W O R L D W I D E
wwPDB
P R O T E I N D A T A B A N K

wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 31, 2023 – 04:41 AM EDT

PDB ID	:	1PIG
Title	:	PIG PANCREATIC ALPHA-AMYLASE COMPLEXED WITH THE OLIGOSACCHARIDE V-1532
Authors	:	Machius, M.; Vertesy, L.; Huber, R.; Wiegand, G.
Deposited on	:	1996-06-15
Resolution	:	2.20 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 ⓘ 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:

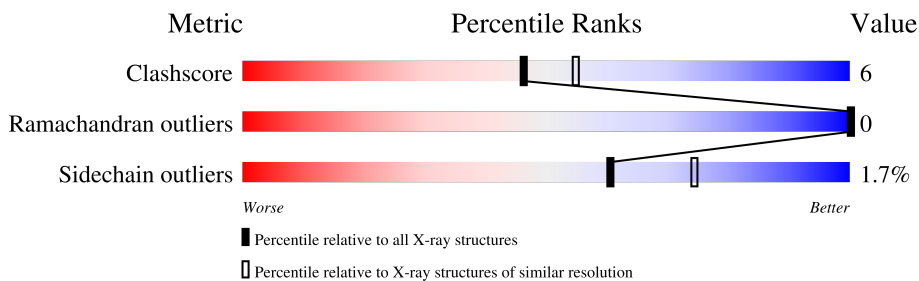|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.34

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 2.20 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	Percentile Ranks	Value
Clashscore		6
Ramachandran outliers		0
Sidechain outliers		1.7%

Worse   *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)

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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	496	87% 11% •
2	B	2	100%
3	C	3	33% 33% 33%
4	D	2	50% 50%
5	E	2	50% 50%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
7	BGC	A	1002	-	-	X	-

2 Entry composition ⓘ

There are 10 unique types of molecules in this entry. The entry contains 4257 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 ALPHA-AMYLASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	496	Total 3904	C 2466	N 688	O 729	S 21	19	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	49	VAL	ILE	conflict	UNP P00690
A	243	SER	GLN	conflict	UNP P00690
A	310	SER	ALA	conflict	UNP P00690
A	404	GLN	GLU	conflict	UNP P00690
A	451	ASN	ASP	conflict	UNP P00690
A	484	GLN	GLU	conflict	UNP P00690

- Molecule 2 is an oligosaccharide called 4-amino-4,6-dideoxy-alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose.

4N

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	2	Total 21	C 12	N 1	O 8	0	0	0

- Molecule 3 is an oligosaccharide called 4-amino-4,6-dideoxy-alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose.

4N
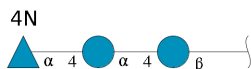
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	C	3	33	18	1	14	0	0	0

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

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
			Total	C	O			
4	D	2	23	12	11	0	0	0

- Molecule 5 is an oligosaccharide called beta-D-glucopyranose-(1-4)-alpha-D-glucopyranose.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
			Total	C	O			
5	E	2	23	12	11	0	0	0

- Molecule 6 is 5-HYDROXYMETHYL-CHONDURITOL (three-letter code: HMC) (formula: $C_7H_{12}O_5$).
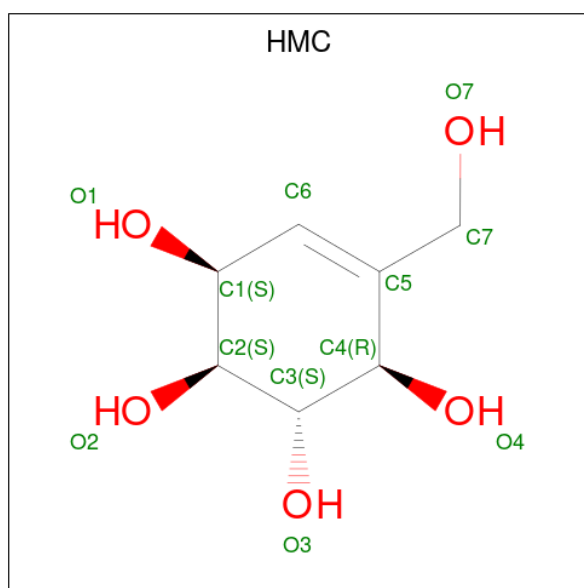
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total 11	C 7	O 4	0	0

- Molecule 7 is beta-D-glucopyranose (three-letter code: BGC) (formula: $C_6H_{12}O_6$).

BGC
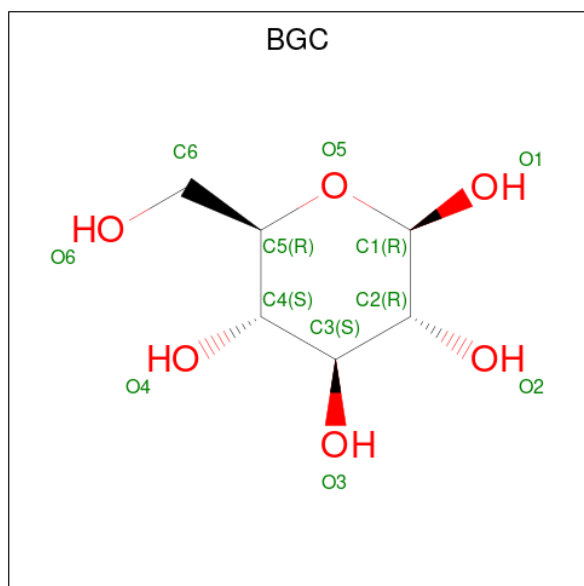
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total 12	C 6	O 6	0	0

- Molecule 8 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

- Molecule 9 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	1	Total 1	Cl 1	0	0

- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	228	Total 228	O 228	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. 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. 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.

Note EDS was not executed.

• Molecule 1: ALPHA-AMYLASE

Chain A:
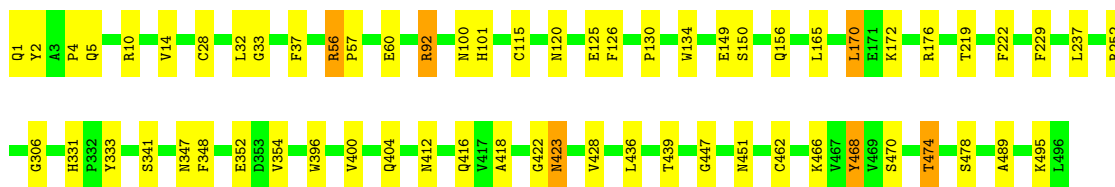
• Molecule 2: 4-amino-4,6-dideoxy-alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose

Chain B:

• Molecule 3: 4-amino-4,6-dideoxy-alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose-(1-4)-beta-D-glucopyranose

Chain C:

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

Chain D:

• Molecule 5: beta-D-glucopyranose-(1-4)-alpha-D-glucopyranose

Chain E:

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	70.50Å 114.80Å 118.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	7.00 – 2.20	Depositor
% Data completeness (in resolution range)	93.7 (7.00-2.20)	Depositor
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	X-PLOR 3.1	Depositor
R, R_{free}	0.178 , 0.210	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4257	wwPDB-VP
Average B, all atoms (Å^2)	29.0	wwPDB-VP

5   Model quality ⓘ

5.1   Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: AGL, BGC, PCA, CL, CA, HMC, 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	0.36	0/4007	0.63	0/5447

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
1	A	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	468	TYR	Sidechain

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	3904	0	3681	49	0
2	B	21	0	20	0	0

Continued on next page...

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	33	0	29	1	0
4	D	23	0	21	0	0
5	E	23	0	21	0	0
6	A	11	0	9	0	0
7	A	12	0	12	8	0
8	A	1	0	0	0	0
9	A	1	0	0	0	0
10	A	228	0	0	1	0
All	All	4257	0	3793	49	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 6.

The worst 5 of 49 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:470:SER:HB3	7:A:1002:BGC:H2	1.63	0.80
1:A:474:THR:HG22	7:A:1002:BGC:H4	1.74	0.69
1:A:149:GLU:H	1:A:156:GLN:NE2	1.93	0.66
1:A:100:ASN:HD22	1:A:101:HIS:HD2	1.44	0.66
1:A:120:ASN:HB3	1:A:125:GLU:CG	2.27	0.64

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	494/496 (100%)	476 (96%)	18 (4%)	0	100 100

There are no Ramachandran outliers to report.

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	412/412 (100%)	405 (98%)	7 (2%)	60	74

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	170	LEU
1	A	341	SER
1	A	474	THR
1	A	423	ASN
1	A	115	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	355	ASN
1	A	423	ASN
1	A	435	GLN
1	A	156	GLN
1	A	161	GLN

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	PCA	A	1	1	7,8,9	0.48	0	9,10,12	4.53	3 (33%)

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	PCA	A	1	1	-	0/0/11/13	0/1/1/1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	1	PCA	OE-CD-CG	10.70	145.44	126.76
1	A	1	PCA	OE-CD-N	-7.84	106.57	124.86
1	A	1	PCA	O-C-CA	-2.23	118.94	124.78

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 ⓘ

9 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	B	1	2,6	11,11,12	0.69	0	15,15,17	1.07	2 (13%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z|$ > 2	Counts	RMSZ	#$|Z|$ > 2
2	AGL	B	2	2	9,10,11	3.02	1 (11%)	14,14,16	0.99	2 (14%)
3	BGC	C	1	3	12,12,12	0.35	0	17,17,17	0.88	1 (5%)
3	GLC	C	2	3	11,11,12	0.49	0	15,15,17	0.75	0
3	AGL	C	3	3,6	9,10,11	0.56	0	14,14,16	0.93	1 (7%)
4	GLC	D	1	4	12,12,12	0.43	0	17,17,17	0.83	1 (5%)
4	GLC	D	2	4	11,11,12	0.41	0	15,15,17	0.70	0
5	GLC	E	1	5	12,12,12	0.68	0	17,17,17	1.26	1 (5%)
5	BGC	E	2	5	11,11,12	0.43	0	15,15,17	0.66	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
2	GLC	B	1	2,6	-	0/2/19/22	0/1/1/1
2	AGL	B	2	2	-	-	0/1/1/1
3	BGC	C	1	3	-	0/2/22/22	0/1/1/1
3	GLC	C	2	3	-	0/2/19/22	0/1/1/1
3	AGL	C	3	3,6	-	-	0/1/1/1
4	GLC	D	1	4	-	0/2/22/22	0/1/1/1
4	GLC	D	2	4	-	0/2/19/22	0/1/1/1
5	GLC	E	1	5	-	0/2/22/22	0/1/1/1
5	BGC	E	2	5	-	0/2/19/22	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	2	AGL	C4-N4	-8.96	1.33	1.47

The worst 5 of 8 bond angle outliers are listed below:

| Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	E	1	GLC	C4-C3-C2	-3.57	104.58	110.82
2	B	1	GLC	C1-C2-C3	2.68	112.97	109.67
4	D	1	GLC	C3-C4-C5	-2.41	105.93	110.24
3	C	3	AGL	C1-C2-C3	2.33	112.53	109.67
2	B	1	GLC	C1-O5-C5	2.16	115.12	112.19

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	1	BGC	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.

Oligosaccharide Chain B

Bond lengths   Bond angles

Torsions   Rings

Oligosaccharide Chain C

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain D

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain E

Bond lengths   Bond angles

Torsions   Rings
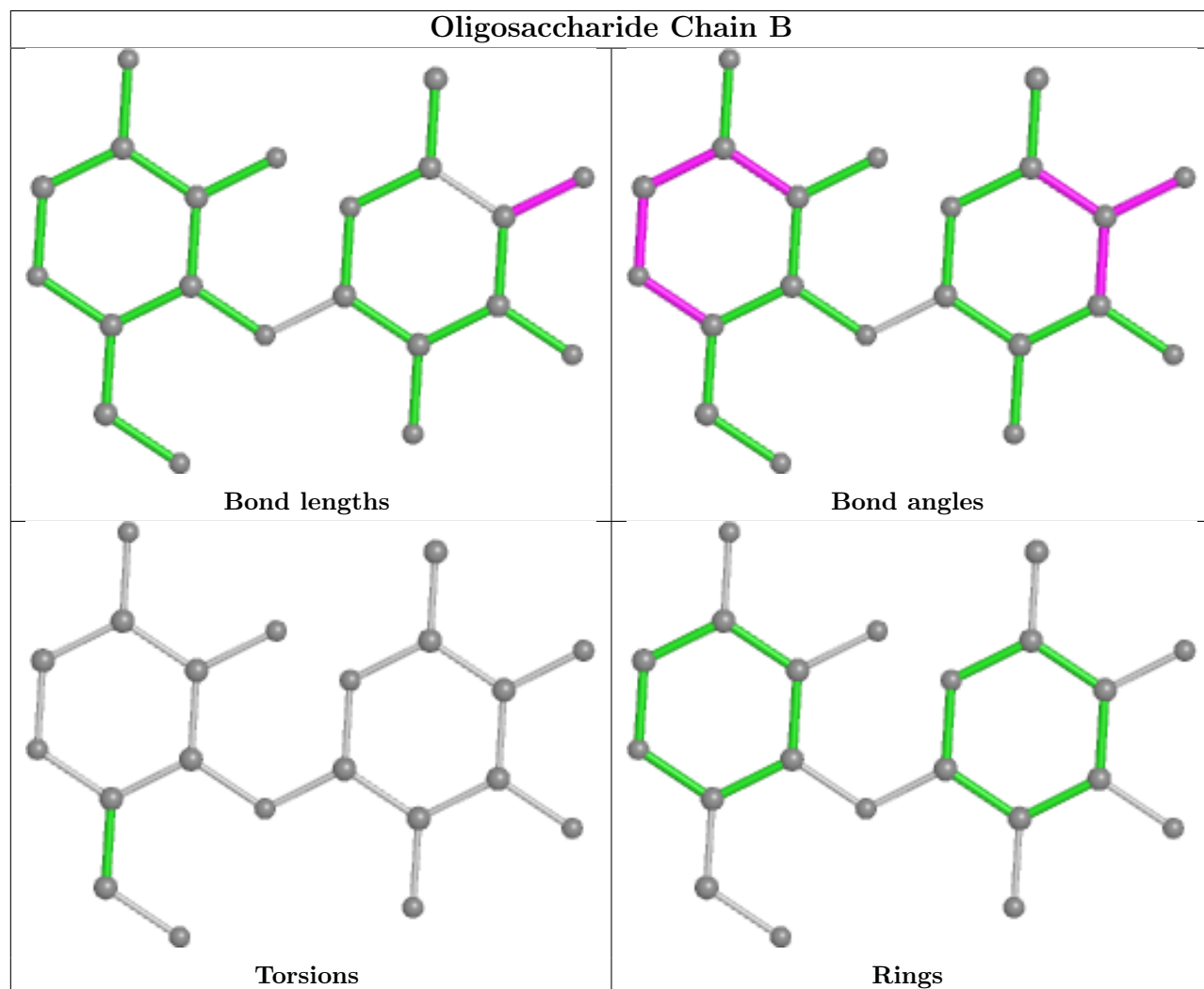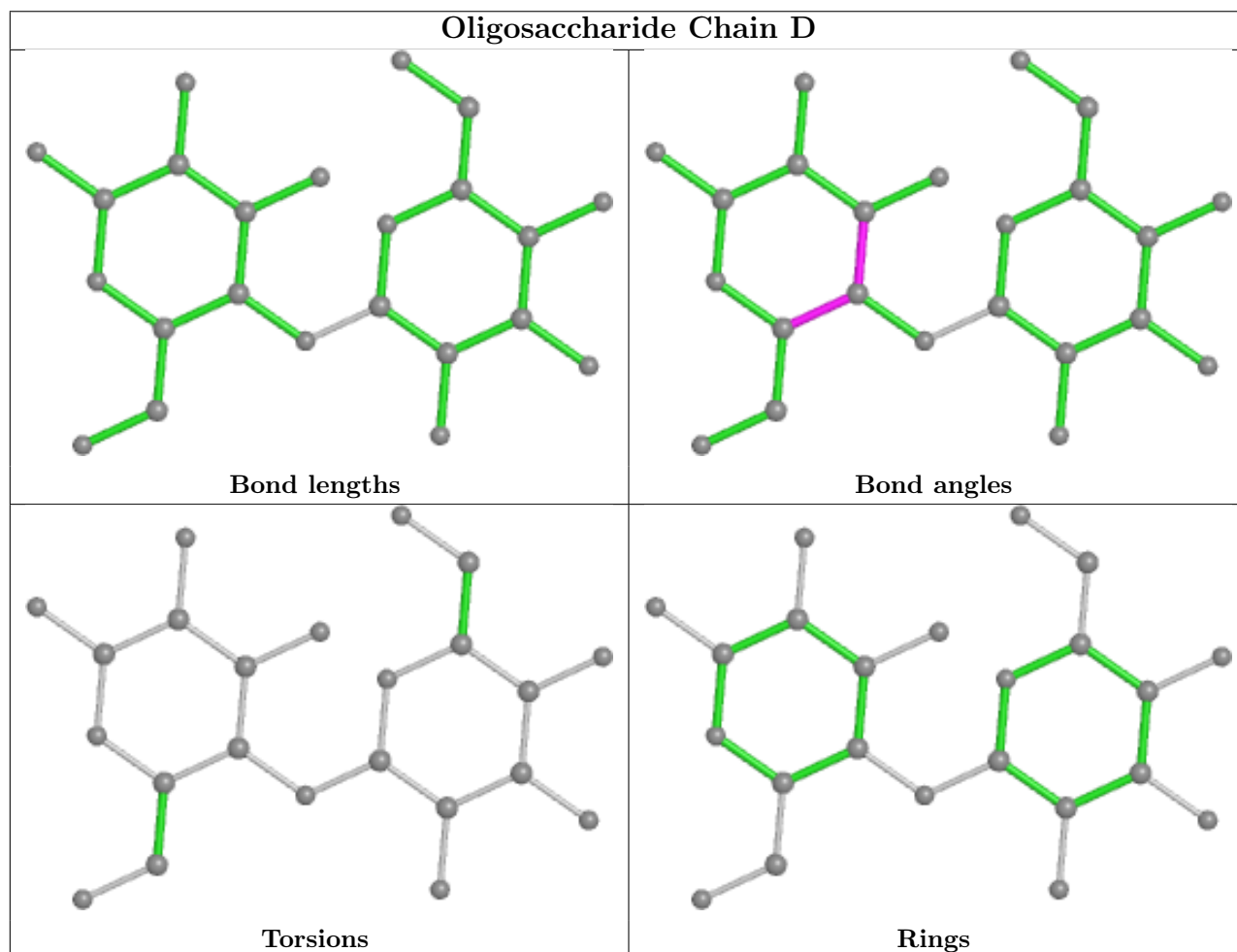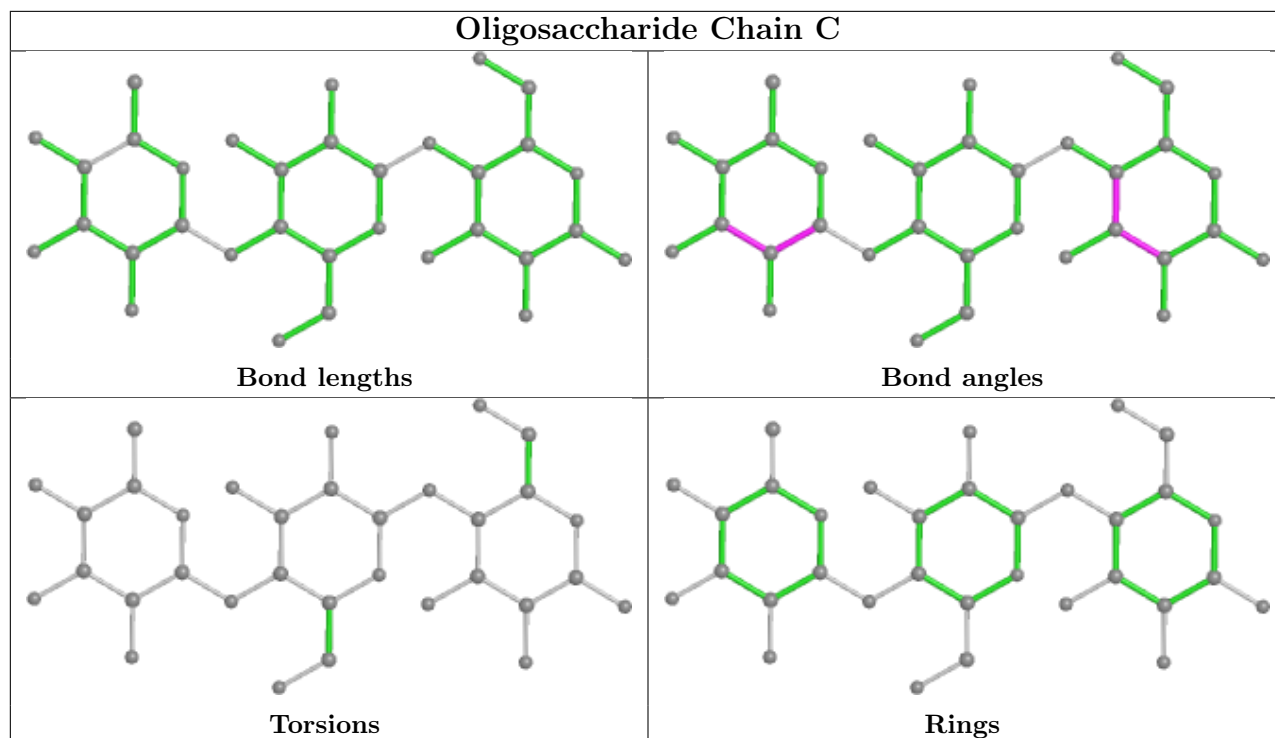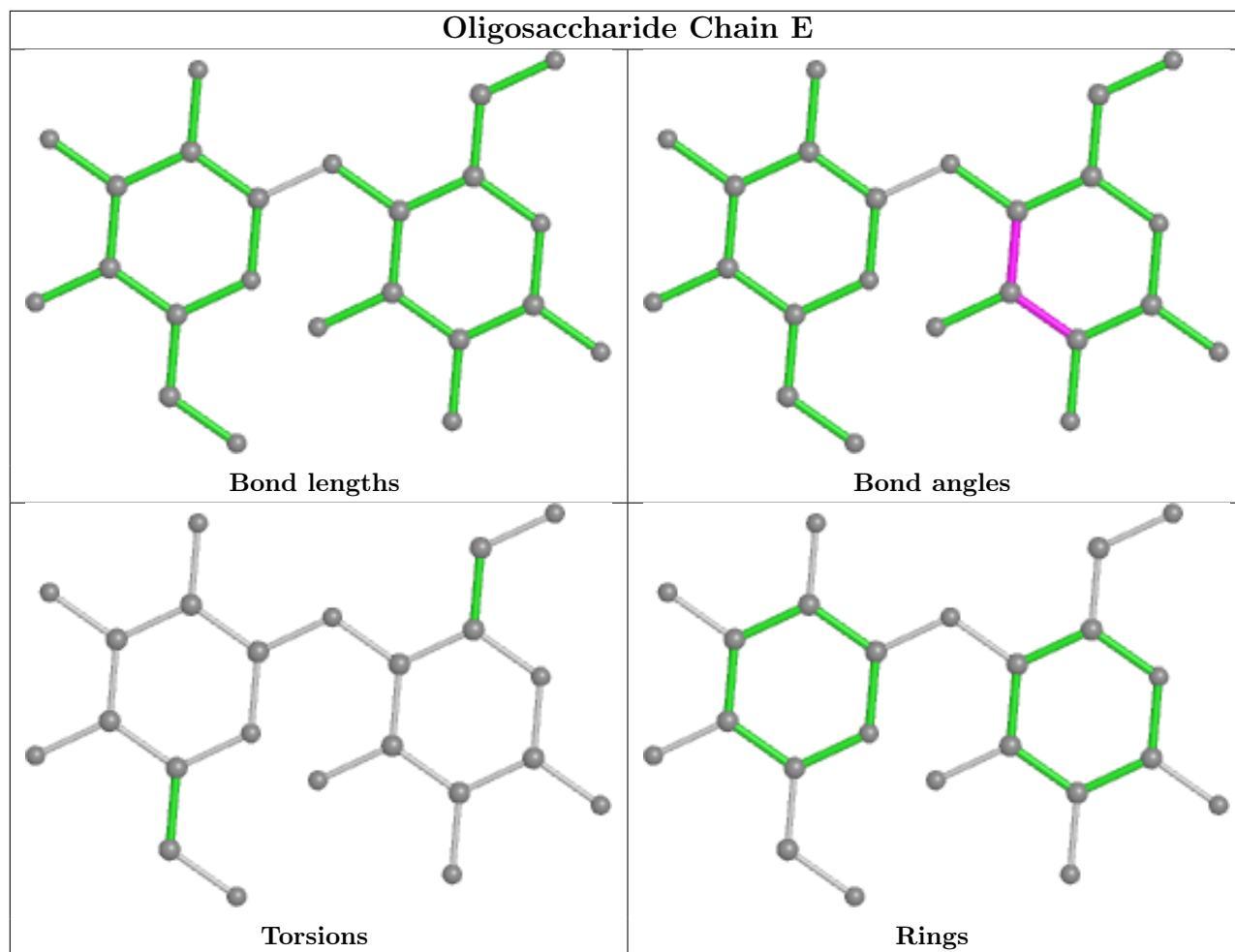
5.6   Ligand geometry ⓘ

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 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
6	HMC	A	703	2,3	11,11,12	2.13	3 (27%)	9,15,17	1.42	1 (11%)
7	BGC	A	1002	-	12,12,12	0.54	0	17,17,17	1.26	2 (11%)

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	HMC	A	703	2,3	-	1/2/19/22	0/1/1/1
7	BGC	A	1002	-	-	0/2/22/22	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	703	HMC	C4-C5	4.73	1.55	1.51
6	A	703	HMC	C3-C4	4.00	1.58	1.53
6	A	703	HMC	C1-C2	2.20	1.54	1.52

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	A	1002	BGC	C4-C3-C2	-3.23	105.18	110.82
7	A	1002	BGC	O1-C1-O5	-2.71	102.25	110.38
6	A	703	HMC	O7-C7-C5	-2.06	107.57	112.50

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	703	HMC	C4-C5-C7-O7

There are no ring outliers.

1 monomer is involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	A	1002	BGC	8	0

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.