Full wwPDB X-ray Structure Validation Report (i)

Aug 9, 2020 – 12:05 PM BST

PDB ID	:	5J3J
Title	:	Crystal Structure of human DPP-IV in complex with HL1
Authors	:	Wu, F.; Li, H.; Zhao, Z.; Zhu, L.; Xu, H.; Li, S.
Deposited on	:	2016-03-31
Resolution	:	2.75 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 following versions of software and data (see references (i)) were used in the production of this report:

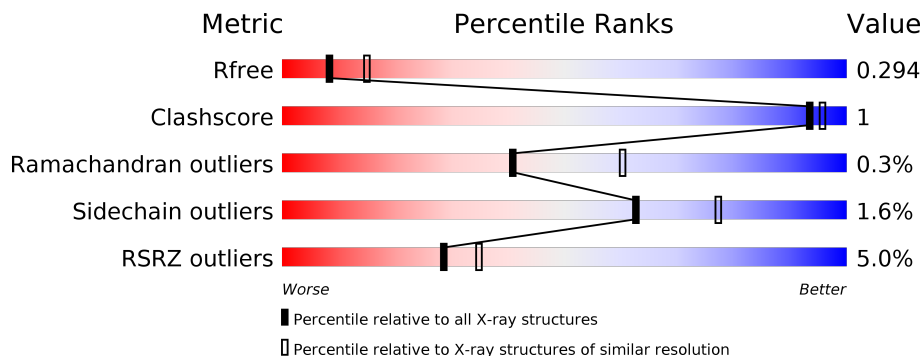|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

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 2.75 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		0.294
Clashscore		1
Ramachandran outliers		0.3%
Sidechain outliers		1.6%
RSRZ outliers		5.0%

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}	130704	1235 (2.78-2.74)
Clashscore	141614	1277 (2.78-2.74)
Ramachandran outliers	138981	1257 (2.78-2.74)
Sidechain outliers	138945	1257 (2.78-2.74)
RSRZ outliers	127900	1207 (2.78-2.74)

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% 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	734	5% 94% 5% ·
1	B	734	5% 92% 7% ·
2	C	4	100%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 12141 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 Dipeptidyl peptidase 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	727	Total 5957	C 3824	N 981	O 1126	S 26	0	0	0
1	B	727	Total 5957	C 3824	N 981	O 1126	S 26	0	0	0

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	767	GLY	-	expression tag	UNP P27487
A	768	HIS	-	expression tag	UNP P27487
A	769	HIS	-	expression tag	UNP P27487
A	770	HIS	-	expression tag	UNP P27487
A	771	HIS	-	expression tag	UNP P27487
A	772	HIS	-	expression tag	UNP P27487
A	773	HIS	-	expression tag	UNP P27487
B	767	GLY	-	expression tag	UNP P27487
B	768	HIS	-	expression tag	UNP P27487
B	769	HIS	-	expression tag	UNP P27487
B	770	HIS	-	expression tag	UNP P27487
B	771	HIS	-	expression tag	UNP P27487
B	772	HIS	-	expression tag	UNP P27487
B	773	HIS	-	expression tag	UNP P27487

- Molecule 2 is an oligosaccharide called alpha-D-mannopyranose-(1-3)-alpha-D-mannopyran ose-(1-3)-alpha-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.
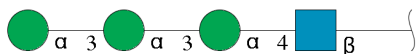
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	4	Total 47	C 26	N 1	O 20	0	0	0

- Molecule 3 is (2 {S},3 {R})-8,9-dimethoxy-3-[2,4,5-tris(fluoranyl)phenyl]-2,3-dihydro-1 {H} -benzo[f]chromen-2-amine (three-letter code: HL1) (formula: $C_{21}H_{18}F_3NO_3$).
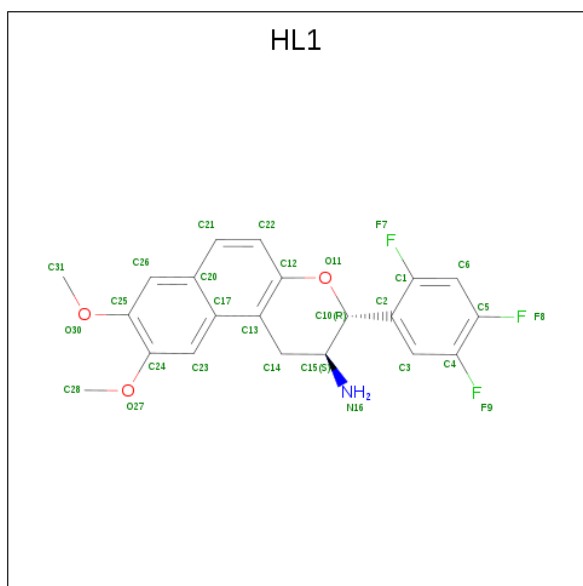
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total 28	C 21	F 3	N 1	O 3	0	0
3	B	1	Total 28	C 21	F 3	N 1	O 3	0	0

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).

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

• Molecule 5 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	59	Total 59	O 59	0	0
5	B	37	Total 37	O 37	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.
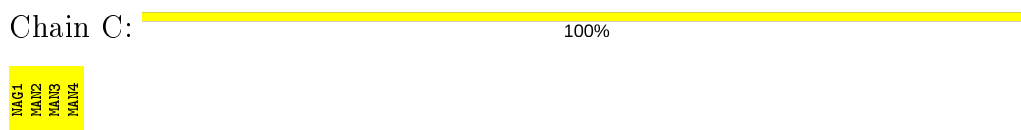
- Molecule 1: Dipeptidyl peptidase 4
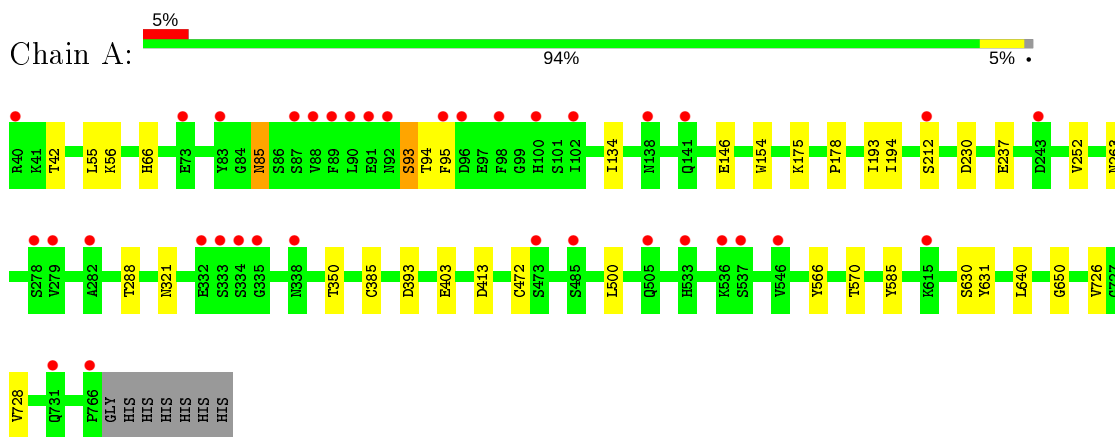
- Molecule 1: Dipeptidyl peptidase 4
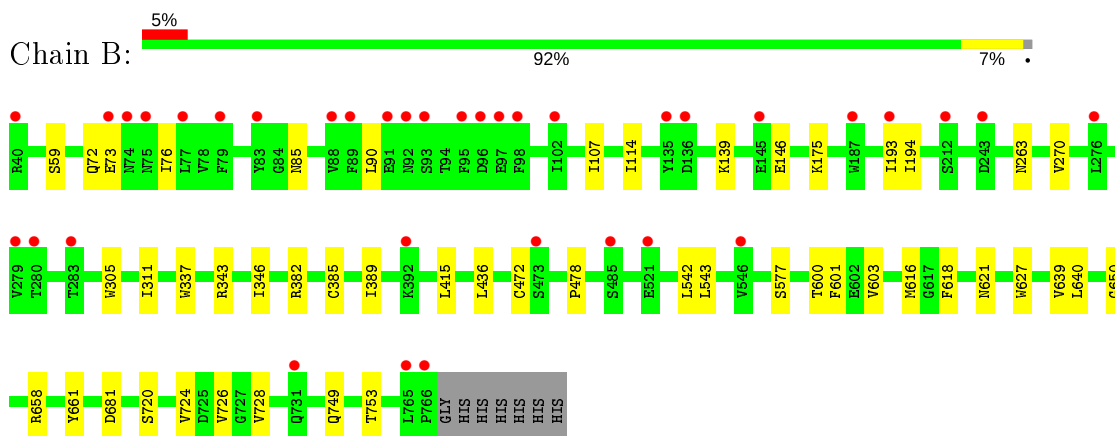
- Molecule 2: alpha-D-mannopyranose-(1-3)-alpha-D-mannopyranose-(1-3)-alpha-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	65.36Å 126.82Å 112.62Å 90.00° 99.95° 90.00°	Depositor
Resolution (Å)	55.11 – 2.75 55.05 – 2.75	Depositor EDS
% Data completeness (in resolution range)	94.7 (55.11-2.75) 94.7 (55.05-2.75)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	1.93 (at 2.77Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.235 , 0.298 0.236 , 0.294	Depositor DCC
R_{free} test set	2248 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	35.5	Xtriage
Anisotropy	0.134	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 37.8	EDS
L-test for twinning[2]	$<|L|> = 0.49$, $<L^2> = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.86	EDS
Total number of atoms	12141	wwPDB-VP
Average B, all atoms (Å^2)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.59% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, HL1, MAN

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$|Z|$>5	RMSZ	#$|Z|$>5
1	A	0.36	1/6129 (0.0%)	0.49	0/8336
1	B	0.30	0/6129	0.50	0/8336
All	All	0.33	1/12258 (0.0%)	0.49	0/16672

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	631	TYR	CE1-CZ	-6.70	1.29	1.38

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	5957	0	5679	12	0
1	B	5957	0	5678	20	0
2	C	47	0	40	0	0
3	A	28	0	0	0	0
3	B	28	0	0	0	0
4	A	14	0	13	0	0
4	B	14	0	13	0	0
5	A	59	0	0	0	0

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	B	37	0	0	0	0
All	All	12141	0	11423	32	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 1.

All (32) 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:658:ARG:HD3	1:B:661:TYR:CZ	2.31	0.65
1:A:640:LEU:HD11	1:A:650:GLY:HA3	1.83	0.60
1:A:726:VAL:HG23	1:A:728:VAL:HG23	1.84	0.59
1:B:76:ILE:HD12	1:B:90:LEU:HD23	1.86	0.57
1:B:472:CYS:O	1:B:478:PRO:HA	2.07	0.54
1:B:726:VAL:HG23	1:B:728:VAL:HG23	1.92	0.51
1:B:415:LEU:HB2	1:B:436:LEU:HD21	1.93	0.49
1:B:346:ILE:N	1:B:346:ILE:HD12	2.28	0.49
1:B:640:LEU:HD11	1:B:650:GLY:HA3	1.94	0.48
1:B:749:GLN:O	1:B:753:THR:OG1	2.26	0.46
1:B:305:TRP:CE2	1:B:311:ILE:HD12	2.51	0.46
1:A:55:LEU:HD23	1:A:500:LEU:CD2	2.45	0.46
1:A:193:ILE:HG22	1:A:194:ILE:HG12	1.97	0.45
1:B:270:VAL:HG11	1:B:337:TRP:CZ2	2.51	0.45
1:A:146:GLU:O	1:A:175:LYS:NZ	2.47	0.44
1:A:42:THR:HG21	1:A:570:THR:HG23	1.99	0.44
1:B:107:ILE:HD11	1:B:114:ILE:HD12	1.99	0.44
1:B:603:VAL:HG13	1:B:639:VAL:HG23	1.98	0.44
1:B:616:MET:HB3	1:B:618:PHE:CE2	2.53	0.44
1:B:193:ILE:HG22	1:B:194:ILE:HG12	1.99	0.44
1:B:343:ARG:NE	1:B:389:ILE:HD12	2.33	0.43
1:B:146:GLU:O	1:B:175:LYS:NZ	2.49	0.43
1:A:154:TRP:CE2	1:A:212:SER:HB3	2.54	0.42
1:A:403:GLU:OE1	1:A:585:TYR:HA	2.19	0.42
1:A:134:ILE:HG21	1:A:178:PRO:HB3	2.02	0.41
1:B:542:LEU:HD23	1:B:543:LEU:N	2.35	0.41
1:B:72:GLN:HG2	1:B:73:GLU:HG2	2.03	0.41
1:A:237:GLU:HA	1:A:252:VAL:O	2.20	0.41
1:A:93:SER:O	1:A:95:PHE:N	2.50	0.41
1:B:600:THR:OG1	1:B:601:PHE:N	2.54	0.41
1:A:413:ASP:N	1:A:413:ASP:OD1	2.55	0.40
1:B:720:SER:O	1:B:724:VAL:HG23	2.21	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	725/734 (99%)	677 (93%)	44 (6%)	4 (1%)	25	42
1	B	725/734 (99%)	682 (94%)	42 (6%)	1 (0%)	51	75
All	All	1450/1468 (99%)	1359 (94%)	86 (6%)	5 (0%)	41	60

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	85	ASN
1	A	85	ASN
1	A	93	SER
1	A	630	SER
1	A	94	THR

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	652/658 (99%)	640 (98%)	12 (2%)	59	75
1	B	652/658 (99%)	643 (99%)	9 (1%)	67	79
All	All	1304/1316 (99%)	1283 (98%)	21 (2%)	62	77

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	56	LYS
1	A	66	HIS
1	A	85	ASN
1	A	230	ASP
1	A	263	ASN
1	A	288	THR
1	A	321	ASN
1	A	350	THR
1	A	385	CYS
1	A	393	ASP
1	A	472	CYS
1	A	566	TYR
1	B	59	SER
1	B	139	LYS
1	B	263	ASN
1	B	382	ARG
1	B	385	CYS
1	B	577	SER
1	B	621	ASN
1	B	627	TRP
1	B	681	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	85	ASN
1	A	281	ASN
1	A	321	ASN
1	A	435	GLN
1	A	520	ASN
1	B	138	ASN
1	B	196	ASN
1	B	227	GLN
1	B	263	ASN
1	B	363	HIS
1	B	388	GLN
1	B	435	GLN
1	B	572	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)

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	NAG	C	1	1,2	14,14,15	0.56	0	17,19,21	1.33	3 (17%)
2	MAN	C	2	2	11,11,12	0.97	0	15,15,17	1.71	2 (13%)
2	MAN	C	3	2	11,11,12	0.53	0	15,15,17	1.19	3 (20%)
2	MAN	C	4	2	11,11,12	0.61	0	15,15,17	1.19	3 (20%)

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	NAG	C	1	1,2	-	3/6/23/26	0/1/1/1
2	MAN	C	2	2	-	2/2/19/22	0/1/1/1
2	MAN	C	3	2	-	2/2/19/22	0/1/1/1
2	MAN	C	4	2	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	2	MAN	C1-C2-C3	5.66	116.63	109.67
2	C	3	MAN	O5-C5-C6	2.58	111.25	107.20
2	C	3	MAN	C1-O5-C5	2.58	115.69	112.19
2	C	1	NAG	C8-C7-N2	2.57	120.46	116.10

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	4	MAN	C1-O5-C5	2.46	115.52	112.19
2	C	2	MAN	O5-C5-C4	-2.45	104.88	110.83
2	C	1	NAG	C1-O5-C5	2.33	115.35	112.19
2	C	4	MAN	C3-C4-C5	2.32	114.38	110.24
2	C	1	NAG	O5-C1-C2	-2.26	107.73	111.29
2	C	3	MAN	C1-C2-C3	2.26	112.44	109.67
2	C	4	MAN	O5-C1-C2	-2.17	107.43	110.77

There are no chirality outliers.

All (7) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	2	MAN	C4-C5-C6-O6
2	C	1	NAG	C8-C7-N2-C2
2	C	1	NAG	O7-C7-N2-C2
2	C	2	MAN	O5-C5-C6-O6
2	C	3	MAN	C4-C5-C6-O6
2	C	1	NAG	O5-C5-C6-O6
2	C	3	MAN	O5-C5-C6-O6

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.

Oligosaccharide Chain C

Bond lengths Bond angles Torsions Rings
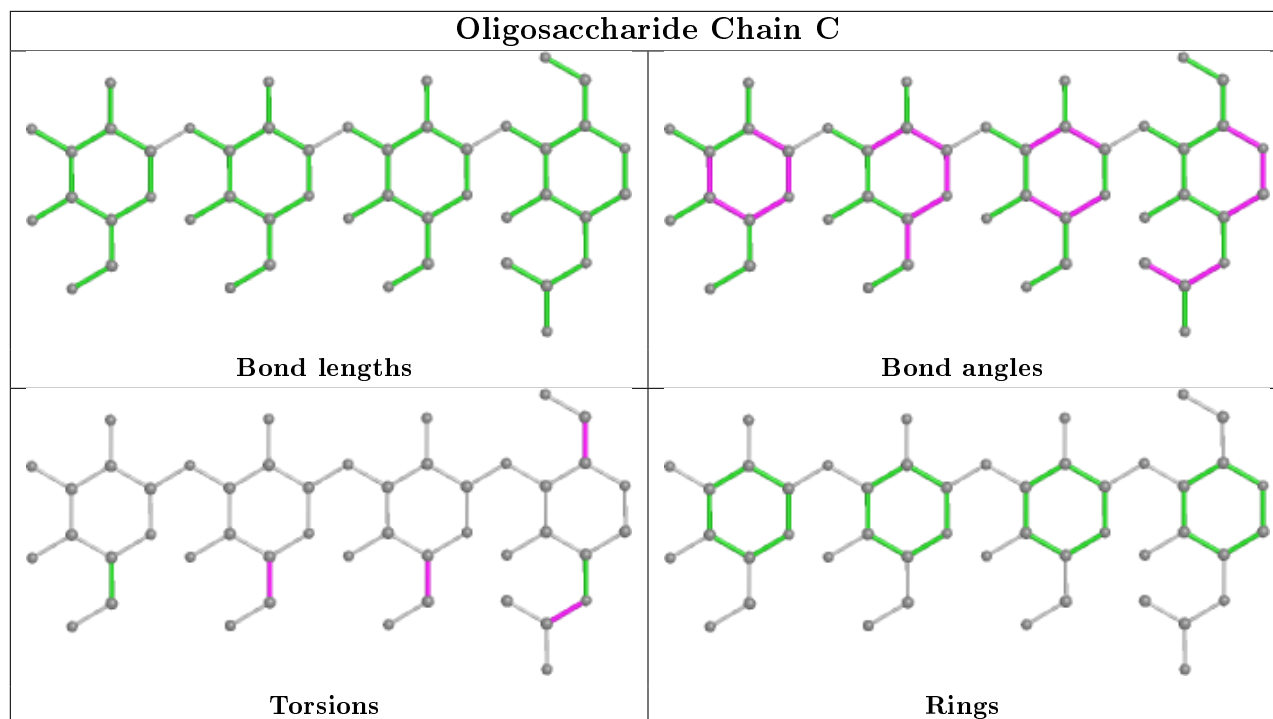
5.6 Ligand geometry (i)

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	HL1	A	801	-	31,31,31	1.84	4 (12%)	43,46,46	1.66	6 (13%)
3	HL1	B	801	-	31,31,31	1.87	6 (19%)	43,46,46	1.61	8 (18%)
4	NAG	A	802	1	14,14,15	0.46	0	17,19,21	1.45	1 (5%)
4	NAG	B	802	1	14,14,15	0.48	0	17,19,21	1.29	1 (5%)

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	HL1	A	801	-	-	2/8/20/20	0/4/4/4
3	HL1	B	801	-	-	0/8/20/20	0/4/4/4
4	NAG	A	802	1	-	0/6/23/26	0/1/1/1
4	NAG	B	802	1	-	2/6/23/26	0/1/1/1

All (10) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	801	HL1	C12-C13	6.54	1.49	1.38
3	B	801	HL1	C12-C13	6.25	1.48	1.38
3	B	801	HL1	C25-C24	4.01	1.49	1.40
3	A	801	HL1	C25-C24	3.71	1.48	1.40
3	B	801	HL1	C17-C20	3.55	1.49	1.43
3	A	801	HL1	C13-C17	3.52	1.49	1.43
3	A	801	HL1	C17-C20	3.45	1.49	1.43
3	B	801	HL1	C13-C17	3.41	1.48	1.43
3	B	801	HL1	C26-C25	2.25	1.40	1.36
3	B	801	HL1	C23-C24	2.10	1.40	1.36

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	801	HL1	C3-C2-C1	4.66	120.36	116.48
3	A	801	HL1	C31-O30-C25	4.56	124.41	117.53
3	B	801	HL1	C31-O30-C25	4.43	124.22	117.53
4	B	802	NAG	C1-O5-C5	4.43	118.19	112.19
4	A	802	NAG	C1-O5-C5	4.05	117.68	112.19
3	B	801	HL1	C3-C2-C1	4.05	119.85	116.48
3	B	801	HL1	C6-C1-C2	-3.79	119.29	123.83
3	A	801	HL1	C6-C1-C2	-3.61	119.51	123.83
3	A	801	HL1	C28-O27-C24	3.23	122.40	117.53
3	B	801	HL1	C28-O27-C24	2.95	121.98	117.53
3	A	801	HL1	O30-C25-C26	-2.65	121.81	125.24
3	A	801	HL1	O27-C24-C23	-2.48	122.03	125.24
3	B	801	HL1	O11-C12-C22	2.47	120.69	116.29
3	B	801	HL1	O27-C24-C23	-2.23	122.36	125.24
3	B	801	HL1	O11-C12-C13	-2.10	118.64	122.39
3	B	801	HL1	C5-C6-C1	2.04	120.71	117.10

There are no chirality outliers.

All (4) torsion outliers are listed below:

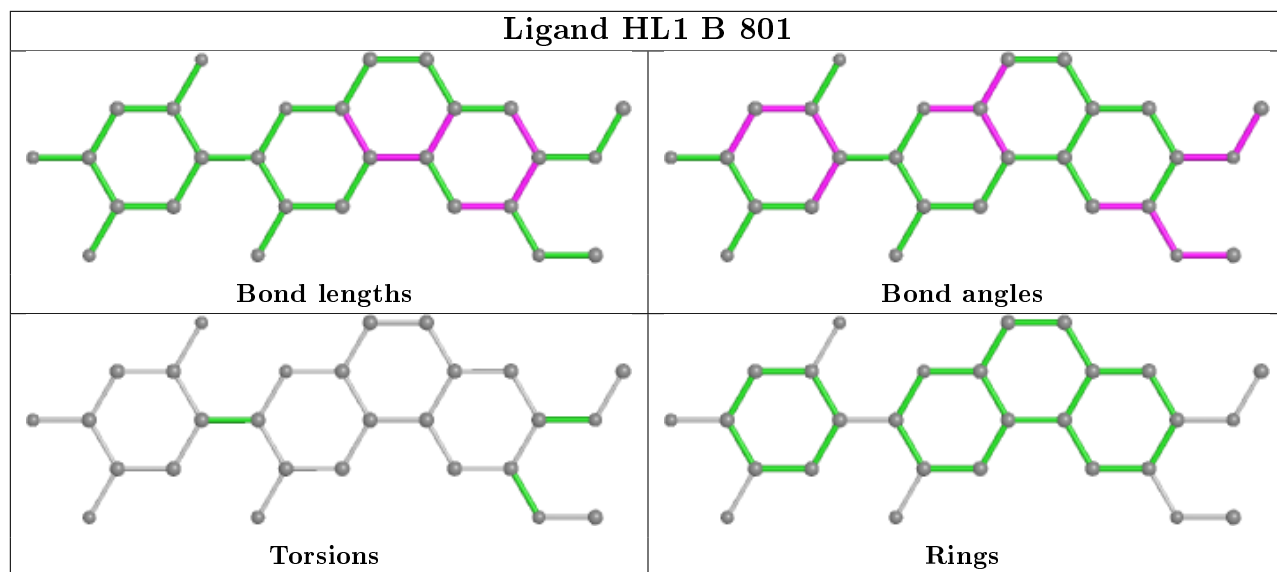| Mol	Chain	Res	Type	Atoms
4	B	802	NAG	O5-C5-C6-O6
3	A	801	HL1	C24-C25-O30-C31
4	B	802	NAG	C4-C5-C6-O6
3	A	801	HL1	C26-C25-O30-C31

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 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 HL1 A 801

Bond lengths   Bond angles   Torsions   Rings

Ligand HL1 B 801

Bond lengths

Bond angles

Torsions

Rings
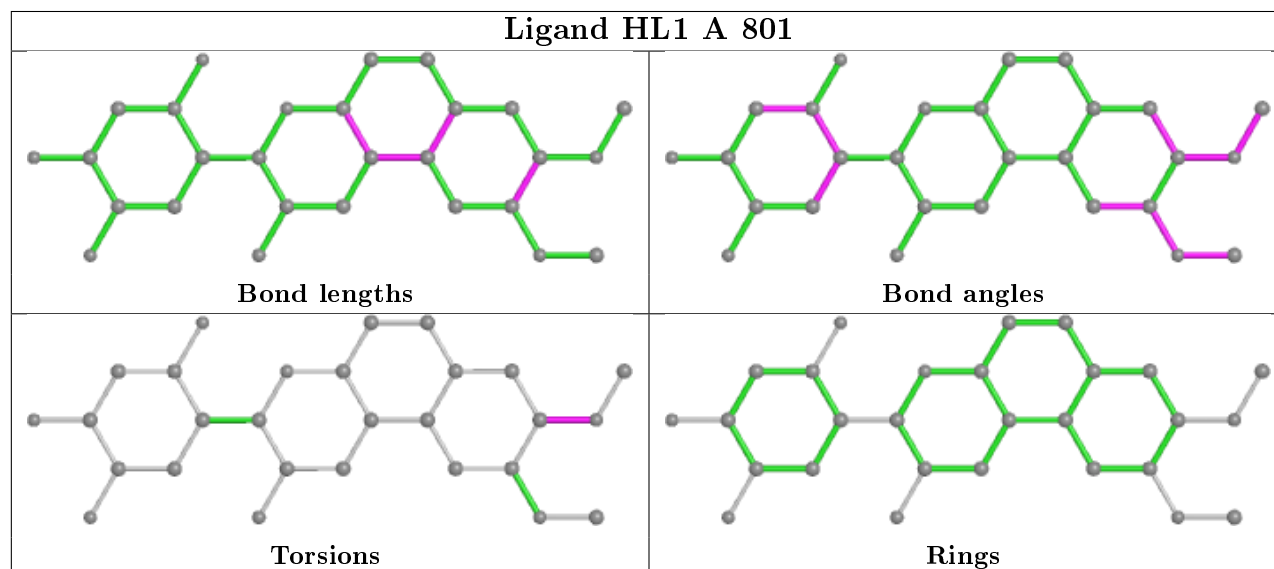
5.7 Other polymers (i)

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(Å²)	Q<0.9
1	A	727/734 (99%)	0.32	36 (4%)	28	35	16, 32, 59, 73	8 (1%)
1	B	727/734 (99%)	0.28	36 (4%)	28	35	16, 30, 64, 82	8 (1%)
All	All	1454/1468 (99%)	0.30	72 (4%)	28	35	16, 31, 61, 82	16 (1%)

All (72) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	88	VAL	5.4
1	B	766	PRO	5.2
1	A	766	PRO	4.6
1	A	87	SER	4.3
1	A	96	ASP	4.2
1	A	282	ALA	3.9
1	B	77	LEU	3.8
1	A	731	GLN	3.5
1	B	280	THR	3.5
1	B	95	PHE	3.4
1	A	333	SER	3.4
1	B	765	LEU	3.4
1	B	89	PHE	3.3
1	A	546	VAL	3.3
1	A	138	ASN	3.3
1	B	98	PHE	3.2
1	B	83	TYR	3.2
1	A	91	GLU	3.2
1	B	546	VAL	3.2
1	A	102	ILE	3.2
1	A	83	TYR	3.0
1	A	334	SER	2.8
1	A	40	ARG	2.8
1	B	75	ASN	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	97	GLU	2.8
1	A	243	ASP	2.8
1	B	40	ARG	2.7
1	B	74	ASN	2.7
1	B	92	ASN	2.7
1	A	473	SER	2.7
1	A	90	LEU	2.7
1	B	135	TYR	2.6
1	A	485	SER	2.6
1	A	89	PHE	2.6
1	A	98	PHE	2.6
1	B	392	LYS	2.6
1	A	95	PHE	2.5
1	B	91	GLU	2.5
1	B	279	VAL	2.5
1	B	136	ASP	2.5
1	B	93	SER	2.4
1	B	212	SER	2.4
1	A	100	HIS	2.4
1	B	473	SER	2.4
1	A	278	SER	2.4
1	A	615	LYS	2.4
1	B	243	ASP	2.4
1	A	88	VAL	2.4
1	B	187	TRP	2.4
1	B	731	GLN	2.3
1	A	505	GLN	2.3
1	B	73	GLU	2.3
1	B	283	THR	2.2
1	B	193	ILE	2.2
1	B	79	PHE	2.2
1	A	92	ASN	2.2
1	B	485	SER	2.2
1	A	141	GLN	2.2
1	A	537	SER	2.2
1	A	533	HIS	2.1
1	A	212	SER	2.1
1	A	279	VAL	2.1
1	B	145	GLU	2.1
1	A	338	ASN	2.1
1	A	332	GLU	2.1
1	A	335	GLY	2.1

Continued on next page...

Continued from previous page...

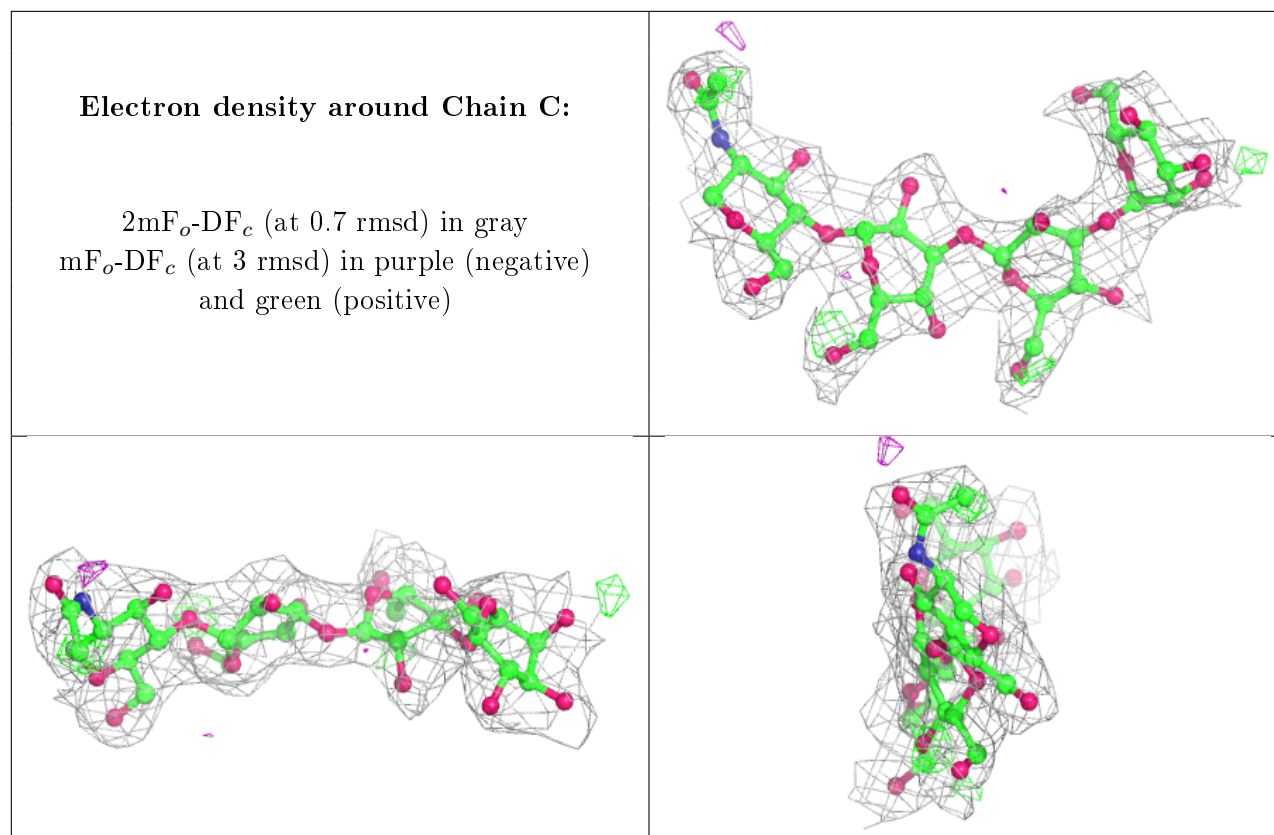| Mol	Chain	Res	Type	RSRZ
1	B	96	ASP	2.1
1	B	102	ILE	2.1
1	B	521	GLU	2.1
1	B	276	LEU	2.0
1	A	73	GLU	2.0
1	A	536	LYS	2.0

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)

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	MAN	C	2	11/12	0.77	0.21	46,48,52,53	0
2	MAN	C	3	11/12	0.79	0.28	54,56,57,59	0
2	MAN	C	4	11/12	0.80	0.24	58,59,60,60	0
2	NAG	C	1	14/15	0.91	0.17	36,39,42,44	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:

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

6.4 Ligands (i)

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.

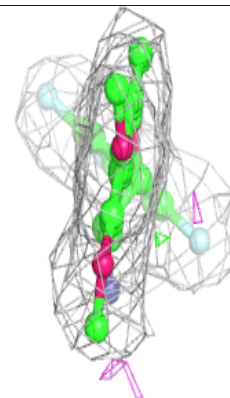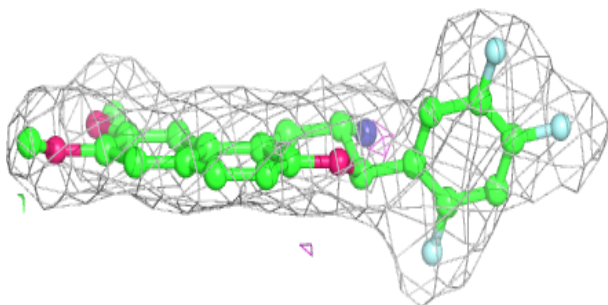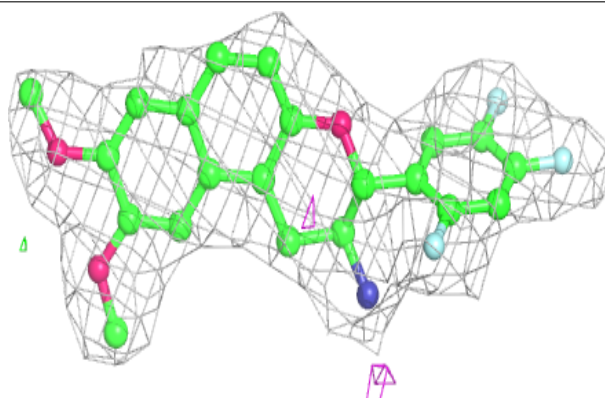| Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	NAG	A	802	14/15	0.89	0.23	35,35,37,37	0
4	NAG	B	802	14/15	0.89	0.28	52,53,54,54	0
3	HL1	A	801	28/28	0.94	0.17	23,24,26,26	0
3	HL1	B	801	28/28	0.96	0.17	18,18,18,19	0

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 HL1 A 801:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
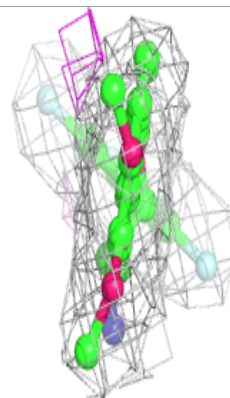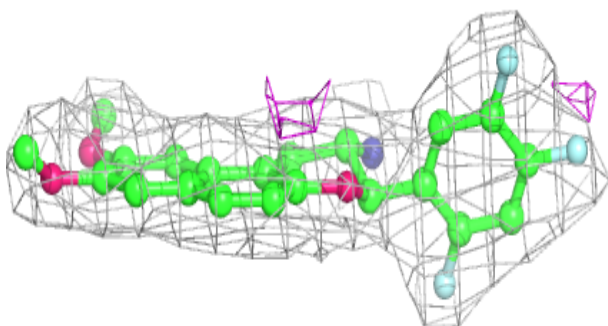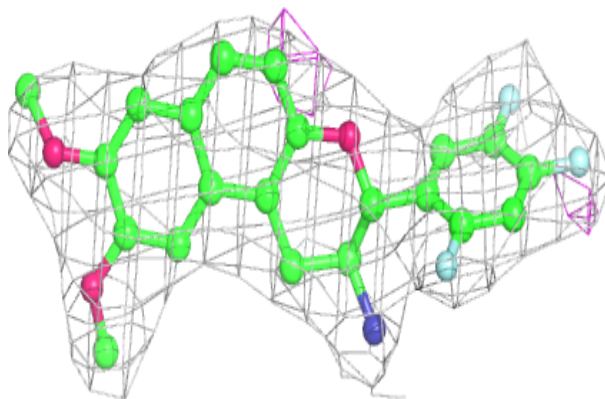
Electron density around HL1 B 801:

$2mF_o$-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.