



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 12, 2020 – 09:59 PM BST

PDB ID : 2IG6  
Title : Crystal structure of a nimc/nima family protein (ca\_c2569) from clostridium acetobutylicum at 1.80 Å resolution  
Authors : Joint Center for Structural Genomics (JCSG)  
Deposited on : 2006-09-22  
Resolution : 1.80 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 : 2.14.4.dev1  
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) : 2.14.4.dev1

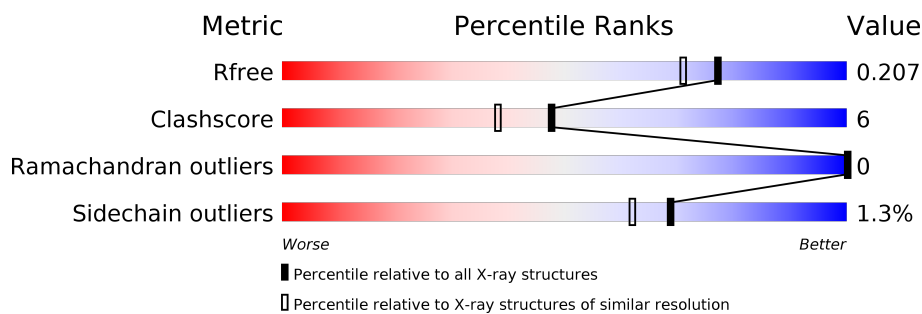
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.80 Å.

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



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

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  $\geq 3$ , 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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	150	
1	B	150	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	UNL	A	340	-	-	X	-

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 2631 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 NimC/NimA family protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	143	Total	C	N	O	S	Se	0	3	0
			1149	736	195	210	4	4			
1	B	139	Total	C	N	O	S	Se	0	6	0
			1139	732	188	211	4	4			

There are 46 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-18	MSE	-	LEADER SEQUENCE	UNP Q97G05
A	-17	GLY	-	LEADER SEQUENCE	UNP Q97G05
A	-16	SER	-	LEADER SEQUENCE	UNP Q97G05
A	-15	ASP	-	LEADER SEQUENCE	UNP Q97G05
A	-14	LYS	-	LEADER SEQUENCE	UNP Q97G05
A	-13	ILE	-	LEADER SEQUENCE	UNP Q97G05
A	-12	HIS	-	LEADER SEQUENCE	UNP Q97G05
A	-11	HIS	-	LEADER SEQUENCE	UNP Q97G05
A	-10	HIS	-	LEADER SEQUENCE	UNP Q97G05
A	-9	HIS	-	LEADER SEQUENCE	UNP Q97G05
A	-8	HIS	-	LEADER SEQUENCE	UNP Q97G05
A	-7	HIS	-	LEADER SEQUENCE	UNP Q97G05
A	-6	GLU	-	LEADER SEQUENCE	UNP Q97G05
A	-5	ASN	-	LEADER SEQUENCE	UNP Q97G05
A	-4	LEU	-	LEADER SEQUENCE	UNP Q97G05
A	-3	TYR	-	LEADER SEQUENCE	UNP Q97G05
A	-2	PHE	-	LEADER SEQUENCE	UNP Q97G05
A	-1	GLN	-	LEADER SEQUENCE	UNP Q97G05
A	0	GLY	-	LEADER SEQUENCE	UNP Q97G05
A	1	MSE	MET	MODIFIED RESIDUE	UNP Q97G05
A	53	MSE	MET	MODIFIED RESIDUE	UNP Q97G05
A	64	MSE	MET	MODIFIED RESIDUE	UNP Q97G05
A	99	MSE	MET	MODIFIED RESIDUE	UNP Q97G05
B	-18	MSE	-	LEADER SEQUENCE	UNP Q97G05
B	-17	GLY	-	LEADER SEQUENCE	UNP Q97G05

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-16	SER	-	LEADER SEQUENCE	UNP Q97G05
B	-15	ASP	-	LEADER SEQUENCE	UNP Q97G05
B	-14	LYS	-	LEADER SEQUENCE	UNP Q97G05
B	-13	ILE	-	LEADER SEQUENCE	UNP Q97G05
B	-12	HIS	-	LEADER SEQUENCE	UNP Q97G05
B	-11	HIS	-	LEADER SEQUENCE	UNP Q97G05
B	-10	HIS	-	LEADER SEQUENCE	UNP Q97G05
B	-9	HIS	-	LEADER SEQUENCE	UNP Q97G05
B	-8	HIS	-	LEADER SEQUENCE	UNP Q97G05
B	-7	HIS	-	LEADER SEQUENCE	UNP Q97G05
B	-6	GLU	-	LEADER SEQUENCE	UNP Q97G05
B	-5	ASN	-	LEADER SEQUENCE	UNP Q97G05
B	-4	LEU	-	LEADER SEQUENCE	UNP Q97G05
B	-3	TYR	-	LEADER SEQUENCE	UNP Q97G05
B	-2	PHE	-	LEADER SEQUENCE	UNP Q97G05
B	-1	GLN	-	LEADER SEQUENCE	UNP Q97G05
B	0	GLY	-	LEADER SEQUENCE	UNP Q97G05
B	1	MSE	MET	MODIFIED RESIDUE	UNP Q97G05
B	53	MSE	MET	MODIFIED RESIDUE	UNP Q97G05
B	64	MSE	MET	MODIFIED RESIDUE	UNP Q97G05
B	99	MSE	MET	MODIFIED RESIDUE	UNP Q97G05

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



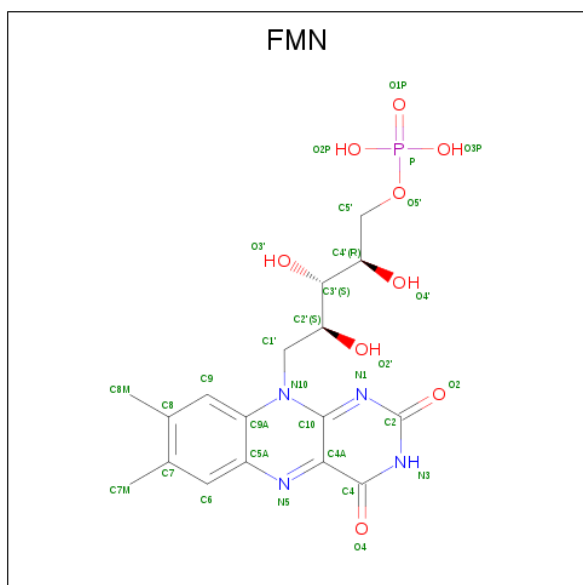
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

Continued from previous page...

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

- Molecule 3 is FLAVIN MONONUCLEOTIDE (three-letter code: FMN) (formula: C<sub>17</sub>H<sub>21</sub>N<sub>4</sub>O<sub>9</sub>P).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			31	17	4	9	1		

- Molecule 4 is UNKNOWN LIGAND (three-letter code: UNL) (formula: ).

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

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



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

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	132	Total	O	0	2
			134	134		
6	B	143	Total	O	0	7
			150	150		

**i**

- Molecule 1: NimC/NimA family protein



- Molecule 1: NimC/NimA family protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	52.43 Å   65.31 Å   108.79 Å 90.00°   90.00°   90.00°	Depositor
Resolution (Å)	26.86 – 1.80 26.86 – 1.80	Depositor EDS
% Data completeness (in resolution range)	95.9 (26.86-1.80) 95.9 (26.86-1.80)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	0.05	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.98 (at 1.80 Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.168   ,   0.198 0.181   ,   0.207	Depositor DCC
$R_{free}$ test set	1706 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	26.0	Xtriage
Anisotropy	0.394	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 46.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	2631	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.99% 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: FMN, UNL, SO4, EDO

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	0.56	0/1180	0.66	0/1581
1	B	0.66	1/1175 (0.1%)	0.72	0/1571
All	All	0.61	1/2355 (0.0%)	0.69	0/3152

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	1	MSE	SE-CE	-5.27	1.64	1.95

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	1149	0	1115	11	0
1	B	1139	0	1130	17	0
2	A	5	0	0	0	0
2	B	10	0	0	0	0
3	A	31	0	19	1	0
4	A	9	0	0	4	0
5	B	4	0	6	1	0
6	A	134	0	0	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	B	150	0	0	0	0
All	All	2631	0	2270	28	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 6.

All (28) 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)
4:A:340:UNL:O9	4:A:340:UNL:O4	1.64	1.15
4:A:340:UNL:O9	4:A:340:UNL:O8	1.74	1.03
1:A:19:ASN:ND2	1:B:19:ASN:OD1	1.94	0.99
1:B:5:LEU:HD12	1:B:32[B]:VAL:CG1	2.19	0.72
1:B:44:ASN:HB2	1:B:106[A]:ILE:HG22	1.77	0.65
1:B:32[A]:VAL:CG2	1:B:39:LEU:HD11	2.35	0.57
1:B:5:LEU:HD12	1:B:32[B]:VAL:HG12	1.89	0.54
1:A:35:TYR:CD2	1:A:88:LEU:HD22	2.42	0.54
1:A:87:GLU:HG2	1:A:102[B]:VAL:HG11	1.92	0.51
1:B:94:PRO:CB	1:B:97:LYS:HZ2	2.25	0.49
1:A:99:MSE:HE1	4:A:340:UNL:O2	2.13	0.48
1:B:94:PRO:CB	1:B:97:LYS:NZ	2.77	0.48
1:B:94:PRO:HB3	1:B:97:LYS:HZ2	1.80	0.47
1:A:102[A]:VAL:HG23	6:A:360:HOH:O	2.17	0.45
1:B:32[A]:VAL:HG21	1:B:39:LEU:HD11	1.98	0.45
1:A:99:MSE:HE1	4:A:340:UNL:O3	2.17	0.45
1:B:1:MSE:HG2	1:B:39:LEU:HB2	1.99	0.44
1:B:35:TYR:HD2	1:B:88:LEU:HD23	1.81	0.44
1:A:35:TYR:CD2	1:A:88:LEU:CD2	3.01	0.44
1:B:87:GLU:HG3	1:B:102:VAL:HG11	1.99	0.44
1:B:32[A]:VAL:HG12	5:B:134:EDO:H22	2.01	0.43
1:A:-9:HIS:O	1:A:-6:GLU:HB2	2.19	0.42
1:A:13[B]:VAL:HG22	3:A:339:FMN:HM72	2.00	0.42
1:A:19:ASN:HD21	1:B:19:ASN:CG	2.03	0.41
1:B:5:LEU:CD1	1:B:32[B]:VAL:CG1	2.96	0.41
1:B:64:MSE:HG2	1:B:70:TRP:HB3	2.03	0.41
1:A:99:MSE:HE2	1:A:100:TYR:CE2	2.56	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	144/150 (96%)	140 (97%)	4 (3%)	0	100	100
1	B	141/150 (94%)	136 (96%)	5 (4%)	0	100	100
All	All	285/300 (95%)	276 (97%)	9 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	121/125 (97%)	120 (99%)	1 (1%)	81	78
1	B	123/125 (98%)	121 (98%)	2 (2%)	62	54
All	All	244/250 (98%)	241 (99%)	3 (1%)	69	65

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	27	ARG
1	B	27	ARG
1	B	44	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	-11	HIS
1	A	19	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](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 1 is unknown - leaving 5 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	$\# Z  > 2$
3	FMN	A	339	-	31,33,33	1.36	5 (16%)	40,50,50	2.12	6 (15%)
2	SO4	B	133	-	4,4,4	0.19	0	6,6,6	0.45	0
2	SO4	B	132	-	4,4,4	0.24	0	6,6,6	0.28	0
5	EDO	B	134	-	3,3,3	0.39	0	2,2,2	0.36	0
2	SO4	A	132	-	4,4,4	0.18	0	6,6,6	0.14	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	FMN	A	339	-	-	1/18/18/18	0/3/3/3
5	EDO	B	134	-	-	0/1/1/1	-

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	339	FMN	C4A-N5	4.14	1.39	1.33
3	A	339	FMN	C10-N1	3.43	1.37	1.33
3	A	339	FMN	C4-N3	3.03	1.38	1.33
3	A	339	FMN	C1'-N10	2.24	1.50	1.48
3	A	339	FMN	C5A-N5	2.03	1.38	1.35

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	339	FMN	C4-N3-C2	7.48	121.46	115.14
3	A	339	FMN	C1'-N10-C9A	7.13	123.90	118.29
3	A	339	FMN	C5A-C9A-N10	4.62	121.06	117.72
3	A	339	FMN	C4A-C4-N3	-3.73	118.33	123.43
3	A	339	FMN	C4A-N5-C5A	2.51	119.28	116.77
3	A	339	FMN	C9A-N10-C10	-2.35	118.83	121.91

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	339	FMN	C4'-C5'-O5'-P

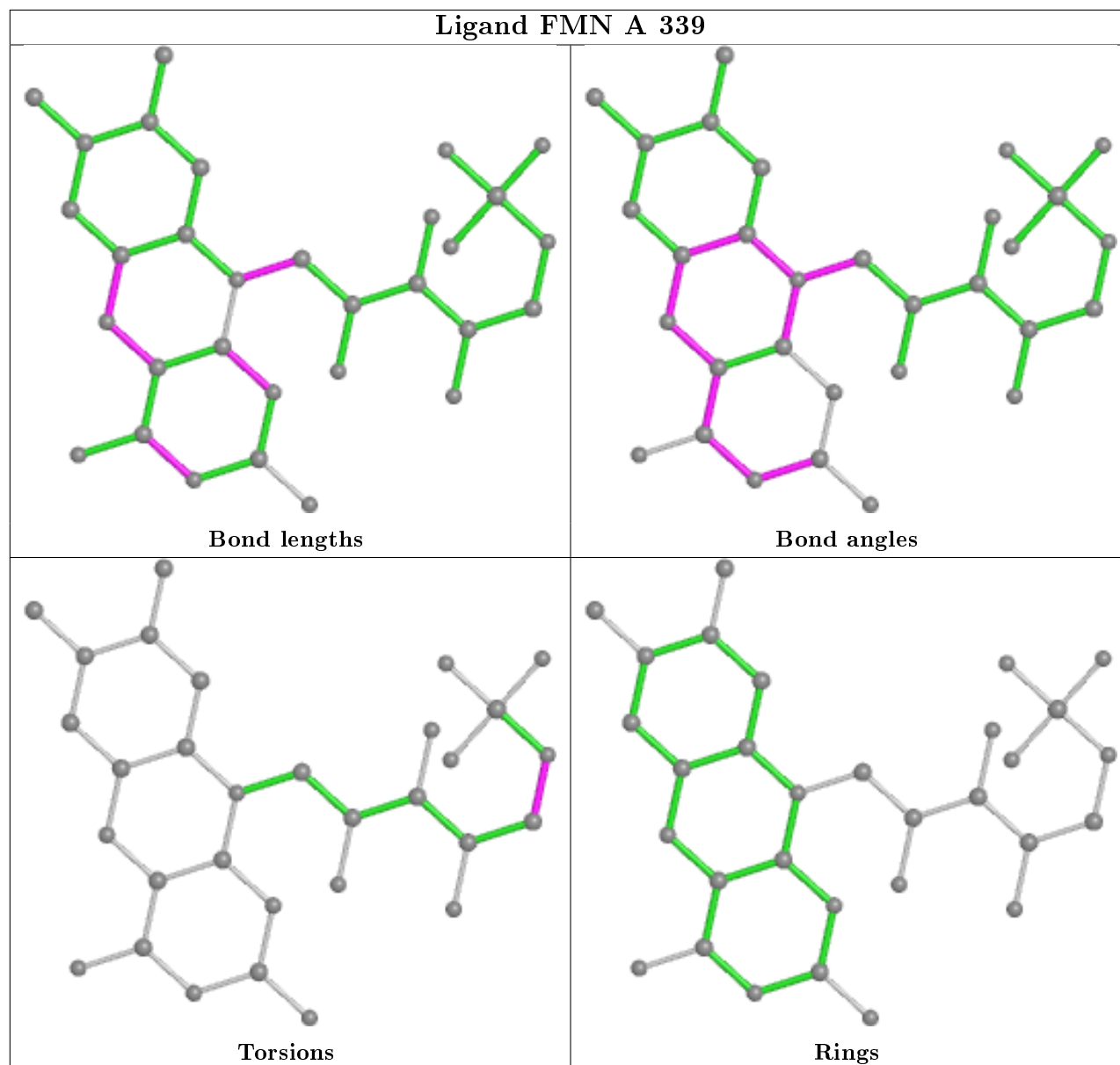
There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	339	FMN	1	0
5	B	134	EDO	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.



## 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

### 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.

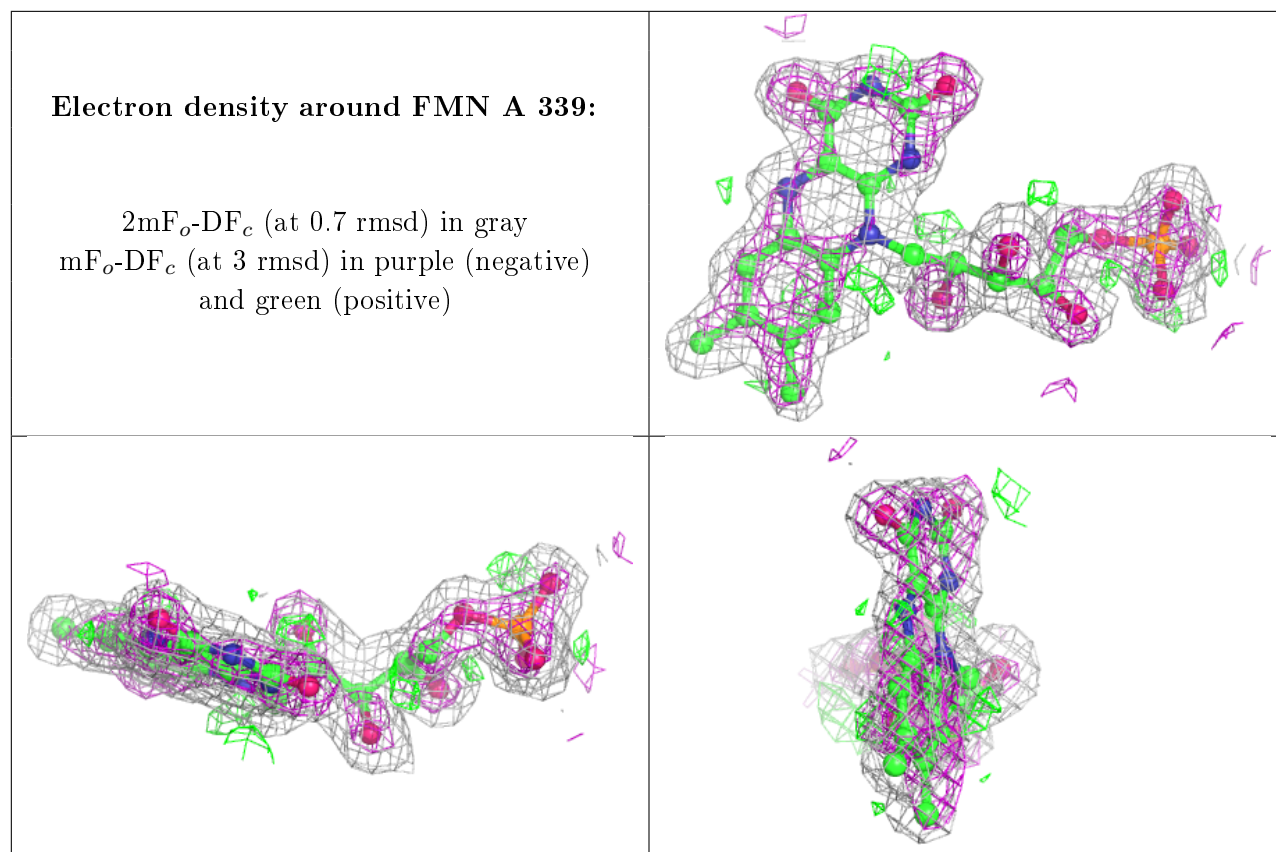
### 6.3 Carbohydrates

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.



## 6.5 Other polymers [i](#)

Unable to reproduce the depositor's R factor - this section is therefore empty.