

# wwPDB X-ray Structure Validation Summary Report (i)

Dec 10, 2023 - 09:52 am GMT

PDB ID	:	2WN3
Title	:	Crystal structure of Discoidin I from Dictyostelium discoideum in complex
		with the disaccharide GalNAc beta 1-3 galactose, at 1.6 A resolution.
Authors	:	Mathieu, S.; Imberty, A.; Varrot, A.
Deposited on	:	2009-07-07
Resolution	:	1.59  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 A user guide is available at https://www.wwpdb.org/validation/2017/XrayValidationReportHelp with specific help available everywhere you see the (i) 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 (1)) were used in the production of this report:

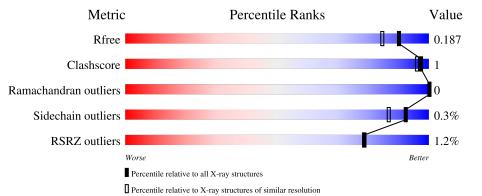
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as $541$ be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

# 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:  $X\text{-}RAY \, DIFFRACTION$ 

The reported resolution of this entry is 1.59 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	$\begin{array}{c} \textbf{Whole archive} \\ (\#\textbf{Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$
$R_{free}$	130704	3398 (1.60-1.60)
Clashscore	141614	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)
RSRZ outliers	127900	3321 (1.60-1.60)

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 >=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 <=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	А	254	% 
1	В	254	96% ·
1	С	254	98% ···
2	D	2	50% 50%
2	Е	2	100%

Continued on next page...



Continued from previous page...

Mol	Chain	Length	Quality of chain
2	$\mathbf{F}$	2	100%



# 2 Entry composition (i)

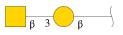
There are 7 unique types of molecules in this entry. The entry contains 7281 atoms, of which 0 are hydrogens and 0 are deuteriums.

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.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
1	Δ	254	Total	С	Ν	0	S	0	5	0
	A	204	2019	1267	353	389	10	0		
1	Р	253	Total	С	Ν	0	S	0	4	0
	D	200	2008	1258	351	389	10			
1	C	252	Total	С	Ν	0	S	0	1	0
	I C	202	2002	1254	351	388	9	0	4	

• Molecule 1 is a protein called DISCOIDIN-1 SUBUNIT A.

• Molecule 2 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-galactopyranose-(1-3)-b eta-D-galactopyranose.



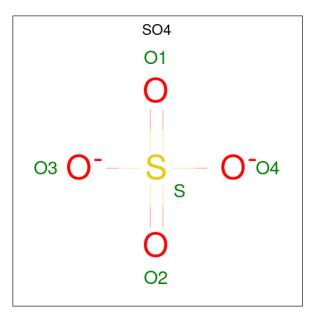
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
2	D	2	Total         C         N         O           26         14         1         11	0	0	0
2	Е	2	Total         C         N         O           26         14         1         11	0	0	0
2	F	2	Total         C         N         O           26         14         1         11	0	0	0

• Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	А	1	Total Ca 1 1	0	0
3	В	1	Total Ca 1 1	0	0
3	С	1	Total Ca 1 1	0	0

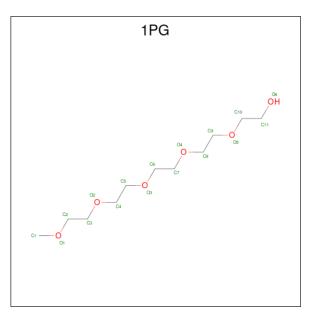


• Molecule 4 is SULFATE ION (three-letter code: SO4) (formula:  $O_4S$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	А	1	$\begin{array}{ccc} \text{Total} & \text{O} & \text{S} \\ 5 & 4 & 1 \end{array}$	0	0
4	А	1	$\begin{array}{ccc} \text{Total} & \text{O} & \text{S} \\ 5 & 4 & 1 \end{array}$	0	0
4	В	1	$\begin{array}{ccc} \text{Total} & \text{O} & \text{S} \\ 5 & 4 & 1 \end{array}$	0	0

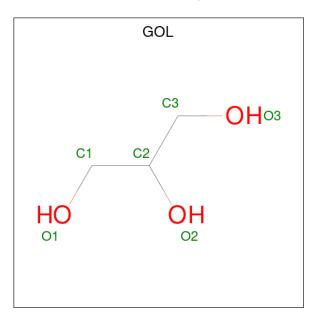
• Molecule 5 is 2-(2-{2-[2-(2-METHOXY-ETHOXY)-ETHOXY]-ETHOXY}-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)-ETHOXY)





Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	А	1	Total 17	C 11	O 6	0	0

• Molecule 6 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	В	1	$\begin{array}{ccc} \text{Total} & \text{C} & \text{O} \\ 6 & 3 & 3 \end{array}$	0	0
6	С	1	$\begin{array}{ccc} \text{Total} & \text{C} & \text{O} \\ 6 & 3 & 3 \end{array}$	0	0
6	С	1	$\begin{array}{ccc} \text{Total} & \text{C} & \text{O} \\ 6 & 3 & 3 \end{array}$	0	0

• Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	А	409	Total O 409 409	0	0
7	В	369	Total O 369 369	0	0
7	С	343	Total O 343 343	0	0



# 3 Residue-property plots (i)

• Molecule 1: DISCOIDIN-1 SUBUNIT A

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.

Chain E:	100%
GAL1 NGA2	
• Molecule	2: 2-acetamido-2-deoxy-beta-D-galactopyranose-(1-3)-beta-D-galactopyranose
Chain F:	100%



GAL1 NGA2



# 4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants	72.58Å 274.00Å 85.76Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $90.00^{\circ}$ $90.00^{\circ}$	Depositor
Resolution (Å)	137.00 - 1.59	Depositor
Resolution (A)	43.73 - 1.59	EDS
% Data completeness	99.0 (137.00-1.59)	Depositor
(in resolution range)	99.0(43.73-1.59)	EDS
R <sub>merge</sub>	0.06	Depositor
R <sub>sym</sub>	(Not available)	Depositor
$< I/\sigma(I) > 1$	$2.17 (at 1.58 \text{\AA})$	Xtriage
Refinement program	REFMAC 5.5.0072	Depositor
B B.	0.152 , $0.180$	Depositor
$R, R_{free}$	0.161 , $0.187$	DCC
$R_{free}$ test set	5754 reflections $(5.02\%)$	wwPDB-VP
Wilson B-factor $(Å^2)$	13.7	Xtriage
Anisotropy	0.583	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$ , $B_{sol}(Å^2)$	0.35 , $49.6$	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	7281	wwPDB-VP
Average B, all atoms $(Å^2)$	15.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 4.35% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;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.



<sup>&</sup>lt;sup>1</sup>Intensities estimated from amplitudes.

# 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: NGA, GOL, 1PG, CA, GAL, 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	Bo	nd lengths	Bond angles		
MOI	Unam	RMSZ	# Z  > 5	RMSZ	# Z  > 5	
1	А	0.82	0/2079	0.83	0/2834	
1	В	0.80	2/2064~(0.1%)	0.82	1/2814~(0.0%)	
1	С	0.80	0/2058	0.84	2/2806~(0.1%)	
All	All	0.81	2/6201~(0.0%)	0.83	3/8454~(0.0%)	

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	В	4[A]	GLN	CG-CD	-5.31	1.38	1.51
1	В	4[B]	GLN	CG-CD	-5.31	1.38	1.51

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
1	С	83	ASP	CB-CG-OD1	5.77	123.49	118.30
1	В	96	ASP	CB-CG-OD1	5.49	123.24	118.30
1	С	169	ASP	CB-CG-OD1	5.27	123.04	118.30

All (3) bond angle outliers are listed below:

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	А	2019	0	1948	2	0
1	В	2008	0	1933	7	0
1	С	2002	0	1922	1	0
2	D	26	0	24	0	0
2	Е	26	0	24	0	0
2	F	26	0	24	0	0
3	А	1	0	0	0	0
3	В	1	0	0	0	0
3	С	1	0	0	0	0
4	А	10	0	0	0	0
4	В	5	0	0	0	0
5	А	17	0	24	0	0
6	В	6	0	8	0	0
6	С	12	0	16	0	0
7	А	409	0	0	1	0
7	В	369	0	0	4	0
7	С	343	0	0	1	0
All	All	7281	0	5923	10	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 1.

The worst 5 of 10 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:175[A]:THR:HG22	7:B:2274:HOH:O	2.01	0.60
1:B:175[B]:THR:HG23	7:B:2118:HOH:O	2.08	0.53
1:A:1:MET:O	1:A:4[A]:GLN:HG2	2.11	0.50
1:B:175[B]:THR:HG22	1:B:176:GLY:N	2.31	0.46
1:B:162:GLY:HA3	1:B:247:TYR:CE1	2.51	0.45

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.



Mol	Chain	Analysed	Favoured	Allowed	Outliers	Perce	$\mathbf{ntiles}$
1	А	257/254~(101%)	252~(98%)	5(2%)	0	100	100
1	В	255/254~(100%)	251 (98%)	4 (2%)	0	100	100
1	С	254/254~(100%)	249 (98%)	5 (2%)	0	100	100
All	All	766/762~(100%)	752 (98%)	14 (2%)	0	100	100

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

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	Percentil	$\mathbf{les}$
1	А	223/218~(102%)	221~(99%)	2(1%)	78 65	
1	В	222/218~(102%)	221 (100%)	1 (0%)	88 80	
1	С	221/218 (101%)	221 (100%)	0	100 10	0
All	All	666/654~(102%)	663 (100%)	3~(0%)	92 80	

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	А	177[A]	SER
1	А	177[B]	SER
1	В	177	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

#### 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)

6 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).

Mal	Mol Type Chain		vpe Chain Res Link		Bo	Bond lengths			Bond angles		
IVIOI	Type	Chain	nes	nes		Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z >2
2	GAL	D	1	2	$12,\!12,\!12$	0.52	0	$17,\!17,\!17$	0.88	0	
2	NGA	D	2	2	$14,\!14,\!15$	0.46	0	$17,\!19,\!21$	1.38	1 (5%)	
2	GAL	Е	1	2	12,12,12	0.55	0	17,17,17	1.05	1 (5%)	
2	NGA	Е	2	2	14,14,15	0.93	1 (7%)	17,19,21	2.19	7 (41%)	
2	GAL	F	1	2	12,12,12	0.68	0	17,17,17	1.16	2 (11%)	
2	NGA	F	2	2	14,14,15	0.51	0	17,19,21	1.29	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
2	GAL	D	1	2	-	0/2/22/22	0/1/1/1
2	NGA	D	2	2	-	0/6/23/26	0/1/1/1
2	GAL	Е	1	2	-	1/2/22/22	0/1/1/1
2	NGA	Е	2	2	-	0/6/23/26	0/1/1/1
2	GAL	F	1	2	-	0/2/22/22	0/1/1/1
2	NGA	F	2	2	-	0/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	Ε	2	NGA	C8-C7	2.07	1.54	1.50



Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
2	Е	2	NGA	O5-C1-C2	-4.19	104.67	111.29
2	Е	2	NGA	C1-C2-N2	3.91	117.17	110.49
2	Е	2	NGA	C4-C3-C2	2.97	115.38	111.02
2	Е	2	NGA	C8-C7-N2	-2.93	111.14	116.10
2	Е	1	GAL	C3-C4-C5	-2.83	105.19	110.24

The worst 5 of 12 bond angle outliers are listed below:

There are no chirality outliers.

All (1) torsion outliers are listed below:

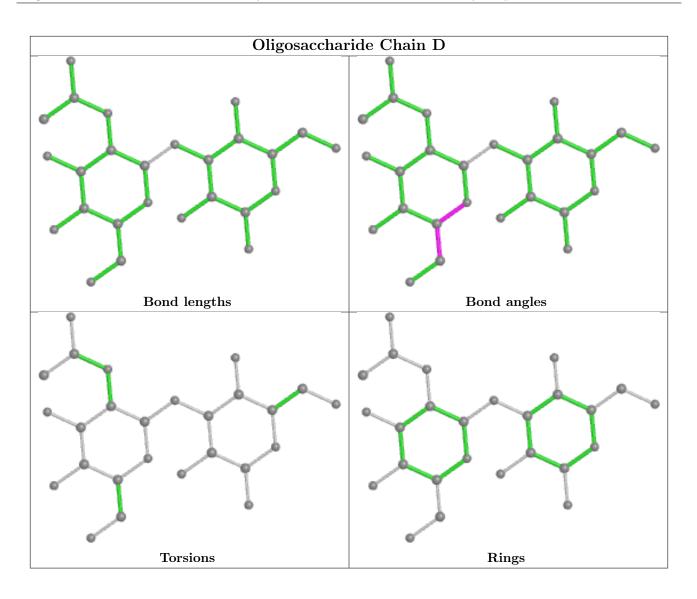
Mol	Chain	Res	Type	Atoms
2	Ε	1	GAL	O5-C5-C6-O6

There are no ring outliers.

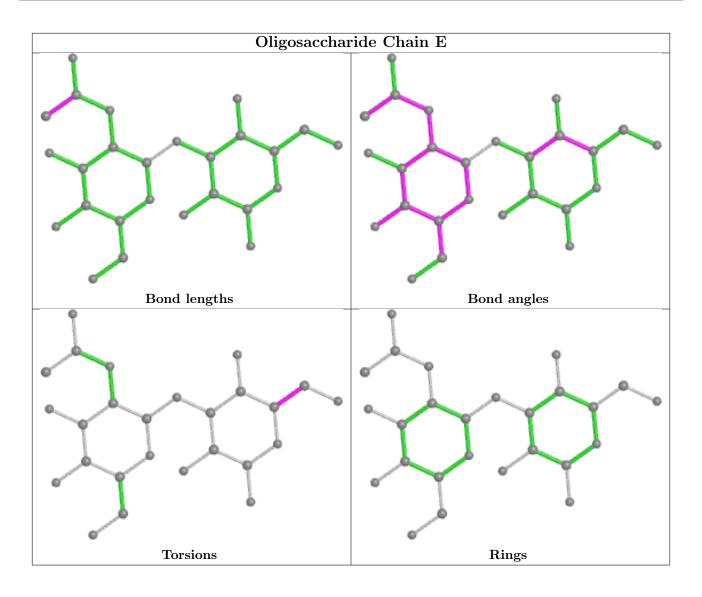
No monomer is involved in short contacts.

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

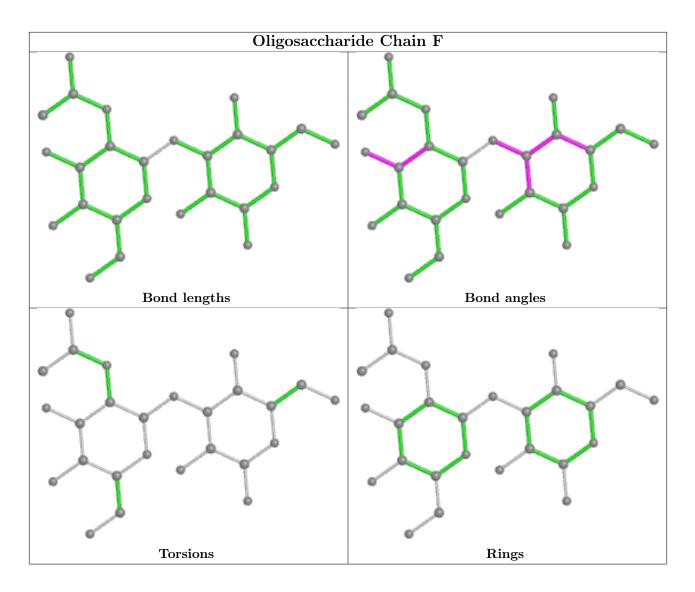












## 5.6 Ligand geometry (i)

Of 10 ligands modelled in this entry, 3 are monoatomic - leaving 7 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).

Mal	Mol Type Chain		Res	Link	Bond lengths			Bond angles		
	туре	Chain	nes	LIIIK	Counts	RMSZ	# Z >2	Counts	RMSZ	# Z  > 2
4	SO4	А	401	-	4,4,4	0.24	0	6,6,6	0.58	0
4	SO4	В	402	-	4,4,4	0.20	0	6,6,6	0.30	0
6	GOL	В	500	-	$5,\!5,\!5$	0.55	0	$5,\!5,\!5$	0.69	0
6	GOL	С	501	-	$5,\!5,\!5$	0.44	0	$5,\!5,\!5$	0.91	0



Mol	Type	Chain	Res	Link	Bo	ond leng	$\mathbf{ths}$	Bond angles		
IVIOI	туре	Unam	nes	LIIIK	Counts	RMSZ	# Z >2	Counts	RMSZ	# Z >2
4	SO4	А	403	-	4,4,4	0.28	0	$6,\!6,\!6$	0.64	0
5	1PG	А	500	-	16, 16, 16	0.46	0	$15,\!15,\!15$	0.72	0
6	GOL	С	500	-	$5,\!5,\!5$	0.22	0	$5,\!5,\!5$	0.76	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
6	GOL	С	501	-	-	0/4/4/4	-
6	GOL	В	500	-	-	2/4/4/4	-
5	1PG	А	500	-	-	1/14/14/14	-
6	GOL	С	500	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (5) torsion outliers are listed below:

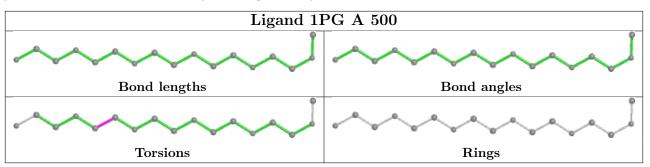
Mol	Chain	Res	Type	Atoms
6	В	500	GOL	C1-C2-C3-O3
6	С	500	GOL	O2-C2-C3-O3
6	С	500	GOL	C1-C2-C3-O3
6	В	500	GOL	O2-C2-C3-O3
5	А	500	1PG	O4-C8-C9-O5

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





equivalents in the CSD to analyse the geometry.

### 5.7 Other polymers (i)

There are no such residues in this entry.

#### 5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.



# 6 Fit of model and data (i)

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

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median,  $95^{th}$  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	$\langle RSRZ \rangle$	# RSRZ > 2	$OWAB(Å^2)$	Q<0.9
1	А	254/254~(100%)	-0.17	2 (0%) 86 86	8, 11, 18, 29	1 (0%)
1	В	253/254~(99%)	-0.14	5 (1%) 65 64	8, 12, 19, 35	3 (1%)
1	С	252/254~(99%)	-0.14	2 (0%) 86 86	8, 13, 21, 29	4 (1%)
All	All	759/762~(99%)	-0.15	9 (1%) 79 78	8, 12, 19, 35	8 (1%)

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	В	1	MET	3.4
1	С	2	SER	3.2
1	А	16[A]	HIS	3.0
1	С	4[A]	GLN	2.8
1	А	2	SER	2.6

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

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

### 6.3 Carbohydrates (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median,  $95^{th}$  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	$\mathbf{B} ext{-factors}(\mathbf{\AA}^2)$	Q<0.9
2	GAL	Е	1	12/12	0.74	0.29	$27,\!34,\!38,\!39$	0
2	GAL	F	1	12/12	0.80	0.27	27,36,40,44	0
2	GAL	D	1	12/12	0.82	0.30	20,30,35,35	0

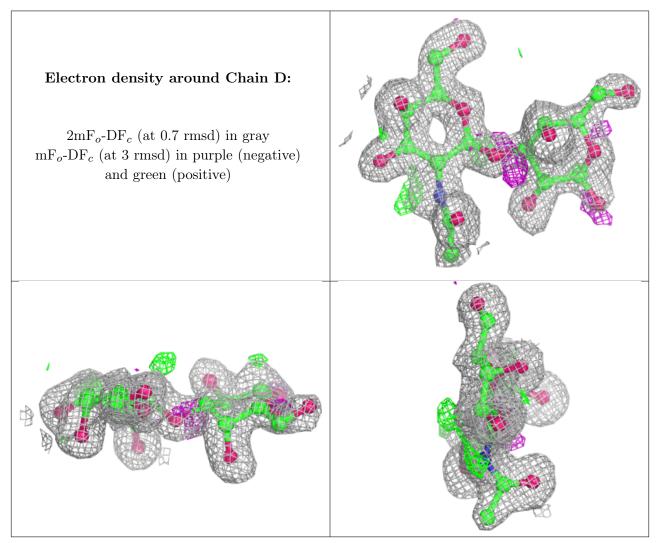
Continued on next page...



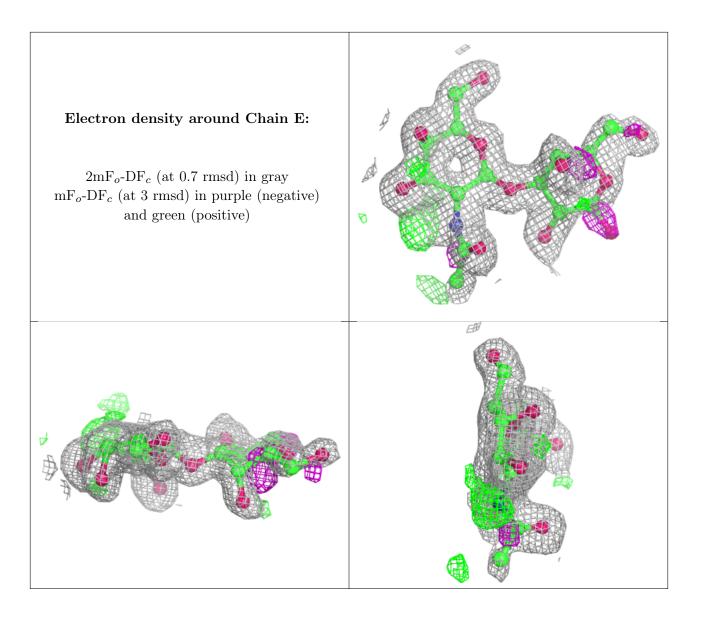
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	$B-factors(Å^2)$	Q < 0.9
2	NGA	Е	2	14/15	0.88	0.14	$19,\!21,\!26,\!28$	0
2	NGA	D	2	14/15	0.90	0.16	15,18,21,24	0
2	NGA	F	2	14/15	0.92	0.13	17,20,23,28	0

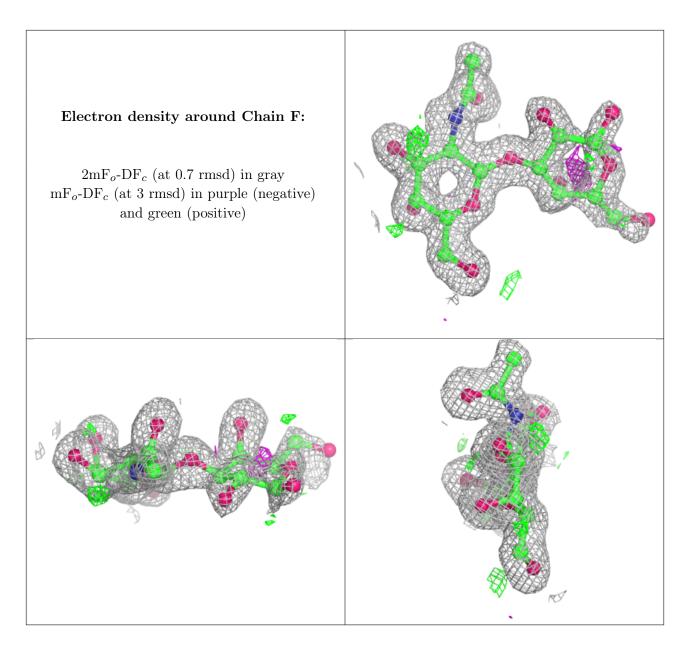
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.











### 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,  $95^{th}$  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	$\mathbf{B} ext{-factors}(\mathbf{\AA}^2)$	Q < 0.9
6	GOL	С	500	6/6	0.82	0.10	17,20,27,27	0
6	GOL	С	501	6/6	0.86	0.12	22,29,33,33	0
4	SO4	А	403	5/5	0.91	0.19	32,32,33,35	0
6	GOL	В	500	6/6	0.93	0.12	24,29,31,34	0
5	1PG	А	500	17/17	0.96	0.08	12,15,18,22	0

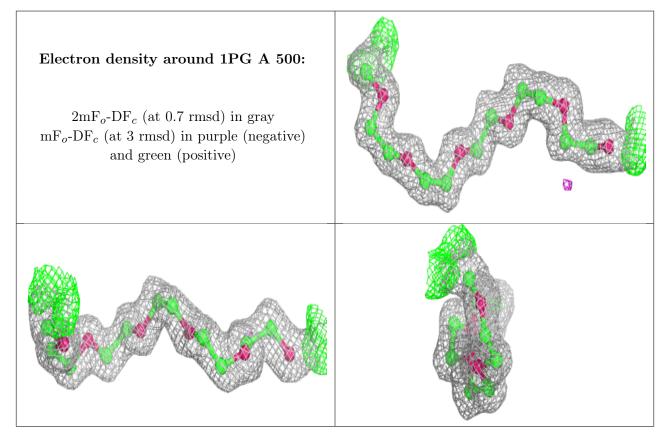
Continued on next page...



e chundada ji chu presto ac pagetti								
Mol	Type	Chain	Res	Atoms	RSCC	RSR	${f B} ext{-factors}({ m \AA}^2)$	Q<0.9
4	SO4	В	402	5/5	0.96	0.17	$28,\!29,\!30,\!30$	5
4	SO4	А	401	5/5	0.97	0.12	$15,\!16,\!17,\!18$	0
3	CA	В	400	1/1	1.00	0.05	8,8,8,8	0
3	CA	С	400	1/1	1.00	0.06	7,7,7,7	0
3	CA	А	400	1/1	1.00	0.07	8,8,8,8	0

Continued from previous page...

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.



### 6.5 Other polymers (i)

There are no such residues in this entry.

