interaction with appropriate ACP-linked substrates stimulates a relatively modest level of activity. Control of thioesterase activity in this manner would help to limit improper hydrolysis of metabolically important acyl-CoA or acyl-ACP substrates by an otherwise promiscuous enzyme.

**Works Cited**

1. Harvey, B. M., Hong, H., Jones, M. A., Hughes-Thomas, Z. A., Goss, R. M., Heathcote, M. L., Bolanos-Garcia, V. M., Kroutil, W., Staunton, J., Leadlay, P. F., and Spencer, J. B. (2006) Evidence that a novel thioesterase is responsible for polyketide chain release during biosynthesis of the polyether ionophore monensin. *Chembiochem* **7**, 1435-1442

2. Chakravarty, B., Gu, Z., Chirala, S. S., Wakil, S. J., and Quiocho, F. A. (2004) Human fatty acid synthase: structure and substrate selectivity of the thioesterase domain. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 15567-15572

3. Roongsawang, N., Washio, K., and Morikawa, M. (2007) In vivo characterization of tandem C-terminal thioesterase domains in arthrofactin synthetase. *ChemBioChem* **8**, 501-512

4. Scholz-Schroeder, B. K., Soule, J. D., and Gross, D. C. (2003) The sypA, sypB, and sypC Synthetase Genes Encode Twenty-Two Modules Involved in the Nonribosomal Peptide Synthesis of Syringopeptin by Pseudomonas syringae pv. syringae B301D. *Molecular Plant-Microbe Interactions* **16**, 271-280

5. Libertini, L. J., and Smith, S. (1978) Purification and properties of a thioesterase from lactating rat mammary gland which modifies the product specificity of fatty acid synthetase. *J Biol Chem* **253**, 1393-1401

6. Tsai, S. C., Lu, H., Cane, D. E., Khosla, C., and Stroud, R. M. (2002) Insights into Channel Architecture and Substrate Specificity from Crystal Structures of Two Macrocycle-Forming Thioesterases of Modular Polyketide Synthases. *Biochemistry* **41**, 12598-12606

7. Tsai, S.-C., Miercke, L. J. W., Krucinski, J., Gokhale, R., Chen, J. C. H., Foster, P. G., Cane, D. E., Khosla, C., and Stroud, R. M. (2001) Crystal structure of the macrocycle-forming thioesterase domain of the erythromycin polyketide synthase: Versatility from a unique substrate channel. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 14808-14813

8. Bruner, S. D., Weber, T., Kohli, R. M., Schwarzer, D., Marahiel, M. A., Walsh, C. T., and Stubbs, M. T. (2002) Structural basis for the cyclization of the lipopeptide antibiotic surfactin by the thioesterase domain SrfTE. *Structure* **10**, 301-310

9. Samel, S. A., Wagner, B., Marahiel, M. A., and Essen, L.-O. (2006) The Thioesterase Domain of the Fengycin Biosynthesis Cluster: A Structural Base for the Macrocyclization of a Non-ribosomal Lipopeptide. *J. Mol. Biol.* **359**, 876-889

10. Frueh, D. P., Arthanari, H., Koglin, A., Vosburg, D. A., Bennett, A. E., Walsh, C. T., and Wagner, G. (2008) Dynamic thiolation-thioesterase structure of a non-ribosomal peptide synthetase. *Nature* **454**, 903-906

11. Koglin, A., Lohr, F., Bernhard, F., Rogov, V. V., Frueh, D. P., Strieter, E. R., Mofid, M. R., Guntert, P., Wagner, G., Walsh, C. T., Marahiel, M. A., and Dotsch, V. (2008) Structural basis for the selectivity of the external thioesterase of the surfactin synthetase. *Nature* **454**, 907-911

12. Yeh, E., Kohli, R. M., Bruner, S. D., and Walsh, C. T. (2004) Type II thioesterase restores activity of a NRPS module stalled with an aminoacyl-S-enzyme that cannot be elongated. *ChemBioChem* **5**, 1290-1293

13. Jiralerspong, S., Rangaswamy, V., Bender, C. L., and Parry, R. J. (2001) Analysis of the enzymatic domains in the modular portion of the coronafacic acid polyketide synthase. *Gene* **270**, 191-200

14. Armougom, F., Moretti, S., Poirot, O., Audic, S., Dumas, P., Schaeli, B., Keduas, V., and Notredame, C. (2006) Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. *Nucl. Acids Res.* **34**, W604-608

15. Clamp, M., Cuff, J., Searle, S. M., and Barton, G. J. (2004) The Jalview Java alignment editor. *Bioinformatics* **20**, 426-427

16. Geoffroy, V. A., Fetherston, J. D., and Perry, R. D. (2000) Yersinia pestis YbtU and YbtT are involved in synthesis of the siderophore yersiniabactin but have different effects on regulation. *Infect Immun* **68**, 4452-4461

17. Xue, Y., Zhao, L., Liu, H. W., and Sherman, D. H. (1998) A gene cluster for macrolide antibiotic biosynthesis in Streptomyces venezuelae: architecture of metabolic diversity. *Proc Natl Acad Sci U S A* **95**, 12111-12116

18. Butler, A. R., Bate, N., and Cundliffe, E. (1999) Impact of thioesterase activity on tylosin biosynthesis in Streptomyces fradiae. *Chemistry & Biology* **6**, 287-292

19. Schneider, A., and Marahiel, M. A. (1998) Genetic evidence for a role of thioesterase domains, integrated in or associated with peptide synthetases, in non-ribosomal peptide biosynthesis in Bacillus subtilis. *Arch Microbiol* **169**, 404-410

20. Yu, F.-M., Qiao, B., Zhu, F., Wu, J.-C., and Yuan, Y.-J. (2006) Functional Analysis of type II Thioesterase of *Streptomyces lydicus* AS 4.2501. *Appl. Biochem. Biotechnol.* **135**, 145-158

21. Reimmann, C., Patel, H. M., Walsh, C. T., and Haas, D. (2004) PchC thioesterase optimizes nonribosomal biosynthesis of the peptide siderophore pyochelin in Pseudomonas aeruginosa. *J Bacteriol* **186**, 6367-6373

22. Hu, Z., Pfeifer, B. A., Chao, E., Murli, S., Kealey, J., Carney, J. R., Ashley, G., Khosla, C., and Hutchinson, C. R. (2003) A specific role of the Saccharopolyspora erythraea thioesterase II gene in the function of modular polyketide synthases. *Microbiology* **149**, 2213-2225

23. Chen, H., and Walsh, C. T. (2001) Coumarin formation in novobiocin biosynthesis: beta-hydroxylation of the aminoacyl enzyme tyrosyl-S-NovH by a cytochrome P450 NovI. *Chem. Biol.* **8**, 301-312

24. Schwarzer, D., Mootz, H. D., Linne, U., and Marahiel, M. A. (2002) Regeneration of misprimed nonribosomal peptide synthetases by type II thioesterases. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 14083-14088

25. Schwarzer, D., Mootz, H. D., Linne, U., and Marahiel, M. A. (2002) Regeneration of misprimed nonribosomal peptide synthetases by type II thioesterases. *Proc Natl Acad Sci U S A* **99**, 14083-14088

26. Kotowska, M., Pawlik, K., Butler, A. R., Cundliffe, E., Takano, E., and Kuczek, K. (2002) Type II thioesterase from Streptomyces coelicolor A3(2). *Microbiology* **148**, 1777-1783

27. Heathcote, M. L., Staunton, J., and Leadlay, P. F. (2001) Role of type II thioesterases: evidence for removal of short acyl chains produced by aberrant decarboxylation of chain extender units. *Chemistry & Biology* **8**, 207-220

28. Kim, B. S., Cropp, T. A., Beck, B. J., Sherman, D. H., and Reynolds, K. A. (2002) Biochemical Evidence for an Editing Role of Thioesterase II in the Biosynthesis of the Polyketide Pikromycin. *J. Biol. Chem.* **277**, 48028-48034

29. Tang, Y., Koppisch, A. T., and Khosla, C. (2004) The acyltransferase homologue from the initiation module of the R1128 polyketide synthase is an acyl-ACP thioesterase that edits acetyl primer units. *Biochemistry* **43**, 9546-9555

30. Witkowski, A., Naggert, J., Witkowska, H. E., Randhawa, Z. I., and Smith, S. (1992) Utilization of an active serine 101----cysteine mutant to demonstrate the proximity of the catalytic serine 101 and histidine 237 residues in thioesterase II. *Journal of Biological Chemistry* **267**, 18488-18492

31. Claxton, H. B., Akey, D. L., Silver, M. K., Admiraal, S. J., and Smith, J. L. (2009) Structure and functional analysis of RifR, the type II thioesterase from the rifamycin biosynthetic pathway. *J Biol Chem* **284**, 5021-5029

32. Akey, D. L., Kittendorf, J. D., Giraldes, J. W., Fecik, R. A., Sherman, D. H., and Smith, J. L. (2006) Structural Basis for Macrolactonization by the Pikromycin Thioesterase. *Nature Chem. Biol.* **2**, 537-542

33. Linne, U., Schwarzer, D., Schroeder, G. N., and Marahiel, M. A. (2004) Mutational analysis of a type II thioesterase associated with nonribosomal peptide synthesis. *Eur. J. Biochem.* **271**, 1536-1545

34. Chen, H., Hubbard, B. K., O'Connor, S. E., and Walsh, C. T. (2002) Formation of beta-hydroxy histidine in the biosynthesis of nikkomycin antibiotics. *Chem. Biol.* **9**, 103-112

35. Fujimori, D. G., Hrvatin, S., Neumann, C. S., Strieker, M., Marahiel, M. A., and Walsh, C. T. (2007) Cloning and characterization of the biosynthetic gene cluster for kutznerides. *Proceedings of the National Academy of Sciences* **104**, 16498-16503

36. Patel, J., Hoyt, J. C., and Parry, R. J. (1998) Investigations of Coronatine Biosynthesis. Overexpression and Assay of CmaT, a Thioesterase Involved in Coronamic Acid Biosynthesis. *Tetrahedron* **54**, 15927-15936

37. Sambrook, J., Fritsch, E. F., and and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY., Cold Spring Harbor, NY.

38. Hoover, D. M., and Lubkowski, J. (2002) DNAWorks: an automated method for designing oligonucleotides for PCR-based gene synthesis. *Nucleic Acids Res.* **30**, e43

39. Gokhale, R. S., Tsuji, S. Y., Cane, D. E., and Khosla, C. (1999) Dissecting and exploiting intermodular communication in polyketide synthases. *Science* **284**, 482-485

40. Tsuji, S. Y., Cane, D. E., and Khosla, C. (2001) Selective protein-protein interactions direct channeling of intermediates between polyketide synthase modules. *Biochemistry* **40**, 2326-2331

41. Walker, J. M. (2005) *The Proteomics Protocols Handbook*, Humana Press, Totowa, NJ.

42. Guerrero, S. A., Hecht, H. J., Hofmann, B., Biebl, H., and Singh, M. (2001) Production of selenomethionine-labelled proteins using simplified culture

conditions and generally applicable host/vector systems. *Appl. Microbiol. Biotechnol.* **56**, 718-723

43.    Deutsch, J., Rapoport, S. I., and Rosenberger, T. A. (2002) Coenzyme A and short-chain acyl-CoA species in control and ischemic rat brain. *Neurochem Res* **27**, 1577-1582

44.    Lambalot, R. H., Gehring, A. M., Flugel, R. S., Zuber, P., LaCelle, M., Marahiel, M. A., Reid, R., Khosla, C., and Walsh, C. T. (1996) A new enzyme superfamily - the phosphopantetheinyl transferases. *Chem Biol* **3**, 923-936

45.    Quadri, L. E., Weinreb, P. H., Lei, M., Nakano, M. M., Zuber, P., and Walsh, C. T. (1998) Characterization of Sfp, a Bacillus subtilis phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases. *Biochemistry* **37**, 1585-1595

46.    Otwinowski, Z., and Minor, V. (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode. in *Methods in Enzymology*, Academic Press New York, NY

47.    Terwilliger, T. C., and Berendzen, J. (1999) Automated MAD and MIR structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **55 ( Pt 4)**, 849-861

48.    Terwilliger, T. C. (2003) Automated main-chain model building by template matching and iterative fragment extension. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **59**, 38-44

49.    Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 2126-2132

50.    Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **53**, 240-255

51.    Painter, J., and Merritt, E. (2006) TLSMD Webserver for the Generation of Multi-group TLS Models. *J. Appl. Cryst.* **39**, 109-111

52.    Navaza, J. (1994) AMoRe: an automated package for molecular replacement. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **50**, 157-163

53.    Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall, W. B., 3rd, Snoeyink, J., Richardson, J. S., and Richardson, D. C. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35**, W375-383

54.    Frishman, D., and Argos, P. (1995) Knowledge-based protein secondary structure assignment. *Proteins* **23**, 566-579

55.    Heinig, M., and Frishman, D. (2004) STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. *Nucleic Acids Res.* **32**, W500-502

# Chapter 3

# Ketoreductase Domains

Ketoreductase (KR) domains incorporated into PKS modules use NADPH to reduce the β-carbonyl of the growing intermediate to an alcohol (Fig 3.1).   This reduction reaction is stereospecific at the C-3 and the C-2 of the polyketide.  Crystal structures of KR domains and substrate modeling into the active site have provided clues as to how conserved motifs and residues may proctor stereospecificity, however, no structure has been solved with a native substrate in the active site.  The position of a substrate in the active site and the residues that affect stereospecificity remain unknown. Determining these residues and the placement of substrates in the active site are crucial for PKS reprogramming efforts.
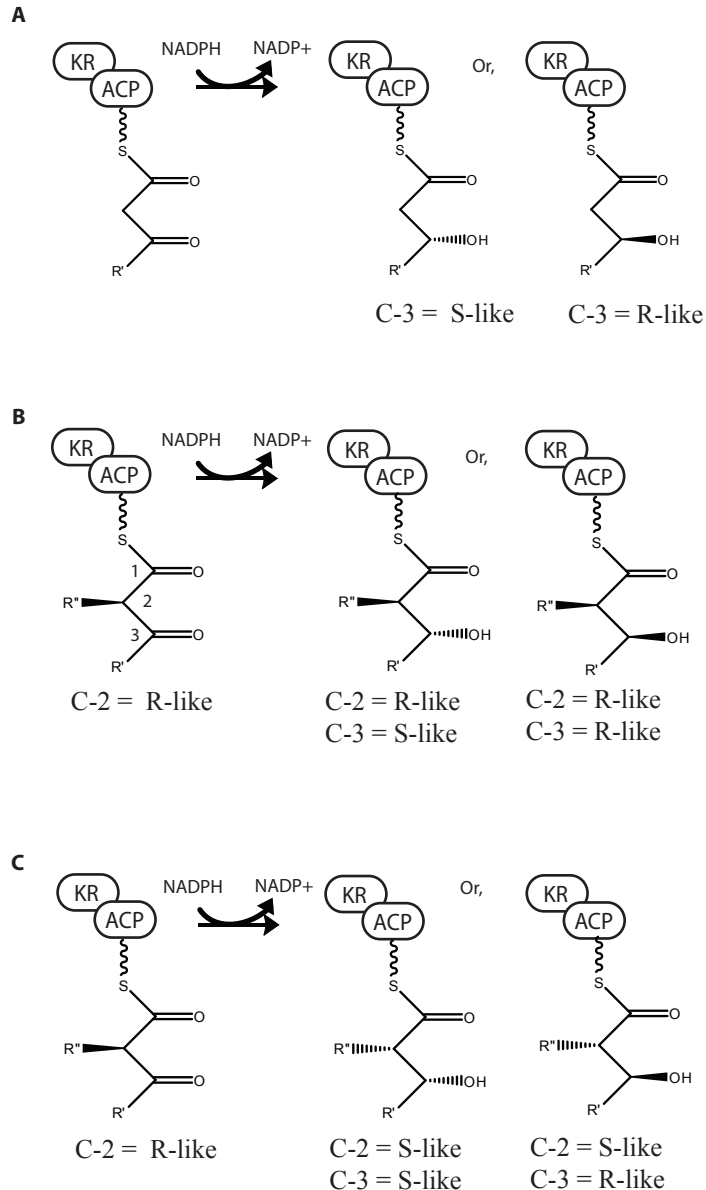
**A**

KR ACP  NADPH  NADP+  KR ACP  Or,  KR ACP

C-3 = S-like        C-3 = R-like

**B**

KR ACP  NADPH  NADP+  KR ACP  Or,  KR ACP

C-2 = R-like        C-2 = R-like     C-2 = R-like
                    C-3 = S-like     C-3 = R-like

**C**

KR ACP  NADPH  NADP+  KR ACP  Or,  KR ACP

C-2 = R-like        C-2 = S-like     C-2 = S-like
                    C-3 = S-like     C-3 = R-like

**Figure 3.1 The ketoreductase domain uses NADPH to reduce the β-carbonyl to an alcohol.** **A**. For ease of communication in this field, the 'R-like' or 'S-like' configuration of C-3 is determined by presuming the C-2 has higher priority than C-4, unless otherwise noted (1-4). **B**. For ease of communication in this field, the R-like or S-like configuration of C-2 is determined by presuming the C-2 substituent has the lowest priority after the C-2 hydrogen, unless otherwise noted. **C**. Some authors have suggested the KR domain may be able to epimerize the C-2 substituent either before or during C-3 reduction (3,5-7).

Polyketide ketoreductases are a part of the short-chain dehydrogenase/reductase family (SDR) and consist of two subdomains, each composed of a Rossmann fold. The catalytic subdomain contains the catalytic tyrosine, an invariant serine, and an NADPH binding site. The structural subdomain is thought to stabilize the catalytic subdomain and contains few conserved residues (3). In PKS modules that contain an enoyl reductase (ER), the ER domain is located between the structural and catalytic subdomains of the KR.

Type I PKS KR domains have been classified into two types— an A-type that reduces the substrate to a hydroxyl product in the S-like configuration, and a B-type that reduces the substrate to a hydroxyl product in the R-like configuration (8). Though *in vitro* experiments with non-natural substrates sometimes lead to a mix of S-like and R-like products (4,9), it is assumed that the S-like configuration is produced exclusively by the A-type KR and the R-like product exclusively by the B-type KR domain *in vivo*. The A-type and B-type KR are classified based on two sequence motifs. B-type KR domains contain a Leu-Asp-Asp sequence in the "LDD" motif, while A-type KR domains have a conserved tryptophan in a separate "W" motif (8).

The acyl transferase domain selectively chooses acyl units that contain the C-2 carbon substituent in the S-like configuration (10). The ketosynthase domain epimerizes the C-2 substituent during condensation to the R-like configuration. C-2 stereochemistry may arise from the KR domain selectively choosing substrates with the appropriate C-2 stereochemistry. Reduction at C-3 then locks in the stereochemistry at C-2 (Fig 3.1B). Alternatively, some authors have suggested the KR domain may be able to catalyze epimerization of C-2 before, or during the reduction of C-3. (Fig. 3.1C) (3,5-7).

Crystal structures of six ketoreductases from polyketide synthases or fatty acid synthases are available. Three crystal structures of type II PKS KR domains are available, including structures of the actinorhodin KR with and without the inhibitor emodin (PDB 2RH4, 1X7G) (11,12), and one crystal structure of the KR domain of the R1128 PKS, found in *S. coelicolor* A3(2) (PDB 2NM0) (13). The crystal structure of the tropinone KR bound with NADPH (PDB 1IPF) (14) has been used to model the

interactions of NADPH in the binding site (1). The structure of the bacterial FAS KR, FabG (PDB 1Q7B) (15), is also available.

Crystal structures of three type I PKS KR domains are available. Two structures are available of the KR of the first module of the erythromycin PKS, EryKR1. Both EryKR1 crystal structures show NADPH bound in the NADPH binding site (PDB 2FR0, 2FR1) (3). One crystal structure is available of the KR of the first module of the tylosin PKS, TylKR1 (PDB 2Z5L) (1). Both EryKR1 and TylKR1 are B-type KR domains. Three crystal structures of the KR of the second module of the amphotericin PKS, an A-type KR domain, are available. These include a free enzyme structure, an $NADP^+$-bound structure, and an $NADP^+$-bound structure with the substrate mimic malic acid in the active site (AmpKR2, PDB 3MJC, 3MJE, 3MJS) (16).

The lid structure found in SDR enzymes comprises a helix and loop that forms part of the substrate binding site (Fig 3.2) (1). The NADPH binding sites of the EryKR1 structures, the TylKR1 structure, and the AmpKR2 structures are well ordered, with the exception of a lid helix and a phosphate binding P-loop. Though well ordered in the TylKR1 free enzyme structure, the lid helix and P-loop of EryKR1 are invisible in one crystal structure, and have higher temperature factors in the other (3). Similarly, the N-terminus of the lid helix of AmpKR2 is invisible in the $NADP^+$-bound structure, but ordered in the AmpKR2 free enzyme structure (16). The lid helix and P-loop of the cofactor-bound EryKR1 were found to be closer to the active site when compared to the same motifs of the TylKR1 free enzyme structure (1,3). In contrast, the lid helix of AmpKR2 was shifted away from the active site when both $NADP^+$ and malic acid were present (16). This suggests that $NADP^+$ binding promotes a conformation change that makes the lid helix more flexible (16). The active site residues, a tyrosine, a serine, and a lysine are located in a groove proximal to the location of the nicotinamide ring of NADPH.

**Figure 3.2 Structures of an A-type and B-type ketoreductase.**

**A**. Structure of the B-type ketoreductase from the first module of the erythromycin PKS (PDB 2FR0) (3). The catalytic subdomain is shown in green, the structural subdomain in blue, the catalytic tyrosine and serine are in cyan, the lid helix is in beige, and the 'LDD' motif is shown in yellow. The bridge that connects the structural and catalytic domains is shown in red, and the P-loop of the NADPH binding site is shown in purple. **B**. The structure of the A-type ketoreductase from the second module of the amphotericin PKS (PDB 3MJS) (16). The color designations are the same as above except the A-type ketoreductase contains a 'W' motif shown in yellow instead of an 'LDD' motif.
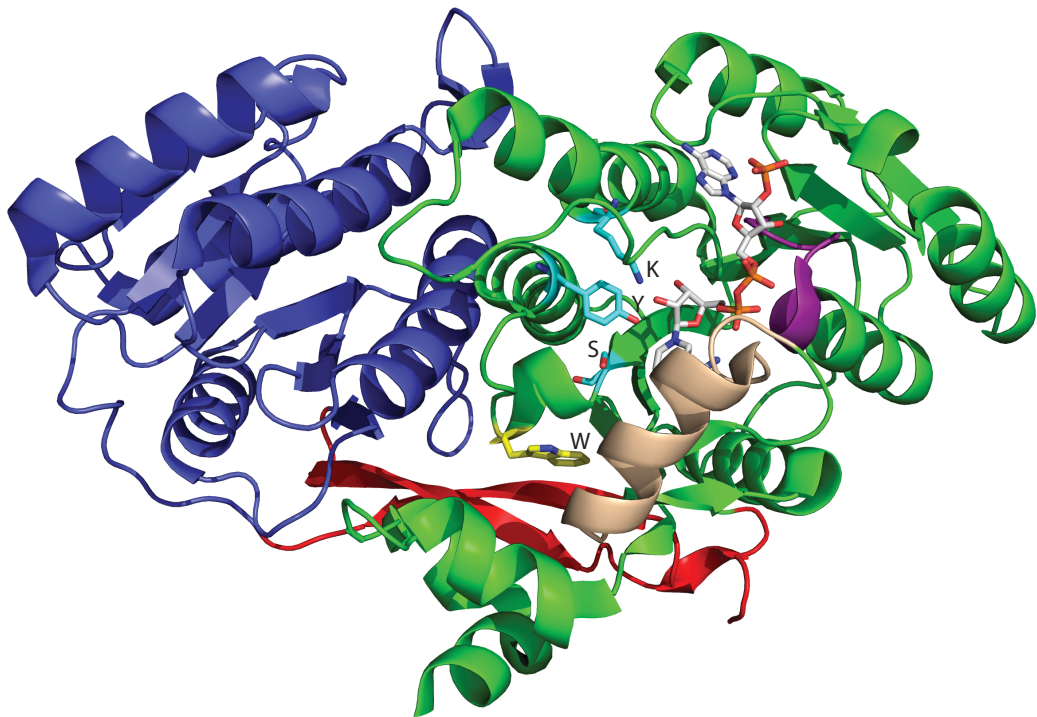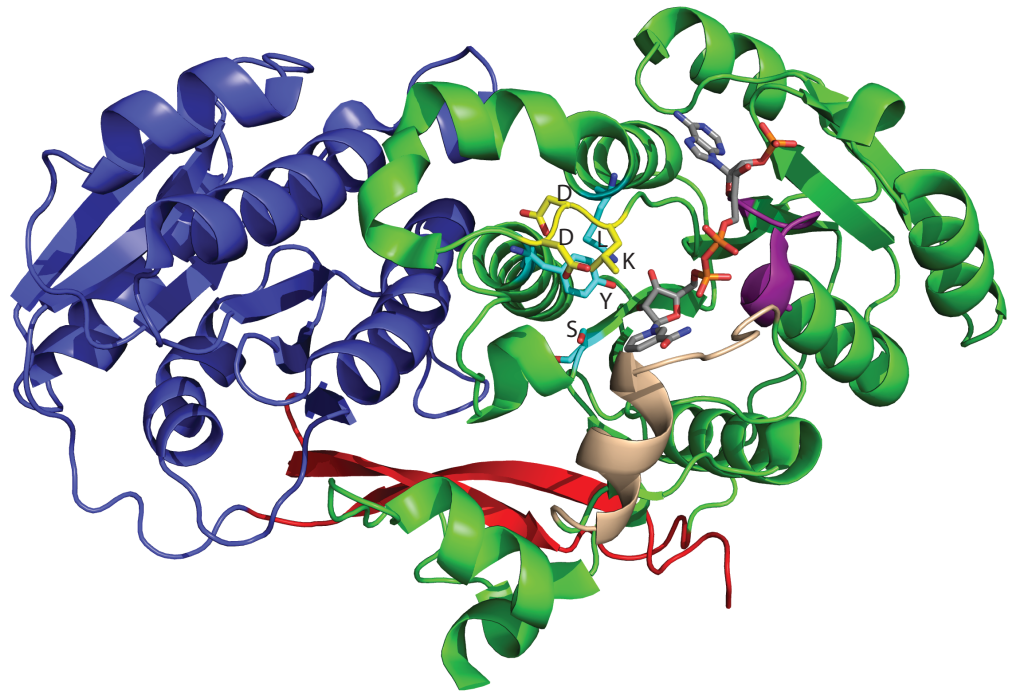
**Figure 3.2 Structures of an A-type and B-type ketoreductase.**

The stereospecificity of KR domains vary from PKS to PKS and even from module to module. Excised KR domains as well as modular KR domains have been tested with a variety of substrates, including natural acyl-ACP substrates and non-natural acyl-CoA or acyl-SNAC substrates.

Replacement of a KR domain in a module with another from a different type or pathway results in an equivalent exchange of acyl-ACP stereoproducts. For example the presence of a substituent at the C-2 position did not affect the ability of the rapamycin KR2, when it was substituted for EryKR2, to reduce the triketide of the DEBS pathway (17). Other modular KR domains are highly specific and will reduce only substrates with the correct C-2 stereochemistry when presented as an acyl-ACP as in the case of DEBS KR2 (18).

When excised from the module, EryKR1 and TylKR1 yield only one stereoproduct when presented with a racemic mixture of acyl-SNACs (4). However, the excised KR domains, EryKR2, EryKR5, and EryKR6, produce a mixture of stereoproducts from an acyl-SNAC substrate (4). The EryKR5 and EryKR6 domains, which act on longer substrates, appeared to show more tolerance to SNACs with differing C-2 stereoconfigurations than the excised domains, EryKR1, EryKR2 and EryKR3 (4,9).

To investigate the stereospecificity of ketoreductase domains by trapping substrate or product complexes for crystallization, nine didomain constructs were created as described in Table 3.1. These didomain constructs comprise a KR domain with the respective ACP domain from modules of the tylosin, pikromycin, and erythromycin biosynthetic pathways. These constructs were chosen based on the variety of natural substrates, and the lack of an ER domain located between the catalytic and structural subunits of the KR didomain.

Substrate mimics, referred to as Fecik labels, were synthesized and provided by our collaborators in the Robert Fecik lab of the University of Minnesota (19). (Fig 3.3) To create these Fecik labels, polyketide intermediates from the tylosin, pikromycin and erythromycin pathways were synthesized and attached to the Ppant arm of CoA through a thioether bond. Attachment through a thioether bond was preferred over attachment

through the less stable thioester bond, as the thioether bond should increase the stability of the Fecik label under protein crystallization conditions.

The Ppant arm and polyketide intermediate of the Fecik labels were transferred to the ACP of the didomain via a phosphopantetheinyl transferase. Competitive assays were performed with the loaded didomains and the non-natural substrates *trans*-1-decalone and cyclohexanone. Fecik labels that were able to compete with *trans*-1-decalone or cyclohexanone were presumed to do so by entering and occluding the active-site.

Unfortunately, efforts to crystallize the tylosin KR1ACP1 didomain with the EKL-103 substrate mimic did not yield diffracting crystals. However, kinetic investigations into these didomains showed that all didomains retained KR activity in the KR-ACP construct. Specific Fecik labels tethered to the ACP domain of the KR-ACP constructs were shown to compete with free standing unnatural substrates, providing insights into KR didomain substrate specificity.

Table 3.1 Didomain product stereochemistry.

| | β-hydroxyl stereochemistry | α-substituent stereochemistry | α-substituent Anti/Syn to β-hydoxyl |
|---|---|---|---|
| DEBSI KR1ACP1 | R | S | Syn |
| DEBSI KR2ACP2 | S | R | Syn |
| PikAI KR1ACP1 | R | S | Syn |
| PikAI KR2ACP2 | ?* | None | |
| TylG1 KR1ACP1 | R | R | Anti |
| TylG1 KR2ACP2 | ?* | ?* | ?* |
| TylG1 KR3ACP3 | ?* | None | |
| DEBS III KR5ACP5 | S | R | Syn |
| DEBS III KR6ACP6 | S | R | Syn |
| PikAIII KR5ACP5 | S | R | Syn |
| TylGIV KR6ACP6 | S | R | Syn |
| TylGIV KR7ACP7 | S-like** | None | |

*Stereochemistry cannot be determined due to the removal of the alcohol functional group by the DH domain in the module.

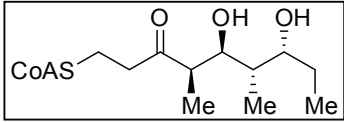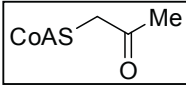** For ease of communication in this field, C-2 is always presumed to have a higher priority than C-4.
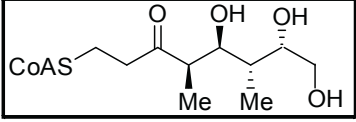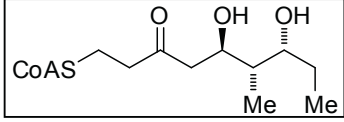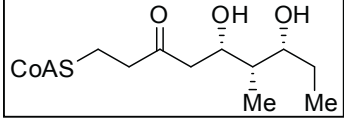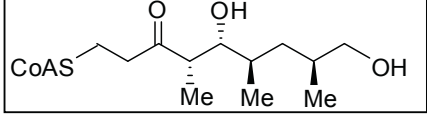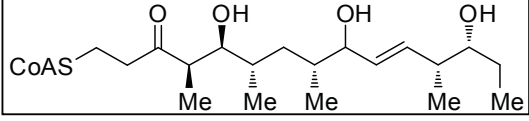
| Label | Name |
|-------|------|
|  | D-KR2-P1 |
|  | EKL-103 |
|  | D-KR2-P1 |
|  | P-KR2-P1 |
|  | P-KR2-P2 |
|  | P-KR5-P1 |
|  | P-KR5-P3 |

**Figure 3.3 Fecik labels use to investigate KR-ACP stereospecificity.**
With the exception of EKL-103, the name represents the KR domain product that the Fecik label is meant to mimic. (P=pikromycin, D=erythromycin)

**Methods and Materials**

*Expression and Purification of Proteins.* Expression plasmids for each KR-ACP didomain as described in Table 3.2 were created by Jamie Razelun. For each didomain, *E. coli* strain BL21 Star (DE3) (Invitrogen) was transformed with the appropriate expression plasmid. 500-mL cultures were grown at 37°C in 2-L flasks containing Terrific Broth with 4% glycerol, trace metal components (20) according to Table 3.3, and 1 mg/L ampicillin. The trace metal solution was added to the 500-mL cultures to stop the post-translational addition of a Ppant arm to the ACP of the KR-ACP didomain. KR-ACP didomains that do not contain the Ppant arm are referred to as 'apo'. KR-ACP didomains that contain the Ppant arm are referred to as 'holo'. Expression was induced with 100 µM IPTG at an optical density at 600 nm of 1.0. After induction, incubation was continued for 20 h at 20°C. The cells were then harvested by centrifugation at 5,670g. The cell pellet was frozen and stored at -20°C.

Thawed cell pellets were resuspended in 40 mL of 300 mM NaCl, 50 mM Tris pH 7.5, and 10% glycerol. The resuspended cells were disrupted by sonnication and the lysate was collected by centrifugation at 26,890g and loaded onto a Histrap HP column (5 mL, GE HealthCare) previously equilibrated with 300 mM NaCl, 50 mM Tris pH 7.5, and 10% glycerol. The column was washed with 15 mL of 20 mM imidazole with 300 mM NaCl, 50 mM Tris 7.5 and 10% glycerol. Proteins were eluted with an imidazole gradient (20 mM-400 mM) in the same solution.

Pooled fractions containing the didomain were concentrated to ~ 2 mg/mL and loaded onto a HiLoad 16/60 Superdex 200 column previously equilibrated with 300 mM NaCl, 50 mM Tris pH 7.5, 2 mM DTT, and 10% glycerol. Proteins were eluted with 300 mM NaCl, 50 mM Tris pH 7.5, 2 mM DTT and 10% glycerol. Pooled fractions containing the didomain were concentrated with an Amicon Ultra-15 centrifugal filter unit (Millipore) to concentrations of 3-10 mg/mL.

Purified proteins were flash-frozen in liquid nitrogen and stored at -80°C. Protein concentrations were determined from absorbance at 280 nm using calculated extinction coefficients (21) (Table 3.2).
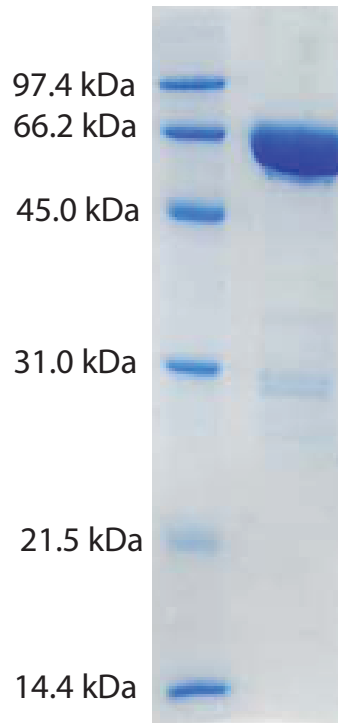
**Figure 3.4  Production of Tyl KR1ACP1**
TylKR1ACP1 is shown at the expected molecular weight (~ 65 kDa) after purification with a HiLoad 16/60 S200 Superdex column.

Table 3.2 KR-ACP didomain plasmids

| Target | MW kDa | Plasmid | Extinction Coefficients at 280 nm (21) ($M^{-1}$ $cm^{-1}$) |
|---|---|---|---|
| Pik KR1ACP1 | 60.0 | pJRR130 | 73450 |
| Pik KR2ACP2 | 61.6 | pJRR156 | 47440 |
| Pik KR5ACP5 | 67.4 | pJRR150 | 89950 |
| DEBS KR2ACP2 | 68.0 | pJRR209 | 82960 |
| DEBS KR5ACP5 | 62.1 | pJRR017 | 62450 |
| DEBS KR6ACP6 | 58.4 | pJRR036 | 56950 |
| Tyl KR1ACP1 | 65.2 | pJRR023 | 62450 |
| Tyl KR2ACP2 | 62.7 | pJRR200 | 41940 |
| Tyl KR3ACP3 | 59.2 | pJRR026 | 47440 |
| Tyl KR6ACP6 | 61.7 | pJRR157 | 73450 |
| Tyl KR7ACP7 | 62.9 | pJRR161 | 81930 |

Table 3.3 Trace metal solution (20)

| Metal | Conc. ($\mu M$) |
|---|---|
| Fe | 50 |
| Ca | 20 |
| Mn | 10 |
| Zn | 10 |
| Cu | 2 |
| Ni | 2 |
| Mo | 2 |
| Se | 2 |
| $H_3BO_3$ | 2 |

*Sfp Didomain Labeling Reaction.* The Tyl KR1ACP1 didomain (30 $\mu M$) was labeled with free CoA (0-100 $\mu M$), utilizing the nonspecific phosphopantetheinyl transferase Sfp (8.5 $\mu M$ – 10.6 $\mu M$) (22) in a reaction that consisted of 100 mM Tris-HCL pH 8.0, 10 mM $MgCl_2$, 5 mM DTT, and 6% glycerol. Free CoA concentrations were varied to determine optimal labeling concentrations. The reaction proceeded for 30 min at room temperature and was quenched by adding an equivalent volume of 500 mM EDTA to the reaction. The reaction was diluted to a KR-ACP concentration of 0.1 mg/mL and a 5 $\mu L$ injection of each sample was analyzed by LC-MS with a Shimadzu LCMS-2010EV after separation on a PLRP-S 22 × 50 mm 4000 Å, 8 $\mu m$ polymeric RP-HPLC column

(Varian, Palo Alto, CA) heated to 50°C.  Samples were desalted online for 5 min with 95% buffer A (98.9% water, 1% acetonitrile, 0.1% formic acid) and 5% buffer B (1% water, 98.9% acetonitrile, and 0.1% formic acid), followed by gradient elution to 100% Buffer B over 20 min. Profile mode data was gathered from m/z 400–2000 utilizing electrospray ionization in positive ion mode. (23)

*Svp Didomain Labeling Reaction:* The KR-ACP didomains (30 µM) were labeled with either a Fecik label or free CoA, utilizing the phosphopantetheinyl transferase Svp (24). Control reactions consisted of apo didomain in which no label was added to the reaction mix.  The labeling reaction consisted of 100 mM Tris-HCL pH 8.0, 10 mM $MgCl_2$, 6% glycerol, Fecik label (35-43 µM ) or CoA (35-47 µM), and didomain (18 µM-22 µM ).  The ratio between Fecik label or CoA and didomain was kept above 1.75.  The reaction proceeded for 2 hrs at room temperature.

        To determine labeling efficiency, fluorescent CoA 547 (66 µM) from New England Biolabs was added to the labeling reaction, which continued for another 1.5-2 hrs.  The reaction was then quenched with EDTA (200 µM), diluted, and analyzed by HPLC using a PLRP-S  22 × 50 mm 4000 Å, 8 µm polymeric RP-HPLC column (Varian, Palo Alto, CA).   A linear gradient of buffer A (water, 0.1% triflouroacetic acid) and buffer B (acetonitrile, 0.1% triflouroacetic acid) at a flow rate of 0.33 mL/min was used to separate Svp and didomain.  Initial conditions consisted of  5% buffer B for 2 mins, then a linear gradient over 12 mins to 60% buffer B.  Buffer B was then set to 100% and held constant for 5 mins and finally brought back down to 5% over 4 mins.  The column was then allowed to equilibrate for 9 mins at 5% buffer B.  To control for any difference in injection volume, a ratio was taken between the integrated peak at 550 nm, where the the fluorophore absorbs, and the integrated peak at 220 nm, where the didomain absorbs.  To determine the percentage of didomain that was unlabeled in the reaction, the ratio of fluorophore 547: Fecik-labeled didomain was divided by the fluorophore547:didomain ratio of a parallel sample containing unlabeled didomain with apo-ACP.

*KR Activity Assay:*  *Trans*-1-decalone  and NADPH were obtained from Sigma Aldrich. *Trans*-1-decalone, which is not miscible with water, was dissolved in 3% DMSO.  The didomain (labeled with free CoA or an apo control) underwent pre-incubation with

NADPH for 20 minutes at 30°C to ensure that the NADPH binding sites were saturated upon addition of substrate.  The pre-incubation mix contained: 31 mM HEPES pH 7.0, 1.6 mM NADPH, 2.5 µM didomain.   After the pre-incubation, 0 – 50 mM *trans*-1-decalone was added, creating the reaction mix composed of 25 mM HEPES pH 7.0, 1.3 mM NADPH, 0-50 mM decalone, 0.2% DMSO and 2-20 µM didomain.  The absorabance at 340 nm was measured for 55 minutes at 30°C. A negative control composed of 20 µM TylKR1ACP1 didomain, 25 mM HEPES pH 7.0, 50 mM *trans*-1-decalone, and 0.2% DMSO showed these components did not affect absorbance at 340 nm.

Alternatively, *trans*-1-decalone was dissolved in 100% DMSO.  No pre-incubation reaction was performed and the reaction was initiated upon addition of *trans*-1-decalone.  The reaction mix consisted of 25 mM HEPES pH 7.0, 1.3 mM NADPH, 0-50 mM decalone, 4.5% DMSO and 2 µM didomain.  Activity assays utilizing cyclohexanone dissolved in water did not include pre-incubation with NADPH.  The reaction consisted of 25 mM HEPES pH 7.0, 1.3 mM NADPH, 0-50 mM *trans*-1-decalone, 0.2% DMSO and 2-8 µM didomain.

For all KR didomains, activity was monitored by loss of NADPH. The absorbance at 340 nm was observed and recorded for 50-60 min after addition of substrate using a Softmax Pro5 Flexstation 3 (Molecular Devices).  Triplicate reactions were run in parallel in a 96 well format.  Data analysis was performed using Kaleidagraph (Synergy Software).  Initial velocities were extracted by fitting the progress plot to the equation: absorbance = $1 - (1 - \text{CoA-fraction}_0)e^{-t(kobs)}$, where absorbance was fit to a 1 cm pathlength, t = time, and kobs = $V_{max}/K_M$. initial velocity. At initial velocities, the rate of substrate consumption is equivalent to $(V_{max}/K_M)[S_0]$ where $S_0$ is initial substrate concentration.  The absorbance was fit to a 1 cm pathlength by dividing the raw absorbance by a 'pathcheck coefficient' provided by the Softmax Pro5 Flexstation 3 software.


**Results**

*Didomain Labeling Assay*. Addition of the trace metal solution to bacterial cultures produced KR-ACP didomain that lacked the post-translation modification of a Ppant arm

(apo didomain).(20)  Acyl moieties attached to the Ppant arm of an acyl-CoA were loaded onto the ACP of the didomain through transfer of the acyl-Ppant arm of the acyl-CoA to the ACP using a phosphopantetheinyl transferase. To ensure phosphopantetheinyl transferase conditions yielded complete labeling of the didomains, labeling conditions were optimized using Tyl KR1ACP1 with CoA and the phosphopantetheinyl transferase Sfp. The amounts of apo and holo Tyl KR1ACP1 were analyzed using mass spectroscopy.  By increasing the concentration of the phosphopantetheinyl transferase, Sfp, to 11 µM, the ratio of CoA to protein could be reduced to nearly 1:1 and still observe complete labeling according to mass spectrometry  (Fig 3.5).  These conditions were used as a guide for didomain labeling reactions utilizing the phosphopantetheinyl transferase Svp and Fecik labels.

*Determine the Labeling Efficiency of Fecik Labels*.  Didomains were labeled with Fecik labels using the phosphopantetheinyl transferase Svp.  To determine labeling efficiency, CoA 547 was added to the labeling reaction.  Any unlabeled didomain that remained from the initial labeling reaction would then be loaded with Ppant-fluorophore 547.  The amount of unlabeled, and subsequently fluorophore-labeled, protein in the reaction was identified by comparing the absorbance at 550 nm with the absorbance of a parallel apo-ACP sample (Fig 3.6).  The ratio between the two samples yielded the percentage of didomain left unlabeled in the initial labeling reaction.  Labeling efficiencies of the various didomains with various Fecik labels are shown in table 3.4

Table 3.4 Labeling efficiencies of didomains and Fecik labels

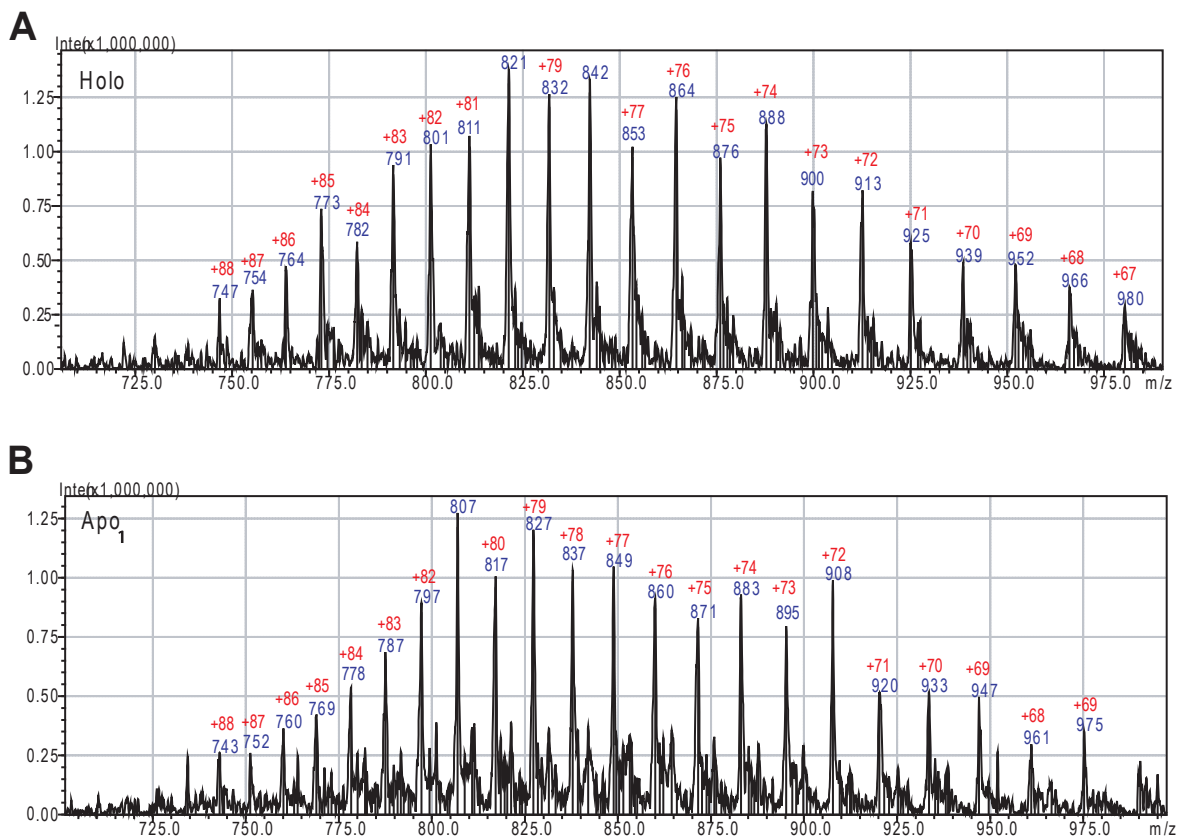| Label | Efficiency per didomain | | |
|---|---|---|---|
| P-KR2-P2 | 98% | DEBS | KR2ACP2 |
| | 96% | Pik | KR2ACP2 |
| | 63-75% | Tyl | KR1ACP1 |
| | 82 -92% | Tyl | KR7ACP7 |
| P-KR2-P1 | 37% | DEBS | KR2ACP2 |
| | 64% | Pik | KR2ACP2 |
| D-KR2-P1 | 100% | Pik | KR2ACP2 |
| | 84% | Tyl | KR1ACP1 |
| P-KR5-P3 | 63-95% | Tyl | KR1ACP1 |
| | 90-90% | Tyl | KR7ACP7 |
| EKL-103 | 71-99% | Tyl | KR1ACP1 |
| | 87-94% | Tyl | KR7ACP7 |
| P-KR5-P1 | 87-91% | Tyl | KR1ACP1 |
| | 80% | Pik | KR2ACP2 |

**Figure 3.5 Mass spectrum of Apo and Holo Tyl KR1ACP1.**
The blue numbers represent the m/z ratio of each peak. The red numbers represent the
charge state of the protein at each peak. A peak shift equivalent to the m/z ratio of a
Ppant arm (MW=340) can be observed between **A**, Tyl KR1ACP1 labeled with CoA, and
**B**, unlabeled Tyl KR1ACP1. For example, $340/80 = 4.25$ in agreement with the shift of the
+80 ion from an m/z of 821 to 817.

A



B
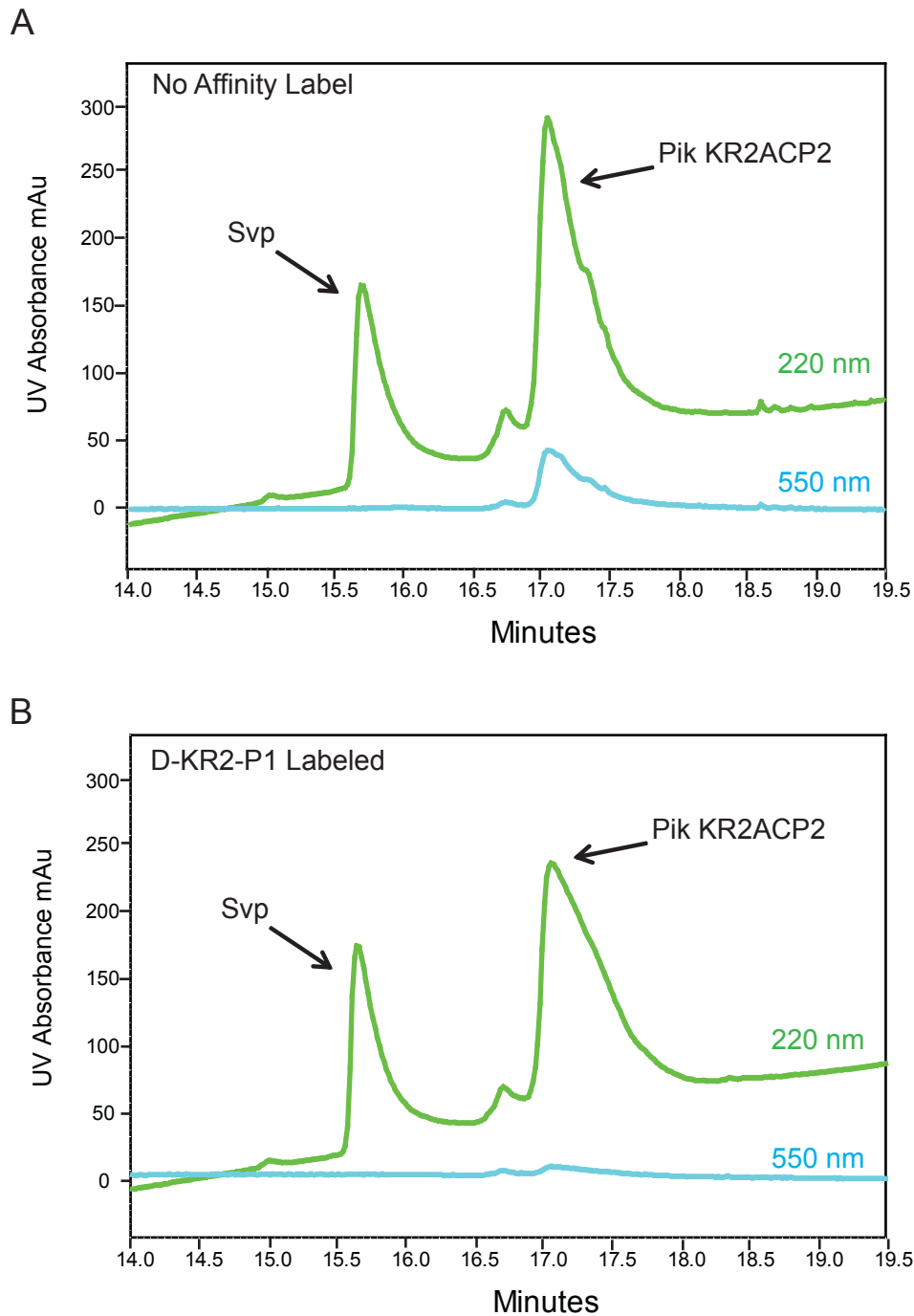


**Figure 3.6 Svp labeling efficiency.**
**A**. Negative Control. CoA-547 (New England Biolabs) was added to a mock labeling reaction. The Apo Pik KR2ACP2 was labeled with fluorophore-547 as denoted by correlating peaks at 550 nm where the fluorophore-547 absorbs, and at 220 nm where the didomain absorbs. **B**. Labeling Reaction. CoA 547 was added to a labeling reaction utilizing Fecik label P-KR5-P1 and Pik KR2ACP2. Due to near complete labeling of the didomain by P-KR5-P1, very little fluorophore-547 was tethered to the didomain.

*Didomain Activity Assay*. To determine whether a ketoreductase domain remains active in a KR-ACP didomain construct, the reductive activity of each didomain was tested with the unnatural substrate *trans*-1-decalone (Fig 3.7A). The activity of the didomain was monitored by consumption of NADPH through the decrease of absorbance at 340 nm (Fig 3.7B). All nine didomains were able to reduce *trans*-1-decalone (Fig 3.7C-E). Due to issues with *trans*-1-decalone solubility, discovered after experiment completion, the actual amount of *trans*-1-decalone in the reactions is unknown. However, a reliable comparison of enzyme activity is still possible as the same *trans*-1-decalone stock solution was used to make the serial dilutions for all activity assays and all enzymes display concentration-dependent activity.

For eight of the nine didomains, the presence or absence of the Ppant arm loaded onto the ACP did not affect the reduction rate of *trans*-1-decalone (Fig 3.7 C, D). For the didomain Tyl KR1ACP1, the reduction rate of *trans*-1-decalone increased in the presence of a Ppant arm. A control experiment entitled 'apo no Sfp labeling rxn', in which the TylKR1ACP1 didomain did not go through the mock labeling reaction to produce the Apo sample, shows a similar reduction rate to the 'apo no CoA' sample, suggesting the conditions of the non-CoA components of the labeling reaction do not affect the activity of the enzyme (Fig 3.7F).

The activity of DEBS KR2ACP2 in the apo and holo forms was tested with *trans*-1-decalone solubilized in 100% DMSO (final reaction DMSO concentration was 4.5%) (Fig 3.7E). Subsequent experiments, in which the scale of the activity assay was increased, showed that *trans*-1-decalone immediately coagulated upon addition to the reaction mix and eventually dissolved over two minutes as the reduction proceeded. Despite this discrepancy, it can still be concluded that DEBS KR2ACP2 is able to reduce *trans*-1-decalone.

**Figure 3.7 Didomain activity with unnatural substrates.**

**A.** The unnatural substrates *trans*-1-decalone and cyclohexanone were used to test the activity of nine didomains. **B.** Progress curves of apo and holo Tyl KR1ACP1 with three initial concentrations of *trans*-1-decalone solubilized in 3% DMSO.  Progress curves were fit with the equation:  absorbance $= 1 - (1 - \text{absorbance}_0)e^{-t(kobs)}$, where absorbance was corrected to a 1 cm pathlength, t = time, and kobs $= V_{max}/K_M$.  The 0 mM *trans*-1-decalone curves were linear fit.   **C., D.,**  *Trans*-1-decalone, solubilized in 3% DMSO, was used to test the activity of the holo and apo forms of eight didomains.  After completion of all assays, it was found that *trans*-1-decalone was not sufficiently soluble in 3% DMSO, resulting in a stock solution containing an unknown amount of *trans*-1-decalone. Thus the concentrations of *trans*-1-decalone listed in these figure are inaccurate.  Despite this inaccuracy, all tested didomains reduced *trans*-1-decalone in a concentration-dependent manner. Data points and error bars represent the average and spread, respectively, of at least 3 measurements. **E.**  The activity of the apo and holo forms of DEBS KR2ACP2 was tested with *trans*-1-decalone solubilized in 100% DMSO. **F.** Components of the labeling reaction do not affect the reducing ability of apo Tyl KR1ACP1
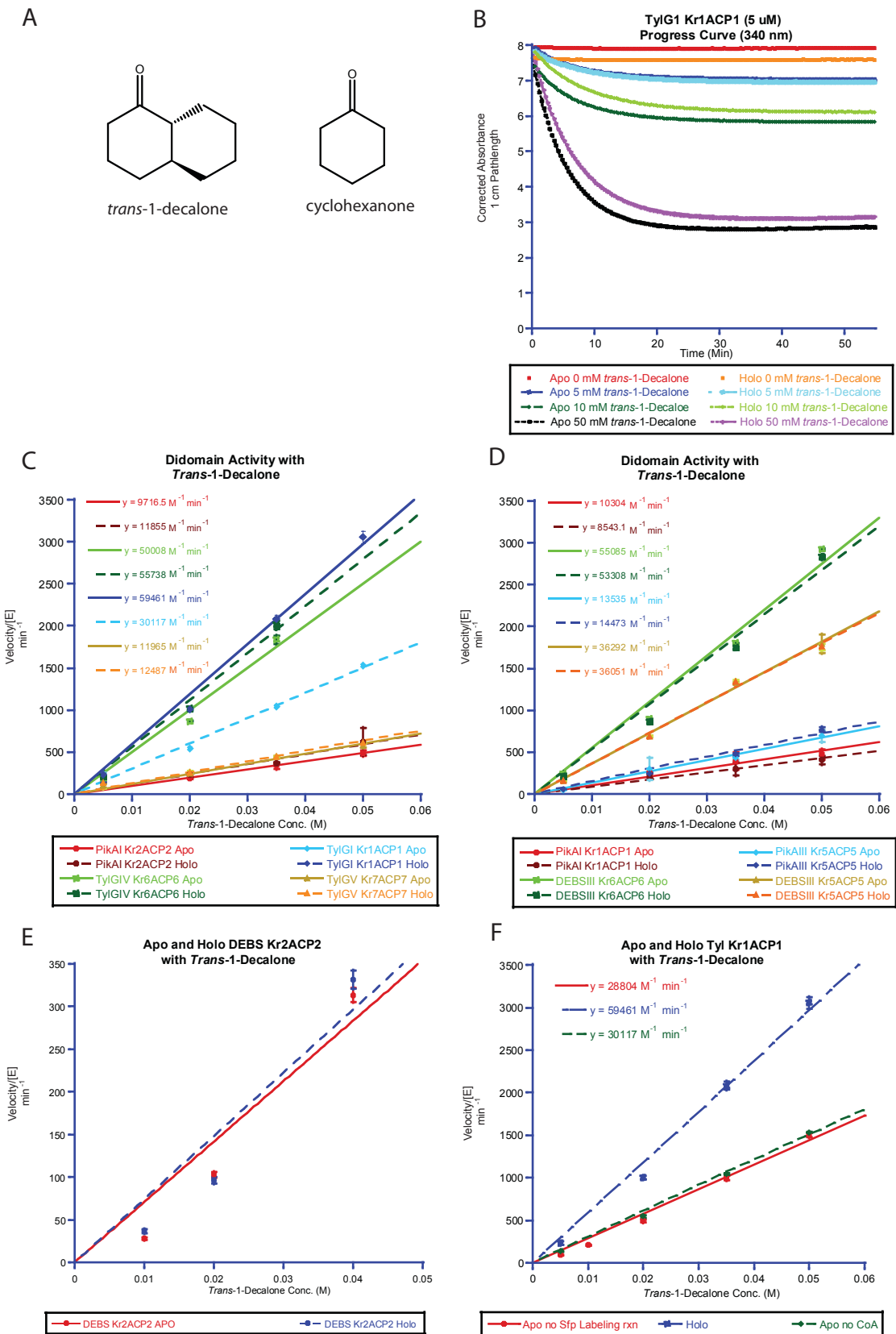
**Figure 3.7  Didomain activity with unnatural substrates.**

*Fecik Label Competition with Unnatural Substrates*: Tyl KR1ACP1 was labeled with CoA and four Fecik labels: EKL-103, D-KR2-P1, P-KR5-P1 and P-KR5-P3. The activity of labeled Tyl KR1ACP1 was assayed using *trans*-1-decalone solubilized in 100% DMSO (final reaction DMSO concentration was 4.5%) and cyclohexanone. Because *trans*-1-decalone was insoluble under starting conditions, it is not possible to obtain kinetic values of the reaction. However, a reliable trend emerged where the EKL-103-Tyl KR1ACP1 has significantly reduced activity when compared to the other labeled and unlabeled Tyl KR1ACP1 didomains (Fig 3.8A). This experiment was repeated utilizing the unnatural substrate cyclohexanone, without DMSO in the final reaction, and Tyl KR1CP1 labeled with CoA and the three Fecik labels: EKL-103, P-KR2-P2 andP-KR5-P3. A similar, though less prominent, trend was observed where the EKL-103-Tyl KR1ACP1 had reduced activity in comparison to other labeled and unlabeled Tyl KR1ACP1. (Fig 3.8B).

Tyl KR7ACP7 was labeled with CoA and three Fecik labels: EKL-103, P-KR2-P2, P-KR5-P3. The activity of Tyl KR7ACP7 labeled with different Fecik labels was compared utilizing the unnatural substrate cyclohexanone. (Fig 3.8C) EKL-103-Tyl Kr7ACP7 was the slowest enzyme followed by P-KR5-P3-Tyl KR7ACP7. The remaining P-KR2-P2 Tyl KR7ACP7, holo Tyl KR7ACP7, and apo Tyl KR7ACP7 have similar rates within experimental error.

For both the Tyl Kr1ACP1 and Tyl Kr7ACP7 competition reactions, a lag was observed in the Michaelis-Menten plots. This lag may be due to cooperativity between the substrate and NADPH binding sites, or to experimental issues in determining the initial velocities used to create the Michaelis-Menten graphs.
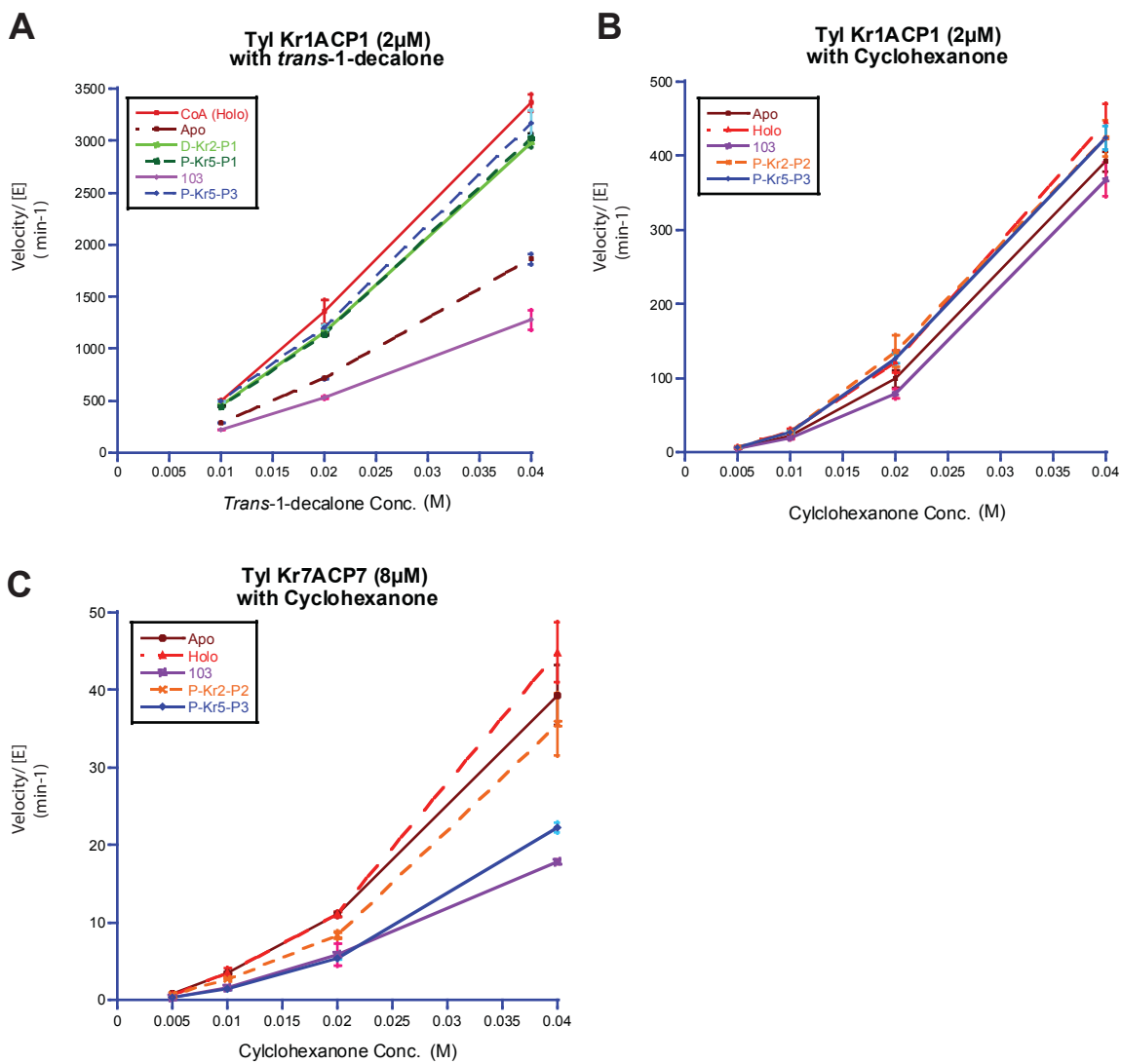
**A** Tyl Kr1ACP1 (2μM) with *trans*-1-decalone

Legend:
- CoA (Holo)
- Apo
- D-Kr2-P1
- P-Kr5-P1
- 103
- P-Kr5-P3

X-axis: *Trans*-1-decalone Conc. (M)
Y-axis: Velocity/ [E] ( min-1)

**B** Tyl Kr1ACP1 (2μM) with Cyclohexanone

Legend:
- Apo
- Holo
- 103
- P-Kr2-P2
- P-Kr5-P3

X-axis: Cylclohexanone Conc. (M)
Y-axis: Velocity/ [E] (min-1)

**C** Tyl Kr7ACP7 (8μM) with Cyclohexanone

Legend:
- Apo
- Holo
- 103
- P-Kr2-P2
- P-Kr5-P3

X-axis: Cylclohexanone Conc. (M)
Y-axis: Velocity/ [E] (min-1)

**Figure 3.8 Fecik labels compete with unnatural untethered substrates.**
Tyl Kr1ACP1 or Tyl Kr7ACP7 was labeled with multiple Fecik labels, and CoA. The apo sample did not have a Ppant arm tethered to the ACP domain. The activity of each labeled and unlabeled Tyl Kr1ACP1 was tested with **A**. trans-1-decalone or **B**. cyclohexanone. **C**. The activity of labeled or unlabeled Tyl Kr7ACP7 with cyclohexanone.

89

**Discussion**

*Loading Fecik Labels onto KR-ACP Didomains.* Though optimized labeling conditions were originally determined using the phosphopantetheinyl transferase Sfp and CoA, the phosphopantetheinyl transferase Svp was used to load the Fecik labels on the ACP of KR-ACP didomains. Compared to Sfp, Svp is more with long chain substrates and was used with the Fecik labels.

To determine the completeness of the Svp labeling reaction with a variety of Fecik labels and didomain combinations, a new method was developed. CoA 547 contains a flourophore that absorbs at 550 nm and can be transferred to the ACP domain with a phosphopantetheinyl transferase. Addition of CoA 547 to the initial labeling reaction yielded KR-ACP didomains labeled with fluorophore 547. Thus, any didomain left unlabeled by the initial reaction was subsequently labeled and identified through absorbance at 550 nm (Fig. 3.5).

*Activity of KR-ACP Didomains.* Despite the low solubility of *trans*-1-decalone, it is apparent the KR domain is active when produced as a KR-ACP didomain. The benefit of such a didomain includes the presentation of substrates by a natural carrier—the ACP. Because the direction of entry into the active site may play a key role in determining the stereochemistry of the product, presentation of substrate by an adjoining ACP domain provides more pertinent stereochemical information than that provided by a free substrate. These didomains will be useful for structural studies and placement of the substrate in the active site of the KR.

Lags were observed for the activity assay of DEBS KR2ACP2 with *trans*-1-decalone and the competition assays with Tyl KR1ACP1 and Tyl KR7ACP7 and *trans*-1-decalone or cyclohexanone (Fig 3.7). Non-Michaelis-Menten kinetics are often seen with protein complexes that have two or more binding sites. The cooperativity seen in these cases yields a sigmoidal curve. Attempts to test the activity of Tyl KR1ACP1 with cyclohexanone at higher concentrations (>320 mM ) resulted in precipitation of the substrate. The experimental curve continued to 320 mM without reaching a plateau. It is possible that didomain cooperativity may occur if substrate binding enhances NADPH binding or vice versa. Under this situation, a plateau would be seen only when the

substrate reached saturating levels, which was difficult to achieve with this unnatural substrate.

Alternatively, the lag may be due to experimental error. Samples were loaded onto the 96-well plate at room temperature and then placed in a 30°C plate reader upon start of the reaction. Neither the plate, nor the reactions were heated to 30°C prior to the start of the reaction. Thus, it is possible the rising temperature difference caused an error when determining initial velocities through curve fitting. Additionally, evaporation of the reaction mixtures occurs at 30°C yet the pathlength for each well was determined after the reaction, further obscuring absorbance readings. Therefore, a more appropriate experimental setup may yield more accurate initial values.

Of the nine didomains tested, only Tyl KR1ACP1 showed a significant difference in the reduction rates of the apo form and holo form. How the presence of the phosphopantetheine arm increases the reduction rate is unknown. It is possible that the Ppant arm is able to stabilize a more reactive conformation of the KR, allowing easier entry for either the substrate or NADPH. While the difference in reaction rates support a conformation change, a solved structure in which the substrate is present would be needed to verify such a change.

*Fecik Label Competition with Unnatural Substrates* The reduction rate of both Tyl KR1ACP1 and Tyl KR7ACP7 were affected by the presence of specific Fecik labels loaded onto the ACP domain. It is presumed that the decrease in reduction activity was due to the Fecik label competing with the free substrate for space within the active site. For both didomains, the reduction rate was decreased in the presence of loaded a EKL-103—the smallest of the available Fecik labels. Fecik label EKL-103 comprises three carbons beyond the sulfur group of the Ppant arm. This is shorter than the natural substrate, which comprises five carbons beyond the sulfur group of the Ppant arm. It is possible the other Fecik labels, P-KR2-P2 and P-KR5-P3 are too large to fit into the active site and are unable to occlude the active sites of Tyl KR1ACP1.

Alternatively, the EKL-103 Fecik label may be more effective at occluding the active site of Tyl KR1ACP1 because it has fewer carbons between the sulfur of the Ppant arm and carbonyl compared to the other labels P-KR5-P3 and P-KR2-P2, which contain

two connecting carbons. It is possible that active site binding is helped by proper placement of the thioether bond. The carbonyls of labels P-KR5-P3 and P-KR2-P2 may be too far away to bind properly.

Tyl KR7ACP7 shows decreased activity in the presence of both EKL-103 and P-KR5-P3. It is interesting that both the short chain label and the long chain label were able to compete with cyclohexanone while the medium chain label was not. This suggests that size is not the only determinant of inhibition. P-KR5-P3 is similar to the product of Tyl KR7ACP7 in that the β-alcohol is in the S-like position, unlike the P-KR2-P2 in which the β-alcohol is in the R-like position. It is possible the stereochemistry of the β-alcohol of P-KR2-P2 prevents it from entering the active site. No structure is available for Tyl KR7ACP7, thus the structure of the active site must be inferred from others.

## Works Cited

1. Keatinge-Clay, A. T. (2007) A Tylosin Ketoreductase Reveals how Chirality is Determined in Polyketides. *Chem Biol* **14**, 898-908

2. Caffrey, P. (2005) The Stereochemistry of Ketoreduction. *Chemistry & Biology* **12**, 1060-1062

3. Keatinge-Clay, A. T., and Stroud, R. M. (2006) The Structure of a Ketoreductase Determines the Organization of the β-Carbon Processing Enzymes of Modular Polyketide Synthases. *Structure* **14**, 737-748

4. Siskos, A. P., Baerga-Ortiz, A., Bali, S., Stein, V., Mamdani, H., Spiteller, D., Popovic, B., Spencer, J. B., Staunton, J., Weissman, K. J., and Leadlay, P. F. (2005) Molecular Basis of Celmer's Rules: Stereochemistry of Catalysis by Isolated Ketoreductase Domains from Modular Polyketide Synthases. *Chem Biol* **12**, 1145-1153

5. Valenzano, C. R., Lawson, R. J., Chen, A. Y., Khosla, C., and Cane, D. E. (2009) The Biochemical Basis for Stereochemical Control in Polyketide Biosynthesis. *Journal of the American Chemical Society* **131**, 18501-18511

6. Antonio, S., Marcel, J., John, C., Daslav, H., and Paul, F. L. (2007) Predicting the Nature and Timing of Epimerisation on a Modular Polyketide Synthase. *ChemBioChem* **8**, 28-31

7. Castonguay, R., He, W., Chen, A. Y., Khosla, C., and Cane, D. E. (2007) Stereospecificity of Ketoreductase Domains of the 6-Deoxyerythronolide B Synthase. *Journal of the American Chemical Society* **129**, 13758-13769

8. Caffrey, P. (2003) Conserved Amino Acid Residues Correlating with Ketoreductase Stereospecificity in Modular Polyketide Synthases. *Chembiochem* **4**, 654-657

9. Holzbaur, I. E., Ranganathan, A., Thomas, I. P., Kearney, D. J., Reather, J. A., Rudd, B. A., Staunton, J., and Leadlay, P. F. (2001) Molecular Basis of Celmer's Rules: Role of the Ketosynthase Domain in Epimerisation and Demonstration that Ketoreductase Domains can have Altered Product Specificity with Unnatural Substrates. *Chem Biol* **8**, 329-340

10. Cane, D. E. (1986) Stereochemical Studies of Natural Products Biosynthesis. *Ann N Y Acad Sci* **471**, 130-137

11. Korman, T. P., Tan, Y. H., Wong, J., Luo, R., and Tsai, S. C. (2008) Inhibition kinetics and emodin cocrystal structure of a type II polyketide ketoreductase. *Biochemistry* **47**, 1837-1847

12. Korman, T. P., Hill, J. A., Vu, T. N., and Tsai, S. C. (2004) Structural Analysis of Actinorhodin Polyketide Ketoreductase: Cofactor Binding and Substrate Specificity. *Biochemistry* **43**, 14529-14538

13. Tang, Y., Lee, H. Y., Kim, C. Y., Mathews, I., and Khosla, C. (2006) Structural and Functional Studies on SCO1815: a β-Ketoacyl-Acyl Carrier Protein Reductase from Streptomyces Coelicolor A3(2). *Biochemistry* **45**, 14085-14093

14. Yamashita, A., Endo, M., Higashi, T., Nakatsu, T., Yamada, Y., Oda, J., and Kato, H. (2003) Capturing Enzyme Structure Prior to Reaction Initiation: Tropinone Reductase-II-Substrate Complexes. *Biochemistry* **42**, 5566-5573

15. Price, A. C., Zhang, Y. M., Rock, C. O., and White, S. W. (2004) Cofactor-Induced Conformational Rearrangements Establish a Catalytically Competent Active Site and a Proton Relay Conduit in FabG. *Structure* **12**, 417-428

16. Zheng, J., Taylor, C. A., Piasecki, S. K., and Keatinge-Clay, A. T. (2010) Structural and Functional Analysis of A-Type Ketoreductases from the Amphotericin Modular Polyketide Synthase. *Structure* **18**, 913-922

17. Kao, C. M., McPherson, M., McDaniel, R. N., Fu, H., Cane, D. E., and Khosla, C. (1998) Alcohol Stereochemistry in Polyketide Backbones Is Controlled by the β-Ketoreductase Domains of Modular Polyketide Synthases. *Journal of the American Chemical Society* **120**, 2478-2479

18.    Bohm, I., Holzbaur, I. E., Hanefeld, U., Cortes, J., Staunton, J., and Leadlay, P. F. (1998) Engineering of a Minimal Modular Polyketide Synthase, and Targeted Alteration of the Stereospecificity of Polyketide Chain Extension. *Chem Biol* **5**, 407-412

19.    Leggans, E. K., Akey, D. L., Smith, J. L., and Fecik, R. A. A general scheme for synthesis of substrate-based polyketide labels for acyl carrier proteins. *Bioorganic and Medicinal Chemistry Letters* **20**, 5939-5942

20.    Studier, F. W. (2005) Protein Production by Auto-Induction in High-Density Shaking Cultures. *Protein Expression and Purification* **41**, 207-234

21.    Walker, J. M. (2005) *The Proteomics Protocols Handbook*, Humana Press, Totowa, NJ.

22.    Quadri, L. E., Weinreb, P. H., Lei, M., Nakano, M. M., Zuber, P., and Walsh, C. T. (1998) Characterization of Sfp, a Bacillus subtilis phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases. *Biochemistry* **37**, 1585-1595

23.    Lopanik, N. B., Shields, J. A., Buchholz, T. J., Rath, C. M., Hothersall, J., Haygood, M. G., Håkansson, K., Thomas, C. M., and Sherman, D. H. (2008) In Vivo and In Vitro Trans-Acylation by BryP, the Putative Bryostatin Pathway Acyltransferase Derived from an Uncultured Marine Symbiont. *Chemistry & Biology* **15**, 1175-1186

24.    Sánchez, C., Du, L., Edwards, D. J., Toney, M. D., and Shen, B. (2001) Cloning and characterization of a phosphopantetheinyl transferase from Streptomyces verticillus ATCC15003, the producer of the hybrid peptide polyketide antitumor drug bleomycin. *Chemistry & Biology* **8**, 725-738

# Chapter 4

# Curacin Dehydratases

Dehydratase (DH) domains located in a PKS module are responsible for the dehydration of the β alcohol of the growing polyketide, leaving a double bond between the α and β carbons of the dehydration product (Scheme 4.1).  Structures are available for five type I polyketide DH domains, as well as for the mammalian type I fatty acid synthase (FAS) and bacterial type II FAS.  However, very few kinetic data are available for type I PKS dehydratase domains (1).  It is presumed the DH domains do not require great ACP specificity because the ACP is an integral part of the module. However, investigations into the curacin A pathway of *L. majuscula* found that placement of the DH domains in the biosynthase based on protein sequence alignments did not agree with the placement of DH domains based on the chemical structure of curacin A. This suggests dehydratase domains may be able to act on external ACPs or they may be able to act on an alcohol moiety not in the β position.  By investigating and comparing the activities of the four curacin DH domains when presented with substrates attached to different ACPs, it will be possible to determine if the curacin DH domains have a preference for substrates presented by the natural ACP over substrates presented by an ACP from a neighboring module.
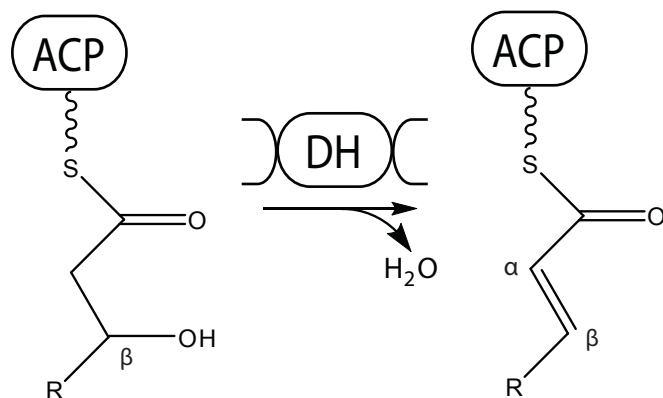
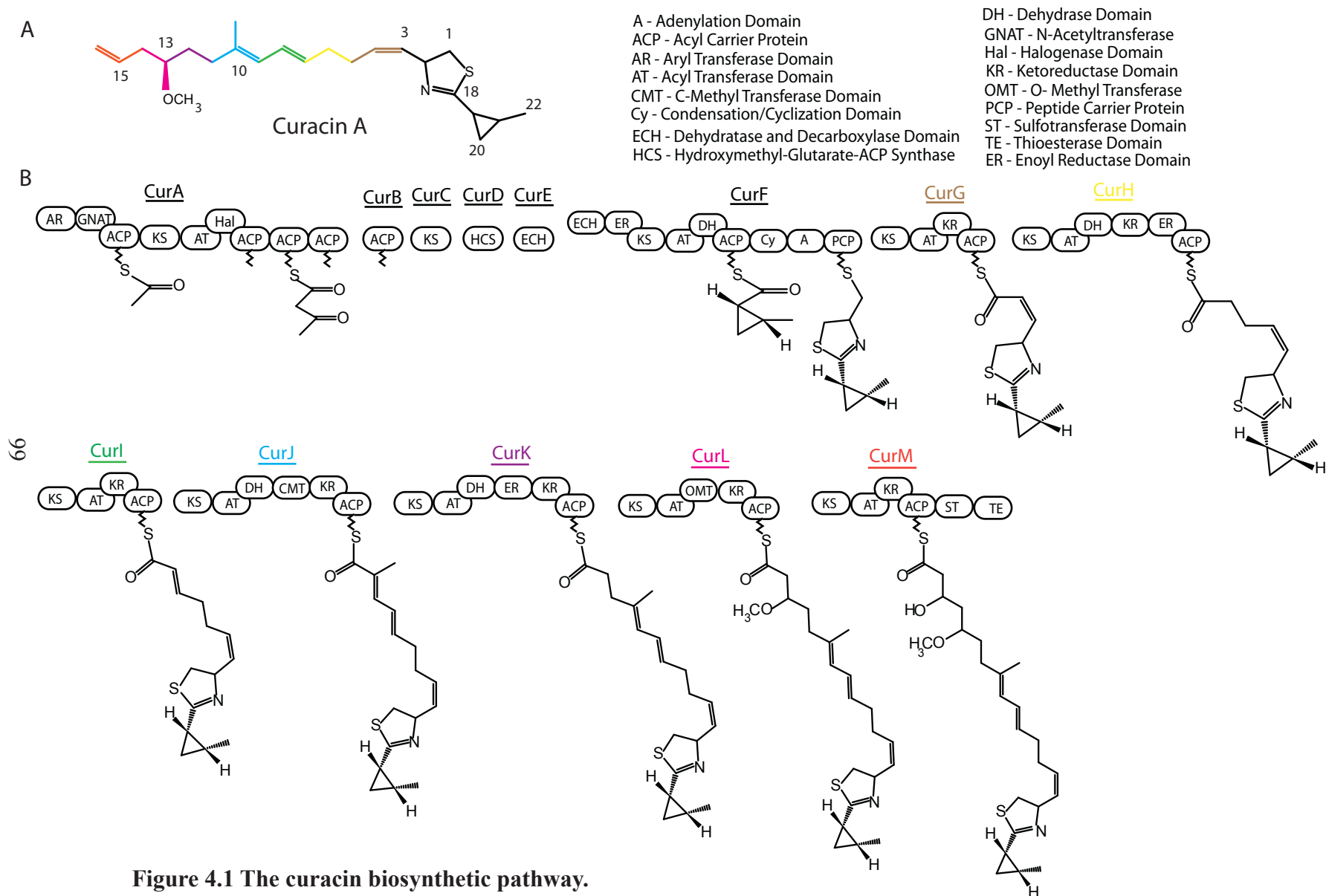**Scheme 4.1  Dehydration Reaction.**
The dehydratase is responsible for the dehydration of the β hydroxyl to a
*cis* or *trans* double bond at the C2 position.

Curacin A, a compound that has been found to inhibit microtubule polymerization and may therefore be a valuable anticancer agent (2,3), is produced by a PKS/NRPS hybrid biosynthetic pathway found in the marine cyanobacterium *L. majuscula*. Based on the chemical structure of curacin A (Fig. 4.1A), the curacin biosynthetic pathway was predicted to have five DH domains located in modules CurG, CurH, CurI, CurJ and CurK (4). The sequence identity among type I PKS dehydratases is generally low, even among domains found in the same biosynthetic pathway (5), and DH domain predictions are based on the structure of the polyketide compound and by the signature motif HxxxGxxxxP found in DH domains. Further investigations into the sequences of the curacin modules found that module CurF contains a DH domain, while modules CurG and CurI do not (4,5). It has been proposed that adjacent DH-containing modules may be able to step in to catalyze the dehydration previously thought to be performed by modules CurG and CurI (4,5).

Another unusual feature of the curacin A pathway is the production of a *cis*-double bond in curacin A (4) (Fig. 4.1A). Production of a *cis*-double bond by a DH domain is rare in polyketide synthases (6,7), and is thought to occur when the DH domain is presented with a ketoreductase product with the β-hydroxyl in the S-like configuration. As described in chapter 3, a β-hydoxyl in the S-like configuration is produced by an A-type ketoreductase, such as the ketoreductase of the CurG module. (6) Typically, DH domains that follow a B-type ketoreductase, which are much more common and produce a β-hydroxyl in the R-like configuration, yield a *trans*-double bond in the β position of the growing polyketide. Two catalytic residues can be found in the active site, a catalytic aspartate and a catalytic histidine. The stereochemistry of the β-hydroxyl group may position the polyketide such that the catalytic histidine will only be able to abstract the pro-S proton, forming a *cis*-double bond, or pro-R proton, creating a *trans*-double bond (5,8,9).

**Figure 4.1 The curacin biosynthetic pathway.**
**A**. The structure of Curacin A.  **B**. The Curacin PKS/NRPS pathway is composed of 41
catalytic domains on 13 polypeptides.  (5) Module names are color-coded to correspond
with the associated segments of curacin A.

**A**

Curacin A

A - Adenylation Domain
ACP - Acyl Carrier Protein
AR - Aryl Transferase Domain
AT - Acyl Transferase Domain
CMT - C-Methyl Transferase Domain
Cy - Condensation/Cyclization Domain
ECH - Dehydratase and Decarboxylase Domain
HCS - Hydroxymethyl-Glutarate-ACP Synthase

DH - Dehydrase Domain
GNAT - N-Acetyltransferase
Hal - Halogenase Domain
KR - Ketoreductase Domain
OMT - O- Methyl Transferase
PCP - Peptide Carrier Protein
ST - Sulfotransferase Domain
TE - Thioesterase Domain
ER - Enoyl Reductase Domain

**B**

**Figure 4.1 The curacin biosynthetic pathway.**

Crystal structures are available for five DH domains excised from type I PKS modules. These include DH domains from the curacin A pathway (5), CurF DH (PDB 3KG6), CurH DH (PDB 3KG7), CurJ DH (PDB 3KG8), CurK DH (PDB 3KG9), and from the 6-deoxyerythronolide pathway Ery DH4 (7) (PDB 3EL6). Crystal structures of three type II bacterial fatty acid synthase DH domains have been determined, ecFabA (8) (PDB 1MKA) from *E. coli*, paFabZ (10) (PDB 1U1Z ) from *P. aeruginosa* and pfFabZ (11) (PDB 1Z6B) from *P. falciparum*. A crystal structure of the mammalian fatty acid synthase, mFAS (12) (PDB 2vz9) provides information not only on the structure of a mammalian mFAS DH, but also on how the DH domain interacts with the other domains in the synthase.

The Cur DH domains are dimers in solution with each monomer consisting of a double hotdog fold (Fig. 4.2). A short helix and an extended linker about 20 residues long connect the two hotdog motifs, which make up the monomer (5). A cap motif covers the active site, which is located in a tunnel created by the two hotdog motifs and the cap (13). The catalytic residues consist of a histidine located on a cap strand (β2a) and an aspartate from the C-terminal hotdog helix (αHD2) (Fig. 4.2). Dehydration occurs as the catalytic histidine located on the β2a strand of the cap abstracts a proton from the alpha carbon of the substrate (14) and the aspartic acid abstracts a hydroxyl from the β-carbon of the substrate (8). The positive dipole of the N terminus of αHD2 may stabilize the enolate intermediate (8).

Investigations into the activity of FAS DH domains have shown a preference for ACP-linked substrates over the pantetheine, SNAC, and CoA substrates (15). When presented with 2-octenoyl-ACP, 2-hexenoyl-ACP, and crotonyl-ACP, the excised rat FASDH domain was found to prefer longer chain acyl-substrates over shorter chain substrates (8 > 6 > 4) for both the acyl-pantetheine substrates and acyl-ACP substrates (15). When assayed with S-β-hydroxy-butyryl-CoA, the rat FAS module reached equilibrium at 20% crotonyl-CoA after 10 mins (16). This shows the *in vitro* conditions favor the hydration substrate over the dehydration product.
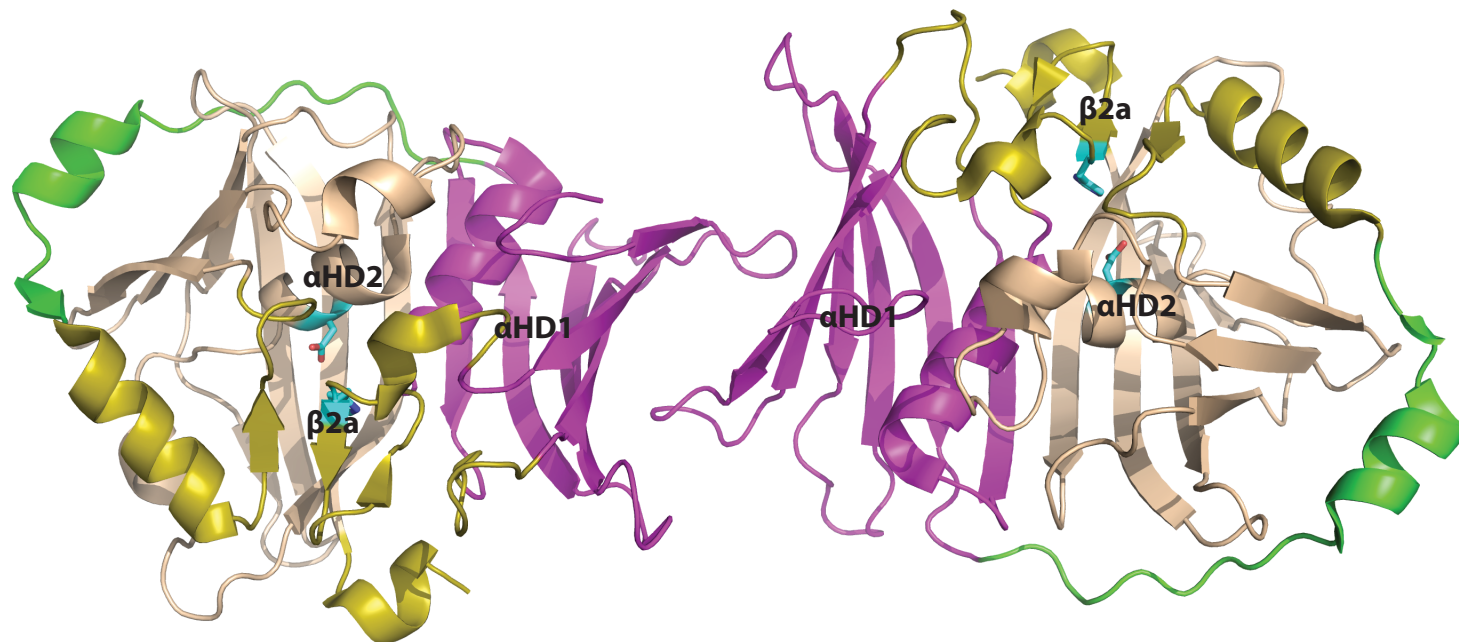
**Figure 4-2. Structure of CurH DH.**
The structure of CurH DH consists of an N-terminal hot dog motif (pink), a C-terminal hot dog motif (beige), a linker region (green) connecting the two hot dog motifs, and a cap region (gold) covering the active site. View is along the dimer axis. The catalytic His and Asp are shown as cyan sticks.

Investigations of the stereochemical specificity of the DH domain of the second module of the pikromycin biosynthetic pathway utilized a Pik2-TE fusion protein, which consists of the pikromycin module 2 (KS-AT-DH-KR-ACP) and the thioesterase domain of DEBS. Pik2-TE preferred the acyl-SNAC substrate **1** 40 times more than the enantiomeric substrate **2**, and 20 times more than the diasteriometric substrate **3** (1) (Fig 4.3). Substrate **1** contains the natural stereochemistry of the triketide product of the pikromycin module 1. The Pik2-TE fusion protein showed no activity towards the diasteriometric substrate anti-(2R,3R)-diketide-SNAC **4** (1). This suggests that stereochemistry of both the β-hydroxyl and the α-substituent are important for the activity of Pik2 DH. The Pik2-TE fusion protein was able to catalyze the reverse reaction on the triketide product **5**, though at significantly decreased rate (<5%) as compared to the rate of dehydration of substrate **1** (1).

To determine the role of the ACP in substrate specificity, six of the curacin ACPs were produced and used to present substrates to the four curacin DH domains. Of great interest is curacin DH activity on substrates presented by the CurI ACP and CurG ACPs, as the growing intermediate associated with these modules contains a dehydrated product, yet neither module contains a DH domain.

Unfortunately, initial DH activity assays utilizing a number of acyl-CoA and acyl-ACP substrates resulted in no activity. Eventually, curacin DH activity was observed when a phosphate-buffered system was utilized. The necessity for phosphate in the reaction mix remains a mystery. With the phosphate-buffered system, all four curacin DH domains were able to dehydrate a small amount (<6%) of β-hydroxyl-butyryl-CoA to crotonyl-CoA. Additionally, all four curacin DHs were able convert crotonyl-CoA to β-hydroxyl-butyryl-CoA with differing efficiencies.
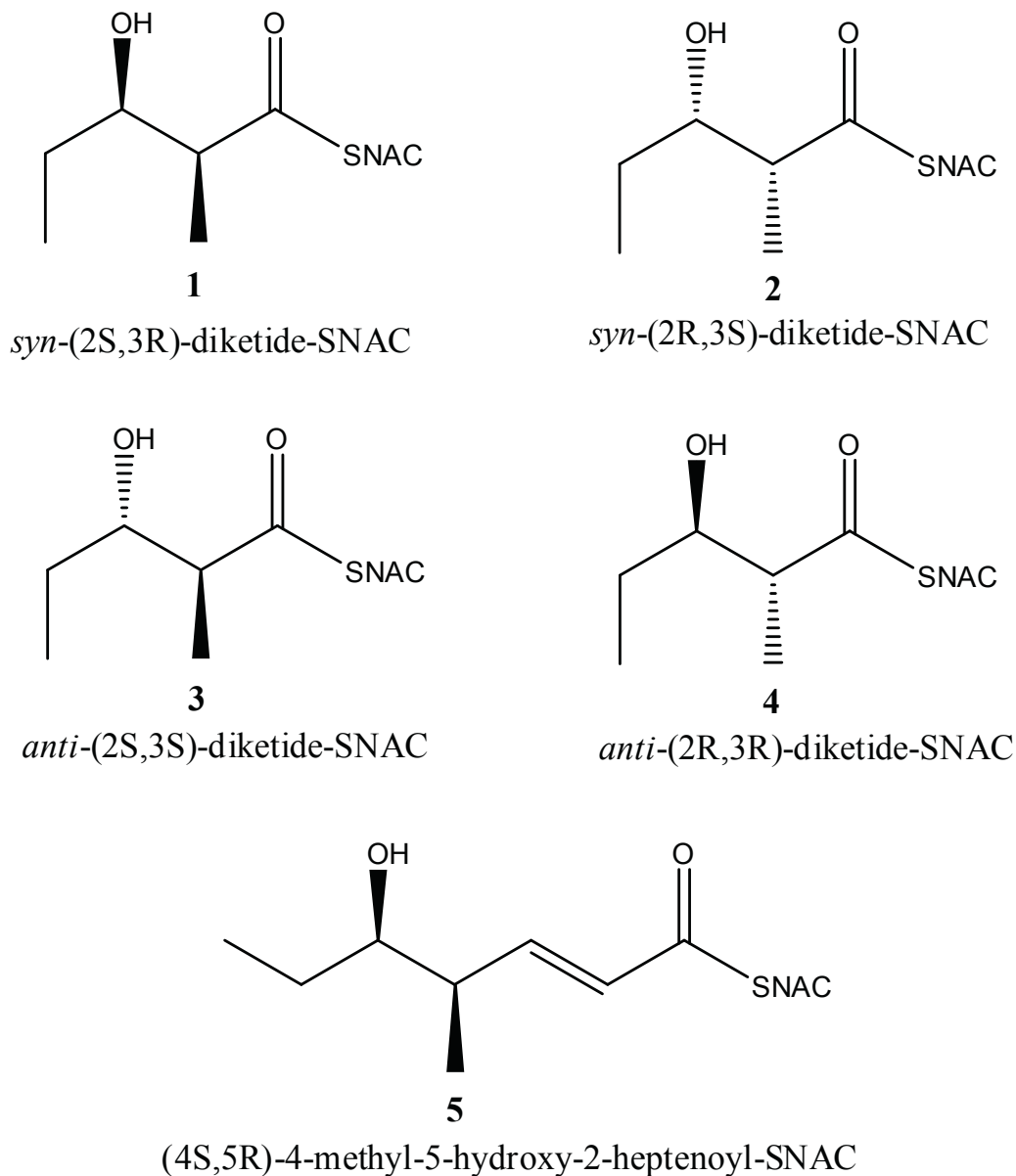
**1**

*syn*-(2S,3R)-diketide-SNAC

**2**

*syn*-(2R,3S)-diketide-SNAC

**3**

*anti*-(2S,3S)-diketide-SNAC

**4**

*anti*-(2R,3R)-diketide-SNAC

**5**

(4S,5R)-4-methyl-5-hydroxy-2-heptenoyl-SNAC

**Figure 4-3. Substrates used to assay stereospecifity of Pik-TE.**
N-acetycysteamine (SNAC) substrates mimic the phosphopantetheinyl arm of the
natural acyl-ACP substrates. The diketide substrates mimic the natural
ketoreductase triketide product presented to the dehydratase of the pikromycin
module 2. Compound 4 has the natural stereochemistry of the triketide product.
The heptenol-SNAC substrate mimics the natural product of the dehydratase
of pikromycin module 2 (1).

## Methods and Materials

*Construction of Expression Vectors for CurACPs.* 6His-Mocr-TEV-ACP fusion expression plasmids encoding the six Cur ACP genes, CurF ACP, CurG ACP, CurH ACP, CurJ ACP and CurK ACP, were created by PCR amplification from the pLM9 cosmid (4) (Table 4.1). PCR products were inserted into the expression vector pMCSG7 modified to contain a Mocr fusion tag (6His-Mocr-TEV) (17) via ligation-independent cloning (18). The resulting plasmids are listed in table 4.1. Expression plasmids for Cur DHs, Human FAS DH (hFAS DH), CurB ACP, and the multidomains (CurH DH-ER-KR, and CurJ DH-MT-KR) were created by Jamie Razelun.

*Solubility of Cur ACPs.* For each CurACP, *E. coli* strain BL21 Star (DE3) (Invitrogen) were transformed with the appropriate plasmid, and grown in a 5 mL culture at 37°C overnight. The next morning, 3 mL of culture was added to a 500 mL culture at 37°C in 2-L flasks containing Terrific Broth with 4% glycerol, trace metal components (19) according to Table 3.3 and 1 mg/L ampicillin. When the optical density reached 1.0 at 600 nm, the culture was transferred to 20°C. A 1 mL sample was transferred to an eppendorf tube and labeled the 'non-induced' sample. 100 µM IPTG was added to the 500 mL culture and both the culture and non-induced sample were allowed to shake at 20°C overnight. The next morning, a 1 mL sample of the 500 mL culture was taken and labeled the 'induced' sample. The 500 mL culture was centrifuged at 5,670g for 45 min. The cell pellet was frozen and stored at -20°C. 0.07g of Cell Lytic (Sigma-Aldrich) was added to the non-induced and induced samples. After incubating at room temperature for 10 min, the induced sample was centrifuged at 18,00g for 10 min. The supernatant of the induced sample was labeled the 'soluble' sample, and the pellet of the induced sample was labeled the 'insoluble' sample. 2x SDS-PAGE loading dye was added to a 10 µL aliquot of the soluble, insoluble and induced samples. 10 µL of the samples were loaded onto a 12.5% SDS-PAGE gel and run at 200 V for 45 min (Fig 4.4).

Table 4.1 Primers for Cur ACPs

| Plasmid | Gene | nt of pLM9 | Primers |
|---|---|---|---|
| pHC003 | Cur HACP | 18512-18829 | CurH ACP_for : TAC TTC CAA TCC AAT GCA CTA GAA GCA GCG GCT GTT T<br>CurH ACP_rev: TTA TCC ACT TCC AAT GTT TA TTT CCT CTT GTT CTA A |
| pHC004 | CurJ ACP | 30419-30739 | CurJ ACP_for: TAC TCC CAA TCC AAT GCT CTT GAA GCA ACA GCA CC<br>CurJ ACP_rev: TTA TCC ACT TCC AAT GTT AAT CTA CCA TAT CAT CTA GTG TTT C |
| pHC007 | CurF ACP | 3825-4142 | Start_CurF: TAC TCC CAA TCC AAT GCT TTA AAA TCA GCA TCT GTT TCA<br>CurF_Stop: TTA TCC ACT TCC AAT GTT AAT ACC TTT GCT CTG GAT GA |
| pHC008 | CurG ACP | 11904-12221 | CurG_Start: TAC TTC CAA TCC AAT GCT TTG GAA GCT GTA ACC GAA AC<br>CurG_Stop: TTA TCC ACT TCC AAT GTT ATG AAT CGT CTT CTT GTT GAT CT |
| pHC009 | CurI ACP | 23476-23793 | CurI_Start: TAC TTC CAA TCC AAT GCT TTA GAT AAG GCA TCA GCA G<br>CurI_Stop: TTA TCC ACT TCC AAT GTT AAT CTA CCT CTT TTC GTT CAG |
| pHC010 | CurK ACP | 37188-37505 | CurK_Start: TAC TTC CAA TCC AAT GCA CTA GAA GCA GCA CTG GTT<br>CurK_Stop: TTA TCC ACT TCC AAT GTT AAT CTT CTT CAT CTA GTG CTT T |
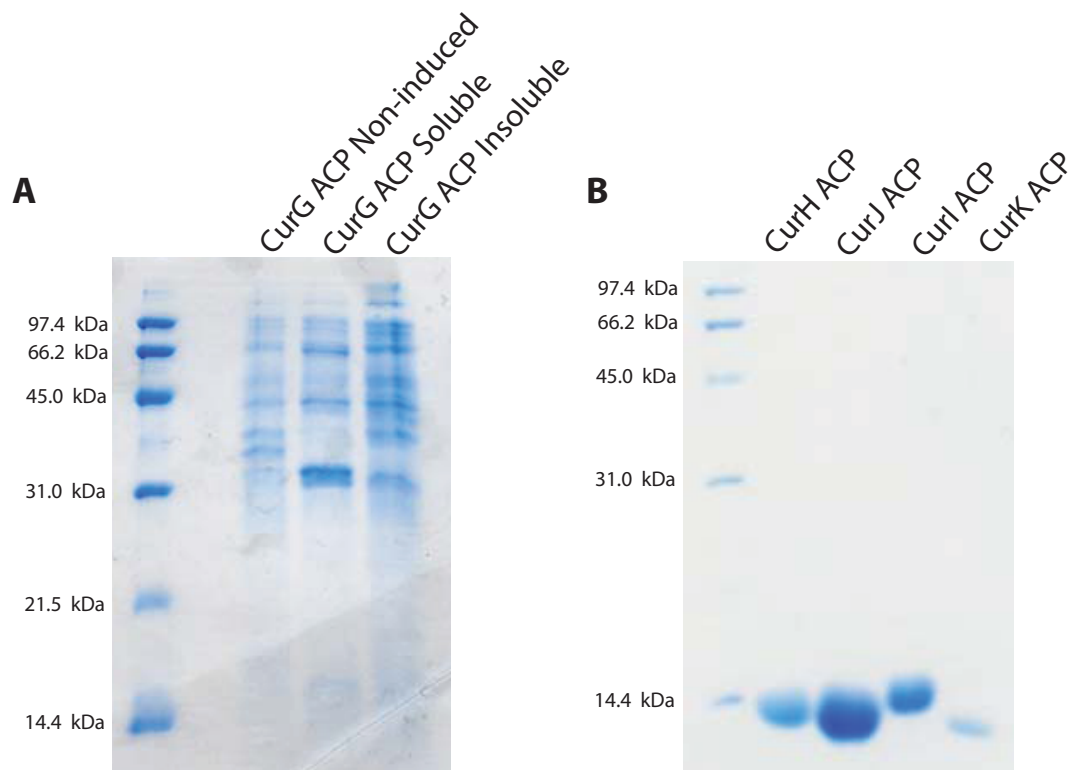
**Figure 4.4 Production of the CurACPs.**
**A**. 6His-Mocr-TEV-CurG ACP appears in the soluble fraction of the cell lysate at
the expected molecular weight (28 kDa). **B**. After ACP production, the Mocr peptide
was cleaved off resulting in purified stand-alone Cur ACPs. The Cur ACPs are
shown at the expected molecular weight (~ 12 kDa) after one freeze-thaw cycle
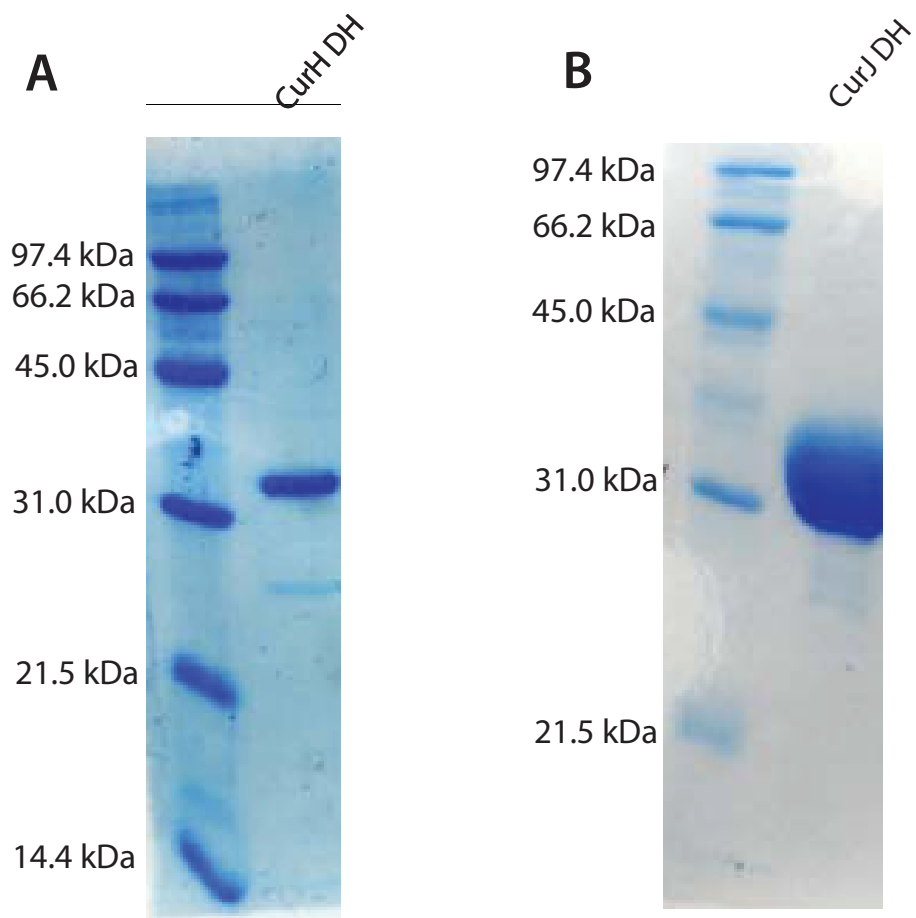after purification.

**Figure 4.5 Production of the CurH DH and CurJ DH.**
**A**. CurH DH is shown at the expected molecular weight (~ 34 kDa) after purification with a HiLoad 16/60 S200 Superdex column. **B**. CurJ DH is shown at the expected molecular weight (~36 kDa) after purification with a HiLoad 16/60 Superdex S200 column.

*Expression and Purification of Proteins.* For each Cur ACP and Cur DH, *E. coli* strain BL21 Star (DE3) (Invitrogen) was transformed with the appropriate expression plasmid. 500-mL cultures were grown at 37°C in 2-L flasks containing Terrific Broth with 4% glycerol, trace metal components (19) according to Table 3.3 and 1 mg/L ampicillin. After the cell cultures reached an optical density at 600 nm of 1.0, cultures were transferred to 20°C for 2 hrs. Afterwards, the cultures were induced with 100 µM IPTG and allowed to shake for 20 hrs at 20°C. The cells were then harvested by centrifugation at 5,670g and cell pellets were frozen and stored at -20°C.

Thawed cell pellets were resuspended in 40 mL of 300 mM NaCl, 50 mM Tris pH 7.5, and 10% glycerol. The resuspended cells were disrupted by sonication and the lysate was collected by centrifugation at 26,890g. Lysate supernatant was loaded onto a Histrap HP column (5 mL, GE HealthCare), previously equilibrated with 15 mL of 20 mM imidazole with 300 mM NaCl, 50 mM Tris 7.5 and 10% glycerol. The proteins were eluted with an imidazole gradient (20 mM-400 mM) in the same solution.

For purification of 6His-Mocr-TEV-ACP fusion protein, fractions containing the fusion protein were pooled and underwent TEV cleavage to remove the 6His-Mocr tag. 0.02 mg of TEV protein was used for each mg of fusion protein. DTT or TCEP was added to the TEV cleavage solution for a final concentration of 1 mM DTT or 0.6 mM TCEP. The TEV cleavage continued for 4 hrs at room temperature, after which, the protein solution was dialyzed in 300 mM NaCl, 50 mM Tris 7.5 and 10% glycerol to remove the imidazole. The protein solution was then reloaded onto a Histrap HP column (5 mL, GE HealthCare) previously equilibrated with 300 mM NaCl, 50 mM Tris 7.5, 10% glycerol and 30 mM imidazole. Flow-through fractions containing the TEV-cleaved ACP protein were collected and pooled.

Fractions containing the TEV-cleaved Cur ACP protein were pooled and concentrated to ~ 2 mg/mL and loaded onto a HiLoad 16/60 Superdex 75 column which had previously been equilibrated with 300 mM NaCl, 50 mM Tris pH 7.5, and 10% glycerol, or with 150 mM NaCl, 20 mM Tris pH 7.5, and 10% glycerol. Proteins were eluted with the same buffer system used to equilibrate the column. Pooled fractions containing the Cur ACP protein were concentrated with an Amicon Ultra-15 centrifugal

filter unit (Millipore) to concentrations of 3-10 mg/mL.  Purified proteins were flash-frozen in liquid nitrogen and stored at -80°C. Protein concentrations were determined using the absorbance at 280 nm and calculated extinction coefficients (20).  Protein purity was determined by loading Cur ACPs onto a 3% SDS-PAGE gel.  The gel was run for 45 min at 200V (Fig 4.4B).

For purification of Cur DHs, fractions containing the Cur DH protein were pooled, concentrated to ~ 2 mg/mL and loaded onto a HiLoad 16/60 Superdex 200 column which had previously been equilibrated with 300 mM NaCl, 50 mM Tris pH 7.5, and 10% glycerol.  In the case of CurF DH, CurK DH, and the multidomains CurJ multidomain (DH-MT-KR) and CurH multidomain (DH-ER-KR), the equilibrating buffer contained 2 mM DTT.  Proteins were eluted with the same buffer system used to equilibrate the column.  Pooled fractions containing the DH protein were concentrated with an Amicon Ultra-15 centrifugal filter unit (Millipore) to concentrations of 3-10 mg/mL.  Purified proteins were flash-frozen in liquid nitrogen and stored at -80°C. Protein concentrations were determined using the calculated extinction coefficients (20). (Fig 4.5)

Frozen aliquots of CurF DH were thawed, and loaded onto a HiLoad 16/60 Superdex 75 analytical column which had previously been equilibrated with 300 mM NaCl, 50 mM Tris pH 7.5, and 10% glycerol to remove DTT.  Pooled DH protein were concentrated with an Amicon Ultra-15 centrifugal filter unit (Millipore) to concentrations of  7 mg/mL.  Purified protein was flash-frozen in liquid nitrogen again and stored at -80°C. Protein concentrations were determined using the absorbance at 280 nm and calculated extinction coefficients (20).  Before use in assays, CurK DH, CurJ multidomain (DH-MT-KR) and CurH multidomain (DH-ER-KR) were dialyzed overnight in 300 mM NaCl, 50 mM Tris pH 7.5, and 10% glycerol to remove DTT.


*ACP Labeling Reaction.*   The labeling reaction consisted of 100 mM Tris pH 8.0, 10 mM MgCl$_2$, 300 µM acyl-CoA or CoA,11µM phosphopantetheinyl transferase protein (Svp,

AAG43513), and 18 μM ACP protein.   The labeling reaction proceeded for at least 1 hr at 30°C or room temperature.

In instances where the ACP was loaded with an acyl unit and refrozen, the labeling reaction was run over a Histrap HP column (5 mL, GE HealthCare) to remove the phosphopantetheinyl transferase.   Fractions containing the labeled ACP protein were collected, pooled , concentrated to 1-3 mg/mL with an Amicon Ultra-4 centrifugal filter unit (Millipore), frozen in liquid nitrogen and stored at -80°C.  Protein concentrations were determined using the calculated extinction coefficients (20).

*Activity Assay with Acyl-CoAs*. The initial acyl-CoA dehydratase assay mix consisted of 100 mM Tris-HCl pH 7.0, 100 μM acyl-CoA, 10-25 μM curacin DH.  The dehydratase assay proceeded for 1-24 hrs before the reaction was quenched with a final concentraton of 5% tricholoracetic acid.  The samples were spun down at 4°C to remove the DH protein, and the supernatant containing the acyl-CoAs was analyzed using high pressure liquid chromatography.  Supernatant samples were loaded onto a C18 reverse phase column (Polymer Labs) and were monitored at 259 nm.  Separation of samples were performed using a linear gradient of buffer A (water with 0.1% triflouroacetic Acid) and buffer B (acetonitrile with 0.1% trifluoroacetic acid) with a constant flow rate of 1.0 mL/min.  Initial conditions included 5% buffer B over 5 min. Then, a linear gradient to 100% B over 20 min. Then 100% B for 5 min, and 5% B for 10 mins.

Additional acyl-CoA dehydratase assays were performed using a phosphate buffering system.  These reactions consisted of 100 mM $NaPO_4$ pH 7.5, 1 mM crotonyl-CoA or R/S-β-hydroxyl-butyryl-CoA, and 25 μM of curacin DH.  The reactions proceeded for up to 32 hrs.  18 μL of reaction mixture were quenched with an equivalent volume of 10% trichloroacetic acid.  Quenched reactions were spun down at 4°C to remove the DH protein.  25 μL of supernatant was removed and diluted to 125 μL with filtered water.  100 μL of diluted sample was loaded onto a C18 reverse phase column (Polymer Labs) and were monitored at 259 nm.  Separation of samples was performed using a linear gradient of buffer A (water with 0.1% triflouroacetic Acid) and buffer B (acetonitrile with 0.1% trifluoroacetic acid) with a constant flow rate of 1.0 mL/min.  Initial conditions included 10% buffer B over 2 min. Then, a linear gradient to 30% B

over 12 min. Buffer B was increased to 100% over 1 min, and then held steady at 100% for 5 min. After which , buffer B was brought back to 10% over 1 min, and the column equilibrated to 10% buffer B for 8 mins.

*Mass Spectrometry Experiments with Acyl-Cur ACPs.* CurB ACP, CurH ACP, and CurJ ACP were loaded with either R,S-β-hydroxy-butyryl-CoA or crotonyl-CoA in an ACP labeling reaction as described above utilizing Sfp as the phosphopantetheinyl transferase. The labeling reaction proceeded for 2 hrs before the appropriate curacin DH (10-30 μM final concentration) was added to the solution. Samples were then analyzed using LC-MS. 5 uL of each sample was loaded onto a PLRPS Column (Polymer Labs) at a flow rate of 0.33 mL/min with a linear gradient utilizing HPLC grade buffer A (water with 0.1% formic acid), and HPLC grade buffer B (acetonitrile with 0.1% formic acid). Initial conditions included 5% buffer B over 5 min. Then, a linear gradient to 50% B over 13 min. Then 100% B for 5 min, and 5% B for 10 mins.

**Results:**

*Dehydratase Activity Assays.* Initial attempts to observe DH activity were performed in Tris-HCl buffered solutions. Despite the variety of substrates presented to the curacin DHs, neither dehydration nor hydration was observed in these assays. Later attempts to observe DH activity proved successful when a sodium phosphate buffering system was utilized instead.

In a Tris-HCl buffered system, neither CurJ DH, CurH DH, nor CurF DH catalyzed the dehydration of R,S-β-hydroxyl-butyryl-CoA, or the hydration of crotonyl-CoA. Hypothesizing that the DHs were specific towards ACP bound substrates, the crotonyl and β-hydroxyl-butyryl moieties were loaded onto CurB, a stand-alone ACP of the curacin pathway. Mass spectrometry was used to identify an 18 Da shift, equivalent to the loss or gain of a water molecule, however, no DH activity was observed. Changing the pH of the Tris-HCl buffered reaction solution had no effect on DH activity as varying

the pH from 6.0-9.0 in 0.5 step increments failed to yield observable dehydration or hydration.

DH domains function as part of an entire module *in vivo*. It was hypothesized that the surrounding domains were needed for DH function. Thus, acyl-CurB substrates were presented to the CurJ multidomain (DH-MT-KR) and CurH multidomain (DH-ER-KR). However, no DH activity was observed under these conditions.

CurB has low sequence identity to the other ACPs of the curacin pathway (Table 4.2) and it was hypothesized that the curacin DHs may have evolved to not recognize acyl-CurB substrates. Excised ACPs of the CurF, CurG, CurH, CurI, CurJ and CurK modules were cloned, produced and purified. CurJ ACP and CurH ACP were labeled with β-hydroxy-butyryl or crotonyl moieties and presented as substrates to the respective curacin DH. Despite what should have been a more similar substrate, no dehydration or hydration activity was observed with the Tris-HCl buffering system.

Table 4.2 Sequence Identity of the curacin DH domains

|  | CurHACP 2054-2169 | CurKACP 2082-2197 | CurFACP 1970-2085 | CurJACP 2144-2259 | CurBACP 1-79 |
|---|---|---|---|---|---|
| CurHACP/ 2054-2169 | ---- | 80.2 | 69.8 | 37.9 | 19 |
| CurKACP/ 2082-2197 | 93 | ---- | 63.8 | 38.8 | 17.7 |
| CurKACP/ 2082-2197 | 93 | ---- | 63.8 | 38.8 | 17.7 |
| CurFACP/ 1970-2085 | 81 | 74 | ---- | 30.2 | 16.5 |
| CurJACP/ 2144-2259 | 44 | 45 | 35 | ---- | 17.7 |
| CurBACP/ 1-79 | 15 | 14 | 13 | 14 | ---- |

The presumed intermediates of the curacin DHs are longer than the four carbon hydroxyl butyryl or crotonyl moieties. To test whether the curacin DHs were specific towards long chained substrates, the R and S form of **6** (Fig. 4.6A) (21) was loaded onto CurH ACP and presented as a substrate to CurH DH. The reactions utilized a potassium

phosphate buffering system (pH 7.0), or a Tris-HCl buffering system (pH 8.0). After two hours, no dehydration activity by CurH DH was observed in either buffering system.

As a positive control for dehydration, the excised DH domain of the human fatty acid synthase (hFAS DH) was purified. Reaction conditions were obtained from the literature (22) and hydration or dehydration activity with R,S-β-hydroxyl-butyryl-CoA or crotonyl-CoA in a sodium phosphate buffered solution was tested for all four curacin DHs and hFAS. HPLC was utilized to quantify the amounts of acyl-CoA in each reaction (Fig 4.6B, C). After 2 hrs, hydration activity was observed for CurK DH, CurJ DH and hFAS DH (Fig 4.7A). After 32 hrs, hydration activity was observed for all tested DHs. hFAS DH appeared to reach equilibrium at 8% crotonyl-CoA after 4 hrs, and CurJ DH appeared to reach equilibrium at 30% crotonyl-CoA after 8 hrs. CurF DH and CurH DH continued to approach equilibrium after 32 hrs.

**Figure 4.6 Assay of CurJ DH hydration and dehydration activity.**
Reactions containing curacin DH with acyl-CoA (red) were run in parallel with control reactions that did not contain curacin DH (blue). After 32 hrs, a sample of each reaction was taken and acyl-CoAs were identified by HPLC. **A**. Compounds used to test the activity of the Cur DHs. **B.** Dehydration of β-hydroxyl-butyryl-CoA to crotonyl-CoA by CurJ DH in 32 hrs. **C**. Hydration of crotonyl-CoA to β-hydroxyl butyryl-CoA by CurJ DH in 32 hrs.
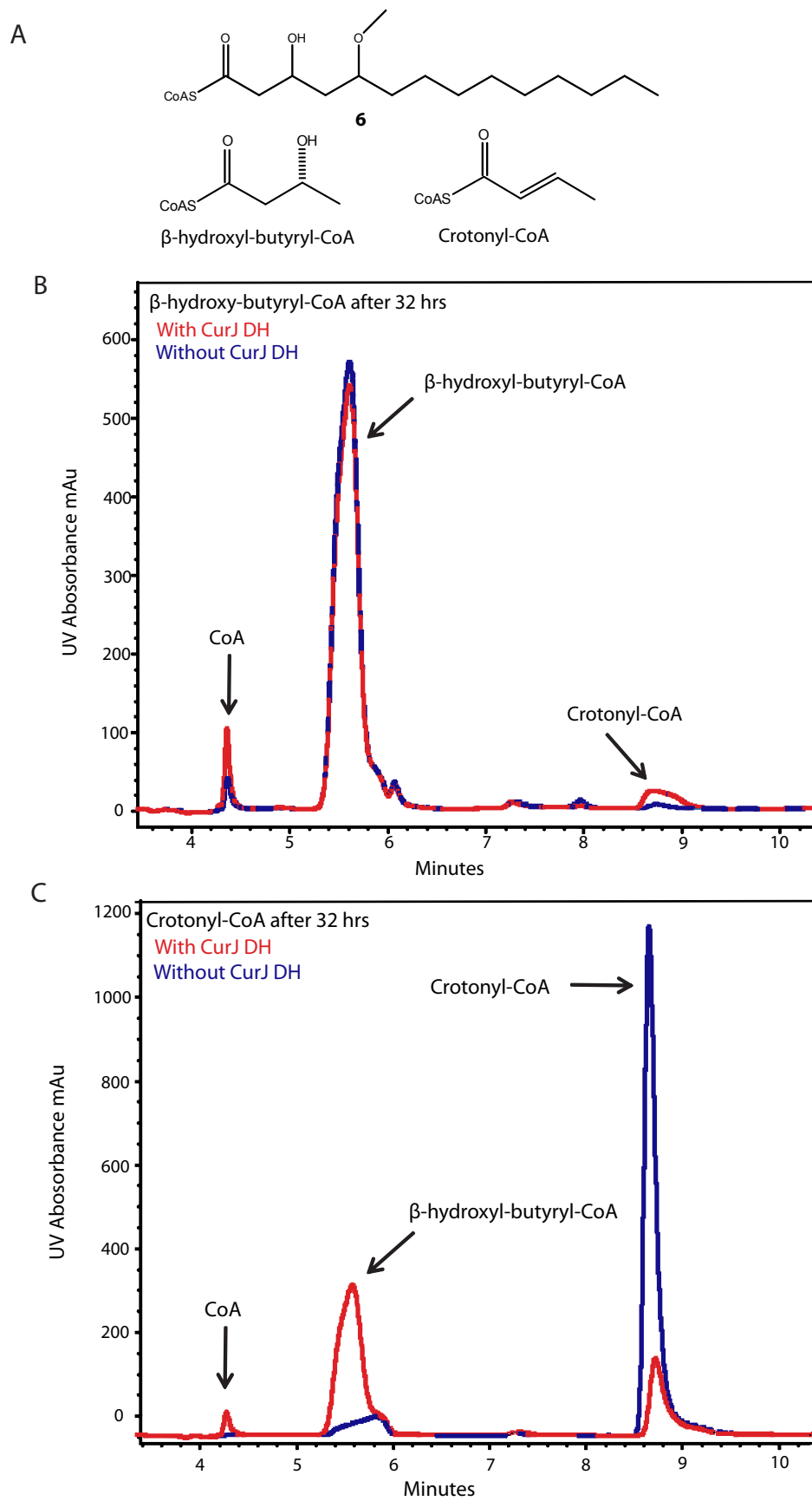
**Figure 4.6 Assay of CurJ DH hydration and dehydration activity.**

In addition to hydration activity, a small amount of CoA (<2%) accumulated in all but the CurH DH hydration assay (Fig. 4.7B, C) after 32 hrs.  CoA was observed after 30 min in both the hFAS and CurJ DH reactions.  CoA did not appear to accumulate in the CurF DH and CurK DH reactions until the 24 hr time point.   In a negative control, which contained crotonyl-CoA without a DH, no CoA accumulation was observed.   Free CoA is not an expected substrate of the Cur DHs and it is likely that any CoA accumulation *in vivo* is not significant under *in vivo* time-frames.

A small amount of dehydration activity by hFAS DH and all four curacin DHs was observed after 32 hrs (Fig. 4.7D).   After 32 hrs, the hFAS DH did not appear to reach equilibrium.  A small amount of CoA was observed in the no DH control (<2.5%) (data not shown) of the dehydration reaction, suggesting that the β-hydroxyl-butyryl-CoA was not stable in the reaction conditions for long periods of time.
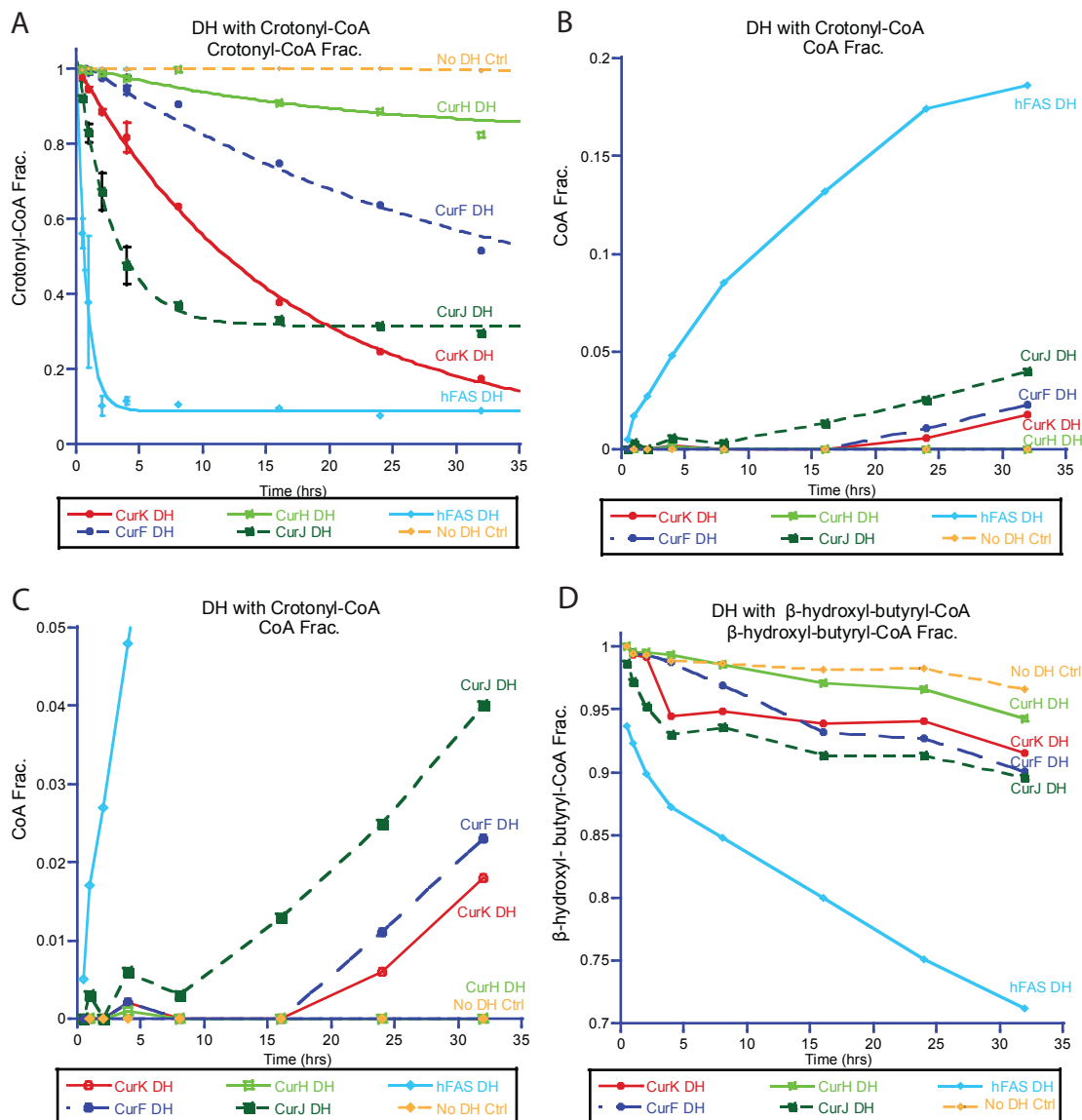
**Figure 4.7 Hydration and dehydration of acyl-CoAs.**
Hydration reactions contained crotonyl-CoA and a dehydratase. Dehydration reactions contained β-hydroxyl butyryl-CoA and a dehydratase. The 'no DH' samples contained only acyl-CoA and were used to control for any degradation of the acyl-CoA. **A**. Hydration assay of crotonyl-CoA to β-hydroxyl-butyryl-CoA as observed by the decrease of the crotonyl-CoA fraction. **B**. Increase of the CoA fraction during the hydration assay. **C**. Enlarged view of plot B to show the increase of the CoA fraction for the CurK DH and CurF DH hydration reactions. **D**. Dehydration of β-hydroxyl-butyryl-CoA to crotonyl-CoA as observed by the decrease of the β-hydroxyl-butyryl-CoA fraction.

**Discussion**

*Buffering System Preference.* The curacin DH domain dependence on the sodium phosphate buffering system needs further characterization. Studies of Pik DH2 activity were performed in a potassium phosphate buffer as were studies of rat FAS DH and hFAS DH, all of which showed hydration and dehydration activity (1,15,16). Studies of Ery DH4 in a Tris buffered reaction system yielded no DH activity (7). Thus a preference in buffering systems may not be exclusive to the curacin DH domains and should be investigated further. For a phosphate ion to be present in the active site of a DH, one would expect to find a highly-charged positive area within the active site to attract such a highly charged negative ion as phosphate. No such positively charged area can be observed in the active sites of the four curacin DHs. Thus, if phosphate buffer is affecting catalysis, it is probably doing so in an indirect manner. Or, it is also possible the observed inactivity of the enzymes in a Tris-HCl buffer is due to an unfortunate interaction between the Tris molecule and the enzyme, such as the Tris molecule occluding the active site entrance.

*Curacin Dehydratase Equilibrium. In vitro* reactions containing rat FAS DH domains have typically found the equilibrium of the hydration and dehydration reactions to lie towards the hydrated substrates instead of the dehydrated products (16). This may be due to the high concentration of water, a product of dehydration, in the *in vitro* reactions, which pushes the equilibrium towards the hydrated agents. It is possible that *in vivo* conditions are more conducive to producing an equilibrium where the dehydration products are favored. Alternatively, equilibrium may not have to favor the products for efficient catalysis, as the dehydrated products are quickly removed through further modification and transfer to downstream modules, resulting in a constant pull towards dehydration (16).

*CurF DH Activity.* The presumed final product of the CurF module does not contain an α-β double bond, suggesting that dehydration is not carried out on the intermediate attached to the CurF module. However, results show that excised CurF DH is able to catalyze both the dehydration and hydration at the β carbon, albeit at a slower rate than CurJ DH and CurK DH. It is possible the CurF DH domain may dehydrate an intermediate

attached to another module, such as the CurG or the CurI modules. Such a preference may be determined by presenting the β-hydroxy-butyryl- or crotonyl- moieties on multiple Cur ACPs.

*Production of a Cis Double Bond.* Curacin A is unusual in that it contains a *cis* double bond. As mentioned previously, a *cis* double bond is thought to be the product of the dehydration of an S-like β-hydroxyl. As described in chapter 3, an S-like β-hydroxyl is produced by an A-type KR. The structure of curacin A suggests the *cis* double bond is produced by the CurG module and the KR of the CurG module is an A-type KR according to sequence. However, the CurG module lacks a DH domain. Thus, the DH responsible for producing the *cis* double bond remains a mystery. It has previously been hypothesized that a curacin DH from another module may act on the CurG intermediate (5). Further testing to determine any ACP substrate specificity of the four curacin DH domains may shed light on this question.

Whichever dehydratase is responsible for the dehydration to produce the CurG intermediate must accept an alcohol substrate in the S-like configuration. Chiral HPLC should be utilized to determine the if any Cur DH is able to convert a crotonyl- moiety to an S-β-hydroxyl-butyryl- moiety. Alternatively, chiral HPLC may be used to separate R-β-hydroxyl butyryl-CoA, from S- β-hydroxyl butyryl-CoA and Cur DH activities may be tested with the stereochemically pure substrates. A preference for the S-like configuration over the R-like configuration would suggest that particular DH is responsible for the *cis* double bond of the CurG intermediate.

With a working assay to determine curacin DH dehydration and hydration activity, acyl-ACP substrate specificity may be characterized. Prior to this, substrate specificity for a PKS dehydratase domain has been available only for the Pik DH2 (1). Additionally, a number of active-site mutants for the various curacin DHs, created by Jamie Razelun and Jason Tehranisa, are available and may be used to determine the importance of various residues within the active-site for dehydration or hydration activity.

**Works cited**

1. Wu, J., Zaleski, T. J., Valenzano, C., Khosla, C., and Cane, D. E. (2005) Polyketide Double Bond Biosynthesis. Mechanistic Analysis of the Dehydratase-Containing Module 2 of the Picromycin/Methymycin Polyketide Synthase. *Journal of the American Chemical Society* **127**, 17393-17404

2. Blokhin, A. V., Yoo, H. D., Geralds, R. S., Nagle, D. G., Gerwick, W. H., and Hamel, E. (1995) Characterization of the Interaction of the Marine Cyanobacterial Natural Product Curacin A with the Colchicine Site of Tubulin and Initial Structure-Activity Studies with Analogues. *Molecular Pharmacology* **48**, 523-531

3. Verdier-Pinard, P., Lai, J.-Y., Yoo, H.-D., Yu, J., Marquez, B., Nagle, D. G., Nambu, M., White, J. D., Falck, J. R., Gerwick, W. H., Day, B. W., and Hamel, E. (1998) Structure-Activity Analysis of the Interaction of Curacin A, the Potent Colchicine Site Antimitotic Agent, with Tubulin and Effects of Analogs on the Growth of MCF-7 Breast Cancer Cells. *Molecular Pharmacology* **53**, 62-76

4. Chang, Z., Sitachitta, N., Rossi, J. V., Roberts, M. A., Flatt, P. M., Jia, J., Sherman, D. H., and Gerwick, W. H. (2004) Biosynthetic Pathway and Gene Cluster Analysis of Curacin A, an Antitubulin Natural Product from the Tropical Marine Cyanobacterium Lyngbya majuscula *Journal of Natural Products* **67**, 1356-1367

5. Akey, D. L., Razelun, J. R., Tehranisa, J., Sherman, D. H., Gerwick, W. H., and Smith, J. L. (2010) Crystal Structures of Dehydratase Domains from the Curacin Polyketide Biosynthetic Pathway. *Structure* **18**, 94-105

6. Reid, R., Piagentini, M., Rodriguez, E., Ashley, G., Viswanathan, N., Carney, J., Santi, D. V., Hutchinson, C. R., and McDaniel, R. (2003) A Model of Structure and Catalysis for Ketoreductase Domains in Modular Polyketide Synthases. *Biochemistry* **42**, 72-79

7. Keatinge-Clay, A. (2008) Crystal Structure of the Erythromycin Polyketide Synthase Dehydratase. *Journal of Molecular Biology* **384**, 941-953

8. Leesong, M., Henderson, B. S., Gillig, J. R., Schwab, J. M., and Smith, J. L. (1996) Structure of a Dehydratase-Isomerase from the Bacterial Pathway for Biosynthesis of Unsaturated Fatty Acids: Two Catalytic Activities in One Active Site. *Structure* **4**, 253-264

9. Schwab, J. M., Habib, A., and Klassen, J. B. (1986) A Thorough Study of the Stereochemical Consequences of the Hydration/Dehydration Reaction Catalyzed by β-Hydroxydecanoyl Thioester Dehydrase. *Journal of the American Chemical Society* **108**, 5304-5308

10. Kimber, M. S., Martin, F., Lu, Y., Houston, S., Vedadi, M., Dharamsi, A., Fiebig, K. M., Schmid, M., and Rock, C. O. (2004) The Structure of (3R)-Hydroxyacyl-Acyl Carrier Protein Dehydratase (FabZ) from Pseudomonas Aeruginosa. *Journal of Biological Chemistry* **279**, 52593-52602

11. Dirk, K., Fritz, K. W., Gerd, F., Leonardo, S., and Remo, P. (2005) The Crystal Structure of PfFabZ, the Unique β-Hydroxyacyl-ACP Dehydratase Involved in Fatty Acid Biosynthesis of *Plasmodium Falciparum. Protein Science* **14**, 1570-1580

12. Maier, T., Leibundgut, M., and Ban, N. (2008) The Crystal Structure of a Mammalian Fatty Acid Synthase. *Science* **321**, 1315-1322

13. Akey, D. L., Kittendorf, J. D., Giraldes, J. W., Fecik, R. A., Sherman, D. H., and Smith, J. L. (2006) Structural Basis for Macrolactonization by the Pikromycin Thioesterase. *Nature Chem. Biol.* **2**, 537-542

14. Helmkamp, G. M., and Bloch, K. (1969) β-Hydroxydecanoyl Thioester Dehydrase. *Journal of Biological Chemistry* **244**, 6014-6022

15. Pasta, S., Witkowski, A., Joshi, A. K., and Smith, S. (2007) Catalytic Residues Are Shared Between Two Pseudosubunits of the Dehydratase Domain of the Animal Fatty Acid Synthase. *Chemistry & Biology* **14**, 1377-1385

16. Witkowski, A., Joshi, A. K., and Smith, S. (2004) Characterization of the β-Carbon Processing Reactions of the Mammalian Cytosolic Fatty Acid Synthase: Role of the Central Core *Biochemistry* **43**, 10458-10466

17. DelProposto, J., Majmudar, C. Y., Smith, J. L., and Brown, W. C. (2009) Mocr: A Novel Fusion Tag for Enhancing Solubility that is Compatible with Structural Biology Applications. *Protein Expression and Purification* **63**, 40-49

18. Aslanidis, C., and de Jong, P. J. (1990) Ligation-Independent Cloning of PCR Products (LIC-PCR). *Nucl. Acids Res.* **18**, 6069-6074

19. Studier, F. W. (2005) Protein Production by Auto-Induction in High-Density Shaking Cultures. *Protein Expression and Purification* **41**, 207-234

20. Walker, J. M. (2005) *The Proteomics Protocols Handbook*, Humana Press, Totowa, NJ.

21. Gu, L., Wang, B., Kulkarni, A., Gehret, J. J., Lloyd, K. R., Gerwick, L., Gerwick, W. H., Wipf, P., HÃ¥kansson, K., Smith, J. L., and Sherman, D. H. (2009) Polyketide Decarboxylative Chain Termination Preceded by O-Sulfonation in Curacin A Biosynthesis. *Journal of the American Chemical Society* **131**, 16033-16035

22. Stoops, J. K., and Wakil, S. J. (1981) Animal fatty acid synthetase. A novel arrangement of the beta-ketoacyl synthetase sites comprising domains of the two subunits. *The Journal of biological chemistry* **256**, 5128

# Chapter 5

# Conclusions and Further Directions

**Thioesterase IIs**

RifR, the thioesterase II of the rifamycin PKS, was shown to remove acyl moieties from both acyl-CoAs and acyl-ACPs. Preference was shown for acyl-ACP substrates over acyl-CoAs substrates, and for decarboxylated acyl moieties over carboxylated moieties. However, the preference for decarboxylated acyl moieties became less pronounced when the acyl moieties were presented as acyl-ACPs instead of as acyl-CoAs. This suggests the ACP may play a role in substrate specificity. Genetic sequences of the rifamycin ACPs are available, and future plans to further characterize the substrate specificity of RifR should include producing excised ACPs of the rifamycin pathway to test for ACP specificity.

Investigations into the ACP preference of other type I PKS TEIIs have been limited (1). The TEII of the erythromycin pathway showed little preference for acyl moieties presented by the erythromycin module 1 ACP over those presented by the erythromycin module 2 ACP (1). Because TEIIs are predicted to scan the biosynthetic pathway to remove acyl units from multiple modules, it is presumed these TEIIs will show little ACP specificity. Therefore, RifR should be able to recognize and hydrolyze acyl moieties presented by each ACP of the rifamycin pathway. Testing the activity of

RifR on acyl moieties presented on various Rif ACPs is needed to validate this hypothesis. If RifR is highly specific for one or two ACPs of the rifamycin biosynthetase, it would suggest that RifR is not able to scan the entire biosynthetase and the house keeping function associated with TEIIs, as described in chapter 2, may be inaccurate.

Like the TEII of the NRPS surfactin pathway, and the TEIs of the NRPS pathways surfactin, fengycin, and enterobactin, RifR appears to undergo a conformation change to allow substrates to enter the active site. We hypothesized the flexible loop region of the lid allows for an expansion of the substrate chamber, creating room for the medium chain acyl moieties (C4-C10) (2). To verify this hypothesis, co-crystallization trials of RifR with various acyl-CoA substrates were performed. However our efforts did not yield crystals with an acyl-CoA substrate in the active site, possibly because dithiothreitol (DTT) was included in the crystallization solution. The thiols of DTT were later found to interact with some acyl-CoAs, causing hydrolysis of the acyl unit from the Ppant arm of the CoA. It is possible the precipitation observed upon addition of the acyl-CoA to the crystallization drop may be due to the removal of the acyl moiety from the Ppant of the acyl-CoA by DTT. Therefore, future co-crystallization attempts should be performed in the absence of DTT.

Substrate specificity of TEIIs can vary from pathway to pathway, and though there have been numerous investigations into the substrate specificity of many TEIIs (1-9), there is little data that allows a direct comparison of the substrate specificity of TEIIs (Tables 2.4, 2.5). This has made it difficult to identify, by sequence alignment, possible active site features or conserved residues that may contribute to the substrate specificity of a TEII. Without knowing the position of the acyl moiety in the active site, it is difficult to determine the structural features that are responsible for substrate specificity. Therefore, further investigations into the acyl-specificity of the thioesterase IIs will rely on the information garnered from a co-crystal structure with a small and medium length substrate in the active-site.

**Type I Ketoreductases**

All nine KR-ACP didomains were shown to be active through reduction of the non-natural substrates *trans*-1-decalone and cyclohexanone. In later reactions, with more accurate amounts of non-natural substrate, a lag was observed in the Michaelis-Menten plots. As described in chapter 3, two hypotheses were provided for this lag: experimental error or cooperativity between the NADPH and substrate binding sites.

Experimental error from inaccurate pathlength measurements and/or inaccurate initial reaction temperatures may be responsible for the lag observed in the Michaelis-Menten plots. Activity assays were performed in 96-well plates in a Softmax Pro5 Flexstation 3 (Molecular Devices). The reaction was initiated by addition of *trans*-1-decalone or cyclohexanone. After the absorbance was measured at 340 nm for 55 min at 30°C, the pathlength of each reaction well was measured. It is possible that significant evaporation at 30°C, enough to affect the pathlength of the reaction well, may lead to an inaccurate absorbance values of NADPH after the raw data has been corrected to a pathlength of 1 cm. Reactions in which either enzyme or substrate were withheld showed a slight decrease in absorbance over time—presumably due to the instability of NADPH at high temperatures. Additionally, later reactions did not include a pre-incubation of the reactants at 30°C, nor were *trans*-1-decalone or cyclohexanone heated to 30°C before addition to the reaction. Both of these factors may produce inaccurate initial velocities, skewing the Michaelis-Menten plots. To address these errors in the future, the activity assays should be pre-incubated at 30°C and should be performed in single cuvette format, instead of in 96-well format. The cuvette will keep a constant pathlength throughout the reaction and pre-heating the assay components will provide more accurate initial velocities. If a lag is still observed in the Michaelis-Menten plots of KRACP didomain activity after the above mentioned experimental errors have been addressed, it may be concluded that the lag is not due to such experimental factors.

Preliminary data suggest that some Fecik labels are able to compete with the non-natural substrates *trans*-1-decalone and cyclohexanone. Presumably, this competition is the result of the Fecik label entering and occluding the entrance to the active site. According to this hypothesis, increased competition should correlate to increased

occupancy of the Fecik label in the active site. These data will be especially helpful for future co-crystallization trials as Fecik labels with higher occupancy would be more likely to produce crystals with the Fecik label in the active site. By obtaining the position of the Fecik label in the active site through crystallization, we hope to identify how stereospecificity is conferred by the KR domain. Specifically, we are interested in how the 'LDD' or the 'W' motifs discussed in chapter 3 affect the stereochemistry of the product.

While crystallization trials are underway, the Fecik labels may provide additional information about the nature of KR active sites. EKL-103, the smallest of the Fecik labels, (Fig 3.3) competed with the non-natural substrates in both Tyl KR1ACP1 and Tyl KR7ACP7. Presumably, EKL-103 competed because its small size allowed it to enter the active site of Tyl Kr1ACP1, while other Fecik labels were too large. To test this hypothesis, medium and long chain Fecik labels (P-Kr2-P2, P-Kr5-P3) should be modeled into the active site of TylKR1 to determine any features that might hinder these substrate mimics from entering the active site. If such features are identified in this manner, further competition assays with Fecik labels and KRACP didomains that have no solved structure available would prove a useful tool in predicting the success of various Fecik label-protein co-crystallization combinations.

For TylKR7ACP7 both EKL-103 and P-Kr5-P3 competed with the non-natural substrate cyclohexanone. These results were unexpected in that both the smallest and longest Fecik labels competed with cyclohexanone, while the middle chain length Fecik label (P-Kr2-P2) did not. Unlike TylKR1ACP1, size does not appear to be a factor in determining whether a Fecik label is able to enter the KR active site. P-Kr5-P3 and P-Kr2-P2 differ not only in chain length, but also in the stereochemistry of the β alcohol. The β-alcohol of P-Kr5-P3 is in the S-like configuration while the β-alcohol of P-Kr2-P2 is in the R-like configuration. The stereoconfiguration of KR products is discussed in chapter 3. TylKR7ACP7 is an A-type ketoreductase domain, which means it produces an alcohol in the S-like configuration (10). To determine if the configuration of the β-alcohol affects whether a Fecik label can enter the active site of TylKR7ACP7, more competition assays should be performed. Specifically, a competition assay testing the

ability of the Fecik label P-Kr2-P1 to compete with cyclohexanone is needed. P-KR2-P1 is similar to the non-competing Fecik label, P-Kr2-P2, except the P-KR2-P2 Fecik label has the β-alcohol in the opposite configuration. If the S-like P-KR2-P1 competes with cyclohexanone, while R-like P-KR2-P2 does not, it would suggest that TylKR7ACP7 is only able to accommodate products of the appropriate stereochemistry in the active site.

The active sites of the A-type ketoreductase AmpKR2, and the B-type ketoreductases TylKR1, and EryKR1 are very similar (11). The 'LDD' and 'W' motifs are predicted to control the stereochemical outcome of the reduction by controlling the placement of the substrate in the active site (11,12). However, there are no obvious structural features that would stop a substrate of the opposite stereoconfiguration from entering the active site. If further study of the TylKR7ACP7 finds that only Fecik labels of the S-like configuration are able to enter the active site, it would suggest that some KR domains do contain structural features that preclude a product of the opposite configuration from entering the active site.

**Curacin Dehydratases**

Now that a working assay to measure curacin dehydratase activity is available, a number of questions regarding substrate specificity, stereochemistry, and active site interactions, may be investigated. Further optimization of the assay may provide information about the reaction mechanism of these dehydratases. For instance, the reason dehydratase activity was observed in a phosphate buffered system but not in Tris-HCl buffered system remains unknown. Investigating whether phosphate plays a direct or indirect role in catalysis would garner information about the type I PKS dehydratase reaction mechanism.

To determine optimal activity conditions for the curacin DHs, the activities of the fastest, CurJ DH, and slowest, CurH DH, should be tested under a variety of conditions. Specifically, curacin DH activity should be tested in different buffering system such as HEPES, sodium phosphate, sodium cacodylate, sodium barbital, Bis-Tris Propane, TAPSO and Tris-HCl. If the curacin DHs are only active in the strongly ionic buffers (sodium phosphate, sodium cacodylate, or sodium barbital) it is probable that DH

domains are dependent on the ionic strength of the solution. If the DH domains are inactive in Tris-HCl, but active in the other buffers, it suggests that the Tris molecule is interfering with the dehydratase reaction. For instance, Tris may be entering and occluding the active site. If the DH domains are only active in the phosphate buffers, or active in solutions that have phosphate added, it would suggest phosphate plays a critical, previously unknown role in catalysis. Once a favored buffering system has been determined, the pH and ionic strength of the solution may also be varied to determine optimal conditions for the dehydratase assay.

The product of the curacin PKS/NRPS suggests five dehydration reactions occur, yet the curacin PKS contains only four dehydratases, one of which is located in a module that is not predicted to catalyze a dehydration reaction (13). Two hypotheses have been provided for the orphan dehydration reactions. One hypothesis suggests the curacin DHs are able to act on intermediates in trans, meaning a Cur DH is able to act on an intermediate tethered to another ACP in the biosynthase (13). Another hypothesis suggests the curacin DHs are able to dehydrate an alcohol not located at the β-position. To test these hypotheses, information on the ACP specificity of the Cur DHs is needed. Excised ACPs of the CurG, CurJ, CurH, CurI, and CurK modules are available, allowing an investigation into Cur DH acyl-ACP specificity.

β-hydroxyl-butyryl or crotonyl moieties may be loaded onto Cur ACPs and used as substrates to test DH acyl-ACP specificity. If all the Cur DHs were found to be highly specific for acyl-substrates tethered to the ACP of the respective module, it would disprove the first hypothesis that a Cur DH may dehydrate intermediates in trans. If the Cur DHs show no preference for acyl moieties presented by one ACP over another, it would support the hypothesis that Cur DHs may recognize and act on ACPs of different modules. Two modules, CurI and CurG do not contain a Cur DH, yet dehydration is predicted to occur on intermediates tethered to these modules. Identification of the Cur DH that is able to recognize and dehydrate a substrate attached to one of these modules would support the hypothesis that Cur DH domains are able to work on the intermediates in trans.

If no Cur DH is able to act on substrates presented by CurIACP, or CurGACP, it would suggest that dehydration is not carried out while the intermediate is tethered to the CurI or CurG module. In this instance, the Cur DH domains should be tested for dehydratase activity on alcohol groups not in the β position as the remaining alcohol may be dehydrated while attached to a subsequent module. To test this hypothesis, acyl moieties with a δ-alcohol should be loaded onto CurH ACP or CurJ ACP. If either CurH DH or CurJ DH is able to dehydrate the δ-alcohol of these acyl-ACP substrates, it would suggest these subsequent Cur DH domains are responsible for the double bonds initially associated with the CurI and CurG modules. If no Cur DH is able to recognize acyl-CurG ACP or acyl-CurI ACP substrates, and no Cur DH is able to dehydrate an alcohol in the δ-position, the dehydration reactions predicted to occur on the CurG and CurI intermediates must be performed by a different protein, previously unassociated with the PKS/NRPS.

Curacin A contains an uncommon *cis* double bond, presumably created on the CurG intermediate. Previous hypotheses have suggested that a *cis* bond is produced through dehydration of a β-alcohol in the S-like configuration (14). This hypothesis was further supported when substrates of both stereoconfiguration were modeled into the active sites of the Cur DHs (13). The configuration of the β-hydroxyl positioned the substrate such that the S-like substrate adopted a *cis*-like conformation, and the R-like substrate adopted a *trans*-like conformation (13). By sequence, the CurG KR is predicted to be an A-type KR, which are known to exclusively produce β-alcohols in the S-like configuration as described in chapter 3. To test the hypothesis that a dehydratase may accommodate and dehydrate a substrate of either stereoconfiguration, all four Cur DHs should be tested for activity with an S-like β-hydroxyl-acyl-CurG ACP or β-hydroxyl-acyl-CoA. Chiral HPLC may be used to identify a *cis* vs *trans* double bond, or an R-like vs S-like β-hydroxyl-acyl-CoA. The results of these experiments will determine if a Cur DH is able to accommodate both an S-like β–hydroxyl-acyl moiety, and an R-like β–hydroxyl-acyl moiety, and to verify the stereoconfiguration of the β-hydroxyl controls the stereoconfiguration of the product.

To further investigate active site interactions of the curacin DHs, a variety of plasmids encoding Cur DH active site mutants, created by Jamie Razelun and Jason Tehranisa, are available (Table 5.1). Testing the activity of the numerous active-site mutants with acyl-ACPs or acyl-CoA would verify the putative catalytic residues of each DH. Substrate modeling in the active site suggests the β-alcohol forms a hydrogen bond to the catalytic Asp (13). The catalytic His abstracts the α-proton resulting in either a *cis* or *trans* double bond, depending on whether the pro-S or pro-R proton is removed (13,15). By comparing the activities of the various DH mutants, insights into the reaction mechanism may be obtained.

Table 5.1 Cur DH active site mutants

| CurF DH D1755N | CurH DH H971A | CurJ DH H978A |
|---|---|---|
| CurF DH D1893N | CurH DH D1000N | CurJ DH D1012N |
| CurF DH D1949N | CurH DH D1199N | CurJ DH D1156N |
| CurF DH H1720N | CurH DH D1136N | CurJ DH D1220N |
| CurF DH H1720Q | CurH DH H971Q | CurJ DH H978N |
|  | CurH DH H971N | CurJ DH H978Q |

**Applications for Designer Biosynthases**

Identifying structural features that confer substrate specificity of a type 1 PKS domain is an initial step towards creating designer biosynthases. For instance, the structure of RifR suggests the broad substrate range and ability to accept medium chain length substrates is due to the flexible lid loop and a lack of highly charged residues in the active site. By tweaking either of these structural features, one may potentially change the substrate specificity of RifR to remove only the desired acyl units from a PKS. With a working DH activity assay, further investigations into the substrate specificity of type I PKS DH domains may commence. And, a better understanding of

how ketoreductase domains are able to control the configuration of their products may result in ketoreductases with lax stereochemical control, something that will be useful when creating libraries of compounds that differ only in the configuration at each chiral center. Further characterization of these domains will promote development of newer, more potent drugs and a wider variety of natural product derivatives.

**Work Cited**

1. Hu, Z., Pfeifer, B. A., Chao, E., Murli, S., Kealey, J., Carney, J. R., Ashley, G., Khosla, C., and Hutchinson, C. R. (2003) A specific role of the Saccharopolyspora erythraea thioesterase II gene in the function of modular polyketide synthases. *Microbiology* **149**, 2213-2225

2. Claxton, H. B., Akey, D. L., Silver, M. K., Admiraal, S. J., and Smith, J. L. (2009) Structure and functional analysis of RifR, the type II thioesterase from the rifamycin biosynthetic pathway. *J Biol Chem* **284**, 5021-5029

3. Witkowski, A., Naggert, J., Witkowska, H. E., Randhawa, Z. I., and Smith, S. (1992) Utilization of an active serine 101----cysteine mutant to demonstrate the proximity of the catalytic serine 101 and histidine 237 residues in thioesterase II. *Journal of Biological Chemistry* **267**, 18488-18492

4. Heathcote, M. L., Staunton, J., and Leadlay, P. F. (2001) Role of type II thioesterases: evidence for removal of short acyl chains produced by aberrant decarboxylation of chain extender units. *Chemistry & Biology* **8**, 207-220

5. Yeh, E., Kohli, R. M., Bruner, S. D., and Walsh, C. T. (2004) Type II thioesterase restores activity of a NRPS module stalled with an aminoacyl-S-enzyme that cannot be elongated. *ChemBioChem* **5**, 1290-1293

6. Kim, B. S., Cropp, T. A., Beck, B. J., Sherman, D. H., and Reynolds, K. A. (2002) Biochemical Evidence for an Editing Role of Thioesterase II in the Biosynthesis of the Polyketide Pikromycin. *J. Biol. Chem.* **277**, 48028-48034

7. Tang, Y., Koppisch, A. T., and Khosla, C. (2004) The acyltransferase homologue from the initiation module of the R1128 polyketide synthase is an acyl-ACP thioesterase that edits acetyl primer units. *Biochemistry* **43**, 9546-9555

8. Linne, U., Schwarzer, D., Schroeder, G. N., and Marahiel, M. A. (2004) Mutational analysis of a type II thioesterase associated with nonribosomal peptide synthesis. *Eur. J. Biochem.* **271**, 1536-1545

9. Schwarzer, D., Mootz, H. D., Linne, U., and Marahiel, M. A. (2002) Regeneration of misprimed nonribosomal peptide synthetases by type II thioesterases. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 14083-14088

10. Caffrey, P. (2003) Conserved Amino Acid Residues Correlating with Ketoreductase Stereospecificity in Modular Polyketide Synthases. *Chembiochem* **4**, 654-657

11. Zheng, J., Taylor, C. A., Piasecki, S. K., and Keatinge-Clay, A. T. (2010) Structural and Functional Analysis of A-Type Ketoreductases from the Amphotericin Modular Polyketide Synthase. *Structure* **18**, 913-922

12. Keatinge-Clay, A. T. (2007) A Tylosin Ketoreductase Reveals how Chirality is Determined in Polyketides. *Chem Biol* **14**, 898-908

13. Akey, D. L., Razelun, J. R., Tehranisa, J., Sherman, D. H., Gerwick, W. H., and Smith, J. L. (2010) Crystal Structures of Dehydratase Domains from the Curacin Polyketide Biosynthetic Pathway. *Structure* **18**, 94-105

14. Reid, R., Piagentini, M., Rodriguez, E., Ashley, G., Viswanathan, N., Carney, J., Santi, D. V., Hutchinson, C. R., and McDaniel, R. (2003) A Model of Structure and Catalysis for Ketoreductase Domains in Modular Polyketide Synthases. *Biochemistry* **42**, 72-79

15.    Schwab, J. M., Habib, A., and Klassen, J. B. (1986) A Thorough Study of the Stereochemical Consequences of the Hydration/Dehydration Reaction Catalyzed by β-Hydroxydecanoyl Thioester Dehydrase. *Journal of the American Chemical Society* **108**, 5304-5308