



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 9, 2020 – 07:46 AM BST

PDB ID : 4YH7  
Title : Crystal structure of PTPdelta ectodomain in complex with IL1RAPL1  
Authors : Yamagata, A.; Fukai, S.  
Deposited on : 2015-02-27  
Resolution : 4.40 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

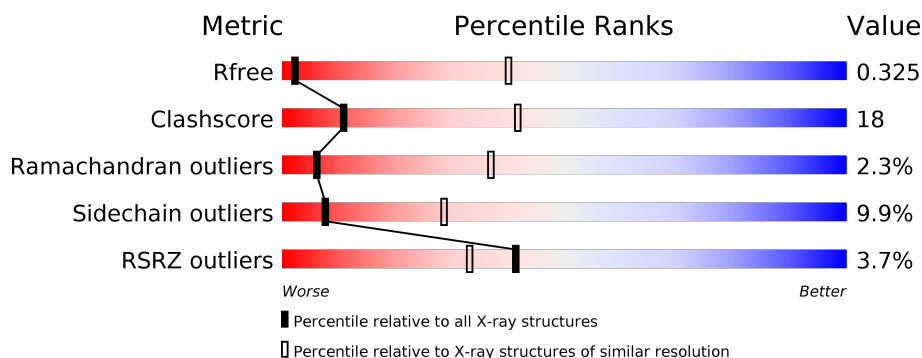
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 4.40 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	1043 (5.00-3.80)
Clashscore	141614	1111 (5.00-3.80)
Ramachandran outliers	138981	1059 (5.00-3.80)
Sidechain outliers	138945	1041 (5.00-3.80)
RSRZ outliers	127900	1095 (5.08-3.70)

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	841	<div> <div>4%</div> <div> <div></div> <div>39%</div> <div>27%</div> <div>•</div> <div>30%</div> </div> </div>
2	B	348	<div> <div>%</div> <div> <div></div> <div>58%</div> <div>28%</div> <div>7%</div> <div>7%</div> </div> </div>
3	C	2	<div> <div></div> <div>100%</div> </div>
4	D	5	<div> <div></div> <div>60%</div> <div>40%</div> </div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 7240 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 Receptor-type tyrosine-protein phosphatase delta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	585	Total	C	N	O	S	0	0	0
			4515	2820	784	894	17			

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	862	LYS	-	expression tag	UNP Q64487
A	863	HIS	-	expression tag	UNP Q64487
A	864	HIS	-	expression tag	UNP Q64487
A	865	HIS	-	expression tag	UNP Q64487
A	866	HIS	-	expression tag	UNP Q64487
A	867	HIS	-	expression tag	UNP Q64487
A	868	HIS	-	expression tag	UNP Q64487

- Molecule 2 is a protein called Interleukin-1 receptor accessory protein-like 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	324	Total	C	N	O	S	0	0	0
			2608	1658	436	498	16			

There are 14 discrepancies between the modelled and reference sequences:

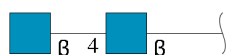
Chain	Residue	Modelled	Actual	Comment	Reference
B	12	ALA	-	expression tag	UNP P59823
B	13	GLN	-	expression tag	UNP P59823
B	14	PRO	-	expression tag	UNP P59823
B	15	ALA	-	expression tag	UNP P59823
B	16	ALA	-	expression tag	UNP P59823
B	17	ARG	-	expression tag	UNP P59823
B	18	ASP	-	expression tag	UNP P59823
B	353	LYS	-	expression tag	UNP P59823
B	354	HIS	-	expression tag	UNP P59823

*Continued on next page...*

*Continued from previous page...*

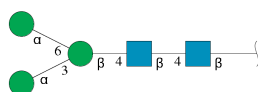
Chain	Residue	Modelled	Actual	Comment	Reference
B	355	HIS	-	expression tag	UNP P59823
B	356	HIS	-	expression tag	UNP P59823
B	357	HIS	-	expression tag	UNP P59823
B	358	HIS	-	expression tag	UNP P59823
B	359	HIS	-	expression tag	UNP P59823

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



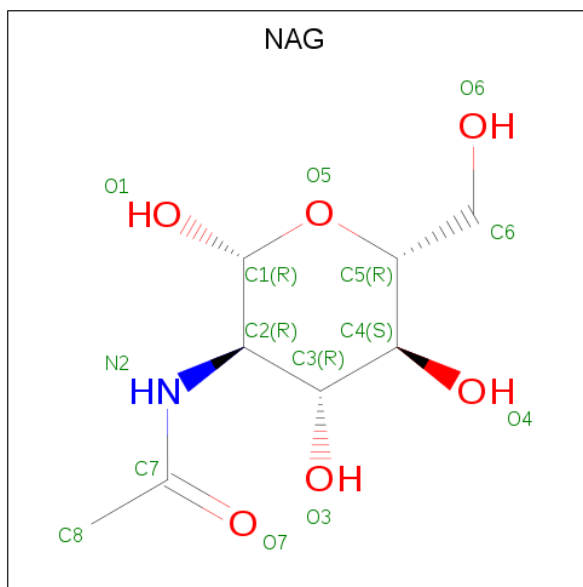
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 4 is an oligosaccharide called alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	D	5	Total	C	N	O	0	0	0
			61	34	2	25			

- Molecule 5 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
5	B	1	Total	C	N	O	0	0
			14	8	1	5		
5	B	1	Total	C	N	O	0	0
			14	8	1	5		





## 4 Data and refinement statistics

Property	Value	Source
Space group	I 41	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	286.30 Å   286.30 Å   70.18 Å 90.00°   90.00°   90.00°	Depositor
Resolution (Å)	39.09 – 4.40 49.36 – 4.40	Depositor EDS
% Data completeness (in resolution range)	95.4 (39.09-4.40) 95.4 (49.36-4.40)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.83 (at 4.45 Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, $R_{free}$	0.290 , 0.325 0.290 , 0.325	Depositor DCC
$R_{free}$ test set	896 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	175.3	Xtriage
Anisotropy	0.101	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.19 , 22.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.35$ , $\langle L^2 \rangle = 0.18$	Xtriage
Estimated twinning fraction	0.149 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.83	EDS
Total number of atoms	7240	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	147.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.41% 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: BMA, NAG, MAN

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	0.36	0/4615	0.78	5/6296 (0.1%)
2	B	0.28	0/2666	0.56	0/3600
All	All	0.33	0/7281	0.71	5/9896 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	B	0	1

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	358	VAL	CG1-CB-CG2	-9.03	96.45	110.90
1	A	518	THR	N-CA-C	6.12	127.53	111.00
1	A	452	ASN	N-CA-C	5.75	126.52	111.00
1	A	302	ARG	N-CA-C	-5.61	95.86	111.00
1	A	474	MET	N-CA-CB	5.48	120.46	110.60

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	88	PRO	Peptide

## 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	4515	0	4454	196	0
2	B	2608	0	2540	72	0
3	C	28	0	25	0	0
4	D	61	0	52	2	0
5	B	28	0	26	1	0
All	All	7240	0	7097	261	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 18.

The worst 5 of 261 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:462:TYR:HE1	1:A:476:HIS:HB2	1.37	0.90
1:A:392:PRO:HB2	1:A:421:GLU:OE2	1.74	0.88
1:A:98:ARG:NH2	2:B:291:GLU:OE2	2.09	0.85
1:A:124:ARG:O	1:A:127:GLN:NE2	2.13	0.82
1:A:424:PRO:HA	1:A:452:ASN:HD22	1.44	0.80

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	583/841 (69%)	523 (90%)	48 (8%)	12 (2%)	7 39

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	320/348 (92%)	274 (86%)	37 (12%)	9 (3%)	5	33
All	All	903/1189 (76%)	797 (88%)	85 (9%)	21 (2%)	6	37

5 of 21 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	70	LYS
1	A	85	SER
1	A	178	ASN
1	A	474	MET
1	A	482	GLN

### 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	508/736 (69%)	472 (93%)	36 (7%)	14	41
2	B	289/309 (94%)	246 (85%)	43 (15%)	3	17
All	All	797/1045 (76%)	718 (90%)	79 (10%)	8	29

5 of 79 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	596	SER
2	B	84	ASP
2	B	347	VAL
2	B	40	LYS
2	B	50	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	452	ASN
2	B	357	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	456	GLN
1	A	323	GLN
2	B	68	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

7 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	3,2	14,14,15	0.37	0	17,19,21	0.44	0
3	NAG	C	2	3	14,14,15	0.26	0	17,19,21	0.45	0
4	NAG	D	1	2,4	14,14,15	0.29	0	17,19,21	0.60	0
4	NAG	D	2	4	14,14,15	0.62	1 (7%)	17,19,21	0.58	0
4	BMA	D	3	4	11,11,12	1.61	3 (27%)	15,15,17	1.72	2 (13%)
4	MAN	D	4	4	11,11,12	1.22	2 (18%)	15,15,17	0.95	1 (6%)
4	MAN	D	5	4	11,11,12	0.99	1 (9%)	15,15,17	1.38	3 (20%)

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	-	0/6/23/26	0/1/1/1
3	NAG	C	2	3	-	2/6/23/26	0/1/1/1
4	NAG	D	1	2,4	-	4/6/23/26	0/1/1/1
4	NAG	D	2	4	-	2/6/23/26	0/1/1/1
4	BMA	D	3	4	-	2/2/19/22	0/1/1/1
4	MAN	D	4	4	-	2/2/19/22	1/1/1/1
4	MAN	D	5	4	-	2/2/19/22	0/1/1/1

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	3	BMA	O3-C3	3.14	1.50	1.43
4	D	5	MAN	C1-C2	2.72	1.58	1.52
4	D	4	MAN	C1-C2	2.53	1.58	1.52
4	D	4	MAN	O5-C5	2.46	1.48	1.43
4	D	3	BMA	C4-C5	2.44	1.58	1.53

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	D	3	BMA	C1-C2-C3	-4.25	104.44	109.67
4	D	3	BMA	O3-C3-C2	3.43	116.56	109.99
4	D	5	MAN	C1-O5-C5	3.36	116.74	112.19
4	D	4	MAN	O2-C2-C3	-2.49	105.14	110.14
4	D	5	MAN	O5-C1-C2	2.05	113.94	110.77

There are no chirality outliers.

5 of 14 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	D	1	NAG	O5-C5-C6-O6
3	C	2	NAG	O5-C5-C6-O6
4	D	1	NAG	C4-C5-C6-O6
4	D	5	MAN	C4-C5-C6-O6
4	D	4	MAN	C4-C5-C6-O6

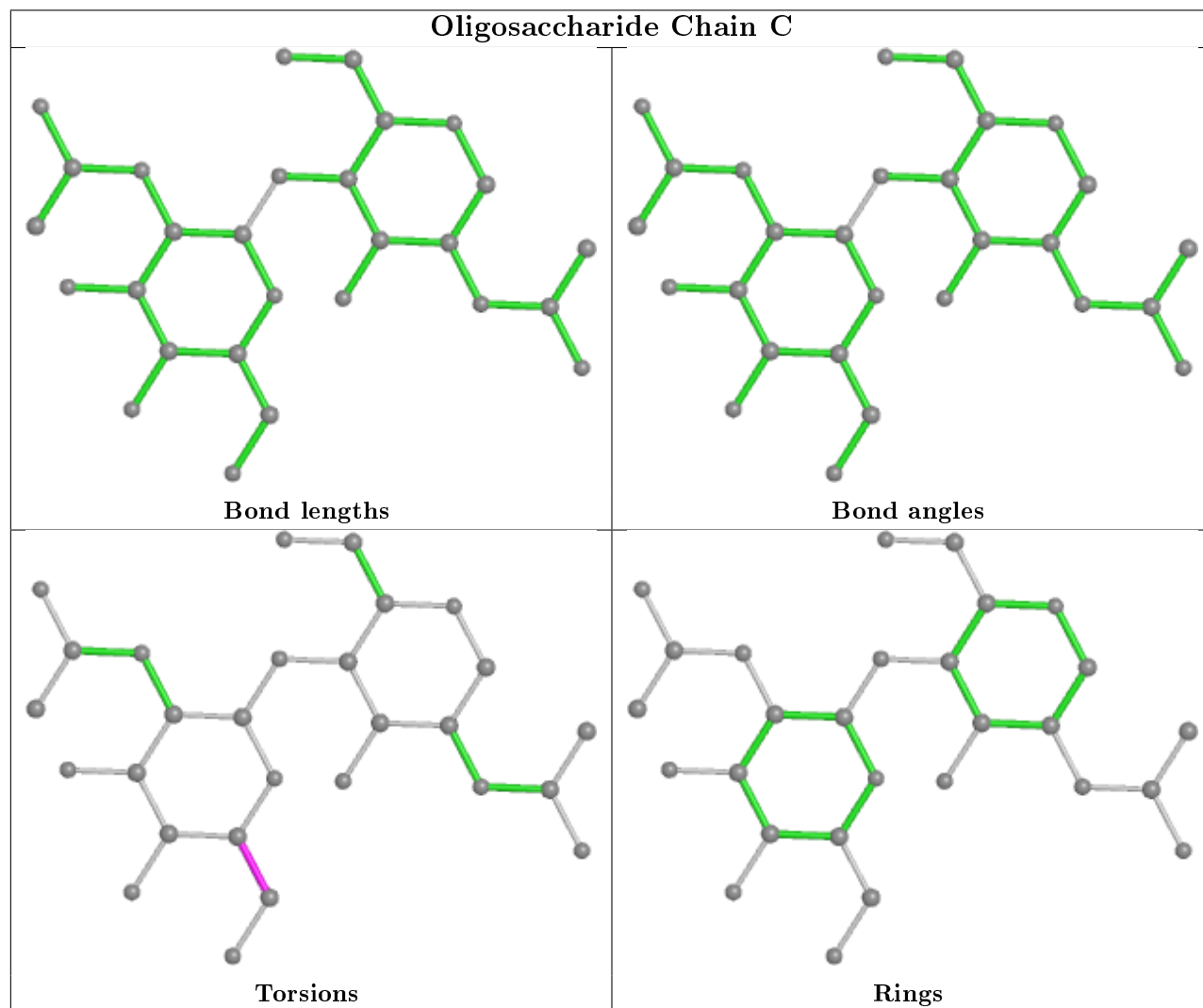
All (1) ring outliers are listed below:

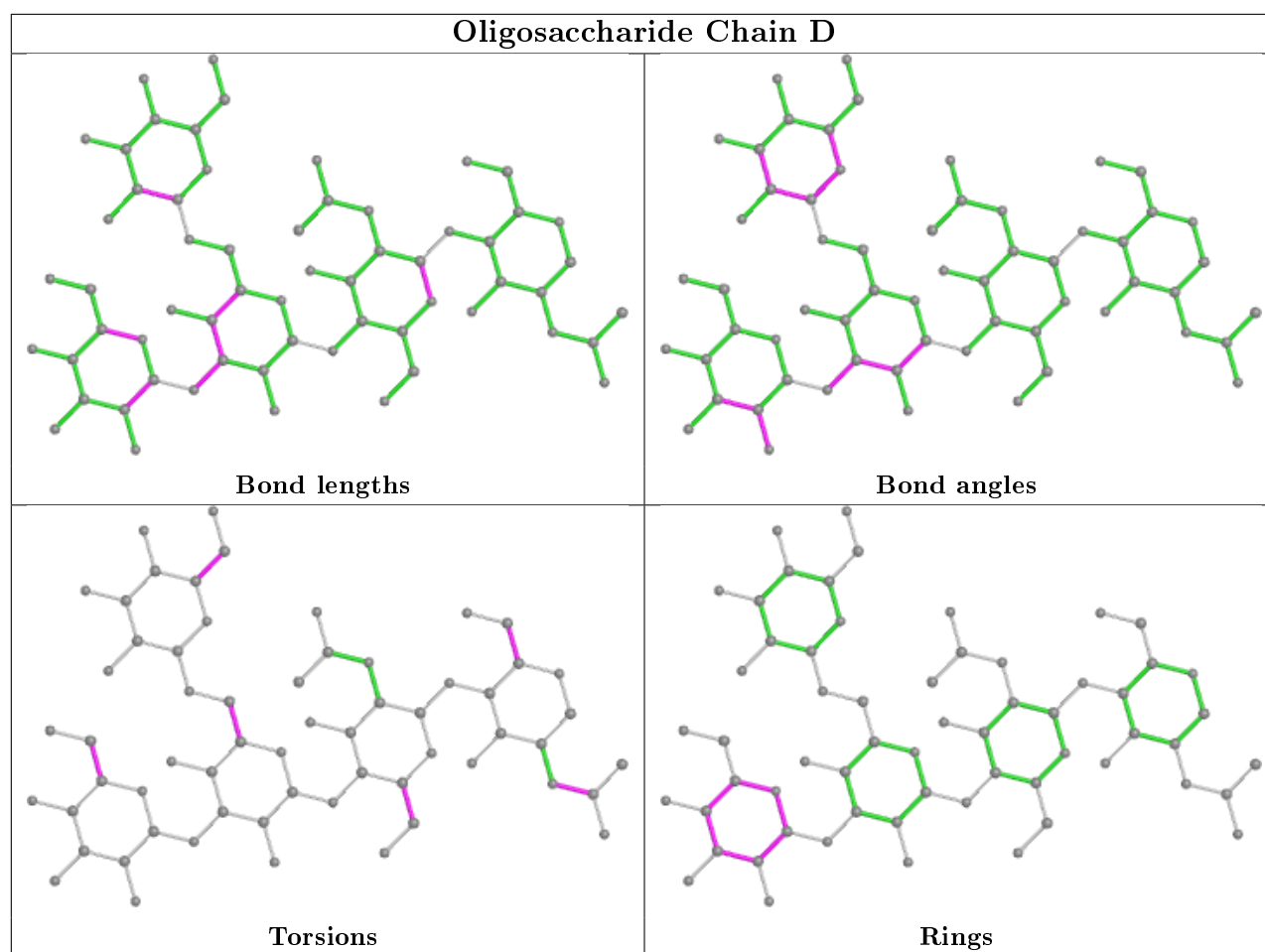
Mol	Chain	Res	Type	Atoms
4	D	4	MAN	C1-C2-C3-C4-C5-O5

3 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	D	3	BMA	1	0
4	D	1	NAG	1	0
4	D	5	MAN	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](#)

2 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
5	NAG	B	401	2	14,14,15	0.71	0	17,19,21	0.57	0
5	NAG	B	402	2	14,14,15	0.44	0	17,19,21	0.40	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
5	NAG	B	401	2	-	3/6/23/26	0/1/1/1
5	NAG	B	402	2	-	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 (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	401	NAG	O5-C5-C6-O6
5	B	401	NAG	C4-C5-C6-O6
5	B	401	NAG	C3-C2-N2-C7

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	401	NAG	1	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.



## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	585/841 (69%)	0.06	30 (5%) 28 24	70, 180, 242, 273	0
2	B	324/348 (93%)	-0.16	4 (1%) 79 70	52, 98, 137, 206	0
All	All	909/1189 (76%)	-0.02	34 (3%) 41 33	52, 137, 234, 273	0

The worst 5 of 34 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	576	SER	5.4
1	A	433	ALA	4.8
1	A	113	GLU	4.2
1	A	501	ALA	3.7
1	A	398	PHE	3.7

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

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.

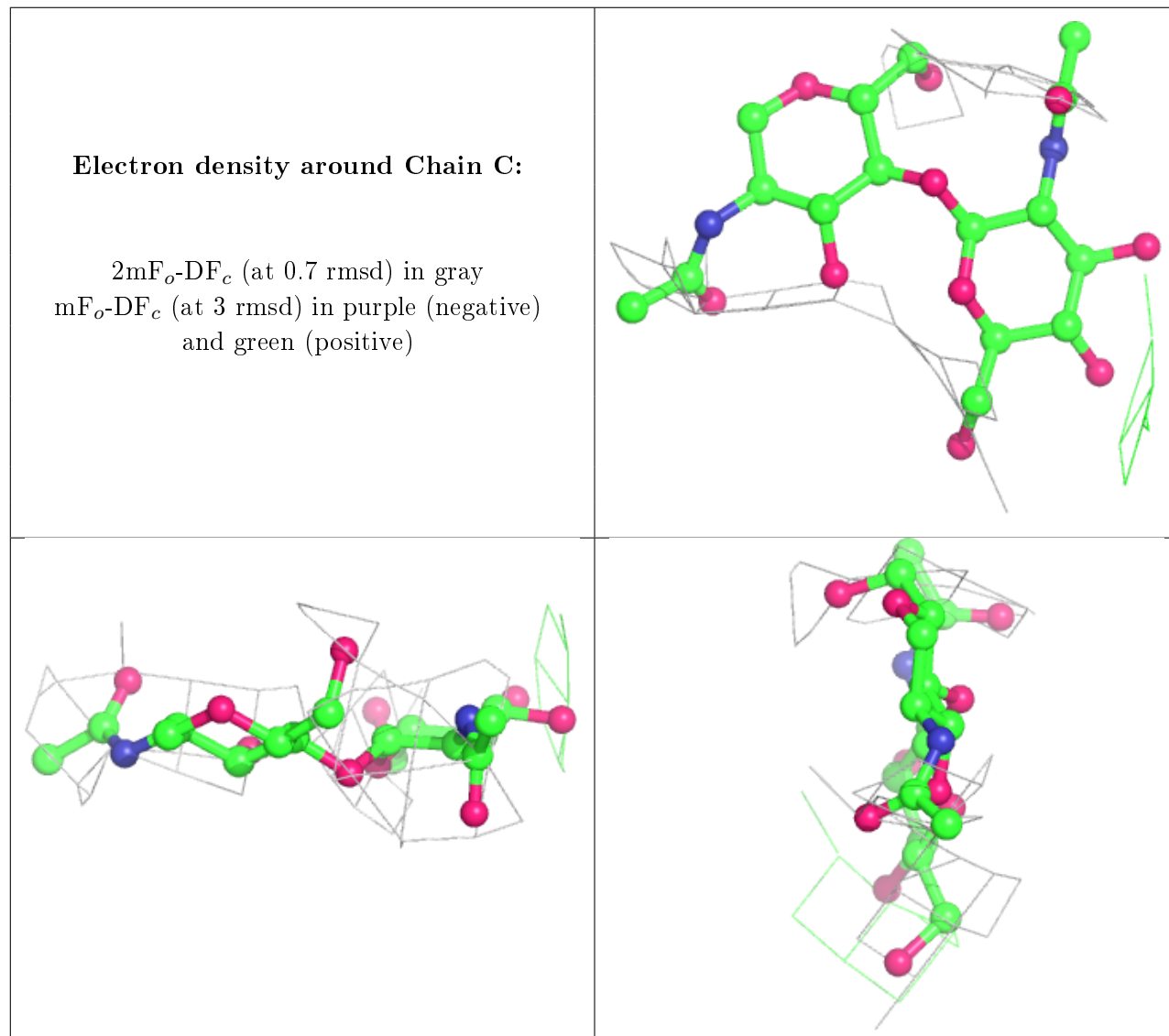
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	MAN	D	4	11/12	0.70	0.18	146,185,203,205	0
3	NAG	C	2	14/15	0.73	0.30	119,138,152,155	0
4	NAG	D	2	14/15	0.81	0.31	118,148,173,176	0
4	BMA	D	3	11/12	0.83	0.07	174,176,182,197	0
4	MAN	D	5	11/12	0.87	0.26	169,185,192,196	0

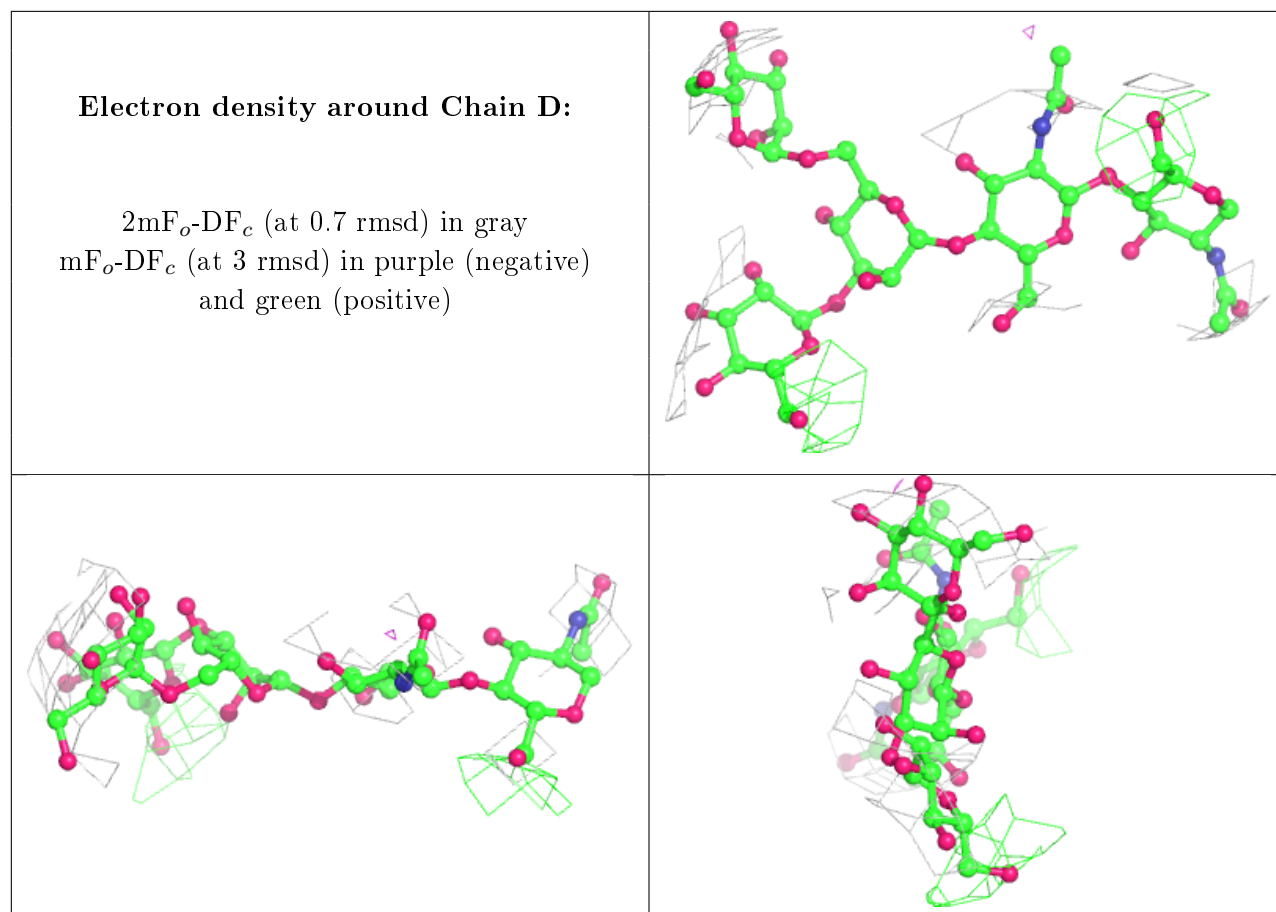
*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	NAG	D	1	14/15	0.87	0.16	70,113,125,126	0
3	NAG	C	1	14/15	0.90	0.25	107,121,136,138	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
5	NAG	B	401	14/15	0.76	0.28	132,151,158,159	0
5	NAG	B	402	14/15	0.91	0.23	109,135,145,148	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.