



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 11, 2021 – 06:59 AM EDT

PDB ID : 3CCZ  
Title : Thermodynamic and structure guided design of statin hmg-coa reductase inhibitors  
Authors : Pavlovsky, A.; Sarver, R.W.; Harris, M.S.; Finzel, B.C.  
Deposited on : 2008-02-26  
Resolution : 1.70 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](mailto: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

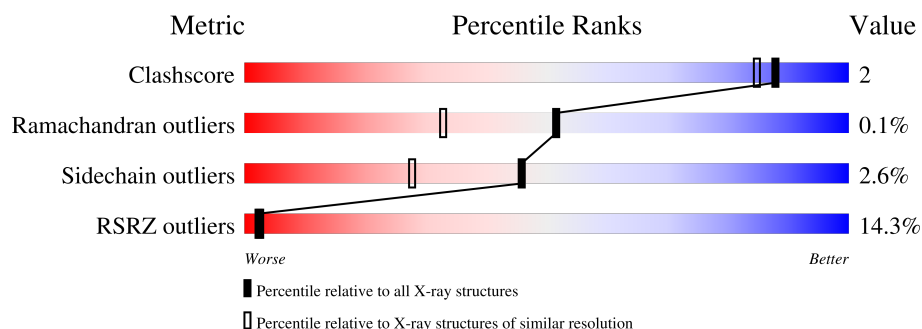
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.70 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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  $\geq 3$ , 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	441	<div> <div>14%</div> <div>91%</div> <div>5%</div> </div>
1	B	441	<div> <div>14%</div> <div>87%</div> <div>5%</div> <div>8%</div> </div>
1	C	441	<div> <div>14%</div> <div>87%</div> <div>8%</div> </div>
1	D	441	<div> <div>11%</div> <div>84%</div> <div>5%</div> <div>11%</div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 13469 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 3-hydroxy-3-methylglutaryl-coenzyme A reductase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	420	Total	C	N	O	S	0	3	0
			3128	1946	549	601	32			
1	B	405	Total	C	N	O	S	0	3	0
			3014	1874	529	580	31			
1	C	404	Total	C	N	O	S	0	3	0
			2997	1862	524	579	32			
1	D	394	Total	C	N	O	S	0	4	0
			2922	1815	512	565	30			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	435	HIS	-	expression tag	UNP P04035
A	436	HIS	-	expression tag	UNP P04035
A	437	HIS	-	expression tag	UNP P04035
A	438	HIS	-	expression tag	UNP P04035
A	439	HIS	-	expression tag	UNP P04035
A	440	HIS	-	expression tag	UNP P04035
A	485	ILE	MET	engineered mutation	UNP P04035
B	435	HIS	-	expression tag	UNP P04035
B	436	HIS	-	expression tag	UNP P04035
B	437	HIS	-	expression tag	UNP P04035
B	438	HIS	-	expression tag	UNP P04035
B	439	HIS	-	expression tag	UNP P04035
B	440	HIS	-	expression tag	UNP P04035
B	485	ILE	MET	engineered mutation	UNP P04035
C	435	HIS	-	expression tag	UNP P04035
C	436	HIS	-	expression tag	UNP P04035
C	437	HIS	-	expression tag	UNP P04035
C	438	HIS	-	expression tag	UNP P04035
C	439	HIS	-	expression tag	UNP P04035
C	440	HIS	-	expression tag	UNP P04035
C	485	ILE	MET	engineered mutation	UNP P04035

*Continued on next page...*

Continued from previous page...

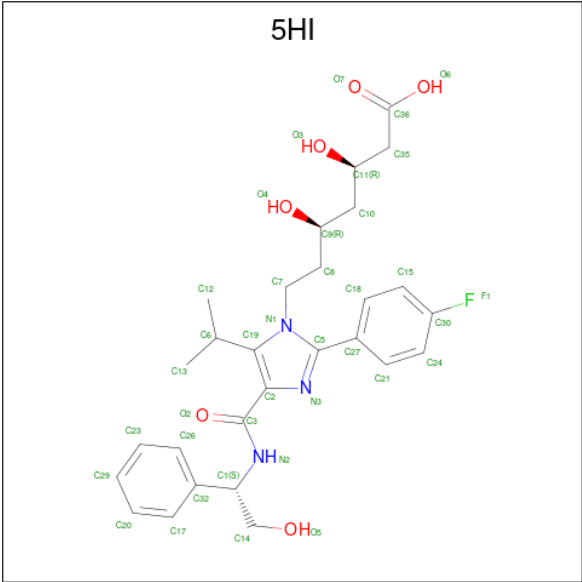
Chain	Residue	Modelled	Actual	Comment	Reference
D	435	HIS	-	expression tag	UNP P04035
D	436	HIS	-	expression tag	UNP P04035
D	437	HIS	-	expression tag	UNP P04035
D	438	HIS	-	expression tag	UNP P04035
D	439	HIS	-	expression tag	UNP P04035
D	440	HIS	-	expression tag	UNP P04035
D	485	ILE	MET	engineered mutation	UNP P04035

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	C	1	Total	O	S	0	0
			5	4	1		
2	D	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is (3R,5R)-7-[2-(4-fluorophenyl)-4-{[(1S)-2-hydroxy-1-phenylethyl]carbamoyl}-5-(1-methylethyl)-1H-imidazol-1-yl]-3,5-dihydroxyheptanoic acid (three-letter code: 5HI) (formula: C<sub>28</sub>H<sub>34</sub>FN<sub>3</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	F	N	O	0	0
			38	28	1	3	6		
3	B	1	Total	C	F	N	O	0	0
			38	28	1	3	6		
3	C	1	Total	C	F	N	O	0	0
			38	28	1	3	6		
3	D	1	Total	C	F	N	O	0	0
			38	28	1	3	6		

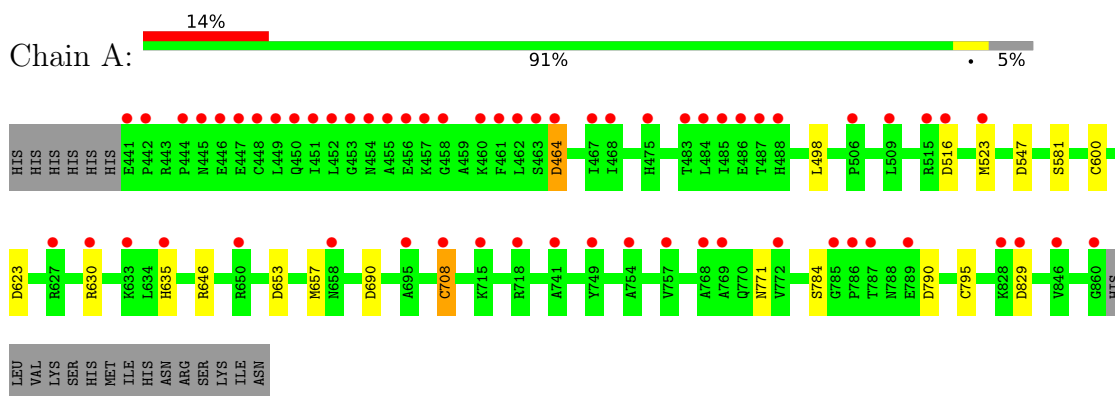
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	335	Total	O	0	0
			335	335		
4	B	316	Total	O	0	0
			316	316		
4	C	293	Total	O	0	0
			293	293		
4	D	292	Total	O	0	0
			292	292		

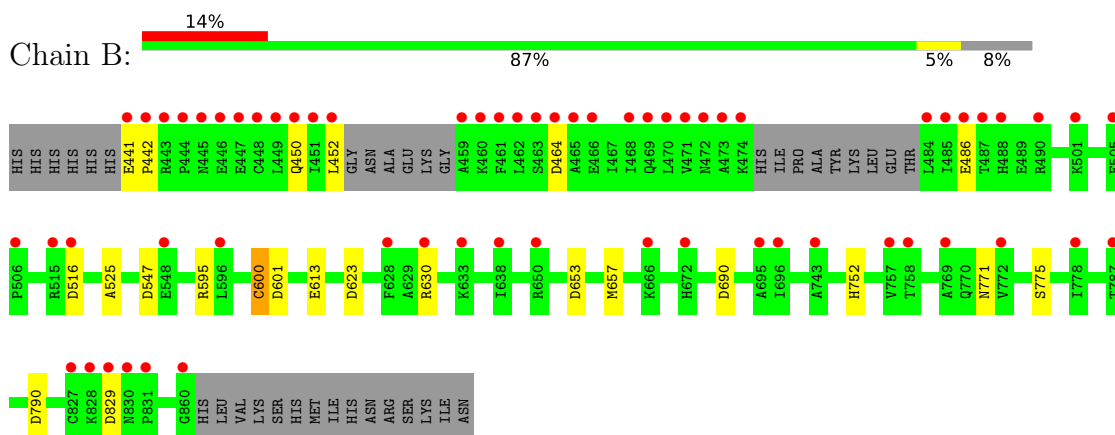
### 3 Residue-property plots

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.

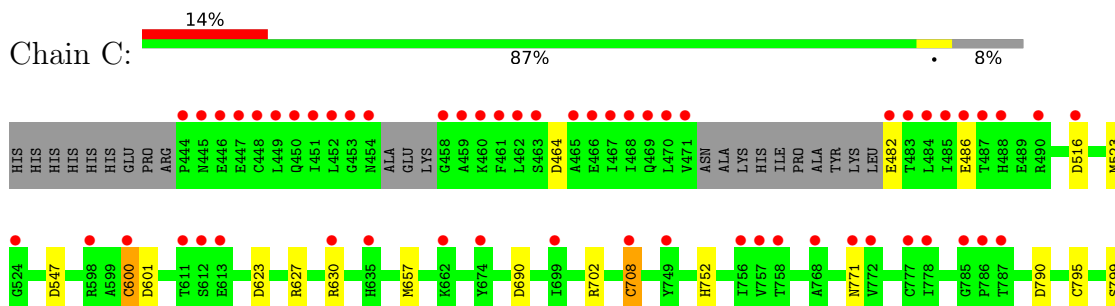
- Molecule 1: 3-hydroxy-3-methylglutaryl-coenzyme A reductase

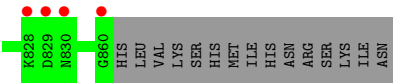


- Molecule 1: 3-hydroxy-3-methylglutaryl-coenzyme A reductase

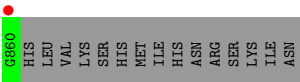
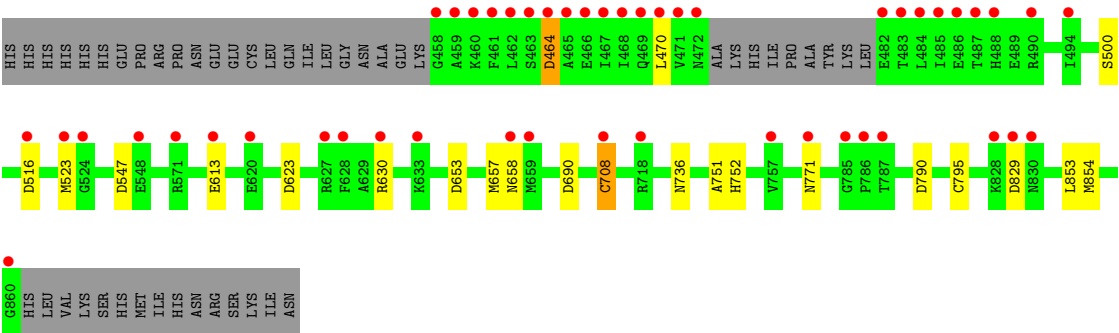
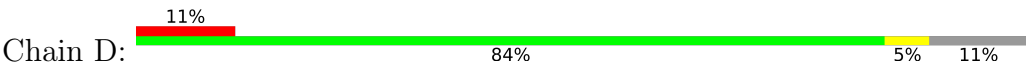


- Molecule 1: 3-hydroxy-3-methylglutaryl-coenzyme A reductase





● Molecule 1: 3-hydroxy-3-methylglutaryl-coenzyme A reductase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	82.86Å 135.66Å 83.15Å 90.00° 97.16° 90.00°	Depositor
Resolution (Å)	50.00 – 1.70 41.11 – 1.70	Depositor EDS
% Data completeness (in resolution range)	93.6 (50.00-1.70) 91.0 (41.11-1.70)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	7.72 (at 1.70Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.232 , 0.257 0.237 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	19.8	Xtriage
Anisotropy	0.152	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 46.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.054 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	13469	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.64% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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 ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: 5HI, SO4

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	0.26	0/3188	0.57	10/4310 (0.2%)
1	B	0.26	0/3069	0.56	10/4146 (0.2%)
1	C	0.25	0/3051	0.57	8/4121 (0.2%)
1	D	0.26	0/2981	0.57	8/4028 (0.2%)
All	All	0.26	0/12289	0.57	36/16605 (0.2%)

There are no bond length outliers.

All (36) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	547	ASP	CB-CG-OD2	5.65	123.39	118.30
1	A	547	ASP	CB-CG-OD2	5.62	123.35	118.30
1	C	547	ASP	CB-CG-OD2	5.61	123.35	118.30
1	D	547	ASP	CB-CG-OD2	5.48	123.23	118.30
1	A	790	ASP	CB-CG-OD2	5.48	123.23	118.30
1	D	790	ASP	CB-CG-OD2	5.45	123.20	118.30
1	A	600[A]	CYS	CA-CB-SG	-5.35	104.37	114.00
1	A	600[B]	CYS	CA-CB-SG	-5.35	104.37	114.00
1	B	790	ASP	CB-CG-OD2	5.34	123.10	118.30
1	C	790	ASP	CB-CG-OD2	5.33	123.09	118.30
1	D	623	ASP	CB-CG-OD2	5.32	123.09	118.30
1	C	600[A]	CYS	CA-CB-SG	-5.32	104.43	114.00
1	C	600[B]	CYS	CA-CB-SG	-5.32	104.43	114.00
1	A	653	ASP	CB-CG-OD2	5.28	123.06	118.30
1	C	516	ASP	CB-CG-OD2	5.28	123.05	118.30
1	D	516	ASP	CB-CG-OD2	5.26	123.03	118.30
1	A	690	ASP	CB-CG-OD2	5.23	123.00	118.30
1	A	623	ASP	CB-CG-OD2	5.22	123.00	118.30
1	C	623	ASP	CB-CG-OD2	5.22	123.00	118.30
1	A	829	ASP	CB-CG-OD2	5.21	122.99	118.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	600[A]	CYS	CA-CB-SG	-5.21	104.63	114.00
1	B	600[B]	CYS	CA-CB-SG	-5.21	104.63	114.00
1	B	690	ASP	CB-CG-OD2	5.18	122.96	118.30
1	B	653	ASP	CB-CG-OD2	5.15	122.93	118.30
1	C	464	ASP	CB-CG-OD2	5.12	122.91	118.30
1	B	516	ASP	CB-CG-OD2	5.12	122.90	118.30
1	B	829	ASP	CB-CG-OD2	5.10	122.89	118.30
1	D	829	ASP	CB-CG-OD2	5.10	122.89	118.30
1	D	464	ASP	CB-CG-OD2	5.09	122.89	118.30
1	D	653	ASP	CB-CG-OD2	5.09	122.88	118.30
1	C	690	ASP	CB-CG-OD2	5.09	122.88	118.30
1	B	464	ASP	CB-CG-OD2	5.07	122.87	118.30
1	A	464	ASP	CB-CG-OD2	5.07	122.86	118.30
1	A	516	ASP	CB-CG-OD2	5.07	122.86	118.30
1	B	623	ASP	CB-CG-OD2	5.07	122.86	118.30
1	D	690	ASP	CB-CG-OD2	5.06	122.86	118.30

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	3128	0	3164	22	0
1	B	3014	0	3050	22	0
1	C	2997	0	3026	16	0
1	D	2922	0	2953	19	0
2	A	5	0	0	0	0
2	B	5	0	0	0	0
2	C	5	0	0	0	0
2	D	5	0	0	0	0
3	A	38	0	33	0	0
3	B	38	0	33	1	0
3	C	38	0	33	1	0
3	D	38	0	33	0	0
4	A	335	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	316	0	0	0	0
4	C	293	0	0	0	0
4	D	292	0	0	0	0
All	All	13469	0	12325	49	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 (49) 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:771[B]:ASN:OD1	1:B:771[B]:ASN:ND2	1.69	1.25
1:C:771[B]:ASN:HD21	1:D:771[B]:ASN:CG	1.44	1.21
1:A:771[B]:ASN:ND2	1:B:771[B]:ASN:OD1	1.75	1.17
1:C:771[B]:ASN:ND2	1:D:771[B]:ASN:OD1	1.76	1.16
1:A:771[B]:ASN:CG	1:B:771[B]:ASN:HD21	1.68	0.95
1:C:771[B]:ASN:HD21	1:D:771[B]:ASN:ND2	1.64	0.95
1:A:771[B]:ASN:HD21	1:B:771[B]:ASN:CG	1.72	0.93
1:C:771[B]:ASN:CG	1:D:771[B]:ASN:HD21	1.72	0.91
1:C:771[B]:ASN:ND2	1:D:771[B]:ASN:CG	2.26	0.85
1:C:771[B]:ASN:ND2	1:D:771[B]:ASN:ND2	2.24	0.85
1:C:771[A]:ASN:OD1	1:D:771[A]:ASN:ND2	2.09	0.84
1:C:771[B]:ASN:ND2	1:D:771[B]:ASN:HD21	1.76	0.84
1:C:771[A]:ASN:OD1	1:D:771[A]:ASN:OD1	2.03	0.77
1:D:751:ALA:HB2	1:D:854:MET:CE	2.20	0.71
1:A:771[A]:ASN:HD21	1:B:771[A]:ASN:CG	1.95	0.69
1:C:771[A]:ASN:OD1	1:D:771[A]:ASN:CG	2.32	0.68
1:A:771[A]:ASN:CG	1:B:771[A]:ASN:HD21	1.97	0.67
1:A:771[A]:ASN:OD1	1:B:771[A]:ASN:OD1	2.13	0.67
1:A:771[A]:ASN:ND2	1:B:771[A]:ASN:ND2	2.46	0.64
1:D:751:ALA:HB2	1:D:854:MET:HE2	1.80	0.62
1:C:771[B]:ASN:OD1	1:D:771[B]:ASN:ND2	2.28	0.59
1:A:771[A]:ASN:HD21	1:B:771[A]:ASN:ND2	2.00	0.59
1:A:771[A]:ASN:ND2	1:B:771[A]:ASN:OD1	2.35	0.59
1:D:751:ALA:HB2	1:D:854:MET:HE3	1.84	0.58
1:A:771[A]:ASN:OD1	1:B:771[A]:ASN:ND2	2.36	0.57
1:A:771[A]:ASN:ND2	1:B:771[A]:ASN:HD21	2.01	0.57
1:C:708[B]:CYS:SG	1:C:795:CYS:HB2	2.47	0.54
1:C:771[A]:ASN:ND2	1:D:771[A]:ASN:OD1	2.43	0.52
1:D:708[B]:CYS:SG	1:D:795:CYS:HB2	2.49	0.52
1:A:771[B]:ASN:CG	1:B:771[B]:ASN:ND2	2.44	0.51

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:771[A]:ASN:CG	1:B:771[A]:ASN:OD1	2.49	0.51
1:A:771[A]:ASN:CG	1:B:771[A]:ASN:ND2	2.64	0.50
1:A:771[A]:ASN:ND2	1:B:771[A]:ASN:CG	2.63	0.49
1:A:771[B]:ASN:ND2	1:B:771[B]:ASN:CG	2.47	0.49
1:A:771[A]:ASN:OD1	1:B:771[A]:ASN:CG	2.50	0.48
1:A:635:HIS:HB3	1:A:646:ARG:HB3	1.95	0.47
1:A:771[B]:ASN:ND2	1:B:771[B]:ASN:ND2	2.62	0.47
1:D:736:ASN:HD21	1:D:854:MET:HE3	1.79	0.47
3:B:876:5HI:H13A	3:B:876:5HI:O2	2.15	0.47
1:C:771[A]:ASN:CG	1:D:771[A]:ASN:OD1	2.53	0.46
1:D:751:ALA:HB1	1:D:853:LEU:HD23	1.96	0.46
1:A:708[B]:CYS:SG	1:A:795:CYS:HB2	2.57	0.44
1:C:702:ARG:O	1:C:799:SER:HA	2.18	0.44
1:B:600[B]:CYS:SG	1:B:601:ASP:N	2.92	0.43
3:C:876:5HI:H12A	3:C:876:5HI:O2	2.17	0.43
1:A:581:SER:HB3	1:A:708[A]:CYS:SG	2.58	0.43
1:B:441:GLU:N	1:B:442:PRO:CD	2.83	0.42
1:C:600[B]:CYS:SG	1:C:601:ASP:N	2.95	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	421/441 (96%)	405 (96%)	16 (4%)	0	100	100
1	B	402/441 (91%)	390 (97%)	11 (3%)	1 (0%)	47	30
1	C	401/441 (91%)	387 (96%)	14 (4%)	0	100	100
1	D	394/441 (89%)	382 (97%)	12 (3%)	0	100	100
All	All	1618/1764 (92%)	1564 (97%)	53 (3%)	1 (0%)	51	33

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	525	ALA

### 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	337/355 (95%)	329 (98%)	8 (2%)	49	31
1	B	326/355 (92%)	318 (98%)	8 (2%)	47	29
1	C	324/355 (91%)	315 (97%)	9 (3%)	43	25
1	D	316/355 (89%)	304 (96%)	12 (4%)	33	14
All	All	1303/1420 (92%)	1266 (97%)	37 (3%)	46	25

All (37) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	464	ASP
1	A	498	LEU
1	A	523	MET
1	A	630	ARG
1	A	657	MET
1	A	708[A]	CYS
1	A	708[B]	CYS
1	A	784	SER
1	B	450	GLN
1	B	452	LEU
1	B	486	GLU
1	B	595	ARG
1	B	613	GLU
1	B	630	ARG
1	B	657	MET
1	B	752	HIS
1	C	482	GLU
1	C	486	GLU
1	C	523	MET
1	C	627	ARG
1	C	630	ARG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	657	MET
1	C	708[A]	CYS
1	C	708[B]	CYS
1	C	752	HIS
1	D	464	ASP
1	D	470	LEU
1	D	500[A]	SER
1	D	500[B]	SER
1	D	523	MET
1	D	613	GLU
1	D	630	ARG
1	D	657	MET
1	D	658	ASN
1	D	708[A]	CYS
1	D	708[B]	CYS
1	D	752	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	454	ASN
1	A	819	GLN
1	B	819	GLN
1	C	632	GLN
1	C	819	GLN

### 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

8 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	5HI	C	876	-	32,40,40	0.63	0	42,55,55	1.16	6 (14%)
2	SO4	B	2	-	4,4,4	0.14	0	6,6,6	0.04	0
2	SO4	D	4	-	4,4,4	0.15	0	6,6,6	0.06	0
3	5HI	A	876	-	32,40,40	0.62	0	42,55,55	1.15	6 (14%)
2	SO4	C	3	-	4,4,4	0.14	0	6,6,6	0.07	0
2	SO4	A	1	-	4,4,4	0.14	0	6,6,6	0.06	0
3	5HI	D	876	-	32,40,40	0.62	1 (3%)	42,55,55	1.26	6 (14%)
3	5HI	B	876	-	32,40,40	0.62	0	42,55,55	1.15	4 (9%)

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	5HI	C	876	-	-	6/30/35/35	0/3/3/3
3	5HI	D	876	-	-	5/30/35/35	0/3/3/3
3	5HI	A	876	-	-	5/30/35/35	0/3/3/3
3	5HI	B	876	-	-	8/30/35/35	0/3/3/3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	876	5HI	C32-C1	-2.03	1.49	1.52

All (22) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	876	5HI	C14-C1-N2	-4.19	103.66	109.17

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	876	5HI	C14-C1-N2	-3.07	105.13	109.17
3	B	876	5HI	C14-C1-N2	-3.05	105.15	109.17
3	A	876	5HI	C14-C1-N2	-2.55	105.81	109.17
3	A	876	5HI	C1-N2-C3	2.55	125.63	122.34
3	C	876	5HI	C19-C2-C3	-2.46	126.46	131.18
3	A	876	5HI	C32-C1-N2	2.44	116.55	112.00
3	B	876	5HI	N3-C5-N1	-2.39	109.12	115.11
3	C	876	5HI	N3-C5-N1	-2.38	109.13	115.11
3	D	876	5HI	N3-C5-N1	-2.37	109.15	115.11
3	A	876	5HI	N3-C5-N1	-2.34	109.23	115.11
3	D	876	5HI	C32-C1-N2	2.32	116.34	112.00
3	D	876	5HI	C19-C2-C3	-2.25	126.87	131.18
3	C	876	5HI	C32-C1-N2	2.24	116.19	112.00
3	C	876	5HI	C24-C30-C15	-2.18	119.93	122.83
3	B	876	5HI	C24-C30-C15	-2.17	119.95	122.83
3	A	876	5HI	C19-C2-C3	-2.14	127.07	131.18
3	D	876	5HI	C24-C30-C15	-2.12	120.00	122.83
3	B	876	5HI	C32-C1-N2	2.12	115.96	112.00
3	C	876	5HI	C10-C11-C35	-2.10	109.00	113.19
3	D	876	5HI	C1-N2-C3	2.07	125.01	122.34
3	A	876	5HI	C24-C30-C15	-2.04	120.12	122.83

There are no chirality outliers.

All (24) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	876	5HI	C21-C27-C5-N1
3	A	876	5HI	C18-C27-C5-N1
3	A	876	5HI	C18-C27-C5-N3
3	A	876	5HI	C19-C2-C3-O2
3	B	876	5HI	C21-C27-C5-N1
3	B	876	5HI	C18-C27-C5-N1
3	B	876	5HI	C18-C27-C5-N3
3	B	876	5HI	N1-C19-C6-C13
3	B	876	5HI	C2-C19-C6-C13
3	B	876	5HI	C19-C2-C3-O2
3	C	876	5HI	C21-C27-C5-N1
3	C	876	5HI	C18-C27-C5-N1
3	C	876	5HI	C18-C27-C5-N3
3	C	876	5HI	C19-C2-C3-O2
3	D	876	5HI	C21-C27-C5-N1
3	D	876	5HI	C18-C27-C5-N1

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
3	D	876	5HI	C18-C27-C5-N3
3	D	876	5HI	C19-C2-C3-O2
3	A	876	5HI	C21-C27-C5-N3
3	D	876	5HI	C21-C27-C5-N3
3	B	876	5HI	C21-C27-C5-N3
3	C	876	5HI	C21-C27-C5-N3
3	B	876	5HI	C7-C8-C9-O4
3	C	876	5HI	C7-C8-C9-O4

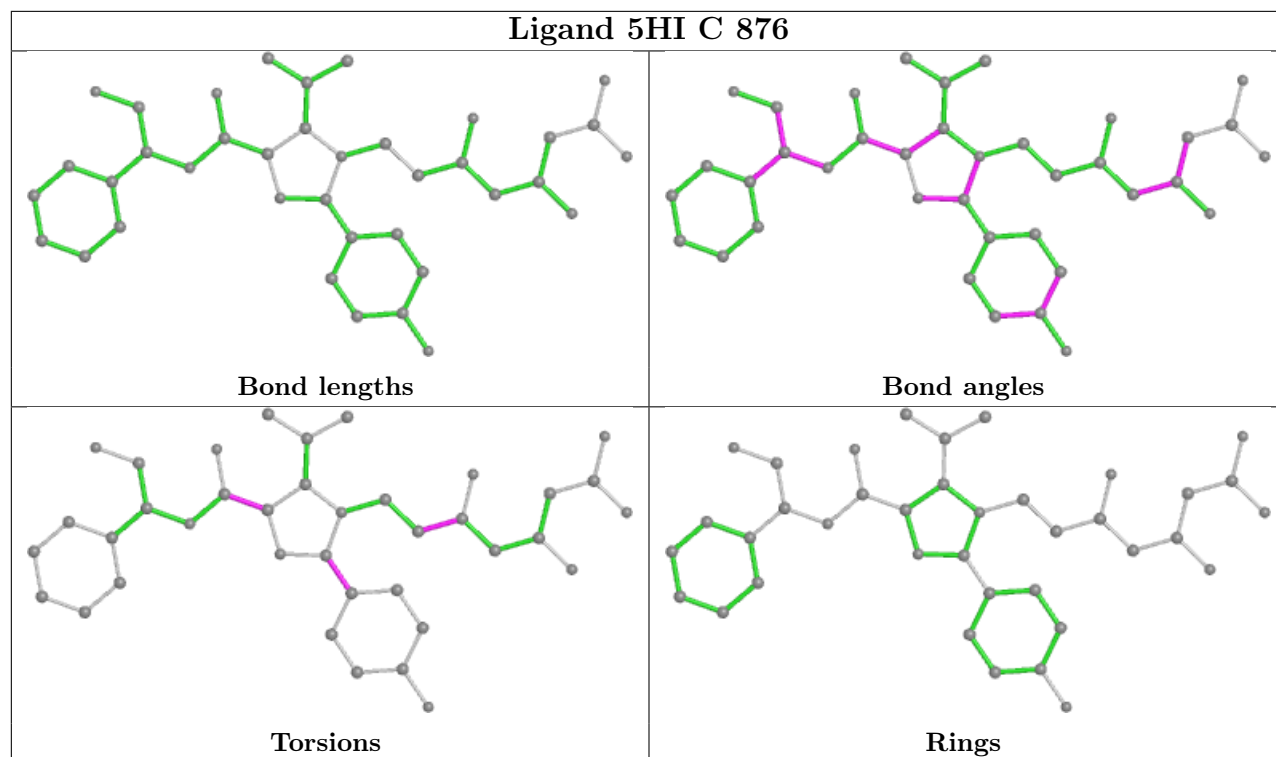
There are no ring outliers.

2 monomers are involved in 2 short contacts:

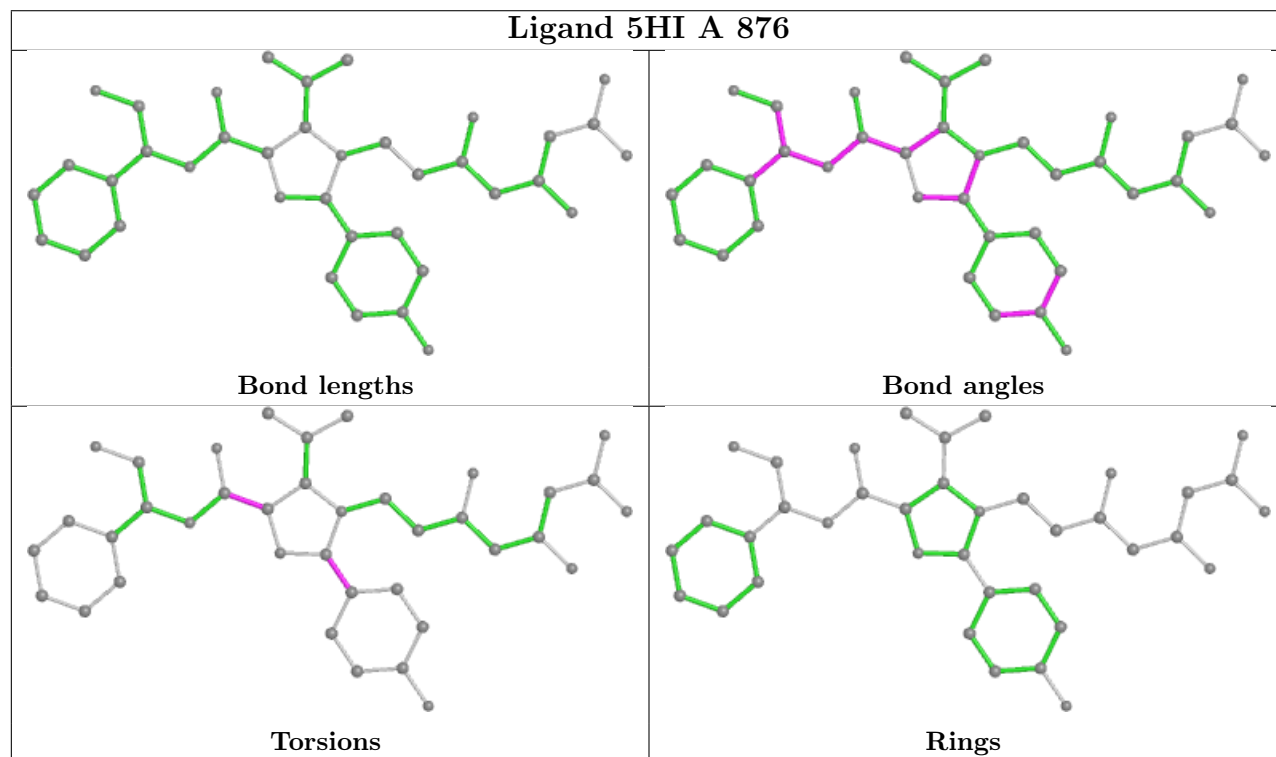
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	876	5HI	1	0
3	B	876	5HI	1	0

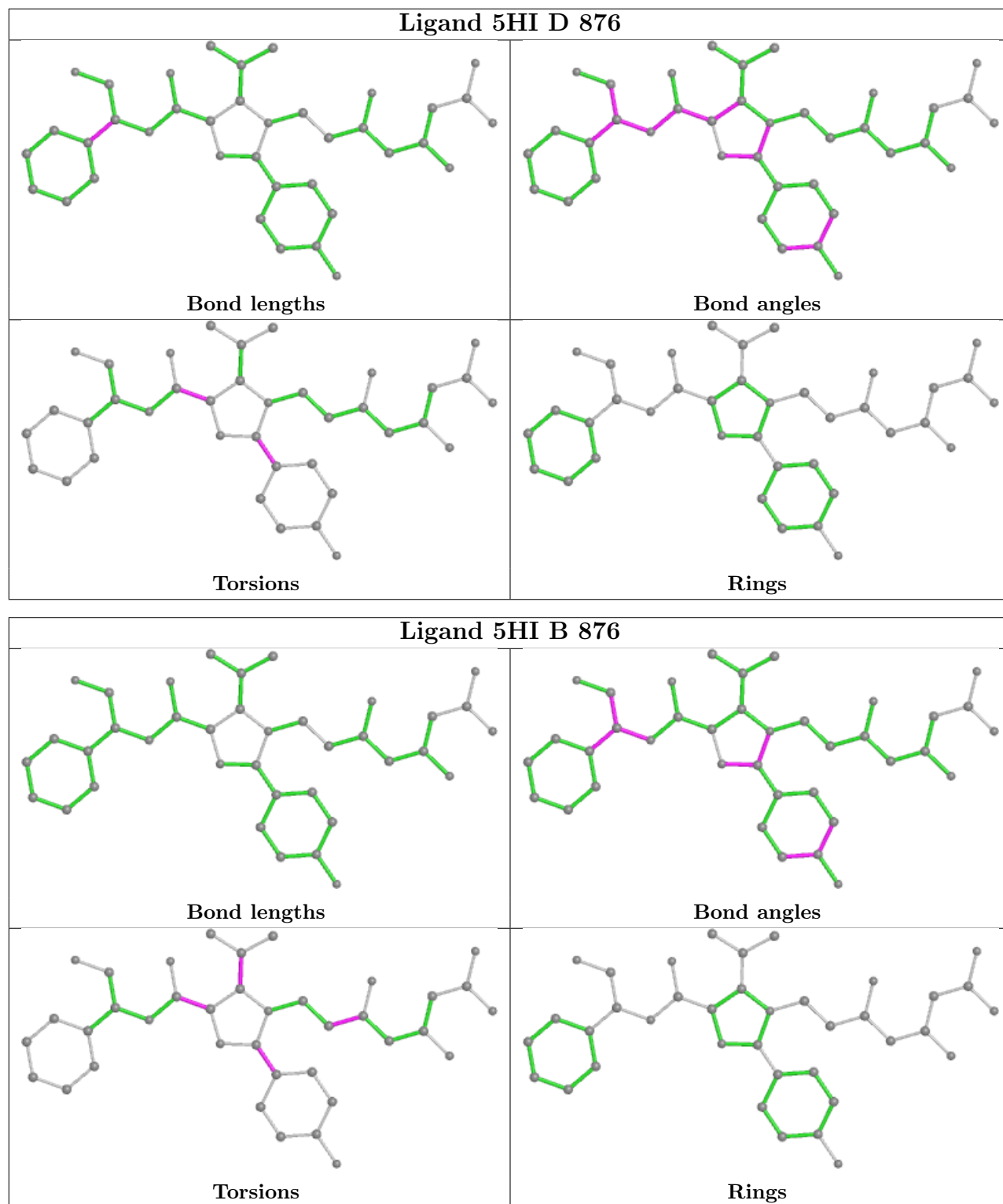
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less 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.

## Ligand 5HI C 876



## Ligand 5HI A 876





## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	420/441 (95%)	1.08	61 (14%) <b>2</b> <b>2</b>	15, 24, 54, 79	0
1	B	405/441 (91%)	1.12	62 (15%) <b>2</b> <b>2</b>	14, 24, 80, 99	0
1	C	404/441 (91%)	1.26	61 (15%) <b>2</b> <b>2</b>	14, 23, 80, 96	0
1	D	394/441 (89%)	1.06	48 (12%) <b>4</b> <b>4</b>	14, 24, 59, 93	0
All	All	1623/1764 (92%)	1.13	232 (14%) <b>2</b> <b>2</b>	14, 23, 69, 99	0

All (232) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	461	PHE	12.9
1	B	461	PHE	12.5
1	D	484	LEU	12.0
1	D	461	PHE	11.9
1	D	458	GLY	11.6
1	B	449	LEU	11.6
1	C	453	GLY	11.5
1	A	453	GLY	11.5
1	C	483	THR	11.1
1	C	449	LEU	10.9
1	D	469	GLN	10.7
1	D	470	LEU	10.4
1	C	484	LEU	10.1
1	A	455	ALA	9.9
1	D	462	LEU	9.8
1	D	483	THR	9.7
1	C	458	GLY	9.6
1	D	472	ASN	9.4
1	B	450	GLN	9.3
1	B	474	LYS	9.1
1	C	452	LEU	9.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	461	PHE	8.3
1	C	462	LEU	8.2
1	C	448	CYS	8.0
1	C	444	PRO	8.0
1	D	468	ILE	7.9
1	B	452	LEU	7.9
1	B	442	PRO	7.7
1	B	451	ILE	7.5
1	C	469	GLN	7.4
1	C	487	THR	7.1
1	C	454	ASN	7.0
1	C	450	GLN	7.0
1	B	444	PRO	7.0
1	A	441	GLU	7.0
1	A	442	PRO	7.0
1	B	448	CYS	6.8
1	C	446	GLU	6.7
1	C	486	GLU	6.7
1	B	462	LEU	6.7
1	C	451	ILE	6.7
1	D	486	GLU	6.6
1	D	471	VAL	6.6
1	D	467	ILE	6.6
1	C	482	GLU	6.6
1	C	471	VAL	6.6
1	D	459	ALA	6.6
1	B	441	GLU	6.5
1	C	470	LEU	6.4
1	A	452	LEU	6.1
1	A	454	ASN	5.9
1	C	445	ASN	5.7
1	D	485	ILE	5.6
1	A	458	GLY	5.6
1	C	447	GLU	5.5
1	A	456	GLU	5.4
1	A	450	GLN	5.3
1	D	487	THR	5.3
1	B	446	GLU	5.2
1	B	484	LEU	5.0
1	C	708[A]	CYS	4.7
1	B	488	HIS	4.6
1	A	635	HIS	4.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	471	VAL	4.5
1	C	488	HIS	4.4
1	B	465	ALA	4.4
1	C	460	LYS	4.4
1	D	786	PRO	4.3
1	A	457	LYS	4.3
1	B	828	LYS	4.3
1	C	786	PRO	4.3
1	B	445	ASN	4.2
1	A	828	LYS	4.1
1	A	786	PRO	4.1
1	A	462	LEU	4.1
1	A	446	GLU	4.1
1	B	447	GLU	4.1
1	A	444	PRO	4.0
1	A	785	GLY	4.0
1	B	473	ALA	4.0
1	B	468	ILE	4.0
1	C	829	ASP	4.0
1	D	460	LYS	3.9
1	D	465	ALA	3.9
1	C	785	GLY	3.9
1	A	787	THR	3.8
1	A	483	THR	3.8
1	B	485	ILE	3.8
1	A	449	LEU	3.8
1	B	486	GLU	3.8
1	A	484	LEU	3.7
1	A	708[A]	CYS	3.7
1	B	460	LYS	3.7
1	B	829	ASP	3.6
1	C	485	ILE	3.6
1	C	787	THR	3.6
1	A	860	GLY	3.5
1	B	470	LEU	3.5
1	C	459	ALA	3.5
1	C	467	ILE	3.5
1	B	516	ASP	3.4
1	C	860	GLY	3.4
1	C	828	LYS	3.4
1	C	630	ARG	3.4
1	D	466	GLU	3.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	463	SER	3.4
1	D	829	ASP	3.4
1	B	459	ALA	3.4
1	B	772	VAL	3.4
1	A	630	ARG	3.3
1	B	506	PRO	3.3
1	B	487	THR	3.2
1	D	708[A]	CYS	3.2
1	D	627	ARG	3.2
1	D	630	ARG	3.2
1	A	718	ARG	3.1
1	A	475	HIS	3.1
1	B	633	LYS	3.1
1	D	633	LYS	3.0
1	C	490	ARG	3.0
1	A	523	MET	3.0
1	A	460	LYS	3.0
1	B	743	ALA	3.0
1	A	789	GLU	3.0
1	B	466	GLU	2.9
1	B	515	ARG	2.9
1	B	464	ASP	2.9
1	D	828	LYS	2.9
1	A	772	VAL	2.9
1	C	463	SER	2.9
1	A	829	ASP	2.8
1	D	463	SER	2.8
1	D	482	GLU	2.8
1	C	465	ALA	2.8
1	D	659	MET	2.8
1	C	468	ILE	2.8
1	A	633	LYS	2.8
1	D	490	ARG	2.8
1	B	501	LYS	2.8
1	B	695	ALA	2.7
1	D	613	GLU	2.7
1	B	831	PRO	2.7
1	A	488	HIS	2.7
1	A	516	ASP	2.7
1	D	620	GLU	2.7
1	B	827	CYS	2.7
1	A	485	ILE	2.7

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	469	GLN	2.7
1	B	490	ARG	2.7
1	A	451	ILE	2.6
1	A	448	CYS	2.6
1	C	612	SER	2.6
1	C	772	VAL	2.6
1	A	515	ARG	2.6
1	C	768	ALA	2.6
1	C	771[A]	ASN	2.6
1	A	715	LYS	2.5
1	A	487	THR	2.5
1	D	488	HIS	2.5
1	B	628	PHE	2.5
1	B	830	ASN	2.5
1	A	650	ARG	2.5
1	D	787	THR	2.5
1	B	778	ILE	2.5
1	D	523	MET	2.5
1	D	771[A]	ASN	2.5
1	C	466	GLU	2.5
1	C	757	VAL	2.5
1	A	506	PRO	2.4
1	D	860	GLY	2.4
1	D	516	ASP	2.4
1	A	741	ALA	2.4
1	D	830	ASN	2.4
1	C	778	ILE	2.4
1	C	613	GLU	2.4
1	D	464	ASP	2.4
1	B	443	ARG	2.4
1	D	718	ARG	2.4
1	B	472	ASN	2.4
1	A	695	ALA	2.4
1	B	787	THR	2.4
1	B	638	ILE	2.3
1	C	600[A]	CYS	2.3
1	C	516	ASP	2.3
1	C	749	TYR	2.3
1	D	658	ASN	2.3
1	D	757	VAL	2.3
1	D	524	GLY	2.3
1	D	785	GLY	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	749	TYR	2.3
1	A	463	SER	2.3
1	A	464	ASP	2.3
1	A	509	LEU	2.3
1	D	628	PHE	2.3
1	B	505	GLU	2.2
1	A	468	ILE	2.2
1	A	658	ASN	2.2
1	A	447	GLU	2.2
1	D	494	ILE	2.2
1	C	674	TYR	2.2
1	C	662	LYS	2.2
1	B	630	ARG	2.2
1	C	598	ARG	2.2
1	D	571	ARG	2.2
1	A	445	ASN	2.2
1	A	467	ILE	2.2
1	B	696	ILE	2.2
1	C	699	ILE	2.2
1	C	611	THR	2.2
1	B	650	ARG	2.1
1	B	757	VAL	2.1
1	C	758	THR	2.1
1	C	756	ILE	2.1
1	A	768	ALA	2.1
1	D	548	GLU	2.1
1	B	672	HIS	2.1
1	A	757	VAL	2.1
1	C	777	CYS	2.1
1	B	666	LYS	2.1
1	B	758	THR	2.1
1	B	860	GLY	2.1
1	A	486	GLU	2.1
1	A	754	ALA	2.1
1	B	548	GLU	2.1
1	A	769	ALA	2.0
1	B	769	ALA	2.0
1	A	627	ARG	2.0
1	B	596	LEU	2.0
1	C	635	HIS	2.0
1	A	846	VAL	2.0
1	C	524	GLY	2.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	830	ASN	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](#)

There are no monosaccharides in this entry.

## 6.4 Ligands [i](#)

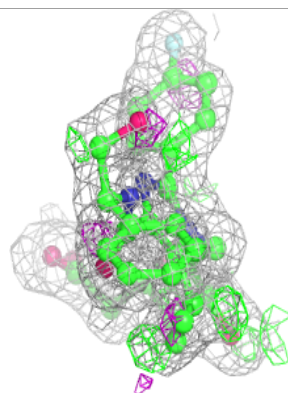
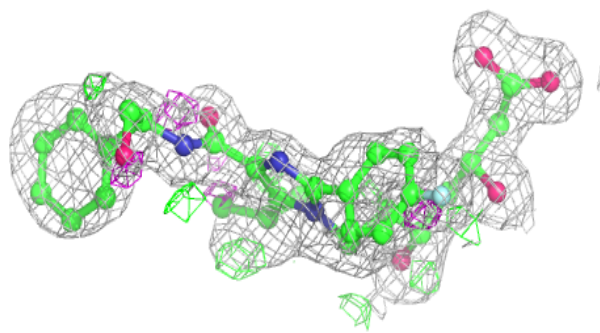
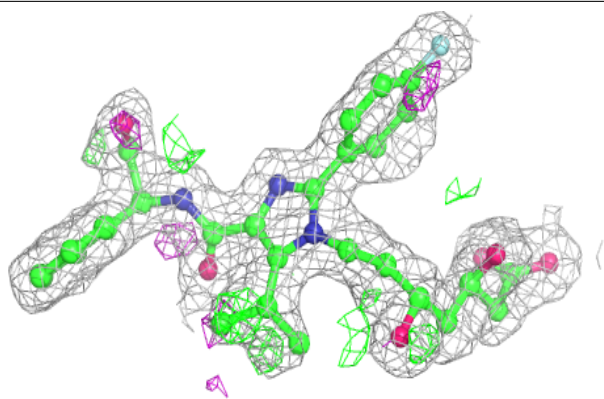
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	SO4	A	1	5/5	0.66	0.26	77,77,77,77	0
2	SO4	B	2	5/5	0.70	0.24	81,81,81,81	0
2	SO4	D	4	5/5	0.75	0.24	69,69,69,69	0
2	SO4	C	3	5/5	0.76	0.28	70,70,70,70	0
3	5HI	B	876	38/38	0.86	0.15	21,29,30,30	0
3	5HI	C	876	38/38	0.86	0.14	20,27,29,30	0
3	5HI	D	876	38/38	0.86	0.13	21,27,28,29	0
3	5HI	A	876	38/38	0.87	0.15	19,25,27,29	0

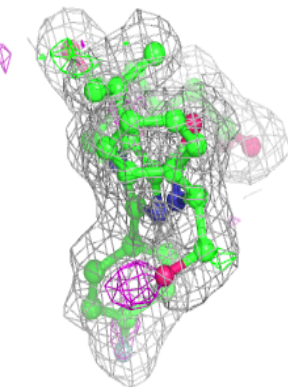
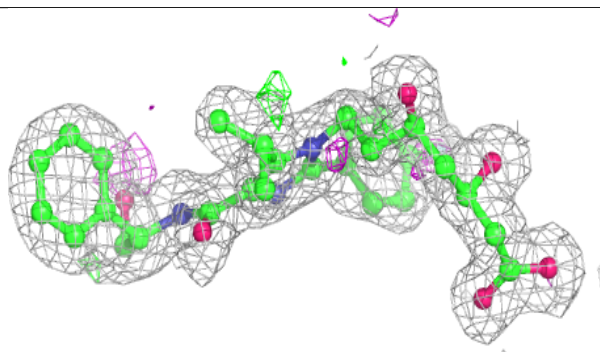
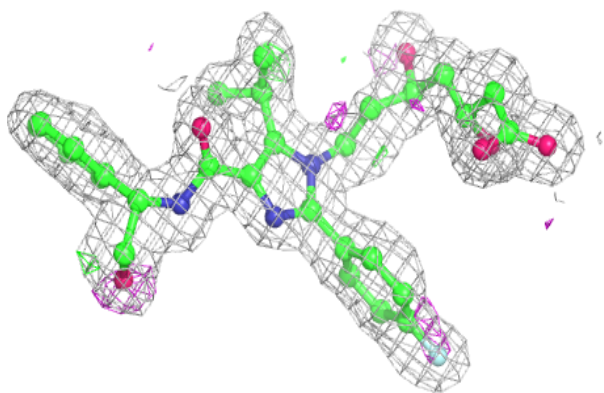
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

**Electron density around 5HI B 876:**

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

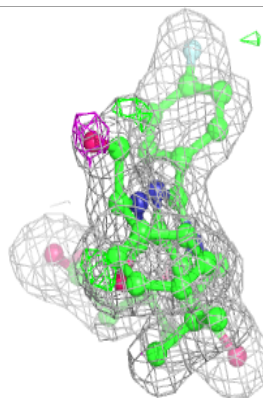
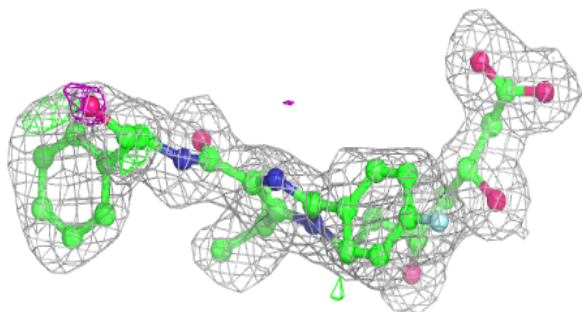
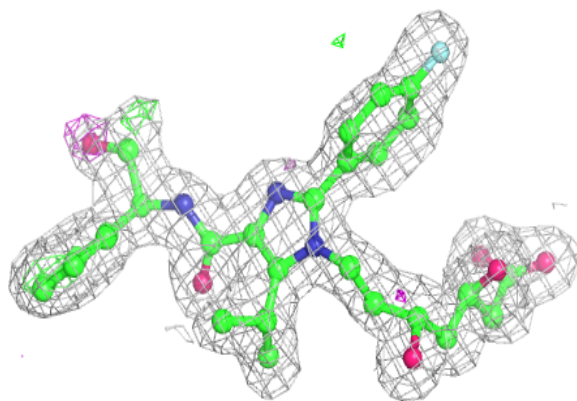
**Electron density around 5HI C 876:**

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

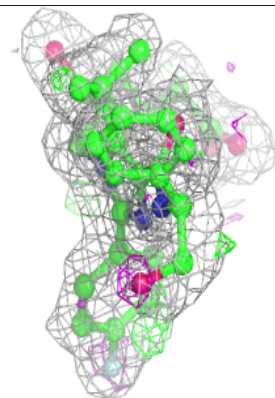
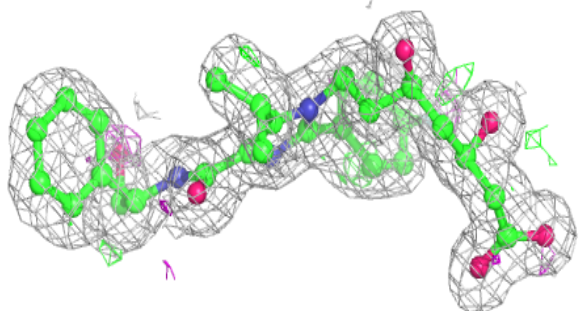
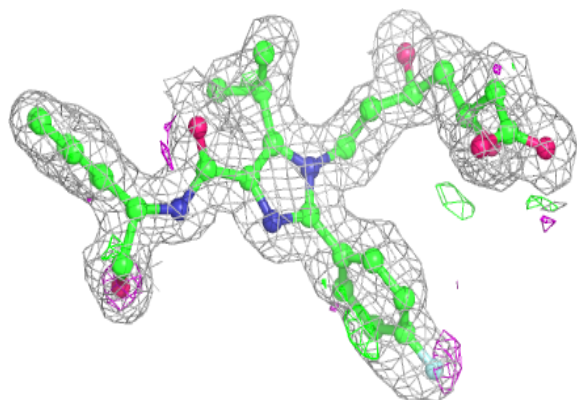


**Electron density around 5HI D 876:**

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

**Electron density around 5HI A 876:**

$2mF_o-DF_c$  (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.