



wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 04:11 pm BST

PDB ID : 6OJM
Title : Crystal structure of 1,4-dihydroxy-2-naphthoyl-CoA synthase Elizabethkingia anophelis NUHP1
Authors : Seattle Structural Genomics Center for Infectious Disease (SSGCID)
Deposited on : 2019-04-11
Resolution : 1.60 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 ⓘ symbol.

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

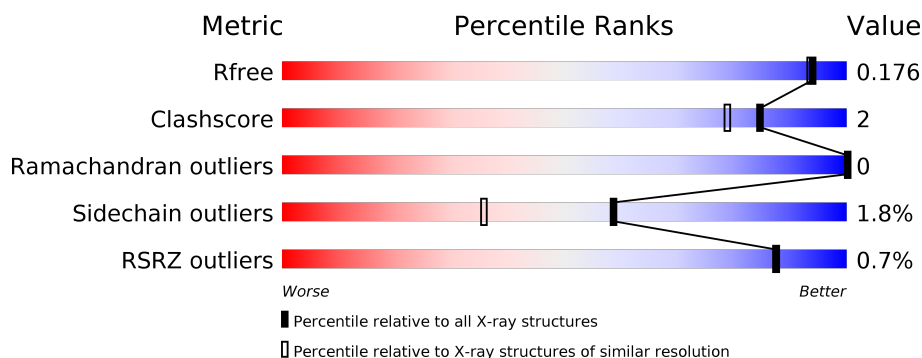
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.60 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	3398 (1.60-1.60)
Clashscore	141614	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)
RSRZ outliers	127900	3321 (1.60-1.60)

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 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 $\leq 5\%$. 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	286	<div> <div>87%</div> <div>5%</div> <div>7%</div> </div>
1	B	286	<div> <div>%</div> <div>92%</div> <div>5%</div> <div>.</div> </div>
1	C	286	<div> <div>%</div> <div>87%</div> <div>6%</div> <div>7%</div> </div>
1	D	286	<div> <div>87%</div> <div>5%</div> <div>8%</div> </div>
1	E	286	<div> <div>87%</div> <div>6%</div> <div>7%</div> </div>
1	F	286	<div> <div>%</div> <div>90%</div> <div>7%</div> <div>.</div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 14286 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 1,4-dihydroxy-2-naphthoyl-CoA synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	266	Total	C	N	O	S	0	11	0
			2125	1358	352	403	12			
1	B	278	Total	C	N	O	S	0	24	0
			2209	1401	381	415	12			
1	C	266	Total	C	N	O	S	0	9	0
			2111	1347	351	401	12			
1	D	263	Total	C	N	O	S	0	6	0
			2070	1317	347	395	11			
1	E	265	Total	C	N	O	S	0	10	0
			2094	1336	351	396	11			
1	F	278	Total	C	N	O	S	0	22	0
			2207	1398	384	413	12			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-7	MET	-	initiating methionine	UNP A0A077EJG6
A	-6	ALA	-	expression tag	UNP A0A077EJG6
A	-5	HIS	-	expression tag	UNP A0A077EJG6
A	-4	HIS	-	expression tag	UNP A0A077EJG6
A	-3	HIS	-	expression tag	UNP A0A077EJG6
A	-2	HIS	-	expression tag	UNP A0A077EJG6
A	-1	HIS	-	expression tag	UNP A0A077EJG6
A	0	HIS	-	expression tag	UNP A0A077EJG6
B	-7	MET	-	initiating methionine	UNP A0A077EJG6
B	-6	ALA	-	expression tag	UNP A0A077EJG6
B	-5	HIS	-	expression tag	UNP A0A077EJG6
B	-4	HIS	-	expression tag	UNP A0A077EJG6
B	-3	HIS	-	expression tag	UNP A0A077EJG6
B	-2	HIS	-	expression tag	UNP A0A077EJG6
B	-1	HIS	-	expression tag	UNP A0A077EJG6
B	0	HIS	-	expression tag	UNP A0A077EJG6
C	-7	MET	-	initiating methionine	UNP A0A077EJG6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-6	ALA	-	expression tag	UNP A0A077EJG6
C	-5	HIS	-	expression tag	UNP A0A077EJG6
C	-4	HIS	-	expression tag	UNP A0A077EJG6
C	-3	HIS	-	expression tag	UNP A0A077EJG6
C	-2	HIS	-	expression tag	UNP A0A077EJG6
C	-1	HIS	-	expression tag	UNP A0A077EJG6
C	0	HIS	-	expression tag	UNP A0A077EJG6
D	-7	MET	-	initiating methionine	UNP A0A077EJG6
D	-6	ALA	-	expression tag	UNP A0A077EJG6
D	-5	HIS	-	expression tag	UNP A0A077EJG6
D	-4	HIS	-	expression tag	UNP A0A077EJG6
D	-3	HIS	-	expression tag	UNP A0A077EJG6
D	-2	HIS	-	expression tag	UNP A0A077EJG6
D	-1	HIS	-	expression tag	UNP A0A077EJG6
D	0	HIS	-	expression tag	UNP A0A077EJG6
E	-7	MET	-	initiating methionine	UNP A0A077EJG6
E	-6	ALA	-	expression tag	UNP A0A077EJG6
E	-5	HIS	-	expression tag	UNP A0A077EJG6
E	-4	HIS	-	expression tag	UNP A0A077EJG6
E	-3	HIS	-	expression tag	UNP A0A077EJG6
E	-2	HIS	-	expression tag	UNP A0A077EJG6
E	-1	HIS	-	expression tag	UNP A0A077EJG6
E	0	HIS	-	expression tag	UNP A0A077EJG6
F	-7	MET	-	initiating methionine	UNP A0A077EJG6
F	-6	ALA	-	expression tag	UNP A0A077EJG6
F	-5	HIS	-	expression tag	UNP A0A077EJG6
F	-4	HIS	-	expression tag	UNP A0A077EJG6
F	-3	HIS	-	expression tag	UNP A0A077EJG6
F	-2	HIS	-	expression tag	UNP A0A077EJG6
F	-1	HIS	-	expression tag	UNP A0A077EJG6
F	0	HIS	-	expression tag	UNP A0A077EJG6

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	261	Total O 265 265	0	5
2	B	238	Total O 241 241	0	5
2	C	221	Total O 226 226	0	6
2	D	219	Total O 220 220	0	1

Continued on next page...

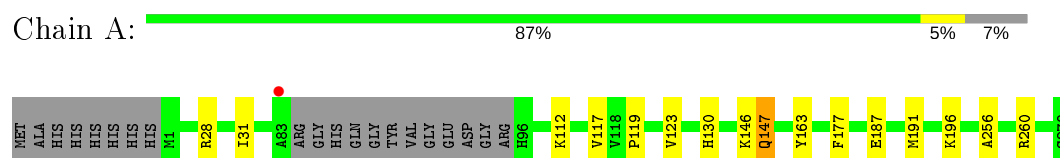
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	E	252	Total 255	O 255	0	3
2	F	258	Total 263	O 263	0	7

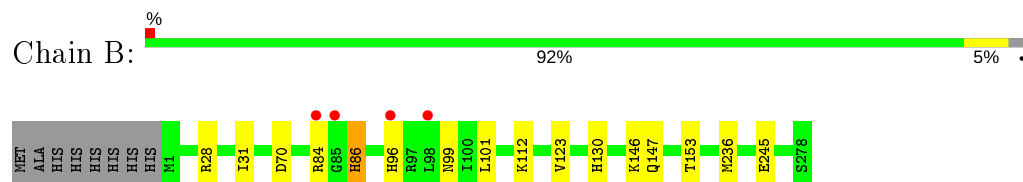
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

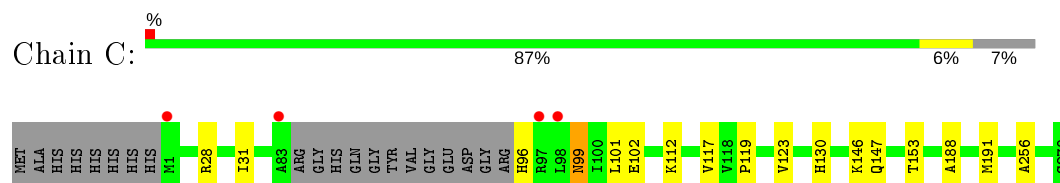
- Molecule 1: 1,4-dihydroxy-2-naphthoyl-CoA synthase



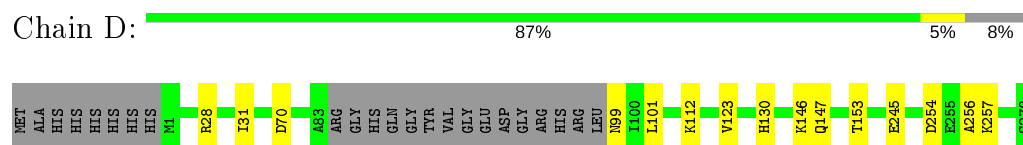
- Molecule 1: 1,4-dihydroxy-2-naphthoyl-CoA synthase



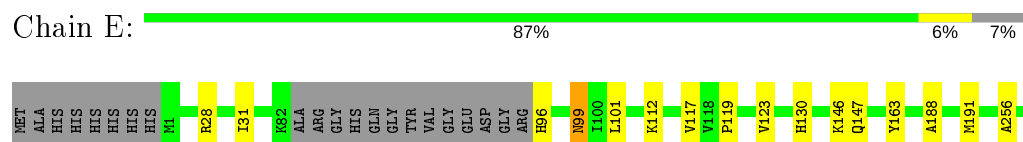
- Molecule 1: 1,4-dihydroxy-2-naphthoyl-CoA synthase



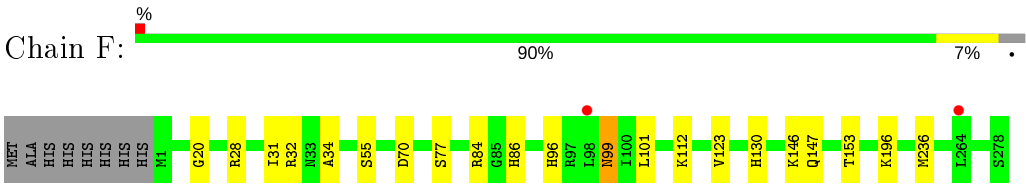
- Molecule 1: 1,4-dihydroxy-2-naphthoyl-CoA synthase



- Molecule 1: 1,4-dihydroxy-2-naphthoyl-CoA synthase



- Molecule 1: 1,4-dihydroxy-2-naphthoyl-CoA synthase



4 Data and refinement statistics

Property	Value	Source
Space group	P 3 2 1	Depositor
Cell constants a, b, c, α , β , γ	138.49 Å 138.49 Å 141.43 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	43.88 – 1.60 45.74 – 1.60	Depositor EDS
% Data completeness (in resolution range)	99.7 (43.88-1.60) 99.7 (45.74-1.60)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.15 (at 1.60 Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.145 , 0.175 0.145 , 0.176	Depositor DCC
R_{free} test set	1983 reflections (0.97%)	wwPDB-VP
Wilson B-factor (Å ²)	16.0	Xtriage
Anisotropy	0.185	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 43.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	0.113 for -h,-k,l	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	14286	wwPDB-VP
Average B, all atoms (Å ²)	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 15.85% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z > 5$	RMSZ	$\# Z > 5$
1	A	0.40	0/2206	0.60	0/2977
1	B	0.38	0/2277	0.58	0/3075
1	C	0.37	0/2186	0.56	0/2953
1	D	0.37	0/2131	0.55	0/2879
1	E	0.40	0/2160	0.59	0/2918
1	F	0.43	0/2269	0.60	0/3062
All	All	0.40	0/13229	0.58	0/17864

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2125	0	2071	11	0
1	B	2209	0	2128	10	3
1	C	2111	0	2043	9	0
1	D	2070	0	1994	8	3
1	E	2094	0	2028	12	0
1	F	2207	0	2132	20	2
2	A	265	0	0	1	0
2	B	241	0	0	2	0
2	C	226	0	0	1	1
2	D	220	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	E	255	0	0	1	0
2	F	263	0	0	2	0
All	All	14286	0	12396	59	6

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.

The worst 5 of 59 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:F:99[A]:ASN:HD22	1:F:99[A]:ASN:H	1.18	0.88
1:B:86[A]:HIS:O	1:B:96[A]:HIS:ND1	2.12	0.80
1:F:86[A]:HIS:O	1:F:96[A]:HIS:ND1	2.17	0.75
1:C:102:GLU:OE1	2:C:301:HOH:O	2.05	0.74
1:D:245:GLU:OE1	2:D:301:HOH:O	2.14	0.66

The worst 5 of 6 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:70:ASP:O	1:D:70:ASP:O[3_454]	0.73	1.47
1:F:70:ASP:C	1:F:70:ASP:O[6_555]	1.25	0.95
1:B:70:ASP:O	1:D:70:ASP:C[3_454]	1.78	0.42
1:B:70:ASP:C	1:D:70:ASP:O[3_454]	1.78	0.42
1:F:70:ASP:C	1:F:70:ASP:C[6_555]	2.10	0.10

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	273/286 (96%)	267 (98%)	6 (2%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	286/286 (100%)	280 (98%)	6 (2%)	0	100	100
1	C	271/286 (95%)	266 (98%)	5 (2%)	0	100	100
1	D	264/286 (92%)	258 (98%)	6 (2%)	0	100	100
1	E	267/286 (93%)	260 (97%)	7 (3%)	0	100	100
1	F	284/286 (99%)	279 (98%)	5 (2%)	0	100	100
All	All	1645/1716 (96%)	1610 (98%)	35 (2%)	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	220/230 (96%)	217 (99%)	3 (1%)	67	47
1	B	223/230 (97%)	219 (98%)	4 (2%)	59	36
1	C	217/230 (94%)	211 (97%)	6 (3%)	43	18
1	D	211/230 (92%)	208 (99%)	3 (1%)	67	47
1	E	215/230 (94%)	211 (98%)	4 (2%)	57	34
1	F	223/230 (97%)	218 (98%)	5 (2%)	52	27
All	All	1309/1380 (95%)	1284 (98%)	25 (2%)	59	34

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

Mol	Chain	Res	Type
1	C	130	HIS
1	D	112	LYS
1	F	130	HIS
1	C	147	GLN
1	D	130	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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](#)

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, 95th 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	266/286 (93%)	-0.93	1 (0%) 92 92	10, 16, 37, 67	1 (0%)
1	B	278/286 (97%)	-0.82	4 (1%) 75 75	11, 18, 36, 48	14 (5%)
1	C	266/286 (93%)	-0.86	4 (1%) 73 73	12, 19, 43, 67	0
1	D	263/286 (91%)	-0.90	0 100 100	12, 20, 39, 57	0
1	E	265/286 (92%)	-0.92	0 100 100	9, 18, 37, 61	1 (0%)
1	F	278/286 (97%)	-0.82	2 (0%) 87 87	8, 14, 39, 53	15 (5%)
All	All	1616/1716 (94%)	-0.87	11 (0%) 87 87	8, 18, 39, 67	31 (1%)

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	83	ALA	4.5
1	B	98[A]	LEU	3.6
1	C	83	ALA	3.4
1	B	96[A]	HIS	3.2
1	C	98	LEU	3.1

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 carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.