wwPDB X-ray Structure Validation Summary Report (i)

Apr 28, 2024 – 06:27 pm BST

PDB ID	:	2VOU
Title	:	Structure of 2,6-dihydroxypyridine-3-hydroxylase from Arthrobacter nicotinovorans
Authors	:	Treiber, N.; Schulz, G.E.
Deposited on	:	2008-02-21
Resolution	:	2.60 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

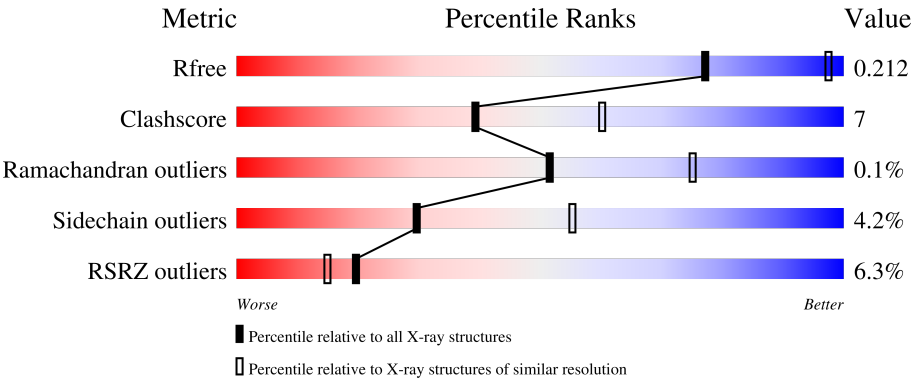
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 2.60 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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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% 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	397	4% 83% 15%
1	B	397	7% 80% 17%
1	C	397	8% 80% 16%

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	ACT	A	1397	-	-	X	-
4	ACT	B	1391	-	-	X	-
4	ACT	C	1391	-	-	X	-

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 9590 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.

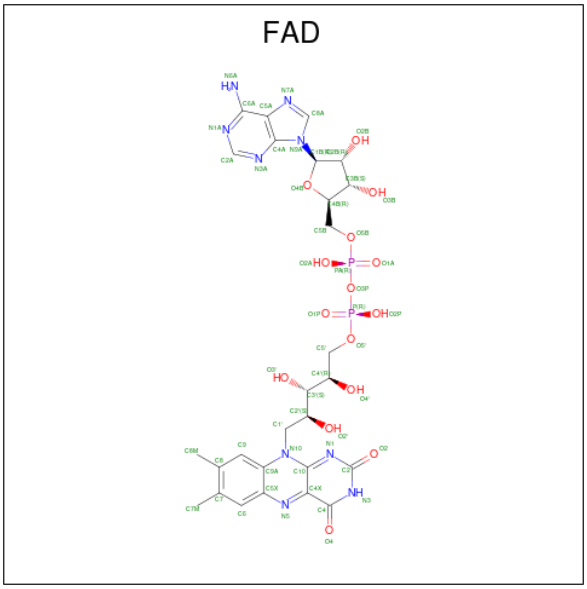
- Molecule 1 is a protein called 2,6-DIHYDROXYPYRIDINE HYDROXYLASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	393	Total	C	N	O	S	0	0	0
			3034	1916	530	581	7			
1	B	387	Total	C	N	O	S	0	0	0
			2990	1889	524	570	7			
1	C	387	Total	C	N	O	S	0	0	0
			2990	1889	524	570	7			

There are 3 discrepancies between the modelled and reference sequences:

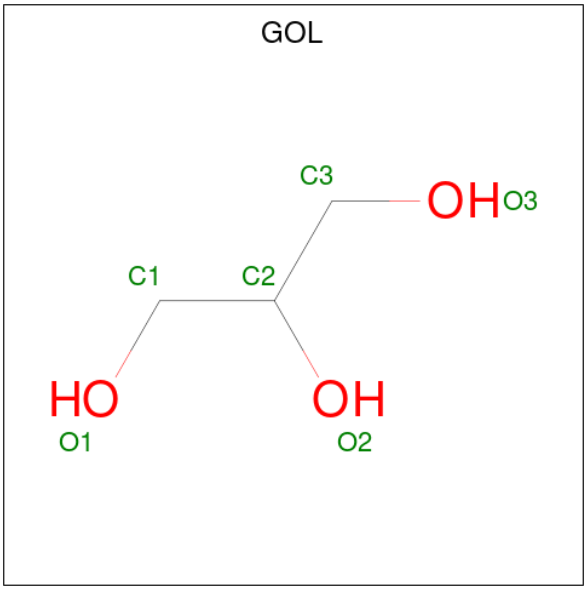
Chain	Residue	Modelled	Actual	Comment	Reference
A	323	SER	CYS	engineered mutation	UNP Q93NG3
B	323	SER	CYS	engineered mutation	UNP Q93NG3
C	323	SER	CYS	engineered mutation	UNP Q93NG3

- Molecule 2 is FLAVIN-ADENINE DINUCLEOTIDE (three-letter code: FAD) (formula: $C_{27}H_{33}N_9O_{15}P_2$).



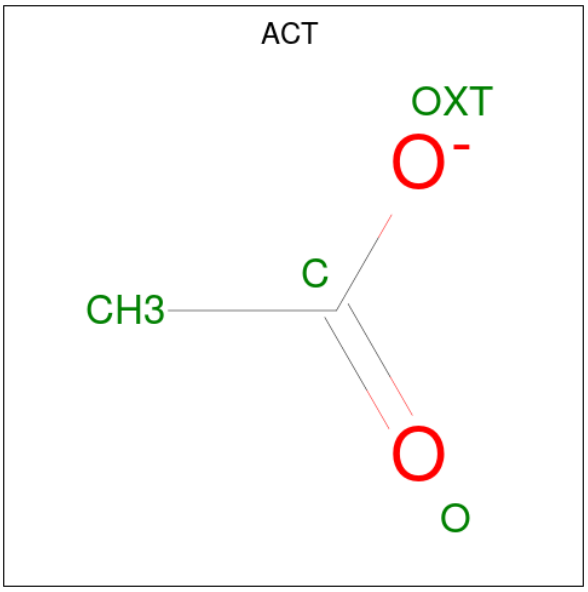
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			53	27	9	15	2		
2	B	1	Total	C	N	O	P	0	0
			53	27	9	15	2		
2	C	1	Total	C	N	O	P	0	0
			53	27	9	15	2		

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	O		0	0
			6	3	3			
3	B	1	Total	C	O		0	0
			6	3	3			
3	C	1	Total	C	O		0	0
			6	3	3			

- Molecule 4 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).



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

- Molecule 5 is water.

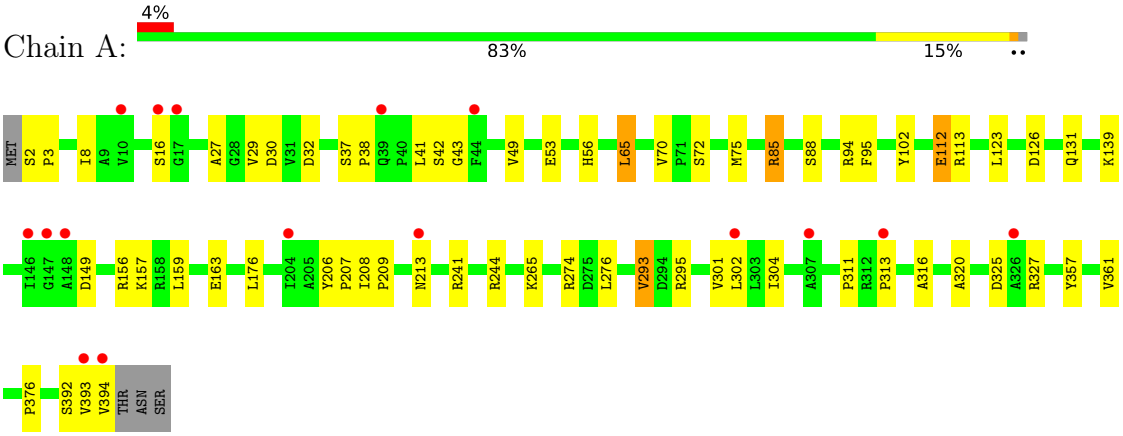
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	192	Total O 192 192	0	0
5	B	127	Total O 127 127	0	0
5	C	68	Total O 68 68	0	0

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 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.

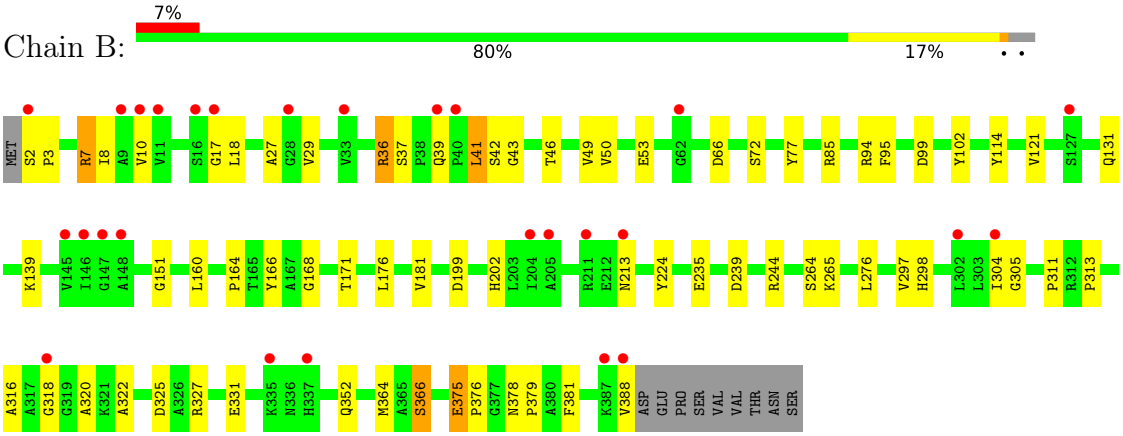
• Molecule 1: 2,6-DIHYDROXYPYRIDINE HYDROXYLASE

Chain A:



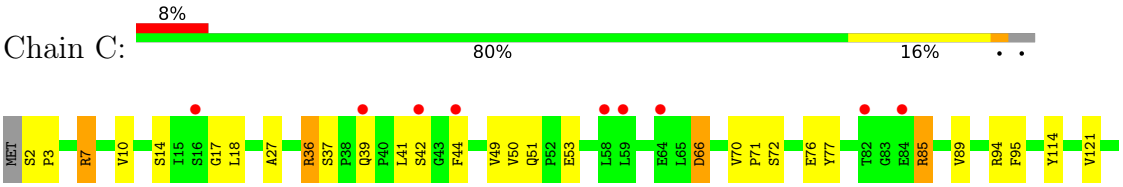
• Molecule 1: 2,6-DIHYDROXYPYRIDINE HYDROXYLASE

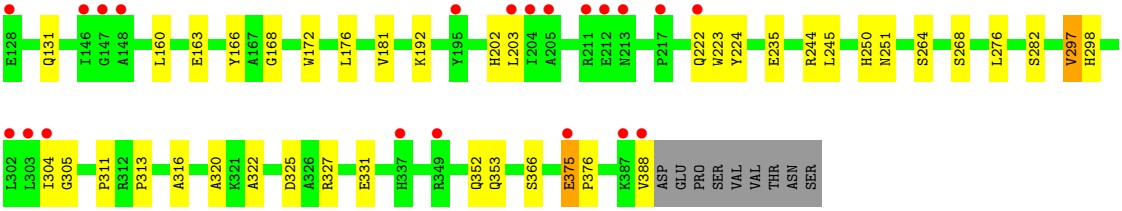
Chain B:



• Molecule 1: 2,6-DIHYDROXYPYRIDINE HYDROXYLASE

Chain C:




E128
Q131
I146
G147
A148
L160
E163
Y166
A167
G168
W172
L176
V181
K192
Y195
H202
L203
I204
A205
R211
E212
N213
P217
Q222
W223
Y224
E235
R244
L245
H250
N251
S264
S268
L276
S282
V297
H298
L302
L303
I304
G305
P311
R312
P313
A316
A320
K321
A322
D325
A326
R327
E331
H337
H349
Q352
Q353
S366
E375
P376
K387
V388
ASP
GLU
PRO
SER
VAL
VAL
THR
ASN
SER

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	185.96Å 185.96Å 104.76Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	27.92 – 2.60 27.92 – 2.60	Depositor EDS
% Data completeness (in resolution range)	100.0 (27.92-2.60) 99.8 (27.92-2.60)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.62 (at 2.61Å)	Xtriage
Refinement program	REFMAC 5.0	Depositor
R, R_{free}	0.185 , 0.222 0.178 , 0.212	Depositor DCC
R_{free} test set	3205 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	59.8	Xtriage
Anisotropy	0.444	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 57.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.022 for -h,-k,l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	9590	wwPDB-VP
Average B, all atoms (Å^2)	67.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.15% 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: GOL, ACT, FAD

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	0.78	0/3110	0.84	6/4235 (0.1%)
1	B	0.65	0/3065	0.70	3/4172 (0.1%)
1	C	0.57	0/3065	0.66	2/4172 (0.0%)
All	All	0.67	0/9240	0.74	11/12579 (0.1%)

There are no bond length outliers.

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	85	ARG	NE-CZ-NH1	-7.79	116.40	120.30
1	A	156	ARG	NE-CZ-NH1	-7.53	116.54	120.30
1	A	43	GLY	N-CA-C	-6.85	95.97	113.10
1	C	85	ARG	NE-CZ-NH1	6.62	123.61	120.30
1	C	85	ARG	NE-CZ-NH2	-6.60	117.00	120.30

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	3034	0	2938	36	0
1	B	2990	0	2898	46	0
1	C	2990	0	2898	46	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	53	0	31	3	0
2	B	53	0	31	2	0
2	C	53	0	31	1	0
3	A	6	0	8	0	0
3	B	6	0	8	0	0
3	C	6	0	8	0	0
4	A	4	0	3	7	0
4	B	4	0	3	7	0
4	C	4	0	3	3	0
5	A	192	0	0	8	0
5	B	127	0	0	7	0
5	C	68	0	0	6	0
All	All	9590	0	8860	133	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 7.

The worst 5 of 133 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:1397:ACT:H2	5:A:2159:HOH:O	1.55	1.03
1:A:42:SER:HB3	5:A:2013:HOH:O	1.69	0.91
1:C:7:ARG:HH11	1:C:7:ARG:HG3	1.36	0.90
1:B:7:ARG:HG3	1:B:7:ARG:HH11	1.34	0.89
4:A:1397:ACT:CH3	5:A:2159:HOH:O	2.20	0.85

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	391/397 (98%)	371 (95%)	19 (5%)	1 (0%)	41 64
1	B	385/397 (97%)	368 (96%)	17 (4%)	0	100 100
1	C	385/397 (97%)	365 (95%)	20 (5%)	0	100 100
All	All	1161/1191 (98%)	1104 (95%)	56 (5%)	1 (0%)	51 75

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	213	ASN

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	324/328 (99%)	314 (97%)	10 (3%)	40 66
1	B	318/328 (97%)	304 (96%)	14 (4%)	28 53
1	C	318/328 (97%)	302 (95%)	16 (5%)	24 47
All	All	960/984 (98%)	920 (96%)	40 (4%)	30 55

5 of 40 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	42	SER
1	C	264	SER
1	C	66	ASP
1	C	131	GLN
1	C	297	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	352	GLN
1	C	250	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	352	GLN
1	B	251	ASN
1	C	220	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)

9 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	$\#\|Z\| > 2$
2	FAD	C	1389	-	53,58,58	1.21	4 (7%)	68,89,89	1.45	11 (16%)
4	ACT	B	1391	-	3,3,3	0.87	0	3,3,3	1.26	0
2	FAD	A	1395	-	53,58,58	1.21	6 (11%)	68,89,89	1.55	10 (14%)
3	GOL	C	1390	-	5,5,5	0.33	0	5,5,5	0.33	0
3	GOL	B	1390	-	5,5,5	0.43	0	5,5,5	0.25	0
4	ACT	A	1397	-	3,3,3	1.12	0	3,3,3	0.84	0
3	GOL	A	1396	-	5,5,5	0.42	0	5,5,5	0.42	0
4	ACT	C	1391	-	3,3,3	0.87	0	3,3,3	1.32	0
2	FAD	B	1389	-	53,58,58	1.18	4 (7%)	68,89,89	1.35	9 (13%)

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	FAD	C	1389	-	-	3/30/50/50	0/6/6/6
2	FAD	A	1395	-	-	1/30/50/50	0/6/6/6
3	GOL	C	1390	-	-	0/4/4/4	-
3	GOL	B	1390	-	-	0/4/4/4	-
3	GOL	A	1396	-	-	2/4/4/4	-
2	FAD	B	1389	-	-	1/30/50/50	0/6/6/6

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	1389	FAD	C2A-N3A	4.28	1.39	1.32
2	B	1389	FAD	C2A-N3A	4.26	1.39	1.32
2	B	1389	FAD	C4X-N5	4.17	1.38	1.30
2	A	1395	FAD	C4X-N5	3.97	1.38	1.30
2	C	1389	FAD	C4X-N5	3.93	1.38	1.30

The worst 5 of 30 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	1395	FAD	N3A-C2A-N1A	-5.94	119.39	128.68
2	B	1389	FAD	N3A-C2A-N1A	-5.71	119.76	128.68
2	C	1389	FAD	N3A-C2A-N1A	-5.45	120.16	128.68
2	C	1389	FAD	P-O3P-PA	-4.11	118.72	132.83
2	A	1395	FAD	C4-C4X-N5	3.34	122.99	118.23

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1396	GOL	C1-C2-C3-O3
2	C	1389	FAD	O4B-C4B-C5B-O5B
2	C	1389	FAD	C3B-C4B-C5B-O5B
3	A	1396	GOL	O2-C2-C3-O3
2	C	1389	FAD	C5'-O5'-P-O3P

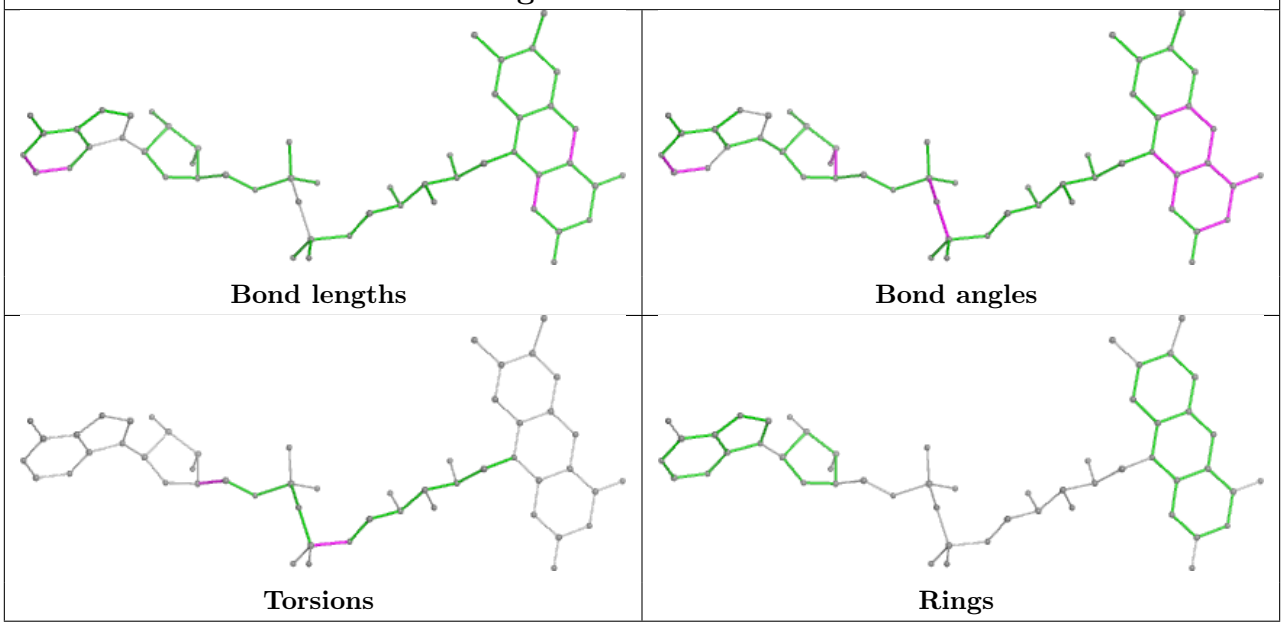
There are no ring outliers.

6 monomers are involved in 18 short contacts:

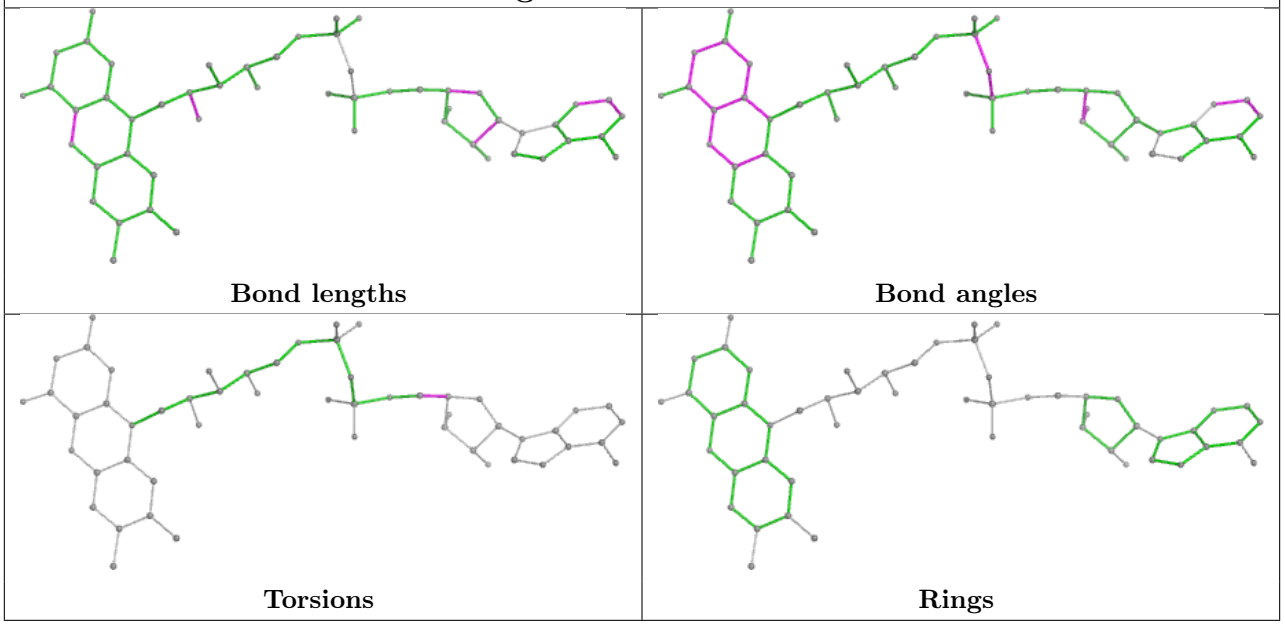
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	1389	FAD	1	0
4	B	1391	ACT	7	0
2	A	1395	FAD	3	0
4	A	1397	ACT	7	0
4	C	1391	ACT	3	0
2	B	1389	FAD	2	0

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.

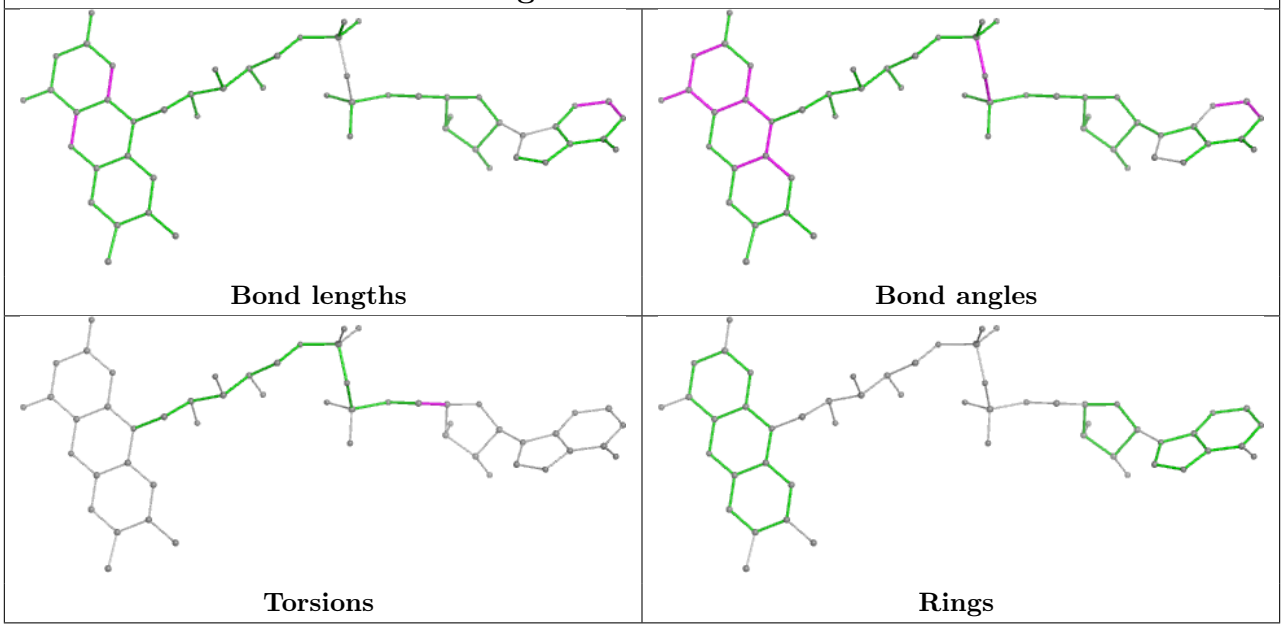
Ligand FAD C 1389



Ligand FAD A 1395



Ligand FAD B 1389



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

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^{th} 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^2)	Q<0.9
1	A	393/397 (98%)	0.12	16 (4%)	37	30	58, 66, 80, 99	0
1	B	387/397 (97%)	0.24	27 (6%)	16	12	64, 66, 70, 76	0
1	C	387/397 (97%)	0.30	30 (7%)	13	9	64, 66, 70, 76	0
All	All	1167/1191 (97%)	0.22	73 (6%)	20	15	58, 66, 73, 99	0

The worst 5 of 73 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	388	VAL	6.1
1	C	388	VAL	5.8
1	C	204	ILE	5.6
1	C	44	PHE	5.1
1	C	148	ALA	4.5

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

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

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

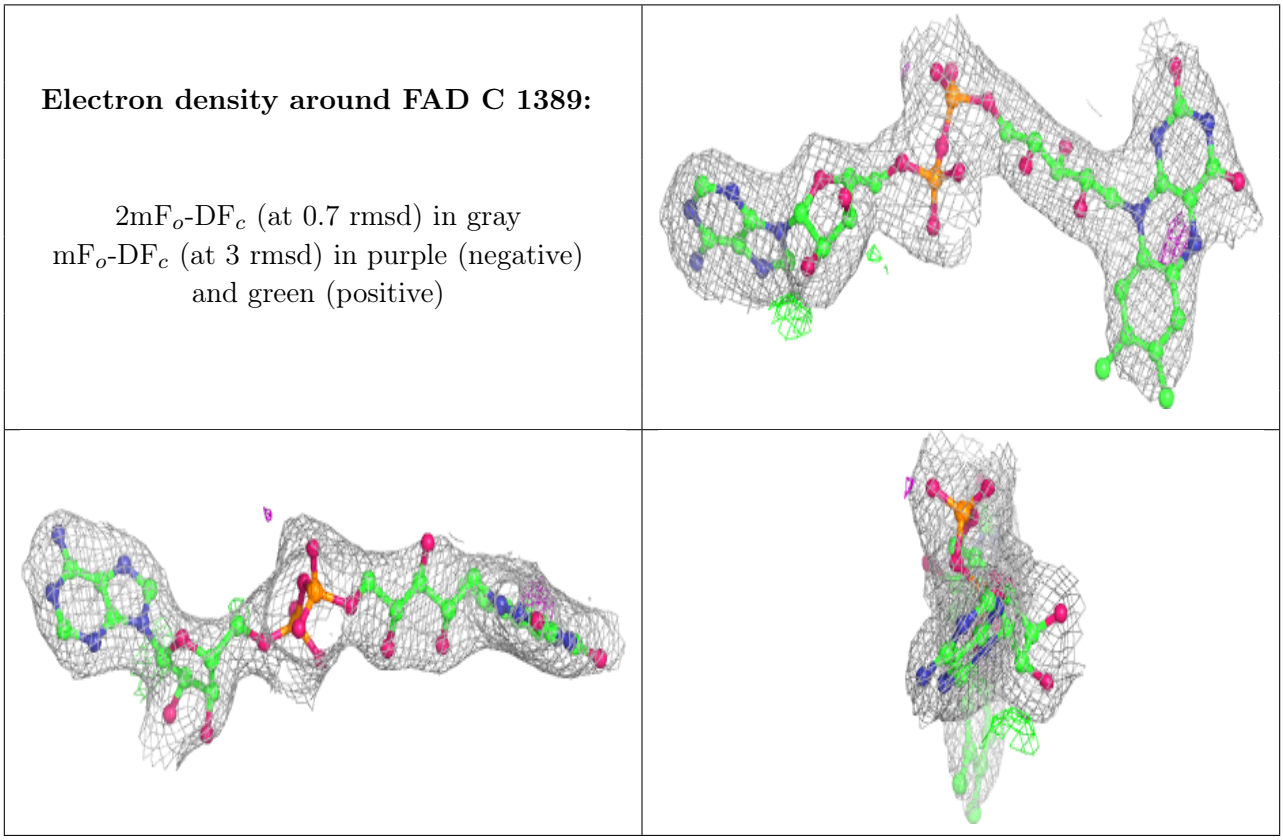
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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	ACT	C	1391	4/4	0.78	0.32	128,129,129,129	0
4	ACT	B	1391	4/4	0.79	0.32	94,94,94,95	0
3	GOL	B	1390	6/6	0.93	0.22	74,77,78,80	0
3	GOL	C	1390	6/6	0.94	0.32	93,95,95,96	0
4	ACT	A	1397	4/4	0.95	0.19	59,60,61,62	0
3	GOL	A	1396	6/6	0.96	0.18	50,60,62,63	0
2	FAD	C	1389	53/53	0.96	0.20	57,68,76,77	0
2	FAD	B	1389	53/53	0.97	0.22	61,70,73,75	0
2	FAD	A	1395	53/53	0.98	0.23	60,66,72,73	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.

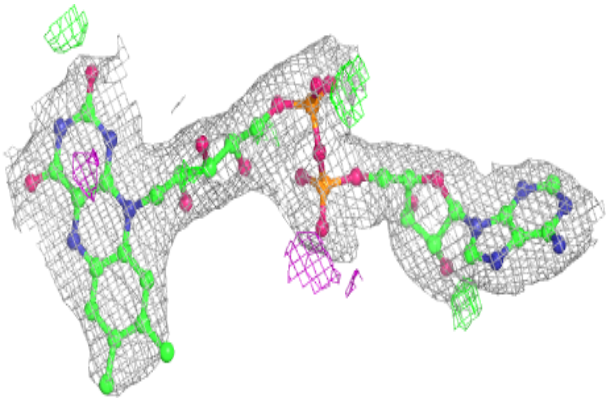
Electron density around FAD C 1389:

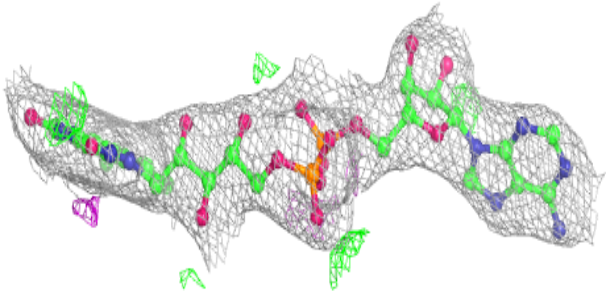
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

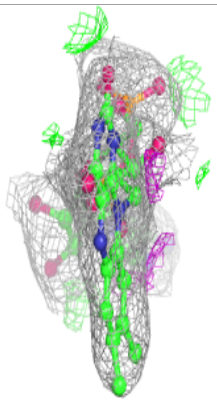


Electron density around FAD B 1389:

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

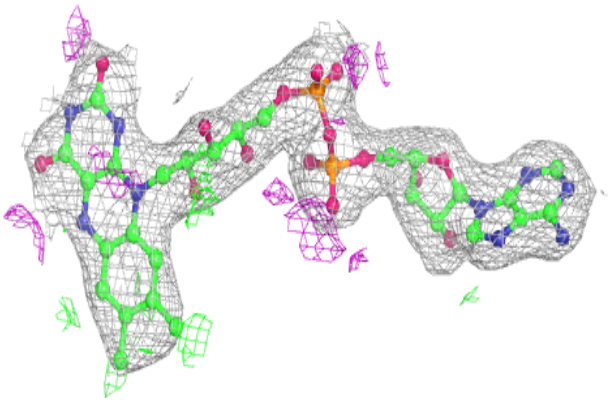


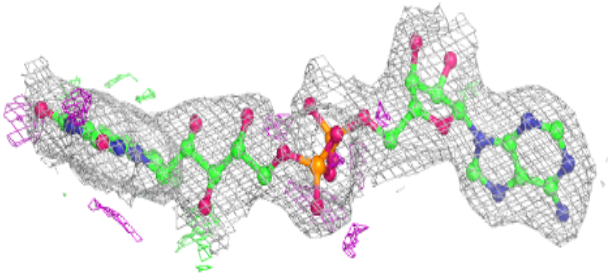


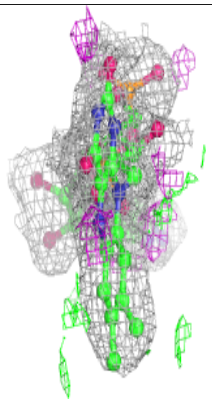


Electron density around FAD A 1395:

$2mF_o$-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.