wwPDB X-ray Structure Validation Summary Report (i)

Sep 20, 2023 – 06:55 PM EDT

PDB ID	:	5K3H
Title	:	Crystals structure of Acyl-CoA oxidase-1 in Caenorhabditis elegans, Apo form-II
Authors	:	Zhang, X.; Li, K.; Jones, R.A.; Bruner, S.D.; Butcher, R.A.
Deposited on	:	2016-05-19
Resolution	:	2.48 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 (1)) were used in the production of this report:

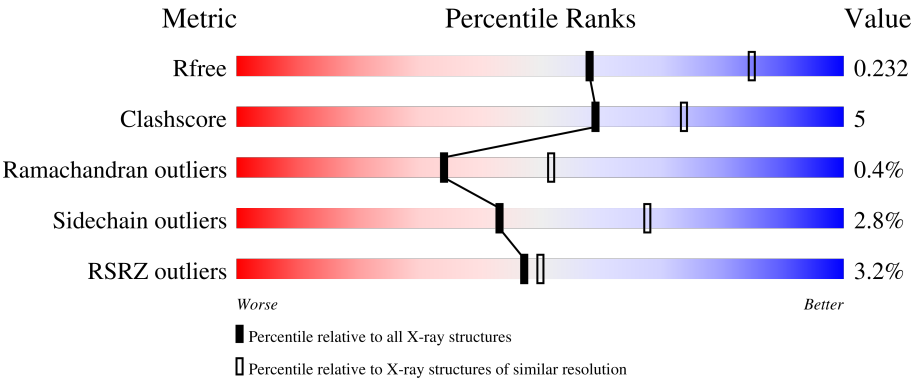
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.48 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	5857 (2.50-2.46)
Clashscore	141614	6594 (2.50-2.46)
Ramachandran outliers	138981	6469 (2.50-2.46)
Sidechain outliers	138945	6471 (2.50-2.46)
RSRZ outliers	127900	5738 (2.50-2.46)

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	684	4% 79% 12% • 7%
1	B	684	2% 82% 10% • 7%
1	C	684	3% 81% 11% 7%
1	D	684	% 83% 8% • 7%
1	E	684	2% 81% 11% • 7%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	684	6% 79% 12% • 7%
1	G	684	3% 82% 11% 7%
1	H	684	2% 81% 10% • 7%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 41742 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 Acyl-coenzyme A oxidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	634	Total	C	N	O	S	0	0	0
			5036	3199	885	928	24			
1	B	635	Total	C	N	O	S	0	0	0
			5039	3198	887	931	23			
1	C	633	Total	C	N	O	S	0	0	0
			5023	3190	885	925	23			
1	D	634	Total	C	N	O	S	0	0	0
			5035	3196	886	930	23			
1	E	635	Total	C	N	O	S	0	0	0
			5036	3197	887	929	23			
1	F	633	Total	C	N	O	S	0	0	0
			5030	3193	885	929	23			
1	G	634	Total	C	N	O	S	0	0	0
			5032	3195	886	928	23			
1	H	634	Total	C	N	O	S	0	0	0
			5035	3196	886	930	23			

There are 80 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	675	HIS	-	expression tag	UNP O62140
A	676	HIS	-	expression tag	UNP O62140
A	677	HIS	-	expression tag	UNP O62140
A	678	HIS	-	expression tag	UNP O62140
A	679	HIS	-	expression tag	UNP O62140
A	680	HIS	-	expression tag	UNP O62140
A	681	HIS	-	expression tag	UNP O62140
A	682	HIS	-	expression tag	UNP O62140
A	683	HIS	-	expression tag	UNP O62140
A	684	HIS	-	expression tag	UNP O62140
B	675	HIS	-	expression tag	UNP O62140
B	676	HIS	-	expression tag	UNP O62140
B	677	HIS	-	expression tag	UNP O62140

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	678	HIS	-	expression tag	UNP O62140
B	679	HIS	-	expression tag	UNP O62140
B	680	HIS	-	expression tag	UNP O62140
B	681	HIS	-	expression tag	UNP O62140
B	682	HIS	-	expression tag	UNP O62140
B	683	HIS	-	expression tag	UNP O62140
B	684	HIS	-	expression tag	UNP O62140
C	675	HIS	-	expression tag	UNP O62140
C	676	HIS	-	expression tag	UNP O62140
C	677	HIS	-	expression tag	UNP O62140
C	678	HIS	-	expression tag	UNP O62140
C	679	HIS	-	expression tag	UNP O62140
C	680	HIS	-	expression tag	UNP O62140
C	681	HIS	-	expression tag	UNP O62140
C	682	HIS	-	expression tag	UNP O62140
C	683	HIS	-	expression tag	UNP O62140
C	684	HIS	-	expression tag	UNP O62140
D	675	HIS	-	expression tag	UNP O62140
D	676	HIS	-	expression tag	UNP O62140
D	677	HIS	-	expression tag	UNP O62140
D	678	HIS	-	expression tag	UNP O62140
D	679	HIS	-	expression tag	UNP O62140
D	680	HIS	-	expression tag	UNP O62140
D	681	HIS	-	expression tag	UNP O62140
D	682	HIS	-	expression tag	UNP O62140
D	683	HIS	-	expression tag	UNP O62140
D	684	HIS	-	expression tag	UNP O62140
E	675	HIS	-	expression tag	UNP O62140
E	676	HIS	-	expression tag	UNP O62140
E	677	HIS	-	expression tag	UNP O62140
E	678	HIS	-	expression tag	UNP O62140
E	679	HIS	-	expression tag	UNP O62140
E	680	HIS	-	expression tag	UNP O62140
E	681	HIS	-	expression tag	UNP O62140
E	682	HIS	-	expression tag	UNP O62140
E	683	HIS	-	expression tag	UNP O62140
E	684	HIS	-	expression tag	UNP O62140
F	675	HIS	-	expression tag	UNP O62140
F	676	HIS	-	expression tag	UNP O62140
F	677	HIS	-	expression tag	UNP O62140
F	678	HIS	-	expression tag	UNP O62140
F	679	HIS	-	expression tag	UNP O62140

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	680	HIS	-	expression tag	UNP O62140
F	681	HIS	-	expression tag	UNP O62140
F	682	HIS	-	expression tag	UNP O62140
F	683	HIS	-	expression tag	UNP O62140
F	684	HIS	-	expression tag	UNP O62140
G	675	HIS	-	expression tag	UNP O62140
G	676	HIS	-	expression tag	UNP O62140
G	677	HIS	-	expression tag	UNP O62140
G	678	HIS	-	expression tag	UNP O62140
G	679	HIS	-	expression tag	UNP O62140
G	680	HIS	-	expression tag	UNP O62140
G	681	HIS	-	expression tag	UNP O62140
G	682	HIS	-	expression tag	UNP O62140
G	683	HIS	-	expression tag	UNP O62140
G	684	HIS	-	expression tag	UNP O62140
H	675	HIS	-	expression tag	UNP O62140
H	676	HIS	-	expression tag	UNP O62140
H	677	HIS	-	expression tag	UNP O62140
H	678	HIS	-	expression tag	UNP O62140
H	679	HIS	-	expression tag	UNP O62140
H	680	HIS	-	expression tag	UNP O62140
H	681	HIS	-	expression tag	UNP O62140
H	682	HIS	-	expression tag	UNP O62140
H	683	HIS	-	expression tag	UNP O62140
H	684	HIS	-	expression tag	UNP O62140

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	150	Total O 150 150	0	0
2	B	180	Total O 180 180	0	0
2	C	199	Total O 199 199	0	0
2	D	215	Total O 215 215	0	0
2	E	188	Total O 188 188	0	0
2	F	156	Total O 156 156	0	0
2	G	203	Total O 203 203	0	0

Continued on next page...

Continued from previous page...

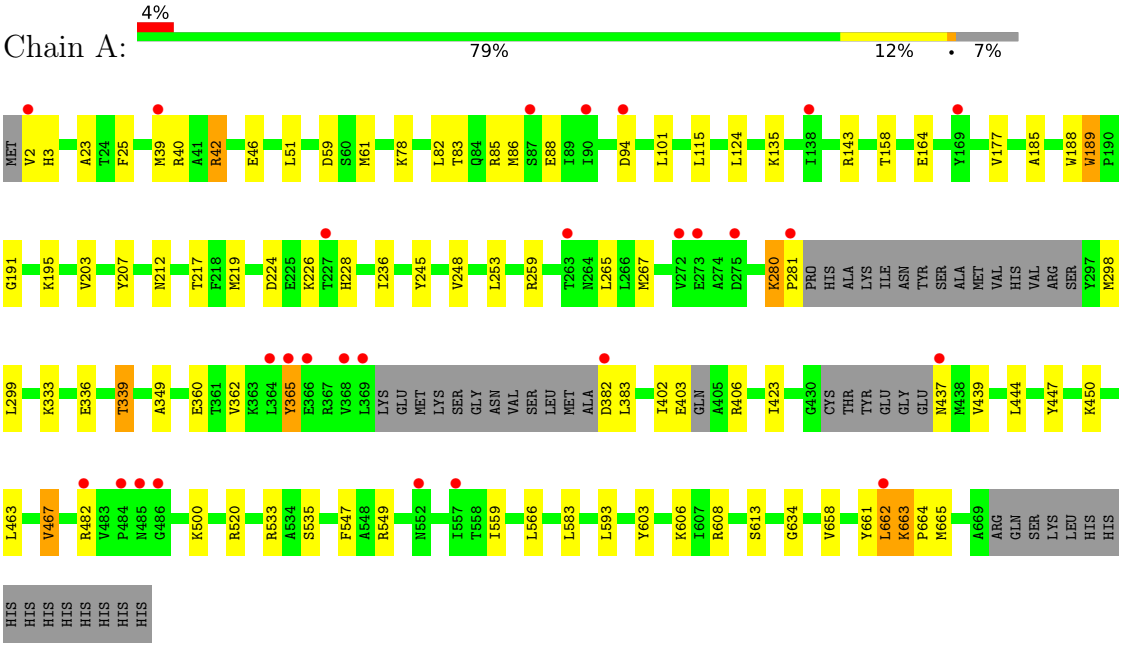
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	H	185	Total 185	O 185	0	0

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.

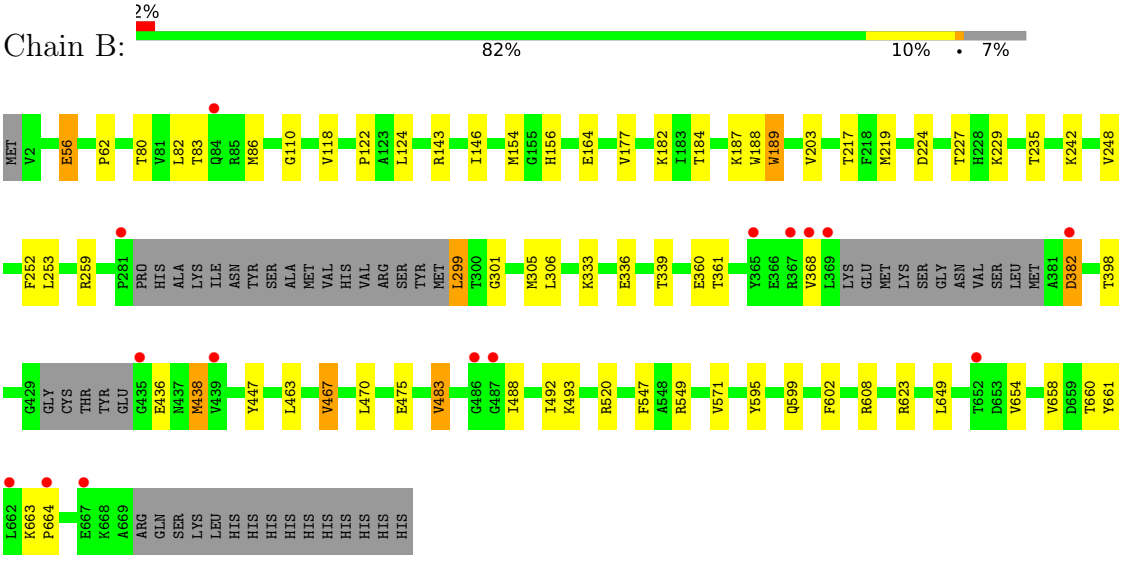
• Molecule 1: Acyl-coenzyme A oxidase

Chain A:



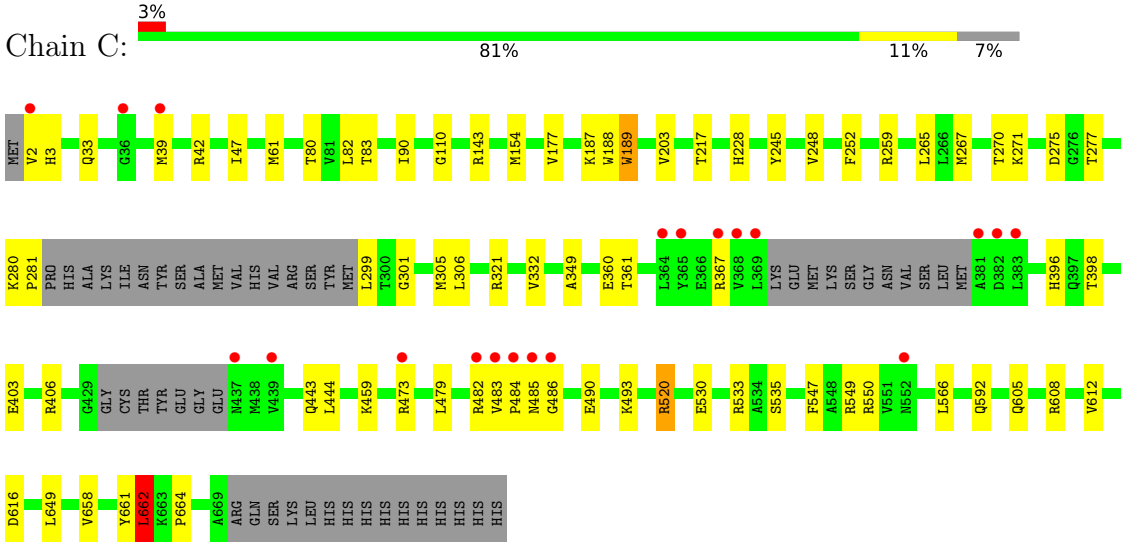
• Molecule 1: Acyl-coenzyme A oxidase

Chain B:



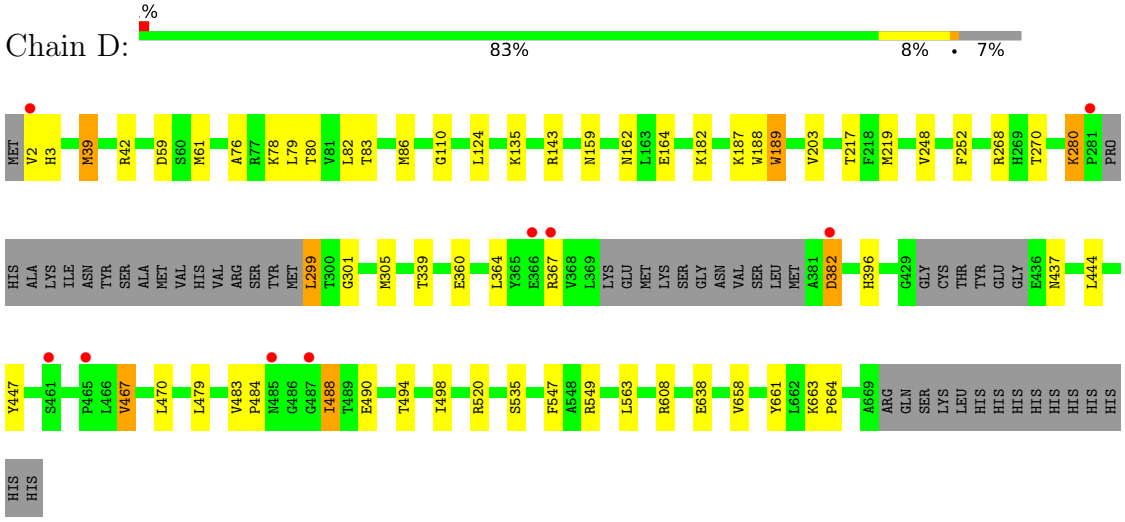
• Molecule 1: Acyl-coenzyme A oxidase

Chain C:



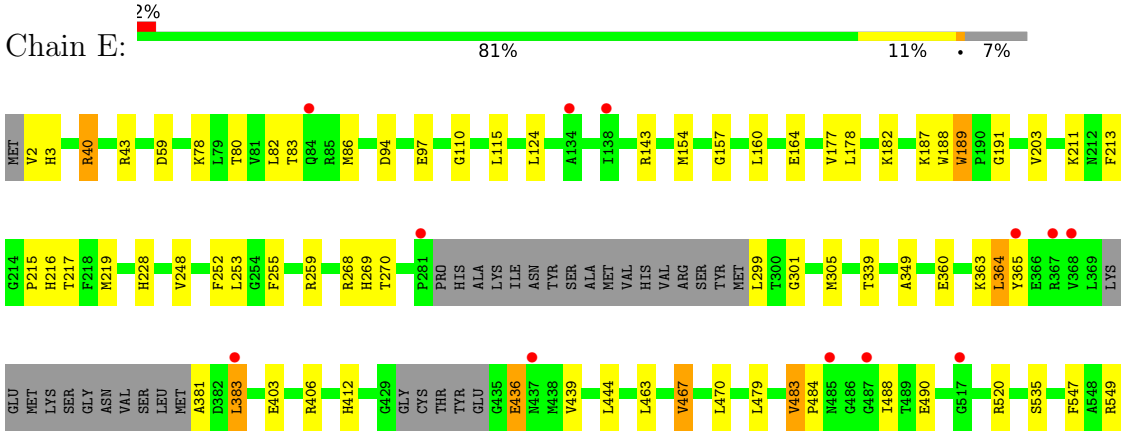
• Molecule 1: Acyl-coenzyme A oxidase

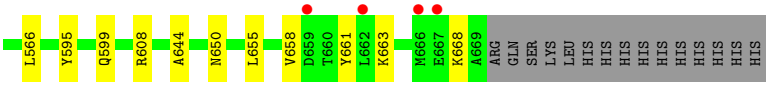
Chain D:



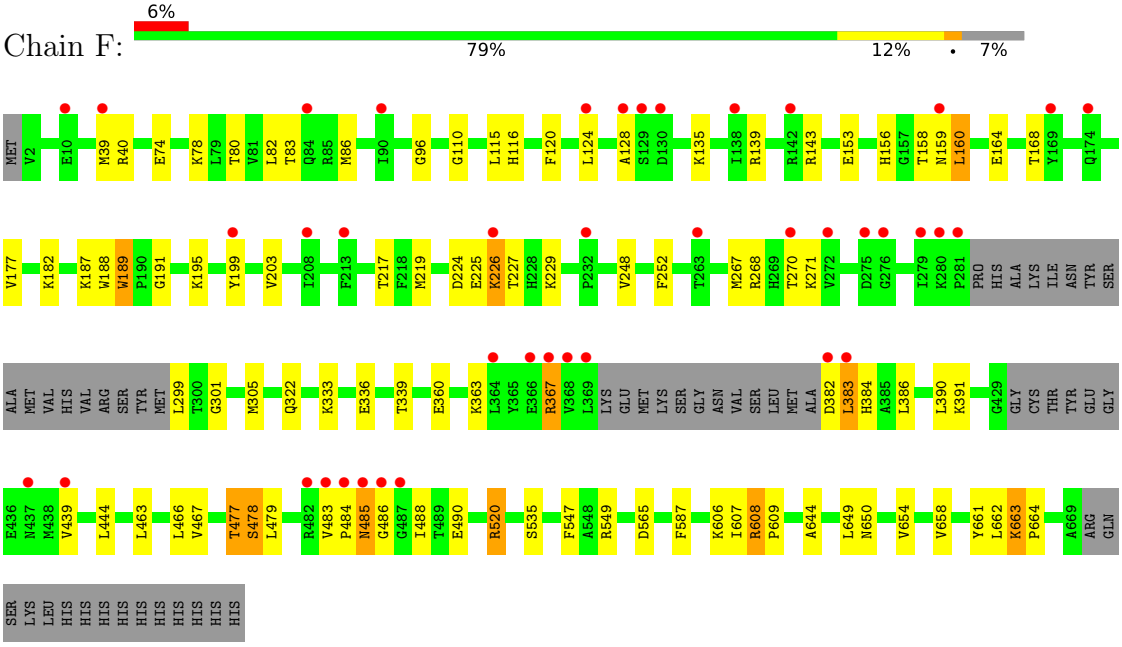
• Molecule 1: Acyl-coenzyme A oxidase

Chain E:

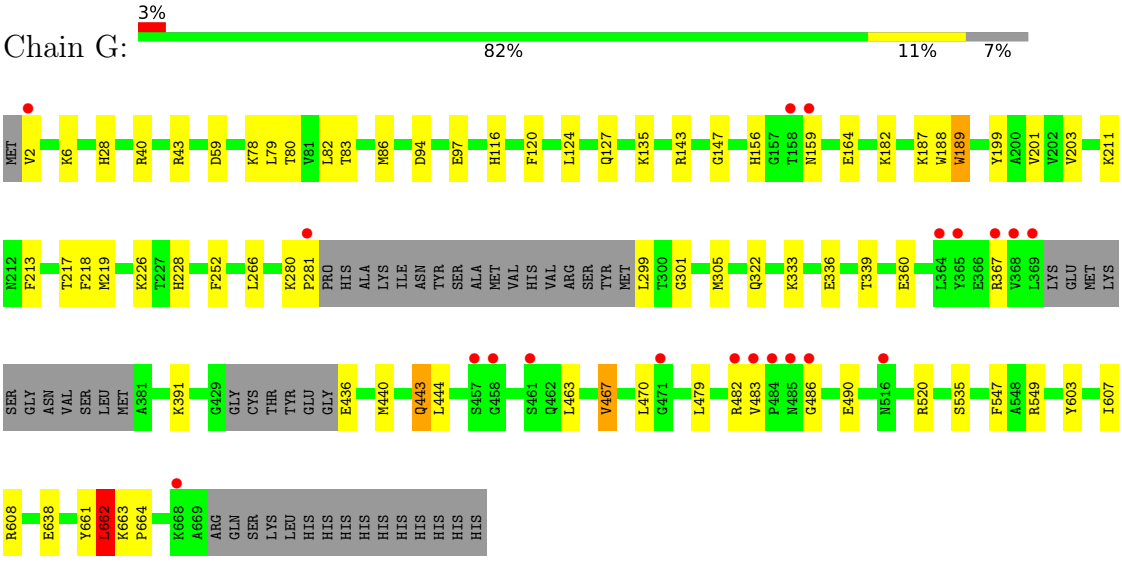




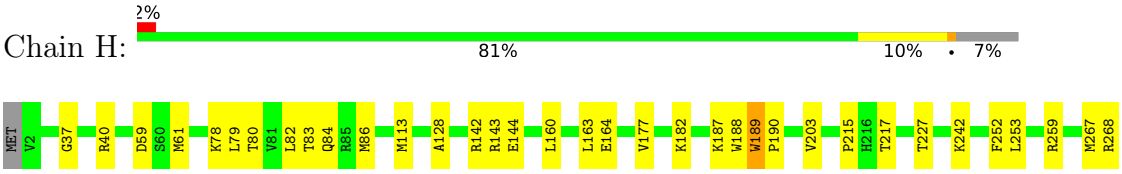
• Molecule 1: Acyl-coenzyme A oxidase

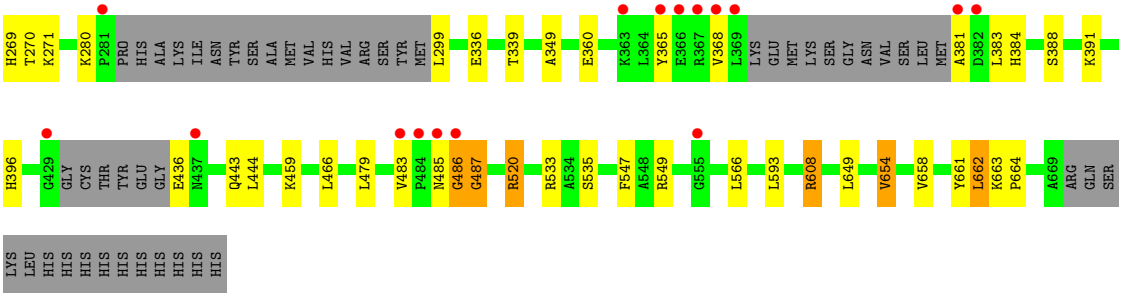


• Molecule 1: Acyl-coenzyme A oxidase



• Molecule 1: Acyl-coenzyme A oxidase




H269 T270 K271 K280 P281 PRO HIS ALA LYS ILE ASN TYR SER ALA MET VAL HIS VAL ARG SER TYR MET L299 E336 T339 A349 E360 K363 L364 Y365 E366 R367 V368 L369 LYS GLU MET LYS SER GLY ASN VAL SER LEU MET A381 D382 L383 H384 S388 K391
H396 G429 GLY CYS THR TYR GLU GLY E436 N437 Q443 L444 K459 L466 L479 V483 P484 N485 G486 G487 R520 R533 A534 S535 F547 A548 R549 G555 L566 L593 R608 L649 V654 V658 Y661 L662 K663 P664 A669 ARG GLN SER
LYS LEU HIS HIS HIS HIS HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	153.07Å 140.79Å 154.93Å 90.00° 117.20° 90.00°	Depositor
Resolution (Å)	49.24 – 2.48 49.24 – 2.48	Depositor EDS
% Data completeness (in resolution range)	99.8 (49.24-2.48) 99.7 (49.24-2.48)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	0.14	Depositor
$< I/\sigma(I) >$ [1]	4.16 (at 2.48Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, R_{free}	0.191 , 0.232 0.193 , 0.232	Depositor DCC
R_{free} test set	10075 reflections (4.89%)	wwPDB-VP
Wilson B-factor (Å^2)	23.3	Xtriage
Anisotropy	0.331	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 46.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.074 for l,-k,h	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	41742	wwPDB-VP
Average B, all atoms (Å^2)	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 55.36 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 3.1995e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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

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	0.44	0/5137	0.61	0/6943
1	B	0.45	0/5140	0.62	0/6949
1	C	0.46	0/5124	0.60	0/6928
1	D	0.48	0/5136	0.62	0/6944
1	E	0.45	0/5137	0.64	0/6945
1	F	0.44	0/5131	0.60	0/6937
1	G	0.47	0/5133	0.62	0/6940
1	H	0.46	0/5136	0.62	0/6944
All	All	0.45	0/41074	0.62	0/55530

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	5036	0	5056	57	0
1	B	5039	0	5058	45	0
1	C	5023	0	5047	44	0
1	D	5035	0	5055	37	0
1	E	5036	0	5056	49	0
1	F	5030	0	5050	73	0
1	G	5032	0	5053	46	0
1	H	5035	0	5055	55	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	150	0	0	5	0
2	B	180	0	0	2	0
2	C	199	0	0	3	0
2	D	215	0	0	2	0
2	E	188	0	0	1	0
2	F	156	0	0	2	0
2	G	203	0	0	1	0
2	H	185	0	0	6	0
All	All	41742	0	40430	376	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 5.

The worst 5 of 376 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:658:VAL:HA	1:A:662:LEU:HB2	1.30	1.13
1:G:661:TYR:O	1:G:664:PRO:HD2	1.56	1.04
1:H:658:VAL:HA	1:H:662:LEU:HB2	1.36	1.03
1:F:658:VAL:HG22	1:F:662:LEU:HD12	1.44	0.99
1:F:658:VAL:HA	1:F:662:LEU:HB2	1.44	0.97

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	624/684 (91%)	610 (98%)	13 (2%)	1 (0%)	47 66
1	B	627/684 (92%)	613 (98%)	12 (2%)	2 (0%)	41 59
1	C	625/684 (91%)	609 (97%)	14 (2%)	2 (0%)	41 59

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	626/684 (92%)	613 (98%)	10 (2%)	3 (0%)	29	46
1	E	627/684 (92%)	614 (98%)	11 (2%)	2 (0%)	41	59
1	F	625/684 (91%)	607 (97%)	15 (2%)	3 (0%)	29	46
1	G	626/684 (92%)	617 (99%)	7 (1%)	2 (0%)	41	59
1	H	626/684 (92%)	609 (97%)	14 (2%)	3 (0%)	29	46
All	All	5006/5472 (92%)	4892 (98%)	96 (2%)	18 (0%)	34	52

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	189	TRP
1	B	189	TRP
1	C	189	TRP
1	D	189	TRP
1	D	488	ILE

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	534/579 (92%)	516 (97%)	18 (3%)	37	61
1	B	534/579 (92%)	521 (98%)	13 (2%)	49	72
1	C	532/579 (92%)	518 (97%)	14 (3%)	46	70
1	D	534/579 (92%)	518 (97%)	16 (3%)	41	65
1	E	533/579 (92%)	520 (98%)	13 (2%)	49	72
1	F	534/579 (92%)	519 (97%)	15 (3%)	43	67
1	G	533/579 (92%)	517 (97%)	16 (3%)	41	65
1	H	534/579 (92%)	520 (97%)	14 (3%)	46	70
All	All	4268/4632 (92%)	4149 (97%)	119 (3%)	43	67

5 of 119 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	520	ARG
1	H	365	TYR
1	E	520	ARG
1	H	339	THR
1	H	662	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	322	GLN
1	G	443	GLN
1	H	269	HIS
1	D	159	ASN
1	A	228	HIS

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)

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

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	634/684 (92%)	0.22	27 (4%)	35	37	18, 30, 55, 71	0
1	B	635/684 (92%)	0.06	15 (2%)	59	61	15, 26, 46, 67	0
1	C	633/684 (92%)	0.09	20 (3%)	47	50	15, 26, 43, 63	0
1	D	634/684 (92%)	-0.04	9 (1%)	75	77	12, 22, 41, 64	0
1	E	635/684 (92%)	0.09	16 (2%)	57	59	14, 26, 46, 65	0
1	F	633/684 (92%)	0.34	41 (6%)	18	18	18, 30, 58, 71	0
1	G	634/684 (92%)	0.02	20 (3%)	47	50	13, 23, 44, 69	0
1	H	634/684 (92%)	0.07	16 (2%)	57	59	15, 27, 45, 81	0
All	All	5072/5472 (92%)	0.10	164 (3%)	47	50	12, 26, 48, 81	0

The worst 5 of 164 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	368	VAL	7.7
1	F	486	GLY	7.4
1	F	485	ASN	6.1
1	B	435	GLY	6.0
1	H	368	VAL	5.5

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

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.