WORLDWIDE
wwPDB
PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

Oct 27, 2025 – 09:13 am GMT

PDB ID	:	9H6F / pdb_00009h6f
Title	:	Bacteroides ovatus GH98 endoxylanase in complex with arabino-xylooligosac charide
Authors	:	Tomlinson, C.W.E.; Cartmell, A.; Bolam, D.
Deposited on	:	2024-10-24
Resolution	:	2.20 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:

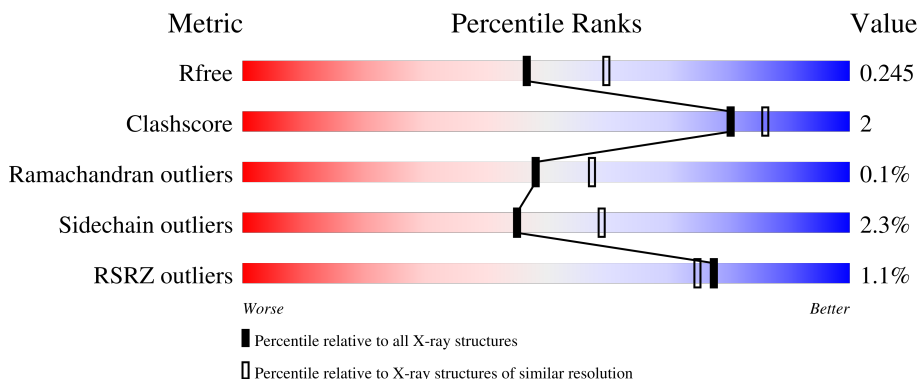|		
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.46

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
Rfree		0.245
Clashscore		2
Ramachandran outliers		0.1%
Sidechain outliers		2.3%
RSRZ outliers		1.1%

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))
R_{free}	164625	5791 (2.20-2.20)
Clashscore	180529	6634 (2.20-2.20)
Ramachandran outliers	177936	6560 (2.20-2.20)
Sidechain outliers	177891	6561 (2.20-2.20)
RSRZ outliers	164620	5791 (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% 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	895	91% 9%
1	B	895	89% 10% •
2	C	4	25% 75%

WORLDWIDE PDB
PROTEIN DATA BANK

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 28675 atoms, of which 13782 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 Glycosyl hydrolase family 98.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
1	A	895	Total 14013	C 4541	H 6865	N 1189	O 1393	P 1	S 24	214	5	0
1	B	895	Total 13959	C 4525	H 6837	N 1184	O 1389	P 1	S 23	214	0	0

- Molecule 2 is an oligosaccharide called beta-D-xylopyranose-(1-3)-beta-D-xylopyranose-(1-4)-beta-D-xylopyranose-(1-4)-beta-D-xylopyranose.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	4	Total 67	C 20	H 30	O 17	6	0	0

- Molecule 3 is alpha-L-arabinofuranose (CCD ID: AHR) (formula: $C_5H_{10}O_5$) (labeled as "Ligand of Interest" by depositor).
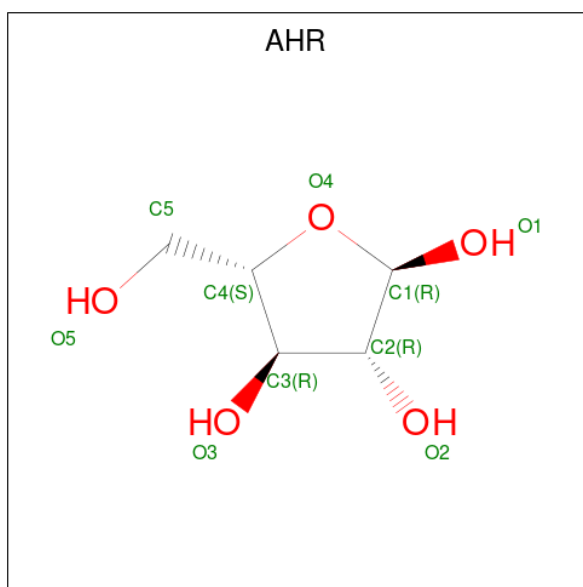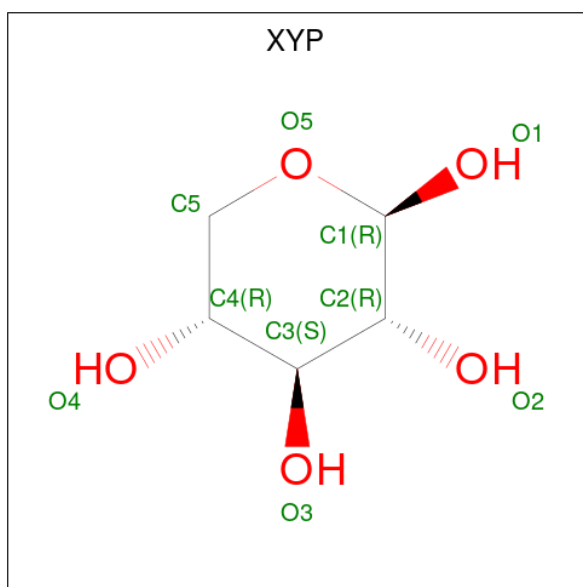
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	H	O	8	1
			40	10	20	10		
3	B	1	Total	C	H	O	8	1
			40	10	20	10		

- Molecule 4 is CALCIUM ION (CCD ID: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Ca	0	0
			1	1		
4	B	1	Total	Ca	0	0
			1	1		

- Molecule 5 is beta-D-xylopyranose (CCD ID: XYP) (formula: $C_5H_{10}O_5$) (labeled as "Ligand of Interest" by depositor).

XYP

O5
O
C5
O1
OH
C1(R)
C4(R) C2(R)
C3(S)
HO
O4
OH
O2
OH
O3

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	B	1	Total 20	C 5	H 10	O 5	4	0

- Molecule 6 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	310	Total 310	O 310	0	0
6	B	224	Total 224	O 224	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 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.

• Molecule 1: Glycosyl hydrolase family 98

Chain A:

• Molecule 1: Glycosyl hydrolase family 98

Chain B:
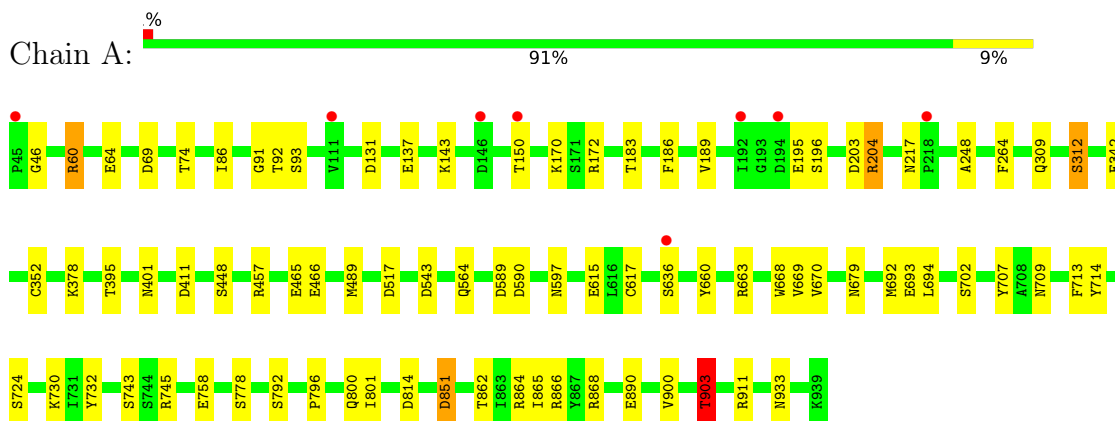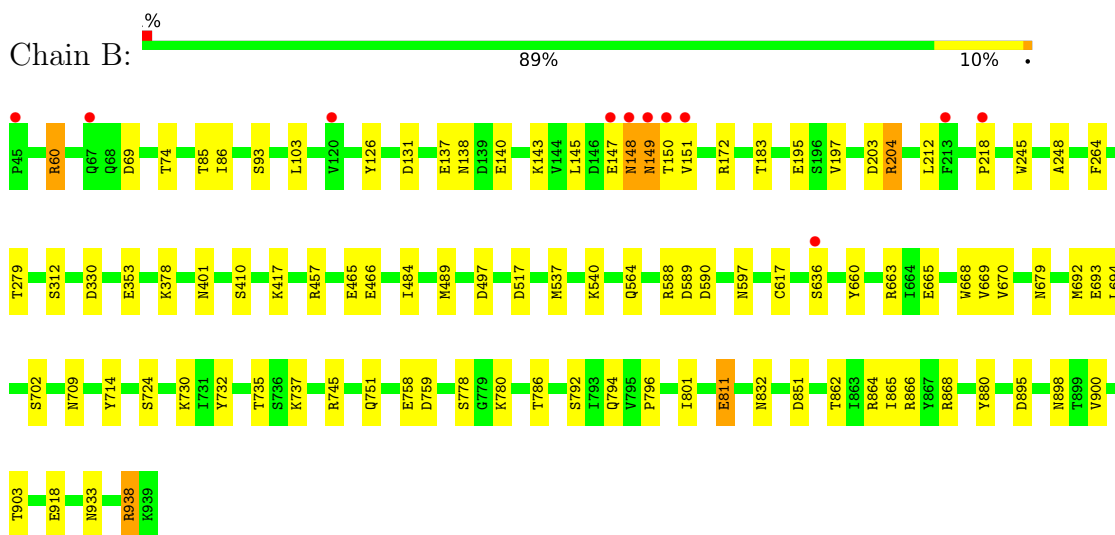
• Molecule 2: beta-D-xylopyranose-(1-3)-beta-D-xylopyranose-(1-4)-beta-D-xylopyranose-(1-4)-beta-D-xylopyranose

Chain C:

4 Data and refinement statistics ⓘ

Property	Value
Space group	C 1 2 1
Cell constants a, b, c, α, β, γ	186.05Å 63.85Å 180.49Å 90.00° 118.69° 90.00°
Resolution (Å)	59.18 – 2.20 59.18 – 2.20
% Data completeness (in resolution range)	99.9 (59.18-2.20) 99.7 (59.18-2.20)
R_{merge}	(Not available)
R_{sym}	(Not available)
$< I/\sigma(I) >$ [1]	1.71 (at 2.20Å)
Refinement program	REFMAC 5.8.0430 (refmacat 0.4.88), REFMAC 5.8.0430 (refmacat 0.4.
R, R_{free}	0.192 , 0.245 0.192 , 0.245
R_{free} test set	4758 reflections (5.02%)
Wilson B-factor (Å^2)	35.6
Anisotropy	0.418
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 25.6
L-test for twinning[2]	$< |L| > = 0.49$, $< L^2 > = 0.32$
Estimated twinning fraction	0.006 for h,-k,-h-l
F_o,F_c correlation	0.95
Total number of atoms	28675
Average B, all atoms (Å^2)	41.0

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 21.04 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 7.6382e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5   Model quality ⓘ

5.1   Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: PHD, CA, XYP, AHR

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	0.64	0/7336	1.11	19/9974 (0.2%)
1	B	0.61	0/7289	1.10	20/9914 (0.2%)
All	All	0.62	0/14625	1.10	39/19888 (0.2%)

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	8
1	B	0	8
All	All	0	16

There are no bond length outliers.

All (39) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	811	GLU	CB-CA-C	-9.24	89.97	110.21
1	B	898	ASN	CA-CB-CG	8.05	120.65	112.60
1	A	312	SER	CA-C-N	-7.79	106.14	121.41
1	A	312	SER	C-N-CA	-7.79	106.14	121.41
1	A	143	LYS	N-CA-CB	-7.56	98.64	110.06
1	B	204	ARG	CB-CA-C	-7.18	99.72	110.06
1	A	617	CYS	CB-CA-C	-7.14	95.33	109.67
1	A	758	GLU	CB-CA-C	6.80	123.96	110.42
1	A	758	GLU	CB-CG-CD	6.68	123.95	112.60
1	A	143	LYS	CB-CA-C	6.49	120.50	109.53
1	A	204	ARG	CB-CA-C	-6.32	98.27	110.35
1	A	590	ASP	CA-CB-CG	6.32	118.92	112.60

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	811	GLU	CB-CG-CD	6.26	123.25	112.60
1	A	890	GLU	CB-CG-CD	6.16	123.07	112.60
1	B	617	CYS	CB-CA-C	-6.06	98.01	110.38
1	B	895	ASP	CA-CB-CG	6.03	118.63	112.60
1	B	794	GLN	CB-CA-C	5.84	119.85	110.16
1	A	814	ASP	CA-CB-CG	5.82	118.42	112.60
1	A	862	THR	CA-CB-OG1	-5.81	100.88	109.60
1	B	330	ASP	CA-CB-CG	5.80	118.40	112.60
1	B	183	THR	CA-CB-OG1	-5.69	101.06	109.60
1	A	183	THR	CA-CB-OG1	-5.64	101.14	109.60
1	A	411	ASP	CB-CA-C	5.61	119.76	110.90
1	B	590	ASP	CA-CB-CG	5.58	118.18	112.60
1	B	862	THR	CA-CB-OG1	-5.57	101.24	109.60
1	A	395	THR	CA-CB-OG1	-5.50	101.35	109.60
1	B	786	THR	CA-CB-OG1	-5.39	101.52	109.60
1	A	903	THR	OG1-CB-CG2	-5.38	98.55	109.30
1	B	279	THR	CA-CB-OG1	-5.38	101.53	109.60
1	B	759	ASP	CB-CA-C	5.38	118.94	111.63
1	A	713	PHE	CA-CB-CG	5.35	119.15	113.80
1	B	517	ASP	CB-CA-C	5.24	119.79	111.51
1	B	751	GLN	CB-CA-C	5.23	118.70	109.80
1	A	543	ASP	CA-CB-CG	5.19	117.79	112.60
1	B	218	PRO	CB-CA-C	5.16	117.19	111.56
1	A	589	ASP	CA-CB-CG	5.08	117.68	112.60
1	B	589	ASP	CA-CB-CG	5.08	117.68	112.60
1	B	417	LYS	CB-CA-C	-5.01	103.01	110.88
1	B	497	ASP	CA-CB-CG	5.01	117.61	112.60

There are no chirality outliers.

All (16) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	172	ARG	Sidechain
1	A	195	GLU	Peptide
1	A	248	ALA	Peptide
1	A	457	ARG	Sidechain
1	A	60[A]	ARG	Sidechain
1	A	745	ARG	Sidechain
1	A	864	ARG	Sidechain
1	A	911	ARG	Sidechain
1	B	145	LEU	Peptide
1	B	172	ARG	Sidechain

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	B	195	GLU	Peptide
1	B	248	ALA	Peptide
1	B	457	ARG	Sidechain
1	B	588	ARG	Sidechain
1	B	745	ARG	Sidechain
1	B	864	ARG	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	7148	6865	6828	30	1
1	B	7122	6837	6824	39	1
2	C	37	30	0	0	0
3	A	20	20	0	1	0
3	B	20	20	0	3	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
5	B	10	10	0	0	0
6	A	310	0	0	3	0
6	B	224	0	0	4	0
All	All	14893	13782	13652	70	2

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 2.

All (70) 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)
3:B:1002[A]:AHR:O3	6:B:1101:HOH:O	1.95	0.85
1:A:489:MET:HE2	1:A:669:VAL:HG23	1.63	0.79
1:B:735:THR:HG21	1:B:780:LYS:HG3	1.63	0.79
1:B:489:MET:HE2	1:B:669:VAL:HG23	1.68	0.75
1:B:149:ASN:HB2	1:B:151:VAL:HG22	1.77	0.67
1:A:60[A]:ARG:NH2	1:A:64:GLU:OE1	2.27	0.67
1:B:60:ARG:HG2	1:B:60:ARG:HH11	1.60	0.66

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:709:ASN:OD1	1:A:778:SER:HB3	2.00	0.62
1:A:131:ASP:OD1	1:A:203:ASP:O	2.19	0.61
1:B:131:ASP:OD1	1:B:203:ASP:O	2.19	0.60
1:B:709:ASN:OD1	1:B:778:SER:HB3	2.02	0.59
1:A:309[B]:GLN:OE1	3:A:1001[B]:AHR:O3	2.22	0.56
1:B:312:SER:OG	1:B:353:GLU:OE2	2.23	0.56
1:A:851:PHD:P	6:A:1102:HOH:O	2.65	0.55
1:B:245:TRP:CD1	3:B:1002[A]:AHR:C5	2.90	0.55
1:A:264:PHE:CE1	1:A:796:PRO:HD3	2.42	0.55
1:B:203:ASP:O	1:B:204:ARG:HB2	2.06	0.54
1:B:60:ARG:HG2	1:B:60:ARG:NH1	2.24	0.52
1:A:866:ARG:NH1	6:A:1106:HOH:O	2.42	0.52
1:A:203:ASP:O	1:A:204:ARG:HB2	2.09	0.52
1:B:60:ARG:HH11	1:B:60:ARG:CG	2.23	0.50
1:A:865:ILE:O	1:A:903:THR:HA	2.12	0.49
1:B:126:TYR:CE2	1:B:143:LYS:HE3	2.47	0.49
1:B:138:ASN:HB2	1:B:140:GLU:HB2	1.94	0.49
1:A:679:ASN:HA	1:A:694:LEU:O	2.13	0.48
1:B:665:GLU:HG3	6:B:1318:HOH:O	2.14	0.48
1:B:865:ILE:O	1:B:903:THR:HA	2.14	0.47
1:B:245:TRP:HD1	3:B:1002[A]:AHR:C5	2.27	0.47
1:B:484:ILE:HG21	1:B:537:MET:HE1	1.97	0.47
1:B:679:ASN:HA	1:B:694:LEU:O	2.14	0.47
1:B:692:MET:HA	1:B:730:LYS:O	2.15	0.46
1:A:465:GLU:O	1:A:466:GLU:HB2	2.15	0.46
1:A:660:TYR:O	1:A:670:VAL:HA	2.15	0.46
1:B:660:TYR:O	1:B:670:VAL:HA	2.16	0.45
1:A:933:ASN:C	1:A:933:ASN:OD1	2.58	0.45
1:B:465:GLU:O	1:B:466:GLU:HB2	2.16	0.45
1:B:103:LEU:HD12	1:B:103:LEU:HA	1.90	0.45
1:B:938:ARG:HH11	1:B:938:ARG:HG3	1.82	0.45
1:A:91:GLY:O	1:A:903:THR:HG23	2.17	0.45
1:B:86:ILE:O	1:B:93:SER:HA	2.18	0.44
1:A:60[A]:ARG:NH2	1:A:64:GLU:CD	2.76	0.44
1:A:203:ASP:O	1:A:204:ARG:CB	2.62	0.44
1:A:93:SER:CB	1:A:866:ARG:HH12	2.31	0.44
1:B:264:PHE:CE2	1:B:796:PRO:HD3	2.53	0.44
1:A:86:ILE:O	1:A:93:SER:HA	2.18	0.44
1:A:692:MET:HA	1:A:730:LYS:O	2.17	0.44
1:B:401:ASN:OD1	1:B:401:ASN:C	2.61	0.43
1:B:665:GLU:CG	6:B:1318:HOH:O	2.65	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:933:ASN:OD1	1:B:933:ASN:C	2.61	0.43
1:B:702:SER:HA	1:B:714:TYR:O	2.19	0.43
1:B:868:ARG:HD2	1:B:900:VAL:O	2.18	0.43
1:B:93:SER:CB	1:B:866:ARG:HH12	2.32	0.43
1:A:868:ARG:HD2	1:A:900:VAL:O	2.19	0.42
1:A:702:SER:HA	1:A:714:TYR:O	2.19	0.42
1:A:186:PHE:HZ	1:A:189:VAL:HG23	1.85	0.42
1:A:707:TYR:OH	6:A:1101:HOH:O	2.20	0.42
1:A:693:GLU:OE1	1:A:732:TYR:OH	2.38	0.42
1:B:148:ASN:OD1	1:B:148:ASN:N	2.50	0.41
1:B:203:ASP:O	1:B:204:ARG:CB	2.62	0.41
1:B:85:THR:CG2	1:B:832:ASN:OD1	2.68	0.41
1:B:663:ARG:HD3	1:B:668:TRP:CZ2	2.55	0.41
1:A:312:SER:O	1:A:352:CYS:O	2.38	0.41
1:B:197:VAL:CG2	1:B:212:LEU:HD13	2.50	0.41
1:A:401:ASN:OD1	1:A:401:ASN:C	2.63	0.41
1:A:663:ARG:HD3	1:A:668:TRP:CZ2	2.55	0.41
1:A:448:SER:HB2	1:A:615:GLU:O	2.21	0.40
1:B:540:LYS:NZ	6:B:1125:HOH:O	2.54	0.40
1:B:693:GLU:OE1	1:B:732:TYR:OH	2.38	0.40
1:B:880:TYR:HB2	1:B:918:GLU:HB2	2.03	0.40

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:378:LYS:HZ1	1:B:792:SER:HG[4_545]	1.20	0.40
1:A:378:LYS:HZ1	1:A:792:SER:HG[4_546]	1.27	0.33

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	897/895 (100%)	861 (96%)	35 (4%)	1 (0%)	48	57
1	B	892/895 (100%)	855 (96%)	37 (4%)	0	100	100
All	All	1789/1790 (100%)	1716 (96%)	72 (4%)	1 (0%)	48	57

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	46	GLY

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	778/773 (101%)	761 (98%)	17 (2%)	47	61
1	B	773/773 (100%)	755 (98%)	18 (2%)	45	59
All	All	1551/1546 (100%)	1516 (98%)	35 (2%)	45	59

All (35) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	69	ASP
1	A	74	THR
1	A	137	GLU
1	A	150	THR
1	A	170	LYS
1	A	196	SER
1	A	217	ASN
1	A	342	GLU
1	A	517	ASP
1	A	564	GLN
1	A	597	ASN
1	A	636	SER
1	A	724	SER
1	A	743	SER
1	A	800	GLN

Continued from previous page...

Mol	Chain	Res	Type
1	A	801	ILE
1	A	903	THR
1	B	60	ARG
1	B	69	ASP
1	B	74	THR
1	B	137	GLU
1	B	147	GLU
1	B	148	ASN
1	B	149	ASN
1	B	150	THR
1	B	410	SER
1	B	564	GLN
1	B	597	ASN
1	B	636	SER
1	B	724	SER
1	B	737	LYS
1	B	758	GLU
1	B	801	ILE
1	B	811	GLU
1	B	938	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	235	GLN
1	A	469	GLN
1	A	641	ASN
1	A	679	ASN
1	A	776	ASN
1	A	817	ASN
1	B	422	ASN
1	B	523	ASN
1	B	679	ASN
1	B	776	ASN
1	B	817	ASN
1	B	898	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues 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	$\#|Z| > 2$
1	PHD	B	851	1	9,11,12	1.75	2 (22%)	10,15,17	2.02	3 (30%)
1	PHD	A	851	1	9,11,12	2.19	2 (22%)	10,15,17	1.07	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
1	PHD	B	851	1	-	2/8/11/13	-
1	PHD	A	851	1	-	4/8/11/13	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	851	PHD	P-OD1	5.23	1.67	1.59
1	B	851	PHD	P-OD1	3.54	1.64	1.59
1	A	851	PHD	P-OP1	3.39	1.61	1.50
1	B	851	PHD	P-OP1	2.79	1.59	1.50

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	851	PHD	OD1-P-OP1	-4.49	93.37	109.32
1	B	851	PHD	OP2-P-OD1	2.80	113.80	105.25
1	B	851	PHD	OP3-P-OD1	2.55	113.02	105.25

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	851	PHD	CA-CB-CG-OD2
1	B	851	PHD	CA-CB-CG-OD2
1	A	851	PHD	CA-CB-CG-OD1
1	B	851	PHD	CA-CB-CG-OD1
1	A	851	PHD	N-CA-CB-CG
1	A	851	PHD	C-CA-CB-CG

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	851	PHD	1	0

5.5 Carbohydrates ⓘ

4 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	XYP	C	1	2	10,10,10	0.78	0	14,14,14	1.18	2 (14%)
2	XYP	C	2	2	9,9,10	0.71	0	10,12,14	1.06	1 (10%)
2	XYP	C	3	2	9,9,10	0.57	0	10,12,14	1.43	2 (20%)
2	XYP	C	4	2	9,9,10	0.52	0	10,12,14	1.08	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	XYP	C	1	2	-	-	0/1/1/1
2	XYP	C	2	2	-	-	0/1/1/1
2	XYP	C	3	2	-	-	0/1/1/1
2	XYP	C	4	2	-	-	0/1/1/1

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	3	XYP	O3-C3-C2	-3.32	103.64	109.99
2	C	2	XYP	C1-C2-C3	2.62	112.89	109.67
2	C	3	XYP	O2-C2-C1	2.27	113.79	109.15
2	C	1	XYP	O3-C3-C2	2.26	115.57	110.35
2	C	1	XYP	O1-C1-C2	2.06	114.83	109.03

There are no chirality outliers.

There are no torsion outliers.

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.
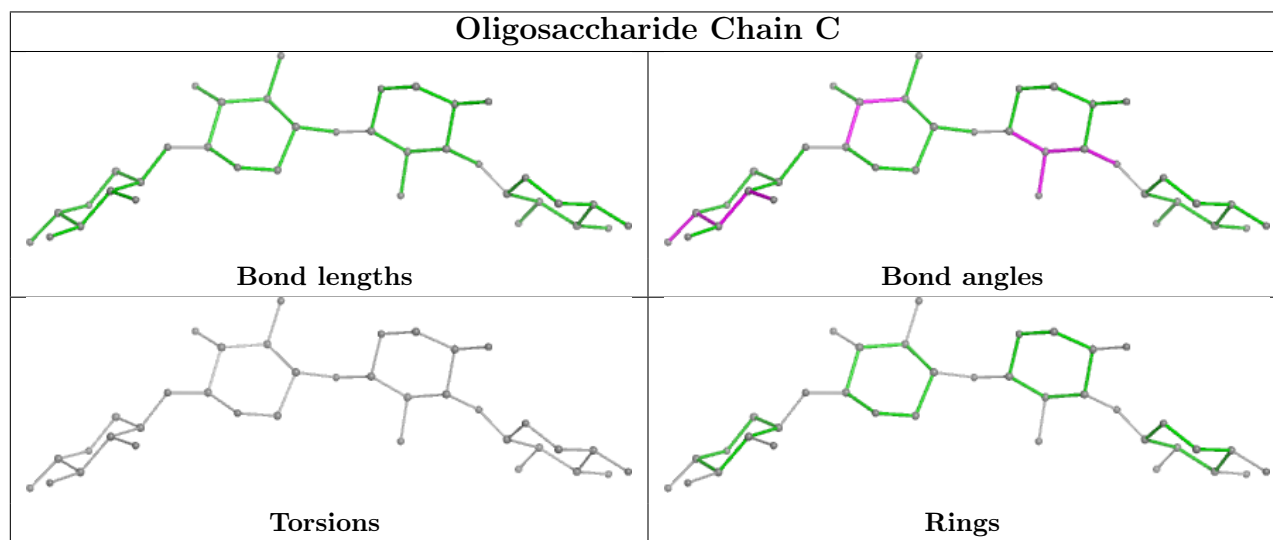
5.6 Ligand geometry ⓘ

Of 7 ligands modelled in this entry, 2 are monoatomic - leaving 5 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
3	AHR	B	1002[A]	-	10,10,10	1.29	2 (20%)	13,14,14	1.66	2 (15%)
3	AHR	B	1002[B]	-	10,10,10	0.46	0	13,14,14	1.11	1 (7%)
5	XYP	B	1001	-	10,10,10	0.47	0	14,14,14	1.32	2 (14%)
3	AHR	A	1001[B]	-	10,10,10	0.80	0	13,14,14	1.93	2 (15%)
3	AHR	A	1001[A]	-	10,10,10	1.72	1 (10%)	13,14,14	1.93	6 (46%)

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	AHR	B	1002[A]	-	-	2/2/18/18	0/1/1/1
3	AHR	B	1002[B]	-	-	0/2/18/18	0/1/1/1
5	XYP	B	1001	-	-	-	0/1/1/1
3	AHR	A	1001[B]	-	-	2/2/18/18	0/1/1/1
3	AHR	A	1001[A]	-	-	2/2/18/18	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1001[A]	AHR	O4-C1	4.76	1.49	1.43
3	B	1002[A]	AHR	O4-C1	3.46	1.47	1.43
3	B	1002[A]	AHR	C1-C2	2.05	1.55	1.52

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1001[B]	AHR	C1-C2-C3	-5.23	95.75	102.30
3	B	1002[A]	AHR	C1-C2-C3	-3.60	97.80	102.30
3	A	1001[A]	AHR	O3-C3-C2	-3.49	100.52	111.82
3	A	1001[A]	AHR	C1-C2-C3	-3.39	98.05	102.30
5	B	1001	XYP	O1-C1-C2	3.09	117.73	109.03
3	B	1002[A]	AHR	O2-C2-C1	3.06	120.28	111.82
3	A	1001[B]	AHR	O4-C1-C2	-2.88	100.93	104.46
3	A	1001[A]	AHR	O3-C3-C4	2.81	119.17	111.05
3	A	1001[A]	AHR	O1-C1-O4	2.48	114.31	111.13
3	B	1002[B]	AHR	C1-C2-C3	2.26	105.13	102.30
5	B	1001	XYP	O2-C2-C1	2.25	114.39	109.16
3	A	1001[A]	AHR	O2-C2-C1	2.16	117.78	111.82

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	1001[A]	AHR	O2-C2-C3	2.03	118.39	111.82

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1001[B]	AHR	O4-C4-C5-O5
3	A	1001[B]	AHR	C3-C4-C5-O5
3	B	1002[A]	AHR	O4-C4-C5-O5
3	B	1002[A]	AHR	C3-C4-C5-O5
3	A	1001[A]	AHR	C3-C4-C5-O5
3	A	1001[A]	AHR	O4-C4-C5-O5

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	1002[A]	AHR	3	0
3	A	1001[B]	AHR	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 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.

Ligand AHR B 1002 (A)

Bond lengths

Bond angles

Torsions

Rings

Ligand AHR B 1002 (B)

Bond lengths

Bond angles

Torsions

Rings

Ligand XYP B 1001

Bond lengths

Bond angles

Torsions

Rings

Ligand AHR A 1001 (B)

Bond lengths

Bond angles

Torsions

Rings

Ligand AHR A 1001 (A)

Bond lengths

Bond angles

Torsions

Rings
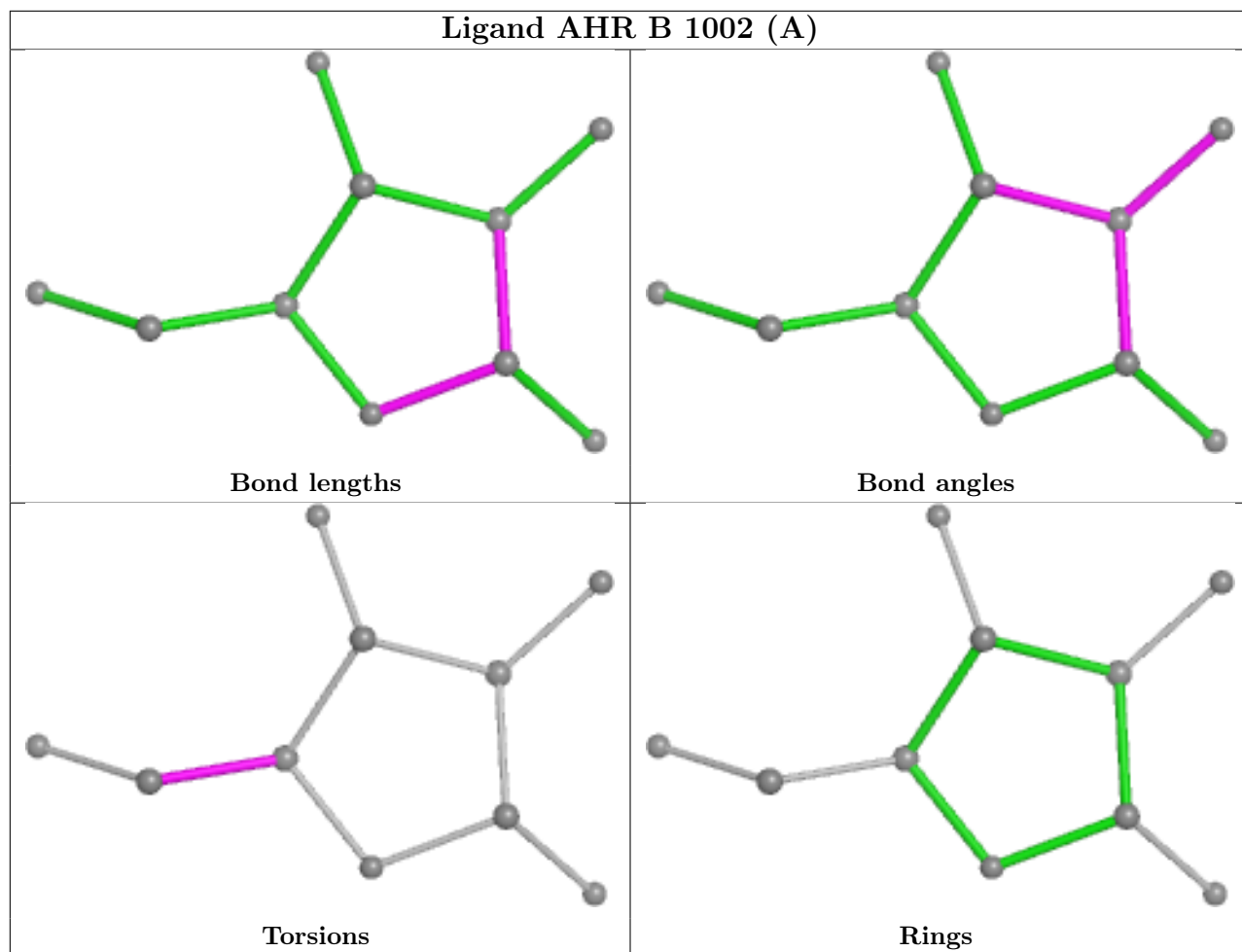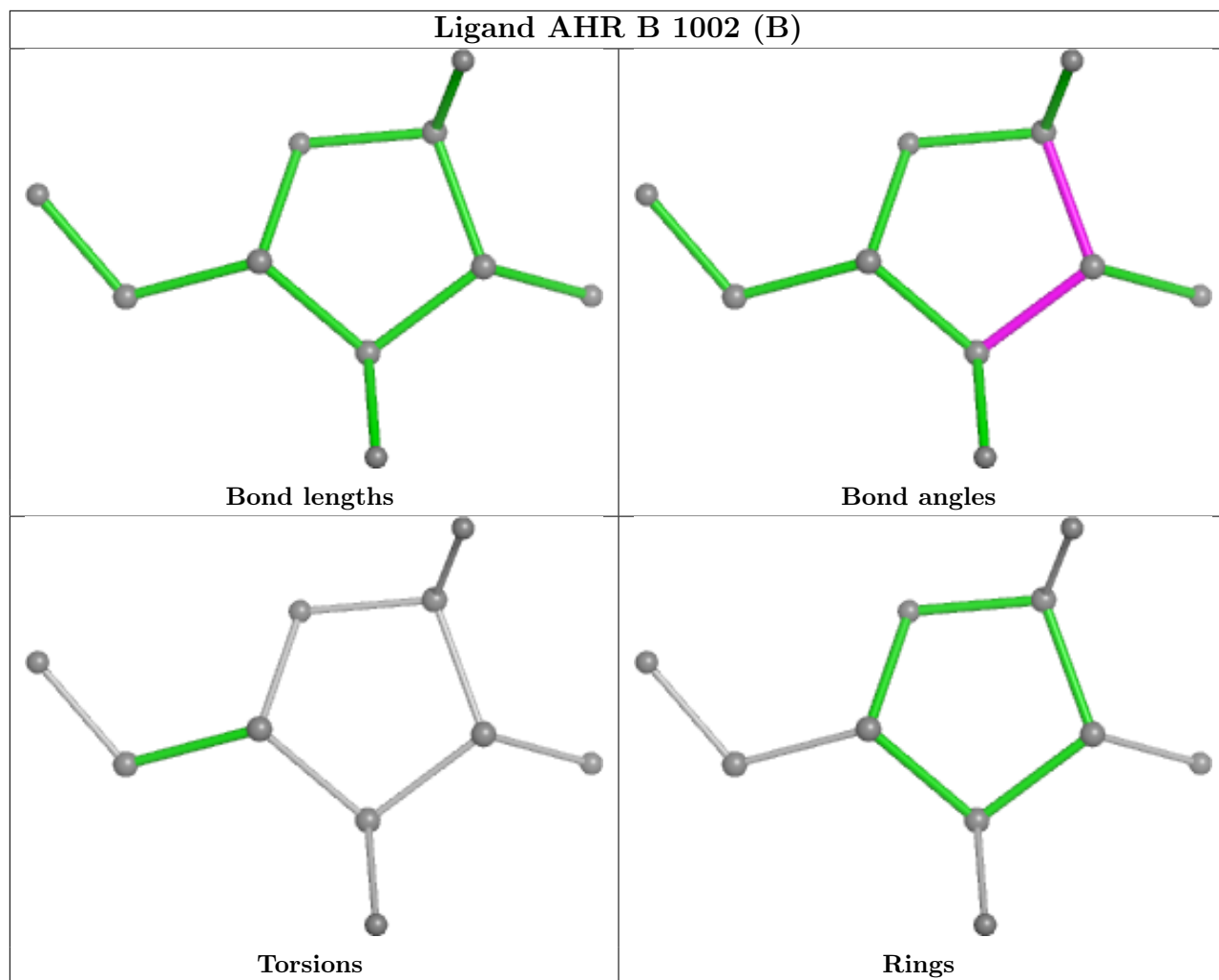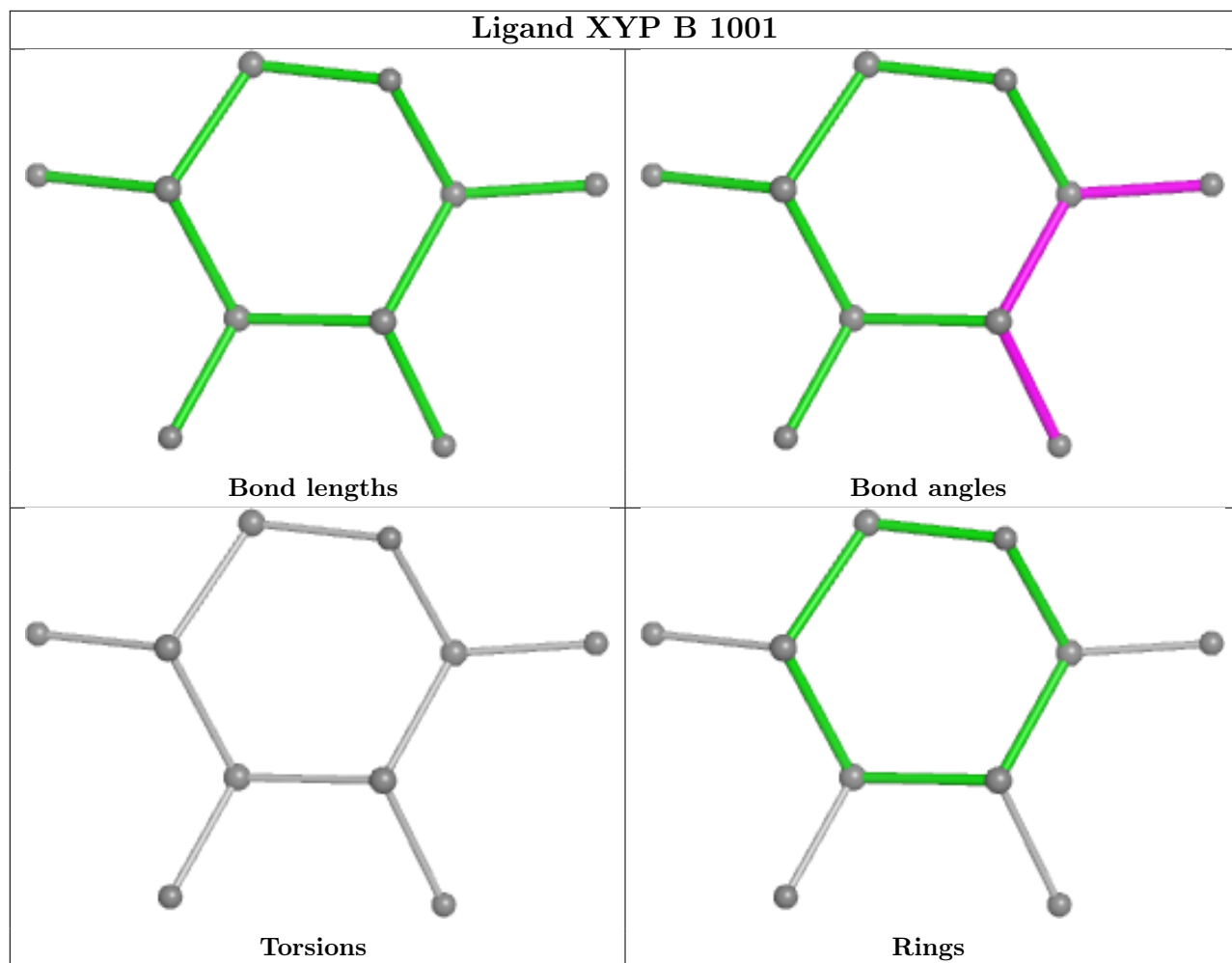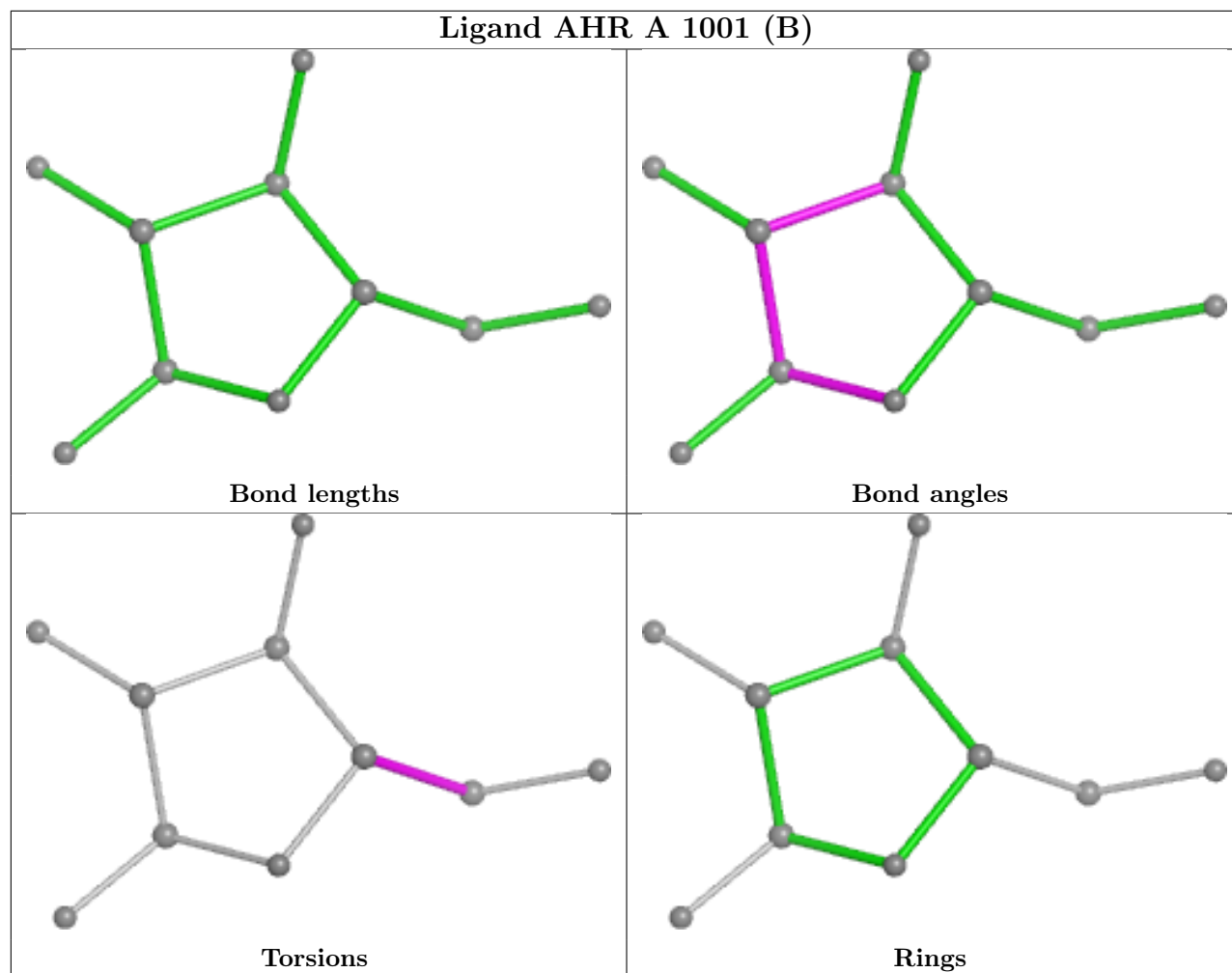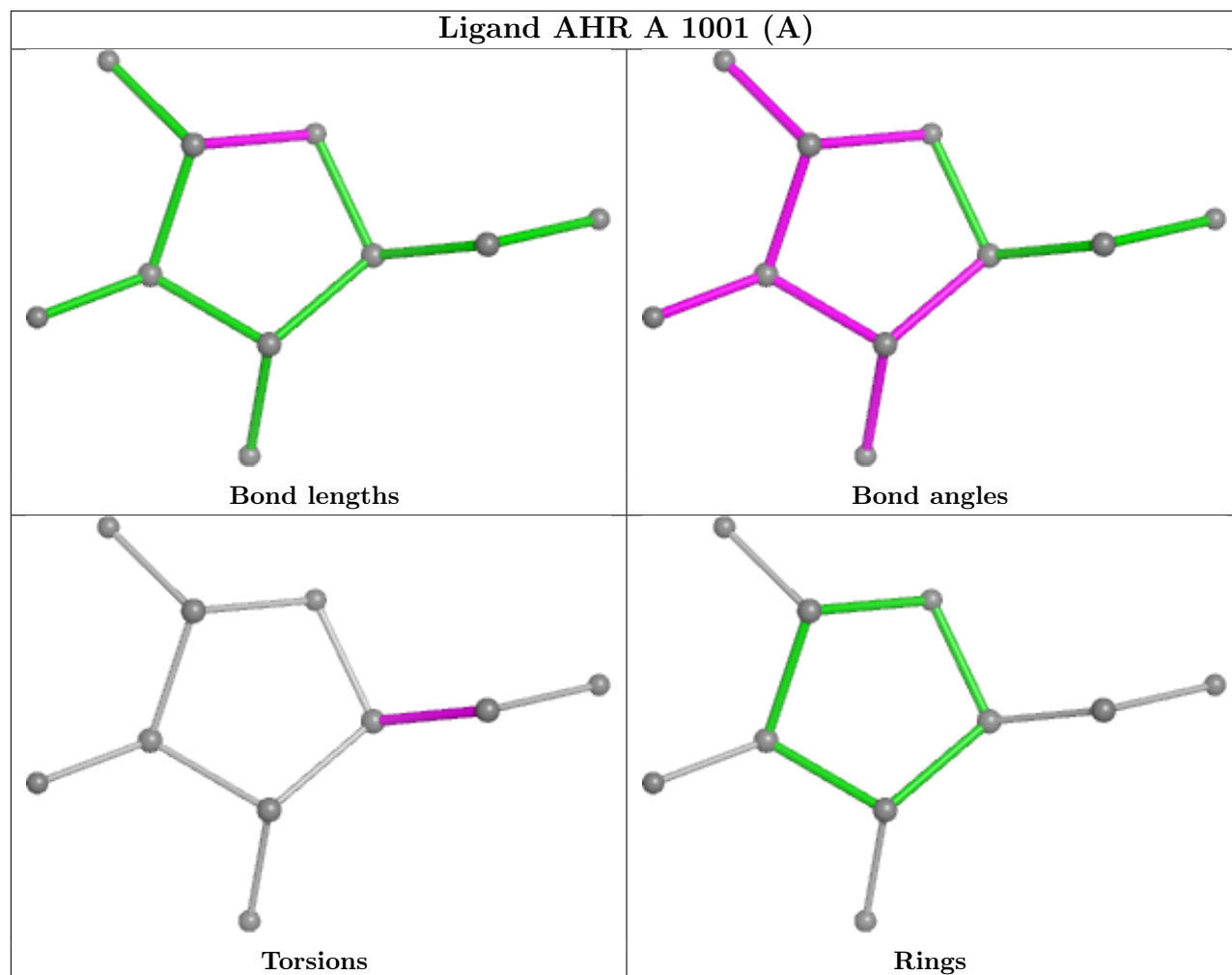
5.7 Other polymers ⓘ

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 ⓘ

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, 95th 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(Å²)	Q<0.9
1	A	894/895 (99%)	-0.28	8 (0%)	81	78	10, 34, 63, 129	3 (0%)
1	B	894/895 (99%)	-0.02	11 (1%)	76	73	23, 41, 73, 140	0
All	All	1788/1790 (99%)	-0.15	19 (1%)	77	75	10, 38, 69, 140	3 (0%)

All (19) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	45	PRO	5.7
1	B	151	VAL	4.0
1	A	150	THR	3.9
1	B	218	PRO	3.4
1	B	150	THR	3.1
1	B	148	ASN	3.0
1	A	218	PRO	2.8
1	A	192	ILE	2.7
1	B	213	PHE	2.7
1	A	111	VAL	2.6
1	B	636	SER	2.5
1	A	636	SER	2.5
1	B	147	GLU	2.5
1	B	120	VAL	2.5
1	B	45	PRO	2.4
1	A	194	ASP	2.4
1	B	67	GLN	2.2
1	A	146	ASP	2.2
1	B	149	ASN	2.1

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

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, 95th 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^2)	Q<0.9
1	PHD	B	851	12/13	0.85	0.11	38,41,46,55	0
1	PHD	A	851	12/13	0.88	0.09	34,35,43,50	0

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, 95th 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^2)	Q<0.9
2	XYP	C	1	10/10	0.80	0.11	30,66,70,76	2
2	XYP	C	2	9/10	0.82	0.12	30,66,72,74	1
2	XYP	C	4	9/10	0.84	0.14	30,65,70,73	3
2	XYP	C	3	9/10	0.89	0.09	55,59,61,63	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.
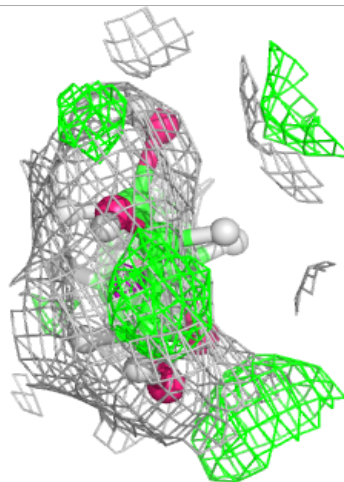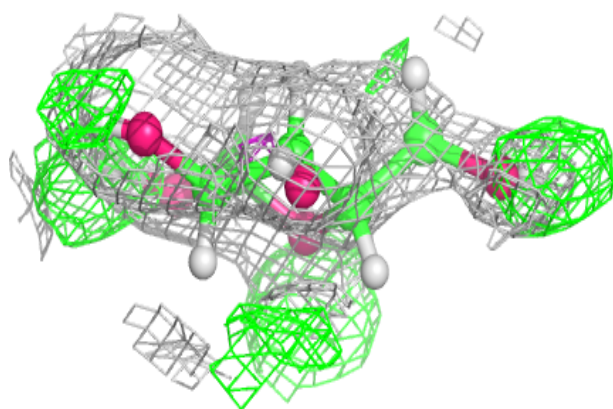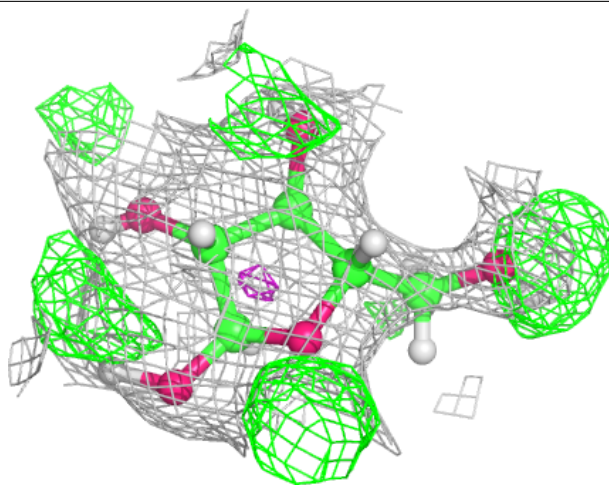

Electron density around Chain C:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)
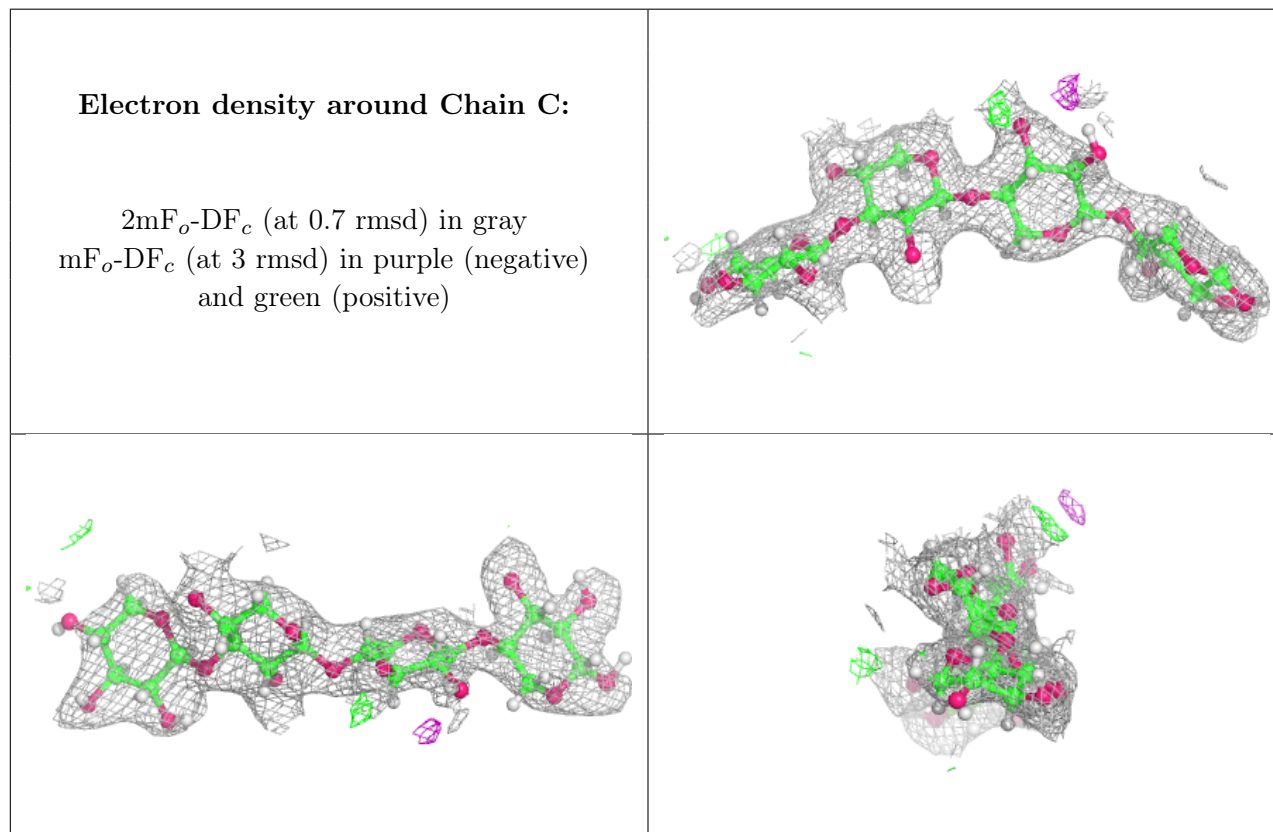
6.4 Ligands ⓘ

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, 95th 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^2)	Q<0.9
3	AHR	A	1001[A]	10/10	0.80	0.13	27,30,31,32	20
3	AHR	A	1001[B]	10/10	0.80	0.13	19,21,30,30	20
3	AHR	B	1002[A]	10/10	0.88	0.14	30,39,44,44	20
3	AHR	B	1002[B]	10/10	0.88	0.14	20,21,30,30	20
5	XYP	B	1001	10/10	0.88	0.11	30,64,67,73	4
4	CA	B	1003	1/1	0.99	0.02	35,35,35,35	0
4	CA	A	1002	1/1	0.99	0.03	27,27,27,27	0

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.

Electron density around AHR A 1001 (A):

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
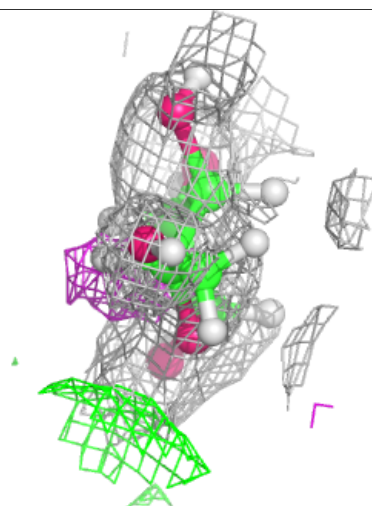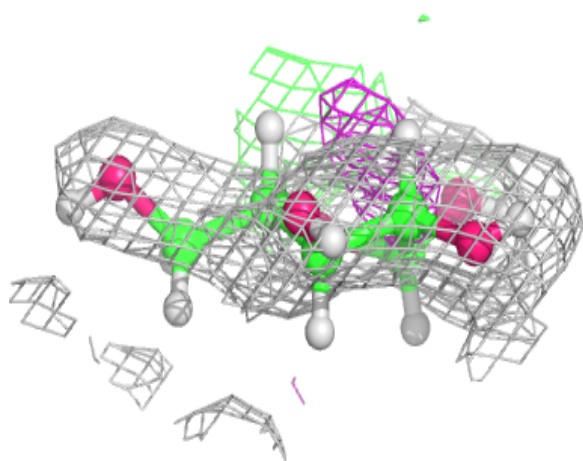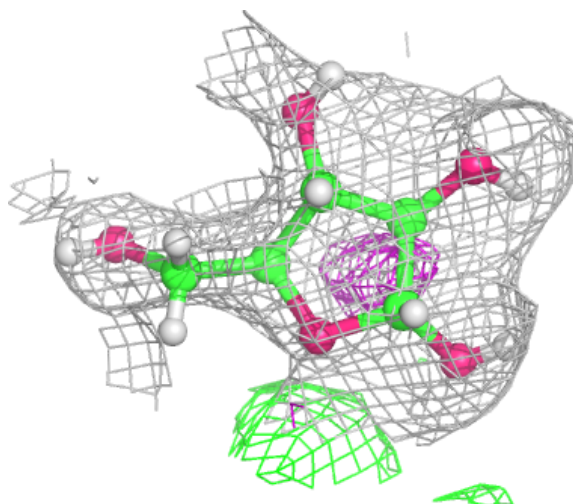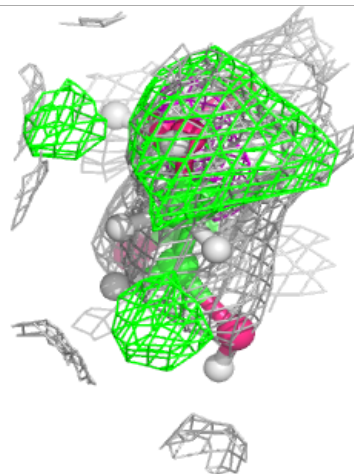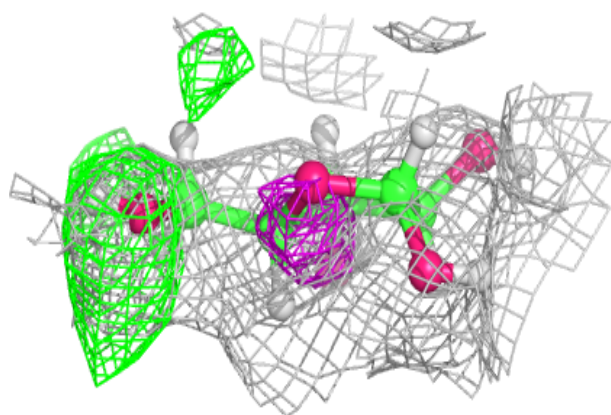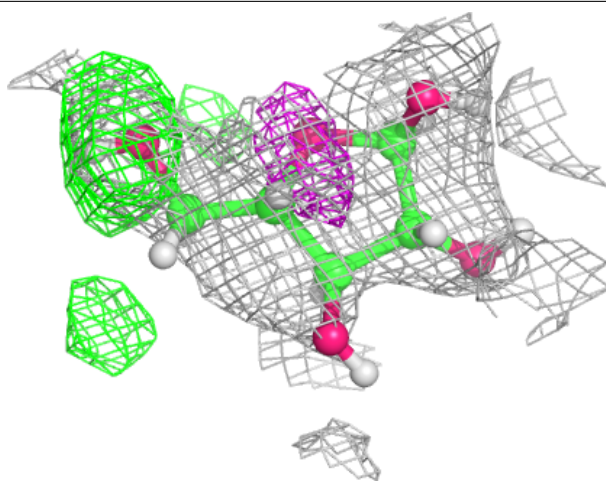and green (positive)
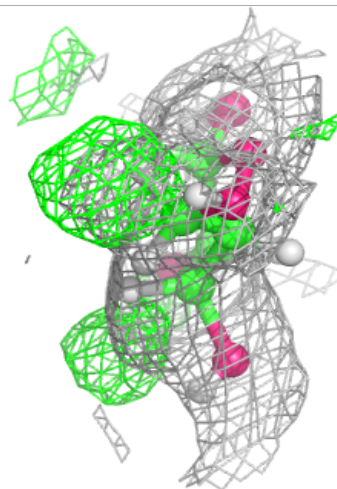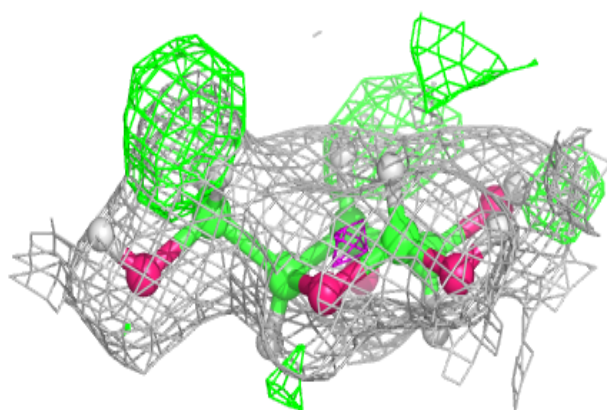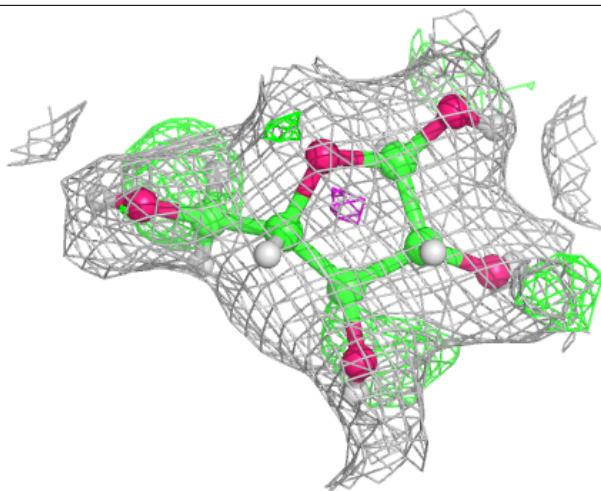
Electron density around AHR A 1001 (B):

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)

Electron density around AHR B 1002 (A):

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
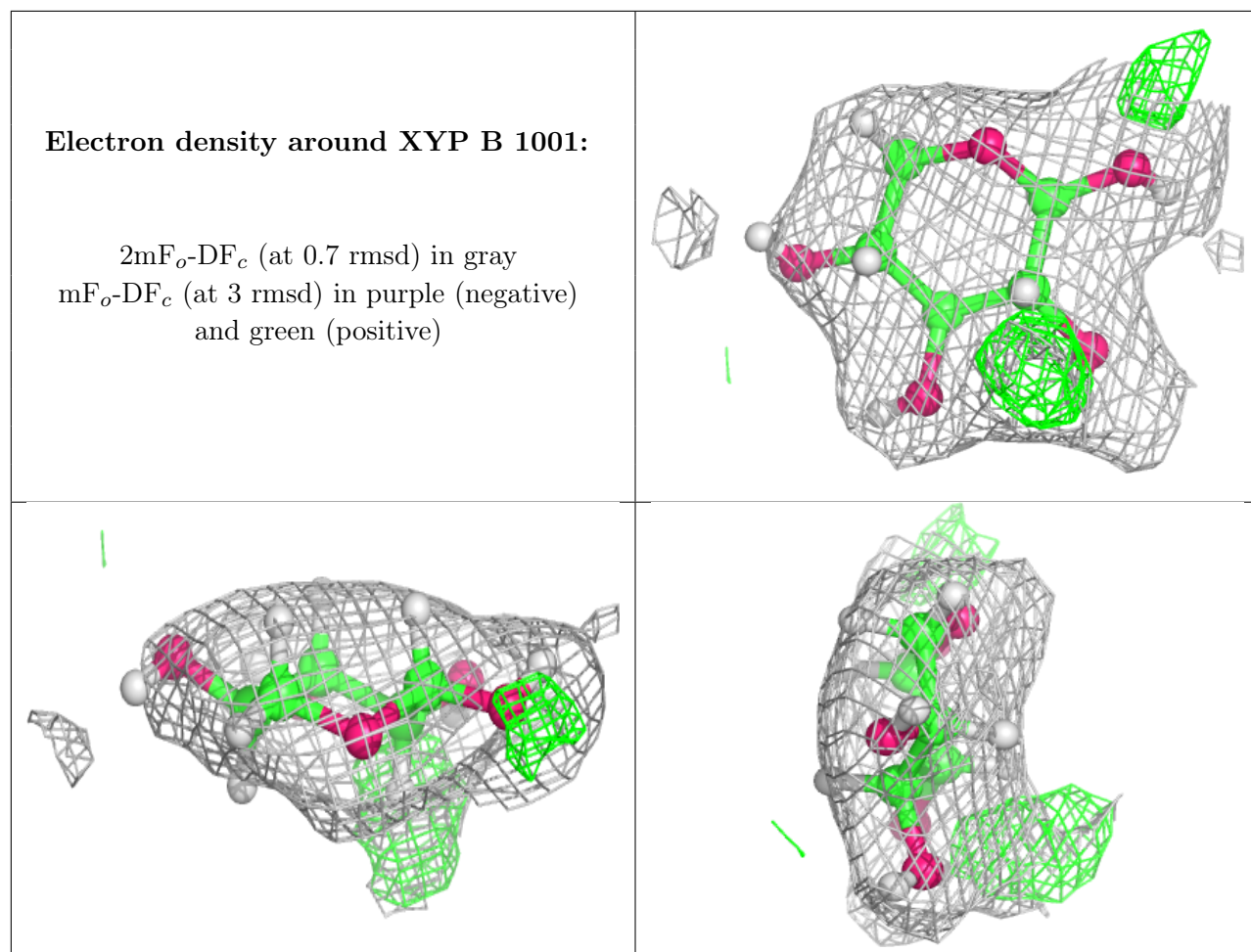
Electron density around AHR B 1002 (B):

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)

Electron density around XYP B 1001:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

6.5 Other polymers ⓘ

There are no such residues in this entry.