Structure and mechanism of ORF36, an Aminosugar Oxidizing Enzyme in Everninomicin Biosynthesis†

Supporting Information

Jessica L. Vey[1], Ahmad Al-Mestarihi[2], Yunfeng Hu[2,&], Michael A. Funk[1,@], Brian O. Bachmann[2,3,], and T. M. Iverson[1,3,]**

Departments of [1]Pharmacology, and [3]Biochemistry, Vanderbilt University Medical Center, Nashville, TN 37232, USA and [2]Department of Chemistry, Vanderbilt University, Nashville, TN 37235, USA.

AUTHOR EMAIL ADDRESS: BOB: brian.o.bachmann@vanderbilt.edu. TMI: tina.iverson@vanderbilt.edu.

RECEIVED DATE

TITLE RUNNING HEAD. Structure and mechanism of nitrososynthase

‡ The atomic coordinates and structure factors for ORF36 (code 3MXL) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[&]Present address: Department of Chemistry, Brown University, Providence, RI 02912

[@]Present address: Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

* To whom correspondence should be addressed. BOB: fax, (615) 322-8865; phone, (615) 322-8865; email: brian.o.bachmann@vanderbilt.edu. TMI: fax, (615) 343-6532; phone, (615) 322-7817; email tina.iverson@vanderbilt.edu.

Supplementary Movie

This movie demonstrates the differences in active site loop conformations between ORF36, human acyl-coA dehydrogenase, and *A. baumannii* 4-hydroxyphenylacetate monooxygenase. The active site loops of interest morph from the conformations observed in ORF36 to those observed in the acyl-coA dehydrogenase, then back to ORF36, then to the conformations observed in 4-hydroxyphenylacetate monooxygenase, and again back to ORF36. The active site loops are labeled L1 to L11, and are displayed and colored as in Figure 7.

Supplementary Figure 1

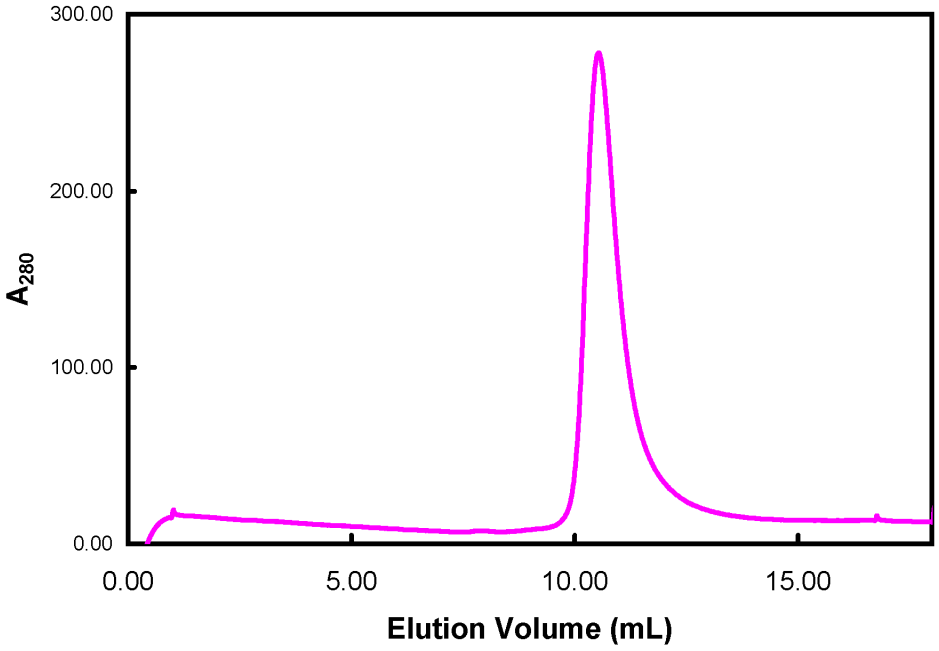

Figure S1. Size exclusion chromatogram of ORF36. Size exclusion chromatography of ORF36 on a Superdex S-200 column resulted in a single, monodisperse peak. The 10.8 mL elution volume corresponds to an approximate molecular weight of 185 kD, which is consistent with the 176 kD calculated for a tetramer.

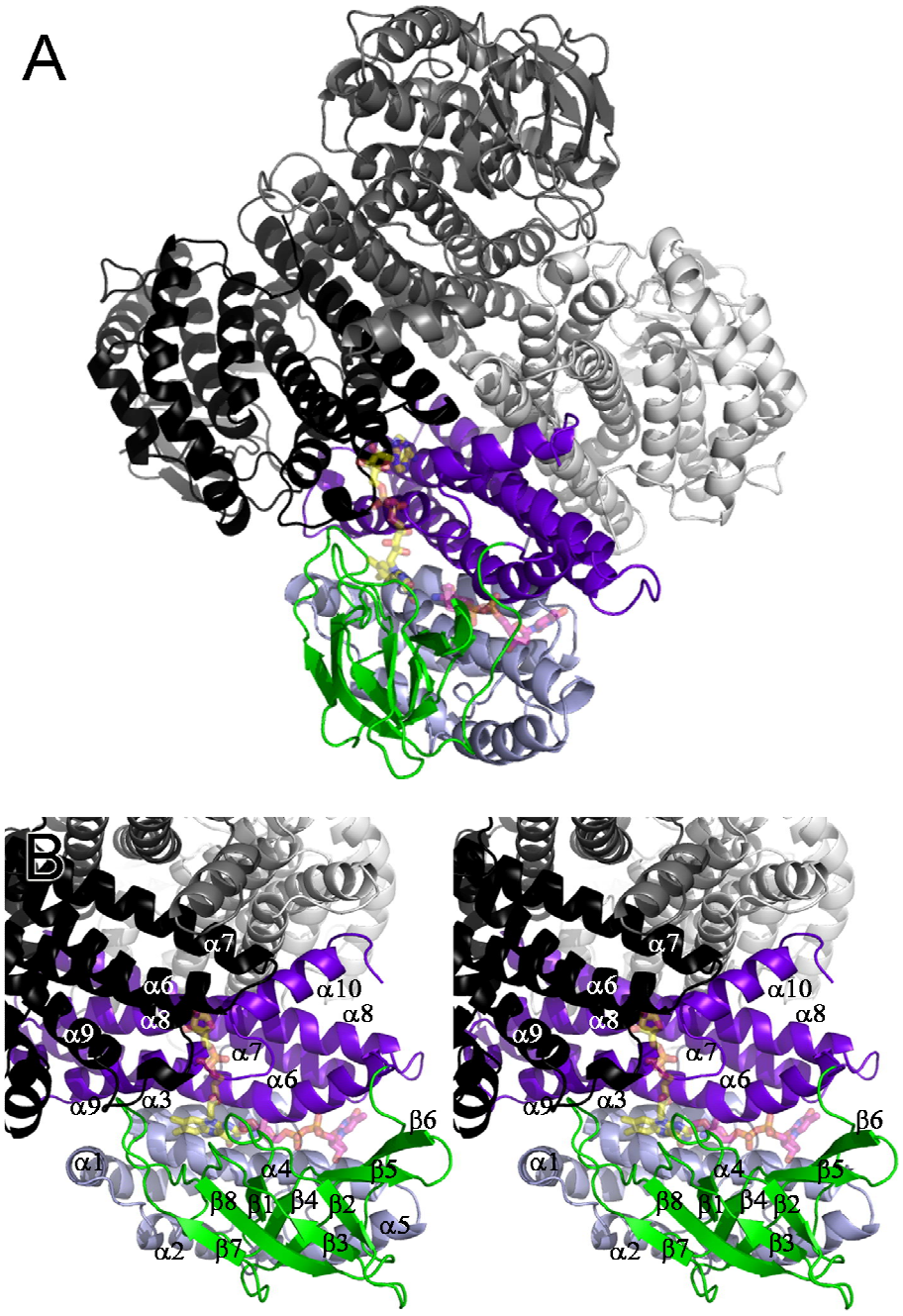

Figure S2. Location of the FAD and TDP-L-evernosamine molecules modeled within ORF36 active site. Chain A of ORF36 is shown colored by domain as in Figure 2. Chains B, C and D are colored white, black and grey, respectively. The modeled small molecules (shown as transparent sticks) are colored as follows: FAD, yellow carbons, TDP-L-evernosamine, magenta carbons, oxygen, red; nitrogen, blue; sulfur, gold; phosphorus, orange. (A) The physiologically relevant tetramer, with

modeled ligands shown at one active site. (B) A closer view of the labeled ORF36 monomer, shown in stereo. The monomer is shown in the same orientation as in Figure 5B, with the other three chains of the tetramer colored as in (A). Selected secondary structural elements from chain C are labeled in white text.

Supplementary Figure 3

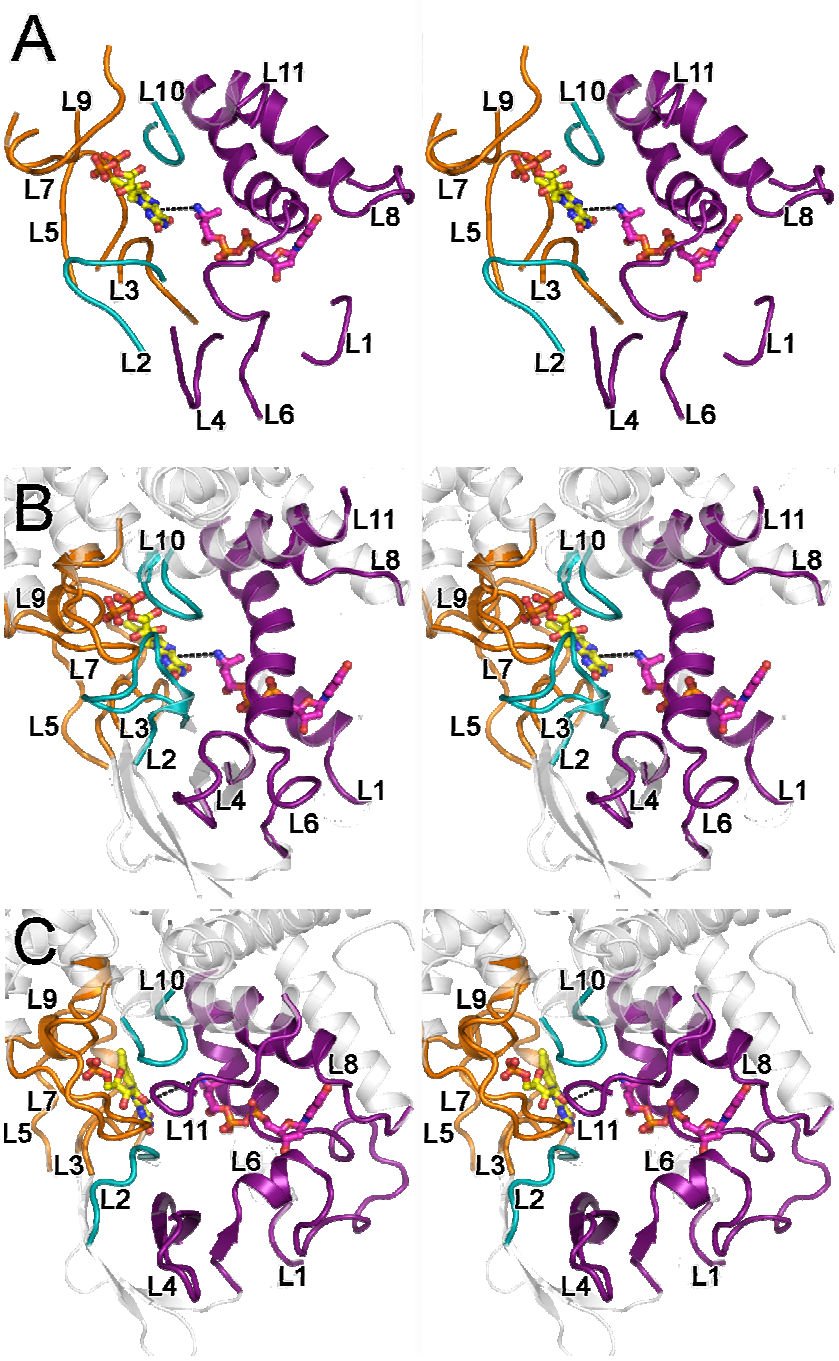

Figure S3. Stereoview diagrams depicting the active site loops. Panel (A) shows the ORF36 active site loops, displayed as in Figure 7a. (B) Human short branched-chain acyl Co-A dehydrogenase active site, displayed as in Figure 7b. (C) *A. baumannii* 4-hydroxyphenylacetate monooxygenase displayed as in Figure 7c.

Supplementary Figure 4

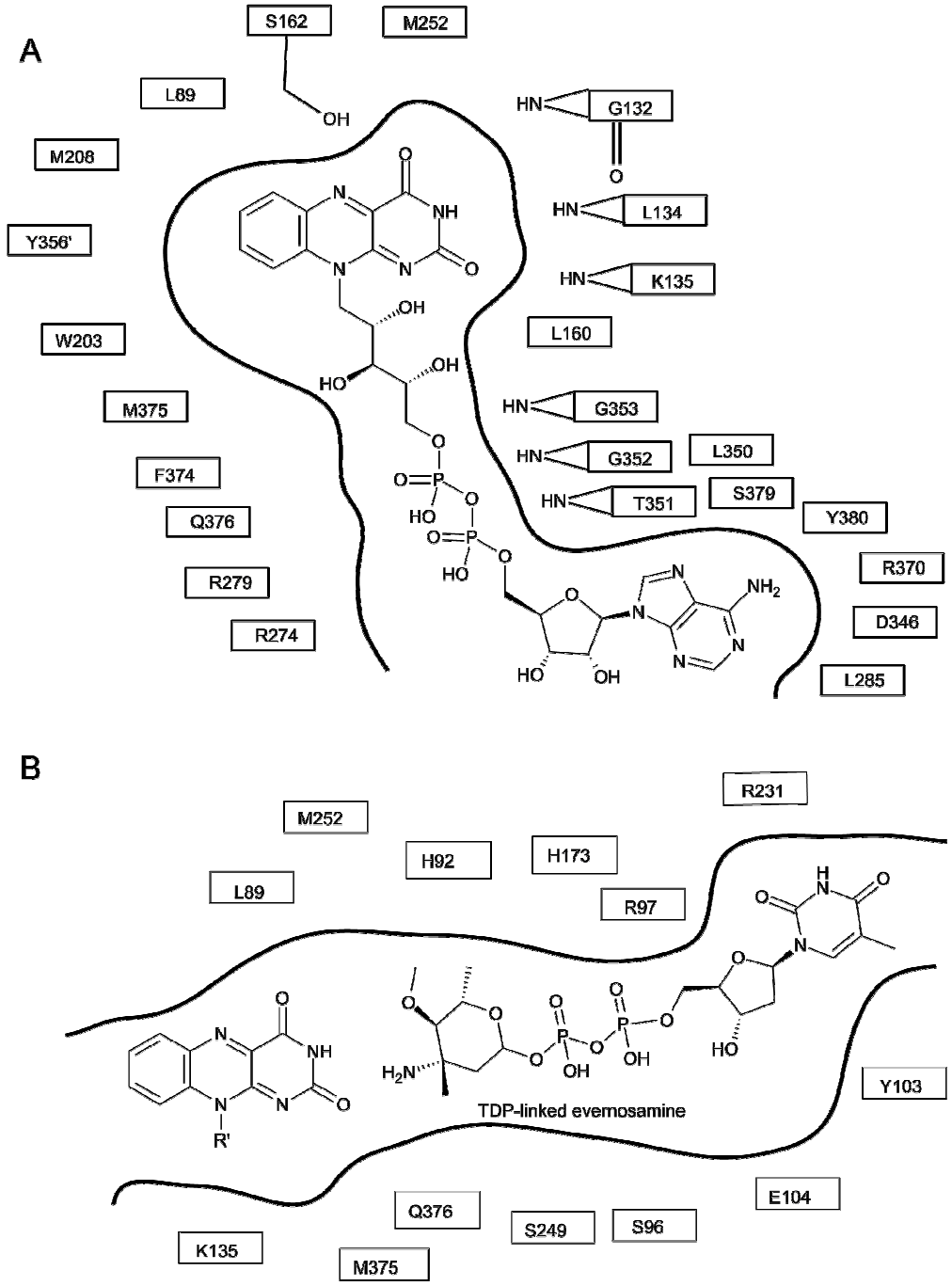

Figure S4. Schematic diagram of the active site. (A) Residues lining the putative FAD binding site are diagrammed. (B) Residues in the vicinity of the modeled FAD cofactor and TDP-L-evernosamine substrate are shown. R' = the ribityl diphosphonucleotide of the modeled FAD cofactor.

Supplementary Figure 5

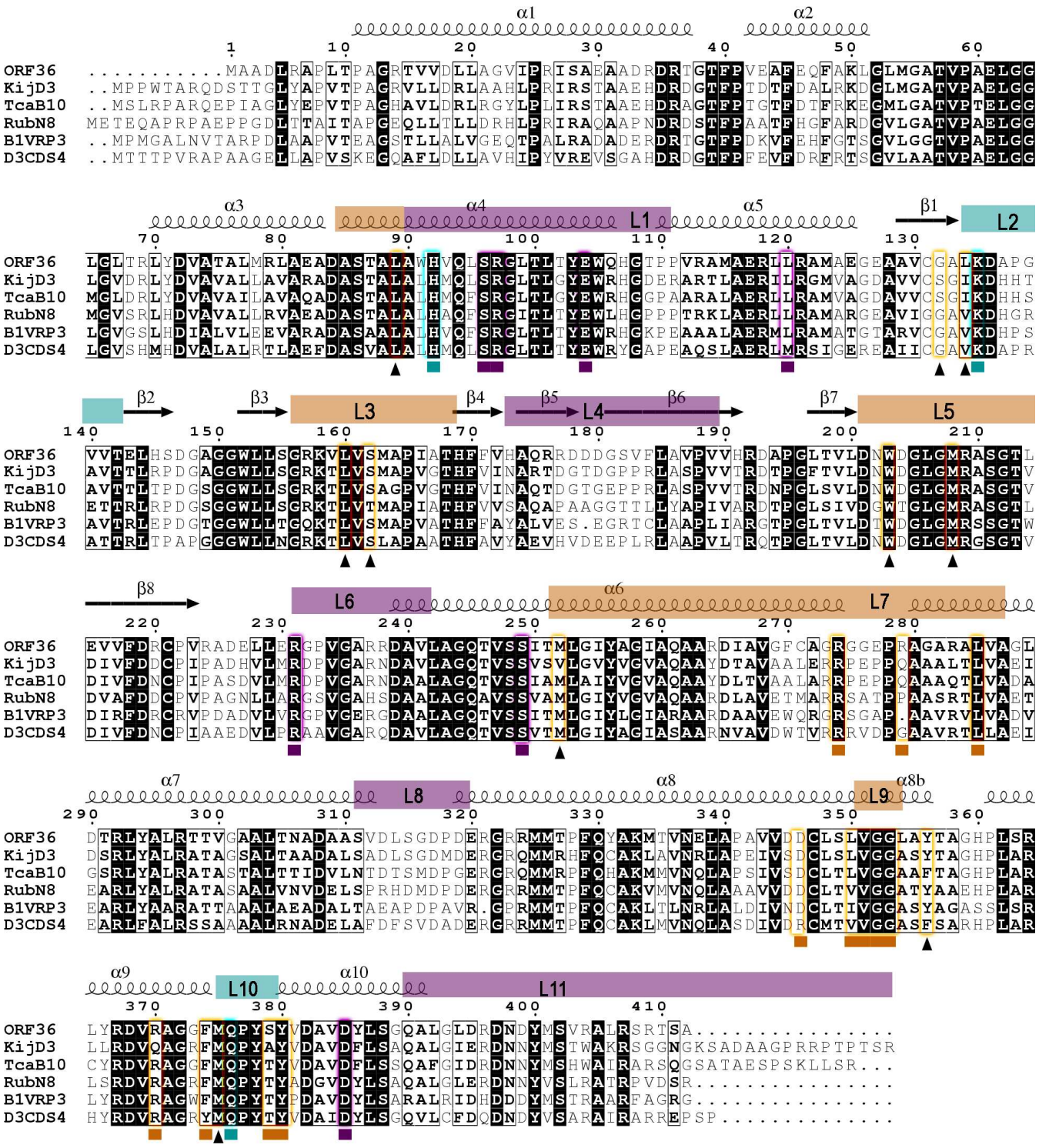

Figure S5. Sequence alignment of ORF36 and putative nitrososynthases. Strictly conserved residues are highlighted in black and residues with >70% similarity are boxed. Residues near the intersection of modeled isoalloxazine and TDP-L-evernosamine are boxed in teal and marked with a teal square;

residues near the isoalloxazine of modeled FAD are boxed in orange and marked with black triangles; residues near the nucleotide of modeled FAD are boxed in orange and marked with orange squares; and residues near the modeled TDP-L-evernosamine are boxed in purple and marked with purple squares. Secondary structural elements and loops L1-L11 are marked above the alignment and highlighted orange (areas involved in flavin binding), purple (areas likely involved in substrate binding) or teal (areas located near the intersection of the putative flavin and substrate binding sites). Sequences aligned with ORF36 are: KijD3 (UniProtKB ID B3TMR1); TcaB10 from the tetrocarcin A biosynthetic pathway of *M. chalcea* (UniProtKB ID B5L6K4); RubN8 from the rubradirin biosynthetic pathway of (UniProtKB ID Q2PC69); B1VRP3, uncharacterized protein from *S. griseus*; D3CDS4, uncharacterized protein from *Micromonospora* sp. L5. These proteins were identified by a BLAST search of the ORF36 sequence (UniProtKB ID B5APQ9) and all had scores > 450 and E-values of less than $1x10^{-125}$. Sequences were aligned using ClustalW (*1*), and the figure was generated using ESPript (*2*).

Supplementary Figure 6

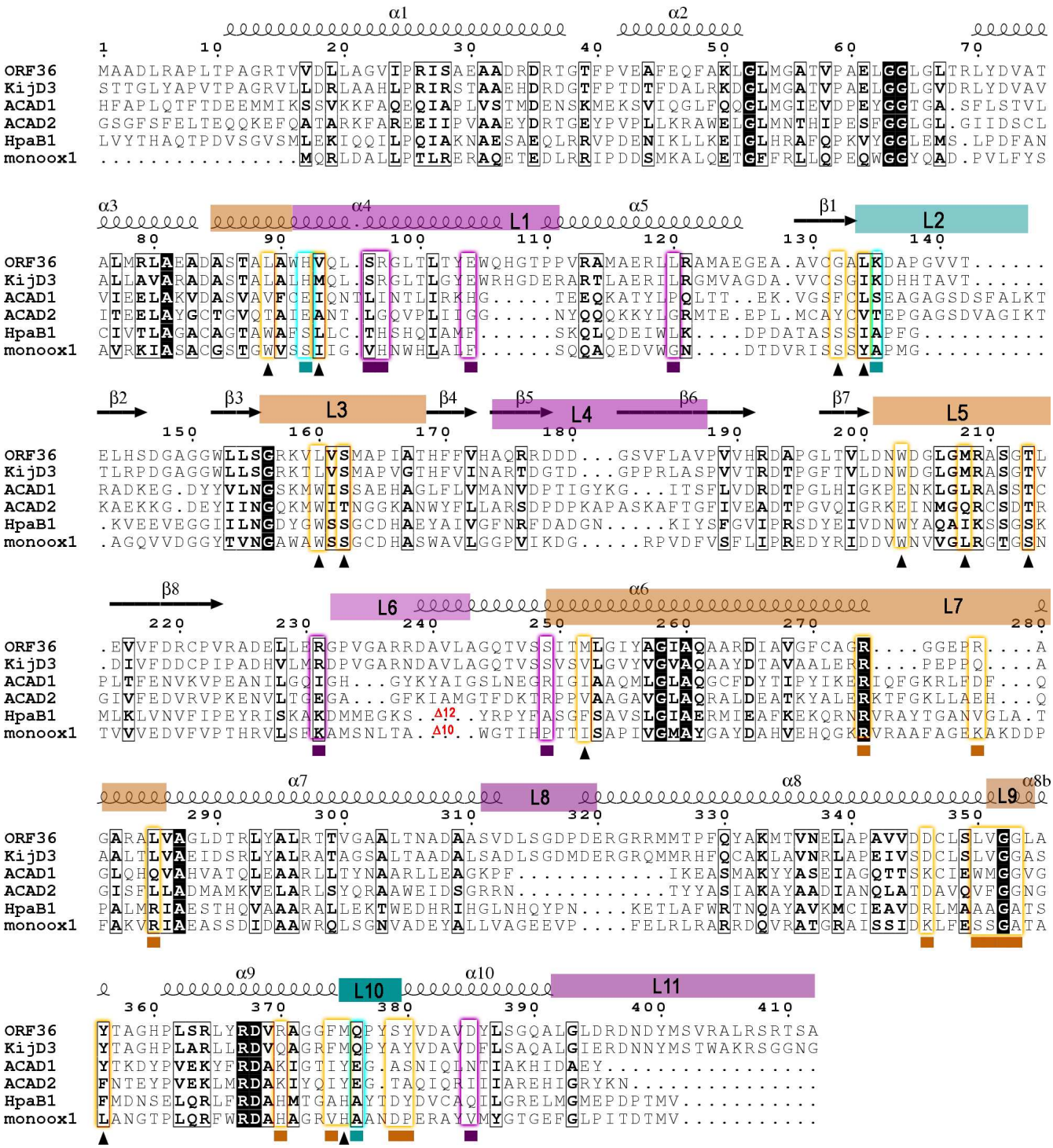

Figure S6. Sequence alignment of ORF36, KijD3, selected acyl-coA dehydrogenases, and selected Class D flavin monooxygenases. Loops and residues at the active site are highlighted as in Supplemental Figure S4. Of particular interest are the residues involved in the isoalloxazine binding site. For clarity, select divergent portions of individual sequences were truncated and are denoted with

red lettering. Sequences aligned with ORF36 are: KijD3 (UniProtKB ID B3TMR1); **ACAD1**, short branched-chain acyl-coA dehydrogenase (*H. sapiens*), 2JIF; **ACAD2**, medium-chain acyl-coA dehydrogenase (*S. scrofa*), PDB entry 3MDE; **HpaB1**, 4-hydroxyphenylacetate monooxygenase (*A. baumannii*), PDB entry 2JBT; **monoox1**, 3-hydroxy-9,10-seco-nandrost-1,3,5(10)-triene-9,17-dione hydroxylase (*Rhodococcus* sp Rha1), PDB entry 2RFQ.

Supplementary Figure 7

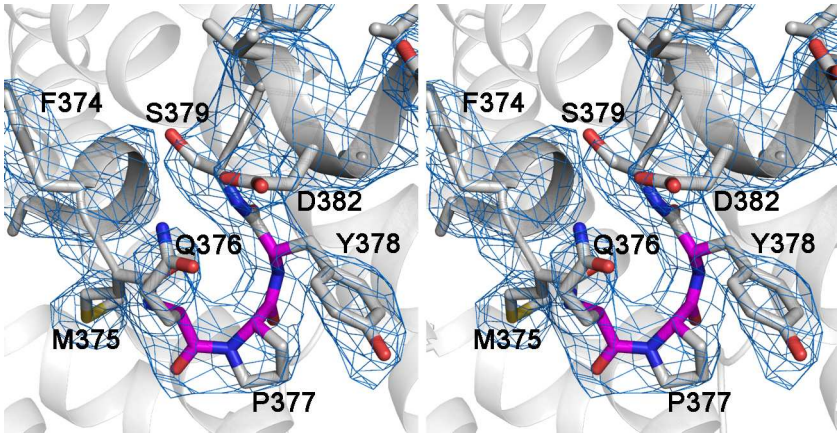

Figure S7. Stereoview of electron density for the tandem cis-peptide in L10 of ORF36. The protein chain is depicted as light grey cartoons. $2|F_o|-|F_c|$ electron density is shown in blue for residues 372-389, which encompasses the two sequential cis-peptides between Gln376-Pro377 and Pro377-Tyr378. Those residues in loop L10 are shown in stick representation, and the main chain carbon atoms of the cis-peptides are highlighted in magenta.

Supplementary Figure 8

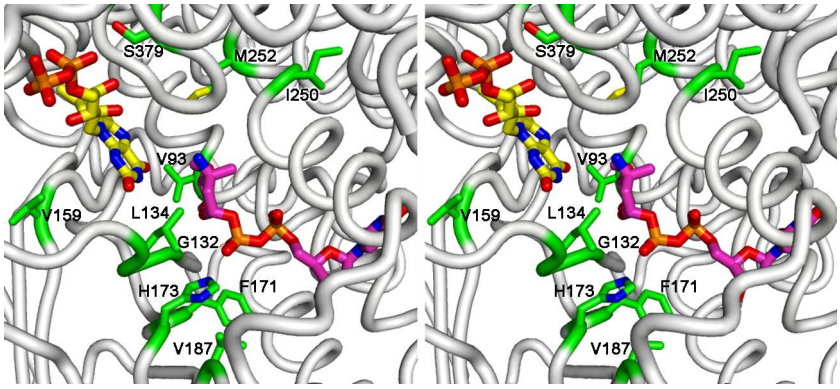

Figure S8. Stereoview of the ORF36 active site, with the protein shown in putty cartoon representation in the same orientation as Figure 7A. Modeled flavin and substrate are shown as thick sticks colored as in Figure 5. Selected residues within 5.5 Å of the flavin cofactor or L-evernosamine that are differ between ORF36 and KijD3 are colored with green carbons, shown in stick representation and labeled according to ORF36 numbering.

Supplementary Table S1

Table S1: Root mean squared deviations of C_α atoms between ORF36, select acyl-CoA dehydrogenases, and class D flavin-containing monooxygenases of known structure. KijD3 is highlighted in blue text, acyl-CoA dehydrogenases and nitroalkane oxidase are in black text, and Class D flavin-containing monooxygenases are in red text. In the table, acyl-CoA dehydrogenase is abbreviated ACAD.

PDB	Protein	RMS deviation (Å)	% identity	Reference
3m9v	KijD3 (*A. kijaniata*)	1.1	64	(*3*)
2jif	short branched-chain ACAD (*H. sapiens*)	2.4	21	(*4*)
1buc	butyryl-CoA dehydrogenase (*M. elsedenii*)	2.6	21	(*5*)
3mde	medium-chain ACAD (*S. scrofa*)	2.8	20	(*6*)
1udy	medium-chain ACAD (*S. scrofa*)	2.8	20	(*7*)
1jqi	rat short-chain ACAD (*R. norvegicus*)	2.6	23	(*8*)
2jbt	4-hydroxyphenylacetate monooxygenase (*A. baumannii*)	2.7	16	(*9*)
2rfq	3-hydroxy-9,10-seco-nandrost-1,3,5(10)-triene-9,17-dione hydroxylase (*Rhodococcus* sp Rha1)	2.6	18	(*10*)
2or0	putative hydroxylase(*Rhodococcus* sp Rha1)	2.8	15	(*11*)
2c0u	nitroalkane oxidase (*F. oxysporum*)	3.0	17	(*12*)
2yyi	4- hydroxyphenylacetate monooxygenase (*T. thermophilus*)	3.5	13	(*13*)
3hwc	chlorophenol-4-monooxygenase (*B. cepacia*)	3.1	13	(*14*)

REFERENCES

1. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007) Clustal W and Clustal X version 2.0, *Bioinformatics (Oxford, England) 23*, 2947-2948.
2. Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999) ESPript: analysis of multiple sequence alignments in PostScript, *Bioinformatics (Oxford, England) 15*, 305-308.
3. Bruender, N. A., Thoden, J. B., and Holden, H. M. (2010) X-ray structure of kijd3, a key enzyme involved in the biosynthesis of D-kijanose, *Biochemistry 49*, 3517-3524.
4. Pike, A. C. W., Hozjan, V., Smee, C., Niesen, F. H., Kavanagh, K. L., Umeano, C., Turnbull, A. P., Von Delft, F., Weigelt, J., Edwards, A., Arrowsmith, C. H., Sundstrom, M., and Oppermann, U. Crystal Structure of Human Short-Branched Chain Acyl-Coa Dehydrogenase., *To be published.*
5. Djordjevic, S., Pace, C. P., Stankovich, M. T., and Kim, J. J. (1995) Three-dimensional structure of butyryl-CoA dehydrogenase from *Megasphaera elsdenii*, *Biochemistry 34*, 2163-2171.
6. Kim, J. J., Wang, M., and Paschke, R. (1993) Crystal structures of medium-chain acyl-CoA dehydrogenase from pig liver mitochondria with and without substrate, *Proc. Natl. Acad. Sci. U. S. A. 90*, 7523-7527.
7. Satoh, A., Nakajima, Y., Miyahara, I., Hirotsu, K., Tanaka, T., Nishina, Y., Shiga, K., Tamaoki, H., Setoyama, C., and Miura, R. (2003) Structure of the transition state analog of medium-chain acyl-CoA dehydrogenase. Crystallographic and molecular orbital studies on the charge-transfer complex of medium-chain acyl-CoA dehydrogenase with 3-thiaoctanoyl-CoA, *J. Biochem. 134*, 297-304.
8. Battaile, K. P., Molin-Case, J., Paschke, R., Wang, M., Bennett, D., Vockley, J., and Kim, J. J. (2002) Crystal structure of rat short chain acyl-CoA dehydrogenase complexed with acetoacetyl-

CoA: comparison with other acyl-CoA dehydrogenases, *J. Biol. Chem. 277*, 12200-12207.

9. Alfieri, A., Fersini, F., Ruangchan, N., Prongjit, M., Chaiyen, P., and Mattevi, A. (2007) Structure of the monooxygenase component of a two-component flavoprotein monooxygenase, *Proc. Natl. Acad. Sci. U. S. A. 104*, 1177-1182.

10. Chang, C., Skarina, T., Kagan, O., Savchenko, A., Edwards, A. M., and Joachimiak, A. Crystal structure of 3-HSA hydroxylase, oxygenase from *Rhodococcus* sp. RHA1., *To be published.*

11. Tan, K., Skarina, T., Kagen, O., Savchenko, A., Edwards, A., and Joachimiak, A. The crystal structure of a putative hydroxylase from *Rhodococcus* sp. RHA1., *To be published.*

12. Nagpal, A., Valley, M. P., Fitzpatrick, P. F., and Orville, A. M. (2006) Crystal structures of nitroalkane oxidase: insights into the reaction mechanism from a covalent complex of the flavoenzyme trapped during turnover, *Biochemistry 45*, 1138-1150.

13. Kim, S. H., Hisano, T., Takeda, K., Iwasaki, W., Ebihara, A., and Miki, K. (2007) Crystal structure of the oxygenase component (HpaB) of the 4-hydroxyphenylacetate 3-monooxygenase from *Thermus thermophilus* HB8, *J. Biol. Chem. 282*, 33107-33117.

14. Webb, B. N., Ballinger, J. W., Kim, E., Belchik, S. M., Lam, K. S., Youn, B., Nissen, M. S., Xun, L., and Kang, C. (2010) Characterization of chlorophenol 4-monooxygenase (TftD) and NADH:FAD oxidoreductase (TftC) of *Burkholderia cepacia* AC1100, *J. Biol. Chem. 285*, 2014-2027.