wwPDB X-ray Structure Validation Summary Report (i)

Aug 8, 2020 – 10:47 PM BST

PDB ID	:	2AAI
Title	:	Crystallographic refinement of ricin to 2.5 Angstroms
Authors	:	Rutenber, E.; Katzin, B.J.; Montfort, W.; Villafranca, J.E.; Ernst, S.R.; Collins, E.J.; Mlsna, D.; Monzingo, A.F.; Ready, M.P.; Robertus, J.D.
Deposited on	:	1993-09-07
Resolution	:	2.50 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (i)) were used in the production of this report:

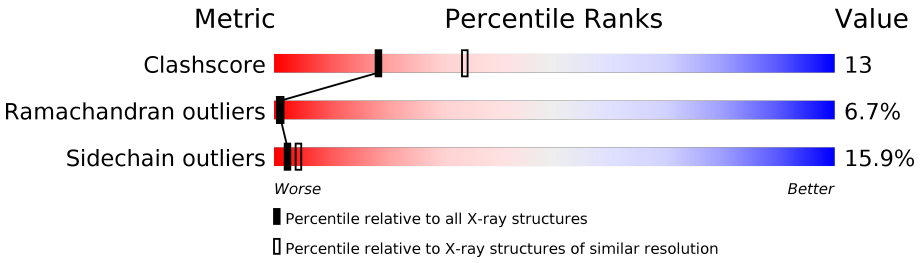
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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))
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)

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 >=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	267	51% 30% 14% 5%
2	B	262	53% 31% 14% .
3	C	2	100%
3	D	2	100%
4	E	5	80% 20%
4	F	5	100%

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
4	NAG	E	1	X	-	-	-
4	NAG	F	1	X	-	-	-

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 4440 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.

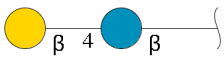
- Molecule 1 is a protein called RICIN (A CHAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	267	2114	1342	372	395	5	0	0	0

- Molecule 2 is a protein called RICIN (B CHAIN).

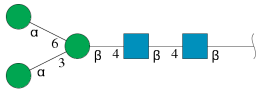
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	262	2035	1273	357	393	12	0	0	0

- Molecule 3 is an oligosaccharide called beta-D-galactopyranose-(1-4)-beta-D-glucopyranose.



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

- 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	E	5	Total	C	N	O	0	0	0
			61	34	2	25			
4	F	5	Total	C	N	O	0	0	0
			61	34	2	25			

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	61	Total	O	0	0
			61	61		
5	B	62	Total	O	0	0
			62	62		

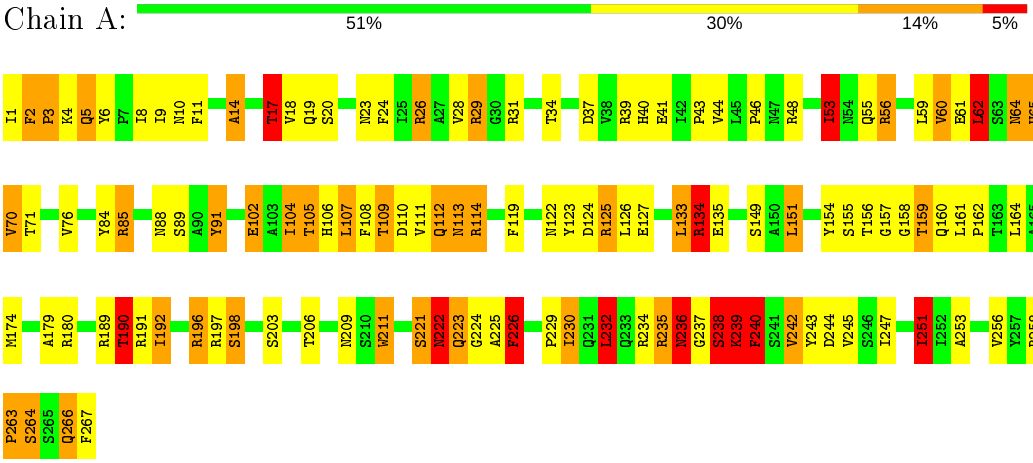
3 Residue-property plots (i)

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.

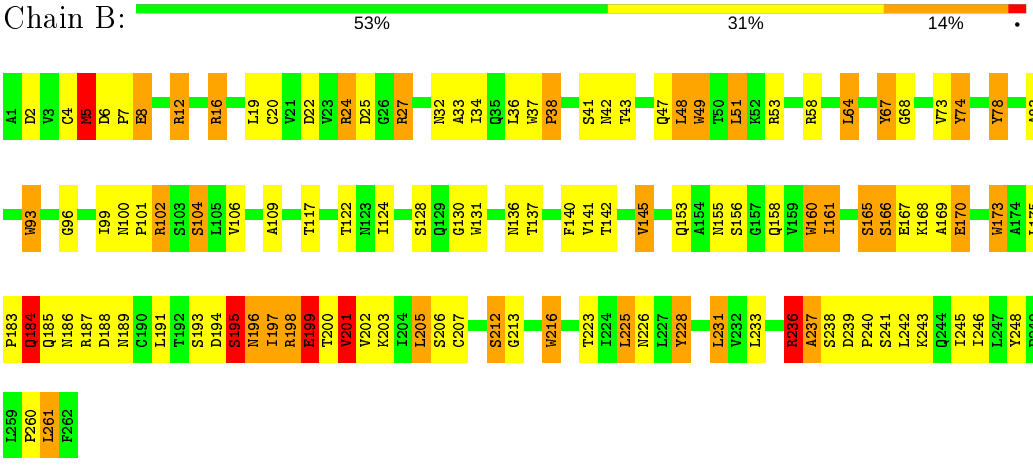
- Molecule 1: RICIN (A CHAIN)

Chain A:



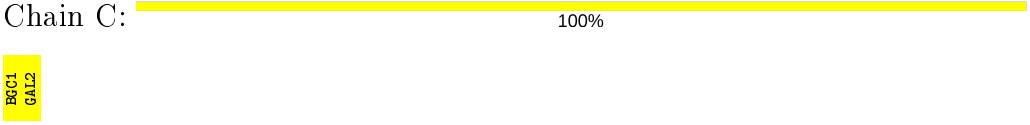
- Molecule 2: RICIN (B CHAIN)

Chain B:



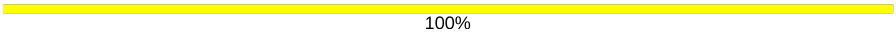
- Molecule 3: beta-D-galactopyranose-(1-4)-beta-D-glucopyranose

Chain C:



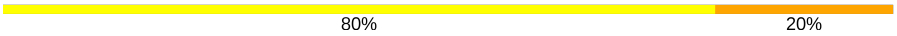
• Molecule 3: beta-D-galactopyranose-(1-4)-beta-D-glucopyranose

Chain D:



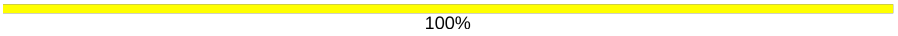
• Molecule 4: alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain E:



• Molecule 4: 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

Chain F:



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	72.74Å 78.49Å 114.34Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	(Not available) – 2.50	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-2.50)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	X-PLOR	Depositor
R, R_{free}	0.212 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4440	wwPDB-VP
Average B, all atoms (Å^2)	26.0	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: MAN, GAL, BGC, BMA, 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	1.04	0/2162	2.00	89/2941 (3.0%)
2	B	1.04	2/2080 (0.1%)	2.12	86/2842 (3.0%)
All	All	1.04	2/4242 (0.0%)	2.06	175/5783 (3.0%)

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	3
2	B	0	3
All	All	0	6

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	131	TRP	CD1-NE1	-5.74	1.28	1.38
2	B	195	SER	CA-CB	5.14	1.60	1.52

The worst 5 of 175 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	200	THR	CA-C-N	-13.18	88.21	117.20
1	A	234	ARG	NE-CZ-NH2	-13.11	113.74	120.30
1	A	234	ARG	NE-CZ-NH1	11.23	125.91	120.30
2	B	37	TRP	CD1-CG-CD2	10.77	114.92	106.30
1	A	258	ARG	NE-CZ-NH1	10.62	125.61	120.30

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	158	GLY	Mainchain
1	A	262	PRO	Peptide
1	A	56	ARG	Sidechain
2	B	170	GLU	Mainchain
2	B	74	TYR	Sidechain

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	2114	0	2083	63	0
2	B	2035	0	1974	51	0
3	C	23	0	21	0	0
3	D	23	0	21	0	0
4	E	61	0	52	1	0
4	F	61	0	52	0	0
5	A	61	0	0	2	0
5	B	62	0	0	4	0
All	All	4440	0	4203	112	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 13.

The worst 5 of 112 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:1:ILE:HG23	1:A:2:PHE:H	1.51	0.76
1:A:236:ASN:HD21	1:A:238:SER:HB2	1.53	0.73
1:A:225:ALA:HB1	1:A:226:PHE:CD1	2.29	0.67
2:B:193:SER:HA	2:B:201:VAL:O	1.94	0.67
2:B:198:ARG:HB3	2:B:199:GLU:OE1	1.94	0.67

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	265/267 (99%)	231 (87%)	17 (6%)	17 (6%)	1 1
2	B	260/262 (99%)	222 (85%)	20 (8%)	18 (7%)	1 1
All	All	525/529 (99%)	453 (86%)	37 (7%)	35 (7%)	1 1

5 of 35 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	3	PRO
1	A	14	ALA
1	A	66	ALA
1	A	67	GLU
1	A	222	ASN

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	226/226 (100%)	184 (81%)	42 (19%)	1 2
2	B	227/227 (100%)	197 (87%)	30 (13%)	4 7
All	All	453/453 (100%)	381 (84%)	72 (16%)	2 4

5 of 72 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	232	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	263	PRO
2	B	225	LEU
1	A	235	ARG
1	A	239	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	233	GLN
1	A	236	ASN
2	B	196	ASN
1	A	209	ASN
2	B	189	ASN

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)

14 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	BGC	C	1	3	12,12,12	1.28	1 (8%)	17,17,17	1.71	3 (17%)
3	GAL	C	2	3	11,11,12	1.00	1 (9%)	15,15,17	1.41	2 (13%)
3	BGC	D	1	3	12,12,12	1.18	1 (8%)	17,17,17	2.06	3 (17%)
3	GAL	D	2	3	11,11,12	1.12	1 (9%)	15,15,17	2.01	7 (46%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
4	NAG	E	1	2,4	14,14,15	1.26	1 (7%)	17,19,21	1.79	5 (29%)
4	NAG	E	2	4	14,14,15	1.76	6 (42%)	17,19,21	3.64	7 (41%)
4	BMA	E	3	4	11,11,12	1.61	2 (18%)	15,15,17	4.15	9 (60%)
4	MAN	E	4	4	11,11,12	1.42	2 (18%)	15,15,17	1.08	0
4	MAN	E	5	4	11,11,12	2.20	5 (45%)	15,15,17	2.38	8 (53%)
4	NAG	F	1	2,4	14,14,15	0.86	0	17,19,21	1.46	4 (23%)
4	NAG	F	2	4	14,14,15	1.08	1 (7%)	17,19,21	2.83	8 (47%)
4	BMA	F	3	4	11,11,12	1.86	2 (18%)	15,15,17	2.04	7 (46%)
4	MAN	F	4	4	11,11,12	1.25	1 (9%)	15,15,17	3.09	11 (73%)
4	MAN	F	5	4	11,11,12	1.42	0	15,15,17	3.95	9 (60%)

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	BGC	C	1	3	-	0/2/22/22	0/1/1/1
3	GAL	C	2	3	-	0/2/19/22	0/1/1/1
3	BGC	D	1	3	-	2/2/22/22	0/1/1/1
3	GAL	D	2	3	-	0/2/19/22	0/1/1/1
4	NAG	E	1	2,4	1/1/5/7	1/6/23/26	0/1/1/1
4	NAG	E	2	4	-	2/6/23/26	0/1/1/1
4	BMA	E	3	4	-	2/2/19/22	0/1/1/1
4	MAN	E	4	4	-	2/2/19/22	0/1/1/1
4	MAN	E	5	4	-	2/2/19/22	0/1/1/1
4	NAG	F	1	2,4	1/1/5/7	2/6/23/26	0/1/1/1
4	NAG	F	2	4	-	4/6/23/26	0/1/1/1
4	BMA	F	3	4	-	1/2/19/22	0/1/1/1
4	MAN	F	4	4	-	0/2/19/22	1/1/1/1
4	MAN	F	5	4	-	1/2/19/22	1/1/1/1

The worst 5 of 24 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	E	5	MAN	C2-C3	3.94	1.58	1.52
4	E	4	MAN	C2-C3	-3.93	1.46	1.52
4	F	3	BMA	O5-C5	3.90	1.51	1.43

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	E	3	BMA	O2-C2	3.32	1.50	1.43
4	F	3	BMA	C2-C3	3.18	1.57	1.52

The worst 5 of 83 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	F	5	MAN	O5-C1-C2	-11.35	93.25	110.77
4	E	3	BMA	C1-C2-C3	-10.34	96.96	109.67
4	E	2	NAG	C1-O5-C5	-7.93	101.44	112.19
4	E	2	NAG	O5-C1-C2	-7.86	98.88	111.29
4	F	2	NAG	C2-N2-C7	-7.01	112.92	122.90

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
4	E	1	NAG	C1
4	F	1	NAG	C1

5 of 19 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	F	2	NAG	C8-C7-N2-C2
4	F	2	NAG	O7-C7-N2-C2
4	E	3	BMA	O5-C5-C6-O6
4	E	4	MAN	C4-C5-C6-O6
4	E	3	BMA	C4-C5-C6-O6

All (2) ring outliers are listed below:

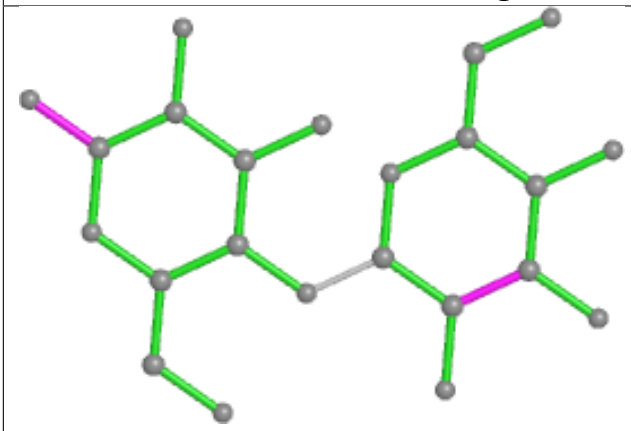
Mol	Chain	Res	Type	Atoms
4	F	5	MAN	C1-C2-C3-C4-C5-O5
4	F	4	MAN	C1-C2-C3-C4-C5-O5

1 monomer is involved in 1 short contact:

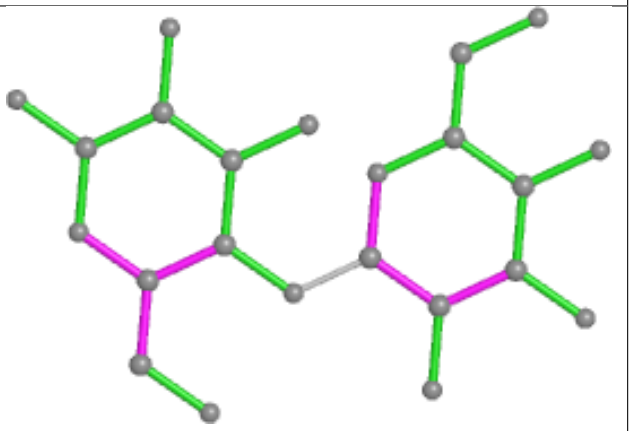
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	E	2	NAG	1	0

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

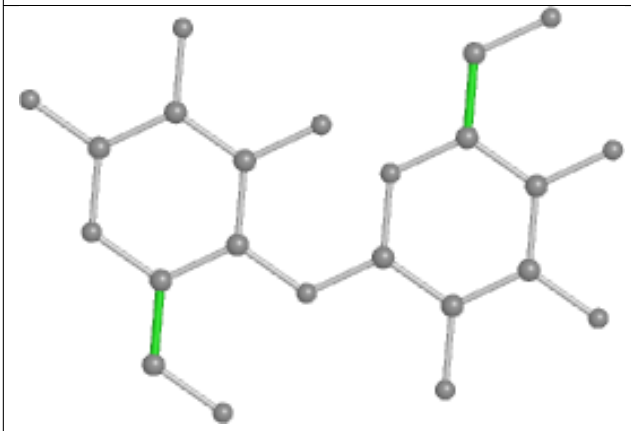
Oligosaccharide Chain C



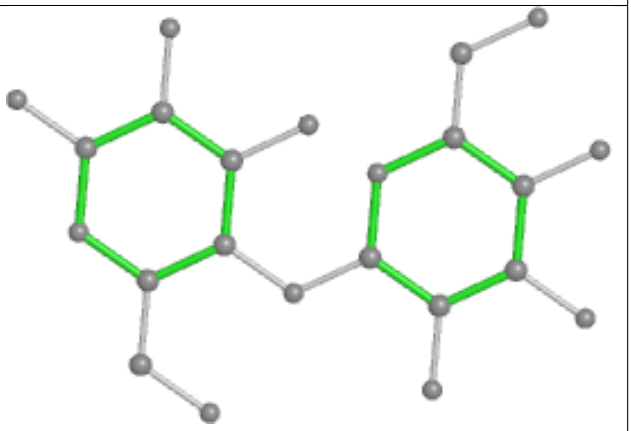
Bond lengths



Bond angles

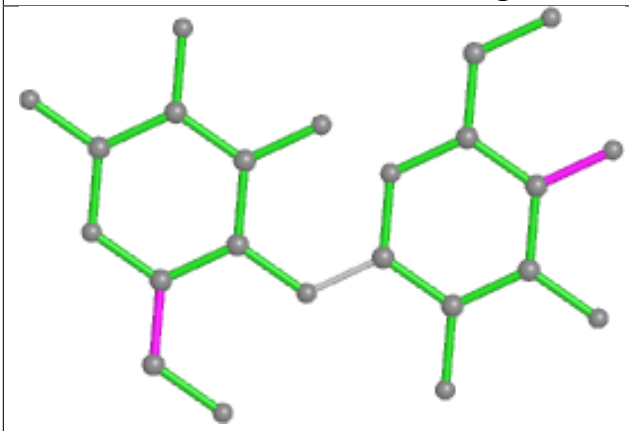


Torsions

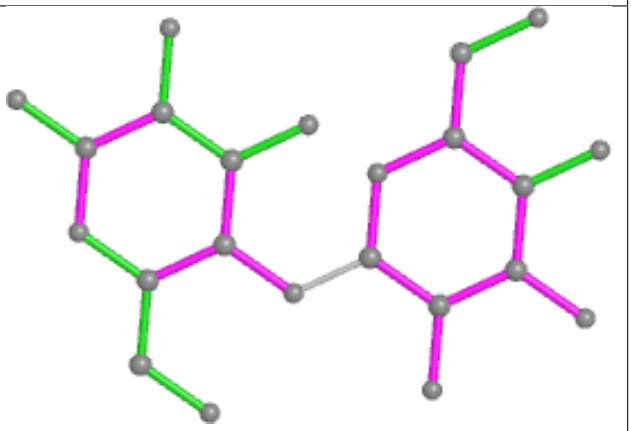


Rings

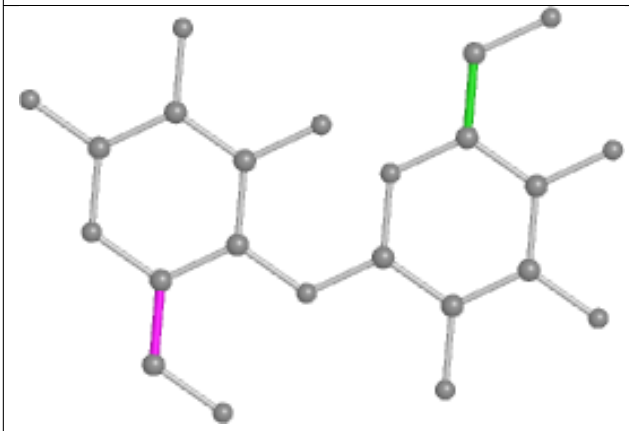
Oligosaccharide Chain D



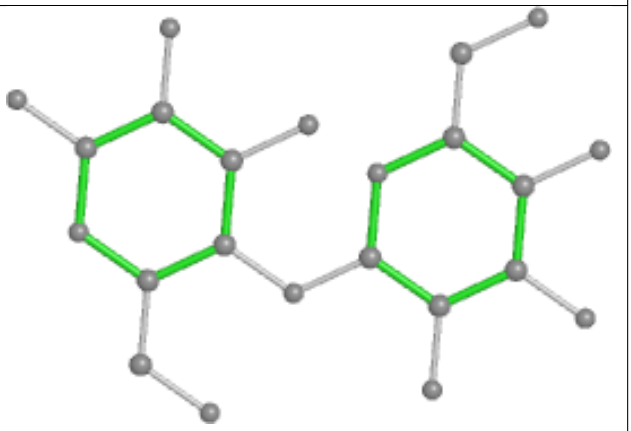
Bond lengths



Bond angles

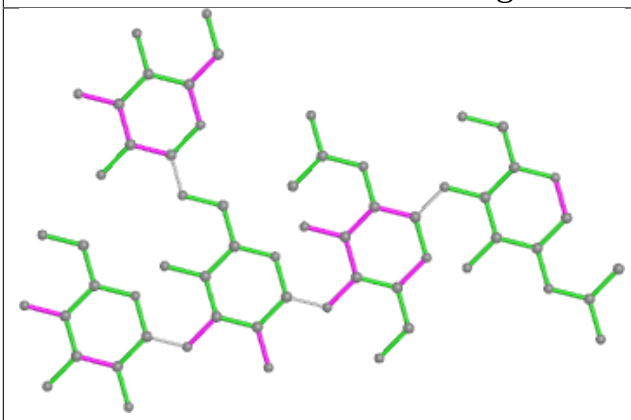


Torsions

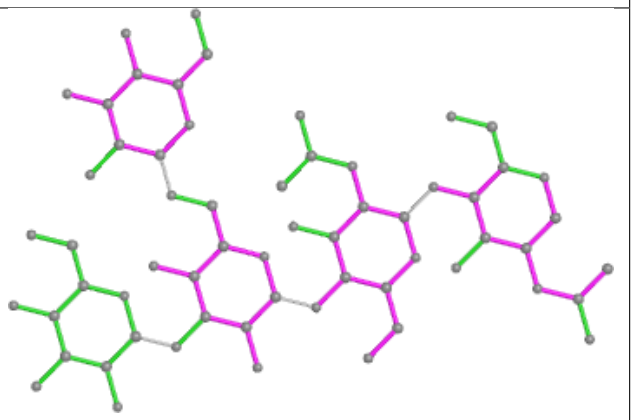


Rings

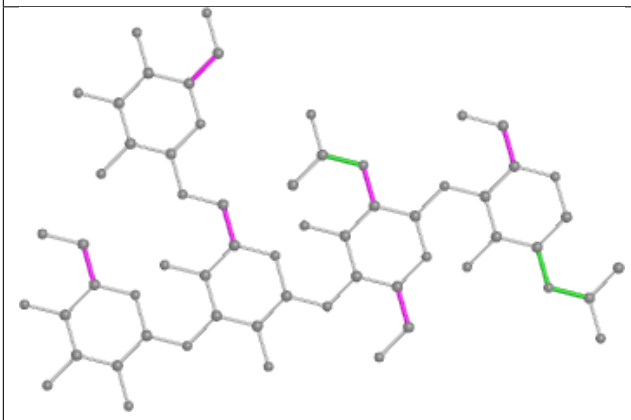
Oligosaccharide Chain E



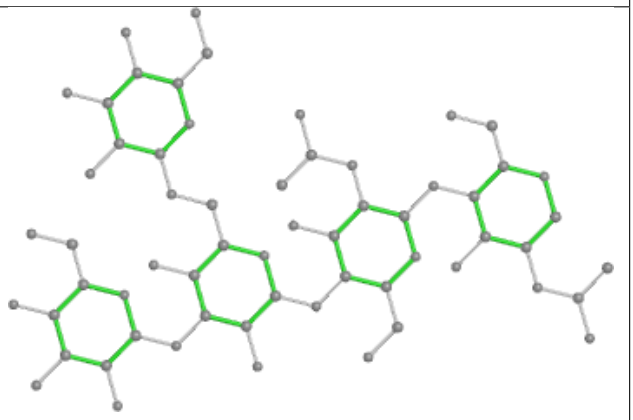
Bond lengths



Bond angles

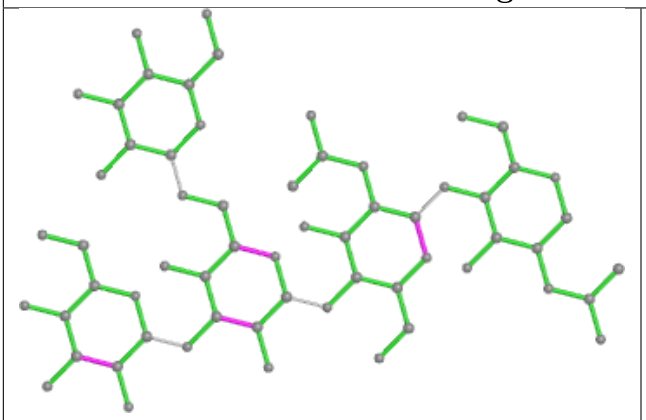


Torsions

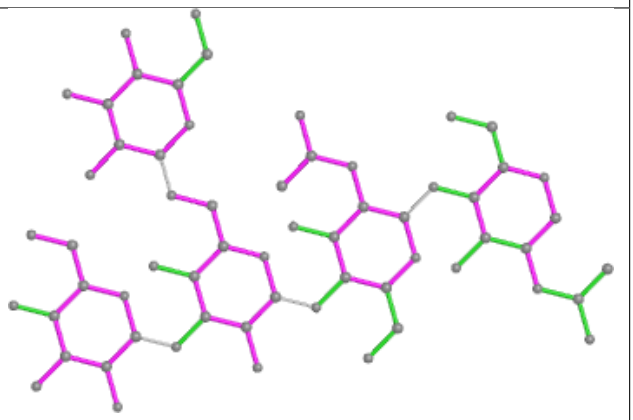


Rings

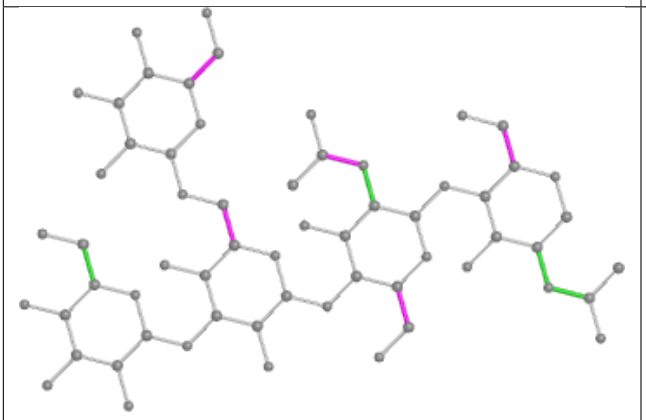
Oligosaccharide Chain F



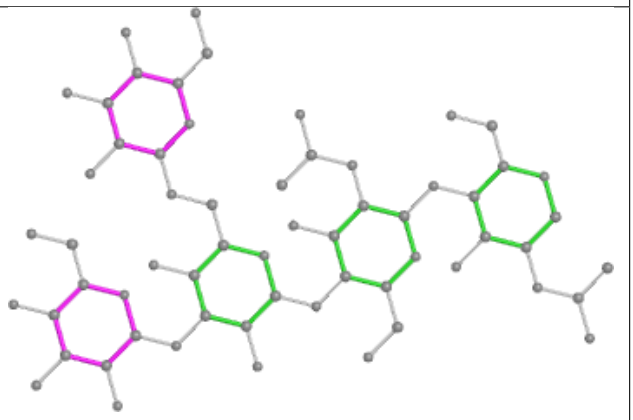
Bond lengths



Bond angles



Torsions



Rings

5.6 Ligand geometry (i)

There are no ligands in this entry.

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

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.