wwPDB X-ray Structure Validation Summary Report (i)

Aug 8, 2020 – 10:00 PM BST

PDB ID	:	2G1Y
Title	:	Ketopiperazine-Based Renin Inhibitors: Optimization of the "C" Ring
Authors	:	Holsworth, D.D.; Jalaiea, M.; Zhanga, E.; Mcconnella, P.
Deposited on	:	2006-02-15
Resolution	:	2.50 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 (i) symbol.

The following versions of software and data (see references (i)) were used in the production of this report:

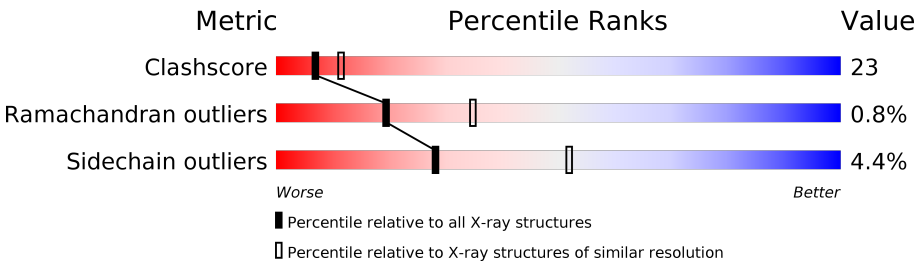
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & 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 2.50 Å.

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



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)

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%

Mol	Chain	Length	Quality of chain
1	A	333	64% 32% •
1	B	333	61% 35% •

2 Entry composition (i)

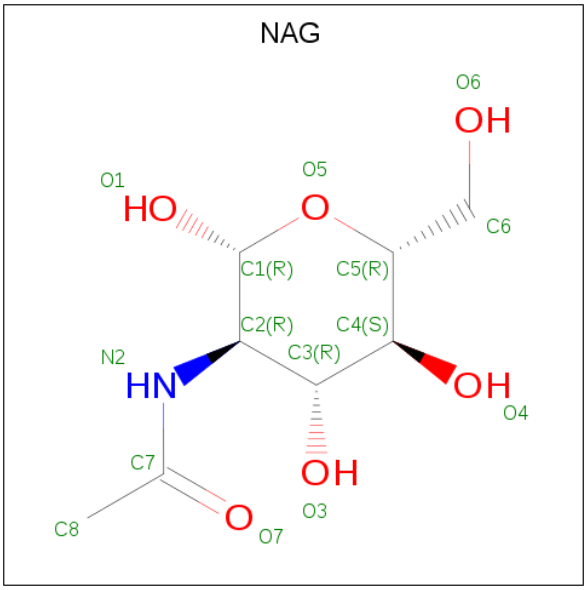
There are 4 unique types of molecules in this entry. The entry contains 5550 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Renin.

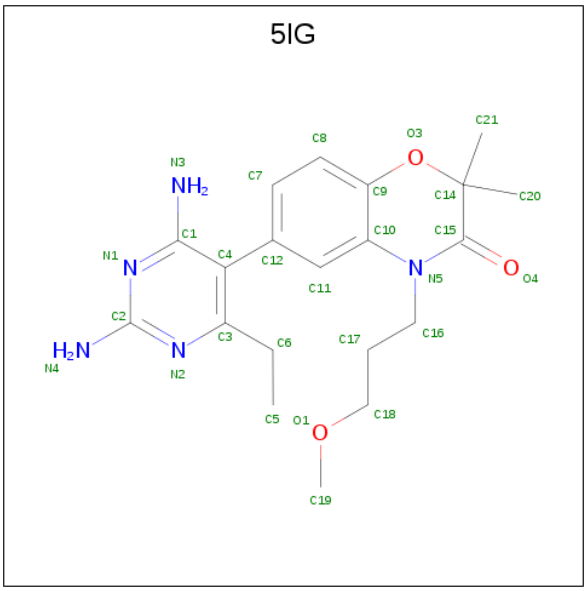
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	333	2572	1642	416	500	14	0	0	0
1	B	333	2572	1642	416	500	14	0	0	0

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



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

- Molecule 3 is 6-(2,4-DIAMINO-6-ETHYLPYRIMIDIN-5-YL)-4-(3-METHOXYPROPYL)-2,2-DIMETHYL-2H-1,4-BENZOXAZIN-3(4H)-ONE (three-letter code: 5IG) (formula: $C_{20}H_{27}N_5O_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	28	20	5	3	0	0
			Total	C	N	O		
3	B	1	28	20	5	3	0	0

- Molecule 4 is water.

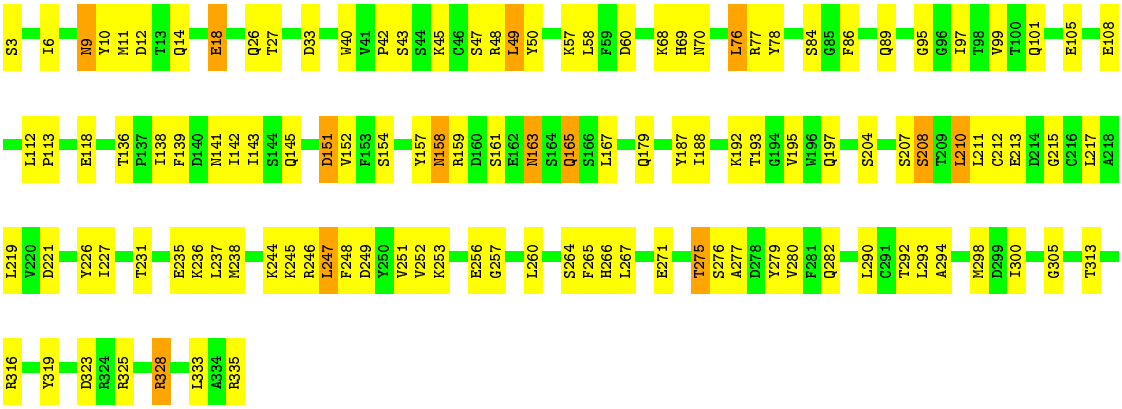
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	198	198	198	0	0
			Total	O		
4	B	124	124	124	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

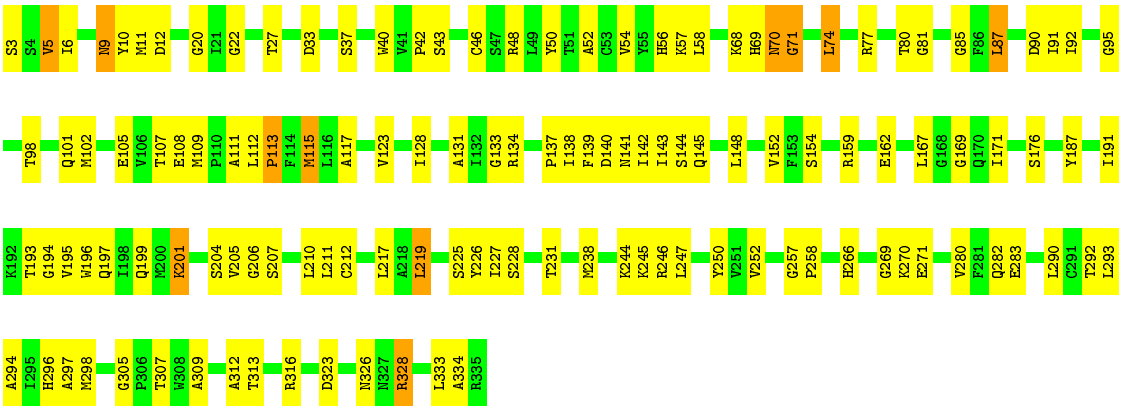
- Molecule 1: Renin

Chain A:



- Molecule 1: Renin

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 3	Depositor
Cell constants a, b, c, α, β, γ	141.04Å 141.04Å 141.04Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.50 49.86 – 2.50	Depositor EDS
% Data completeness (in resolution range)	(Not available) (50.00-2.50) 94.2 (49.86-2.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.14 (at 2.51Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.197 , 0.252 0.203 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	41.5	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 46.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.033 for l,-k,h	Xtriage
F_o,F_c correlation	0.68	EDS
Total number of atoms	5550	wwPDB-VP
Average B, all atoms (Å^2)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.52% 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, 5IG

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	0.33	0/2632	0.66	0/3568
1	B	0.32	0/2632	0.63	0/3568
All	All	0.33	0/5264	0.65	0/7136

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 (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	2572	0	2500	110	0
1	B	2572	0	2500	127	0
2	A	14	0	13	0	0
2	B	14	0	13	5	0
3	A	28	0	27	1	0
3	B	28	0	27	2	0
4	A	198	0	0	18	0
4	B	124	0	0	19	0
All	All	5550	0	5080	237	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 23.

The worst 5 of 237 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:112:LEU:HG	4:A:1158:HOH:O	1.52	1.07
1:A:45:LYS:HD2	4:A:1053:HOH:O	1.62	0.99
1:B:326:ASN:HB2	1:B:328:ARG:HD2	1.54	0.90
1:A:276:SER:O	1:A:280:VAL:HG12	1.74	0.88
1:A:158:ASN:ND2	1:A:159:ARG:H	1.73	0.86

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	331/333 (99%)	312 (94%)	17 (5%)	2 (1%)	25	43
1	B	331/333 (99%)	310 (94%)	18 (5%)	3 (1%)	17	31
All	All	662/666 (99%)	622 (94%)	35 (5%)	5 (1%)	19	35

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	165	GLN
1	B	70	ASN
1	A	208	SER
1	B	246	ARG
1	B	71	GLY

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	284/284 (100%)	270 (95%)	14 (5%)	25	47
1	B	284/284 (100%)	273 (96%)	11 (4%)	32	57
All	All	568/568 (100%)	543 (96%)	25 (4%)	28	52

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	275	THR
1	A	333	LEU
1	B	290	LEU
1	A	328	ARG
1	B	5	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	199	GLN
1	B	9	ASN
1	B	282	GLN
1	A	266	HIS
1	A	282	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

4 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	5IG	B	886	-	29,30,30	1.05	3 (10%)	41,44,44	1.52	7 (17%)
2	NAG	B	986	1	14,14,15	0.62	0	17,19,21	0.50	0
2	NAG	A	985	1	14,14,15	0.50	0	17,19,21	0.75	0
3	5IG	A	885	-	29,30,30	0.93	1 (3%)	41,44,44	1.49	6 (14%)

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	5IG	B	886	-	-	0/11/30/30	0/3/3/3
2	NAG	B	986	1	-	6/6/23/26	0/1/1/1
2	NAG	A	985	1	-	4/6/23/26	0/1/1/1
3	5IG	A	885	-	-	0/11/30/30	0/3/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	886	5IG	C8-C7	2.24	1.42	1.38
3	B	886	5IG	C15-N5	2.09	1.39	1.36
3	B	886	5IG	C10-C9	2.08	1.43	1.40
3	A	885	5IG	C10-C9	2.01	1.43	1.40

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	886	5IG	C2-N2-C3	5.05	120.45	116.24
3	A	885	5IG	C2-N2-C3	4.63	120.10	116.24
3	A	885	5IG	C4-C3-N2	-3.53	119.26	123.61

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	885	5IG	C11-C12-C4	-3.30	115.22	120.57
3	B	886	5IG	C4-C3-N2	-3.26	119.59	123.61

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	985	NAG	C8-C7-N2-C2
2	A	985	NAG	O7-C7-N2-C2
2	B	986	NAG	C3-C2-N2-C7
2	B	986	NAG	C8-C7-N2-C2
2	B	986	NAG	O7-C7-N2-C2

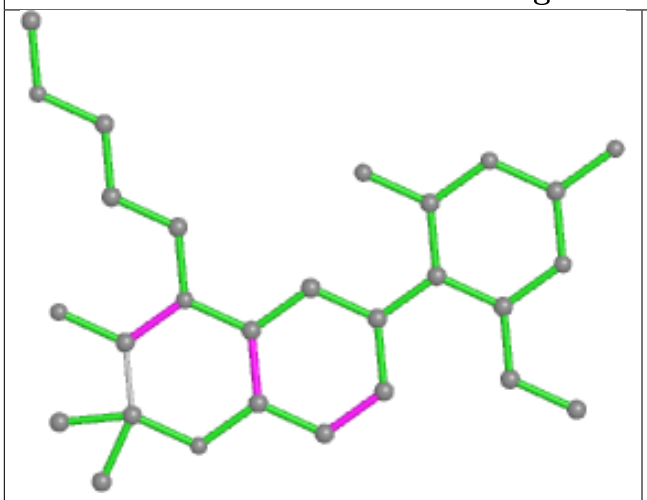
There are no ring outliers.

3 monomers are involved in 8 short contacts:

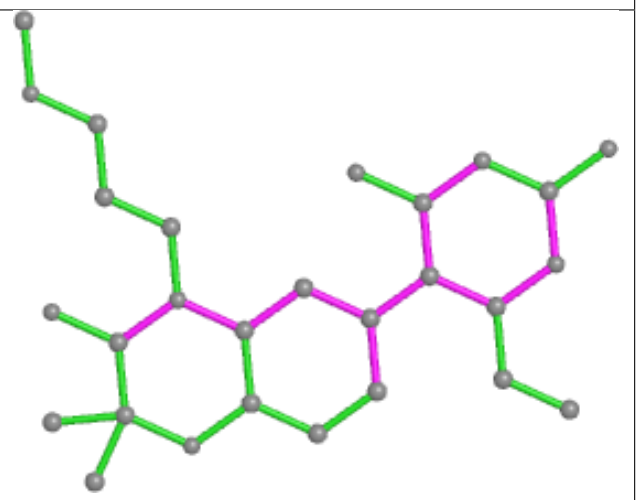
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	886	5IG	2	0
2	B	986	NAG	5	0
3	A	885	5IG	1	0

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.

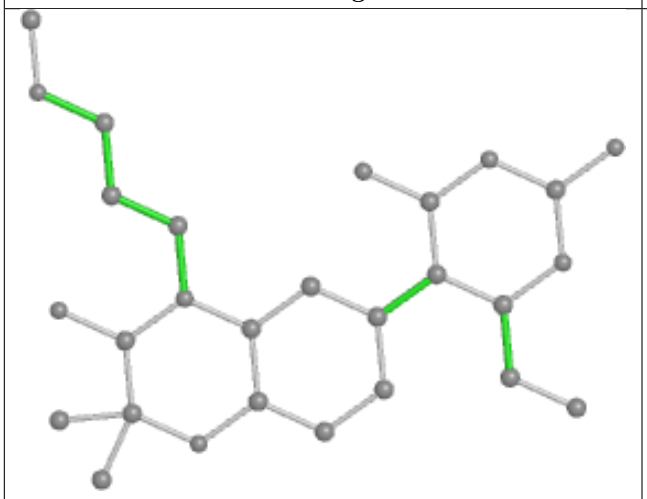
Ligand 5IG B 886



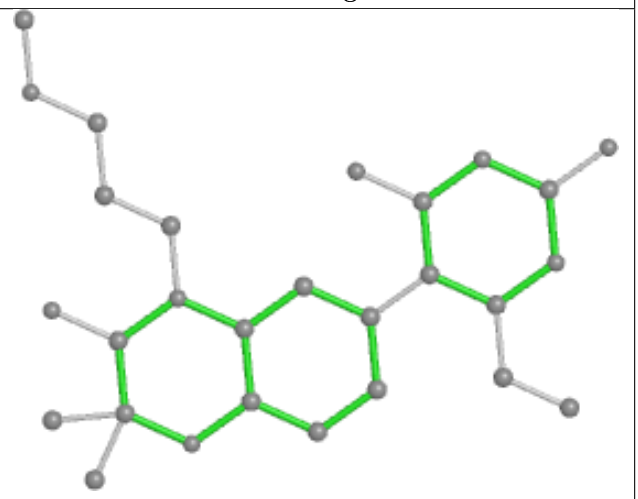
Bond lengths



Bond angles

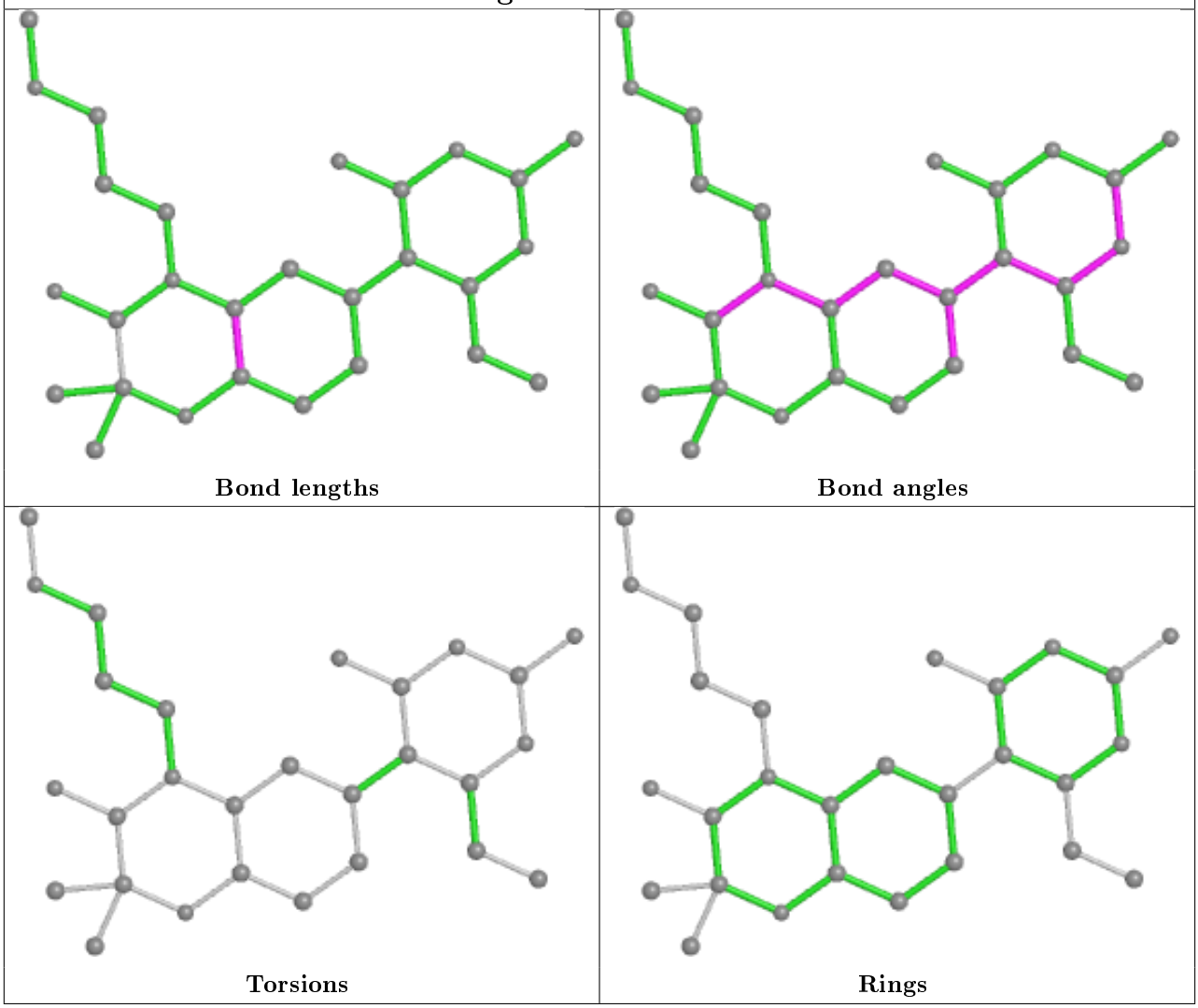


Torsions



Rings

Ligand 5IG A 885



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)

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

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

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

6.3 Carbohydrates (i)

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

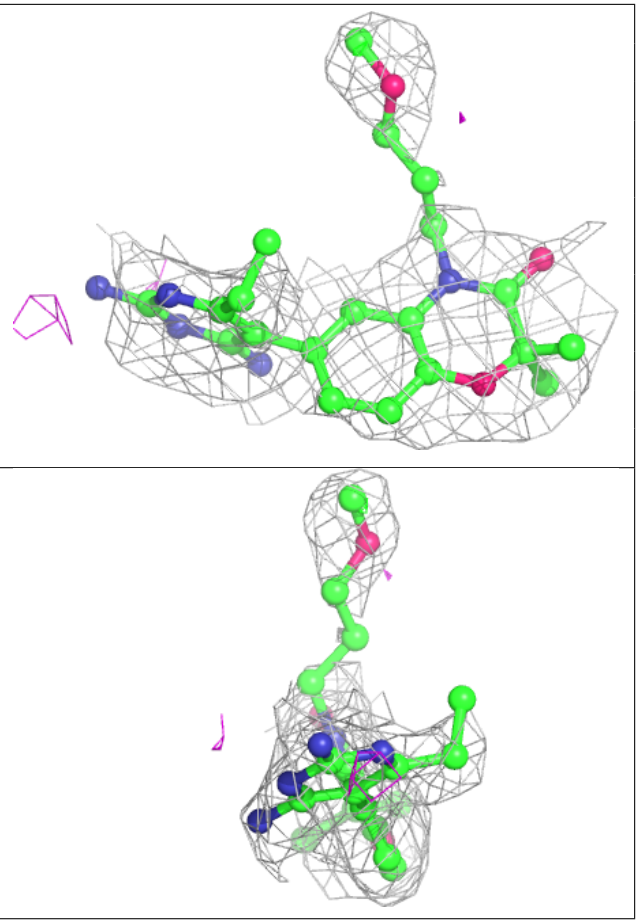
6.4 Ligands (i)

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.

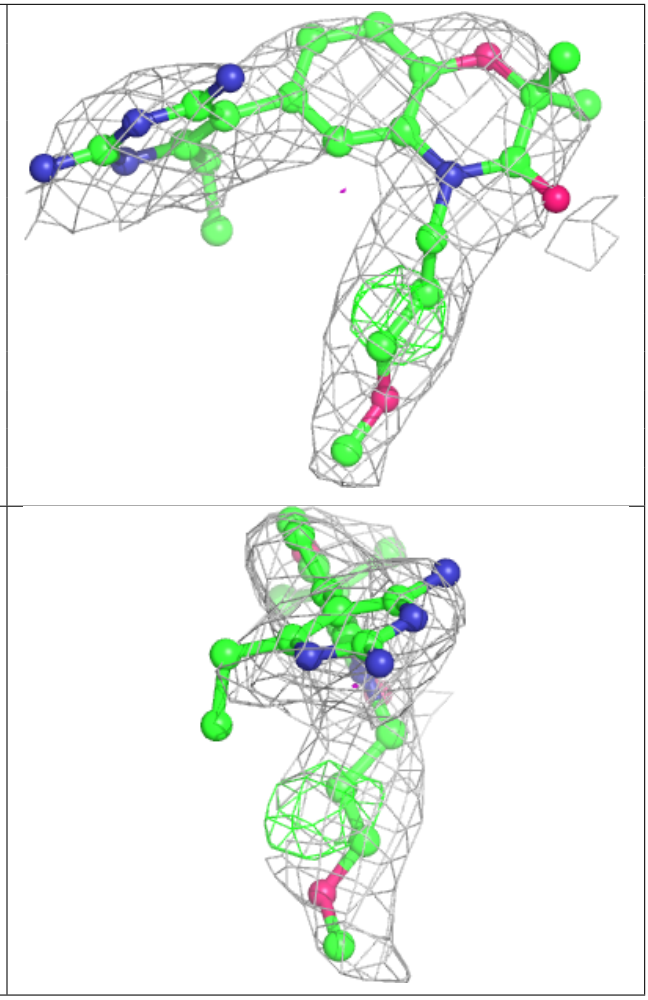
Electron density around 5IG A 885:

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



Electron density around 5IG B 886:

$2mF_o$-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.