Full wwPDB X-ray Structure Validation Report ⓘ

Oct 10, 2023 – 03:49 AM EDT

PDB ID	:	7SUS
Title	:	Crystal structure of Apelin receptor in complex with small molecule
Authors	:	Xu, F.; Yue, Y.; Liu, L.E.; Han, G.W.; Hanson, M.
Deposited on	:	2021-11-18
Resolution	:	2.70 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 ⓘ 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 ⓘ) were used in the production of this report:

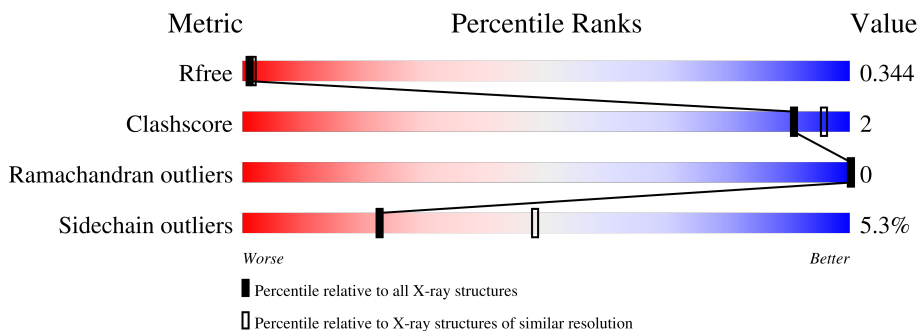|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 2.70 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		0.344
Clashscore		2
Ramachandran outliers		0
Sidechain outliers		5.3%

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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)

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%

Mol	Chain	Length	Quality of chain
1	A	407	78% 7% 15%

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 2738 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 Apelin receptor, with Rubredoxin insertion.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	347	2686	1765	418	475	28	0	0	0

There are 49 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-17	MET	-	initiating methionine	UNP P35414
A	-16	LYS	-	expression tag	UNP P35414
A	-15	THR	-	expression tag	UNP P35414
A	-14	ILE	-	expression tag	UNP P35414
A	-13	ILE	-	expression tag	UNP P35414
A	-12	ALA	-	expression tag	UNP P35414
A	-11	LEU	-	expression tag	UNP P35414
A	-10	SER	-	expression tag	UNP P35414
A	-9	TYR	-	expression tag	UNP P35414
A	-8	ILE	-	expression tag	UNP P35414
A	-7	PHE	-	expression tag	UNP P35414
A	-6	CYS	-	expression tag	UNP P35414
A	-5	LEU	-	expression tag	UNP P35414
A	-4	VAL	-	expression tag	UNP P35414
A	-3	PHE	-	expression tag	UNP P35414
A	-2	ALA	-	expression tag	UNP P35414
A	-1	ASP	-	expression tag	UNP P35414
A	0	TYR	-	expression tag	UNP P35414
A	1	LYS	-	expression tag	UNP P35414
A	2	ASP	-	expression tag	UNP P35414
A	3	ASP	-	expression tag	UNP P35414
A	4	ASP	-	expression tag	UNP P35414
A	5	ASP	-	expression tag	UNP P35414
A	6	LYS	-	expression tag	UNP P35414
A	117	ALA	VAL	engineered mutation	UNP P35414
A	174	CYS	GLU	engineered mutation	UNP P35414
A	177	ASN	THR	engineered mutation	UNP P35414

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	217	CYS	MET	engineered mutation	UNP P35414
A	250	CYS	ILE	engineered mutation	UNP P35414
A	325	LEU	CYS	engineered mutation	UNP P35414
A	326	MET	CYS	engineered mutation	UNP P35414
A	331	LEU	-	expression tag	UNP P35414
A	332	GLU	-	expression tag	UNP P35414
A	333	VAL	-	expression tag	UNP P35414
A	334	LEU	-	expression tag	UNP P35414
A	335	PHE	-	expression tag	UNP P35414
A	336	GLN	-	expression tag	UNP P35414
A	337	GLY	-	expression tag	UNP P35414
A	338	PRO	-	expression tag	UNP P35414
A	339	HIS	-	expression tag	UNP P35414
A	340	HIS	-	expression tag	UNP P35414
A	341	HIS	-	expression tag	UNP P35414
A	342	HIS	-	expression tag	UNP P35414
A	343	HIS	-	expression tag	UNP P35414
A	344	HIS	-	expression tag	UNP P35414
A	345	HIS	-	expression tag	UNP P35414
A	346	HIS	-	expression tag	UNP P35414
A	347	HIS	-	expression tag	UNP P35414
A	348	HIS	-	expression tag	UNP P35414

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
2	A	1	1	1	0	0

- Molecule 3 is (1R,2S)-N-[4-(2,6-dimethoxyphenyl)-5-(6-methylpyridin-2-yl)-1,2,4-triazol-3-yl]-1-(5-methylpyrimidin-2-yl)-1-oxidanyl-propane-2-sulfonamide (three-letter code: 8EH) (formula: $C_{24}H_{27}N_7O_5S$) (labeled as "Ligand of Interest" by depositor).

8EH
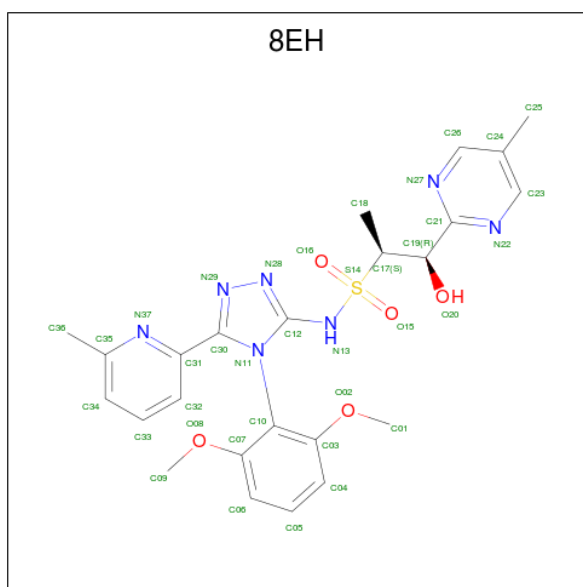
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total 37	C 24	N 7	O 5	S 1	0	0

- Molecule 4 is (2R)-2,3-dihydroxypropyl (9Z)-octadec-9-enoate (three-letter code: OLC) (formula: $C_{21}H_{40}O_4$).

OLC
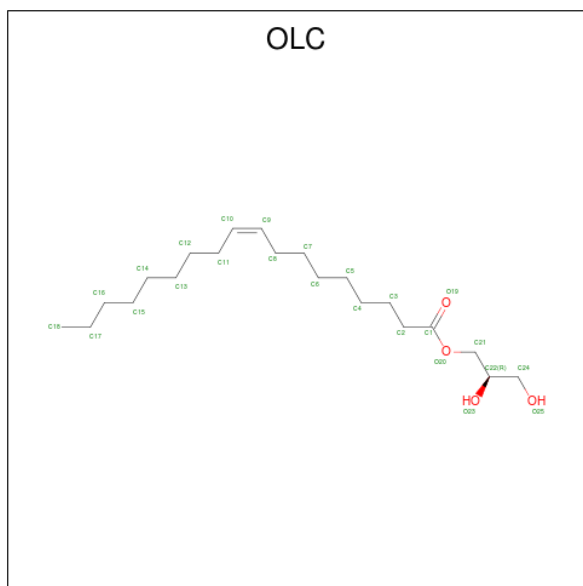
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total 13	C 9	O 4	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total 1	O 1	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. Residues
are color-coded according to the number of geometric quality criteria for which they contain at
least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more
consecutive residues without any outlier are shown as a green connector. Residues present in the
sample, but not in the model, are shown in grey.

• Molecule 1: Apelin receptor, with Rubredoxin insertion

Chain A:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	163.01Å 44.56Å 79.68Å 90.00° 115.62° 90.00°	Depositor
Resolution (Å)	30.00 – 2.70 42.64 – 2.70	Depositor EDS
% Data completeness (in resolution range)	91.7 (30.00-2.70) 91.7 (42.64-2.70)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.41 (at 2.69Å)	Xtriage
Refinement program	BUSTER 2.10.2	Depositor
R, R_{free}	0.252 , 0.274 0.302 , 0.344	Depositor DCC
R_{free} test set	645 reflections (4.85%)	wwPDB-VP
Wilson B-factor (Å^2)	73.3	Xtriage
Anisotropy	0.256	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.23 , 44.5	EDS
L-test for twinning[2]	$< |L| > = 0.43, < L^2 > = 0.25$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	2738	wwPDB-VP
Average B, all atoms (Å^2)	129.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.42% 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 ⓘ

5.1   Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: 8EH, OLC, ZN

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	0.40	0/2764	0.53	0/3784

There are no bond length outliers.

There are no bond angle outliers.

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	2686	0	2580	9	0
2	A	1	0	0	0	0
3	A	37	0	0	0	0
4	A	13	0	15	2	0
5	A	1	0	0	0	0
All	All	2738	0	2595	9	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 2.

All (9) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:201:LEU:HB3	4:A:1103:OLC:H21A	1.89	0.53
1:A:199:VAL:HG23	1:A:272:MET:HB3	1.92	0.51
1:A:137:ASN:HA	1:A:140:LEU:HD12	1.91	0.50
1:A:62:ARG:HD2	1:A:246:LEU:HD21	1.96	0.47
1:A:207:THR:HA	1:A:211:VAL:HB	1.96	0.46
1:A:160:LEU:HD23	4:A:1103:OLC:H5	2.00	0.44
1:A:305:ASN:HA	1:A:308:LEU:HD12	2.01	0.43
1:A:1013:TYR:CZ	1:A:1033:ILE:HD11	2.55	0.42
1:A:1030:PHE:HA	1:A:1033:ILE:HD12	2.00	0.41

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	345/407 (85%)	328 (95%)	17 (5%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains (i)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	285/358 (80%)	270 (95%)	15 (5%)	22	48

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	40	LEU
1	A	48	LEU
1	A	50	LEU
1	A	71	LEU
1	A	82	LEU
1	A	127	ARG
1	A	169	THR
1	A	261	TRP
1	A	268	LYS
1	A	284	ASP
1	A	324	LEU
1	A	326	MET
1	A	328	GLN
1	A	332	GLU
1	A	334	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1014	ASN

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 3 ligands modelled in this entry, 1 is monoatomic - leaving 2 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	OLC	A	1103	-	12,12,24	0.37	0	13,13,25	1.03	2 (15%)
3	8EH	A	1102	-	37,40,40	2.91	26 (70%)	36,58,58	3.50	18 (50%)

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	OLC	A	1103	-	-	5/12/12/24	-
3	8EH	A	1102	-	-	5/24/31/31	0/4/4/4

All (26) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1102	8EH	C10-N11	6.06	1.50	1.45
3	A	1102	8EH	C31-C30	5.18	1.57	1.47
3	A	1102	8EH	C21-C19	4.74	1.59	1.50
3	A	1102	8EH	C12-N13	4.37	1.46	1.36
3	A	1102	8EH	C12-N11	4.37	1.45	1.37
3	A	1102	8EH	C21-N22	4.23	1.41	1.34
3	A	1102	8EH	S14-N13	4.19	1.70	1.62
3	A	1102	8EH	O02-C03	3.69	1.43	1.37
3	A	1102	8EH	O08-C07	3.43	1.42	1.37
3	A	1102	8EH	C33-C32	3.16	1.45	1.38
3	A	1102	8EH	C32-C31	3.07	1.46	1.38
3	A	1102	8EH	C30-N29	2.96	1.40	1.33
3	A	1102	8EH	C12-N28	2.76	1.39	1.33
3	A	1102	8EH	C34-C35	2.61	1.45	1.39
3	A	1102	8EH	C23-N22	2.57	1.39	1.34
3	A	1102	8EH	C05-C06	2.53	1.44	1.38
3	A	1102	8EH	C23-C24	2.45	1.44	1.38
3	A	1102	8EH	C26-C24	2.42	1.44	1.38
3	A	1102	8EH	C36-C35	2.40	1.55	1.50
3	A	1102	8EH	C05-C04	2.35	1.43	1.38
3	A	1102	8EH	C10-C07	2.32	1.47	1.40
3	A	1102	8EH	C10-C03	2.28	1.47	1.40

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1102	8EH	O15-S14	2.20	1.46	1.43
3	A	1102	8EH	C06-C07	2.18	1.44	1.39
3	A	1102	8EH	C04-C03	2.15	1.44	1.39
3	A	1102	8EH	C31-N37	2.09	1.38	1.34

All (20) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1102	8EH	O16-S14-O15	-15.42	107.67	119.24
3	A	1102	8EH	C31-N37-C35	5.35	123.98	118.45
3	A	1102	8EH	C03-C10-N11	4.63	122.07	118.64
3	A	1102	8EH	C07-C10-N11	4.19	121.74	118.64
3	A	1102	8EH	N27-C21-N22	-4.11	120.53	125.95
3	A	1102	8EH	C24-C26-N27	-3.37	120.85	124.81
3	A	1102	8EH	C24-C23-N22	-3.32	120.90	124.81
3	A	1102	8EH	C26-N27-C21	3.22	121.50	116.00
3	A	1102	8EH	O08-C07-C06	-3.15	118.97	124.37
3	A	1102	8EH	O08-C07-C10	2.94	119.33	116.65
3	A	1102	8EH	O02-C03-C04	-2.77	119.63	124.37
3	A	1102	8EH	C12-N28-N29	2.65	109.70	104.83
3	A	1102	8EH	C32-C31-N37	-2.54	118.67	121.97
3	A	1102	8EH	C09-O08-C07	2.54	121.37	117.53
3	A	1102	8EH	C04-C03-C10	2.46	122.38	118.73
4	A	1103	OLC	O20-C1-C2	2.42	119.49	111.91
3	A	1102	8EH	C03-C10-C07	-2.37	116.10	119.54
3	A	1102	8EH	C30-C31-N37	2.33	120.76	116.51
3	A	1102	8EH	C23-N22-C21	2.12	119.62	116.00
4	A	1103	OLC	O20-C1-O19	-2.09	118.31	123.59

There are no chirality outliers.

All (10) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	1103	OLC	O20-C21-C22-C24
4	A	1103	OLC	O19-C1-O20-C21
4	A	1103	OLC	C2-C1-O20-C21
4	A	1103	OLC	O20-C21-C22-O23
3	A	1102	8EH	C10-C03-O02-C01
3	A	1102	8EH	C10-C07-O08-C09
3	A	1102	8EH	C04-C03-O02-C01
3	A	1102	8EH	C06-C07-O08-C09
4	A	1103	OLC	C1-C2-C3-C4

Continued on next page...

Continued from previous page...

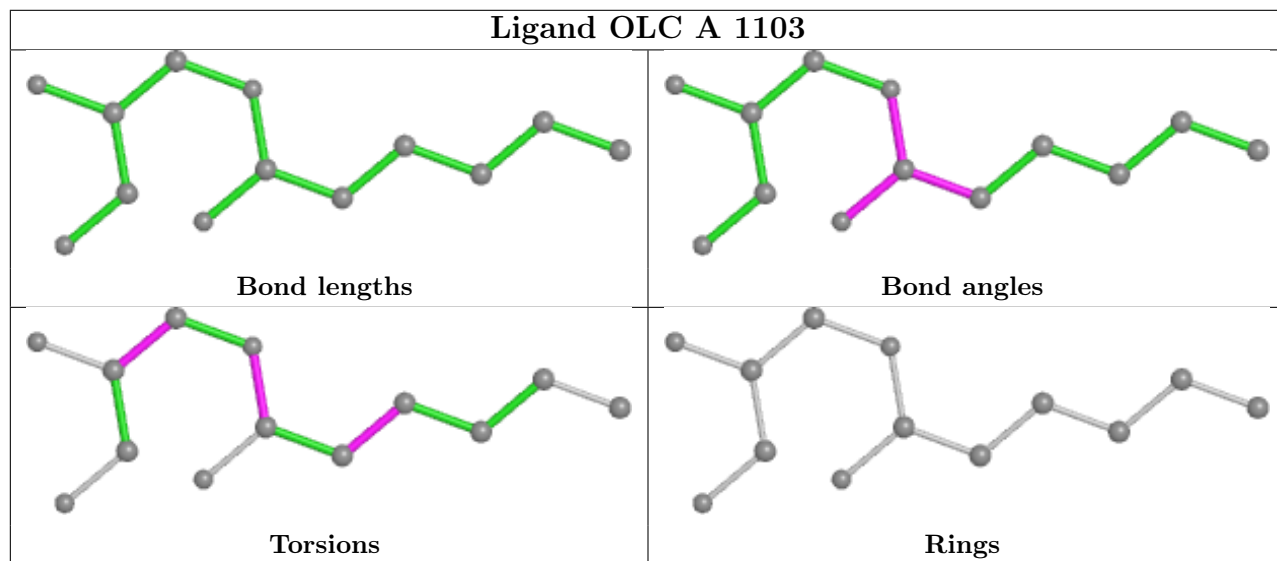| Mol	Chain	Res	Type	Atoms
3	A	1102	8EH	N29-C30-C31-N37

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1103	OLC	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 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 OLC A 1103

Bond lengths

Bond angles

Torsions

Rings

Ligand 8EH A 1102

Bond lengths

Bond angles

Torsions

Rings
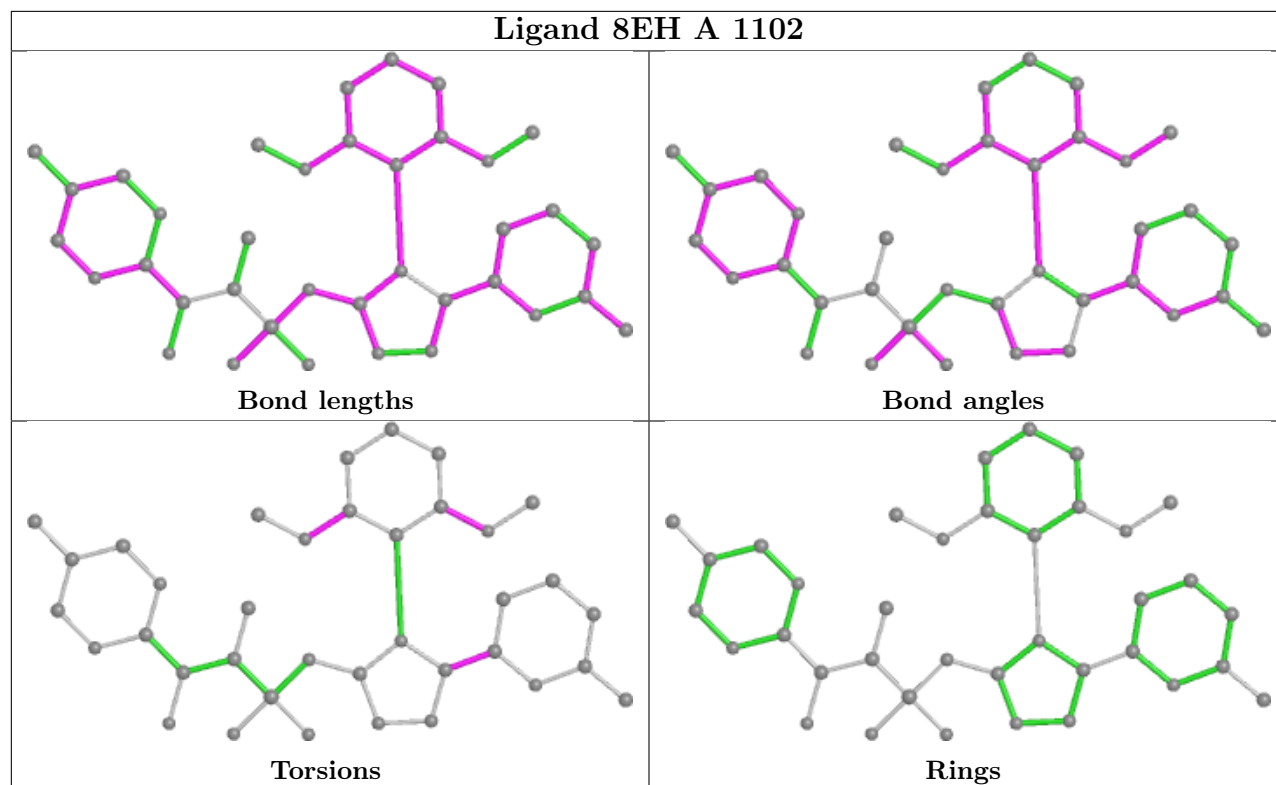
5.7 Other polymers ⓘ

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 ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

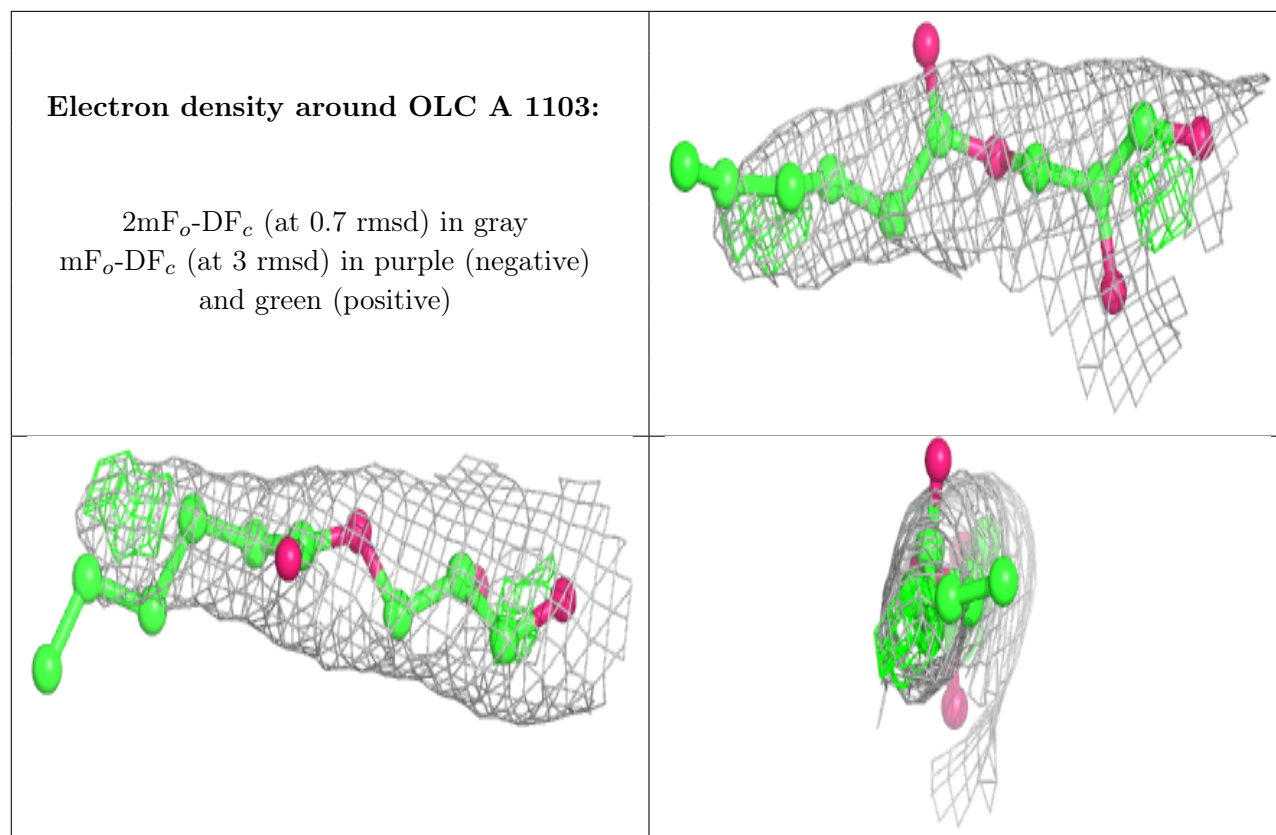Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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 8EH A 1102:

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

Electron density around OLC A 1103:

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

6.5 Other polymers (i)

Unable to reproduce the depositors R factor - this section is therefore empty.