



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 27, 2021 – 10:13 AM JST

PDB ID : 7CJS
Title : structure of aquaporin
Authors : Saitoh, Y.; Ma, J.F.; Suga, M.
Deposited on : 2020-07-13
Resolution : 1.80 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

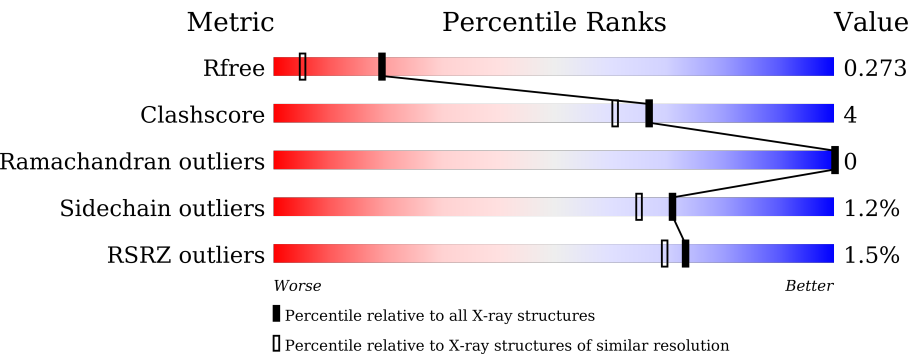
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.80 Å.

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





Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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 $\leq 5\%$. 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	254	<div><div>%</div><div><div></div><div>76%</div><div>7%</div><div>16%</div></div></div>
1	B	254	<div><div>%</div><div><div></div><div>78%</div><div>7%</div><div>14%</div></div></div>
1	C	254	<div><div>%</div><div><div></div><div>78%</div><div>6%</div><div>16%</div></div></div>
1	D	254	<div><div>%</div><div><div></div><div>79%</div><div>7%</div><div>14%</div></div></div>
1	E	254	<div><div>2%</div><div><div></div><div>76%</div><div>8%</div><div>16%</div></div></div>
1	F	254	<div><div>3%</div><div><div></div><div>79%</div><div>6%</div><div>14%</div></div></div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	254	 78% 7% 16%
1	H	254	 78% 7% 14%

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
3	BOG	A	302	-	-	-	X
4	PG4	B	303	-	-	-	X
5	Y01	D	304	-	-	-	X
5	Y01	H	304	-	-	-	X

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 27690 atoms, of which 13508 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 Aquaporin NIP2-1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	213	Total	C	H	N	O	S	0	2	0
			3199	1041	1618	261	271	8			
1	B	219	Total	C	H	N	O	S	0	3	0
			3312	1079	1673	270	282	8			
1	C	214	Total	C	H	N	O	S	0	3	0
			3235	1055	1638	262	272	8			
1	D	219	Total	C	H	N	O	S	0	2	0
			3306	1077	1670	270	281	8			
1	E	214	Total	C	H	N	O	S	0	2	0
			3227	1053	1633	262	271	8			
1	F	219	Total	C	H	N	O	S	0	3	0
			3311	1078	1673	270	282	8			
1	G	214	Total	C	H	N	O	S	0	3	0
			3242	1058	1643	262	271	8			
1	H	219	Total	C	H	N	O	S	0	2	0
			3306	1077	1670	270	281	8			

There are 72 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	45	MET	-	initiating methionine	UNP Q6Z2T3
A	46	ALA	-	expression tag	UNP Q6Z2T3
A	50	ARG	LYS	conflict	UNP Q6Z2T3
A	66	ALA	CYS	conflict	UNP Q6Z2T3
A	93	VAL	THR	conflict	UNP Q6Z2T3
A	139	ALA	CYS	conflict	UNP Q6Z2T3
A	232	ARG	LYS	conflict	UNP Q6Z2T3
A	253	VAL	THR	conflict	UNP Q6Z2T3
A	264	ARG	LYS	conflict	UNP Q6Z2T3
B	45	MET	-	initiating methionine	UNP Q6Z2T3
B	46	ALA	-	expression tag	UNP Q6Z2T3
B	50	ARG	LYS	conflict	UNP Q6Z2T3
B	66	ALA	CYS	conflict	UNP Q6Z2T3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	93	VAL	THR	conflict	UNP Q6Z2T3
B	139	ALA	CYS	conflict	UNP Q6Z2T3
B	232	ARG	LYS	conflict	UNP Q6Z2T3
B	253	VAL	THR	conflict	UNP Q6Z2T3
B	264	ARG	LYS	conflict	UNP Q6Z2T3
C	45	MET	-	initiating methionine	UNP Q6Z2T3
C	46	ALA	-	expression tag	UNP Q6Z2T3
C	50	ARG	LYS	conflict	UNP Q6Z2T3
C	66	ALA	CYS	conflict	UNP Q6Z2T3
C	93	VAL	THR	conflict	UNP Q6Z2T3
C	139	ALA	CYS	conflict	UNP Q6Z2T3
C	232	ARG	LYS	conflict	UNP Q6Z2T3
C	253	VAL	THR	conflict	UNP Q6Z2T3
C	264	ARG	LYS	conflict	UNP Q6Z2T3
D	45	MET	-	initiating methionine	UNP Q6Z2T3
D	46	ALA	-	expression tag	UNP Q6Z2T3
D	50	ARG	LYS	conflict	UNP Q6Z2T3
D	66	ALA	CYS	conflict	UNP Q6Z2T3
D	93	VAL	THR	conflict	UNP Q6Z2T3
D	139	ALA	CYS	conflict	UNP Q6Z2T3
D	232	ARG	LYS	conflict	UNP Q6Z2T3
D	253	VAL	THR	conflict	UNP Q6Z2T3
D	264	ARG	LYS	conflict	UNP Q6Z2T3
E	45	MET	-	initiating methionine	UNP Q6Z2T3
E	46	ALA	-	expression tag	UNP Q6Z2T3
E	50	ARG	LYS	conflict	UNP Q6Z2T3
E	66	ALA	CYS	conflict	UNP Q6Z2T3
E	93	VAL	THR	conflict	UNP Q6Z2T3
E	139	ALA	CYS	conflict	UNP Q6Z2T3
E	232	ARG	LYS	conflict	UNP Q6Z2T3
E	253	VAL	THR	conflict	UNP Q6Z2T3
E	264	ARG	LYS	conflict	UNP Q6Z2T3
F	45	MET	-	initiating methionine	UNP Q6Z2T3
F	46	ALA	-	expression tag	UNP Q6Z2T3
F	50	ARG	LYS	conflict	UNP Q6Z2T3
F	66	ALA	CYS	conflict	UNP Q6Z2T3
F	93	VAL	THR	conflict	UNP Q6Z2T3
F	139	ALA	CYS	conflict	UNP Q6Z2T3
F	232	ARG	LYS	conflict	UNP Q6Z2T3
F	253	VAL	THR	conflict	UNP Q6Z2T3
F	264	ARG	LYS	conflict	UNP Q6Z2T3
G	45	MET	-	initiating methionine	UNP Q6Z2T3

Continued on next page...

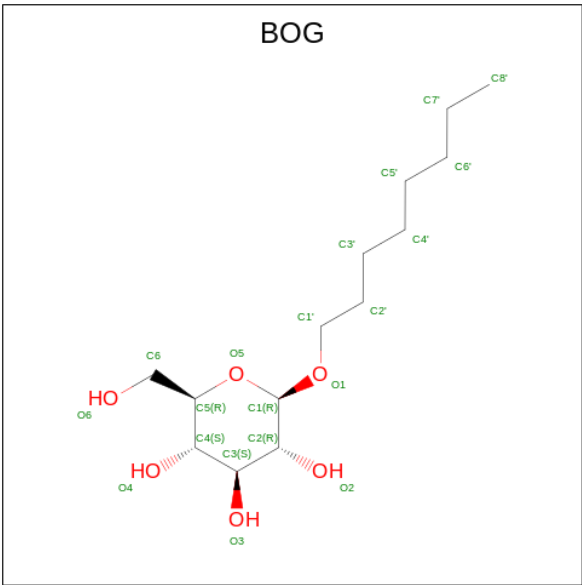
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	46	ALA	-	expression tag	UNP Q6Z2T3
G	50	ARG	LYS	conflict	UNP Q6Z2T3
G	66	ALA	CYS	conflict	UNP Q6Z2T3
G	93	VAL	THR	conflict	UNP Q6Z2T3
G	139	ALA	CYS	conflict	UNP Q6Z2T3
G	232	ARG	LYS	conflict	UNP Q6Z2T3
G	253	VAL	THR	conflict	UNP Q6Z2T3
G	264	ARG	LYS	conflict	UNP Q6Z2T3
H	45	MET	-	initiating methionine	UNP Q6Z2T3
H	46	ALA	-	expression tag	UNP Q6Z2T3
H	50	ARG	LYS	conflict	UNP Q6Z2T3
H	66	ALA	CYS	conflict	UNP Q6Z2T3
H	93	VAL	THR	conflict	UNP Q6Z2T3
H	139	ALA	CYS	conflict	UNP Q6Z2T3
H	232	ARG	LYS	conflict	UNP Q6Z2T3
H	253	VAL	THR	conflict	UNP Q6Z2T3
H	264	ARG	LYS	conflict	UNP Q6Z2T3

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

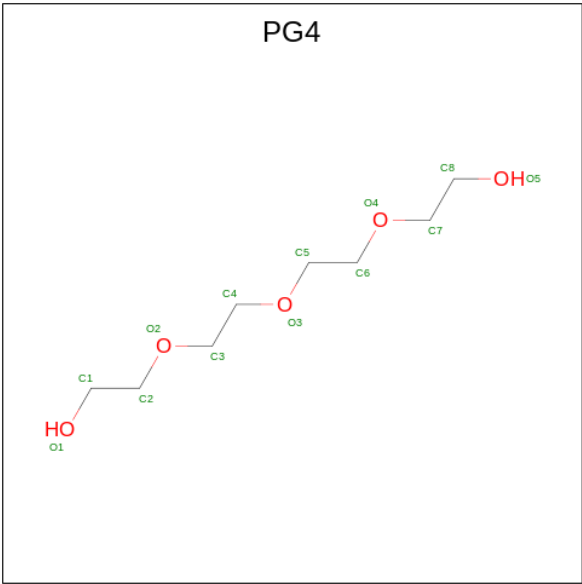
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Na 1 1	0	0
2	B	1	Total Na 1 1	0	0
2	C	1	Total Na 1 1	0	0
2	D	1	Total Na 1 1	0	0
2	E	1	Total Na 1 1	0	0
2	F	1	Total Na 1 1	0	0
2	G	1	Total Na 1 1	0	0
2	H	1	Total Na 1 1	0	0

- Molecule 3 is octyl beta-D-glucopyranoside (three-letter code: BOG) (formula: C₁₄H₂₈O₆).



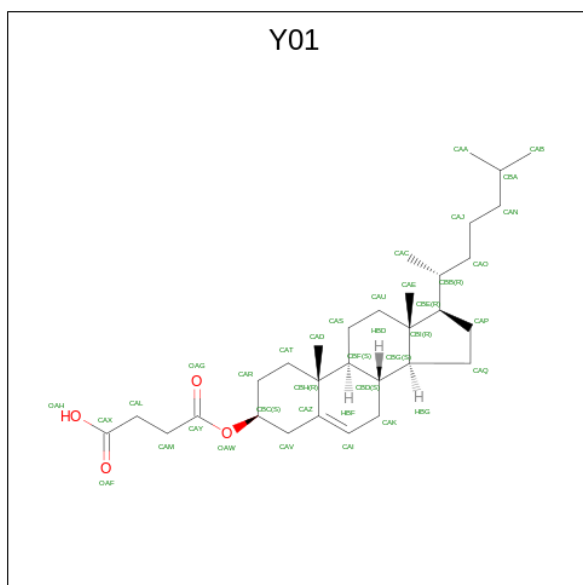
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	H	O	0	0
			29	8	15	6		
3	B	1	Total	C	H	O	0	0
			38	11	21	6		
3	D	1	Total	C	H	O	0	0
			48	14	28	6		
3	H	1	Total	C	H	O	0	0
			48	14	28	6		

- Molecule 4 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: C₈H₁₈O₅).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total 31	C 8	H 18	O 5	0	0
4	A	1	Total 31	C 8	H 18	O 5	0	0
4	B	1	Total 31	C 8	H 18	O 5	0	0
4	B	1	Total 31	C 8	H 18	O 5	0	0
4	C	1	Total 31	C 8	H 18	O 5	0	0
4	D	1	Total 31	C 8	H 18	O 5	0	0
4	E	1	Total 31	C 8	H 18	O 5	0	0
4	F	1	Total 31	C 8	H 18	O 5	0	0
4	G	1	Total 31	C 8	H 18	O 5	0	0
4	G	1	Total 31	C 8	H 18	O 5	0	0
4	H	1	Total 31	C 8	H 18	O 5	0	0

- Molecule 5 is CHOLESTEROL HEMISUCCINATE (three-letter code: Y01) (formula: $C_{31}H_{50}O_4$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	D	1	Total	C	O	0	0
			28	27	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	H	1	Total	C	O	0	0
			28	27	1		

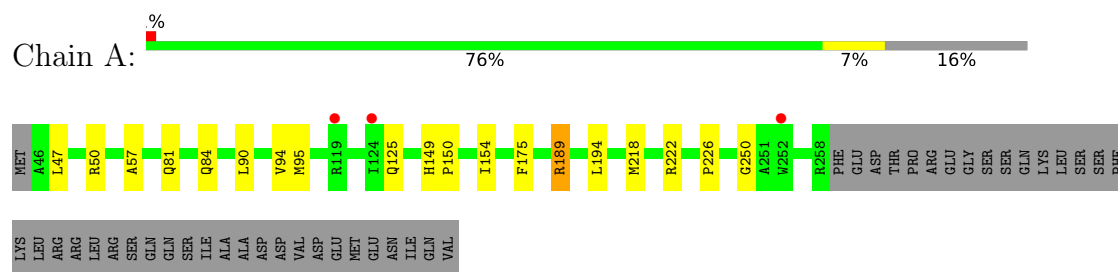
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	116	Total	O	0	0
			116	116		
6	B	117	Total	O	0	0
			117	117		
6	C	137	Total	O	0	0
			137	137		
6	D	117	Total	O	0	0
			117	117		
6	E	113	Total	O	0	0
			113	113		
6	F	111	Total	O	0	0
			111	111		
6	G	144	Total	O	0	0
			144	144		
6	H	129	Total	O	0	0
			129	129		

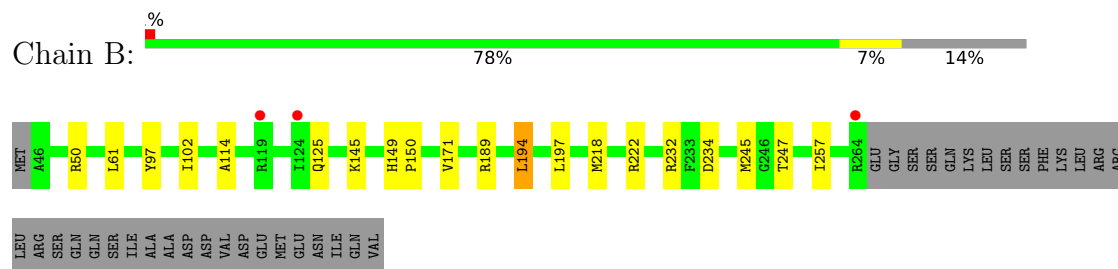
3 Residue-property plots

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.

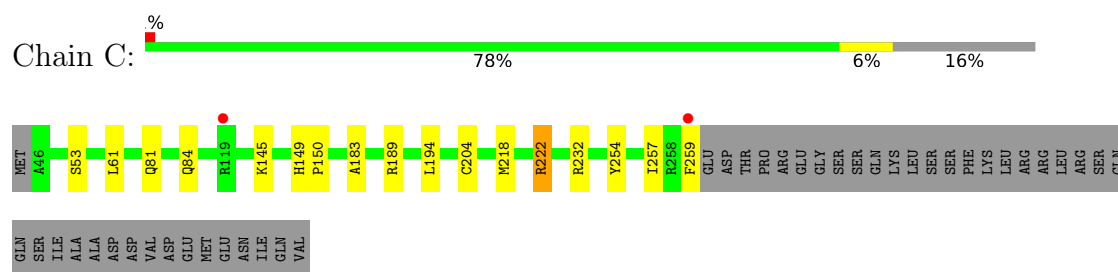
• Molecule 1: Aquaporin NIP2-1



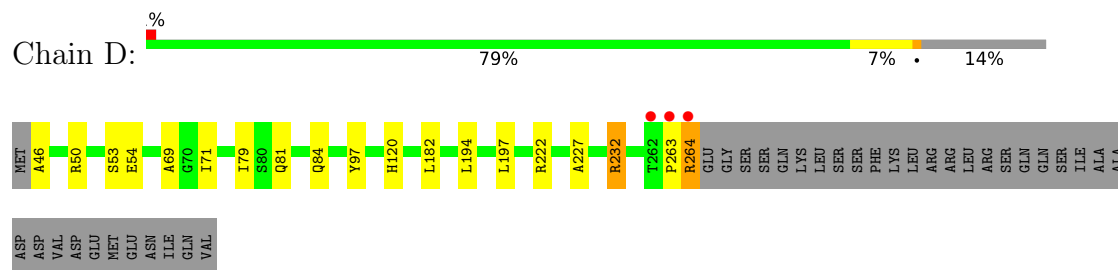
• Molecule 1: Aquaporin NIP2-1



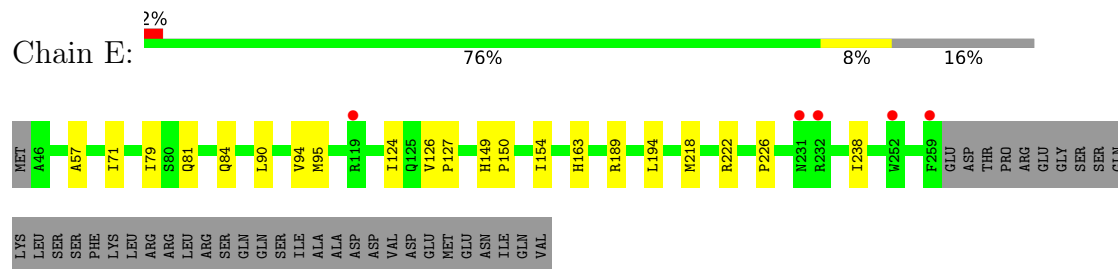
• Molecule 1: Aquaporin NIP2-1



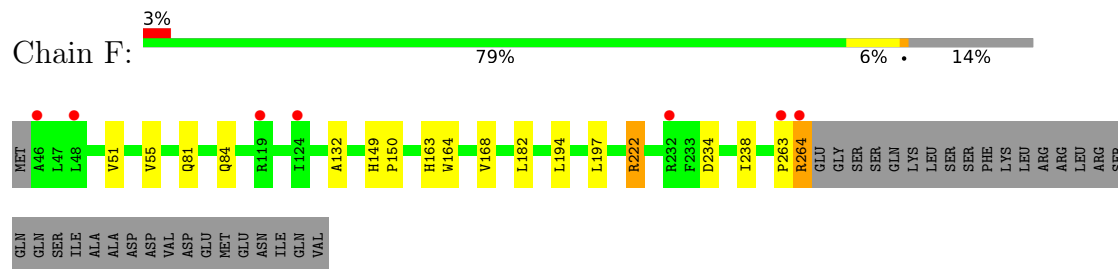
• Molecule 1: Aquaporin NIP2-1



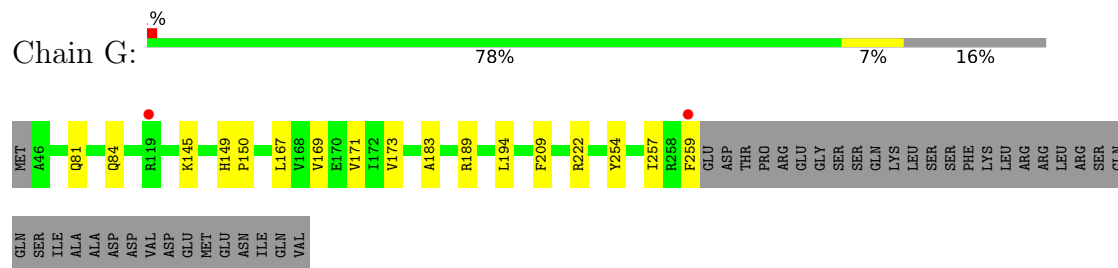
● Molecule 1: Aquaporin NIP2-1



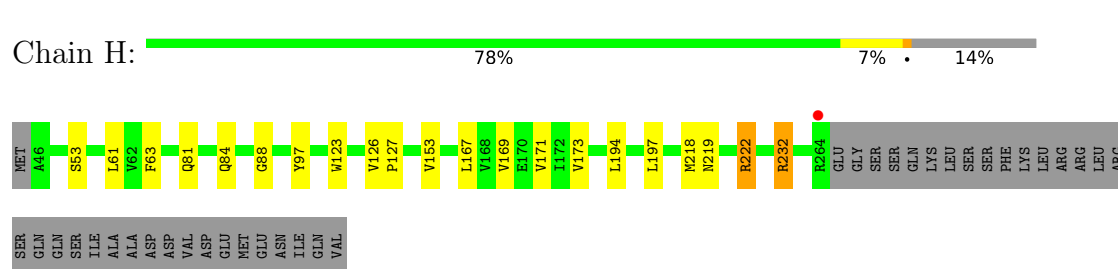
● Molecule 1: Aquaporin NIP2-1



● Molecule 1: Aquaporin NIP2-1



● Molecule 1: Aquaporin NIP2-1



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	89.57Å 92.45Å 166.13Å 90.00° 102.14° 90.00°	Depositor
Resolution (Å)	19.84 – 1.80 19.84 – 1.80	Depositor EDS
% Data completeness (in resolution range)	95.5 (19.84-1.80) 95.4 (19.84-1.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.99 (at 1.80Å)	Xtriage
Refinement program	PHENIX 1.13_2998	Depositor
R, R_{free}	0.245 , 0.273 0.245 , 0.273	Depositor DCC
R_{free} test set	11429 reflections (4.89%)	wwPDB-VP
Wilson B-factor (Å ²)	23.6	Xtriage
Anisotropy	0.321	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.48 , 65.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.53$, $\langle L^2 \rangle = 0.37$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	27690	wwPDB-VP
Average B, all atoms (Å ²)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 45.69 % 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 1.2743e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: BOG, NA, PG4, Y01

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	0.81	0/1629	0.76	2/2229 (0.1%)
1	B	0.75	0/1692	0.75	1/2315 (0.0%)
1	C	0.84	1/1649 (0.1%)	0.81	3/2256 (0.1%)
1	D	0.83	1/1686 (0.1%)	0.77	0/2307
1	E	0.82	0/1643	0.76	2/2248 (0.1%)
1	F	0.82	0/1691	0.77	1/2314 (0.0%)
1	G	0.82	0/1651	0.76	0/2259
1	H	0.83	0/1686	0.76	0/2307
All	All	0.82	2/13327 (0.0%)	0.77	9/18235 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	54	GLU	CD-OE2	-5.59	1.19	1.25
1	C	204	CYS	CB-SG	-5.16	1.73	1.81

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	232	ARG	NE-CZ-NH2	9.28	124.94	120.30
1	A	189	ARG	NE-CZ-NH2	8.32	124.46	120.30
1	C	232	ARG	NE-CZ-NH1	-7.14	116.73	120.30
1	A	189	ARG	NE-CZ-NH1	-6.59	117.00	120.30
1	E	189	ARG	NE-CZ-NH1	5.89	123.25	120.30

There are no chirality outliers.

There are no planarity outliers.

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	1581	1618	1618	14	0
1	B	1639	1673	1675	16	0
1	C	1597	1638	1638	10	0
1	D	1636	1670	1670	12	0
1	E	1594	1633	1633	12	0
1	F	1638	1673	1673	10	0
1	G	1599	1643	1644	11	0
1	H	1636	1670	1670	20	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
2	G	1	0	0	0	0
2	H	1	0	0	0	0
3	A	14	15	13	1	0
3	B	17	21	19	1	0
3	D	20	28	28	0	0
3	H	20	28	28	0	0
4	A	26	36	36	0	0
4	B	26	36	36	0	0
4	C	13	18	18	0	0
4	D	13	18	18	0	0
4	E	13	18	18	0	0
4	F	13	18	18	0	0
4	G	26	36	36	0	0
4	H	13	18	18	2	0
5	D	28	0	45	6	0
5	H	28	0	45	2	0
6	A	116	0	0	2	0
6	B	117	0	0	2	0
6	C	137	0	0	2	0
6	D	117	0	0	2	0
6	E	113	0	0	2	0
6	F	111	0	0	0	0
6	G	144	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	H	129	0	0	6	0
All	All	14182	13508	13597	104	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 4.

The worst 5 of 104 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)
5:D:304:Y01:HAC1	5:D:304:Y01:HAU2	1.71	0.72
4:H:303:PG4:O3	6:H:401:HOH:O	2.08	0.70
1:B:50:ARG:NH2	1:B:125:GLN:OE1	2.27	0.67
1:A:50:ARG:NH2	6:A:402:HOH:O	2.28	0.64
1:D:69:ALA:HB2	6:D:415:HOH:O	1.97	0.64

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	213/254 (84%)	206 (97%)	7 (3%)	0	100	100
1	B	220/254 (87%)	209 (95%)	11 (5%)	0	100	100
1	C	215/254 (85%)	202 (94%)	13 (6%)	0	100	100
1	D	219/254 (86%)	210 (96%)	9 (4%)	0	100	100
1	E	214/254 (84%)	206 (96%)	8 (4%)	0	100	100
1	F	220/254 (87%)	212 (96%)	8 (4%)	0	100	100
1	G	215/254 (85%)	203 (94%)	12 (6%)	0	100	100
1	H	219/254 (86%)	210 (96%)	9 (4%)	0	100	100
All	All	1735/2032 (85%)	1658 (96%)	77 (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	164/200 (82%)	163 (99%)	1 (1%)	86	84
1	B	171/200 (86%)	167 (98%)	4 (2%)	50	37
1	C	166/200 (83%)	165 (99%)	1 (1%)	86	84
1	D	170/200 (85%)	167 (98%)	3 (2%)	59	48
1	E	165/200 (82%)	164 (99%)	1 (1%)	86	84
1	F	171/200 (86%)	168 (98%)	3 (2%)	59	48
1	G	166/200 (83%)	165 (99%)	1 (1%)	86	84
1	H	170/200 (85%)	168 (99%)	2 (1%)	71	65
All	All	1343/1600 (84%)	1327 (99%)	16 (1%)	71	65

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	222	ARG
1	G	222	ARG
1	D	264	ARG
1	F	264	ARG
1	D	232	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	149	HIS
1	C	149	HIS
1	G	149	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 25 ligands modelled in this entry, 8 are monoatomic - leaving 17 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$
4	PG4	A	303	-	12,12,12	0.71	0	11,11,11	0.51	0
3	BOG	B	302	-	17,17,20	0.88	1 (5%)	22,22,25	1.26	2 (9%)
3	BOG	A	302	-	14,14,20	0.68	0	19,19,25	0.71	0
4	PG4	E	302	-	12,12,12	0.60	0	11,11,11	0.80	0
3	BOG	H	302	-	20,20,20	0.86	1 (5%)	25,25,25	1.28	2 (8%)
5	Y01	D	304	-	31,31,38	1.59	7 (22%)	48,48,57	1.58	14 (29%)
5	Y01	H	304	-	31,31,38	1.58	6 (19%)	48,48,57	1.53	7 (14%)
4	PG4	B	303	-	12,12,12	0.55	0	11,11,11	0.27	0
4	PG4	G	303	-	12,12,12	0.76	0	11,11,11	0.88	0
4	PG4	D	303	-	12,12,12	0.58	0	11,11,11	0.56	0
4	PG4	H	303	-	12,12,12	0.58	0	11,11,11	0.50	0
4	PG4	B	304	-	12,12,12	0.52	0	11,11,11	0.49	0
4	PG4	C	302	-	12,12,12	0.66	0	11,11,11	0.76	0
4	PG4	F	302	-	12,12,12	0.69	0	11,11,11	0.42	0
3	BOG	D	302	-	20,20,20	0.62	0	25,25,25	1.06	1 (4%)
4	PG4	A	304	-	12,12,12	0.47	0	11,11,11	0.47	0
4	PG4	G	302	-	12,12,12	0.59	0	11,11,11	0.61	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
4	PG4	A	303	-	-	7/10/10/10	-
3	BOG	B	302	-	-	3/8/28/31	0/1/1/1
3	BOG	A	302	-	-	5/5/25/31	0/1/1/1
4	PG4	E	302	-	-	7/10/10/10	-
3	BOG	H	302	-	-	4/11/31/31	0/1/1/1
5	Y01	D	304	-	-	3/10/68/77	0/4/4/4
5	Y01	H	304	-	-	5/10/68/77	0/4/4/4
4	PG4	B	303	-	-	0/10/10/10	-
4	PG4	G	303	-	-	6/10/10/10	-
4	PG4	D	303	-	-	3/10/10/10	-
4	PG4	H	303	-	-	5/10/10/10	-
4	PG4	B	304	-	-	5/10/10/10	-
4	PG4	C	302	-	-	7/10/10/10	-
4	PG4	F	302	-	-	7/10/10/10	-
3	BOG	D	302	-	-	6/11/31/31	0/1/1/1
4	PG4	A	304	-	-	4/10/10/10	-
4	PG4	G	302	-	-	1/10/10/10	-

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	H	304	Y01	CAQ-CAP	3.89	1.64	1.54
5	D	304	Y01	CAQ-CAP	3.71	1.64	1.54
5	H	304	Y01	CAS-CBF	-3.19	1.48	1.53
5	D	304	Y01	CBH-CBF	2.95	1.61	1.56
5	D	304	Y01	CAS-CBF	-2.81	1.49	1.53

The worst 5 of 26 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	H	302	BOG	O1-C1-C2	4.12	114.73	108.30
3	B	302	BOG	O1-C1-C2	3.85	114.31	108.30
5	H	304	Y01	CAP-CAQ-CBG	-3.65	97.89	105.13
5	H	304	Y01	CAV-CAZ-CBH	3.33	120.84	116.42
3	D	302	BOG	O1-C1-C2	3.24	113.36	108.30

There are no chirality outliers.

5 of 78 torsion outliers are listed below:

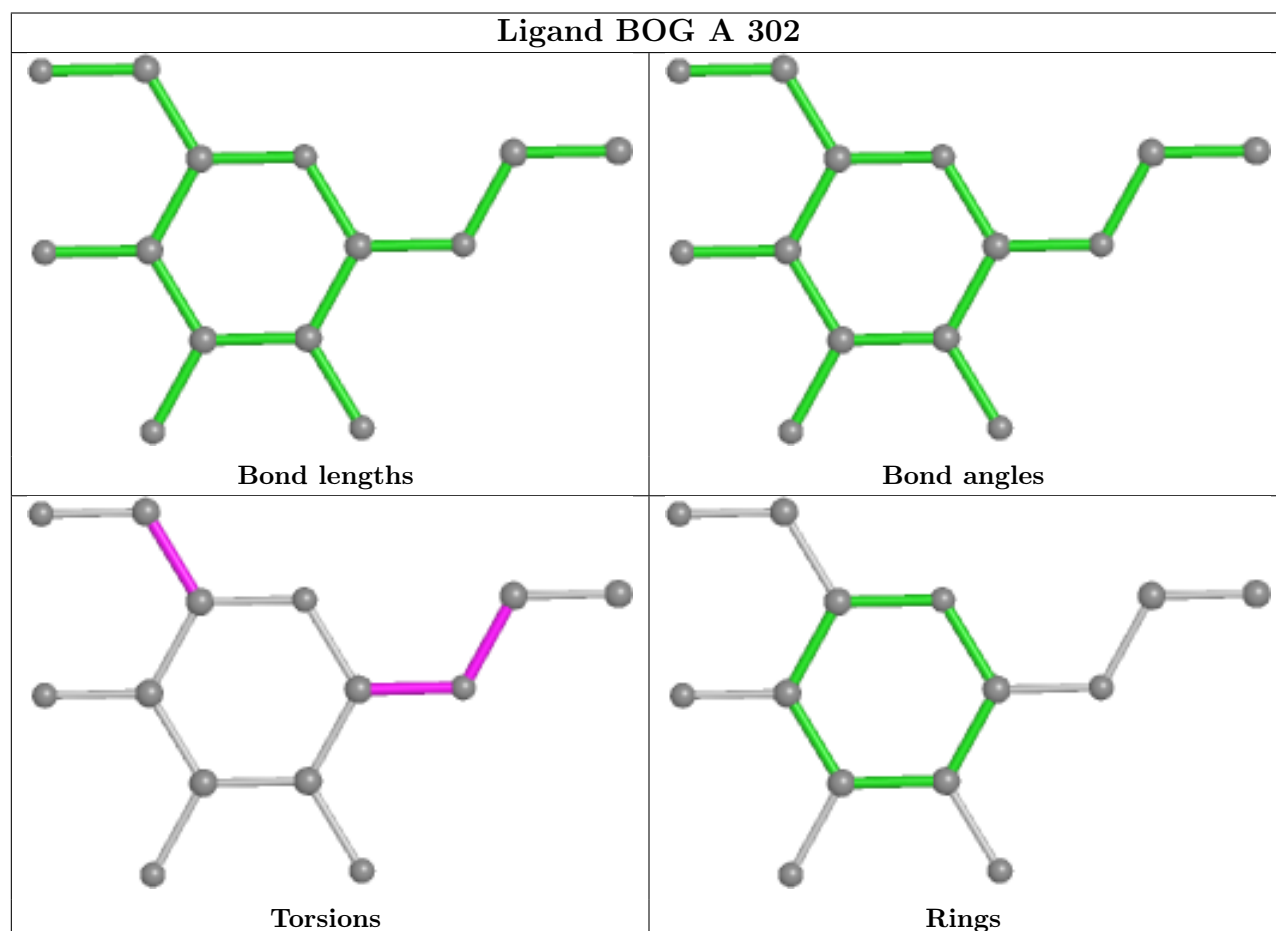
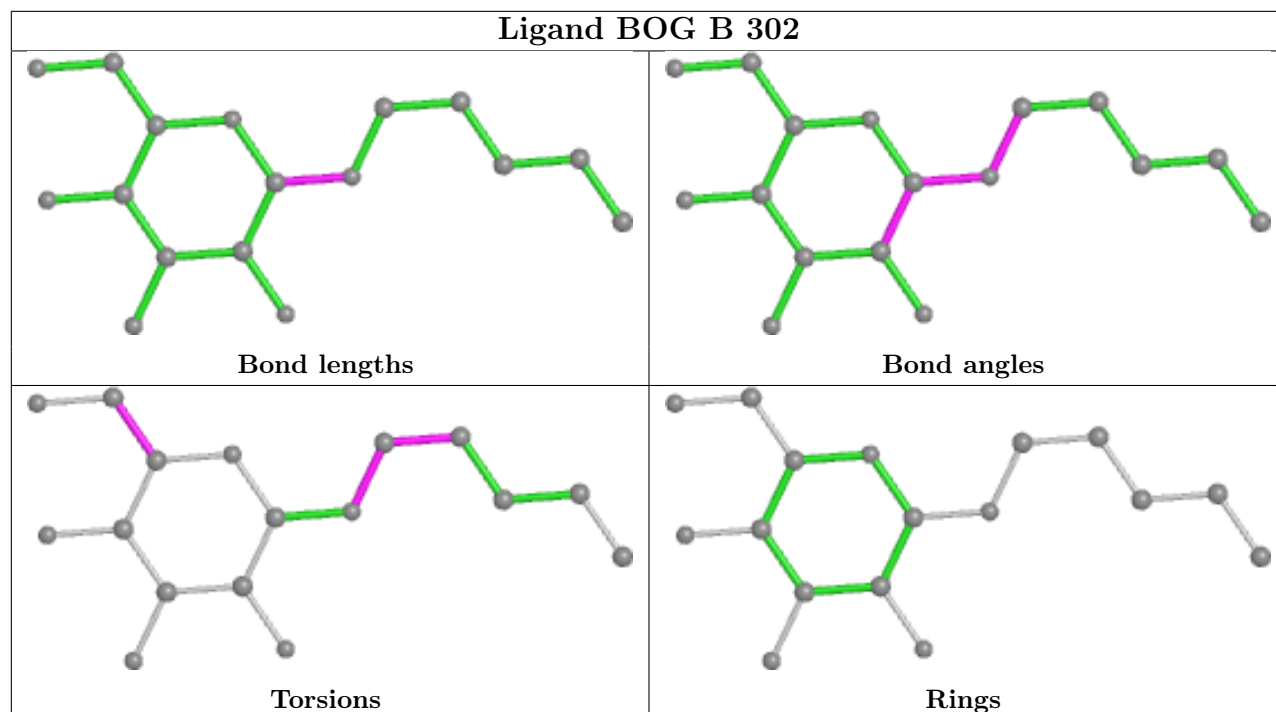
Mol	Chain	Res	Type	Atoms
3	A	302	BOG	C2-C1-O1-C1'
3	A	302	BOG	O5-C1-O1-C1'
3	B	302	BOG	C2'-C1'-O1-C1
5	D	304	Y01	CAJ-CAO-CBB-CBE
4	A	303	PG4	O4-C7-C8-O5

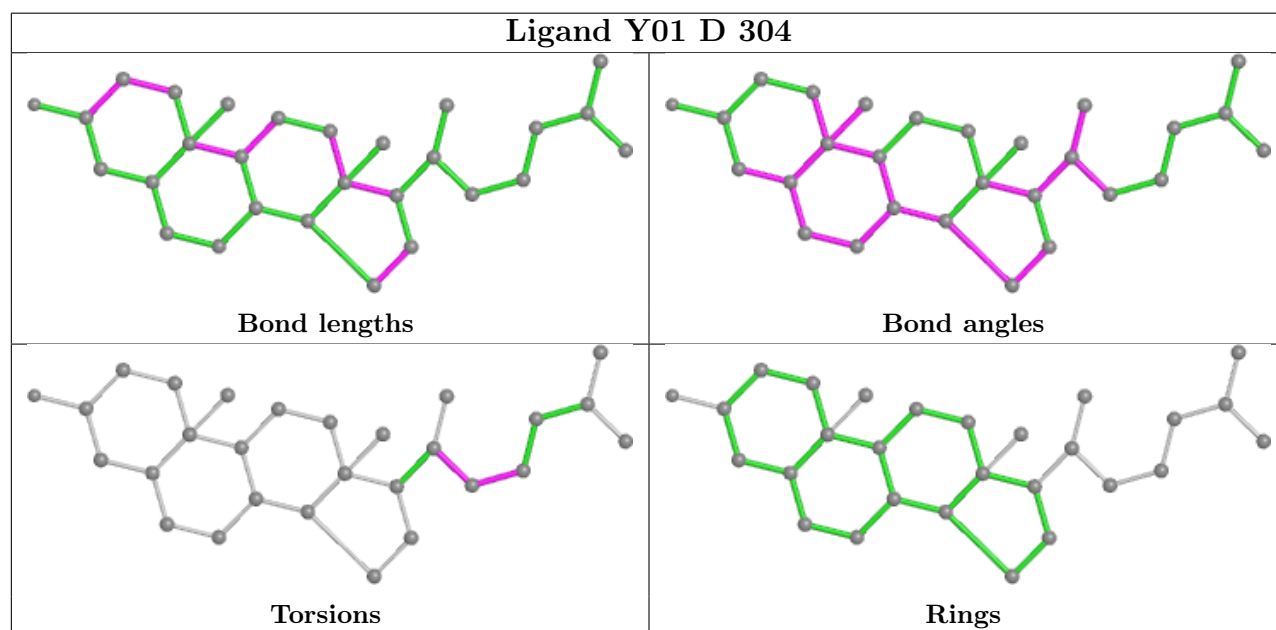
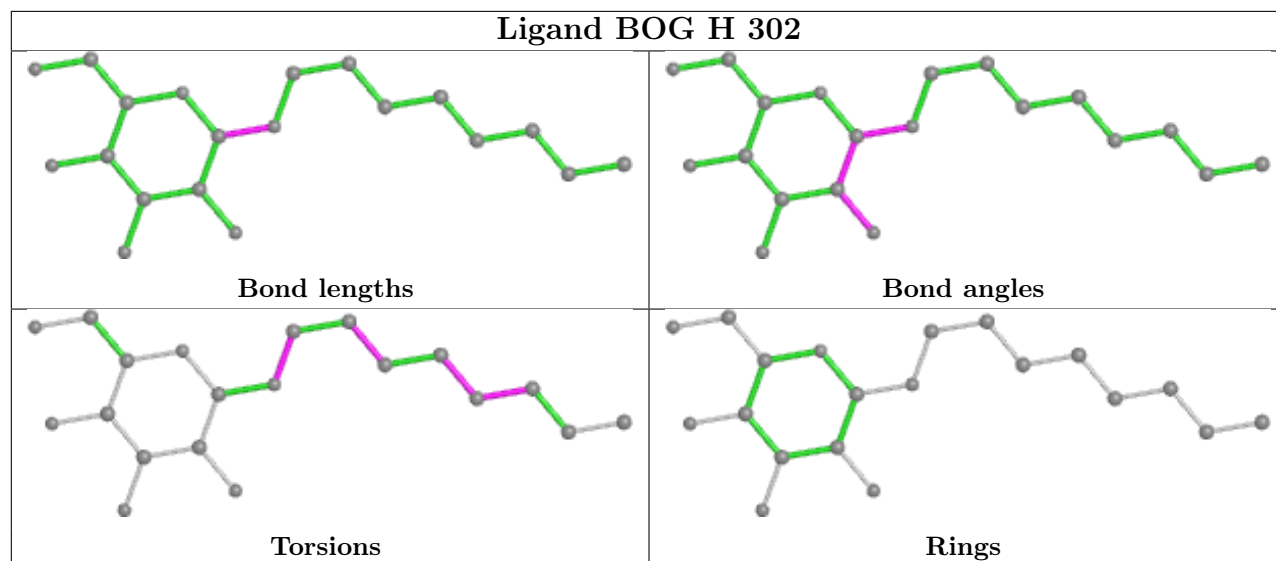
There are no ring outliers.

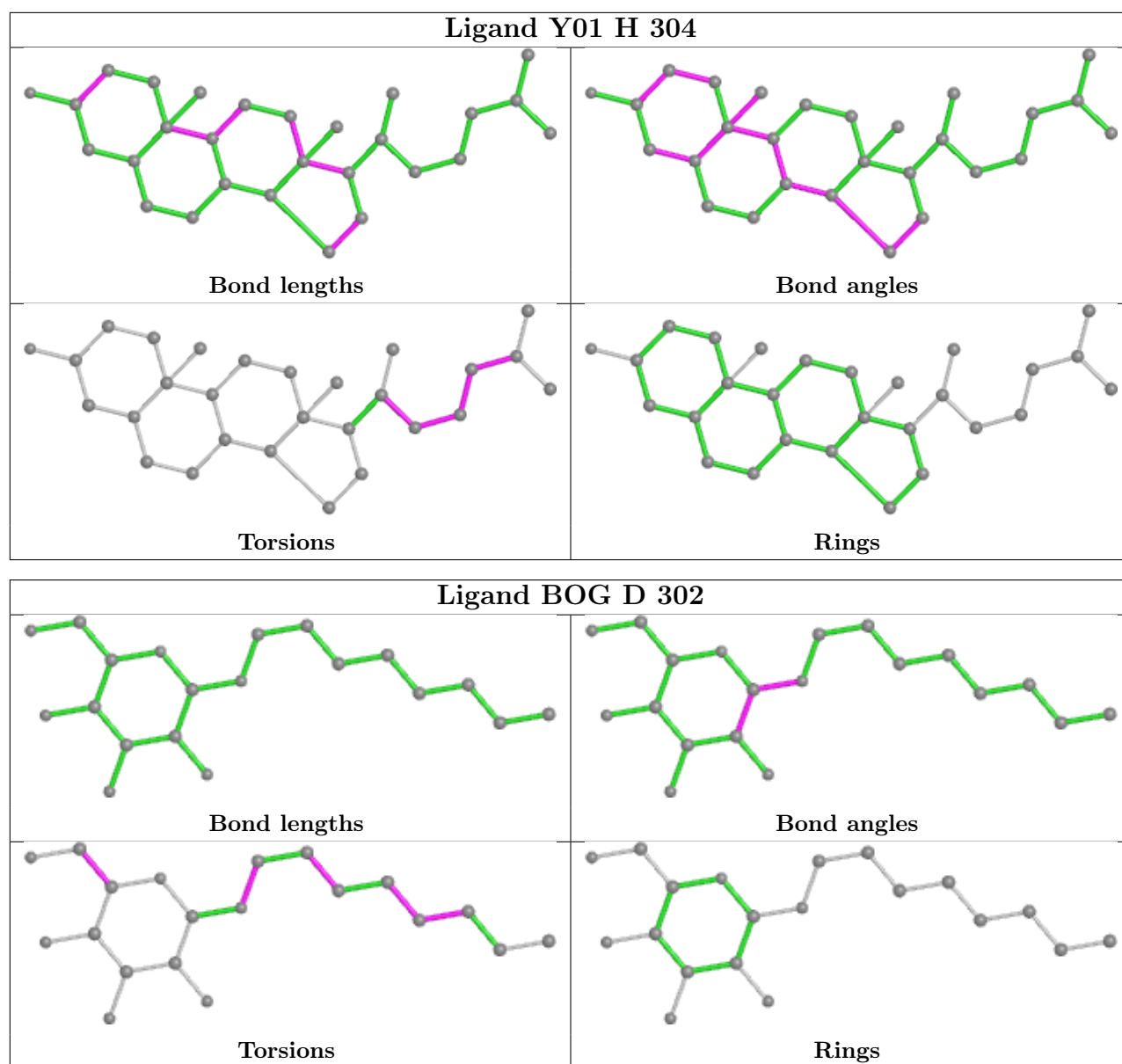
5 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	302	BOG	1	0
3	A	302	BOG	1	0
5	D	304	Y01	6	0
5	H	304	Y01	2	0
4	H	303	PG4	2	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 than 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.







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](#)

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(Å ²)	Q<0.9
1	A	213/254 (83%)	0.01	3 (1%) 75 72	20, 26, 36, 52	0
1	B	219/254 (86%)	0.02	3 (1%) 75 72	22, 27, 41, 76	0
1	C	214/254 (84%)	-0.07	2 (0%) 84 82	20, 24, 34, 58	0
1	D	219/254 (86%)	-0.06	3 (1%) 75 72	19, 25, 37, 77	0
1	E	214/254 (84%)	-0.08	5 (2%) 60 56	18, 25, 36, 92	0
1	F	219/254 (86%)	0.02	7 (3%) 47 41	17, 26, 39, 70	0
1	G	214/254 (84%)	-0.16	2 (0%) 84 82	18, 24, 33, 62	0
1	H	219/254 (86%)	-0.10	1 (0%) 91 89	18, 24, 37, 82	0
All	All	1731/2032 (85%)	-0.05	26 (1%) 73 70	17, 25, 37, 92	0

The worst 5 of 26 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	259	PHE	7.4
1	C	259	PHE	5.1
1	B	124	ILE	4.4
1	H	264	ARG	4.4
1	F	124	ILE	4.2

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

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

6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

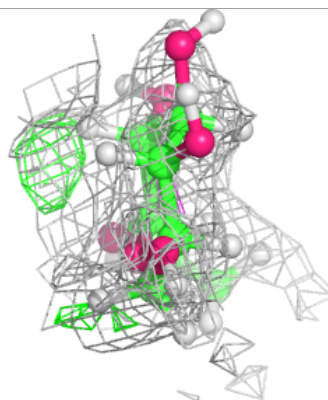
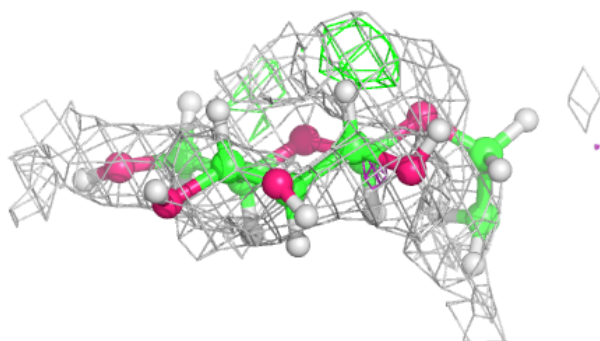
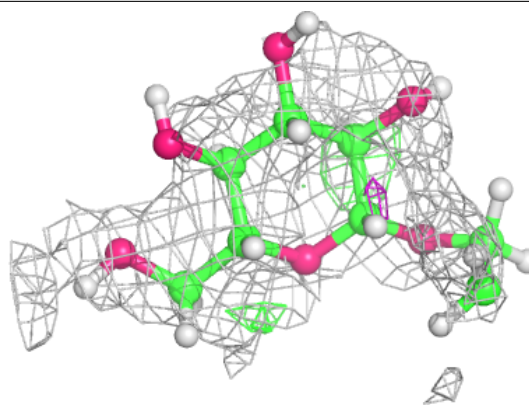
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
4	PG4	F	302	13/13	0.19	0.27	59,71,75,75	0
4	PG4	B	303	13/13	0.27	0.44	84,102,109,109	0
3	BOG	A	302	14/20	0.37	0.40	91,93,112,112	0
4	PG4	D	303	13/13	0.46	0.36	61,74,79,79	0
5	Y01	D	304	28/35	0.50	0.45	70,75,77,77	0
4	PG4	G	302	13/13	0.55	0.24	61,73,76,76	0
4	PG4	A	304	13/13	0.56	0.26	68,81,83,83	0
5	Y01	H	304	28/35	0.56	0.47	78,81,82,82	0
4	PG4	H	303	13/13	0.57	0.26	61,75,76,76	0
3	BOG	B	302	17/20	0.58	0.29	52,64,67,68	0
4	PG4	B	304	13/13	0.62	0.29	64,77,79,79	0
3	BOG	H	302	20/20	0.69	0.32	46,57,60,62	0
3	BOG	D	302	20/20	0.70	0.26	49,60,66,67	0
4	PG4	C	302	13/13	0.74	0.17	49,58,63,63	0
4	PG4	G	303	13/13	0.77	0.18	42,50,58,59	0
4	PG4	E	302	13/13	0.79	0.18	45,55,59,60	0
2	NA	D	301	1/1	0.81	0.19	43,43,43,43	0
2	NA	H	301	1/1	0.81	0.34	44,44,44,44	0
4	PG4	A	303	13/13	0.83	0.15	41,50,51,51	0
2	NA	C	301	1/1	0.90	0.36	45,45,45,45	0
2	NA	E	301	1/1	0.91	0.24	42,42,42,42	0
2	NA	B	301	1/1	0.92	0.21	41,41,41,41	0
2	NA	A	301	1/1	0.93	0.09	48,48,48,48	0
2	NA	G	301	1/1	0.94	0.20	38,38,38,38	0
2	NA	F	301	1/1	0.98	0.18	46,46,46,46	0

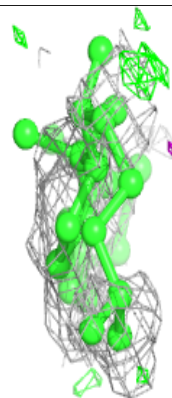
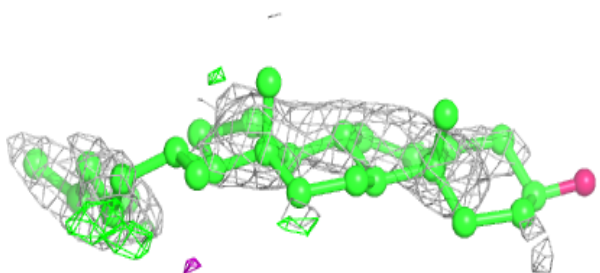
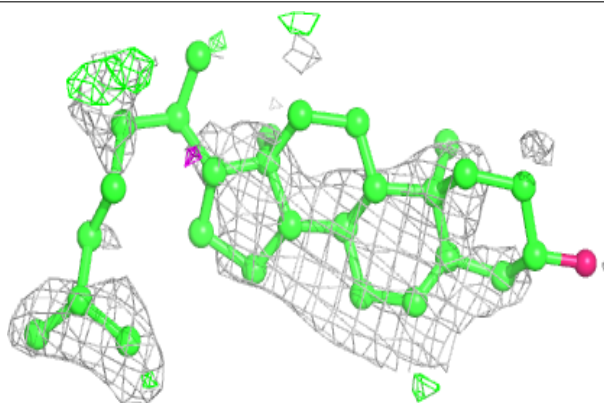
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around BOG A 302:

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

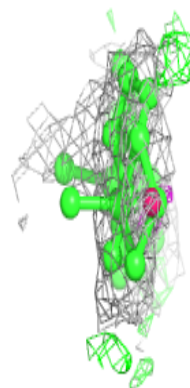
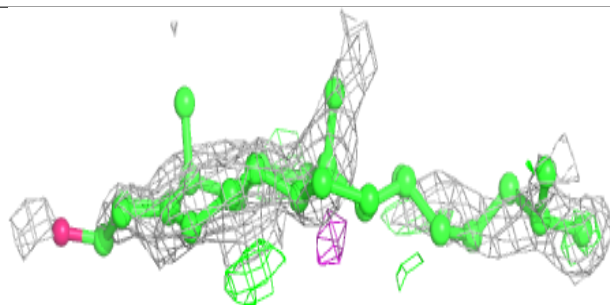
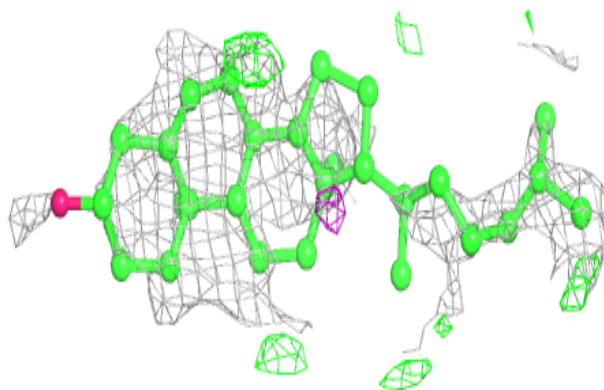
**Electron density around Y01 D 304:**

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



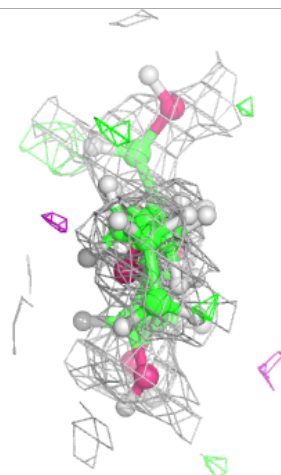
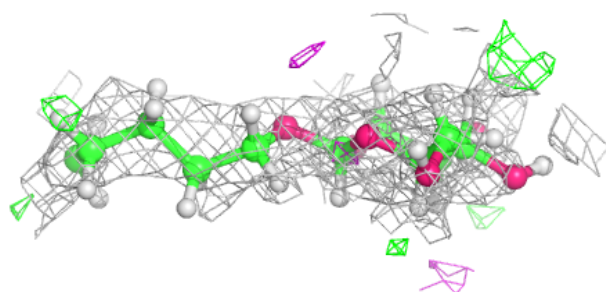
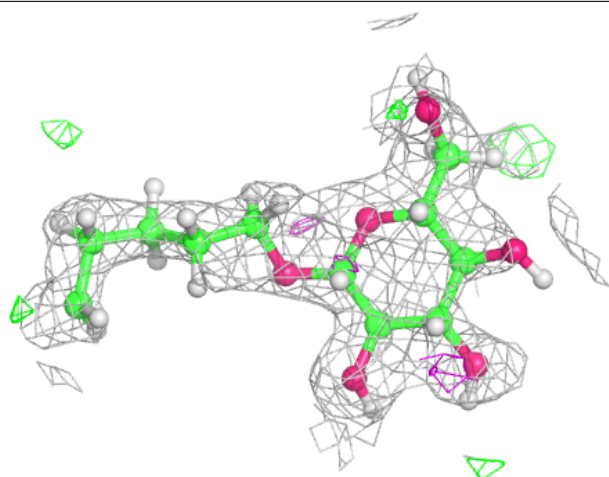
Electron density around Y01 H 304:

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



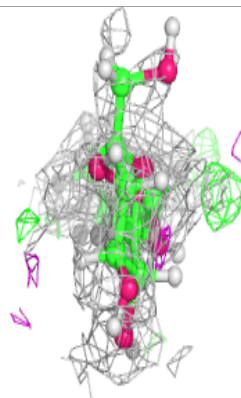
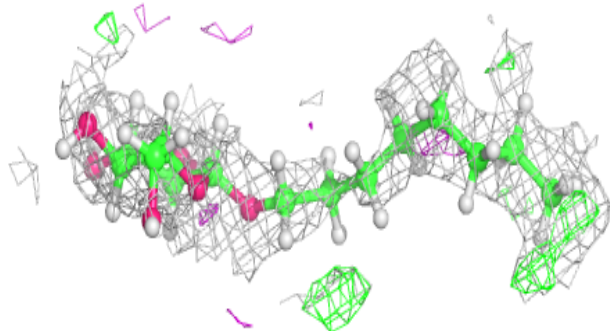
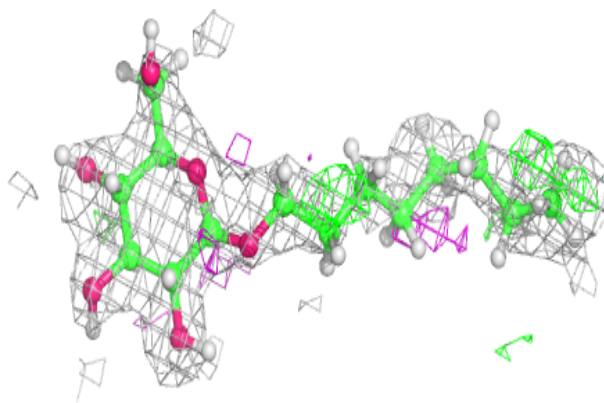
Electron density around BOG B 302:

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

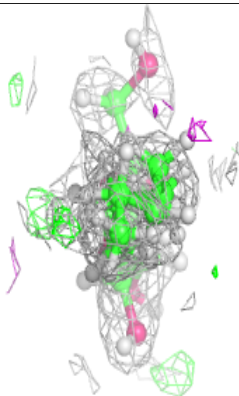
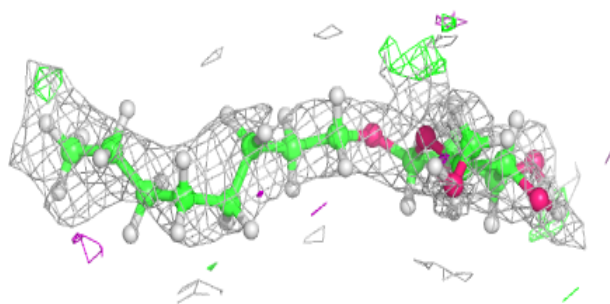
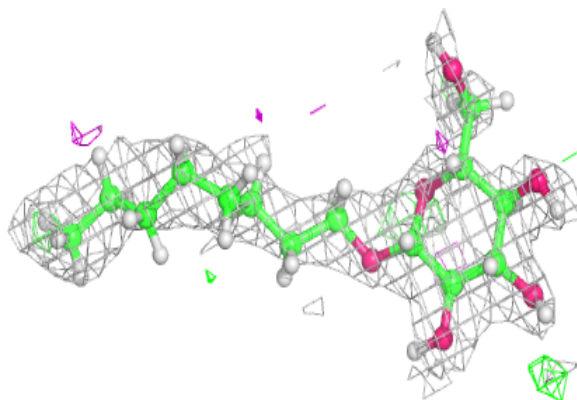


Electron density around BOG H 302:

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

**Electron density around BOG D 302:**

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



6.5 Other polymers [i](#)

There are no such residues in this entry.