



Full wwPDB X-ray Structure Validation Report i

Oct 12, 2024 – 10:00 PM EDT

PDB ID : 2AVI
Title : THREE-DIMENSIONAL STRUCTURES OF AVIDIN AND THE AVIDIN-BIOTIN COMPLEX
Authors : Livnah, O.; Sussman, J.
Deposited on : 1993-04-23
Resolution : 3.00 Å (reported)

This is a Full wwPDB X-ray Structure Validation 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](#) i) were used in the production of this report:

MolProbitiy : 4.02b-467
Mogul : 2022.3.0, CSD as543be (2022)
Xtriaage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.39

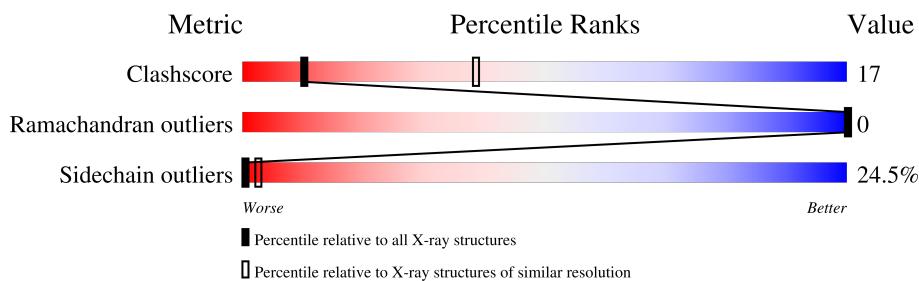
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 Å.

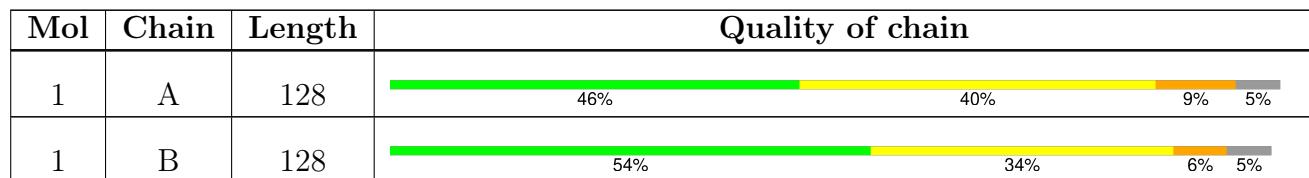
Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	180529	2866 (3.00-3.00)
Ramachandran outliers	177936	2778 (3.00-3.00)
Sidechain outliers	177891	2781 (3.00-3.00)

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 $\leq 5\%$

Note EDS was not executed.



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
2	NAG	A	340	X	-	-	-
2	NAG	B	341	X	-	-	-

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 2026 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 AVIDIN.

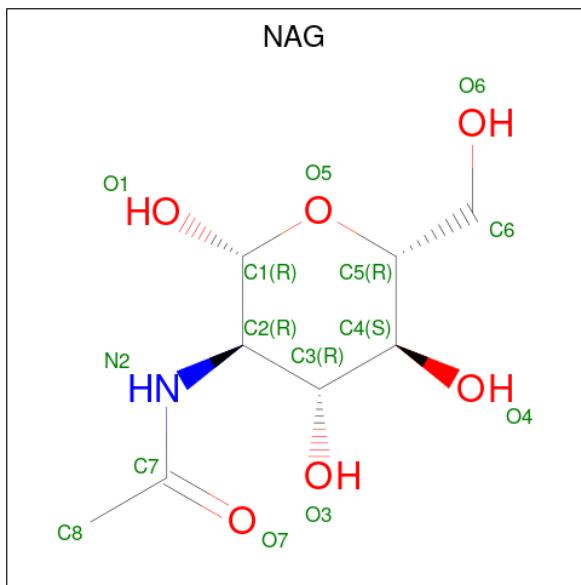
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	A	121	Total	C	N	O	S		
			946	592	165	185	4	0	0

Mol	Chain	Residues	Total	C	N	O	S	ZeroOcc	AltConf	Trace
1	B	121	Total	C	N	O	S			
			946	592	165	185	4	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	34	THR	ILE	conflict	UNP P02701
B	34	THR	ILE	conflict	UNP P02701

- Molecule 2 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C₈H₁₅NO₆).



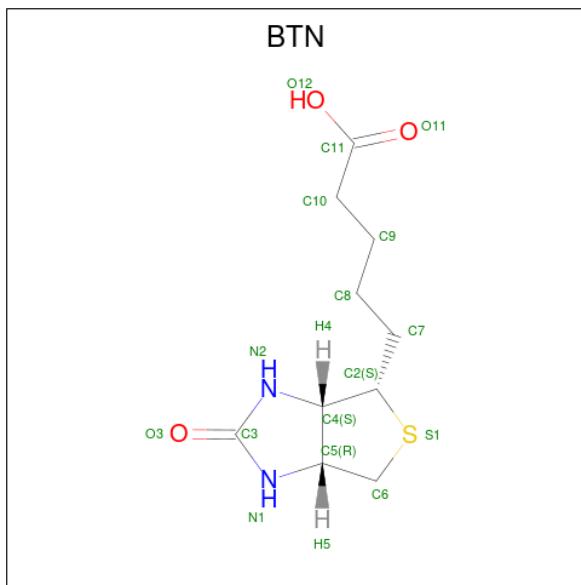
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O		
			14	8	1	5	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total C N O 14 8 1 5	0	0

- Molecule 3 is BIOTIN (three-letter code: BTN) (formula: C₁₀H₁₆N₂O₃S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C N O S 16 10 2 3 1	0	0
3	B	1	Total C N O S 16 10 2 3 1	0	0

- Molecule 4 is water.

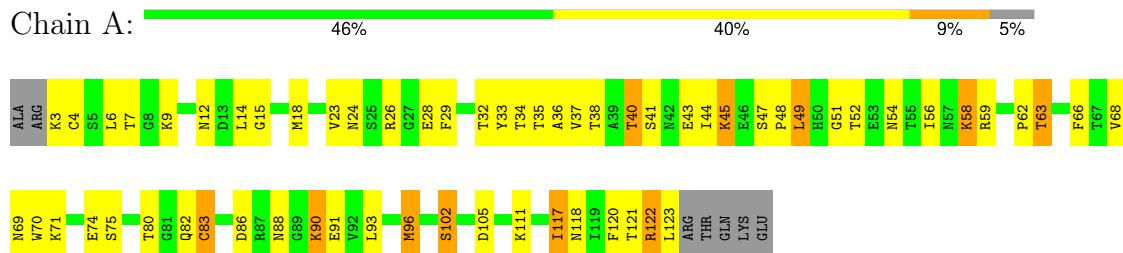
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	34	Total O 34 34	0	0
4	B	40	Total O 40 40	0	0

3 Residue-property plots [\(i\)](#)

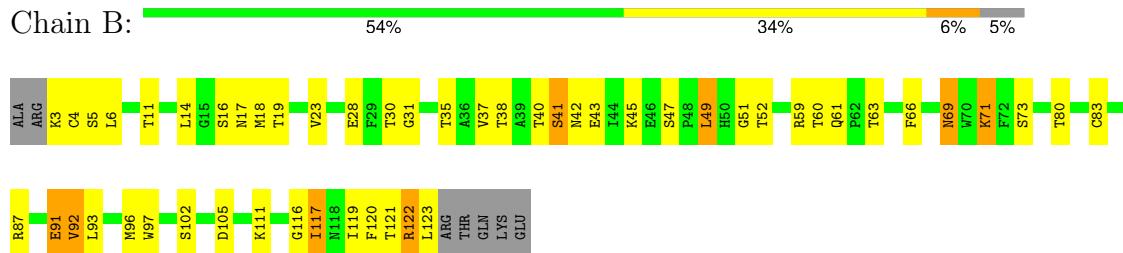
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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: AVIDIN



- Molecule 1: AVIDIN



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 2	Depositor
Cell constants a, b, c, α , β , γ	72.15 \AA 80.37 \AA 43.54 \AA 90.00° 90.00° 90.00°	Depositor
Resolution (\AA)	(Not available) – 3.00	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-3.00)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	X-PLOR	Depositor
R , R_{free}	0.151, (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2026	wwPDB-VP
Average B, all atoms (\AA^2)	11.0	wwPDB-VP

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: BTN, 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	0.66	0/964	0.85	0/1306
1	B	0.64	0/964	0.85	1/1306 (0.1%)
All	All	0.65	0/1928	0.85	1/2612 (0.0%)

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	2
1	B	0	3
All	All	0	5

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	49	LEU	CA-CB-CG	6.85	131.06	115.30

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	33	TYR	Sidechain
1	A	35	THR	Mainchain
1	B	28	GLU	Mainchain
1	B	60	THR	Mainchain
1	B	91	GLU	Mainchain

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	946	0	925	34	0
1	B	946	0	925	27	0
2	A	14	0	13	4	0
2	B	14	0	13	6	0
3	A	16	0	15	1	0
3	B	16	0	15	0	0
4	A	34	0	0	2	0
4	B	40	0	0	8	0
All	All	2026	0	1906	65	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 17.

All (65) 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:4:CYS:SG	1:B:91:GLU:HG2	2.19	0.82
1:A:9:LYS:HE2	1:A:123:LEU:HD23	1.62	0.82
1:A:93:LEU:HB2	1:A:120:PHE:HB2	1.62	0.80
2:B:341:NAG:O7	4:B:542:HOH:O	2.02	0.77
1:A:36:ALA:HB2	2:A:340:NAG:C8	2.16	0.75
1:A:36:ALA:HB2	2:A:340:NAG:H82	1.69	0.74
1:B:123:LEU:HD21	2:B:341:NAG:O6	1.88	0.73
1:A:58:LYS:HB2	1:A:58:LYS:NZ	2.09	0.68
1:B:11:THR:HB	1:B:123:LEU:HD13	1.76	0.67
2:B:341:NAG:H81	4:B:533:HOH:O	1.95	0.66
1:B:30:THR:HG22	1:B:31:GLY:N	2.12	0.65
1:B:16:SER:HB3	1:B:35:THR:HA	1.79	0.64
1:A:7:THR:O	1:A:122:ARG:NH2	2.34	0.61
2:B:341:NAG:H2	4:B:542:HOH:O	2.02	0.59
1:B:38:THR:OG1	1:B:40:THR:HG22	2.01	0.59
1:A:51:GLY:HA3	1:A:66:PHE:HB3	1.85	0.58
1:B:19:THR:O	1:B:31:GLY:HA3	2.02	0.58
1:B:17:ASN:OD1	1:B:123:LEU:HD22	2.04	0.58
1:A:54:ASN:HB2	1:A:63:THR:HG22	1.88	0.56

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:341:NAG:C8	4:B:533:HOH:O	2.53	0.56
1:A:62:PRO:HD2	1:A:83:CYS:HB3	1.87	0.55
1:A:54:ASN:CB	1:A:63:THR:HG22	2.39	0.53
1:B:30:THR:CG2	1:B:31:GLY:N	2.72	0.52
1:B:73:SER:HB2	4:B:520:HOH:O	2.09	0.51
1:B:35:THR:HG22	4:B:540:HOH:O	2.10	0.51
1:A:15:GLY:O	2:A:340:NAG:H82	2.10	0.50
1:B:93:LEU:HB2	1:B:120:PHE:HB2	1.94	0.50
1:A:26:ARG:NH2	4:A:548:HOH:O	2.45	0.49
1:A:62:PRO:O	1:A:82:GLN:HG3	2.13	0.49
1:A:86:ASP:HB3	1:A:88:ASN:OD1	2.13	0.48
1:B:37:VAL:HG22	1:B:38:THR:N	2.29	0.48
1:B:91:GLU:OE1	1:B:122:ARG:NH1	2.46	0.48
1:A:51:GLY:HA2	1:A:66:PHE:HA	1.96	0.48
1:B:97:TRP:CZ2	1:B:116:GLY:HA3	2.49	0.48
1:A:68:VAL:HG11	1:A:70:TRP:CH2	2.50	0.47
2:B:341:NAG:O4	4:B:537:HOH:O	2.20	0.47
1:B:61:GLN:HB3	1:B:83:CYS:O	2.14	0.46
1:B:51:GLY:HA3	1:B:66:PHE:HB3	1.97	0.46
1:A:70:TRP:CZ2	3:A:150:BTN:H82	2.50	0.46
1:B:11:THR:CB	1:B:123:LEU:HD13	2.43	0.46
1:B:69:ASN:HD22	1:B:69:ASN:HA	1.63	0.46
1:A:12:ASN:HB3	1:A:120:PHE:CD1	2.51	0.45
1:A:49:LEU:C	1:A:49:LEU:HD23	2.36	0.45
1:A:9:LYS:HE2	1:A:123:LEU:CD2	2.40	0.44
1:A:96:MET:HE3	1:A:117:ILE:HG13	1.99	0.44
1:A:44:ILE:O	1:A:45:LYS:HE2	2.18	0.44
1:B:47:SER:HB3	1:B:69:ASN:O	2.17	0.44
1:A:4:CYT:SG	1:A:91:GLU:HG2	2.57	0.43
1:A:74:GLU:HB2	1:A:102:SER:OG	2.19	0.43
1:A:36:ALA:CB	2:A:340:NAG:C8	2.93	0.43
1:A:56:ILE:HD13	1:A:56:ILE:HA	1.75	0.43
1:B:3:LYS:N	4:B:534:HOH:O	2.51	0.43
1:B:40:THR:O	1:B:40:THR:HG23	2.20	0.42
1:A:38:THR:HG23	1:A:40:THR:O	2.20	0.42
1:A:117:ILE:HG21	1:B:117:ILE:HD12	2.02	0.42
1:B:40:THR:C	1:B:42:ASN:H	2.23	0.42
1:A:23:VAL:HA	1:A:28:GLU:O	2.20	0.42
1:A:6:LEU:HD23	1:A:29:PHE:CE2	2.55	0.41
1:A:3:LYS:HG3	1:A:4:CYT:N	2.35	0.41
1:B:92:VAL:HG23	1:B:119:ILE:HG23	2.02	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:34:THR:HA	4:A:523:HOH:O	2.20	0.41
1:B:71:LYS:HD2	1:B:71:LYS:HA	1.65	0.41
1:A:88:ASN:ND2	1:A:90:LYS:NZ	2.69	0.40
1:B:41:SER:O	1:B:41:SER:OG	2.38	0.40
1:A:47:SER:HA	1:A:48:PRO:HD3	1.88	0.40

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	119/128 (93%)	112 (94%)	7 (6%)	0	100 100
1	B	119/128 (93%)	109 (92%)	10 (8%)	0	100 100
All	All	238/256 (93%)	221 (93%)	17 (7%)	0	100 100

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	Percentiles
1	A	106/112 (95%)	78 (74%)	28 (26%)	0 2
1	B	106/112 (95%)	82 (77%)	24 (23%)	1 3
All	All	212/224 (95%)	160 (76%)	52 (24%)	0 2

All (52) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	14	LEU
1	A	18	MET
1	A	24	ASN
1	A	32	THR
1	A	37	VAL
1	A	40	THR
1	A	41	SER
1	A	43	GLU
1	A	45	LYS
1	A	49	LEU
1	A	52	THR
1	A	58	LYS
1	A	59	ARG
1	A	63	THR
1	A	69	ASN
1	A	71	LYS
1	A	75	SER
1	A	80	THR
1	A	83	CYS
1	A	90	LYS
1	A	96	MET
1	A	102	SER
1	A	105	ASP
1	A	111	LYS
1	A	117	ILE
1	A	118	ASN
1	A	121	THR
1	A	122	ARG
1	B	5	SER
1	B	6	LEU
1	B	14	LEU
1	B	18	MET
1	B	23	VAL
1	B	41	SER
1	B	43	GLU
1	B	45	LYS
1	B	49	LEU
1	B	52	THR
1	B	59	ARG
1	B	63	THR
1	B	69	ASN
1	B	71	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	80	THR
1	B	87	ARG
1	B	92	VAL
1	B	96	MET
1	B	102	SER
1	B	105	ASP
1	B	111	LYS
1	B	117	ILE
1	B	121	THR
1	B	122	ARG

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

Mol	Chain	Res	Type
1	A	69	ASN
1	B	69	ASN
1	B	82	GLN

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 [\(i\)](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [\(i\)](#)

4 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$
3	BTN	A	150	-	17,17,17	1.00	2 (11%)	23,23,23	1.37	2 (8%)
2	NAG	B	341	1	14,14,15	0.83	0	17,19,21	1.94	4 (23%)
3	BTN	B	150	-	17,17,17	1.14	1 (5%)	23,23,23	0.93	0
2	NAG	A	340	1	14,14,15	0.75	0	17,19,21	1.66	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
3	BTN	A	150	-	-	3/7/28/28	0/2/2/2
2	NAG	B	341	1	1/1/5/7	2/6/23/26	0/1/1/1
3	BTN	B	150	-	-	3/7/28/28	0/2/2/2
2	NAG	A	340	1	1/1/5/7	2/6/23/26	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	150	BTN	C6-C5	2.16	1.57	1.53
3	A	150	BTN	C3-N1	-2.01	1.31	1.35
3	B	150	BTN	O12-C11	-2.01	1.24	1.30

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	341	NAG	O5-C1-C2	6.29	121.02	111.29
2	A	340	NAG	O5-C1-C2	5.51	119.81	111.29
3	A	150	BTN	C5-C6-S1	2.98	110.16	106.06
3	A	150	BTN	C4-C2-S1	2.88	108.30	105.03
2	B	341	NAG	C2-N2-C7	-2.65	119.36	122.90
2	B	341	NAG	C1-C2-N2	2.64	114.59	110.43
2	B	341	NAG	C1-O5-C5	2.63	115.71	112.19
2	A	340	NAG	C1-C2-N2	2.38	114.18	110.43

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	A	340	NAG	C1
2	B	341	NAG	C1

All (10) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	150	BTN	S1-C2-C7-C8
3	B	150	BTN	C4-C2-C7-C8
2	B	341	NAG	O5-C5-C6-O6
2	B	341	NAG	C4-C5-C6-O6
3	A	150	BTN	C11-C10-C9-C8
3	A	150	BTN	C2-C7-C8-C9
2	A	340	NAG	C4-C5-C6-O6
2	A	340	NAG	O5-C5-C6-O6
3	B	150	BTN	C7-C8-C9-C10
3	A	150	BTN	S1-C2-C7-C8

There are no ring outliers.

3 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	150	BTN	1	0
2	B	341	NAG	6	0
2	A	340	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 [\(i\)](#)

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.