5-2013

Regulation of Secondary Metabolism in *Lysobacter enzymogenes*: Studies of Intercellular and Intracellular Signaling

Stephen J. Wright
University of Nebraska – Lincoln, sjwright@huskers.unl.edu

REGULATION OF SECONDARY METABOLISM IN *LYSOBACTER ENZYMOGENES*: STUDIES OF INTERCELLULAR AND INTRACELLULAR SIGNALING

by

Stephen J. Wright

A THESIS

Presented to the faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the degree of Master of Science

Major: Chemistry

Under the Supervision of Professor Liangcheng Du

Lincoln, Nebraska

May, 2013

REGULATION OF SECONDARY METABOLISM IN *LYSOBACTER ENZYMOGENES*: STUDIES OF INTERCELLULAR AND INTRACELLULAR SIGNALING

Stephen J. Wright, M.S.

University of Nebraska, 2013

Advisor: Liangcheng Du

The organisms of the genus *Lysobacter* have been recognized as prolific producers of bioactive secondary metabolites, making them potentially valuable as biocontrol agents and as sources of compounds for drug leads. This study was aimed at understanding the regulatory mechanisms that underlie the production of secondary metabolites in our study organism, *Lysobacter enzymogenes*. Since secondary metabolism is energetically costly, we sought not only to elucidate the biosynthetic chemistry by which the bioactive molecules were constructed, but also the regulation of the biosynthetic machinery. The molecular mechanisms by which *L. enzymogenes* responds to environmental conditions and transduces signals leading to secondary metabolism has hitherto been almost entirely unexplored. In this thesis, we show how the tools of molecular biology and analytical chemistry have been used to investigate the regulatory mechanisms of this valuable organism.

Dedication

They that go down the sea in ships, that do business in the great waters, behold the works of the LORD and His wonders in the deep. For He commandeth, and raiseth the stormy wind, which lifteth up the waves thereof; they mount up to the heaven, they go down again to the depths; their soul is melted because of trouble; they reel to and fro, and stagger like a drunkard, and are at their wits' end. Then they cry unto the LORD in their trouble, and He bringeth them out of their distresses. He maketh the storm a calm, so that the waves thereof are still. Then are they glad because they be quiet; so He bringeth them unto their desired haven. O that they would praise the LORD for His goodness, and for His wonderful works to the children of men! Let them exalt Him also in the congregation, and praise Him in the assembly of the elders. (Psalm 107:23-32, KJV)

Acknowledgments

Sincerest thanks are due to my advisor, Liangcheng Du, for possessing and demonstrating the inimitable combination of patience, encouragement, kindness and leadership throughout this project. Thanks also to the members of his research group for instruction, advice, and warnings; to the staff of Nebraska Center for Mass Spectrometry, especially Dr. Kurt Wulser, for technical assistance; to Dr. Martha Morton and Sara Basiaga of the Department of Chemistry Instrumentation Center; and to Teresa Fangman and the staff of the Morrison Microscopy Core Research Facility.

Table of Contents

List of Figures

List of abbreviations

DSF: diffusible signal factor

cAMP: cyclic adenosine monophosphate

c-di-GMP: cyclic-di-guanosine monophosphate

CRP: cAMP-receptor protein

Clp: CRP-like protein

KR: β-ketoreductase

DH: dehydratase

KS: β-ketosynthase

ACP: acyl carrier protein

AT: acyltransferase/acetyltransferase

HPLC: high-pressure liquid chromatography

NMR: nuclear magnetic resonance

TLC: thin-layer chromatography

PCR: polymerase chain reaction

*Le*C3: *Lysobacter enzymogenes* strain C3

*Le*OH11: *Lysobacter enzymogenes* strain OH11

*Le*DC: clp-deletion mutant of *Le*C3

*Le*DCA: clp- and acetyltransferase-deletion mutant of *Le*C3

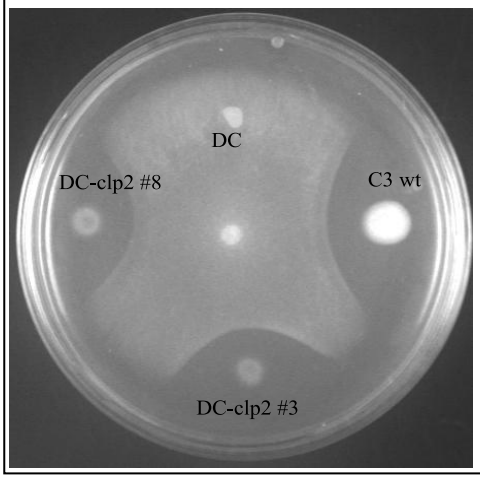
Xcc: *Xanthomonas campestris* pv. *campestris*

Xoo: *Xanthomonas oryzae* pv. *oryzae*

Chapter 1: Introduction

The genus *Lysobacter* was proposed in 1978 to encompass a number of gliding, Gram-negative, high G+C content bacteria which, unlike the myxobacteria, were not known to produce fruiting bodies.[16-17] One of the most notable features of the lysobacters was their production of potent lytic enzymes and secondary metabolites and their resultant capacity to lyse a wide variety of prokaryotic and eukaryotic organisms: both Gram-negative and Gram-positive bacteria (including the Gram-positive actinomycetes), yeasts, filamentous fungi, and nematodes fall within the scope of their predation. [16, 18-19] Although a large number of bioactive secondary metabolites from *Lysobacter* have been isolated [16, 20], their biosynthetic mechanisms have only recently been explored, and very little has been reported on the regulation of *Lysobacter*'s secondary metabolism. Our quest to understand the molecular logic behind *Lysobacter enzymogenes*'s secondary metabolism began with a transposon mutant of *L. enzymogenes* C3, dubbed *Le* 5E4, which exhibited decreased gliding motility, extracellular lytic enzyme production, and antimicrobial activity. It was shown that the transposon's insertion prevented the 5e4's transcription of two genes, separated by only 6 bases, which were predicted to encode a CRP-like protein (Clp) and a Gcn5-like *N*-acetyltransferase (AT), respectively.[3] Complementation of mutant 5E4 with a chromosomally-inserted copy of the *clp* gene mostly restored the wild-type

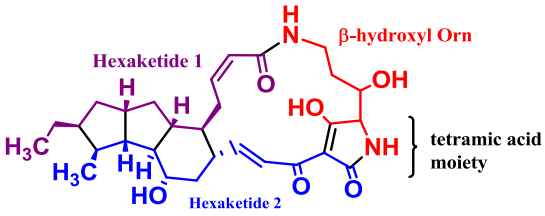
Figure 1: Activity of *Le*C3 wild-type and two *Le*DC Clp complementation strains against *Fusarium verticilliodes*. Clp-deletion mutant *Le*DC [1, 3] shows no activity. (Y. Xie, unpublished data).



phenotype, confirming the central role of the Clp protein in regulating *L. enzymogenes* C3's antimicrobial properties.

L. enzymogenes C3 exhibits potent antifungal activity, which was shown to result at least in part from a heat-stable antifungal factor (HSAF) which Yu *et al*. isolated and identified as *dihydromaltophilin* **(**Figure 2**)**, one of a family of bioactive polycyclic tetramate macrolactams (PTMs) isolated from a variety of marine and terrestrial sources. [15-16, 21] Yu *et al*. also identified dihydromaltophilin's biosynthetic gene cluster in *L. enzymogenes* C3, representing the first biosynthetic genes reported for a natural product in the polycyclic tetramate macrolactam family. Importantly, it was observed that the transposon mutant *L. enzymogenes* 5E4, as well as the Clp-deletion mutant *L. enzymogenes* DC, failed to produce dihydromaltophilin and also lacked concomitant antifungal activity (**Figure 1**). [12] We hypothesized that the up-regulation of dihydromaltophilin's biosynthetic genes, and perhaps the up-regulation of other biosynthetic gene clusters in *L.enzymogenes*, depended on the activity of the transcription factor, Clp.

Figure 2: dihydromaltophilin
(heat-stable antifungal factor, HSAF) [15]

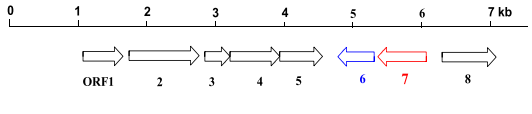


1.1 Overview of studies of transcription factor Clp in the xanthomonads

The *clp* homologous gene in both *L. enzymogenes* C3 (*Le*C3) and *L. enzymogenes* OH11 (*Le*OH11) exhibits strong homology to the cAMP-receptor-like protein (*clp*) gene discovered in the plant pathogen *Xanthomonas campestris* pv. *campestris* (*Xcc*) (**Figure 3**).[22]

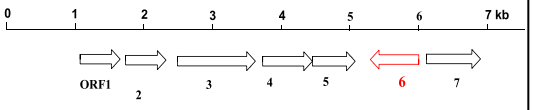
Figure 3: Annotation of *clp* homologues and loci. (A) *clp* and its surrounding genes in *L. enzymogenes* OH11 (B) *clp* and its surrounding genes in *Xanthomonas campestris* pv. *campestris* B100. C) ClustalW alignment of *clp* homologues, with key residues highlighted.[2-4, 10-11]

A) Clp locus in *Le*OH11



ORF	# aa	Blastx Homolog	Identity/similarity	E value
ORF1	193	Anthranilate synthases component II	82/92%	1e-87
ORF2	342	Anthranilate phosphoribosyltransferase	80/89%	3e-149
ORF3	113	Hypothetical protein (putative monooxygenase)	68/78%	5e-35
ORF4	263	Indole-3-glycerol phosphate synthase	71/80%	3e-94
ORF5	250	Possible phosphoserine phosphatase	69/79%	3e-85
ORF6	153	putative *N*-acetyltransferase	41/53%	2e-16
ORF7	229	Clp, cAMP-receptor-like protein	80/86%	2e-96
ORF8	264	*S*-adenosylmethionine decarboxylase	91/97%	1e-135

B) Clp locus in *Xcc* B100



ORF	# aa	predicted function	Genbank accession
ORF1	192	Anthranilate synthases component II	YP_001901903.1
ORF2	210	conserved protein (putative oxygenase)	YP_001901904.1
ORF3	345	Anthranilate phosphoribosyltransferase	YP_001901905.1
ORF4	279	Indole-3-glycerol phosphate synthase	YP_001901906.1
ORF5	238	Possible phosphoserine phosphatase	YP_001901907.1
ORF6	230	Clp, cAMP-receptor-like protein	YP_001901908.1
ORF7	283	*S*-adenosylmethionine decarboxylase	YP_001901909.1

C)

```
OH11 gi|119393855|gb|ABL74444.1|      MSA-NPVAPLTALRRTNSPLLPDGATIERFLAHCHRRRYPSRTDVFRPGD 49
LeC3 gi|32454724|gb|AAP83141.1|       MSA-NPVAPLTALRRTNSPLLPDGATIERFLAHCHRRRYPSRTDVFRPGD 49
Xcc  gi|66766825|ref|YP_241587.1|     MSLGNTTVVTTTVRNATPSLTLDAGTIERFLAHSHRRRYPTRTDVFRPGD 50
                                      **  *...  *::*.:...*  *..********.******:*********

OH11 gi|119393855|gb|ABL74444.1|      PASTLYYVVSGSVSIITEEEDGRELVLGYFGPGEFVGELGLFIESDHREV 99
LeC3 gi|32454724|gb|AAP83141.1|       PASTLYYVVSGSVSIITEEEDGRELVLGYFGPGEFVGELGLFIESDHREV 99
Xcc  gi|66766825|ref|YP_241587.1|     PAGTLYYVISGSVSIIAEEDDDRELVLGYFGSGEFVGEMGLFIESDTREV 10
                                      **.*****:*******:**:*.*********.******:******* ***

OH11 gi|119393855|gb|ABL74444.1|      ILRTRSTCELAEIGHERLYDLLLTRLSLDAPKLLYAIGAQISRRLLDTSR 14
LeC3 gi|32454724|gb|AAP83141.1|       ILRTRSTCELAEIGHERLYDLLLTRLSLDAPKLLYAIGAQISRRLLDTSR 14
Xcc  gi|66766825|ref|YP_241587.1|     ILRTRTQCELAEISYERLQQLFQTSLSPDAPRILYAIGVQLSKRLLDTTR 15
                                      *****: ******.:*** :*: * ** ***::*****.*:*:*****:*

OH11 gi|119393855|gb|ABL74444.1|      KAGRLAFLDVTDRIVRALHDLAKEPEAMSHPQGTQLRVSRQELSRLVGCS 19
LeC3 gi|32454724|gb|AAP83141.1|       KAGRLAFLDVTDRIVRALHDLAKEPEAMSHPQGTQLRVSRQELSRLVGCS 19
Xcc  gi|66766825|ref|YP_241587.1|     KASRLAFLDVTDRIVRTLHDLSKEPEAMSHPQGTQLRVSRQELARLVGCS 20
                                      **.*************:****:**********************:******

OH11 gi|119393855|gb|ABL74444.1|      REMAGRVLKKLQADGKLHARGKTVVLYGTR 229
LeC3 gi|32454724|gb|AAP83141.1|       REMAGRVLKKLQADGKLHARGKTVVLYGTR 229
Xcc  gi|66766825|ref|YP_241587.1|     REMAGRVLKKLQADGLLHARGKTVVLYGTR 230
                                      *************** ***************
```

The protein products of the *clp* homologues are transcription factors which, in *Xcc*, participate in a complex system of regulation involving a small-molecule intracellular diffusible signal factor (DSF), a two-component system, a small-molecule second messenger (cyclic-di-GMP), the transcription factor Clp, and the proteins encoded in its regulon. This remarkable system has been extensively studied and reported in the literature, and has been included in a number of reviews [14, 23-26] describing bacterial intercellular signalling via small molecules. This general process of intercellular communication is often called quorum sensing, as the signals are dependent on the concentration of bacterial cells which secrete the molecular signals into the environment.

In 1997 it was observed that *Xcc*'s wild-type phenotype could be restored to *Xcc* strains with inactivated genes in the *rpf* (regulation of pathogenicity factors) locus, by streaking the mutant strains adjacent to, but not touching, any other strain of *Xcc*, with the exception of *Xcc* mutants of the *rpfF* or *rpfB* gene.[27] These observations were consistent with the synthesis of a small molecule diffusible signal by the gene products of *rpfF* (predicted to be an enoyl-CoA hydratase) and *rpfB* (an acyl-CoA ligase), and the secretion and diffusion of the signal molecule to adjacent cells, where their detection led to the upregulation of genes associated with *Xcc*'s pathogenicity. Later investigations elucidated this signaling pathway in *Xcc*. The diffusible signal factor (DSF) was identified as *cis*-11-methyl-2-dodecenoic acid,[28] and DSF's membrane-bound sensor kinase RpfC transduces the signal by phosphorylating the intracellular effector RpfG, a phosphodiesterase which hydrolyzes the second messenger bis-(3',5')-cyclic dimeric guanosine monophosphate (cyclic-di-GMP or c-di-GMP).[9, 29-32] Decrease in the intracellular concentration of c-di-GMP leads to upregulated transcription of genes involved in production of virulence factors [31] and other genes such as those involved in

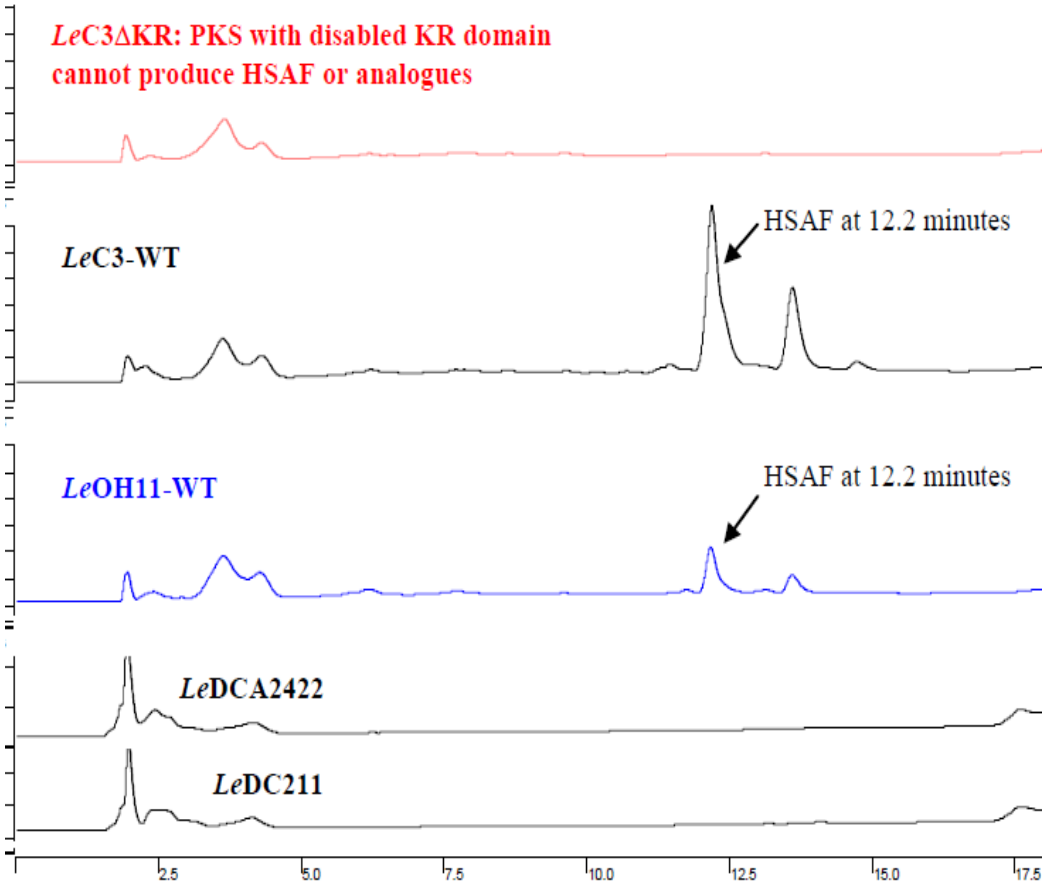
resistance to oxidative stress and in flagellum synthesis.[30] This upregulated transcription is, at least in part, the work of Clp, a novel transcription factor which, like its namesake cAMP-receptor protein (CRP), recognizes a consensus DNA sequence TGTGA-N6-TCACA,[31] but unlike CRP, is negatively regulated by its small molecule second messenger.[1] When bound by c-di-GMP, the Clp dimer adopts an asymmetric conformation which does not bind its cognate DNA sequence; as intracellular c-di-GMP concentrations fall, the Clp dimer is freed to adopt its preferred, symmetrical conformation, in which it can enhance transcription of numerous genes.[1, 31] Key residue differences between Clp and CRP allow the former to adopt its DNA-binding conformation in the absence of any small-molecule effector,[33] and the solved crystal structure of the Clp dimer as well as *in silico* docking studies of Clp with c-di-GMP allowed the identification of the residues involved in c-d-GMP binding and in stabilization of the symmetric dimer in the absence of c-di-GMP.[1] Mutagenesis studies confirmed the roles of these key residues, as did the *in vitro* binding of symmetric Clp dimer to a promoter known to be upregulated by DSF completed the pathway.[1, 34] Additionally, Clp, or one of the transcription factors in its regulon, may upregulate its own transcription, adding an additional layer of autoinduction to the system.[31] One interesting and important aspect of the DSF/Clp signalling system involves the autosuppression of DSF synthesis effected by RpfC, the membrane-bound sensor kinase. In addition to possessing a phosphorelay domain which ultimately phosphorylates the phosphodiesterase RpfG, RpfC includes an intracellular receiver (REC) domain which was definitively shown to interact with the DSF synthase RpfF, isolating RpfF from its presumed enoyl-CoA substrate and preventing DSF production.[9] This interesting

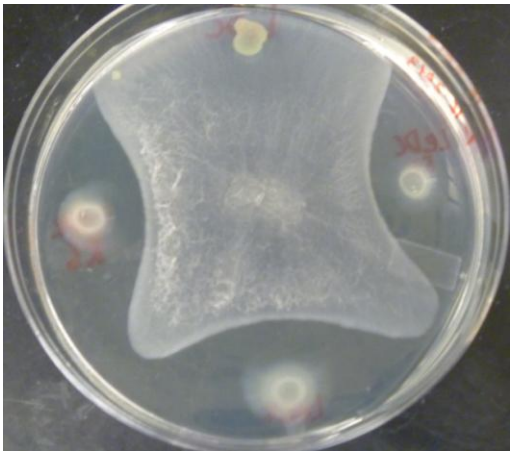
negative regulation of DSF synthesis via RpfC/RpfF interaction was further demonstrated by the crystallization of RpfF, and *in vivo* testing of the RpfC REC domain's interaction with RpfF via a bacterial two-hybrid assay.[8] Mutational analyses of the conserved residues involved in this interaction, as well as overexpression of RpfC's REC domain,[8-9] confirmed that RpfC negatively regulates RpfF and thus DSF synthesis, and explained the observation that RpfC null mutants overproduce DSF.[9]

Although the *rpf*-Clp signaling system has been best studied in *Xcc*, other xanthomonads share conserved homologous genes similar to the *rpf* cluster and employ a similar system of quorum sensing and resulting gene upregulation,[14, 35] and diffusible signal molecules similar to DSF (*cis*-11-methyl-2-dodecenoic acid) have been isolated from species unrelated to the xanthomonads, including *Burkholderia cenocepacia* and *Pseudomonas aeruginosa*,[35] showing that small fatty acids may be a common method of intercellular communication among the Gram-negative bacteria.

On the basis of the signaling system described in the *Introduction* and the observation that loss of its *clp* homologue abolished dihydromaltophilin production and antifungal activity in *Lysobacter enzymogenes* (**Figure 4**), we constructed a hypothesis of intercellular signaling in *Lysobacter enzymogenes* by which DSF, or a similar signaling molecule, served as the initiator for an intracellular cascade that leads to Clp's upregulation of the biosynthetic genes for dihydromaltophilin, either by direct interaction with the promoter of the biosynthetic genes, or by upregulating transcription factors which in turn upregulated the biosynthetic genes.

Figure 4: Effects of loss of *clp* on dihydromaltophilin (HSAF) production and antifungal activity.





Top) Ethyl acetate extractions of dihydromaltophilin from *L. enzymogenes* strains grown in NYGB in which 109 mM glucose replaced glycerol as a carbon source. The two *clp* deletion strains, DCA2422 and DC211,[3] were grown at a later date but under identical conditions. See row 3 of Table 2 for HPLC conditions.

Bottom) antagonism of *L. enzymogenes* C3 strains against *Fusarium verticillioides*. Wild-type and *clp* complementations of DC211 exhibit strong antifungal activity, which *Le*DC211 entirely lacks.

1.2 Genetic evidence for a novel *rpf* system in *Lysobacter enzymogenes*

1.2.1 Clp homologues

The observation that Clp, a transcription factor known to influence antimicrobial activity and production of extracellular lytic enzymes in *Le*C3,[3] also appeared to be essential for the production of the heat-stable antifungal dihydromaltophilin, suggested the existence of an *rpf* system in *L. enzymogenes* which operates in a manner analogous to *Xcc* and the other xanthomonads in which diffusible signal factors had been reported. Since the *rpf* system appears to be essential for *Xcc* and related plant pathogens to carry out their life-cycles of transmission and host infection,[27, 36-37] we postulated that *L. enzymogenes* upregulated genes related to its antimicrobial capabilities in response to signals transduced via the *rpf* two-component system and the transcription factor Clp. Indeed, we considered the possibility that Clp might serve as a "master switch" for *L. enzymogenes*' secondary metabolism, allowing this bacterial predator to coordinate attacks *en masse* on potential food sources.

Genomic data for *Le*C3, in which dihydromaltophilin's biosynthetic genes were initially identified, is limited to small nucleotide sequences surrounding the dihydromaltophilin gene cluster[15] and the *clp* homologue.[3] The genes surrounding *Le*C3's *clp* homologue are remarkably similar to the arrangement surrounding *Xcc*'s *clp* gene (**Figure 1**), with the notable addition to *Le*C3 of a predicted acetyltranferase gene, immediately downstream and transcribed in the same direction as *clp* and shown by Koboyashi *et al*. to be transcribed along with *clp* on the same mRNA transcript, suggesting that these two gene products have some functional link.[3]

In **Figure 5**, a sequence alignment[2] of the Clp homologues from the two *L. enzymogenes* strains[3-4] OH11 and C3 with Clp from *Xcc* ATCC 33913 and *E. coli*'s CRP reveals how strong the resemblance[1] is:

Figure 5: Three *clp* homologues compared to *E. coli*'s cAMP-receptor protein (CRP).[2-6] Residues highlighted in green are predicted to be involved in binding c-di-GMP; residues highlighted in blue participate in H-bonds or salt bridges to stabilize Clp's intrinsic DNA-binding conformation.

```
Lysobacter enzymogenes C3          MSA-NPVAPLTALRRTNSPLLPDGATIERFLAHCHRRRYPSRTDVFRPGD 49
Lysobacter enzymogenes OH11        MSA-NPVAPLTALRRTNSPLLPDGATIERFLAHCHRRRYPSRTDVFRPGD 49
Xcc ATCC 33913                     MSLGNTTVVTTTVRNATPSLTLDAGTIERFLAHSHRRRYPTRTDVFRPGD 50
Escherichia coli F11               ----------MVLGKPQT-----DPTLEWFLSHCHIHKYPSKSTLIHQGE 35
                                             .: .. .       *:* **:*.* ::**::: ::: *:

Lysobacter enzymogenes C3          PASTLYYVVSGSVSIITEEEDGRELVLGYFGPGEFVGELGLFIESDHREV 99
Lysobacter enzymogenes OH11        PASTLYYVVSGSVSIITEEEDGRELVLGYFGPGEFVGELGLFIESDHREV 99
Xcc ATCC 33913                     PAGTLYYVISGSVSIIAEEDDDRELVLGYFGSGEFVGEMGLFIESDTREV 100
Escherichia coli F11               KAETLYYIVKGSVAVLIKDEEGKEMILSYLNQGDFIGELGLFEEGQERSA 85
                                    * ****::.***::: ::::.:*::*.*:. *:*:**:*** *.: *..

Lysobacter enzymogenes C3          ILRTRSTCELAEIGHERLYDLLLTRLSLDAPKLLYAIGAQISRRLLDTSR 149
Lysobacter enzymogenes OH11        ILRTRSTCELAEIGHERLYDLLLTRLSLDAPKLLYAIGAQISRRLLDTSR 149
Xcc ATCC 33913                     ILRTRTQCELAEISYERLQQLFQTSLSPDAPRILYAIGVQLSKRLLDTTR 150
Escherichia coli F11               WVRAKTACEVAEISYKKFRQLIQVN-----PDILMRLSAQMARRLQVTSE 130
                                    :*::: **:***.:::: :*: .       * :*  :..*:::**  *:.

Lysobacter enzymogenes C3          KAGRLAFLDVTDRIVRALHDLAKEPEAMSHPQGTQLRVSRQELSRLVGCS 199
Lysobacter enzymogenes OH11        KAGRLAFLDVTDRIVRALHDLAKEPEAMSHPQGTQLRVSRQELSRLVGCS 199
Xcc ATCC 33913                     KASRLAFLDVTDRIVRTLHDLSKEPEAMSHPQGTQLRVSRQELARLVGCS 200
Escherichia coli F11               KVGNLAFLDVTGRIAQTLLNLAKQPDAMTHPDGMQIKITRQEIGQIVGCS 180
                                   *...*******.**.::* :*:*:*:**:**:* *::::***:.::****

Lysobacter enzymogenes C3          REMAGRVLKKLQADGKLHARGKTVVLYGTR 229
Lysobacter enzymogenes OH11        REMAGRVLKKLQADGKLHARGKTVVLYGTR 229
Xcc ATCC 33913                     REMAGRVLKKLQADGLLHARGKTVVLYGTR 230
Escherichia coli F11               RETVGRILKMLEDQNLISAHGKTIVVYGTR 210
                                   ** .**:** *: :. : *:***:*:****
```

Comparison of the solved crystal structure of *Xcc*'s Clp[1] to the sequence of *Le*'s Clp revealed that residues predicted to be involved in stabilization of *Xcc*Clp's dimeric, symmetrical structure were conserved in *Le*Clp, as were the other key residue differences[1, 34] between the cAMP-receptor protein (CRP) and Clp. A BLAST analysis[38] revealed that *Le*Clp exhibits 79% identity and 85% similarity to *Xcc*Clp, rendering it unlikely that *Le*Clp differs biochemically from *Xcc*Clp.

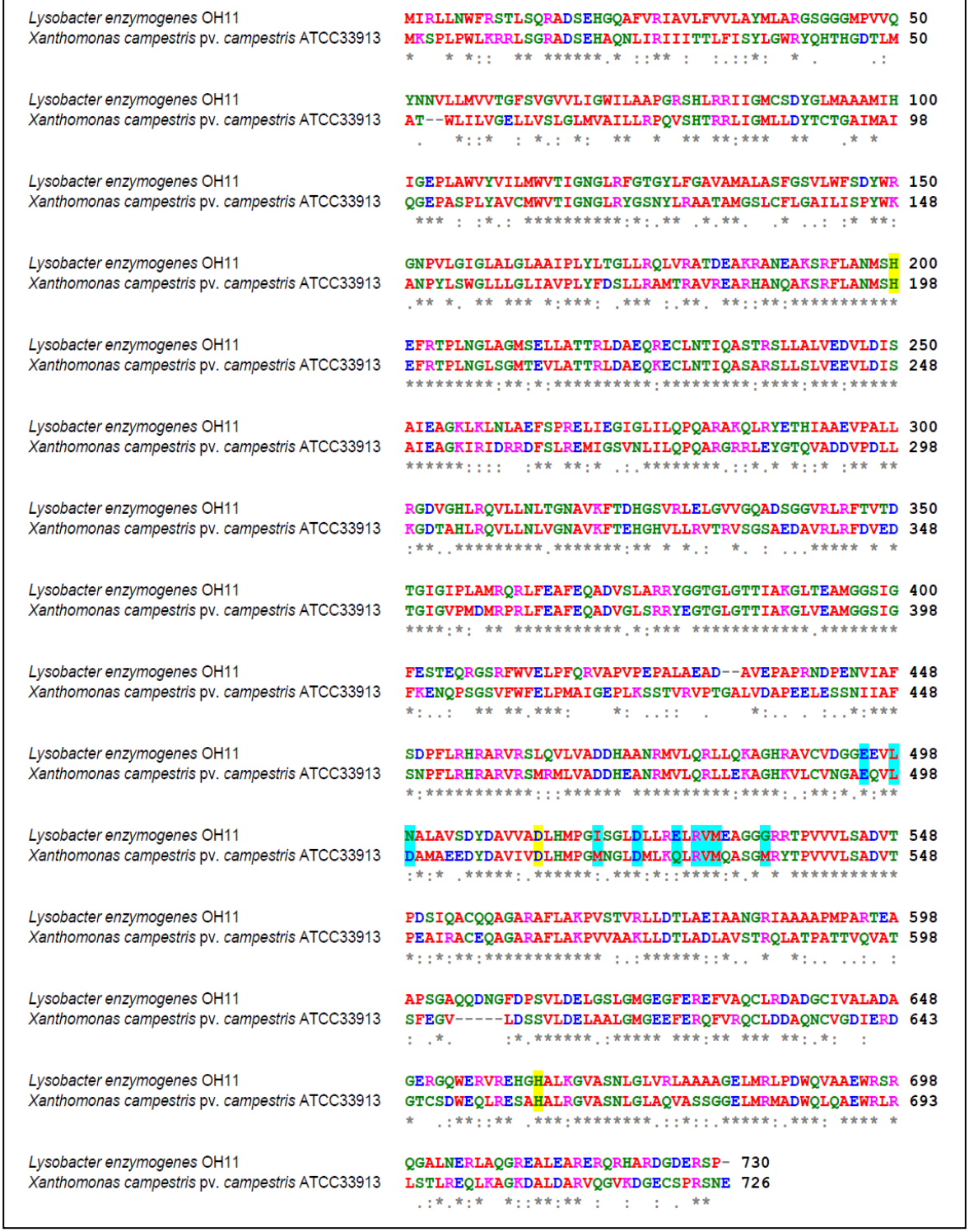
[1] The amino acid sequence of Clp is identical in the three strains of *Xcc* (B100, 8004, and ATCC 33913) found in Genbank. See protein accession numbers YP_001901908.1, NP_635866.1, and YP_241587.1 (http://www.ncbi.nlm.nih.gov/genbank/)

1.2.2 *rpf* homologues

The entire genome of *Lysobacter enzymogenes* OH11 was sequenced by Qian *et al.*,[39-40] enabling a search for the *clp* homologue's hypothetical *rpf* system partners. As expected, genes homologous to those in *Xcc* encoding the membrane-bound histidine kinase RpfC, phosphodiesterase RpfG, acyl-CoA ligase RpfB, and enoyl-CoA hydratase RpfF were discovered in *Le*OH11's genomic sequence.[40] The *rpfC* and *rpfG* homologous genes were predicted to encode the two proteins of the two-component system, while *rpfB* and *rpfF* were proposed to be involved in synthesis of a fatty acid signal molecule similar to *Xcc*'s DSF. Alignments[2] of the gene products of these homologues against the amino acid sequences of *Xcc*'s enzymes revealed a high level of homology between the four pairs of homologues.

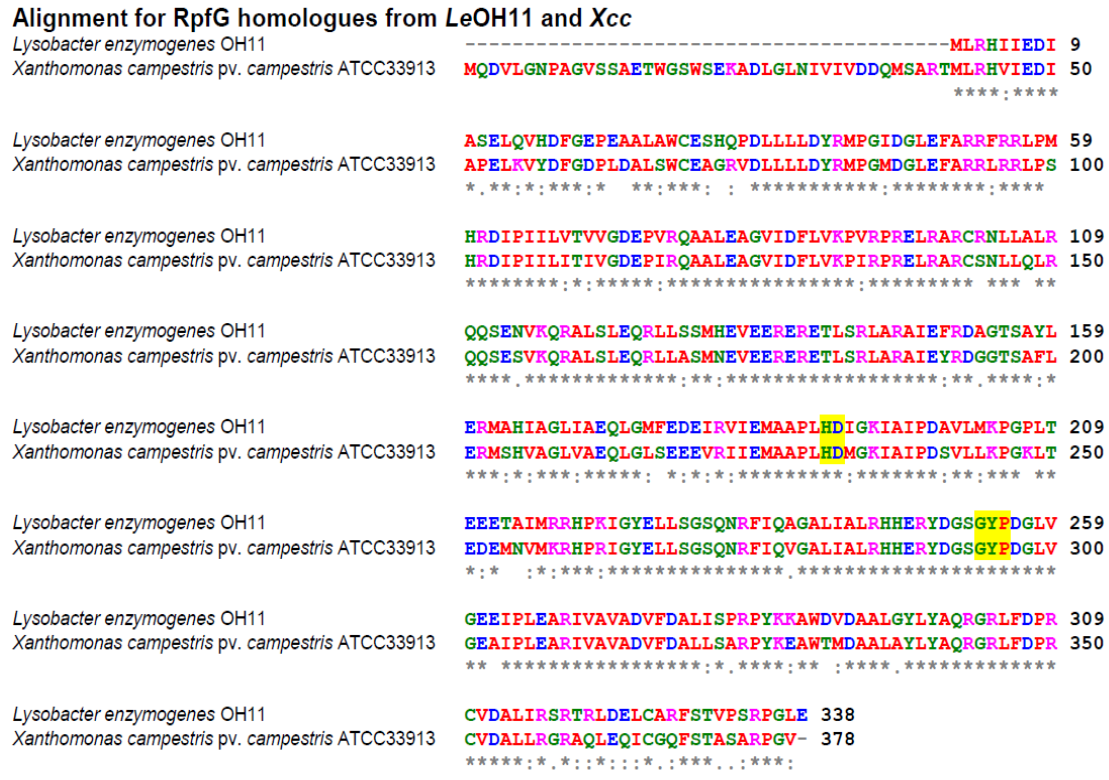
A BLAST analysis of *Le*OH11's homologue of *Xcc*RpfC (Genbank accession NP_637221),[5] the transmembrane sensor which detects extracellular DSF, revealed that the two proteins shared 57% identity and 70% similarity.[38] In addition, comparison of the key residues reported for *Xcc*RpfC's REC domain[8] (highlighted in blue in **Figure 6**) and in its phosphorelay/transfer domains[9] (highlighted in yellow in **Figure 6**) showed that all of the latter were conserved in *Le*OH11's RpfC homologue, and that most of the former were also conserved.

Figure 6: Comparison of *rpfC* homologues in *Le*OH11 and *Xcc*. Highlighted in blue are residues involved in interactions with DSF synthase RpfF; highlighted in green are the active residues of the histidine kinase.[8-9]



RpfG, the intracellular response regulator which partners with RpfC in *Xcc*, has two domains: a CheY-like receiver domain[9] which interacts with RpfC's phosphorelay/transfer domain, and an HD-GYP domain which possesses phosphodiesterase activity[29] against the intracellular second messenger cyclic-di-GMP.[32] Intriguingly, *Xcc*'s RpfG was shown (via yeast two-hybrid studies) to interact

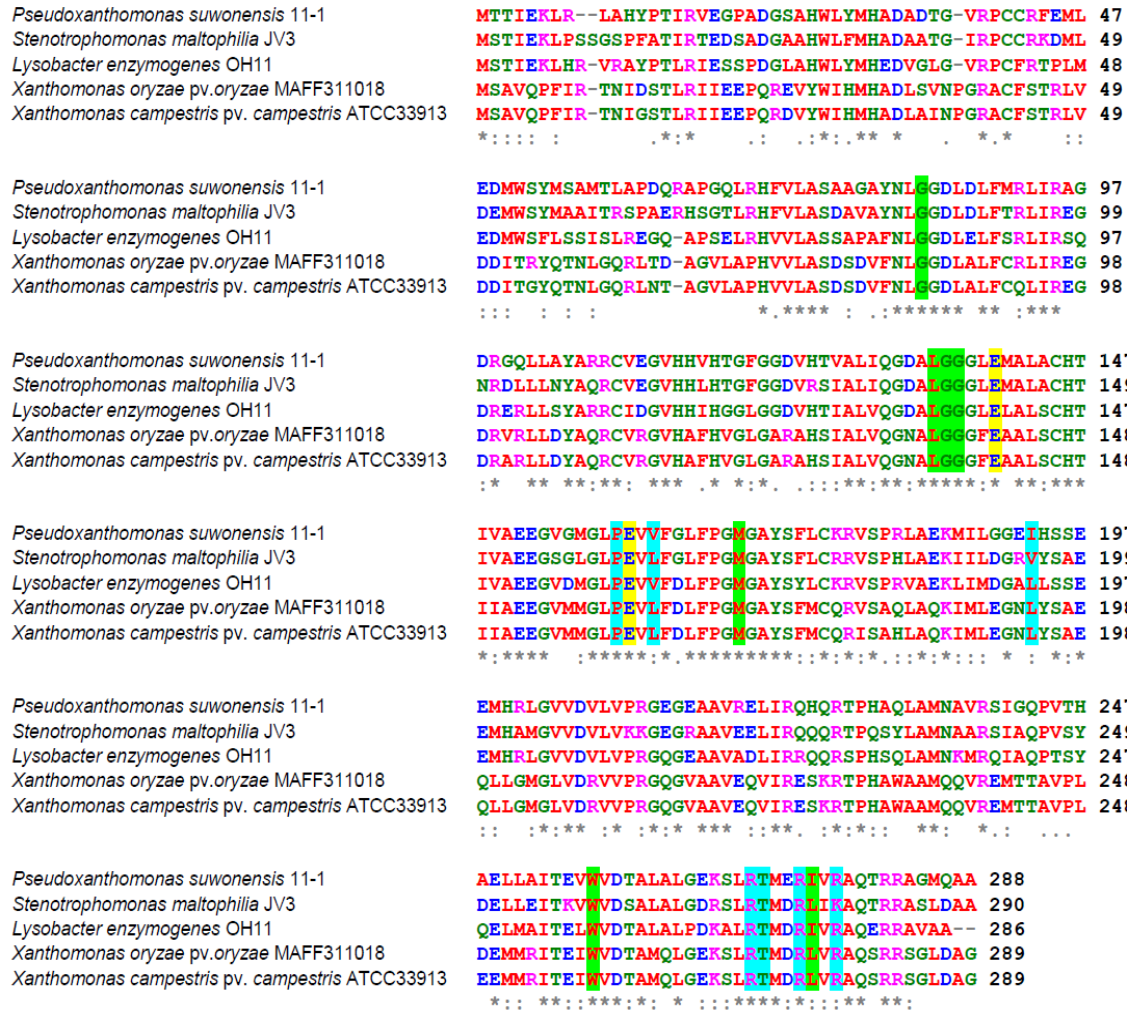
Figure 7: Comparison of *rpfG* homologues in *Le*OH11 and *Xcc*. The conserved HD-GYP residues are highlighted in yellow.



with and possibly inhibit a subset of proteins containing the GGDEF domain, which is associated with diguanylate cyclase (*i.e.* c-di-GMP synthesis) activity, indicating a further layer of regulation.[26, 41] Our BLAST analysis of *Le*OH11's homologue of *Xcc*RpfG (Genbank accession NP_637219) showed a remarkable 80% identity and 92% similarity between the two proteins, and the key HD-GYP residues were conserved (**Figure 7**). Notably, however, the BLAST analysis ignores a long *N*-terminal addition to

*Xcc*RpfG which ends with the methionine residue which corresponds with *Le*OH11's predicted initial Met.

Figure 8: Comparison of *rpfF* homologues from *Le*OH11 and *Xcc*. Two conserved catalytic glutamate residues, characteristic of isomerases, are highlighted in yellow. Residues involved in interactions with RpfC are highlighted in blue, and residues of the predicted hydrophobic pocket for nascent DSF's acyl chain are highlighted in green

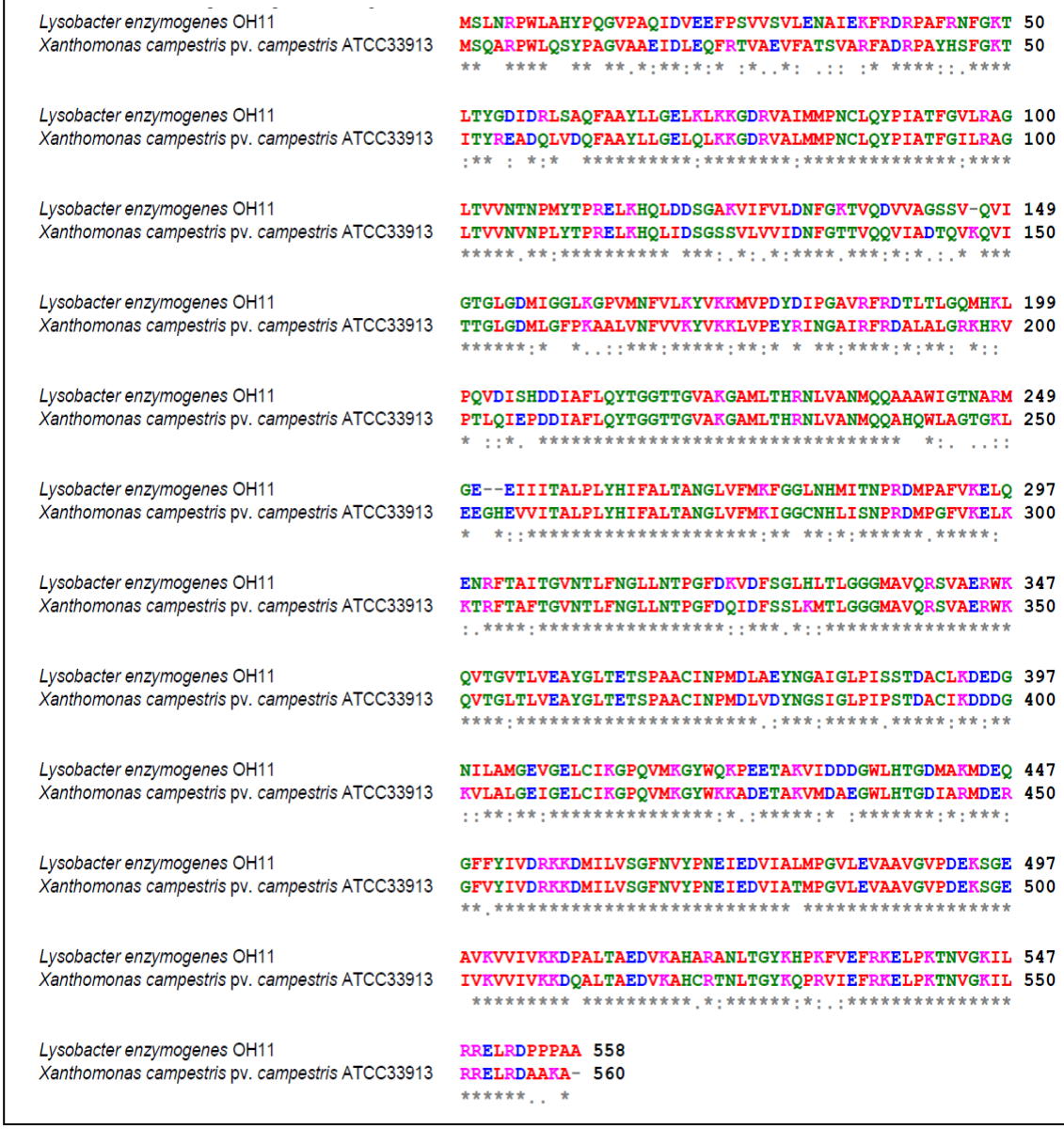


Reported studies of the *rpf* signaling system have focused heavily on *Xcc*RpfF, the essential DSF synthase which, as a predicted enoyl-CoA hydratase, probably modifies an acyl-CoA substrate to produce the fatty acid signal. A crystal structure of *Xcc*RpfF in complex with RpfC's REC domain was published in 2010, providing insight into the important negative regulation which the membrane-bound sensor RpfC exerts upon the

synthesis of its own signal.[8] In addition, it was shown that RpfF possessed the two conserved glutamate residues associated with enoyl-CoA hydratase activity, as opposed to the single Glu residue expected in an enoyl-CoA isomerase.[8] These key catalytic residues, as well as the residues shown by mutagenesis studies to be required for association between RpfG and RpfC, were conserved in *Le*OH11 homologues of *Xcc*RpfF (Genbank accession NP_637222.1), and a BLAST analysis showed 52% identity and 70% similarity between the two proteins. In **Figure 8**, the conservation of the key residues is shown between *Le*OH11 and *Xcc*, as well as with homologues of RpfF found in several related species. The conservation of RpfF-like proteins illustrates the likely widespread distribution of DSF-like signals among bacteria.[14]

In *Xcc*, deletion of the gene *rpfB*, predicted to encode an acyl-CoA ligase, resulted in a mutant strain unable to restore a wild-type phenotype to *rpfF* mutants, showing that *rpfB*'s gene product is also involved in DSF biosynthesis.[27] *Xcc*RpfB (Genbank accession NP_637223) is expected to be an acyl-CoA ligase which supplies RpfF with its acyl-thioester substrate.[5] A BLAST comparison of *Le*OH11's RpfB homologue with *Xcc*RpfB revealed a shared identity of 76% and similarity of 86%, and their strong similarity is illustrated by **Figure 9**.

Figure 9: Comparison between *rpfB* homologues from *Le*OH11 and *Xcc*



As summarized in **Table 1**, the genes of *Xcc*'s well-characterized *rpf* system show a very high degree of similarity to their four homologues in *Le*OH11. Based on this genetic evidence, we hypothesized that cell-cell signaling via a DSF-like molecule might be the key event in the upregulation of dihydromaltophilin's biosynthetic genes by Clp.

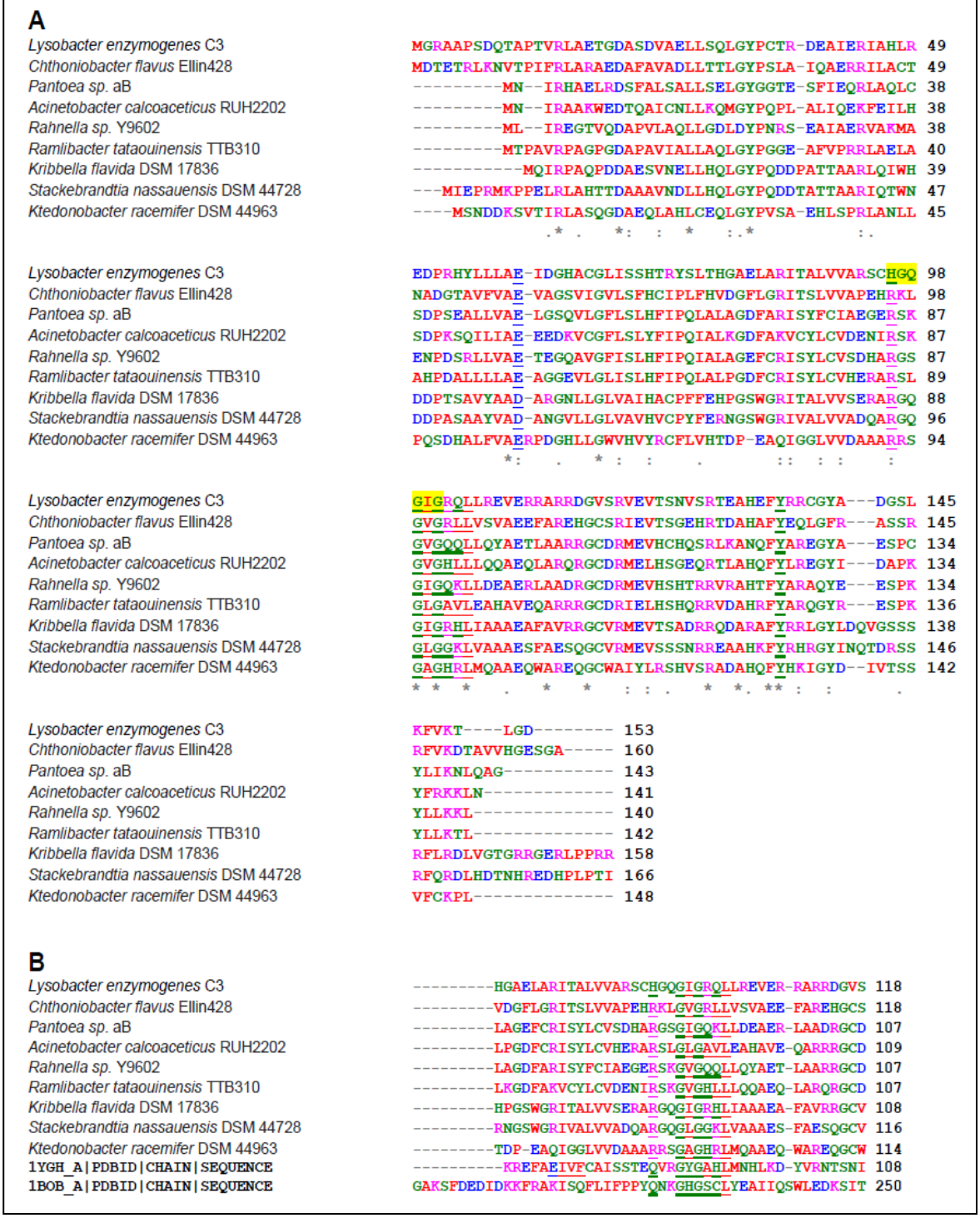
Table 1: BLAST analyses of *clp* and *rpf* homologues of *Le*OH11 and *Xcc* AT33913

gene	# residues	predicted function	identity	Similarity
***Le*Clp**	229	transcription factor	179/227(79%)	193/227(85%)
***Xcc*Clp**	230			
***Le*RpfC**	730	transmembrane sensor	409/721(57%)	511/721(70%)
***Xcc*RpfC**	726			
***Le*RpfG**	338	response regulator	271/337(80%)	313/337(92%)
***Xcc*RpfG**	378			
***Le*RpfF**	286	enoyl-CoA hydratase	142/274(52%)	194/274(70%)
***Xcc*RpfF**	273			
***Le*RpfB**	558	acyl-CoA ligase	425/560(76%)	486/560(86%)
***Xcc*RpfB**	560			

1.2.3 Homologues of Clp's accompanying acetyltransferase

Perhaps most intriguing aspect of *Le*OH11's *clp* and *rpf* homologues is the presence of a putative acetyltranferase gene seven bases downstream from the *clp*'s stop codon and shown by Koboyashi *et al*. to be transcriptionally linked with *clp*.[3] Inactivation of *clp* and the acetyltransferase (by transposon insertion), followed by complementation with Clp, provided a phenotype with all of *Le*C3's wild-type characteristics restored, except for a noticeable increase in gliding motility, [3] suggesting that the *clp*-linked acetyltransferase is involved in modulating this poorly-understood behavior. Some eukaryotic histone acetyltransferases have been implicated in direct modifications of the transcription machinery;[42] furthermore, direct acetylation of transcription factors by GCN5-like acetyltransferases has been reported in *E. coli*, adding yet another regulatory layer to gene expression.[43] We therefore speculated that Clp's

Figure 10: Comparisons between *Le*C3's putative acetyltransferase and some of the most similar sequences returned by BLAST analysis. Conserved residues which we hypothesize are involved in acetyl-CoA binding are highlighted.



acetyltransferase might by some mechanism be involved in modulating expression in Clp's regulon.

The *Le*C3 acetyltransferase shows some homology with known GCN5 (general control non-repressed)-family of *N*-acetyltransferases, which are associated with histone acetylation in eukaryotes.[44] These enzymes generally feature a conserved mechanism: the acetyl-CoA substrate is bound by conserved motif, and acetylation of the accepting amine (usually a lysine residue) is activated by means of proton abstraction by a general base (usually a conserved glutamate residue).[45-46] However, a BLAST search reveals that the *Le*C3's putative transferase shows strongest homology to a number of apparently uncharacterized acetyltransferases in bacteria, and its homology to the well-characterized HAT1 from eukaryote *Saccharomyces cerevisiae* is mostly limited to its putative acetyl-CoA-binding region (**Figure 10**).[47] The alignments shown in Figure 10**B** include two yeast histone acetyltransferases (HATs) with solved structures[46-47], and a careful comparison of the conserved acetyl-CoA-binding motif (R/HXXGXGXXL) between *Le*C3's acetyltransferase and the solved eukaryotic HATs led us to propose a targeted inactivation of the *Le*C3's acetyltransferase by deleting six residues (HGXGIG) involved in binding the enzyme's co-substrate acetyl-CoA. However, we have so far been unable to recover a mutant bearing the expected deletion, which may suggest that expression of Clp without its acetyltransferase has adverse consequences.

Chapter 2: Diffusible signal factors in *Lysobacter enzymogenes*

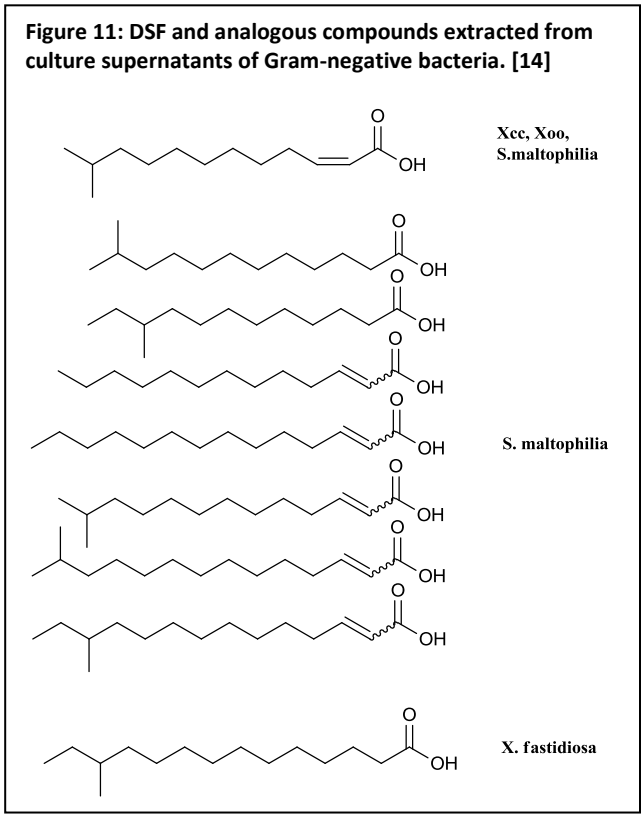


Figure 11: DSF and analogous compounds extracted from culture supernatants of Gram-negative bacteria. [14]

Koboyashi *et al*. noted in 2005 that the effects of Clp's deletion in *Xcc* paralleled those in *Le*C3: production of extracellular enzymes, pigment, and extracellular polysaccharides, as well as pathogenicity, appear to be regulated by Clp in both species.[3] We hypothesized that *Le*C3 and *Le*OH11 similarly use diffusible signal factors to coordinate their own pathogenicity against fungi or bacteria. The first diffusible signal factor identified from *Xcc* was *cis*-11-methyl-2-dodecenoic acid,[28] but the application of improved extraction procedures to *Xcc* and other xanthomonads revealed several similar molecules which provided bioactivities similar, though sometimes to varying degrees, to the first DSF. *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) produces three DSF-like signal molecules, including DSF itself and two fatty acid analogues.[37] *Xylella fastidiosa* secretes a fatty acid with DSF-like activity,[36] but unlike DSF, the signal molecule is an unsaturated fatty acid with a methyl branch on the antepenultimate rather than the penultimate carbon on the acyl chain.[48] *Stenotrophomonas maltophilia* produces eight fatty acid signal-like molecules,[49] including DSF itself which is involved in signaling for bacterial virulence and antibiotic resistance.[50] In all of these organisms, the DSF-like molecules are

associated with *rpf* homologues (Figure 11).[36-37, 50] Non-xanthomonads, such as species of the *Burkholderia cepacia* complex, also produce DSF and its analogues. [14, 51] The use of small fatty acids as intercellular signals seems clearly to be widespread among the Gram-negative bacteria.

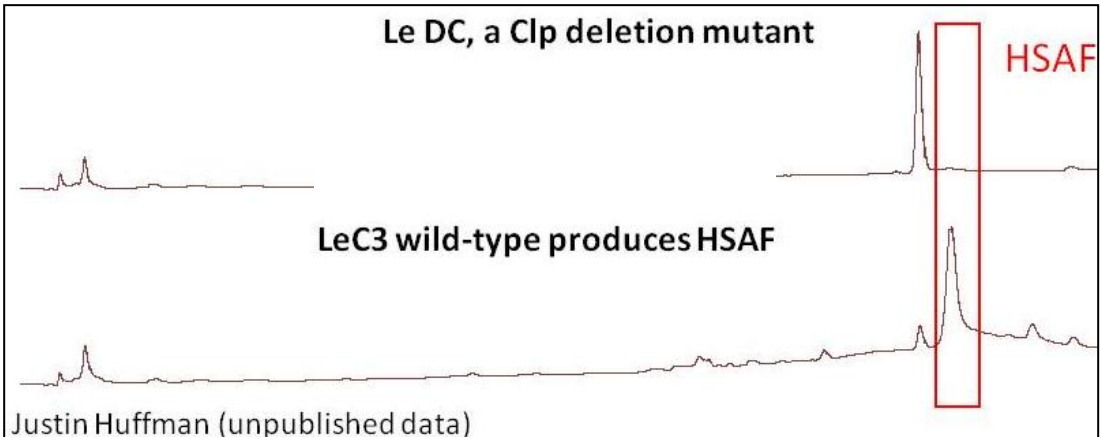
2.1 Culturing conditions

In our search for DSF in *L. enzymogenes*, our basic assumption was that culturing conditions supporting production of dihydromaltophilin would also support production of DSF, since dihydromaltophilin was known to depend upon Clp, which in turn depended upon the DSF

and *rpf* system, according to the hypothetical similarity between *L. enzymogenes* and *Xcc*. We first established that Clp was an absolute requirement for dihydromaltophilin production by culturing the *Le*C3 *clp* deletion mutant, *Le*DC211,[3] in various media (**Figure**

12).

Figure 12: Dihydromaltophilin (HSAF) [12]

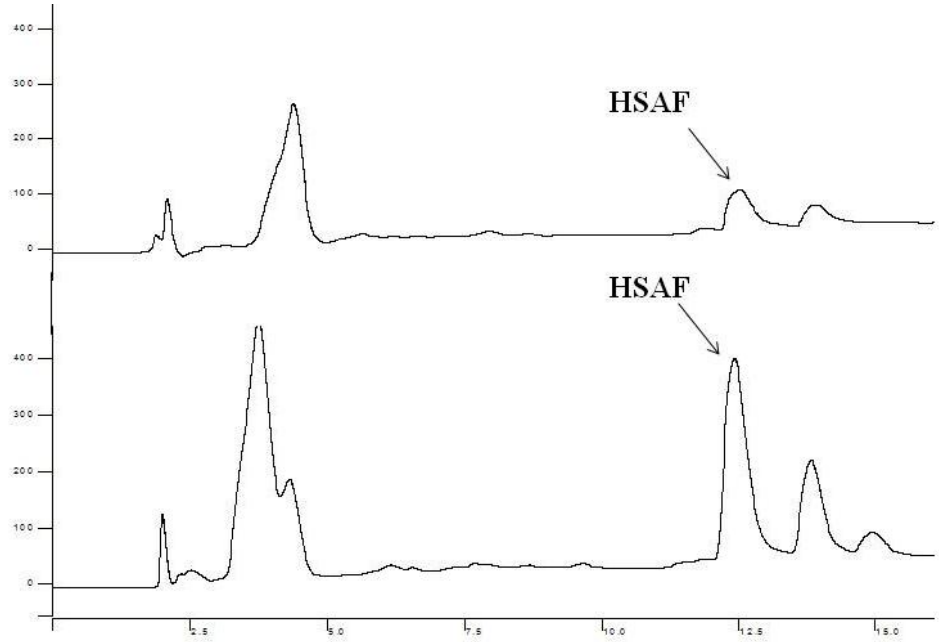


***Le*C3 wild-type**

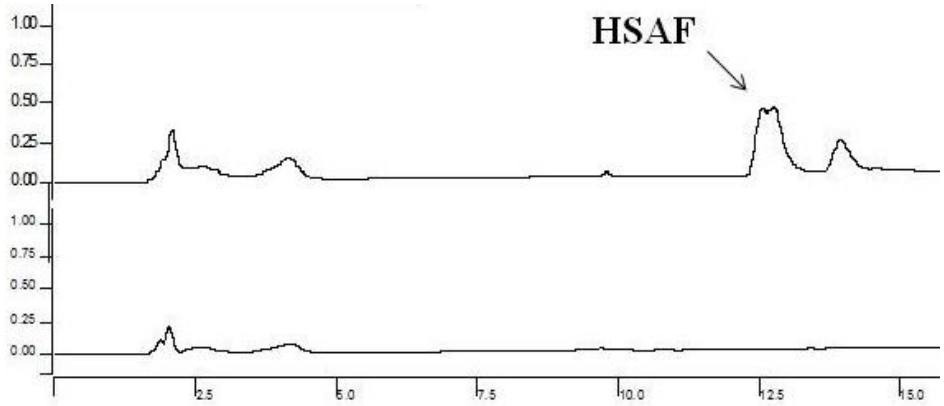
Generally, the medium of choice for dihydromaltophilin production was tryptic soy broth (TSB) (17 g/L casein, 3 g/L soy peptone, 5 g/L sodium chloride, 2.5 g/L potassium phosphate monobasic, 2.5 g/L glucose) diluted tenfold and incubated at 28 °C with

Figure 13: Dihydromaltophilin (HSAF) production in various media. HPLC conditions: row 3 of Table 2

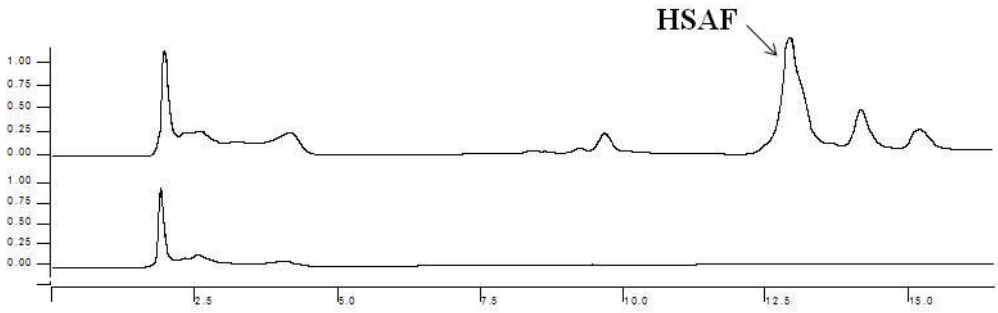
A) Ethyl acetate extraction of *Le*C3 grown in NYGB supplied with 200 m*M* glycerol (top) and in NYGB in which glycerol is replaced with 100 m*M* glucose.



B) Extractions of dihydromaltophilin (HSAF) from *Le*C3 (top) and *Le*DC (bottom) in NYGB in which glycerol is replaced with 40 m*M* maltose.



C) Extractions of dihydromaltophilin (HSAF) from *Le*C3 (top) and *Le*DC (bottom) in NYGB in which glycerol is replaced with 40 m*M* lactose.



shaking for 2-3 days after inoculation with *Le*C3 or *Le*OH11 or mutants thereof.[15] Although 10% TSB was originally chosen because conditions of poor nutrition were thought to favor secondary metabolism, we discovered in the course of our experiments that media of considerably richer nutritional content strongly supported production of dihydromaltophilin and its analogues (**Figure 13**). NYGB (5 g/L peptone; 3 g/L yeast extract; 20 g/L glycerol), the medium used to cultivate *Xcc* and *Xoo* for DSF production,[28, 37] was found to support a high level of dihydromaltophilin production in *L. enzymogenes*. When 20 g/L glycerol (equivalent to 217 m*M*) was replaced with 108 m*M* glucose, dihydromaltophilin production seemed, if anything, to increase **(Figure 13A**). Supplements with various carbohydrates (glycerol, glucose, maltose, and lactose) appeared generally to increase dihydromaltophilin production over that observed in 10% TSB, an observation which seemed to contradict the previous assumption that extremely poor media was necessary to encourage secondary metabolism in *L. enzymogenes*. Another possibility is that supplying an elevated level of carbohydrates induces dihydromaltophilin production by an unknown mechanism.

The observation that rich nutritional conditions, which supported growth to extremely high optical densities, still supported secondary metabolism seemed to support our hypothesis that quorum sensing, mediated by a small DSF-like signal and *L. enzymogenes*'s analogues of the *rpf*/Clp system, led to upregulation of dihydromaltophilin's biosynthetic gene machinery. No medium, poor or rich, was found which supported production of dihydromaltophilin in *Le* 5E4, *Le*DC211, or *Le*DCA2422, the Clp disruption/deletion mutants provided by Koboyashi *et al*,[3] suggesting that the dihydromaltophilin biosynthetic genes lies completely in Clp's regulon, upregulated

either directly or by a downstream transcription factor, and that there is no alternative regulation pathway circumventing Clp that leads to dihydromaltophilin's biosynthesis.

2.2: Extraction of DSF-like molecules using ethyl acetate

Isolation of the diffusible signal factors has generally relied on the difference in polarity between the fatty acid signaling molecules and their aqueous milieu. In the original extraction of DSF from *Xcc*, the bacteria were grown to a high optical density, removed by centrifugation, and the supernatant twice extracted by 0.3 volumes of ethyl acetate (equilibrated with $NaHCO_3$ solution), and the residue redissolved and fractionated on a column of silica gel 60.[27] Sufficient DSF for structural determination was obtained from ethyl acetate extractions of 30 L of an *rpfC* mutant of *Xcc* (which, lacking RpfC's RpfF-repressing domain, overproduces DSF), followed by purification using flash chromatography and HPLC, which yielded ~2 mg of pure material.[28] Because of the fatty acids' relatively low pK_a (<4.5), reducing the pH of the supernatant to ~4.0 ensured that any free fatty acids were fully neutralized, preventing the ionized carboxylate from favoring the aqueous phase over the organic phase during partition.[37]

Our extractions of cultures of *Lysobacter enzymogenes* C3 or OH11 followed this protocol closely. *Le*C3 or OH11 (wild-type or mutant strains) were grown in NYGB to maximum optical density ($OD_{600} \approx 1.8$), which generally took 48-72 hrs. at 28-30 °C, in an incubator with shaking at 200 rpm. Centrifugation was carried out at 4000 rpm in a Sorvall Legend RT centrifuge (radius = 21.3 cm for swinging bucket rotor [52]) at 4 °C, and 37% HCl was added to the supernatant until the pH ≈ 4 (determined visually by pHydrion 1-to-12 litmus paper (MicroEssential Laboratories, New York)). The supernatant was then extracted with an equal volume of ethyl acetate (≥99.5%, Sigma-

Aldrich), and the aqueous and organic phases allowed to partition. The organic phase was then separated from the aqueous phase using a separatory funnel; at this stage, an emulsion often formed that could be minimized by vigorously swirling the ethyl acetate/aqueous phase mixture in the separatory funnel before draining the aqueous phase. The ethyl acetate was then evaporated using rotary evaporation apparatus (Büchi Rotovapor R-200) with a water bath set at 40 °C, and the yellowish residue stored at room temperature while the ethyl acetate was re-used to do a second extraction of the aqueous phase. Once the second extraction was complete, the combined residues (we expected around 200 mg from an ethyl acetate extraction of a 72-hr, 500 mL NYGB culture) were redissolved in small volumes (~1-5 mL) of methanol, ethyl acetate, or dichloromethane, and subjected to further separations or evaluations for biological activity.

2.3 Extract fractionation and isolation of compounds of interest by HPLC and flash chromatography

Numerous variations of standard separation techniques were attempted in our effort to isolate a DSF-like molecule from cultures of *Le*C3 or *Le*OH11. Because we could not be certain that the hypothetical diffusible signal factor produced by *L. enzymogenes* would be identical to any of the known DSFs, we focused on those fractions containing the hydrophobic metabolites similar to the original DSF, *cis*-11-methyl-2-dodecenoic acid. Throughout 2011 and the first half of 2012, we were able to purchase small amounts (~2 mg) of synthetic DSF from Cayman Chemical,[53] but an interruption in the availability of feedstock of DSF synthesis interfered with continuation of experiments using this chemical. However, we were able to use our available supply of DSF to as an HPLC

standard, helping to identify the fractions in which a *Lysobacter* DSF would most likely be present.

After redissolving the residues of the ethyl acetate extractions in the chosen solvent, we usually reserved a small portion of the crude extract for a direct test of biological activity (described below). The remaining crude extract was further fractionated into increasingly hydrophobic portions, using flash chromatography, high-pressure liquid chromatography (HPLC), thin-layer chromatography (TLC), or a combination of these methods. Several different HPLC methods were used, based on the requirements of the extraction and on the availability of HPLC columns. Most often, we used a program developed for separating metabolites extracted from *Le*C3, and specifically for detecting dihydromaltophilin and its analogues. We used water and acetonitrile as solvents A and B, respectively, with each solvent containing 0.025% (v/v) trifluoroacetic acid, which we later replaced with 0.05% formic acid. The separations were generally carried out on a Varian Prostar HPLC system, with some done on an Agilent 1220 Infinity LC system. All solvents were filtered before use. The programs are summarized in **Table 2:**

Table 2: columns and methods used for HPLC of *L. enzymogenes* culture extracts

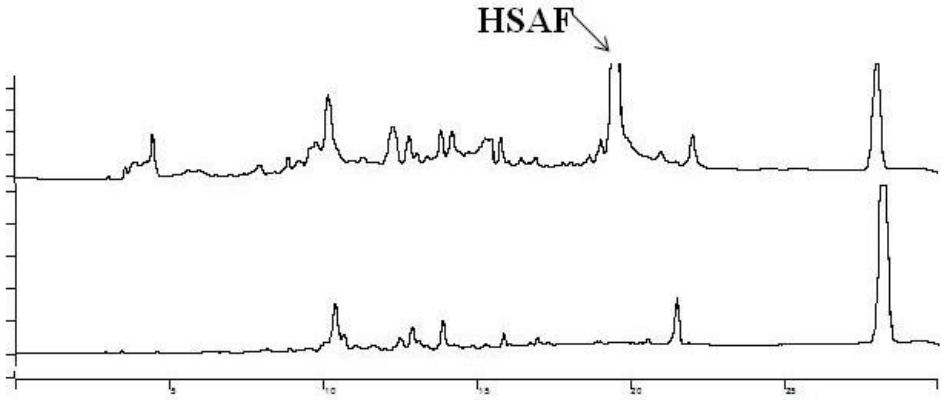
Method	Column	Solvent A	Solvent B	Conditions
1. C3	4.6 mm x 250 mm C18, 5 um particle size, Grace Alltima	H_2O (+0.05% formic acid or 0.025% TFA)	CH_3CN (+0.05% formic acid or 0.025% TFA)	0-10 min, 5-40% B; 10-15 min, 40-80% B; 15-20 min, 80% B; 20-21 min, 80-100% B; 21-23 min, 100% B; 23-25 min, 100-5% B
2. C3-std	4.6 mm x 250 mm C18, 5 um particle size, Grace Alltima	H_2O (+0.05% formic acid or 0.025% TFA)	CH_3CN (+0.05% formic acid or 0.025% TFA)	0-10 min, 5-40% B; 10-15 min, 40-80% B; 15-20 min, 80% B; 20-21 min, 80-100% B; 21-23 min, 100% B; 23-28 min, 100-5% B; 28-30 min, 5% B
3. C3-for150mm	4.6 mm x 150 mm C18, 5 um particle size, Zorbax	H_2O (+0.05% formic acid or 0.025% TFA)	CH_3CN (+0.05% formic acid or 0.025% TFA)	0-5 min, 5-40% B; 5-15 min, 40-60% B; 15-16 min, 60-100% B; 16-17 min, 100-5% B; 17-20 min, 5% B
4. New	4.6 mm x 250 mm C18, 5 um particle size, Grace Alltima	H_2O (+0.05% formic acid or 0.025% TFA)	CH_3CN (+0.05% formic acid or 0.025% TFA)	0-10 min, 5-20% B; 10-13 min, 20-70% B; 13-17 min, 70-95% B; 17-20 min, 95% B; 20-22 min, 95-5% B; 22-27 min, 5% B
5. Dihydro-maltophilin detection in 75mm (DM75)	4.6 mm x 50 mm C18, 2.7 um particle size, Agilent Poroshell	H_2O (+0.05% formic acid or 0.025% TFA)	CH_3CN (+0.05% formic acid or 0.025% TFA)	0-5 min, 5-40% B; 5-15 min, 40-50% B; 15-20 min, 50-90% B; 20-25 min, 90-5% B; 25-30 min, 5% B

Following the reported methods of DSF detection, the detector wavelength was usually 212 nm,[37] but was occasionally set at 220 nm. HPLC was sometimes used for fractionation of the samples, by collecting the peaks as they emerged from the UV detector and evaporating the water-acetonitrile solvent using an air stream at room temperature (**Figure 14**). In all of the methods cited in **Table 2**, synthetic DSF eluted after dihydromaltophilin and its analogues, allowing us to concentrate on compounds which

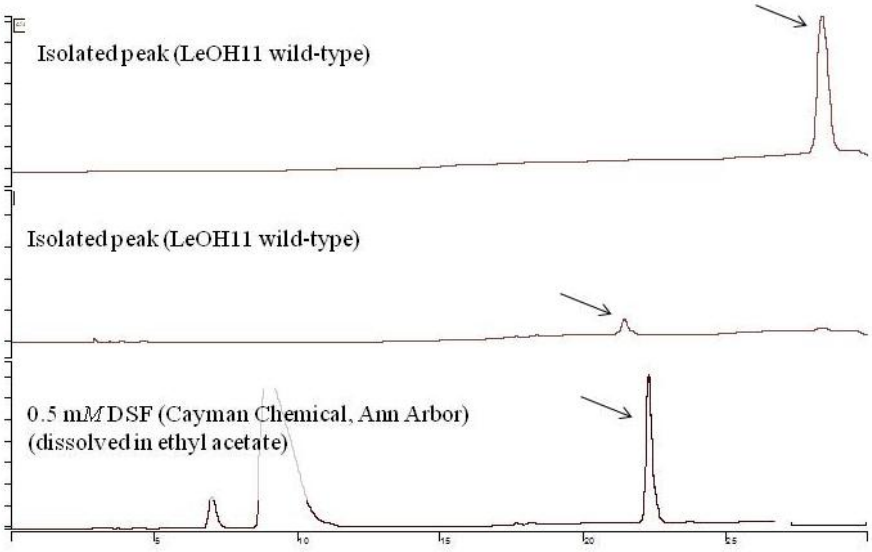
eluted late in each run.

Figure 14: HPLC analysis of ethyl acetate extracts of *L. enzymogenes*. Method 2, Table 2

A) Ethyl acetate extractions of *Le*C3 (top) and *Le*OH11 (bottom) wild-type strains grown in 10% TSB supplemented with 5 mM lactose. Unusually, *Le*OH11 failed to produce dihydromaltophilin (HSAF) in this experiment.



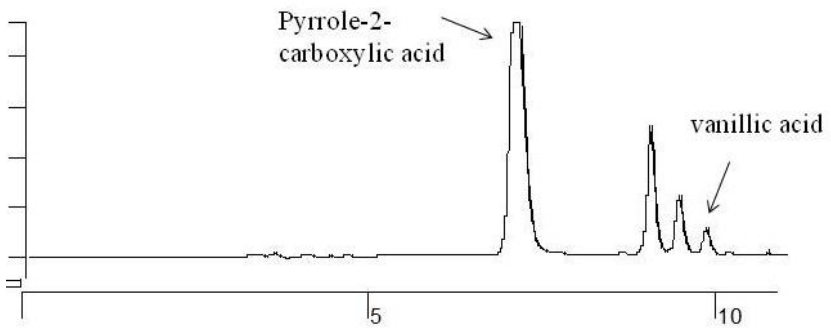
B) Peaks purified by HPLC from *Le*OH11. Synthetic DSF serves as a standard.



Extracts of large cultures of *L. enzymogenes* (0.5-2 L) were initially fractionated using flash chromatography, with silica gel 60 (mass of silica gel = mass of dried residue × 30) packed in 40 mL hexane serving as the stationary phase. The crude residue (~200 mg) was redissolved in dichloromethane and slowly spotted and dried onto 200 mg of

silica gel to allow the sample to be dry-loaded onto the silica gel column. The column was then washed with successive 100-mL volumes of hexane/acetone, in ratios of 4:1, 3:1, 2:1, and finally 100% acetone. Eluents were collected in 10-mL volumes, and the presence of metabolites in each 10-mL elution was monitored by thin-layer chromatography (TLC). The same solvent system used for elution was used for the TLC analysis of each 10-mL fraction. The samples were spotted onto silica gel TLC plates and allowed to develop for 10 minutes at room temperature. After drying, the plates were visualized by immersion is solid I_2 and then in 5% H_2SO_4 in ethyl acetate, followed by drying over a hot plate. Metabolites were then viewed under visible or ultraviolet (λ = 254 nm) light. HPLC analysis clarified the compounds in each fraction (**Figure 15**).

Figure 15: Extract from *Le*C3 fractionated by flash chromatography. Method 2, Table 2



Combined with the ethyl acetate extractions described in Chapter 2.5, these protocols were designed to extract fatty acid-type metabolites from the supernatants of *Lysobacter* cultures. The utility of these measures is demonstrated by the successful extraction, purification, and spectroscopic identification of two relatively nonpolar acids, vanillic acid (**1**) and pyrrole-2-carboxylic acid (**2**) **(Figure 16)** from the supernatant of

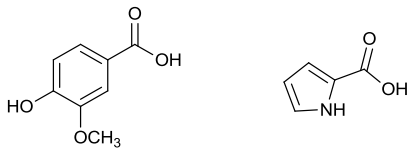


Figure 16: vanillic acid (1) and pyrrole-2-

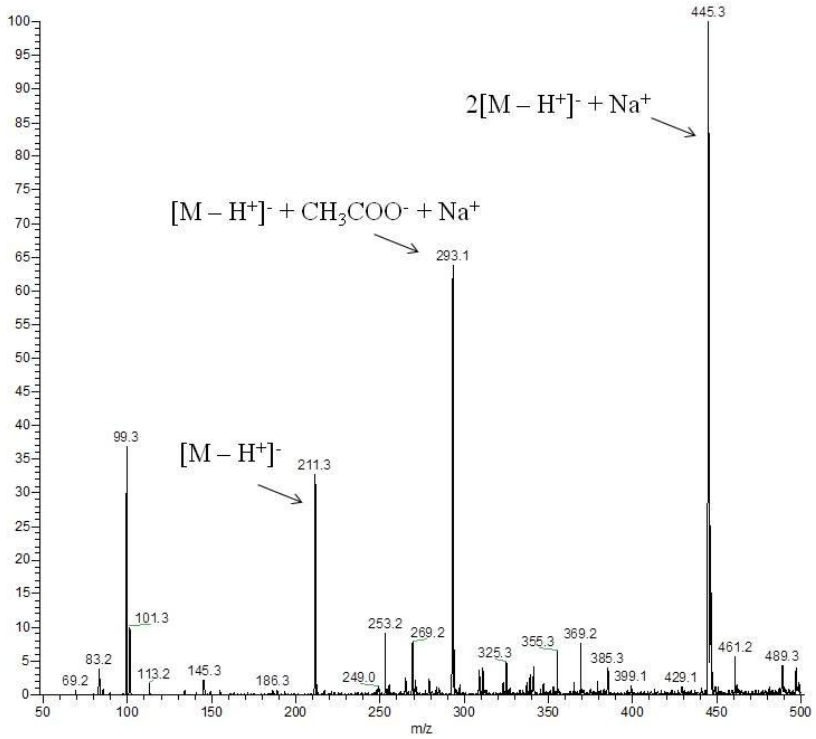
*Le*C3 cultured in NYGB medium. These two metabolites were initially of some interest because of the possibility of their involvement in some type of intercellular signaling, although their dissimilarity to fatty acid-type signal molecules made it seem quite unlikely that either molecule was involved in recognition by a *Lysobacter* homologue of RpfC, the transmembrane DSF sensor.

2.4: MS and LC-MS analysis

For mass analysis of the compounds in the fractionated extracts of *L. enzymogenes* culture, we optimized a method of liquid chromatography-mass spectrometry (LC-MS) using an Agilent 1220 Infinity LC high-pressure liquid chromatography system interfaced with a Finnigan LCQ ion trap mass spectrometer. A 250 mm × 1.0 mm chromatography column (octadecylsilyl (C18) stationary phase, 5 μm particle size) was purchased from Grace Davidson and used for all LC-MS experiments. Reported fatty acid analyses by mass spectrometry usually involved dissolving the samples in methanol or acetonitrile, with 5 mM ammonium acetate to favor deprotonation of the fatty acid.[54] Negative-mode electrospray ionization (ESI) is the method of choice for anionic species like fatty acid conjugates,[54-58] so we expected to observe mostly the $[M - H^+]^-$ peak for acidic analytes. Based on mass spectrometer conditions reported in the literature [57-58] and our own experimentations with synthetic DSF (Cayman Chemical, Ann Arbor, Michigan) and with 10-undecenoic acid, we found that ionization and detection were optimized under the following conditions: the capillary temperature was set at 220 °C and the capillary voltage at -42 V. The tube lens offset was varied between -20 and -30 V and the injection flow rate was usually 10 μL/min when we directly injected pure, or nearly pure,

Figure 17: mass spectra of standard fatty acids

A) 30 μ*M* synthetic DSF (Cayman Chemical, Ann Arbor)



B) 58 μ*M* 10-undecenoic acid (Sigma)

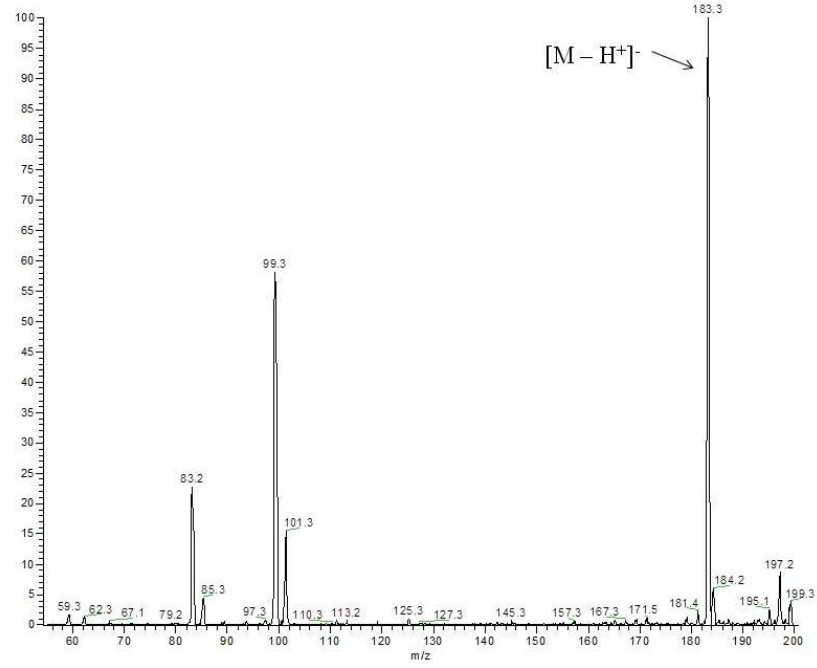
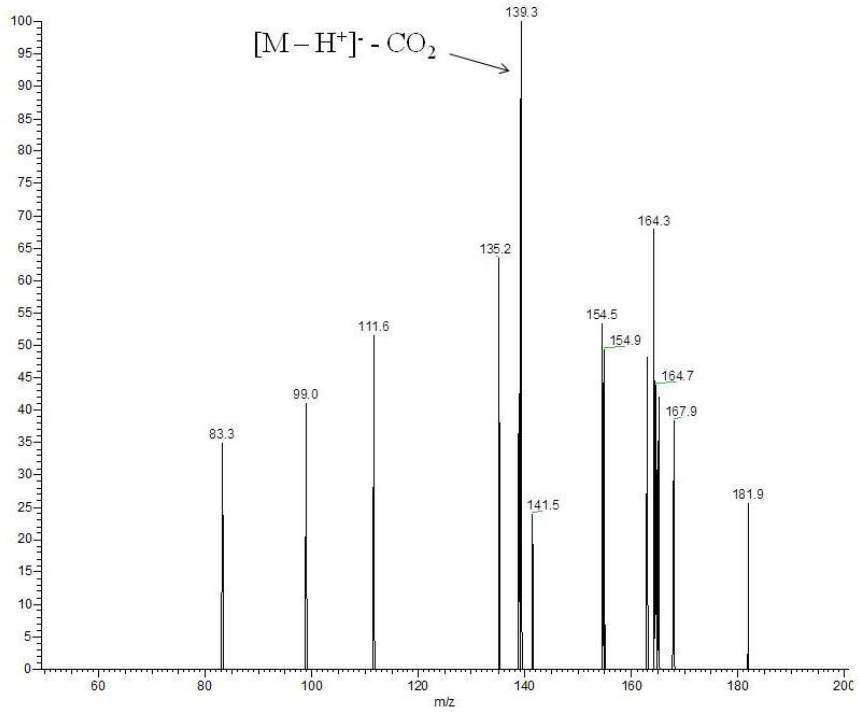


Figure 18: MS/MS analysis of a standard fatty acid (10-undecenoic acid)



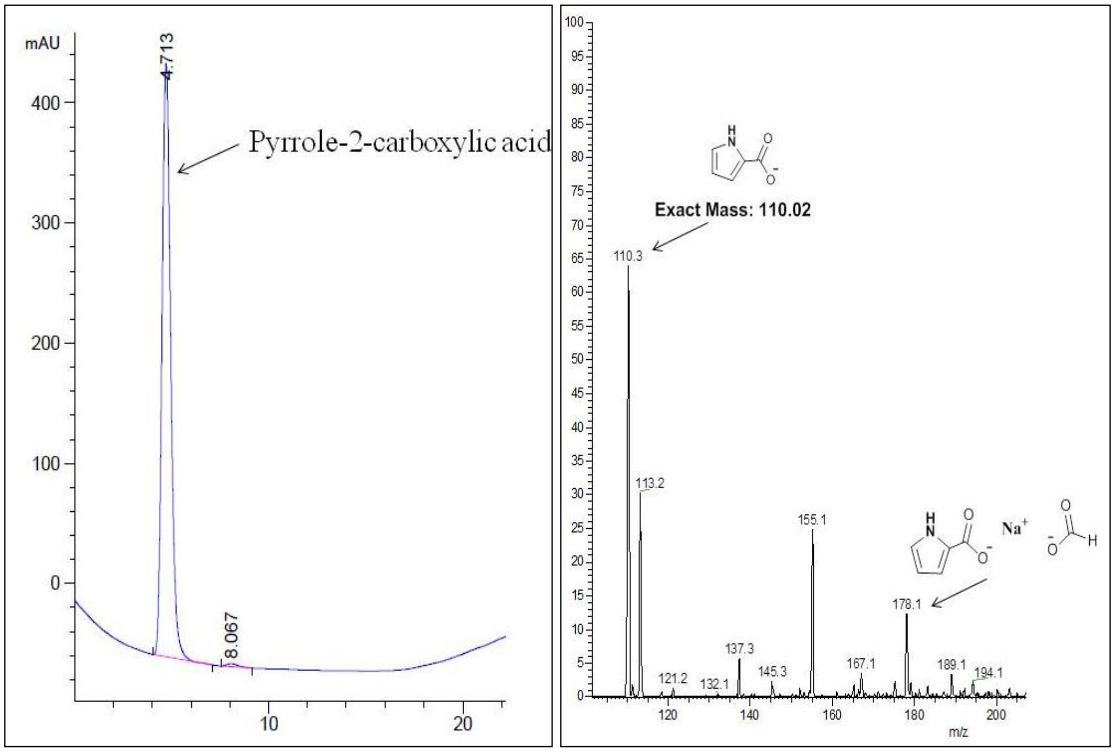
samples into the mass spectrometer. Fatty acid samples were dissolved in methanol with 5 mM ammonium acetate, and sample concentrations were 30-60 μM (**Figure 17**).

When possible, we also performed tandem MS/MS in the same instrument by isolating the $[M - H^+]^-$ fragment and performing collision-induced dissociation (CID) with a collision energy of 40-50%. Appearance of an of intense peak 44 mass units lighter than the isolated $[M - H^+]^-$ peak strongly suggested that the ion of interest had fragmented to carbon dioxide and an alkyl group, confirming that the isolated ion was a deprotonated carboxylic acid (**Figure 18**, **Figure 21**). All mass spectra were analyzed using XCalibur software.

Conditions were also optimized for LC-MS analysis of sample fractions containing several compounds. Combinations of acetonitrile and water, or methanol and water, were tested in the optimization process, Several isocratic programs were tested,

but we were able to obtain adequate separations of the *Lysobacter* culture extracts by switching to a gradient program (**"Gradient4@0.05"**) with the following parameters: from 0-15 minutes solvent B increased from 1% to 40%; 15-25 min., solvent B increased to 80%; 25-30 min., solvent B increased to 100%; 30-45 min., solvent B decreased to 1%; 45-60 min., solvent B was held constant at 1%. Because pure water with 5 m*M* ammonium acetate was found to support rapid bacterial growth, we chose 22% methanol in water, with 5 m*M* ammonium acetate, as Solvent A, with 100% methanol with 5 m*M* ammonium acetate as Solvent B. This program, combined with the mass spectra conditions described above, allowed for the separation and analysis of the fractionated extracts from *L. enzymogenes* cultures (**Figure 19**).

Figure 19: LC-MS analysis of LeC3 extract containing pyrrole-2-carboxylic acid. Formic acid was consistent contaminant under these conditions.



Absorbance at 212 nm **ESI mass spectrum**

For high-resolution mass spectrometry, we submitted pure samples first screened on the Finnigan LCQ ion trap to the Nebraska Center for Mass Spectrometry. Data from a Synapt time-of-flight mass spectrometer provided mass spectra at a resolution of 433 ppm using negative mode ESI. Only samples purified by HPLC were submitted for high-resolution mass spectra analysis. The identity of the isolated vanillic acid was confirmed by its exact mass (**Figure 20**) and by the identical retention times of a synthetic standard and the isolated compound.

Careful analysis of the masses and retention times of compounds from the *L. enzymogenes* culture extracts provided no clear indications of a DSF-like molecule. Our search specifically for a hydrophobic compound with a nominal mass between 200 and 300 with the mass spectral characteristics of a fatty acid did not yield any good candidates.

Figure 20: high-resolution mass spectrum of isolated vanillic acid

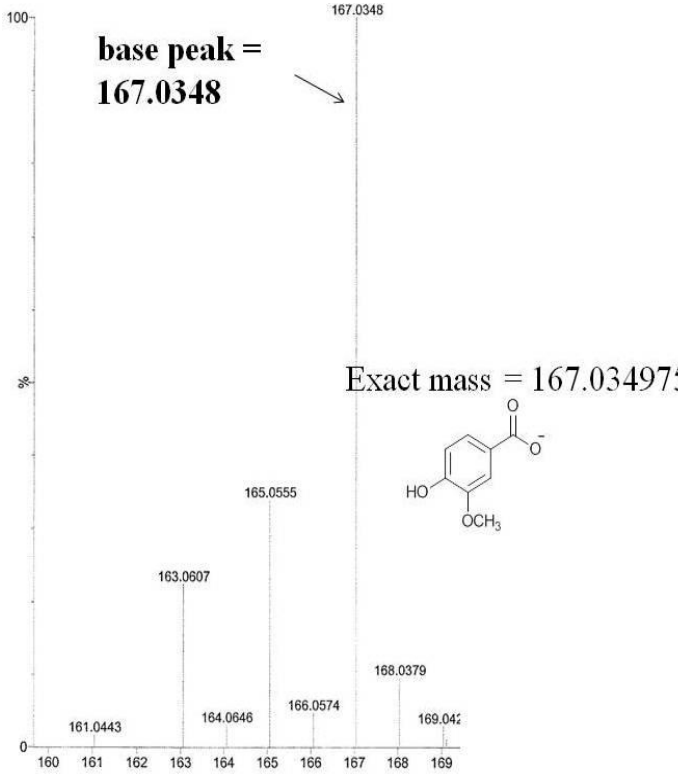
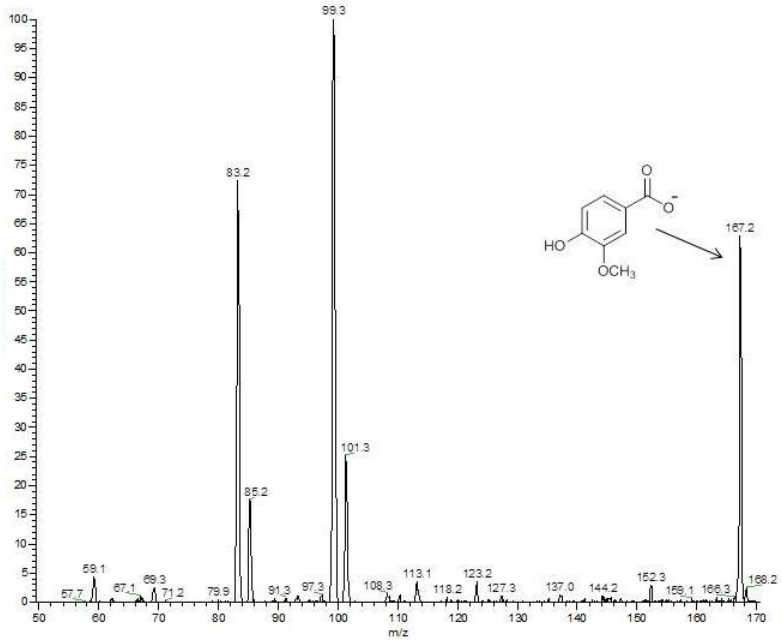
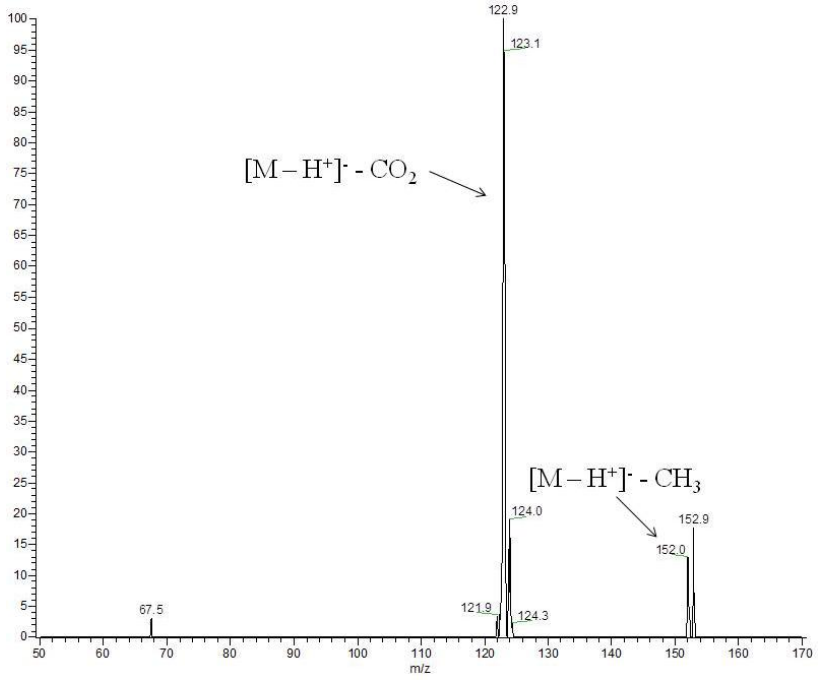


Figure 21: isolated vanillic acid from fractionated extract

A) ESI mass spectrum of isolated vanillic acid



B) Tandem MS/MS of isolated vanillic acid. Loss of CO2 provides the base peak.

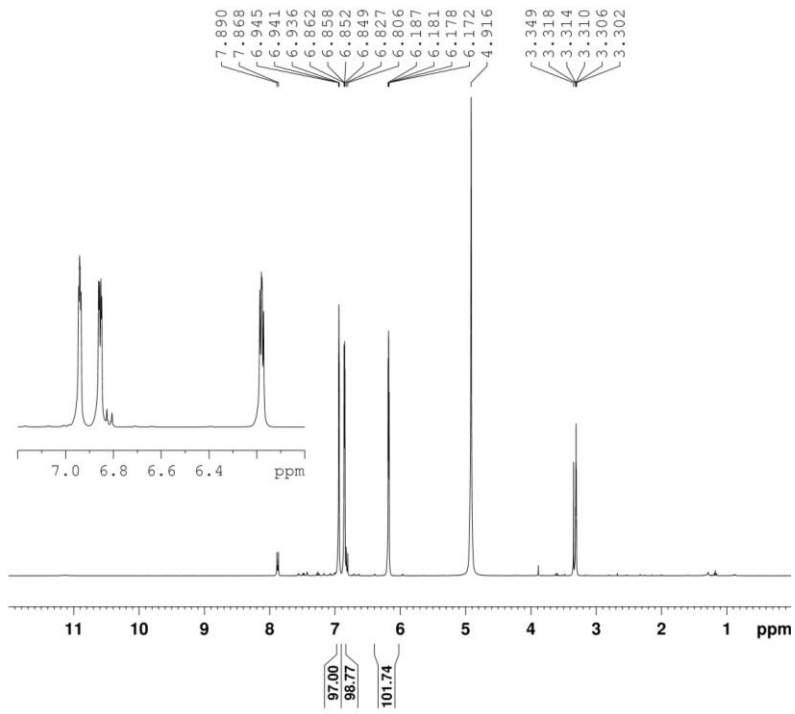


2.5: NMR analysis

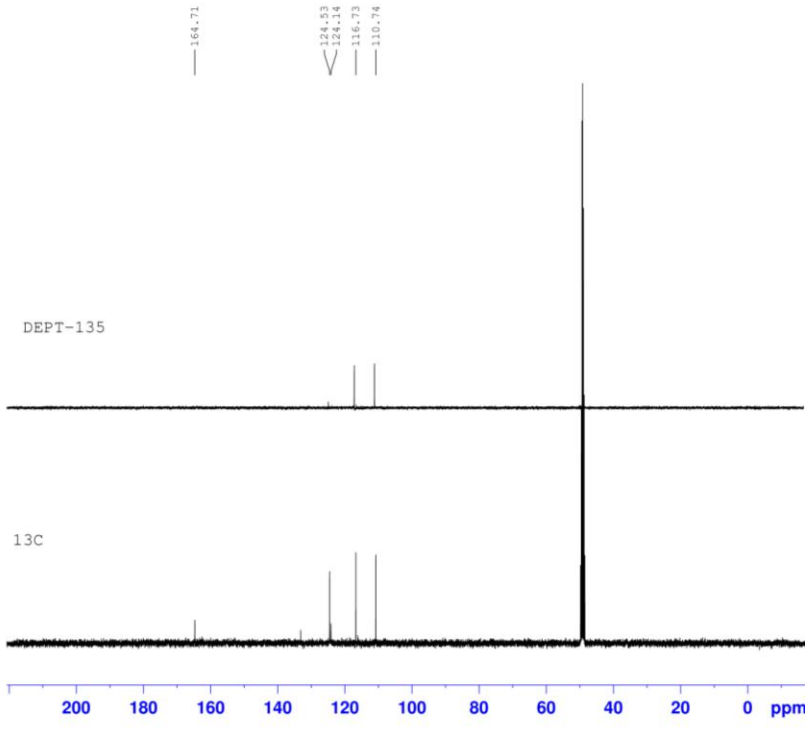
Nuclear magnetic resonance analysis of samples, purified by HPLC as described above and dissolved in deuterated methanol (final concentration 20 mg/mL) were submitted to the Department of Chemistry Research Instrumentation Facility for analysis on a Bruker AVANCE DRX 500 MHz NMR instrument. ^{1}H and ^{13}C NMR data were combined with mass spectral and HPLC retention time information in an effort to identify purified compounds, but only the identity of pyrrole-2-carboxylic acid was

Figure 22: ^{1}H and ^{13}C NMR data for isolated pyrrole-2-carboxylic acid

A) ^{1}H NMR spectrum of pyrrole-2-carboxylic acid



B) ^{13}C NMR spectra of pyrrole-2-carboxylic acid



unambiguously confirmed by NMR spectroscopy (**Figure 22**). None of the isolated compounds yielded spectrometric data which appeared to be similar to fatty acid-type DSF-like molecules.

2.6: Bioassay screening for fractions with DSF-like activity

Several methods of screening crude cellular extracts for DSF-like molecules with signaling actiivty have been reported in the literature. All make use of an *Xcc* transposon-disruption mutant of the *rpfF* gene, *Xcc* 8523, which is unable to synthesize DSF,[27] and therefore lacks several of the phenotypic traits of *Xcc* wild-type. Exogenous DSF, from another bacterial strain co-cultured with *Xcc* 8523, or from extracted or synthesized compounds, can restore these phenotypes in a semi-quantifiable manner. The first bio-detection assay reported by Barber *et al*. involved the restoration of endoglucanase or protease production to *Xcc* 8523 cultured on plates containing 0.5% skimmed milk (to reveal protease production) or carboxymethylcellulose (sensitive to endoglucanases). Quantification of the DSF dose dependence of the restoration required measurement of a "hydrolysis zone" around the DSF-rescued colony of *Xcc* 8523.[27]

An improved bioassay procedure was constructed by Slater *et al*. by fusing a region of DNA containing a Clp-cognate promoter and ribosome binding site (RBS) with the coding sequence of *E. coli*'s β-glucuronidase (*gusA*) and cloning the fusion construct into a broad-host-range vector, pLAFR6, and transforming the resulting plasmid (named pL6engGUS) into the *rpfF* mutant strain *Xcc* 8523, yielding a biosensing strain *Xcc* 8523/pL6engGUS. In the presence of exogenous DSF, *Xcc* 8523/pL6engGUS up-regulates the *gusA* gene and produces β-glucuronidase, which can in turn cleave the chromophore from 5-bromo-4-chloro-3-indolyl-b-D-glucuronide (X-GlcA) and produce a blue halo around the DSF-sensing colony.[29] The level of *gusA* induction is related to the dose of DSF. This biosensing system was employed in the detection of DSF in the

process of its structural elucidation,[28] and to identify DSF and structural analogues from *Xanthomonas oryzae* pv. *oryzae*[37] and the *Burkholderia cepacia* complex.[51]

Notably, the promoter and RBS chosen for constructing the biosensor *Xcc* 8523/pL6engGUS was from the gene for *Xcc*'s major endoglucanase, *engXCA*, the transcription of which was later shown to be up-regulated by the direct binding of Clp to two conserved sites near its -35 sequence.[59] The endoglucanase encoded by *engXCA* is an extracellular 53-kDA protein which is one of wild-type *Xcc*'s major virulence factors,[59] and studies of its regulation contributed to elucidation of the DSF/*rpf* signaling system and its relationship to the transcription factor Clp.

A variation of the DSF biosensor detection method was constructed by Newman *et al*. at UC Berkeley during studies of the insect-vectored plant pathogen *Xylella fastidiosa*. The same promoter-containing sequence (~378 bp) from the *engXCA* gene was fused to a promoterless *gfp* gene on a plasmid conferring spectinomycin and streptomycin resistance, and the resulting construct (pKLN55) was mated into the *rpfF* mutant *Xcc* 8523.[36] The biosensing strain *Xcc* 8523/pKLN55 expresses green fluorescent protein (GFP) in the presence of exogenous DSF, and indeed the plasmid pKLN55 provides a hypothetical method of directly detecting the presence of active Clp by the upregulation of the *gfp* gene and the resulting expression of green fluorescent protein. In 2010, the lab of Steven Lindow at UC-Berkeley graciously provided us with the biosensor strain *Xcc* 8523/pKLN55, and we subsequently developed a similar procedure for exogenous DSF detection.

Following the protocol described by Newman *et al*., we inoculated liquid NYGB medium with biosensing strain *Xcc* 8523/pKLN55 and incubated for 24 hours at 30 °C.

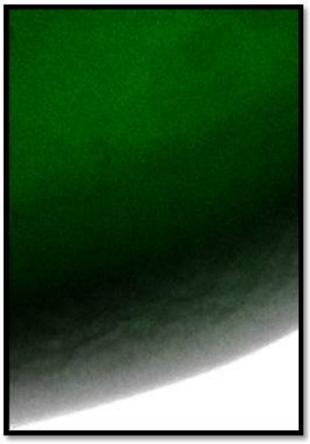
Using a micropipette, 5-10 μL of biosensor culture was spotted onto LB plates along with a control or a putative source of DSF or DSF analogues. The biosensor was incubated at 28-30 °C for a further 48 hours, then visualized using confocal microscopy. An excitation wavelength of 489 nm applied to the biosensor and expressed GFP emitted a wavelength of 509 nm. The method of visualization proved to be important. High magnification of the biosensor made distinguishing between controls difficult, as individual cells occasionally express GFP even in the presence of negative controls, and this background fluorescence could not be distinguished from that induced by DSF detection. We chose to use minimal magnification (10X) of individual "colonies" spotted onto the LB, and visualize the GFP expression of entire colonies against the background of the agar. Negative controls were used to set a zero point of GFP expression, and GFP expression of experimental plates was compared to that of the negative control. For a positive control, to test the continued utility of the biosensor strain under our chosen growth conditions, we placed 1-2 μL of synthetic DSF (Cayman Chemical; diluted to 470 or 47 μM) dissolved in ethyl acetate on a sterile paper disc near the spotted biosensor "colonies."

While some individual bioassays yielded apparently positive results (**Figure 24**) in which crude extracts from *L. enzymogenes* induced levels of GFP in the biosensor *Xcc* 8523 (pKLN55); the result could not be replicated, and seemed to be discredited by moderate levels of GFP in other negative controls (**Figure 25**). It became clear that the level of GFP expression in the biosensing strain could vary between plates, and GFP induction by the same extract could disappear after a few days (**Figure 23**).

Figure 25: Biosensor's GFP expression suggests detection of an extracted *L. enzymogenes* DSF

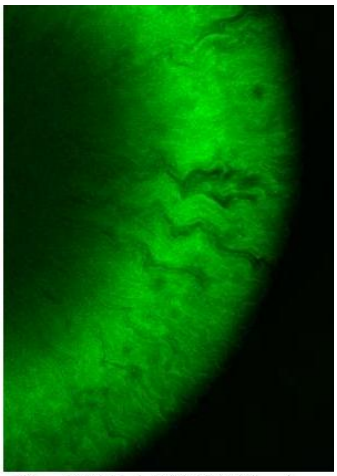
Xcc 8523 (pKLN55) in the presence of synthetic DSF

Xcc 8523 (pKLN55) in the presence of sterile medium



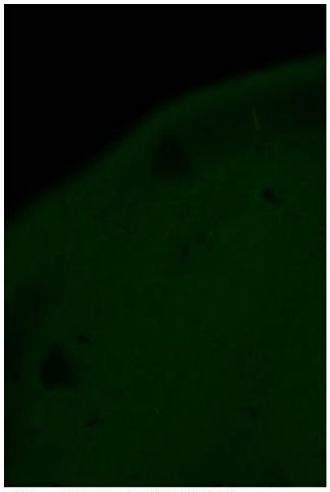
Xcc 8523 (pKLN55) in the presence of *Le*OH11 extract

Figure 23: a false positive

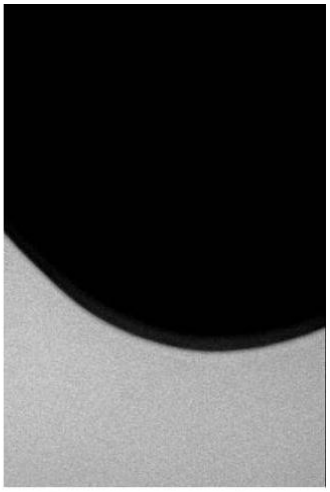


Xcc 8523 (pKLN55) alone on LB agar

Figure 24: Instability of apparently positive bioassay results



Xcc 8523 (pKLN55) near a fractionated extract of *Le*C3



GFP induction vanishes after only a few days of storage at -20 °C

Because the high production of dihydromaltophilin in *L. enzymogenes* C3 and OH11 wild-type strains might interfere with attempts to purify a DSF-like molecule, we carried out several extractions using the LeC3 ΔKR mutant, in which the β-ketoreductase domain of dihydromaltophilin's polyketide synthase had been inactivated by a targeted mutation

(see Chapter 5). Crude extracts of LeC3 ΔKR appeared to induce a moderate level of GFP expression in the biosensor (**Figure 26**) in a manner similar to what had been observed in the wild-type strains, but HPLC analysis of the crude extract did not present any candidate peaks in for a DSF-like molecule. Further fractionation of the extract yielded samples which failed to induce GFP in the biosensor (**Figure 27**), and even the crude extract could not replicate it previous level of GFP induction.

Figure 26: Bioassay of crude extract from *Le*C3 ΔKR

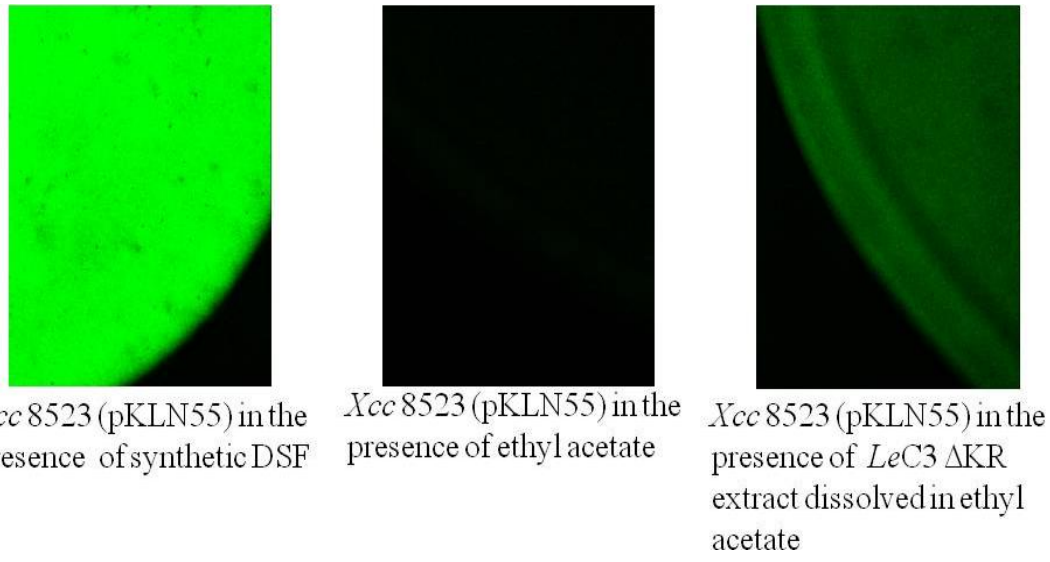
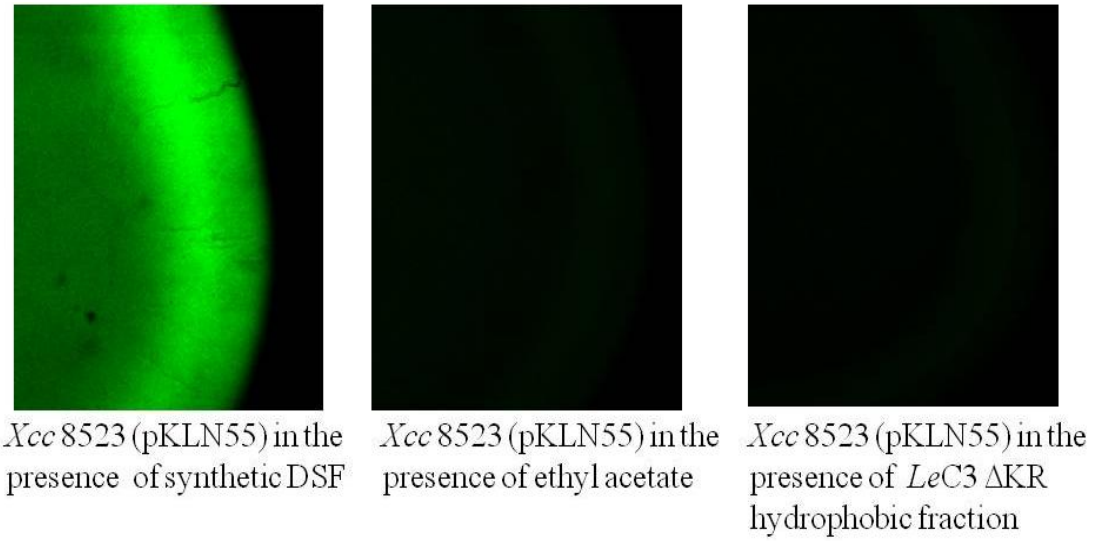
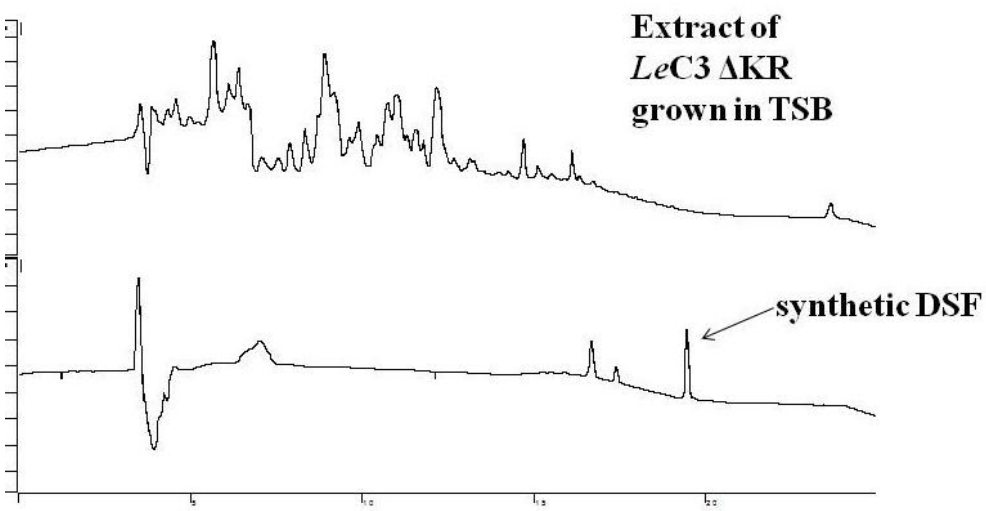


Figure 27: Further fractionations of *Le*C3 ΔKR extract



These results suggested either induction by a highly unstable DSF-like molecule—much more unstable than DSF itself, which retained its bioactivity for many months when stored in ethyl acetate at -20 °C, and has been reported to be chemically and thermally robust[27]—or else that the low and variable levels of GFP induced in the biosensor by *L. enzymogenes* extracts were nonspecific and thus not replicable. A third possibility—that *L. enzymogenes* produces a DSF-like molecule that is both unstable and sufficiently dissimilar to DSF to limit recognition and response by the biosensor *Xcc* 8523 (pKLN55)—cannot be absolutely discounted; but HPLC analysis of the fresh *L. enzymogenes* extracts designed to recover hydrophobic carboxylic acids revealed a conspicuous absence of any notable peaks corresponding to the elution times of fatty acid-type compounds (**Figure 28**). It is also possible that production of DSF-like molecules in *L. enzymogenes* is extremely low under the conditions tested.

Figure 28: HPLC analysis of *Le*C3 ΔKR extract compared with the retention times of known fatty acids (DSF and 10-undecenoic acid)



Although positive controls using synthetic DSF always provided high levels of GFP expression, none of our experimental results showed unambiguously positive results in which extracts of *L. enzymogenes* cultures induced levels of GFP expression comparable with the positive control. Low but varying levels of GFP expression was often observed even in negative controls (i.e. when the biosensor was exposed to sterile media, pure methanol, or pure ethyl acetate), and distinguishing a weakly positive signal from a negative signal became a disturbingly subjective exercise. Analysis of all the results appears to support the conclusion that none of the tested *Lysobacter* strains, either as co-cultures with the biosensor or as crude or fractionated extracts added to the plate, induced GFP expression significantly and unambiguously above background. This conclusion is strengthened by comparison of our *Lysobacter* results with the positive results reported by the Lindow lab using *Xylella*.[36] Although contradicting our hypothesis that *L. enzymogenes* possesses a cell-cell signaling system analogous to the *rpf*/Clp system in *Xcc*, this conclusion concurs with our failure to detect any significant DSF-like compounds during HPLC and MS analysis of *Lysobacter* culture extracts.

2.7: Summary of results

In summary, we believe that these data demonstrate the utility of the isolation and analytical techniques applied to the extracts of *Lysobacter enzymogenes* cultures. Our failure to detect any DSF-like molecules using these techniques do not, of course, preclude their presence in these cultures, perhaps at concentrations below the detection limits; nor can we exclude the possibility that DSFs are present only at specific times in the growth curve, or are produced under different growth conditions than those we tested, supplying some signal entirely unrelated to dihydromaltophilin production. However, the data we have amassed here seems to indicate that production of dihydromaltophilin does not correlate with any accumulation of a DSF-like molecule in the culture supernatant, and therefore suggests that the Clp homologue in *L. enzymogenes* operates independently of the gene products of the *rpf* homologues. Genetic experiments, discussed in Chapter 3, further support this hypothesis, and even suggest that the Clp homologue in *L. enzymogenes* may be constitutive in nature, acting in a [c-di-GMP]-insensitive manner.

The two representative molecules discussed in the chapters on analytical techniques are interesting in their own right, despite the absence of any apparent connection to Clp or dihydromaltophilin production in *L. enzymogenes*. Free pyrrole-2-carboxylic acid has been isolated from several bacterial species,[60] and its biosynthesis from proline, prior to incorporation into larger natural products, has been investigated in a number of biosynthetic pathways.[61-64] We were intrigued by the fact that pyrrole-2-carboxylic acid is known to furnish the precursor for pyrrole moieties in several interesting natural products,[61, 65] including the narrow-spectrum antibiotic hormaomycin which also serves as a signaling molecule, inducing morphological

alterations and secondary metabolism in *Streptomyces* species.[63] The functional parallels between hormaomycin and DSF led us to suspect that pyrrole-2-carboxylic acid might function independently, either (or both) as a signal or an antibiotic; or it might be incorporated into a larger, more labile metabolite with signaling or bioactive properties. A search of *Le*OH11's genomic sequence, using proline dehydrogenase CloN3 (from the biosynthetic gene cluster of the pyrrole-containing antibiotic chlorobiocin [66]) to interrogate the genome, revealed a set of genes which appeared to encode enzymes possibly associated with the biosynthesis of pyrrole-2-carboxylic acid, closely coupled with a nonribosomal peptide synthase (NRPS) predicted to activate cysteine. We speculate that this cluster encodes enzymes for the incorporation of pyrrole-2-carboxylic acid into a larger structure, which might degrade into pyrrole-2-carboxylic acid as the *Lysobacter* culture ages. Inactivation of the key genes in this cluster, and examination of the resulting *Le*OH11 mutants for loss of pyrrole-2-carboxylic acid and possibly other morphological changes, might reveal a pyrrole-containing bioactive or signaling metabolite.

Vanillic acid (**1**) (**Figure 16**) is suspected to contribute the benzoic moiety to *L. enzymogenes*'s distinctive yellow pigment, which is believed to be a non-brominated aryl polyene similar to xanthomonadin.[67] The gene cluster for yellow pigment biosynthesis in *Le*OH11 has been located by our group and confirmed by mutagenesis studies. While the yellow pigments of *Xanthomonas* species, the xanthomonadins, contain a brominated benzoic moiety,[68-69] mass spectral analysis of the *Lysobacter* analogues of xanthomonadin indicated an absence of bromination, and analysis of the *Le*OH11 gene cluster associated with the yellow pigment suggested the likely presence of an *O*-

methylated moiety. While we have not established any connection between the presence of free vanillic acid in the supernatant of *L. enzymogenes* cultures and the yellow pigmentation of *L. enzymogenes* cells, it is interesting to note the similarities between the xanthomonadin precursor, 3-hydroxybenzoic acid (sometimes designated *diffusible factor* (DF) in the literature [70-71]). The diffusible factor, 3-hydroxybenzoic acid (**Figure 29**), in addition to serving as the precursor of xanthomonadin biosynthesis, appears to have some overlapping function with the unsaturated fatty acid DSF, in that DF is reported to "modulate bacterial survival, H_2O_2 resistance, and virulence [7]," although it remains unclear whether DF is serving as a true signal in its own right, or whether the observed phenotypes derive entirely from the loss of DF in DF-deficient mutants. We have not pursued further studies of the effects of vanillic acid on *L. enzymogenes* morphology or secondary metabolism, but it remains a point of interest with possible connection to the intercellular signaling of this species.

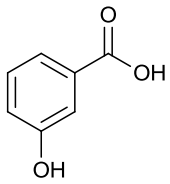


Figure 29: 3-hydroxybenzoic acid, *Xcc*'s diffusible factor (DF) [7]

Chapter 3: Rational engineering of Clp

Because the Clp-dependence of dihydromaltophilin production was so well established, we attempted a genetic approach, complementary to our experiments with the DSF/*rpf* system of *Lysobacter*, to probe the mechanisms of Clp's regulation. The relationship between Clp, the RpfC/RpfG two-component system, and the RpfF/RpfB DSF biosynthetic system has been well studied and largely elucidated in *Xcc*, in which Clp responds to the intracellular concentration of cyclic-di-GMP, an ubiquitous bacterial second messenger which is synthesized by GGDEF protein domains and degraded by EAL or HD-GYP protein domains, thus transducing extracellular stimuli and regulating numerous complex processes.[72-73] Understanding Clp's response to c-di-GMP in *Xcc* led us to attempt *in vivo* modifications of Clp in *L. enzymogenes*, which might support our hypothesis of the *rpf*/Clp relationship in *L. enzymogenes*. By changing the behavior of the transcription factor Clp, we sought to elucidate the nature of the intracellular portion of this regulatory network.

3.1: A proposal for a constitutive Clp transcription factor

A major advance in the understanding of DSF signaling in *Xcc* was the publication of the solved crystal structure of *Xcc*'s Clp in 2010, along with biochemical experiments which demonstrated the effects of c-di-GMP on Clp's activity.[1] Wild-type *Xcc*Clp, along with several mutant versions, were expressed in *E. coli* BL21, purified, and crystallized, and subjected to X-ray diffraction to a resolution near 2.3 Å. The crystal structure of *Xcc*Clp provided a template for modeling the binding of c-di-GMP with Clp *in silico*, and electrophoresis mobility shift assays (EMSA) using the pure protein in the presence of varying concentrations of c-di-GMP revealed that *Xcc*Clp's binding to the

engA promoter DNA[59] is actually inhibited by c-di-GMP.[1] In independent experiments, the Clp homologue from *Xanthomonas axonopodis* pv. *citri* was also expressed *in vitro* and found to be inhibited by elevated c-di-GMP concentrations.[74] EMSA experiments, using wild-type Clp and some targeted alanine mutants of Clp, revealed that binding of the allosteric inhibitor c-di-GMP could be much reduced by replacing with alanines the key residues which *in silico* analysis indicated were involved in interactions with c-di-GMP.[1]

3.1.1 Design and construction of the constitutive Clp mutants

Alignments of the Clp homologues from *Le*C3 and *Le*OH11 with *Xcc*Clp revealed a very high degree of overall similarity and, more importantly, the almost total conservation of those residues which the *Xcc*Clp dimer's crystal structure and *in vitro* mutagenesis studies had shown to be essential for binding to c-di-GMP and for maintaining the Clp dimer's intrinsic DNA-binding conformation in the absence of c-di-GMP (**Figure 5**).[2-6] Key residues involved in Clp's DNA-binding conformation were identified from the crystal structure, which, when compared and contrasted with the crystal structure of the cAMP-receptor protein (CRP) from *E. coli*, revealed the steric and electronic effects which stabilized Clp in its DNA-binding conformation without the presence of any small-molecule effector.[1] The key residue involved in Clp's intrinsic DNA-binding conformation was found to be Glu-99, which corresponded to CRP's Ser-84. In *E. coli*'s CRP, Ser-84 is heavily involved in interactions with CRP's effector, cyclic AMP; but in *Xcc*Clp, the Glu-99 residue presents a larger, negatively-charge side chain which is flipped in the opposite direction, interacting with Arg-150 to form a salt bridge which is not possible in *E. coli*'s CRP, in which Glu-130 is the residue corresponding with

*Xcc*Clp's Arg-150.[1] Indeed, *Xcc* mutants in which Glu-99 was "restored" to serine exhibited altered virulence and decreased upregulation levels of Clp's regulon,[34] supporting the involvement of Glu-99 in stabilizing the DNA-binding conformation. Elsewhere, the replacement of CRP's Ala-145 with Val-165 (Val-164 in *Le*Clp) provides the steric interactions needed to stabilize Clp's α-helices in their DNA-binding position; also, Arg-195 and Asp-162 participate in salt bridges or H-bonding which are not possible in CRP. Indeed, substituting Gly-142 in CRP with an aspartate residue had been previously shown to eliminate CRP's activity dependence upon cAMP; and CRP's Gly-142 corresponds with *Xcc*Clp's Asp-162.[22] In summary, all of the residues implicated in stabilizing the DNA-binding conformation of *Xcc*Clp are conserved in *Le*Clp, and all of the residues involved in binding to the allosteric inhibitor c-di-GMP are conserved between the two proteins, except for the substitution of *Xcc*Clp's key Asp-70 with Glu-69 in *Le*Clp. These comparisons strongly support our hypothesis that *Le*Clp is a c-di-GMP receptor protein which is active when intercellular c-di-GMP concentration is below the binding threshold.

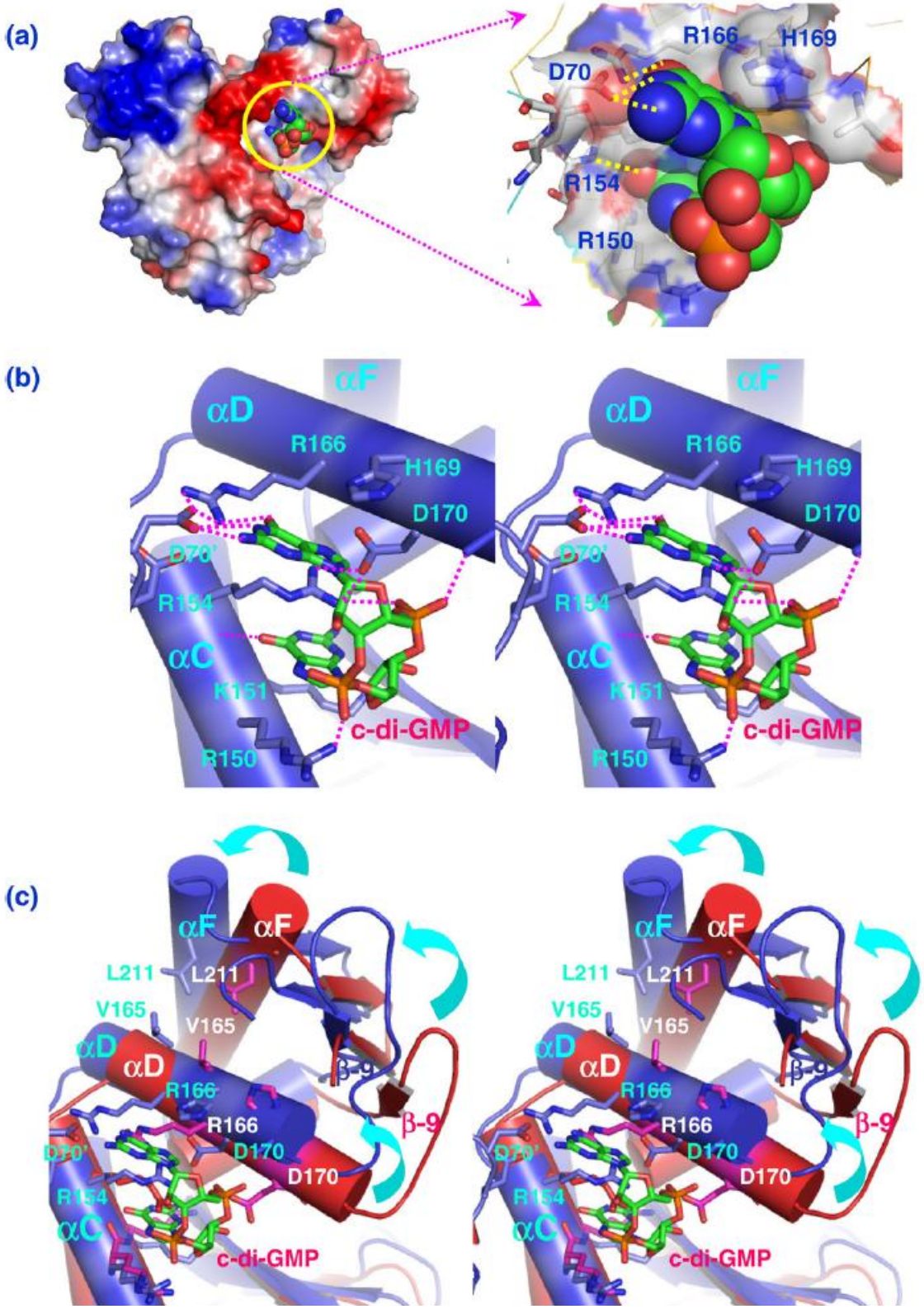
Molecular modeling studies indicated that the key residues involved in interactions between the Clp dimer and c-di-GMP include Asp-70 (corresponding to Glu-69 in *L. enzymogenes*), Arg-154 (Arg-153 in *L. enzymogenes*), Arg-166 (Arg-165 in *L. enzymogenes*), and Asp-170 (Asp-169 in *L. enzymogenes*). Alanine substitutions for these four residues, and the alanine substitution of both R-166 and D-170, generated five mutant versions of Clp which exhibited from 8- to 27-fold elevations in the *in vitro* dissociation constants (K_D) of c-di-GMP with Clp, confirming the *in silico* experiments which showed that these four residues are essential for Clp's binding to its inhibitor.[1]

The alanine-substitution mutants which showed the greatest decrease in binding affinity for c-di-GMP were the D70A mutant (27-fold reduction in binding) and the double mutant R166A/R170A (25-fold reduction in binding). However, all of these alanine mutants also suffered some loss of *in vitro* binding to their cognate DNA sequence (the *engA* promoter [59]).

We reasoned that selected mutations of one or more of these four conserved (except for change of *Xcc*Clp's D70 to E69 in *Le*Clp) residues would result in a constitutively active Clp, which, if expressed in *L. enzymogenes*, would have the same effects on *L. enzymogenes* as of a constant high concentration of its hypothetical diffusible signal factor (which would be detected by the RpfC/RpfG two-component system and lead, by decrease of intercellular [c-di-GMP], to activated Clp). While Chin *et al.* had only expressed their "constitutive" Clp mutants *in vitro*,[1] we decided to attempt *in vivo* expression of similar c-di-GMP insensitive, constitutively active Clp mutants in the hope that such *L. enzymogenes* Clp mutants would mimic the behavior of wild-type *L. enzymogenes* in the presence of its hypothetical DSF.

To choose the best residues to target for substitution in the constitutive Clp mutants, we consulted the data provided by Chin *et al.* from GemDock modeling of c-di-GMP binding to Clp, as well as the biochemical data for Clp and its inhibitor.[1] Asp-70, in addition to forming a salt bridge with Arg-166, interacts via H-bonds with N1 and N2 atoms of one guanine of c-di-GMP, which binds in the *cis*-conformation with the two guanine bases stacked on top of one another. Arg-166 also forms an H-bond with the top guanine base via O6. The atom O6 of the bottom guanine base forms an H-bond with an amide of the peptide backbone, and Arg-154's side chain interacts twice, by an H-bond

Figure 30: GEMDOCK modeling of c-di-GMP docked to *Xcc*Clp. Figure 5 from [1], reprinted by permission (license number 3102080210510, 4 Mar. 2013, Elsevier Ltd. Kidlington, Oxford, UK). A) *Xcc*Clp dimer with c-di-GMP docked. B) Specific interactions between c-di-GMP and *Xcc*Clp dimer. C) Rotation of *Xcc*Clp's helices in response to c-di-GMP docking (apo-*Xcc*Clp in red, *Xcc*Clp/c-di-GMP complex in blue).



between a guanido NH and c-di-GMP's phosphate oxygen, and by π-π interaction between the guanido group and the six-membered ring of the bottom guanine. Asp-170 forms a salt bridge with Arg-154 (**Figure 30**).[1] These interactions "firmly wedge" the c-di-GMP inhibitory ligand between two α-helices of each Clp monomer, setting off a cascade of interactions which alter the dimer's overall conformation and prevent DNA binding. While it was tempting to consider the construction of Clp mutants in which the polarity of the key residues was reversed (for example, by substituting Arg-154 (Arg-153 in *Le*Clp) with glutamate), thus actively setting up repulsions between Clp and its inhibitor, we judged that the danger of thereby disrupting the native DNA-binding conformation outweighed the possible benefits. In addition, by choosing to attempt *in vivo* expression of simple alanine substitutions, we were able to benefit from the *in vitro* characterizations of such mutants by Chin *et al*, which provided data (via isothermal titration calorimetry) for the binding of the alanine mutants both to Clp's cognate DNA and to c-di-GMP, although the two sets of data were obtained under different buffer conditions (low salt (80 mM NaCl, 20 mM Tris (pH 8.0), 20 mM LiCl) and high salt (250 mM NaCl, 20 mM Tris (pH 8.0), 100 mM LiCl), respectively). The data thus reported are summarized in **Table 3**:

Table 3: Comparisons of selected residues' effects on DNA- and inhibitor-binding affinities. [1]

Clp version	*in vitro* K_d (c-di-GMP)	*in vitro* K_d (DNA promoter binding)	K_d-inhibitor/ K_d-DNAsubstrate	analogous *Le*Clp mutation

native Clp	3.5uM (1-fold)	14 nM (1.0-fold)	250	-
D70A	95 uM (27-fold)	28 nM (2.0-fold)	3392	E69A
R154A	30 uM (8-fold)	19 nM (1.4-fold)	1578	R153A
R166A	40 uM (11-fold)	25.6 nM (1.8-fold)	1562	R165A
D170A	38 uM (11-fold)	17 nM (1.2-fold)	2235	D169A
R166A/D170A	88 uM (25-fold)	25.2 nM (1.8-fold)	3492	R165A/ D169A

The ideal *in vivo* constitutive Clp would have maximum affinity for its cognate promoter and minimum affinity for its inhibitor. Table 3 shows that the double alanine mutant combines these two quantities to the greatest extent, but we were put off by the relatively serious effect on DNA affinity and (more especially) by the relative difficulty of simultaneously introducing two single-residue substitutions. We chose to attempt *in vivo* expression of the D70A Clp mutant (E69A in *Le*Clp), which exhibits the lowest affinity for the inhibitor, and of the R154A mutant (R153A in *Le*Clp) which might still exhibit some of c-di-GMP's inhibitory effects.

The effects of such a constitutive Clp on living bacterial cells were uncertain. Should the mutant Clp misfold and lose all activity, we would have expected the mutants

to exhibit the well-documented phenotype of the Clp-deletion mutants (loss antifungal activity and gliding motility, increased yellow pigment production, and increased extracellular polysaccharide production [3]). Another possibility was that heightened Clp activity would prove fatal, perhaps as a result of excessive secondary metabolism, with a concomitant high expenditure of energy on non-essential functions or overproduction of toxic compounds. Thirdly, the presence of an uninhabitable Clp might result in a "super-predator," a *L. enzymogenes* mutant which exhibited very high (but non-fatal) production of bioactive compounds such as dihydromaltophilin and its analogues, but possibly other secondary metabolites as well. We had previously noted, by investigation of the sequenced genome of *Le*OH11, the presence of several gene clusters, seemingly associated with secondary metabolites, whose products were unknown (Liangcheng Du, Guoliang Qian, unpublished data). We also considered the possibility that we might be able to observe the production of new metabolites, or the dramatic increase in metabolites previously observed, in the extraction profiles of the proposed constitutive Clp mutants, which might lead to the identification of novel bioactive natural products.

Construction of the two constitutive Clp mutants of *L. enzymogenes* followed a standard protocol for bacterial mutagenesis. Genomic DNA was isolated from *Lysobacter enzymogenes* C3 using standard procedures and used as template for PCR to amplify the entire 690-bp Clp gene, which was double-digested with *Eco*RI and *Xho*I (New England Biolabs) and ligated into broad-host-range suicide vector pEX18Gm[75] which was cut with the same restriction enzymes. The resulting vector pEX18Gm-Clp served as template for a mutagenesis procedure using QuikChange II site-directed mutagenesis kit (Agilent Technologies, catalog #200523) to change the codon for Glu-69 from GAG to

GCC, and the codon for Arg-153 from CGC to GCG, using primer pairs E69A/E69A-r and R153A/R153-r. Because direct transformations of the provided *E. coli* XL-1 Blue competent cells using the mutagenesis PCR mixture were not successful, we used the mutagenesis PCR mixture as template for amplification of the mutated Clp gene and prepared mutagenic plasmids pEX18Gm-ClpE69A and pEX18Gm-ClpR153A, following the same procedures and conditions previously used to prepare pEX18Gm-Clp. These mutagenic plasmids were transformed into chemically competent *E. coli* S17-1 cells using heat shock, and transformants were selected on LB agar plates supplemented with 25 μg/mL gentamicin. The two mutagenic plasmids were mini-prepped and sequenced to verify the integrity of the base sequence. Plasmid pEX18Gm-ClpE69A contained no errors and the codon for Glu-69 was changed to encode alanine as designed; pEX18Gm-ClpR153A contained the expected codon change for Arg-153 to alanine, but also included a silent mutation in Ala-129's codon, from GCG to GCA. Although aware of the possibility that recombination would take place at the site of the silent mutation instead of at the site of the targeted codon for Arg-153, we proceeded to use both mutagenic plasmids for conjugation between plasmid-containing *E. coli* S17-1 and *Lysobacter enzymogenes* C3. Multiple single colonies of transformed *Le*C3 were selected on 10% TSB agar plates supplemented with 25 μg/mL gentamicin, as well as 25 μg/mL kanamycin which *L. enzymogenes* naturally resists. These putative single-crossover recombinants were confirmed by PCR (using 1-2 μL of raw culture as template) to contain the expected pEX18Gm-ΔClp plasmid inserted into the Clp gene; the single-crossover recombinants were then inoculated into 10% TSB with no antibiotics, incubated for 24 hours at 30 °C with shaking, diluted 1000 to 5000 times in 10% TSB

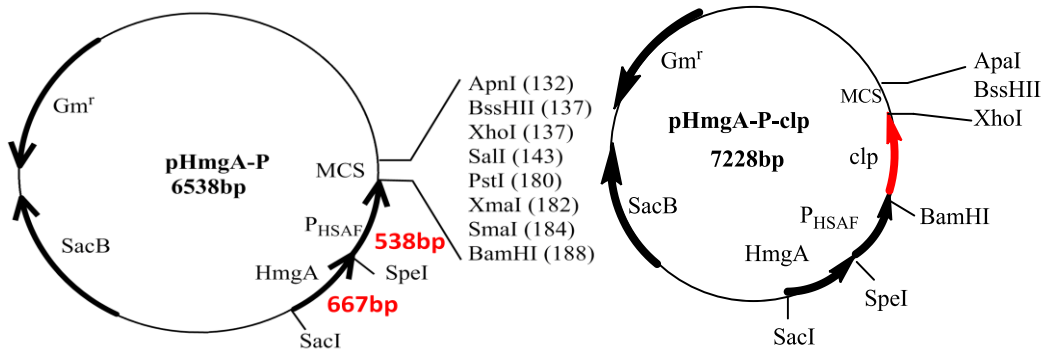
supplemented with 10% sucrose (to encourage double-crossover recombination to take place) and grown a further 6 hours before being plated out on 10% TSB agar plates supplemented with 10% sucrose. Single colonies were putative double-crossovers containing Clp E69A or R153A mutations. These were grown in 10% TSB supplemented with 25 μg/mL kanamycin and 10% sucrose and used for a diagnostic PCR which amplified an 1147-bp region extending upstream and downstream of the 690-bp Clp gene. These 1147-bp sequences were purified and sequenced to check for the presence of the mutant Clp sequence.

Although multiple double-crossovers mutants were isolated by this method, we were unable to locate any that carried the desired mutation, either for Glu-69 or Arg-153 (although several were found which contained the silent Ala-129 mutation). This led us to suspect that the mutant Clp sequences might be sufficiently harmful to the cell that the second recombination event strongly selected against the Clp mutants in favor of reversions to wild-type Clp sequence. Another possibility was that the placement of the two mutations was too near either end of the homologous region, and that the second crossover was thus more likely to simply eliminate the inserted plasmid rather than introduce the desired mutation. To test which of these hypothetical situations was responsible for our failure to isolate any Clp mutants, we subcloned the portion of the Clp gene containing the introduced mutations from the pEX18Gm-ΔClp plasmids into the vector pHmgA-P-Clp2, described below.

The vector pHmgA-P (**Figure 31**) was constructed in our lab to serve as an easily-selected expression vector designed to insert directly into *L. enzymogenes*'s genome. Into the suicide vector pJQ200SK[13] a 667-bp portion of the *L. enzymogenes* homologue of

the homogentisate 1,2-dioxygenase (*hmgA*) gene was cloned by means of *Sac*I/*Spe*I double-digestion and ligation, and a 538-bp region lying upstream from the dihydromaltophilin gene cluster in *L. enzymogenes* OH11 was cloned into the vector by *Sac*I/*Bam*HI digestion and ligation. This 538-bp region contains the putative promoter for *L. enzymogenes*'s dihydromaltophilin biosynthetic genes, and was positioned to lie just upstream from the multiple-cloning-site (MCS) of pJQ200SK. The resulting vector pHmgA-P could therefore accept any genes of interest, which would presumably be under the control of the dihydromaltophilin promoter and would be expressed in *L. enzymogenes* under any conditions which favored production of dihydromaltophilin; indeed, we regarded it as likely that this promoter actually contained a Clp binding site, although several analyses failed to identify any obvious Clp-cognate sequence. Additionally, single-crossover recombination at the HmgA homologue produces a mutant which cannot process homogentisic acid (a catabolite of tyrosine) to maleylacetoacetate,[76] resulting in homogentisate's accumulation within the cells and the eventual appearance of a reddish-brown halo around the single-crossover transformant. Thus, the use of vector pHmgA-P provides a platform for the expression of genes in *L. enzymogenes* with a facile selection of genomic insertion mutants by means of the appearance of the red/brown pigmentation.

Figure 31: Map of pHmgA-P-Clp2, constructed from *Le*C3 inserts cloned into pJQ200SK [13]



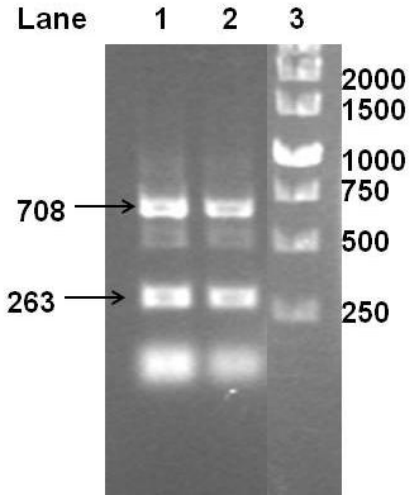
The vector pHmgA-P-Clp2 was constructed in our lab to restore the *clp* gene to the *clp*-deletion mutant of *Le*C3 constructed by Koboyashi *et al*: *Le*DC211 (Yunxuan Xie, unpublished data, [3]). Because the dihydromaltophilin promoter (which is putatively Clp-dependent) aboard pHmgA-P could not be upregulated in Clp-deficient mutant, pHmgA-P was provided with the entire gene sequence employed by Koboyashi *et al*. in their complementation of the *clp*-deletion DC211, which included 200-bp ahead of the start codon and 18-bp following the stop codon, thus supplying vector pHmgA-P-Clp2 which successfully restored antifungal activity, gliding motility, and the other wild-type phenotypes to the *clp* mutant DC211 (see **Figure 1**).

This "DC-clp2" complementation mutant served as the negative control for the expression of the two constitutive Clp genes, Clp E69A and Clp R153A. The two mutation-containing sequences were subcloned from pEX18Gm-ClpE69A or pEX18Gm-ClpR153A into pHmgA-P-Clp2, using the *clp* gene's naturally-occurring BsiWI and PstI sites, and *E. coli* S17-1 strains containing the new pHmgA-P-ClpE69A and pHmgA-P-ClpR153A were used to introduce the vectors into Koboyashi's *clp*-deletion mutant *Le*DC211. Resulting transformants were selected on 10% TSB agar plates containing 25 μg/mL gentamicin and 25 μg/mL kanamycin; single colonies were picked up and grown

in 10% TSB with 40-50 μg/mL gentamicin, and confirmed by PCR, sequencing, and the appearance of the reddish-black pigmentation to be DC211 mutants with the constitutive Clp constructs inserted into the *hmgA* homologue.

Although this approach suffered from the possible obscuring of effects on pigment production (due to the slow increase in the reddish-brown pigment), it had the important advantage of directly selecting for the constitutive Clp mutants on gentamicin-containing plates; by linking the expressible constitutive *clp* constructs with gentamicin resistance, we were able to answer the question of whether the introduction of a constitutive Clp protein would prove fatal to *L. enzymeogenes* cells. This question turned out to be answered in the negative, and we were able to isolate true complementation mutants, confirmed by PCR and sequencing (**Figure 32**), and to partially characterize the bioactive properties of the constitutive Clp mutants.

Figure 32: Electrophoresis confirms generation of mutagenic plasmid inserted into *Le*DC mutants. The amplicons shown were sequenced to confirm their identity.



Lane 1-2: PCR product using primers designed to amplify entire *clp* gene in *clp*-complemented strains
Lane 3: standard DNA fragments

These PCRs demonstrate that the *Le*DC Clp complementations contain two *clp* fragments. The 708-bp amplicon is the gene inserted aboard pHmgA-P-Clp; the 263-bp amplicon is the natural *clp* gene from which 445-bp were deleted by Koboyashi *et al*.

3.1.2 Analysis of constitutive Clp mutants

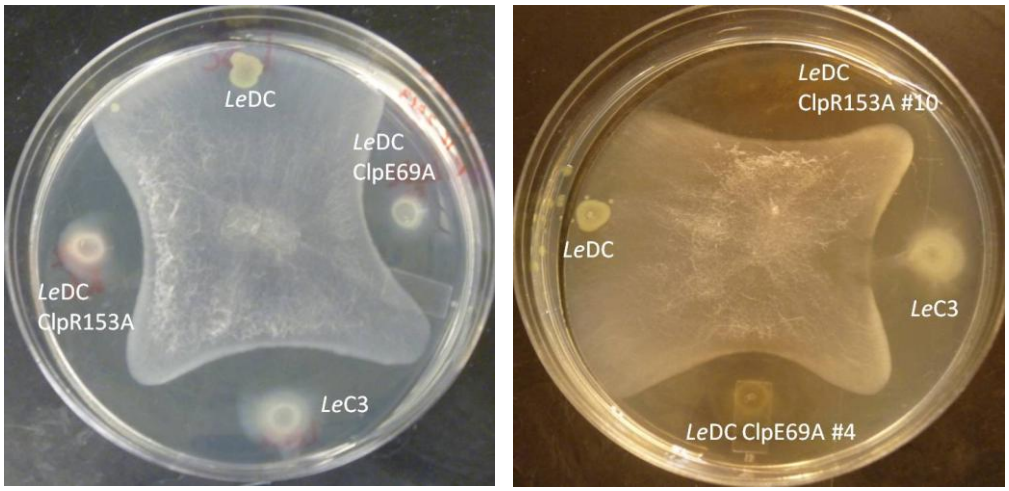
The transformants, once the integrity of their mutant *clp* genes had been confirmed by sequencing, were subjected to a number of bioactivity and biochemical assays to evaluate the effects of the mutations. Since the replacement of Glu-69, or Arg-153, with alanine was expected to result in a Clp transcription factor which could not be inhibited by normal concentrations of cyclic-di-GMP, we expected that any genes in Clp's regulon would be permanently upregulated without regard to environmental stimuli (or the absence thereof).

Our primary interest was in comparing the level of antifungal activity between the *Le*DC211 mutants complemented with the different versions of Clp. Complete restoration of antifungal activity by the restoration of Clp would indicate that the expression of the dihydromaltophilin biosynthetic genes had returned to wild-type levels. On the other hand, the failure of the Clp-complemented *Le*DC strains to exhibit antifungal activity would indicate either a misfolded or inactive Clp protein. Since we were using a construct that had successfully restored Clp to *clp* mutants of *Le*C3 in the experiments of Koboyashi *et al.*,[3] we did not expect a failure of *clp* upregulation to play a role in the mutants' phenotypes.

Antifungal bioactivity assays were carried out using the filamentous fungus *Fusarium verticilliodes*. Mycelia from a previously-cultured plate of *F. verticilliodes* were transferred to the center of a 10% TSB agar plate and allowed to grow at 30 °C for 24-48 hours. *L. enzymogenes* strains were cultured in liquid 10% TSB at 30 °C, with appropriate antibiotics (25 μg/mL kanamycin for all wild-type or mutant strains; 25-50 μg/mL gentamicin for pHmgA-P-ΔClp mutants) for 24 hours. 1-5 μL of the *L.*

enzymogenes culture were pipetted near the edge of the 10% TSB agar plate on which *F. verticilliodes* was growing radially from the center, and the agar plates were incubated at

Figure 33: Antifungal activity of wild-type *Le*C3 compared to *clp*-deletion mutant *Le*DC211 and strains complemented with the constitutive Clp mutants (E69A and R153A).



Left: Wild-type and *clp*-deletion compared to Clp-E69A and Clp-R153A complementations. These mutants did not cross over at the *hmgA* homologue. Right: Clp-E69A and R153A mutants which underwent *hmgA* insertion are compared to wild-type and deletion strains.

30 °C for 2-5 days until fungal growth had reached the edge of the agar plate. *L. enzymogenes* strains possessing antifungal activity traced out a clear zone of inhibition where fungal growth was entirely absent, while strains lacking antifungal activity were tightly surrounded by the advancing mycelia (**Figure 33**).

These assays reveal that, although antifungal activity had clearly been restored to the Clp-complemented strains of *Le*DC211, there was no manifest difference in the bioactivity of those *Le*DC strains complemented with the constitutive Clp (E69A or R153A) and those which were complemented with the wild-type sequence of *clp*; additionally, no major difference could be detected between *Le*C3 wild-type and any of the Clp-complementation strains (**Figure 33, Figure** 34), although the constitutive Clp mutants may have provided a slightly larger inhibition zone than *Le*DC complemented with the wild-type *clp* sequence. Considering the level of c-di-GMP insensitivity displayed by the *Xcc*Clp analogues of these alanine mutant Clp proteins, this lack of any alteration in morphology or bioactivity suggested the possibility that *Le*Clp may already be constitutive (i.e. insensitive to inhibition by c-di-GMP) in *L. enzymogenes*; or, at least, the DNA-binding activity of Clp is, by some mechanism, decoupled from *L. enzymogenes* analogue of *Xcc*'s *rpf* system involving RpfC/RpfG and RpfF/RpfB. In support of this theory, we

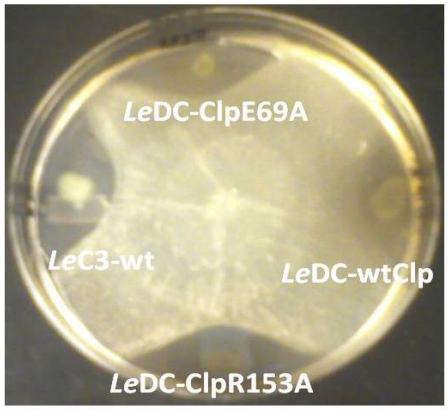
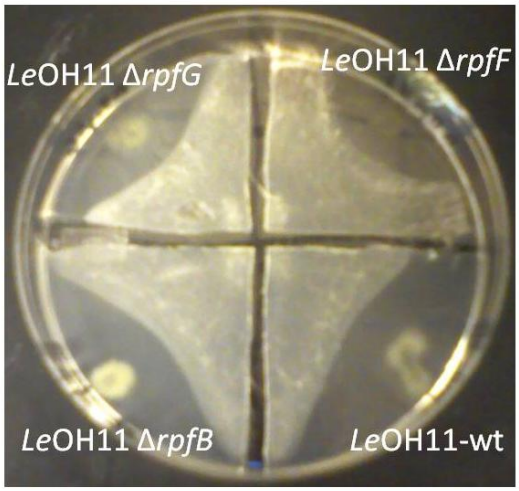


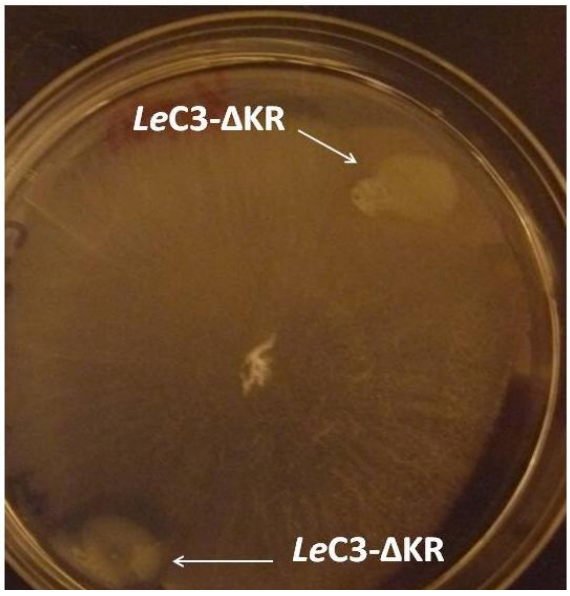
Figure 35: Antifungal activity of *Le*C3 and three strains of *clp*-deletion mutant *Le*DC211 complemented with wild-type Clp, constitutive Clp-E69A, and constitutive Clp-R153A

Figure 34: Antifungal activity of *Le*OH11 compared to three mutants of *rpf* homologues. Loss of *rpfB*, *rpfF*, and *rpfG* does not cause any significant loss of antifungal activity. To prevent diffusion of any DSF-like molecule from wild-type *Le*OH11 to the mutant strains, agar was cut out between each strain.



carried out a further assay of the antifungal bioactivity of *Le*OH11 mutants[2] of *rpfG* (encoding the response regulator phosphodiesterase [32]), *rpfF* (encoding the DSF synthase, an enoyl-CoA hydratase [51]), and *rpfB* (encoding a putative acyl-CoA ligase). Surprisingly, we observed no significant loss of antifungal activity in these mutants which, presumably, lack a functional *rpf*-like signaling system (**Figure 35**). The inhibition zones observed in these antifungal activities can be confidently attributed to dihydromaltophilin production by means of comparison with the *Le*C3 ΔKR mutant, which cannot synthesize dihydromaltophilin but retains an active Clp and all other forms of antimicrobial activity. The ΔKR mutant produces no inhibition zone, but can repel the advance of the fungal mycelia by direct contact, as the *Lysobacter* colony glides across the agar surface (**Figure 36**).

Figure 36: Antifungal activity of *Le*C3 ΔKR compared wild-type *Le*C3 and to *rpfF* and *rpfG* mutants of *Le*OH11.



The successful inhibition of *F. verticilliodes* by the *Le*ΔRpfF deletion mutant lent support to the implications of our previous observation that ethyl-acetate extractions of this same mutant revealed no major decrease in the concentration of dihydromaltophilin. Although this mutant is expected to be incapable of producing a diffusible signal factor similar to *Xcc*'s cis-11-

[2] The *rpfG* and *rpfF* mutants were supplied by Zhou Xue. 40. Guoling Qian, Y.W., Yiru Liu, Feifei Xu, Yawen He, Liangcheng Du, Jiaqin Fan, Baishi Hu, Fengquan Liu, *Lysobacter enzymogenes uses two distinct cell-cell signaling for differential regulation of metabolite biosynthesis.* 2013. The *rpfB* mutant was generated by a gene-disruption strategy in our lab by Haotong Chen.

methyl-dodecen-2-oic acid, it remained capable of a high level of dihydromaltophilin production and concomitant antifungal activity; and so, presumably, its Clp transcription factor had retained its ability to upregulate the dihydromaltophilin gene cluster. Additionally, none of the other phenotypic traits characteristic of the *Le*DC clp-deletion mutant (increased yellow pigment, loss of gliding motility, increased "stickiness" due to high extracellular polysaccharide production [3]) were observed in any of the *rpf* (*rpfG, rpfF, rpfB*) mutants, as would have been expected were there strict coupling between Clp and the *rpf* signaling system.

Notably absent from our experiments is a *Lysobacter enzymogenes* mutant of the *rpfC* homologue, which encodes the putative transmembrane receptor which transduces the signal from extracellular [DSF] to the intracellular response regulator RpfG. The *rpfC* mutant is currently under construction in the Du lab. If its phenotype coincides with that of the other *rpf* homologue mutants in *L. enzymogenes*, it will further confirm our theory that Clp is decoupled from the *rpf* system in *L. enzymogenes*, and may indeed act constitutively, independent of any small-molecule effector. It should be noted, however, that RpfC negatively regulates DSF production in *Xcc* and other DSF-producing xanthomonads,[9, 37] so if a *Lysobacter* DSF-like signaling molecule is ever to be identified, it would be most likely be detected in cultures of the *rpfC*-homologue mutant of *L. enzymogenes*.

A crude extraction of the metabolites present in the Clp-complemented *Le*DC mutants was carried out with the purpose of screening for any new metabolites and for any metabolites which displayed altered production levels. The *Le*DC mutant complemented with pHmgA-P-Clp2 (the wild-type Clp sequence) served as the negative

control, against which the two *Le*DC mutants complemented with either pHmgA-P-ClpE69A or pHmgA-P-ClpR153A were compared for any notable differences. The Clp-complemented *Le*DC cultures were grown under conditions expected to result in dihydromaltophilin production: in NYGB for 60 hours at 30 °C, with shaking at 200 rpm. In an effort to extract as many different metabolites as possible, we extracted the crude cultures with equal volumes of butanol at both high and low pH. The natural pH of mature *Lysobacter* cultures is between 8 and 9, so the first extraction was carried out directly on the crude cultures; 37% HCl was then used to lower the pH to 4, and the butanol extraction repeated. The butanol layers were allowed to separate for ~1 hour, then collected using a separatory funnel and a clean round-bottomed flask. Equal volumes of double-distilled water were added to the butanol and the butanol:water mixture removed by rotary evaporation at 55 °C. Remaining liquid was evaporated using a gentle air stream, and the remaining residue was redissolved as much as possible in 1.5 mL methanol. After centrifugation to pellet any suspended particles, the supernatant was investigated by HPLC. Although we do not regard this experiment as definitive, the absence of any notable difference between the metabolite profiles of the wild-type Clp- and the constitutive Clp-complementation *Le*DC mutants lends support to our theory that either Clp is already constitutive, or else the intracellular concentration of [c-di-GMP] is kept low by some system other than the protein products of the *rpf* homologues, during the conditions in which *L. enzymogenes*'s secondary metabolism is active.

3.2: Summary of results: a qualified conclusion

Taken together, our data suggest that the relationship between the *clp* homologue in *Lysobacter enzymogenes* and the upregulation of the dihydromaltophilin gene cluster is not subject to the upstream control of a signaling system similar to the DSF/*rpf* system in *Xanthomonas campestris* pv. *campestris* and other xanthomonads. This conclusion is directly at odds with our original hypothesis, but none of our data hitherto have supplied unambiguous support for a direct link between a diffusible signal factor and the production of dihydromaltophilin in *L. enzymogenes* C3 or OH11. It remains to form a hypothesis to explain this anomaly: why *L. enzymogenes*, a xanthomonad with highly homologous genes to all of the necessary components of DSF quorum sensing in *Xcc*, should apparently exhibit no connection between the two main portions (*clp* and the four genes of the *rpf* cluster) of that signaling system. While research on this question is still ongoing, and we do not in any sense exclude the possibility that other experimental conditions or improved techniques might yet establish the original hypothesis, we may nonetheless formulate alternative hypotheses to account for our observations hitherto.

The different ecological niches occupied by the plant pathogen *Xcc* and the bacterial predator *L. enzymogenes* may hold the clue to this enigma. While *Xcc* and other phytopathogenic xanthomonads appear to depend heavily upon their DSF-signaling mechanisms to coordinate the different phases of transmission to and infection of their hosts,[77-78] *L. enzymogenes* leads a very different lifestyle and may not require this sort of coordinated behavior; or, if it does, it is possible that we have not identified the phases or conditions under which its DSF-signaling mechanisms come into play.

We cannot exclude that some sort of structural deficiency in *L. enzymogenes* (*i.e.*, mutations in one or more of the genes in the *rpf* system) which alter or abolish the activity of the protein products of the *rpf* homologues. Our analysis of the nucleotide and amino acid sequences failed to uncover any obviously major differences between the two sets of homologues, but we do remark two differences of note between *Xcc*Clp and *Le*Clp. Firstly, as discussed previously, the key aspartate (D70) which in *Xcc*Clp interacts with c-di-GMP inhibitor is replaced with a glutamate (E69) in *Le*Clp. Although the activity of the side-chains is identical, it is possible that the greater size of the glutamate exerts enough of a steric effect to destabilize the binding of the c-di-GMP, making *Le*Clp partially or completely insensitive to intracellular [c-di-GMP]. Secondly, one key difference between *Xcc*Clp and *E. coli*'s cAMP-receptor protein is *Xcc*Clp's substitution of a threonine residue (Thr-149) for the serine (Ser-129) found in *E. coli*'s CRP.[34] In *Le*Clp, Ser-148 is the residue corresponding with *Xcc*Clp's Thr-149 and with *E. coli* CRP's Ser-129 (**Figure 5**). This alteration, at a position that appears to be involved in binding of either c-di-GMP or cAMP to these small-molecule receptor proteins, was regarded as sufficiently important to be tested by mutagenesis studies by Tao *et al.*, but mutation of Thr-149 back to serine did not result in any decrease of inhibition by c-di-GMP of the T149S Clp *in vitro*.[34] Based on this result, we suspect that *Le*Clp's Ser-148 does not render it insensitive to inhibition by c-di-GMP. Other key residues predicted by *Xcc*Clp's crystal structure and molecular modeling to be involved in c-di-GMP binding are strictly conserved in *Le*Clp, as are residues involved in stabilizing Clp's DNA-binding conformation. It remains possible that *Le*Clp contains one or more variant residues, important either to DNA-binding or c-di-GMP inhibition, which have hitherto

escaped attention. Since no crystal structure of *Xcc*Clp bound to c-di-GMP is yet available, we must postpone judgment on the likelihood of this possibility.

Perhaps the greatest deficiency in our investigation to this point has been the absence of any sort of characterization of the *rpfC* homologue in *L. enzymogenes*. Since the transmembrane sensor RpfC negatively regulates DSF biosynthesis in *Xcc*,[8] it would in retrospect have been useful to begin our search for a *Lysobacter* DSF with a mutant containing an inactivated *rpfC* gene. Final judgment on the existence and activity of a DSF-like signal molecule in *L. enzymogenes* probably awaits characterization of such an *rpfC* mutant.

While the energetic expense of secondary metabolism suggests that *L. enzymogenes* should possess some sort of regulatory mechanism for its activation or deactivation, this bacterium's predatory nature may demand a more or less continuous output of the secondary metabolites and lytic enzymes which have been shown to be under the control of Clp.[3] This possibility is supported by our experiments, in which we encountered difficulty in discovering conditions in which *L. enzymogenes* did *not* produce dihydromaltophilin. A number of different media containing different carbon sources were tested, all of which supported at least some level of dihydromaltophilin production and in which, therefore, we may safely assume Clp was active and uninhibited. While our research focused on deciphering the system by which *L. enzymogenes*'s secondary metabolism was regulated, it is an ironic possibility that this predator may have dispensed (partially or completely) with such regulation to accommodate its aggressive lifestyle.

Chapter 4: Future prospects

Confirmation of *Le*Clp's constitutive nature could be easily evaluated via electrophoresis mobility shift assays (EMSA), as has been carried out on *Xcc*Clp and its mutants.[1, 34] Native *Le*Clp has already been expressed as a His_6-tagged protein in *E. coli* BL21 and purified in our lab (Haotong Chen, unpublished data), following the procedure employed by Chin *et al*. prior to their crystallization of *Xcc*Clp.[1] Analysis of this purified protein, along with its (as yet) not purified E69A and R153A "constitutive" mutants, by EMSA would reveal whether or not *Le*Clp is sensitive to inhibition by c-di-GMP *in vitro*. We have proposed several DNA sequences as targets for *Le*Clp binding, including several sites in the 538-bp "HSAF promoter" region which was installed in the vector pHmgA-P and which is predicted by *in silico* analysis [79-80] to contain the promoter for dihydromaltophilin's biosynthetic gene cluster. As a likely positive control, the promoter used so frequently with *Xcc*Clp, that of the *engA* gene in *Xcc*,[29, 36, 59] could be assayed for binding to *Le*Clp *in vitro*. Indeed, the binding of this promoter to *Le*Clp could be directly assayed *in vivo* by expressing a modified version of the Clp-linked DSF reporter pKLN55 constructed by Newman *et al*. in *L. enzymogenes* C3 or OH11, although an appropriate antibiotic selection marker might need to be cloned into pKLN55 to render it suitable for expression in *L. enzymogenes*.[36] EMSA experiments could directly verify that *Le*Clp is c-di-GMP-insensitive and thus, most likely, constitutive in *L. enzymogenes*.

If, as expected based on its similarity to *Xcc*Clp, wild-type *Le*Clp proves to be inhibited by c-di-GMP *in vitro* in a manner similar to *Xcc*Clp, further experiments *in vivo* might reveal to what extent *Le*Clp is inhibited under the conditions in which *L. enzymogenes* has been found to produce dihydromaltophilin. A fascinating experiment

carried out by Tao *et al*. which might be replicated in *L. enzymogenes* was the heterologous expression in *Xcc* of the protein products of genes *PA5487* and *PA3947* from *Pseudomonas aeruginosa* PA14.[34] While *PA5487* encodes a GGEEF-domain-containing protein which was shown to have diguanylate cyclase activity, *PA3947* encodes an EVL-domain-containing protein which serves as a c-di-GMP phosphodiesterase,[81] and *in trans* expression of each enzyme in *L. enzymogenes* would be expected to yield mutants with high or low c-di-GMP concentrations, respectively. A constitutive *Le*Clp would show no response to the elevated c-di-GMP concentrations in the *Le*-PA5487 mutant, and transcription of the dihydromaltophilin gene cluster, dihydromaltophilin production, and antifungal activity would be expected to remain unaltered. A c-di-GMP-sensitive *Le*Clp would likely result in a *Le*-PA5487 mutant which exhibits the same phenotype as the clp-deletion mutant *Le*DC211.[3] Additionally, these two *P. aeruginosa* genes could be co-expressed with a Clp reporter construct similar to pKLN55,[36] which places GFP-expression under the control of the Clp-cognate *engA* promoter, and levels of GFP expression monitored to evaluate the extent of Clp's activity in the presence or absence of c-di-GMP *in vivo.* Finally, the effects of the *in trans* expression of these two genes from *P. aeruginosa* PA14 might also be evaluated in *L. enzymogenes* mutants expressing the constitutive Clp E69A and Clp R153A alleles, revealing whether or not these two mutant versions of Clp are subject to inhibition by c-di-GMP. Direct measurement of the activity of Clp and its mutants *in vivo* could also by achieved by measuring the transcription level of selected portions of the dihydromaltophilin gene cluster, using reverse-transcription PCR. Additionally, a

remarkable method of tracing the intracellular production of c-di-GMP using fluorescent RNA-based biosensors has recently been reported.[82]

One final experiment in the same vein would be the expression of *Xcc*Clp in the clp-deletion mutant *Le*DC211. If some intrinsic feature of *Le*Clp prevents its being subject to c-di-GMP inhibition, expression of *Xcc*Clp might "re-couple" Clp's activity to *L. enzymogenes*'s homologues of the *rpf* system, placing dihydromaltophilin production and its accompanying phenotype under the control of a diffusible signal factor.

There remains the intriguing possibility that the dissimilarity between *L. enzymogenes* and *Xcc*, despite the striking similarity between their *clp* and *rpfC/G/F/B* homologues, can be traced in whole or in part to *Le*Clp's accompanying *N*-acetyltransferase. Direct modification of transcription factors by acetylation has been reported in *E. coli*,[43] and biochemical or genetic experiments directed at determining this acetyltransferase's molecular target might elucidate much that still remains unclear regarding this pathway.

In *Stenotrophomonas maltophilia* WR-C, DSF-like extracellular fatty acids produced by homologues of *rpfF* and *rpfB* were found to be involved in "flagella-independent surface translocation,"[49] and it was previously reported by Koboyashi *et al.* and confirmed in our lab that the *clp*-deletion mutant *Le*DC211 lacked the gliding motility observed in wild-type *Le*C3.[3] It is, therefore, not unprecedented to supposed that the *rpf* homologues in *L. enzymogenes* serve a purpose quite different from intercellular signaling, and the connection between the *clp* and *rpf* homologues in *L. enzymogenes* is much more indirect than in *Xcc*. Although we have not observed any significant morphological differences between the wild-type *Le*OH11 and its *rpf*-

inactivated mutants, experiments designed specifically to evaluate motility might provide clues as to the primary function of the *rpf* homologues in *L. enzymogenes*.

Chapter 5: Appendix

In addition to investigating the relationship between *L. enzymogenes*'s *clp* and *rpf* homologues to dihydromaltophilin production, we carried out investigations on the biosynthetic mechanism of dihydromaltophilin.

5.1 Targeted mutagenesis of dihydromaltophilin polyketide synthase

The biosynthetic gene cluster for dihydromaltophilin in *L. enzymogenes* C3 was identified in 2007, and found to encode a large hybrid nonribosomal peptide synthetase-polyketide synthase (NRPS-PKS) along with four putative oxidoreductases and a putative sterol desaturase,[15] an architecture which was found to be common to other producers of polycyclic tetramic macrolactam natural products.[21] Subsequent gene inactivation and heterologous expression showed that the sterol desaturase was responsible for installing the β-hydroxyl group on the ornithine moiety,[83] and that the NRPS module alone was responsible for the condensation of two amide bonds and a Dieckmann cyclization which yielded the tetramic acid moiety joining two polyketides and the ornithine residue.[39] In contrast to the type I modular PKS usually found in bacteria, in which multiple PKS modules are present and each is responsible for a single round of elongation and processing of the nascent polyketide,[84] the dihydromaltophilin PKS included only a single module, containing predicted β-ketoreductase (KR) and dehydratase (DH) tailoring domains in addition to the essential β-ketosynthase (KS), acyltransferase (AT), and acyl carrier protein (ACP) domains.[15] It appeared that the dihydromaltophilin PKS was an iterative type I, uncommon in bacteria but frequently seen in fungi,[84] but although the three essential domains apparently acted iteratively, it was not clear whether the two tailoring domains (KR and DH) acted only once on the

nascent polyketide, or whether the polyketide was reduced to a polyenoic acid by multiple tailoring steps. As the initially proposed biosynthetic pathway for dihydromaltophilin called for an unknown PKS to supply the second polyketide, which clearly differed from the first,[15] it seemed possible that the dihydromaltophilin PKS produced a polyketide with only one double bond, which then participated in the proposed mechanism shown in **Figure 37**. If this were the case, inactivation of the KR or DH domains of the PKS module might result in analogues of dihydromaltophilin in which an α,β-enoate moiety was replaced with a β-ketone or β-hydroxyl group, respectively(**Figure 38**).

Figure 37: initially proposed mechanism of dihydromaltophilin biosynthesis [15]

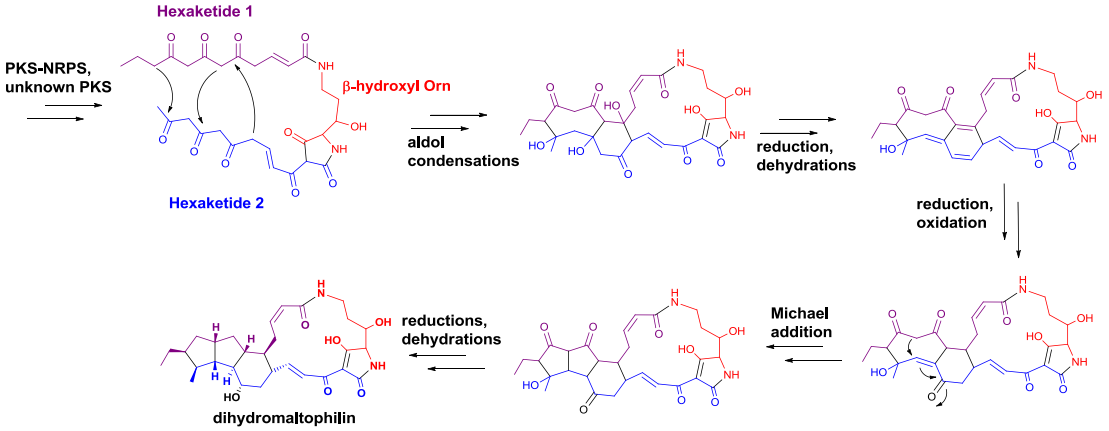
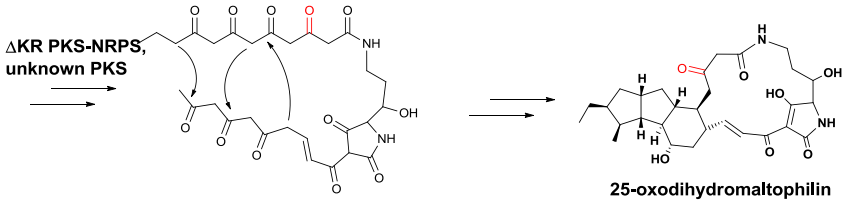
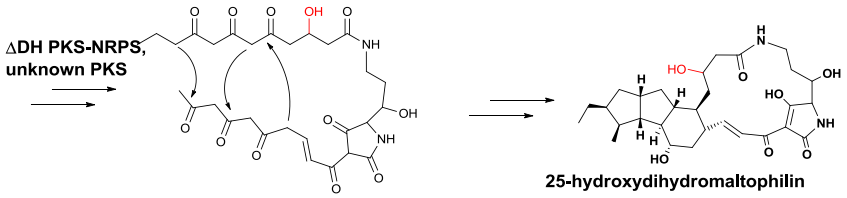


Figure 38: possible dihydromaltophilin analogues

A) proposed analogue resulting from inactivated KR domain



B) proposed analogue resulting from inactivated DH domain

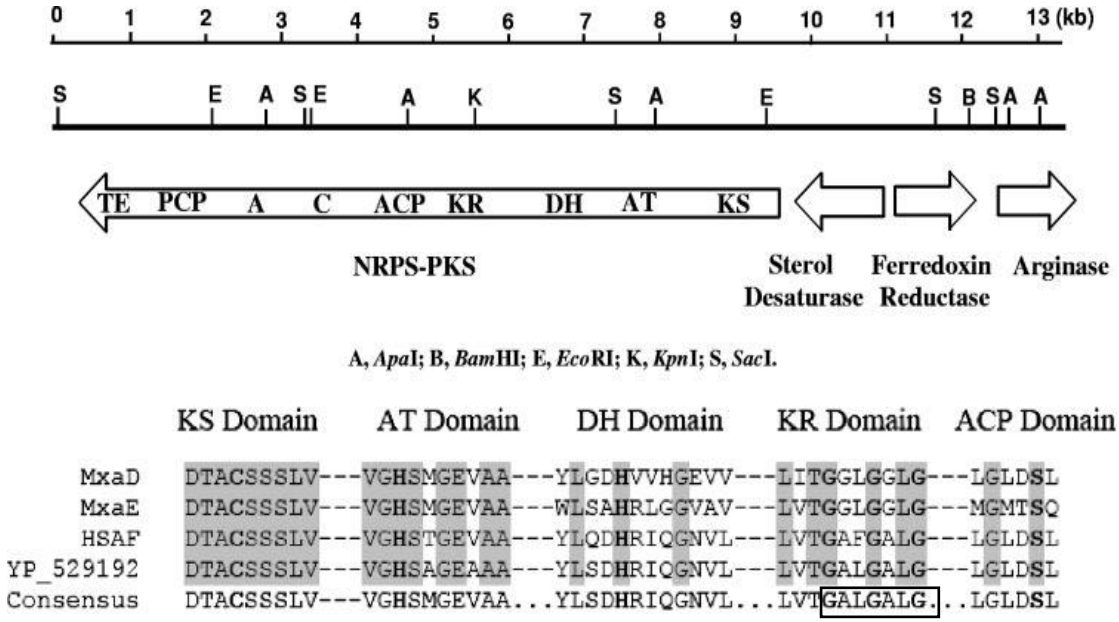


In other modular bacterial polyketide synthases, targeted inactivation of the processing domains had resulted in the production of natural product analogues containing unprocessed keto or hydroxyl groups,[85-86] but such biosynthetic manipulations are not expected to succeed in the case of iterative PKSs. However, the possibility that the dihydromaltophilin PKS's tailoring domains, like those of modular PKSs, acted only once on the nascent polyketide, suggested that genetic manipulations of the NRPS-PKS gene might yield mutants of *L. enzymogenes* C3 which produced the proposed analogues in **Figure 38**.

5.1.1 Construction of *L. enzymogenes* C3 ΔKR mutant

Yu *et al*. had reported the consensus conserved residues of all of the major domains of the hybrid PKS-NRPS, including the PKS's KR and DH tailoring domains, and we proposed a targeted deletion of seven residues of the conserved, glycine-rich NADPH-binding site of the KR domain (**Figure 39**).

Figure 39: Partial map of 13.5-kb region of LeC3 hosting dihydromaltophilin biosynthetic genes (Figure 5 from [15]). Reprinted by permission from the American Society for Microbiology.



To generate the DNA fragment containing the 21-bp deletion, we employed the mutagenesis method "splicing by overlap extension." Briefly, PCR primers were designed upstream and downstream of naturally-occurring *Kpn*I and *Apa*I restriction sites, respectively. Partially-overlapping forward and reverse primers spanning the site targeted for deletion. Two separate PCRs generated the upstream (615-bp) and downstream (455-bp) fragments, both containing the sequence with the 21-bp deletion. These fragments were purified and combined for a slicing PCR, which fused the two fragments together into a single mutagenic fragment (978-bp). This fragment was purified and digested with *Kpn*I and *Apa*I restriction enzymes (New England Biolabs), and ligated into the *Kpn*I/*Apa*I double-digested cloning vector pGem5zf(+) (Promega) and transformed into chemically competent *E. coli* XL-1 Blue. Single colonies which grew on LB agar supplemented with 50 μg/mL ampicillin were picked up and miniprepped, and the insertion of the mutagenic DNA fragment confirmed by digestion

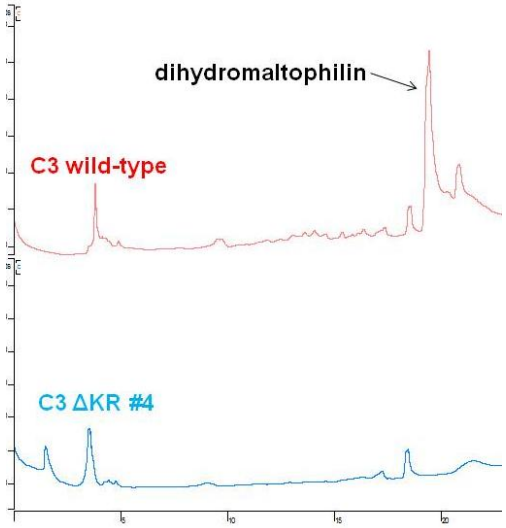
with *Kpn*I and *Apa*I. After sequencing confirmed its fidelity, the DNA insert was subcloned into the suicide vector pJQ200SK,[13] and the resulting mutagenic plasmid pJQ200SK-KR transformed into *E. coli* S17-1 and transferred in *L. enzymogenes* C3 by conjugation. Single colonies which appeared on 10% TSB agar plates supplemented with 20 µg/mL gentamicin and 25 µg/mL kanamycin represented single-crossover recombinants of *Le*C3, which were confirmed by PCR. The single-crossover mutants were grown in liquid 10% TSB supplemented with 25 µg/mL kanamycin and 5% sucrose, and re-streaked onto agar plates supplemented with 5% sucrose to encourage double-crossover heterologous recombination. Double-crossover recombinants were confirmed by loss of gentamicin resistance and by sequencing of the KR region to verify the installation of the 21-bp deletion.

5.1.2 Analysis of *Le*C3 ΔKR mutant

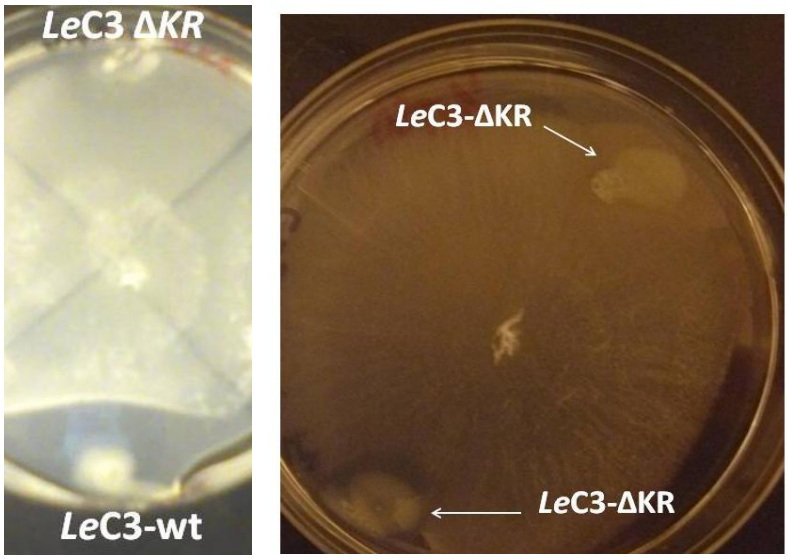
Once confirmed, the in-frame deletion *Le*C3 ΔKR mutant was characterized by assaying its antifungal activity and its level of dihydromaltophilin production. Antifungal assays were carried

Figure 40: Comparisons of *Le*C3 wild-type to *Le*C3 ΔKR mutant

A) Extractions of dihydromaltophilin from the mutant and from wild-type.



B) Antagonism toward *Fusarium verticilliodes*. The ΔKR mutant lacks wild-type's inhibition zone, but appears to retain "close-range" antifungal activity, perhaps the result of secreted lytic enzymes.



out using *Fusarium verticilliodes* as described in Chapter 4, and extraction of dihydromaltophilin by ammonium sulfate precipitation followed the protocol described by Yu *et al*.[15] Cultures of the *L. enzymogenes* strains were grown in liquid 10% TSB supplemented with 5 mM lactose for 4 days at 28 °C with shaking, and the supernatant

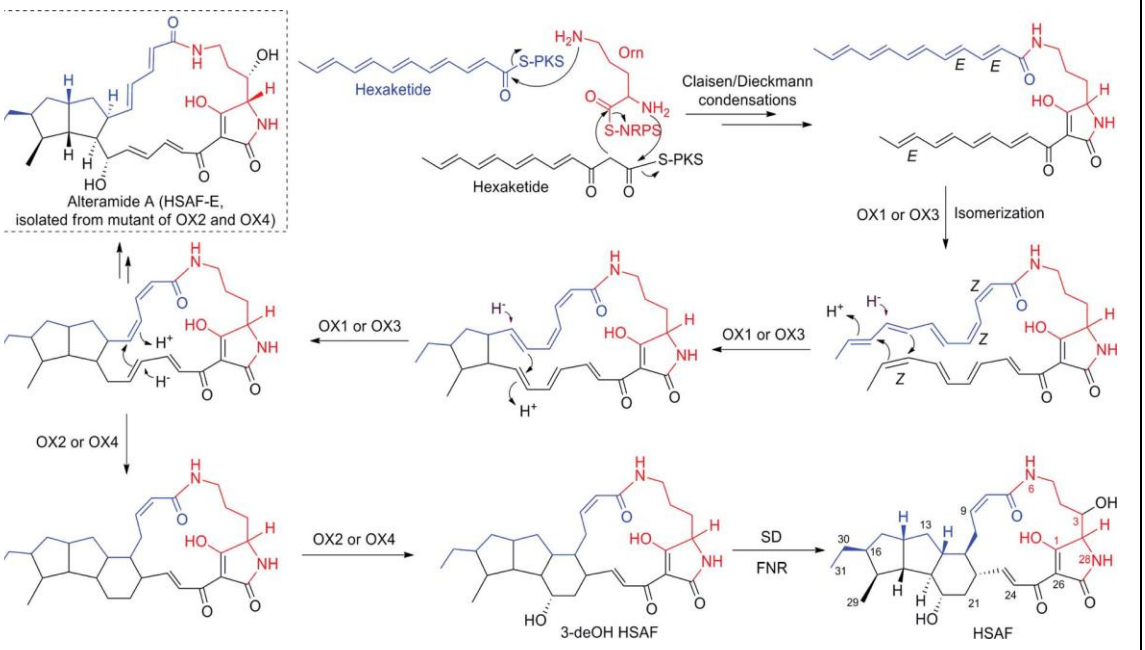
collected by centrifugation (10,000 × *g*, 30 min.) at 4 °C. Ammonium sulfate was added to the supernatants to a final concentration of 0.5 g/mL and vortexed until complete dissolution, and the solution was incubated overnight at 4 °C. The precipitate was collected by centrifugation (10,000 × *g*, 25 min.) at 4 °C and resuspended in 200 µL 100% methanol, and after a short (~30 min.) incubation at room temperature the suspension was centrifuged (10,000 × *g*, 10 min.) at room temperature to separate the precipitate and the methanol fraction. The methanol fraction was collected and concentrated to ~100 µL by evaporation. Results are summarized in **Figure 40**, which showed that the *Le*C3 ΔKR mutant had lost all dihydromaltophilin production, and no new peaks were observed that seemed to correspond with a dihydromaltophilin analogue.

5.1.3 Conclusion and utility of *Le*C3 ΔKR mutant

The inability of the *L. enzymogenes* C3 ΔKR to produce any analogues of dihydromaltophilin, and its complete loss of antifungal activity, supported our suspicions that the dihydromaltophilin PKS was an iterative single-module PKS which probably generates a polyenoic thioester intermediate. The absence of any other PKS anywhere near the dihydromaltophilin biosynthetic gene cluster in *Le*OH11,[39] and a similar absence reported for the frontalamide cluster by Blodgett *et al*,[21] strongly suggests that the single-module PKS possesses the remarkable ability to synthesize both polyketides. If this is the case, the differences between the two polyketide moieties are most likely attributable to the timing of tailoring domains' activities and possibly to the hybrid PKS-NRPS's sensitivity to the increasing chain length of the nascent macrolactam. Although complete elucidation of the biosynthetic mechanism awaits successful biochemical characterization of the polyketide synthase and the accompanying oxidoreductases, an

alternative mechanism in which two polyenoic intermediates appear has already been proposed (**Figure 41**).[16]

Figure 41: latest proposed mechanism of dihydromaltophilin biosynthesis. Copied by permission from [16].



The *Le*C3 ΔKR mutant has proven valuable as a negative control for dihydromaltophilin production. Possessing all of the wild-type's regulatory mechanisms, antibiotic resistance and sensitivities, and other phenotypes, the ΔKR mutant remains our most reliable "wild-type-like" non-producer of dihydromaltophilin.

References

1. Chin, K.H., et al., *The cAMP Receptor-Like Protein CLP Is a Novel c-di-GMP Receptor Linking Cell-Cell Signaling to Virulence Gene Expression in Xanthomonas campestris.* Journal of Molecular Biology, 2010. **396**(3): p. 646-662.
2. Larkin M.A., B.G., Brown N.P., Chenna R., McGettigan P.A., McWilliam H., Valentin F., Wallace I.M., Wilm A., Lopez R., Thompson J.D., Gibson T.J. and Higgins D.G, *ClustalW and ClustalX version 2.* Bioinformatics, 2007. **23**(21): p. 2947-2948.
3. Kobayashi, D.Y., et al., *A clp gene homologue belonging to the Crp gene family globally regulates lytic enzyme production, antimicrobial activity, and biological control activity expressed by Lysobacter enzymogenes strain C3.* Applied and Environmental Microbiology, 2005. **71**(1): p. 261-269.
4. Zhu, H., Hu,B. and Liu,F., *Cloning and functional characterization of clp gene in Lysobacter enzymogenes strain OH11*. 2006, College of Plant Protection of Nanjing Agricultural University: Nanjing, China.
5. da Silva, A.C.R., et al., *Comparison of the genomes of two Xanthomonas pathogens with differing host specificities.* Nature, 2002. **417**(6887): p. 459-463.
6. *catabolite gene activator [Escherichia coli F11]*. 2010 [cited 2013 14 February]; deposited sequence of E. coli CRP protein].
7. He, Y.W., et al., *Xanthomonas campestris Diffusible Factor Is 3-Hydroxybenzoic Acid and Is Associated with Xanthomonadin Biosynthesis, Cell Viability, Antioxidant Activity, and Systemic Invasion.* Molecular Plant-Microbe Interactions, 2011. **24**(8): p. 948-957.
8. Cheng, Z.H., et al., *Structural Basis of the Sensor-Synthase Interaction in Autoinduction of the Quorum Sensing Signal DSF Biosynthesis.* Structure, 2010. **18**(9): p. 1199-1209.
9. He, Y.W., et al., *Dual signaling functions of the hybrid sensor kinase RpfC of Xanthomonas campestris involve either phosphorelay or receiver domain-protein interaction.* Journal of Biological Chemistry, 2006. **281**(44): p. 33414-33421.
10. Qian, W., et al., *Comparative and functional genomic analyses of the pathogenicity of phytopathogen Xanthomonas campestris pv. campestris.* Genome Research, 2005. **15**(6): p. 757-767.
11. Vorholter, F.J., et al., *The genome of Xanthomonas campestris pv. campestris B100 and its use for the reconstruction of metabolic pathways involved in xanthan biosynthesis.* Journal of Biotechnology, 2008. **134**(1-2): p. 33-45.
12. Huffman, J., in *Chemistry*. 2012, University of Nebraska: Lincoln.
13. Quandt, J. and M.F. Hynes, *Versatile Suicide Vectors Which Allow Direct Selection for Gene Replacement in Gram-Negative Bacteria.* Gene, 1993. **127**(1): p. 15-21.
14. Deng, Y., et al., *Listening to a new language: DSF-based quorum sensing in Gram-negative bacteria.* Chem Rev, 2011. **111**(1): p. 160-73.
15. Yu, F.G., et al., *Structure and biosynthesis of heat-stable antifungal factor (HSAF), a broad-spectrum antimycotic with a novel mode of action.* Antimicrobial Agents and Chemotherapy, 2007. **51**(1): p. 64-72.
16. Xie, Y.X., et al., *Bioactive natural products from Lysobacter.* Natural Product Reports, 2012. **29**(11): p. 1277-1287.
17. Reichenbach, H., *The Genus Lysobacter*, in *The Prokaryotes*, M. Dworkin, Editor. 2006, Springer: New York. p. 939-957.

18. Christensen, P. and F.D. Cook, *Lysobacter, a New Genus of Non-Fruiting, Gliding Bacteria with a High Base Ratio.* International Journal of Systematic Bacteriology, 1978. **28**(3): p. 367-393.
19. Emde, K.M.E., N. Best, and S.E. Hrudey, *Production of the Potent Odor Agent, Isopropyl Methoxypyrazine, by Lysobacter-Enzymogenes.* Environmental Technology, 1992. **13**(3): p. 201-206.
20. Nett, M. and G.M. Konig, *The chemistry of gliding bacteria.* Natural Product Reports, 2007. **24**(6): p. 1245-1261.
21. Blodgett, J.A.V., et al., *Common biosynthetic origins for polycyclic tetramate macrolactams from phylogenetically diverse bacteria.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(26): p. 11692-11697.
22. Decrecylagard, V., et al., *A Xanthomonas-Campestris Pv Campestris Protein Similar to Catabolite Activation Factor Is Involved in Regulation of Phytopathogenicity.* Journal of Bacteriology, 1990. **172**(10): p. 5877-5883.
23. Ryan, R.P. and J.M. Dow, *Diffusible signals and interspecies communication in bacteria.* Microbiology, 2008. **154**(Pt 7): p. 1845-58.
24. Ryan, R.P. and J.M. Dow, *Communication with a growing family: diffusible signal factor (DSF) signaling in bacteria.* Trends Microbiol, 2011. **19**(3): p. 145-52.
25. Dufour, N. and R.P. Rao, *Secondary metabolites and other small molecules as intercellular pathogenic signals.* Fems Microbiology Letters, 2011. **314**(1): p. 10-17.
26. Srivastava, D. and C.M. Waters, *A Tangled Web: Regulatory Connections between Quorum Sensing and Cyclic Di-GMP.* Journal of Bacteriology, 2012. **194**(17): p. 4485-4493.
27. Barber, C.E., et al., *A novel regulatory system required for pathogenicity of Xanthomonas campestris is mediated by a small diffusible signal molecule.* Molecular Microbiology, 1997. **24**(3): p. 555-566.
28. Wang, L.H., et al., *A bacterial cell-cell communication signal with cross-kingdom structural analogues.* Molecular Microbiology, 2004. **51**(3): p. 903-912.
29. Slater, H., et al., *A two-component system involving an HD-GYP domain protein links cell-cell signalling to pathogenicity gene expression in Xanthomonas campestris.* Molecular Microbiology, 2000. **38**(5): p. 986-1003.
30. He, Y.W., et al., *Genome scale analysis of diffusible signal factor regulon in Xanthomonas campestris pv. campestris: identification of novel cell-cell communication-dependent genes and functions.* Molecular Microbiology, 2006. **59**(2): p. 610-622.
31. He, Y.W., et al., *Xanthomonas campestris cell-cell communication involves a putative nucleotide receptor protein Clp and a hierarchical signalling network.* Molecular Microbiology, 2007. **64**(2): p. 281-292.
32. Ryan, R.P., et al., *Cell-cell signaling in Xanthomonas campestris involves an HD-GYP domain protein that functions in cyclic di-GMP turnover.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(17): p. 6712-6717.
33. Chou, S.H., *Delicate conformational changes of a protein in the CRP family lead to dramatic functional changes via binding of an alternate secondary messenger molecule.* Virulence, 2011. **2**(2): p. 152-157.
34. Tao, F., et al., *The Cyclic Nucleotide Monophosphate Domain of Xanthomonas campestris Global Regulator Clp Defines a New Class of Cyclic Di-GMP Effectors.* Journal of Bacteriology, 2010. **192**(4): p. 1020-1029.
35. Ryan, R.P. and J.M. Dow, *Communication with a growing family: diffusible signal factor (DSF) signaling in bacteria.* Trends in Microbiology, 2011. **19**(3): p. 145-152.

36. Newman, K.L., et al., *Cell-cell signaling controls Xylella fastidiosa interactions with both insects and plants.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(6): p. 1737-1742.
37. He, Y.W., et al., *Rice bacterial blight pathogen Xanthomonas oryzae pv. oryzae produces multiple DSF-family signals in regulation of virulence factor production.* Bmc Microbiology, 2010. **10**.
38. Altschul, S.F., et al., *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.* Nucleic Acids Research, 1997. **25**(17): p. 3389-3402.
39. Lou, L.L., et al., *Biosynthesis of HSAF, a Tetramic Acid-Containing Macrolactam from Lysobacter enzymogenes.* Journal of the American Chemical Society, 2011. **133**(4): p. 643-645.
40. Guoling Qian, Y.W., Yiru Liu, Feifei Xu, Yawen He, Liangcheng Du, Jiaqin Fan, Baishi Hu, Fengquan Liu, *Lysobacter enzymogenes uses two distinct cell-cell signaling for differential regulation of metabolite biosynthesis.* 2013.
41. Ryan, R.P., et al., *Cell-cell signal-dependent dynamic interactions between HD-GYP and GGDEF domain proteins mediate virulence in Xanthomonas campestris.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(13): p. 5989-5994.
42. Imhof, A., et al., *Acetylation of general transcription factors by histone acetyltransferases.* Current Biology, 1997. **7**(9): p. 689-692.
43. Thao, S., et al., *N-epsilon-Lysine Acetylation of a Bacterial Transcription Factor Inhibits Its DNA-Binding Activity.* Plos One, 2010. **5**(12).
44. Dyda, F., D.C. Klein, and A.B. Hickman, *GCN5-related N-acetyltransferases: A structural overview.* Annual Review of Biophysics and Biomolecular Structure, 2000. **29**: p. 81-103.
45. Vetting, M.W., et al., *Structure and functions of the GNAT superfamily of acetyltransferases.* Archives of Biochemistry and Biophysics, 2005. **433**(1): p. 212-226.
46. Trievel, R.C., et al., *Crystal structure and mechanism of histone acetylation of the yeast GCN5 transcriptional coactivator.* Proceedings of the National Academy of Sciences of the United States of America, 1999. **96**(16): p. 8931-8936.
47. Dutnall, R.N., et al., *Structure of the histone acetyltransferase Hat1: A paradigm for the GCN5-related N-acetyltransferase superfamily.* Cell, 1998. **94**(4): p. 427-438.
48. Simionato, A.V.C., et al., *Characterization of a putative Xylella fastidiosa diffusible signal factor by HRGC-EI-MS.* Journal of Mass Spectrometry, 2007. **42**(10): p. 1375-1381.
49. Huang, T.P. and A.C.L. Wong, *Extracellular fatty acids facilitate flagella-independent translocation by Stenotrophomonas maltophilia.* Research in Microbiology, 2007. **158**(8-9): p. 702-711.
50. Fouhy, Y., et al., *Diffusible signal factor-dependent cell-cell signaling and virulence in the nosocomial pathogen Stenotrophomonas maltophilia.* Journal of Bacteriology, 2007. **189**(13): p. 4964-4968.
51. Deng, Y.Y., et al., *Structural and Functional Characterization of Diffusible Signal Factor Family Quorum-Sensing Signals Produced by Members of the Burkholderia cepacia Complex.* Applied and Environmental Microbiology, 2010. **76**(14): p. 4675-4683.
52. *Sorvall Legend T/RT instruction manual*. 2001, Kendro Laboratory Products
53. *cis-Δ2-11-methyl-Dodecenoic Acid*. 2013 [cited 2013 January 21]; Available from: https://www.caymanchem.com/app/template/Product.vm/catalog/10008123/a/z.

54. Kerwin, J.L., A.M. Wiens, and L.H. Ericsson, *Identification of fatty acids by electrospray mass spectrometry and tandem mass spectrometry.* Journal of Mass Spectrometry, 1996. **31**(2): p. 184-192.
55. Bylund, J., J. Ericsson, and E.H. Oliw, *Analysis of cytochrome P450 metabolites of arachidonic and linoleic acids by liquid chromatography mass spectrometry with ion trap MS2.* Analytical Biochemistry, 1998. **265**(1): p. 55-68.
56. Balazy, M., et al., *Vicinal nitrohydroxyeicosatrienoic acids: Vasodilator lipids formed by reaction of nitrogen dioxide with arachidonic acid.* Journal of Pharmacology and Experimental Therapeutics, 2001. **299**(2): p. 611-619.
57. Oliw, E.H., et al., *Analysis of novel hydroperoxides and other metabolites of oleic, linoleic, and linolenic acids by liquid chromatography mass spectrometry with ion trap MSn.* Lipids, 1998. **33**(9): p. 843-852.
58. Goyal, S., S. Banerjee, and S. Mazumdar, *Oxygenation of Monoenoic Fatty Acids by CYP175A1, an Orphan Cytochrome P450 from Thermus thermophilus HB27.* Biochemistry, 2012. **51**(40): p. 7880-7890.
59. Hsiao, Y.M., et al., *Clp upregulates transcription of engA gene encoding a virulence factor in Xanthomonas campestris by direct binding to the upstream tandem Clp sites.* Febs Letters, 2005. **579**(17): p. 3525-3533.
60. Corpe, W.A., *Extracellular Accumulation of Pyrroles in Bacterial Cultures.* Applied Microbiology, 1963. **11**(2): p. 145-&.
61. Anderle, C., et al., *Biosynthesis of clorobiocin: investigation of the transfer and methylation of the pyrrolyl-2-carboxyl moiety.* Archives of Microbiology, 2007. **187**(3): p. 227-237.
62. Jones, A.C., E.A. Monroe, and W.H. Gerwick, *Elegant Metabolite Biosynthesis.* Chemistry & Biology, 2011. **18**(3): p. 281-283.
63. Hofer, I., et al., *Insights into the Biosynthesis of Hormaomycin, An Exceptionally Complex Bacterial Signaling Metabolite.* Chemistry & Biology, 2011. **18**(3): p. 381-391.
64. Garneau, S., et al., *Characterization of the formation of the pyrrole moiety during clorobiocin and coumermycin A(1) biosynthesis.* Biochemistry, 2005. **44**(8): p. 2770-2780.
65. Walsh, C.T., S. Garneau-Tsodikova, and A.R. Howard-Jones, *Biological formation of pyrroles: Nature's logic and enzymatic machinery.* Natural Product Reports, 2006. **23**(4): p. 517-531.
66. Pojer, F., S.M. Li, and L. Heide, *Molecular cloning and sequence analysis of the clorobiocin biosynthetic gene cluster: new insights into the biosynthesis of aminocoumarin antibiotics.* Microbiology-Sgm, 2002. **148**: p. 3901-3911.
67. Du, Y.W.G.Q.Y.L.Y.W.Y.W.S.W.Y.L.Y.S.F.L.L., *Biosynthetic Mechanism for Sunscreens of the Biocontrol Agent Lysobacter enzymogenes*, in *PLOS Genetics*. 2013.
68. Starr, M.P., et al., *Chemotaxonomic Significance of Xanthomonadins, Novel Brominated Aryl-Polyene Pigments Produced by Bacteria of Genus Xanthomonas.* Archives of Microbiology, 1977. **113**(1-2): p. 1-9.
69. Andrewes, A.G., et al., *Xanthomonas Pigments .2. Xanthomonas Carotenoids - Non-Carotenoid Brominate Aryl-Polyene Esters.* Acta Chemica Scandinavica, 1973. **27**(7): p. 2383-2395.
70. Poplawsky, A.R. and W. Chun, *pigB determines a diffusible factor needed for extracellular polysaccharide slime and xanthomonadin production in Xanthomonas campestris pv campestris.* Journal of Bacteriology, 1997. **179**(2): p. 439-444.

71. Poplawsky, A.R., et al., *A gene for a dioxygenase-like protein determines the production of the DF signal in Xanthomonas campestris pv. campestris.* Molecular Plant Pathology, 2005. **6**(6): p. 653-657.
72. Romling, U. and D. Amikam, *Cyclic di-GMP as a second messenger.* Current Opinion in Microbiology, 2006. **9**(2): p. 218-228.
73. Jenal, U. and J. Malone, *Mechanisms of cyclic-di-GMP signaling in bacteria.* Annual Review of Genetics, 2006. **40**: p. 385-407.
74. Leduc, J.L. and G.P. Roberts, *Cyclic di-GMP Allosterically Inhibits the CRP-Like Protein (Clp) of Xanthomonas axonopodis pv. citri.* Journal of Bacteriology, 2009. **191**(22): p. 7121-7122.
75. Hoang, T.T., et al., *A broad-host-range Flp-FRT recombination system for site-specific excision of chromosomally-located DNA sequences: application for isolation of unmarked Pseudomonas aeruginosa mutants.* Gene, 1998. **212**(1): p. 77-86.
76. Arias-Barrau, E., et al., *The homogentisate pathway: A central catabolic pathway involved in the degradation of L-phenylalanine, L-tyrosine, and 3-hydroxyphenylacetate in Pseudomonas putida.* Journal of Bacteriology, 2004. **186**(15): p. 5062-5077.
77. He, Y.W. and L.H. Zhang, *Quorum sensing and virulence regulation in Xanthomonas campestris.* Fems Microbiology Reviews, 2008. **32**(5): p. 842-857.
78. Ryan, R.P., et al., *Pathogenomics of Xanthomonas: understanding bacterium-plant interactions.* Nat Rev Microbiol, 2011. **9**(5): p. 344-55.
79. Reese, M.G., *Application of a time-delay neural network to promoter annotation in the Drosophila melanogaster genome.* Computers & Chemistry, 2001. **26**(1): p. 51-56.
80. Reese, M.G. 1999, Berkeley Drosophila Genome Project
p. prokaryote promoter predictor.
81. Kulesekara, H., et al., *Analysis of Pseudomonas aeruginosa diguanylate cyclases and phosphodiesterases reveals a role for bis-(3 '-5 ')-cyclic-GMP in virulence.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(8): p. 2839-2844.
82. Kellenberger, C.A., et al., *RNA-Based Fluorescent Biosensors for Live Cell Imaging of Second Messengers Cyclic di-GMP and Cyclic AMP-GMP.* Journal of the American Chemical Society, 2013.
83. Li, Y.Y., et al., *3-Hydroxylation of the polycyclic tetramate macrolactam in the biosynthesis of antifungal HSAF from Lysobacter enzymogenes C3.* Medchemcomm, 2012. **3**(8): p. 982-986.
84. Sattely, E.S., M.A. Fischbach, and C.T. Walsh, *Total biosynthesis: in vitro reconstitution of polyketide and nonribosomal peptide pathways.* Natural Product Reports, 2008. **25**(4): p. 757-793.
85. Donadio, S., et al., *Modular Organization of Genes Required for Complex Polyketide Biosynthesis.* Science, 1991. **252**(5006): p. 675-679.
86. Tang, L., et al., *Elucidating the mechanism of cis double bond formation in epothilone biosynthesis.* Journal of the American Chemical Society, 2004. **126**(1): p. 46-47.