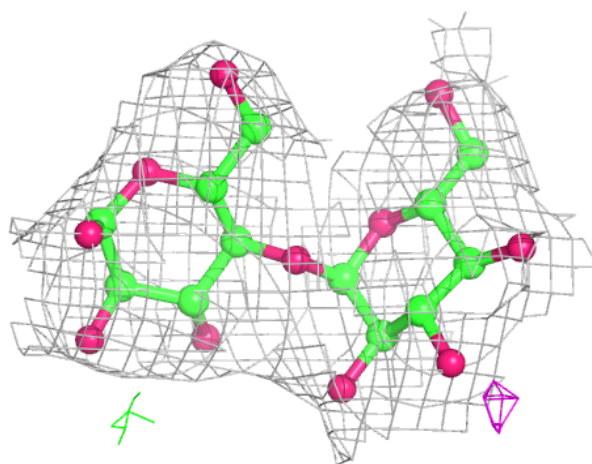
**Electron density around Chain Q:**

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



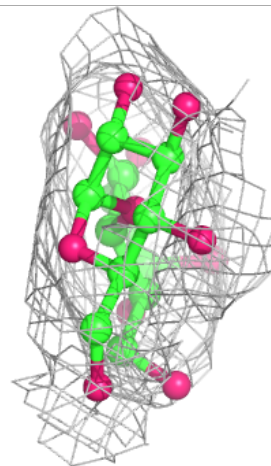
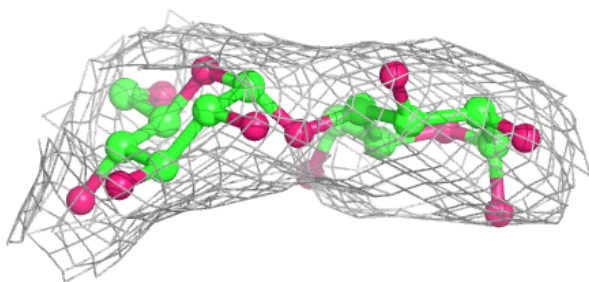
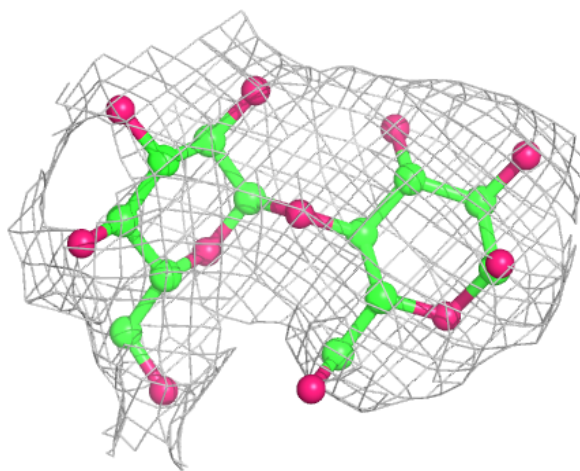
**Electron density around Chain R:**

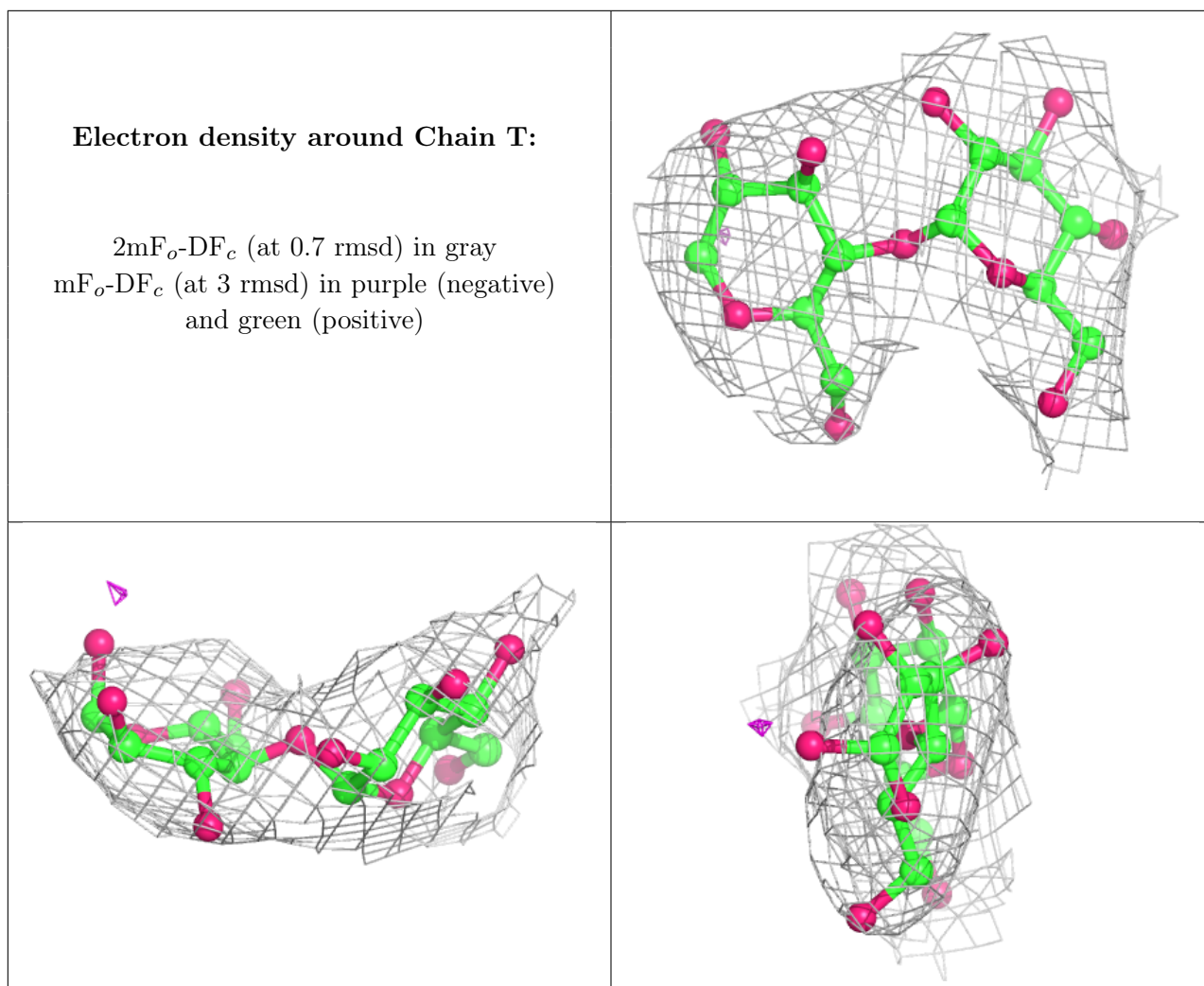
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around Chain S:**

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





## 6.4 Ligands [i](#)

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	NA	E	1502	1/1	0.92	0.18	59,59,59,59	0
3	NA	B	1502	1/1	0.96	0.10	51,51,51,51	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.