



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 6, 2020 – 07:28 PM BST

PDB ID : 6AZP  
Title : A Structurally Dynamic N-terminal Region Drives Function of the Staphylococcal Peroxidase Inhibitor (SPIN)  
Authors : Ramyar, K.X.; Geisbrecht, B.V.  
Deposited on : 2017-09-11  
Resolution : 2.29 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 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

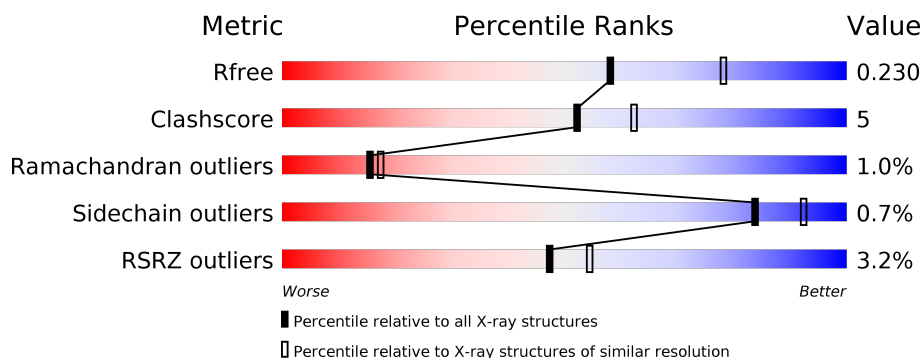
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 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 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	577	<div> <div>3%</div> <div> <div></div> <div>90%</div> <div>9%</div> <div></div> </div> </div>
2	B	60	<div> <div>5%</div> <div> <div></div> <div>85%</div> <div>8%</div> <div>5%</div> </div> </div>
3	C	2	<div> <div></div> <div> <div>50%</div> <div>50%</div> </div> </div>

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 5432 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Myeloperoxidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	577	Total	C	N	O	S	0	0	0
			4620	2912	845	831	32			

- Molecule 2 is a protein called Staphylococcal Peroxidase Inhibitor.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	57	Total	C	N	O	0	0	0
			462	292	77	93			

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



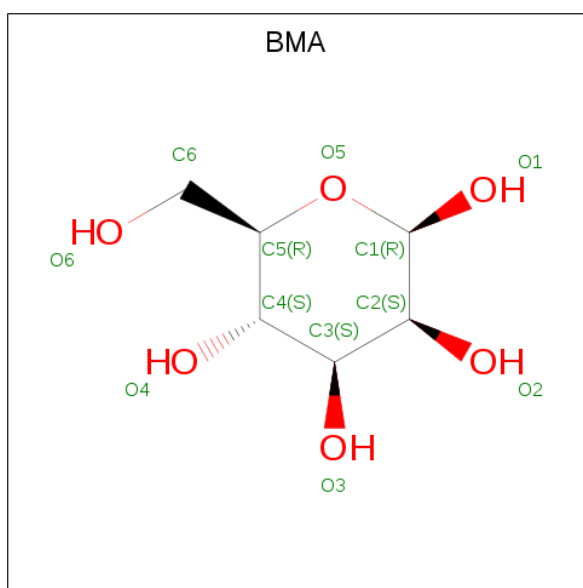
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



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

- Molecule 5 is beta-D-mannopyranose (three-letter code: BMA) (formula:  $C_6H_{12}O_6$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			11	6	5		

- Molecule 6 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

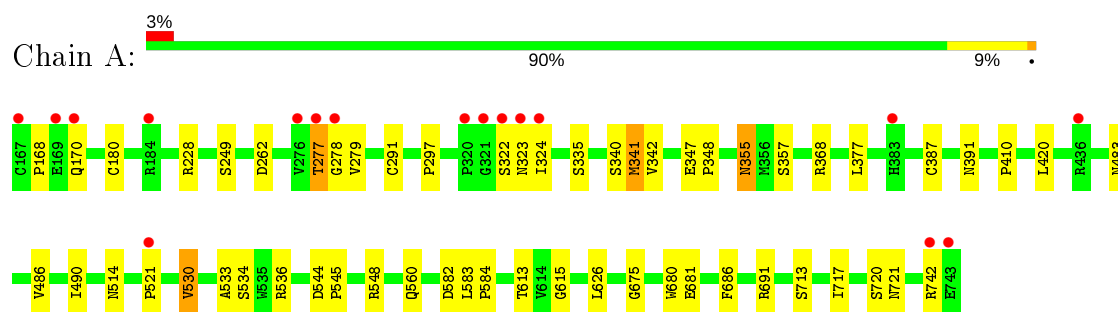
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	269	Total 269	O 269	0	0
7	B	13	Total 13	O 13	0	0

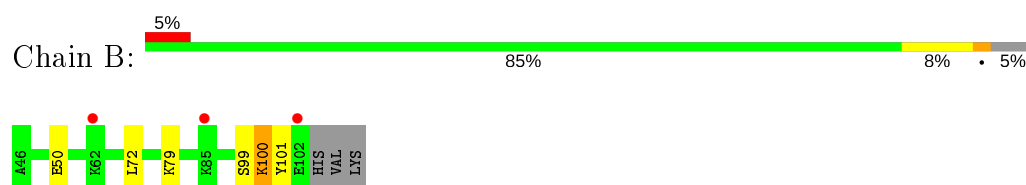
### 3 Residue-property plots

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: Myeloperoxidase



- Molecule 2: Staphylococcal Peroxidase Inhibitor



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



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	128.85Å 92.88Å 80.46Å 90.00° 119.91° 90.00°	Depositor
Resolution (Å)	38.96 – 2.29 39.81 – 2.29	Depositor EDS
% Data completeness (in resolution range)	98.5 (38.96-2.29) 91.0 (39.81-2.29)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.47 (at 2.29Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.179 , 0.229 0.185 , 0.230	Depositor DCC
$R_{free}$ test set	2000 reflections (5.48%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	30.7	Xtriage
Anisotropy	0.635	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 38.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	5432	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.94% 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, BMA, CA, NAG

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	0.54	0/4725	0.69	2/6415 (0.0%)
2	B	0.46	0/469	0.58	0/629
All	All	0.53	0/5194	0.68	2/7044 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	341	MET	CB-CG-SD	5.38	128.53	112.40
1	A	228	ARG	NE-CZ-NH1	5.14	122.87	120.30

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	4620	0	4576	52	0
2	B	462	0	447	6	0
3	C	28	0	24	6	0
4	A	28	0	26	9	0
5	A	11	0	10	0	0
6	A	1	0	0	0	0
7	A	269	0	0	5	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	B	13	0	0	2	0
All	All	5432	0	5083	56	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 5.

The worst 5 of 56 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:A:391:ASN:HD21	4:A:805:NAG:C1	1.00	1.59
1:A:483:ASN:HD21	4:A:801:NAG:C1	1.11	1.55
1:A:391:ASN:ND2	4:A:805:NAG:C1	1.81	1.39
1:A:483:ASN:ND2	4:A:801:NAG:C1	1.96	1.28
1:A:355:ASN:OD1	3:C:1:NAG:C1	1.86	1.22

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	574/577 (100%)	543 (95%)	26 (4%)	5 (1%)	17	20
2	B	55/60 (92%)	52 (94%)	2 (4%)	1 (2%)	8	7
All	All	629/637 (99%)	595 (95%)	28 (4%)	6 (1%)	15	17

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	277	THR
1	A	514	ASN
2	B	100	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	323	ASN
1	A	521	PRO

### 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	505/505 (100%)	501 (99%)	4 (1%)	81	91
2	B	50/53 (94%)	50 (100%)	0	100	100
All	All	555/558 (100%)	551 (99%)	4 (1%)	84	92

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	297	PRO
1	A	340	SER
1	A	355	ASN
1	A	530	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	391	ASN
1	A	483	ASN
2	B	71	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	A	316	1	3,6,7	0.40	0	0,6,8	0.00	-

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	A	316	1	-	0/1/5/7	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates

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	C	1	3	14,14,15	0.29	0	17,19,21	0.63	0
3	NAG	C	2	3	14,14,15	0.58	0	17,19,21	0.42	0

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	C	1	3	-	2/6/23/26	0/1/1/1
3	NAG	C	2	3	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

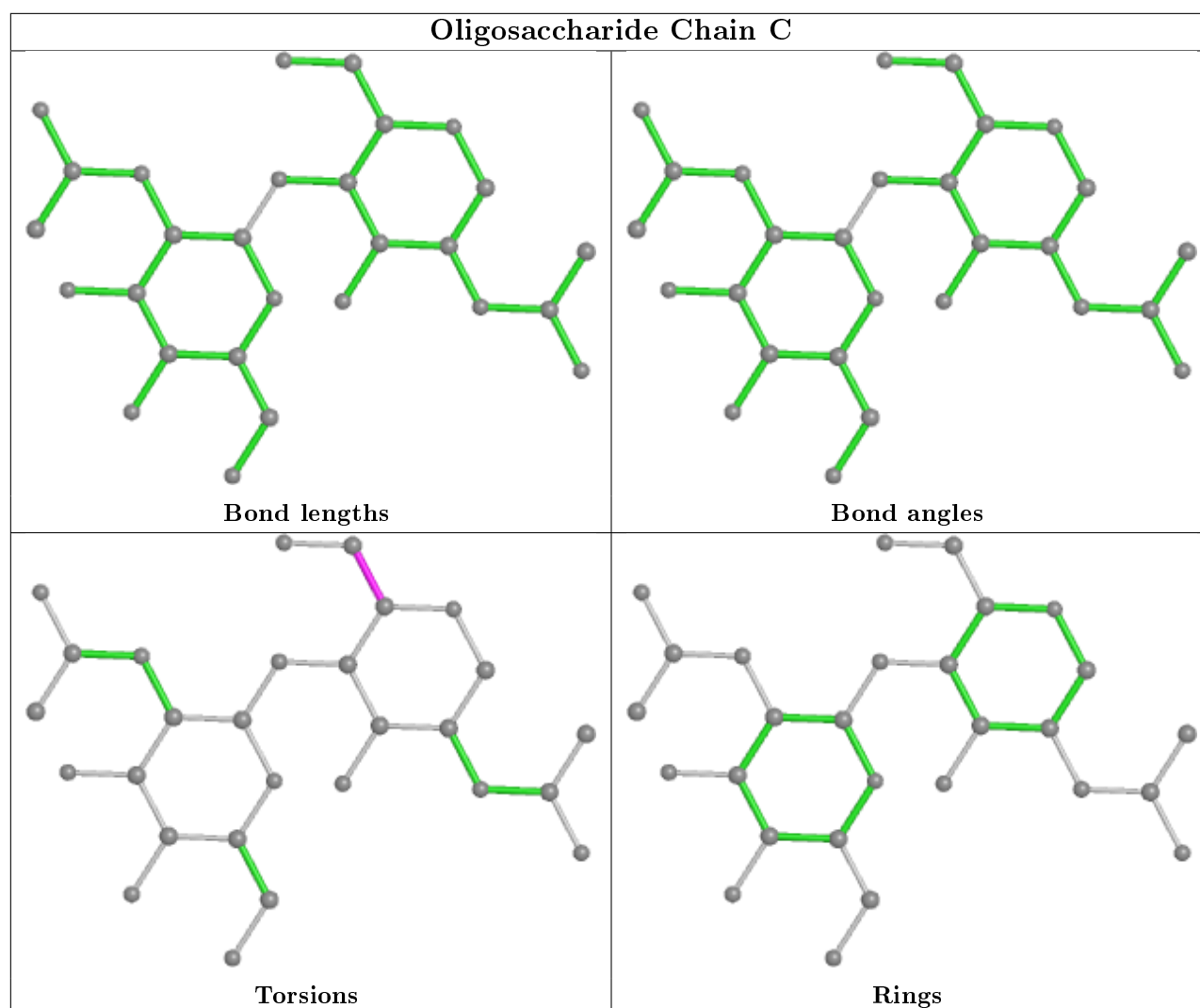
Mol	Chain	Res	Type	Atoms
3	C	1	NAG	C4-C5-C6-O6
3	C	1	NAG	O5-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	1	NAG	6	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 4 ligands modelled in this entry, 1 is 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	805	-	14,14,15	0.42	0	17,19,21	0.58	0
5	BMA	A	804	-	11,11,12	1.98	5 (45%)	15,15,17	1.09	2 (13%)
4	NAG	A	801	-	14,14,15	0.34	0	17,19,21	0.89	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	805	-	-	2/6/23/26	0/1/1/1
5	BMA	A	804	-	-	0/2/19/22	0/1/1/1
4	NAG	A	801	-	-	2/6/23/26	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	804	BMA	C2-C3	3.78	1.58	1.52
5	A	804	BMA	C4-C3	2.53	1.58	1.52
5	A	804	BMA	O3-C3	2.50	1.48	1.43
5	A	804	BMA	C1-C2	2.50	1.57	1.52
5	A	804	BMA	C4-C5	2.37	1.58	1.53

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	801	NAG	C1-O5-C5	2.92	116.15	112.19
5	A	804	BMA	C1-O5-C5	2.13	115.08	112.19
5	A	804	BMA	O5-C1-C2	2.01	113.88	110.77

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	805	NAG	O5-C5-C6-O6
4	A	805	NAG	C4-C5-C6-O6
4	A	801	NAG	C4-C5-C6-O6
4	A	801	NAG	O5-C5-C6-O6

There are no ring outliers.

2 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	805	NAG	5	0
4	A	801	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 [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	576/577 (99%)	-0.19	17 (2%) 50 57	26, 37, 68, 118	0
2	B	57/60 (95%)	-0.15	3 (5%) 26 33	36, 49, 74, 90	0
All	All	633/637 (99%)	-0.18	20 (3%) 47 54	26, 38, 70, 118	0

The worst 5 of 20 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	321	GLY	11.0
1	A	322	SER	10.7
1	A	324	ILE	9.9
1	A	323	ASN	6.2
1	A	276	VAL	6.2

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

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

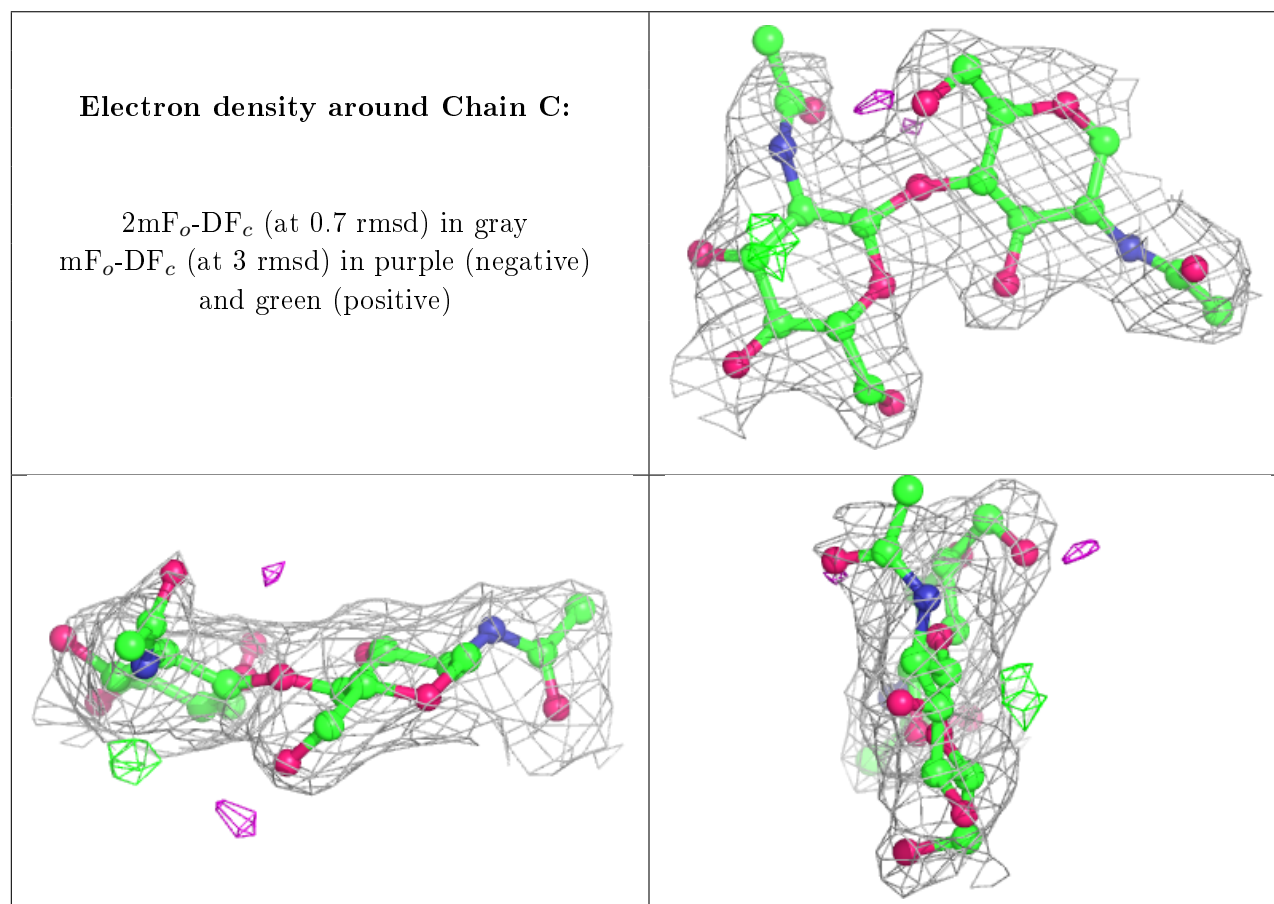
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
1	CSO	A	316	7/8	0.98	0.09	29,34,39,42	0

### 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	C	2	14/15	0.90	0.17	56,69,75,77	0
3	NAG	C	1	14/15	0.95	0.08	40,46,53,54	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( $\text{\AA}^2$ )	Q<0.9
5	BMA	A	804	11/12	0.75	0.19	65,71,82,83	0
4	NAG	A	801	14/15	0.77	0.19	62,75,83,86	0
4	NAG	A	805	14/15	0.92	0.16	33,53,64,65	0
6	CA	A	806	1/1	0.99	0.08	32,32,32,32	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.