The LuxC reductase of the marine bacterium *Photobacterium phosphoreum* may be part of the aldehyde dehydrogenase extended family of enzymes.

By

Nicolas Skouris

Department of Biochemistry

McGill University

Montreal, Canada

April 2004

A thesis submitted to the faculty of graduate studies and research and the biochemistry department for partial fulfillment of the requirements for the degree of Master of Science.

Your file Votre référence
ISBN: 0-494-06456-0
Our file Notre référence
ISBN: 0-494-06456-0

Abstract

The acyl-CoA reductase, LuxC, isolated from *Photobacterium phosphoreum*, has been found to have a low sequence identity to aldehyde dehydrogenases, particularly in the more conserved regions (i.e. motifs) containing the active site residues. Mutational studies on residues involved in cofactor binding and in catalysis conserved in both the LuxC enzyme and in *Vibrio harveyi* aldehyde dehydrogenase have shown that these residues may play similar roles as the effects of mutation on the kinetic parameters of the two enzymes were found to be very similar. Moreover, preliminary X-ray structural characterization of LuxC shows that its overall structure and fold is very similar to that of the *V. harveyi* aldehyde dehydrogenase. Both enzymes are able to bind their NAD(P)(H) cofactors in much the same way and the β-sheet extension that is seen upon dimer formation in the *V. harveyi* enzyme is also seen in the *P. phosphoreum* enzyme. Given these results, and the fact that PpLuxC and Vh-ALDH have polypeptides of very similar size that can catalyze similar reactions, albeit in opposite directions, we are proposing that the acyl-CoA reductase of *P. phosphoreum* is part of the aldehyde dehydrogenase extended family, making it the first member of this family to preferentially catalyze the reduction of fatty acids to aldehydes rather than the oxidation of fatty aldehydes to acids. Finally, through the study of a critical glutamate residue of *V. harveyi* aldehyde dehydrogenase, we have provided direct evidence that this residue may be the critical determinant in deciding whether aldehyde oxidation or acid reduction is catalyzed preferentially.

Abstrait

Il a été trouvé que la réductase des acyl-CoA, LuxC, isolée de l'organisme *Photobacterium phosphoreum,* a une faible identité de séquence avec les aldéhyde déhydrogénases, particulièrement dans les motifs hautement conservés qui entourent le site actif. L'analyse des effets de mutations sur les résidus impliqués dans le mécanisme d'action de l'enzyme et dans sa liaison avec le cofacteur, dans LuxC et dans l'aldéhyde déhydrogénase de l'organisme *Vibrio harveyi*, ont demontré qu'ils peuvent accomplir la même fonction étant donné que les effets sur les paramètres cinétiques sont très similaires. De plus, la structure préliminaire de LuxC à été résolue et démontre qu'elle ressemble beaucoup à la structure des aldéhyde déhydrogénase, en particulier celle de *V. harveyi.* Les deux enzymes en question ont la capacité de se lier au cofacteur NAD(P)(H) de la même manière et que les contacts qui aident à la dimerization de l'aldéhyde déhydrogénase de *V. harveyi* sont aussi présent dans LuxC. Étant donné ces résultats, et le fait que PpLuxC et Vh-ALDH ont approximativement la même masse et qu'ils peuvent catalyser les mêmes reactions, en sense inverse, nous proposons que la réductase des acyl-CoA de *P. phosphoreum* fait parti de la famille étendue des aldéhyde déhydrogénases et qu'elle est la première enzyme de cette catégorie à catalyser preferentiellement la réduction des acides à des aldéhydes et non l'oxidation d'aldéhydes à des acides. Finalement, à travers des études menées sur un acide glutamique critique de Vh-ALDH, nous avons démontré que cet acide aminé peut être un facteur déterminant la fonction prédominante (oxidation versus réduction) des aldéhyde déhydrogénases.

Table of Contents

ACKNOWLEDGEMENTS

I would first and foremost like to thank my supervisor, Dr. Edward A. Meighen, for allowing me to pursue my graduate career under his guidance. His encouragement of independent thought has allowed me to learn a great deal, and his insight in discussions has proven to be invaluable.

I am also very grateful for the technical assistance, advice and friendship of Leo Lin, Rose Szittner and in particular Carol Miyamoto, who has been there for matters both professional and personal.

During this project, I also had the privilege of working under the supervision of Dr. Alice Vrielink, who has taught me a great deal, and whom I would like to thank for allowing me to work in her laboratory. I would also thank the members of her lab, Dr. Paula Lario, Dr. Louis Lim, Art Lyubimov and Scott Foster, for helping me and for making the experience enjoyable and memorable.

I would like to thank my parents for their constant encouragement in continuing my education and for their unwavering support, belief and love, without which I could not have come so far. I will be forever thankful.

And to my dear friends, thank you for being there for all the other little things, and to Matthew Farrell in particular, thank you for your patience, understanding and friendship.

Original contributions to knowledge

1- The preliminary structural characterization of *Photobacterium phosphoreum* LuxC has shown that the overall structure of this enzyme is very similar to the structures obtained so far for aldehyde dehydrogenase. We can clearly see the existence of three separate domains, that the binding mode of the NAD(P)H cofactor resembles that of aldehyde dehydrogenases, and that contacts which encourage dimer formation occur through a β-sheet extension of the catalytic domain of the first protomer by the oligomerization domain of the second protomer.

2- Mutagensis studies have shown that the residues responsible for cofactor binding and catalysis in *P. phosphoreum* LuxC may be the same ones that have been found to accomplish these functions in aldehyde dehydrogenases as their mutation caused similar changes to the kinetic parameters of both PpLuxC and to the *Vibrio harveyi* aldehyde dehydrogenase.

3- Through investigation of the role of the E253 residue of *Vibrio harveyi* aldehyde dehydrogenase, we have found that this single residue may be one of the critical determinants of enzyme function, causing the enzyme to preferentially catalyze either aldehyde oxidation or acid reduction.

List of Tables and Figures

List of abbreviations

Å	=Angstrom
ADP	=Adenosine diphosphate
ALDH	=Aldehyde dehydrogenase
AMP	=Adenosine monophosphate
ATP	=Adenosine triphosphate
CoA	=Coenzyme A
DEAE	=Diethyl aminoethyl
DTT	=Dithiothreitol
FMN(H_2)	=Flavin mononucleotide (reduced)
HEPES	=N-2-Hydroxyethylpiperazine-N'-2-Ethanesulfonic acid
IPTG	=Isopropyl- β-D-thiogalactopyranoside
K	=Kelvin
MMSALDH	=Methylmalonyl-semialdehyde dehydrogenase
NAD(H)	=Nicotinamide adenine dinucleotide (reduced)
NADP(H)	=Nicotinamide adenine dinucleotide phosphate (reduced)
PEG	=Poly-ethelene glycol
PIPES	=1,4-piperazine bis(ethane sulfonic acid)
PpLuxC	=*Photobacterium phosphoreum* LuxC
SDS-PAGE	=Sodium dodecyl sulphate polyacrylamide gel electrophoresis
Vh-ALDH	=*Vibrio harveyi* aldehyde dehydrogenase

Chapter 1-Introduction

1.1-Bioluminescence

The phenomenon of bioluminescence has been of interest to mankind for thousands of years, with the first observations being made by Anaximenes of Miletus, going back to ancient Greece and dating back to 500 B.C. (1). In the 17th century, experimentation was carried out by Sir Robert Boyle, showing that the luminescence reaction was dependent on the presence of air, as cultures grown anaerobically would not luminesce (2), and the phenomenon was often noted by sailors, who would often see what they described as the "phosphorescence of the sea", which we now know was due to bioluminescent algae. Since that time, and especially in the last 25-30 years, much more has been discovered regarding this phenomenon, and it has been found to occur in a multitude of organisms, of both marine and terrestrial origin, and is not limited to bacteria, appearing in higher organisms such as fish, earthworms, beetles and the well-known firefly (3,4,5). The luminescent systems in all of these organisms differ greatly, but a basic description involving two types of molecules can be applied to all.

All systems involve a luciferin and a luciferase. The luciferin is the light emitting substrate, whose name literally means "light-bearing". The luciferase is the enzyme that is responsible for the catalysis of this reaction, or more precisely the oxidation of the substrate, as all luciferases are mixed function oxidases and require oxygen to accomplish their function (3,7).

Many luciferases and luciferins have been identified, and to date there are at least thirty different and independent systems that have been found (3). The luciferases themselves have been isolated and their amino acid sequences have been obtained, and some, three dimensional crystal structures have been solved. Sequence alignments have been performed between many members of this family, and there has been no sequence identity found between them. At the structural level, a comparison of the *Vibrio harveyi* and the *Photinus pyralis* structures, the two luciferase structures that have been solved so far, shows no conserved elements, differing even in the number of domains (Figure 1.1, PDB accession numbers 1LUC and 1LCI, respectively).

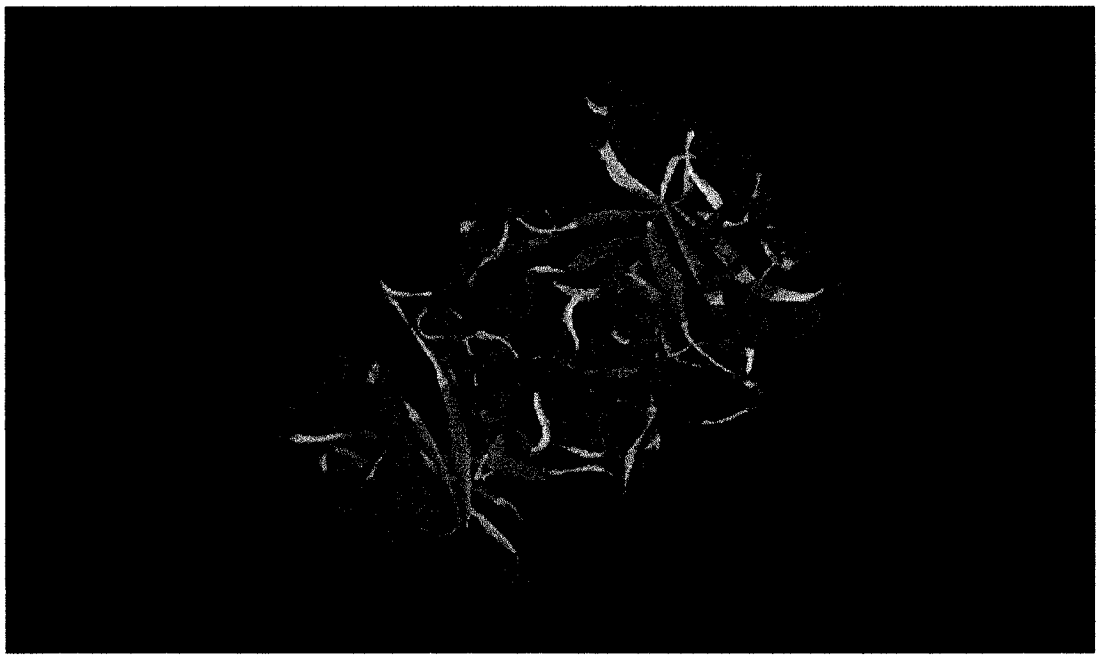
The chemical reactions that the various luciferases are able to catalyze are also very different, especially with regard to the nature and chemical structure of the substrates that they utilize (3,6). Coelenterate organisms, such as the sea pansy *Renilla reniformis* and the *Aequorea* jelly fish, use coelenterazine, an imidazolopyrazine, as a substrate (Figure 1.2, panel A). These luciferases are able to catalyze the oxidation reaction, going through an energy-rich dioxetanone intermediate, and producing carbon dioxide and the oxidized, excited-state product, coelenteramide, which returns to its ground state through photon emission. The firefly luciferases catalyze a different reaction, requiring both ATP and a benzothiazoyl-thiazole luciferin (Figure 1.2, panel D). In this case, the first step is a condensation reaction between ATP and the luciferin, producing a luciferin-AMP intermediate. Reaction with oxygen then takes place, leading to a cyclization of the peroxide to produce, again, a dioxetanone moiety which breaks down and releases energy in the form of light. The mechanism of the dinoflagellates

luciferase, such as that of *Gonyaulax polyedra*, is slightly different, as the substrate, a linear tetrapyrrole (Figure 1.2, panel C), is very sensitive to auto-oxidation and only the luciferase-catalyzed oxidation yields light, as the point of oxidation is different in the chemical (dark) and enzymatic (light) reactions. In either case, the reaction product is non-fluorescent, and since the reaction product is non-fluorescent, it has been proposed that light emission is due to some transient intermediate in the reaction, such as that proposed for bacterial luminescence, but this has yet to be resolved.

The mechanism of light production in bacterial cells requires the presence of two distinct substrates (7,8,9), a long-chain aliphatic aldehyde and $FMNH_2$ (Figure 1.2, panels E and B, respectively).The luciferase first binds the flavin molecule in a hydroperoxide state, followed by reaction with the fatty-aldehyde to form a peroxyhemiacetal. This peroxide bond is then cleaved, a proton is transferred from the aldehyde to the flavin, producing the enzyme-bound 4-a-hydroxyflavin emitter, and the excitation step for this would seem to be a charge annihilation between the radicals on the carbonyl carbon of the aldehyde intermediate and the nitrogen of the isoalloxazine ring system of the flavin, though this is still theoretical as none of the intermediates proposed to occur are stable enough to permit study.

Figure 1.1- Structures of the luciferase of *Vibrio harveyi* (panel A) and *Photinus pyralis* (panel B), obtained from the Protein Data Bank (PDB accession numbers 1LUC and 1LCI, respectively).

A



B

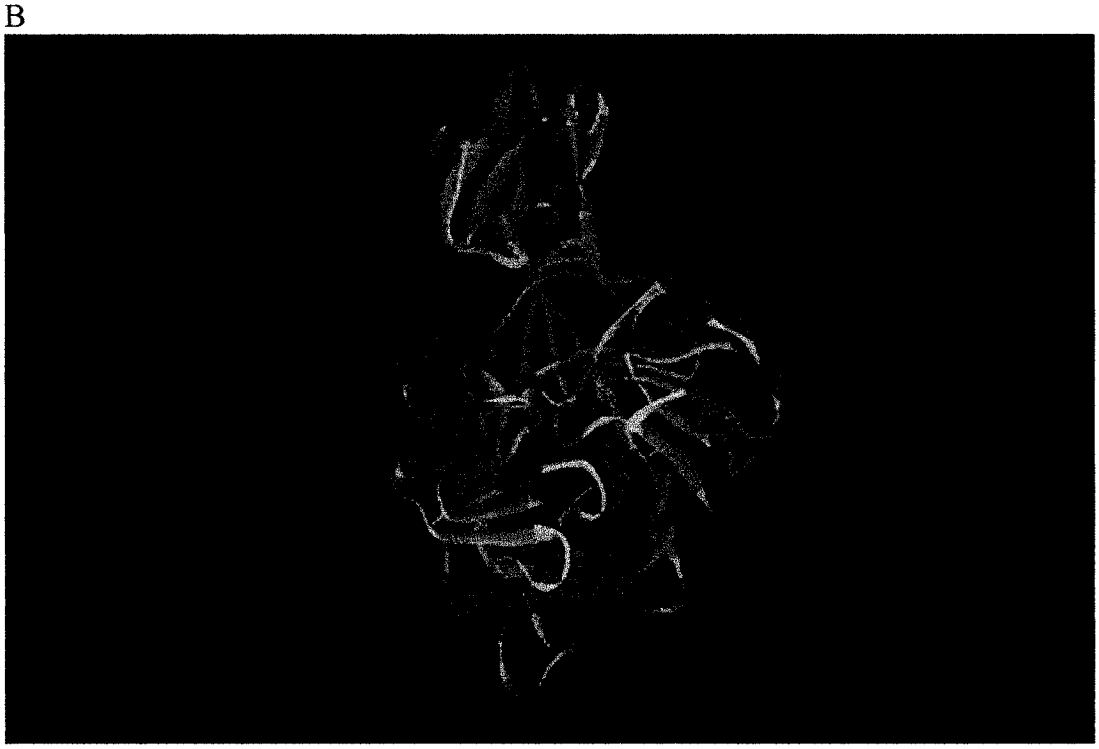
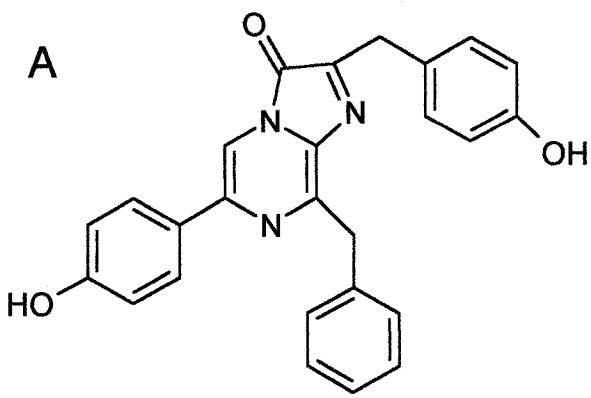
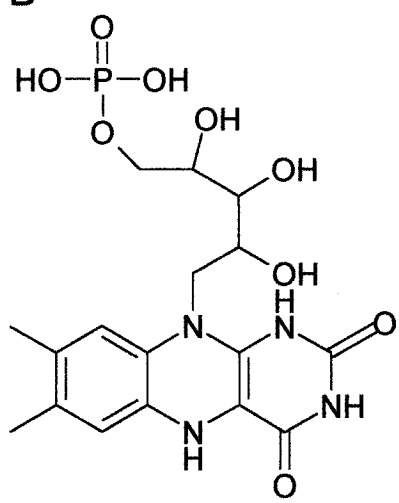


Figure 1.2 – Structures of the luciferins from A-coelenterate jellyfish (coelenterazine), B-bacteria ($FMNH_2$), C-unicellular algae (linear tetrapyrrole), D-firefly (benzothiazoyl-thiazole) and E- bacteria (aliphatic aldehyde).

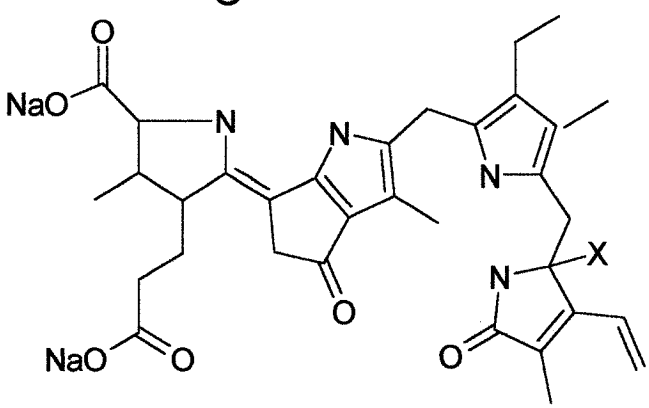
A



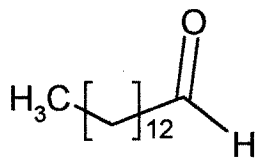
B



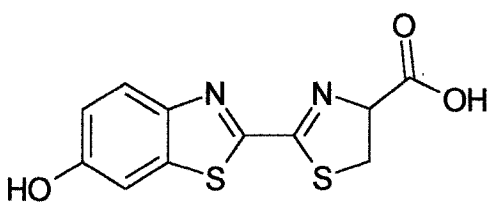
C



E



D



With regard to the control mechanisms that govern the induction and expression patterns of luminescence in different organisms, there is also a very high degree of variability that is seen, even within a single genus, as will be discussed later. In the firefly *P.pyralis* specialized cells have evolved for the luminescence function, termed photocytes, which contain both the luciferase and luciferin (3,4). The photocytes are located in the lantern, the light emitting organ of the firefly, and are arranged in a rosette-like manner around a central cylinder that is composed of nervous tissue and tracheae. It has been proposed, though not proven, that the firefly controls its light emission through nerve impulses sent to the lantern that trigger oxygen entry to the photocytes.

In dinoflagellates, which are unicellular algae, a specific organelle, termed the scintillon, has been found to control luminescence (3,10). These organelles are small vesicles that project into vacuolar space and have a continuous membrane. Activation then depends on an impulse sent through the vesicular membrane, which activates a proton pump and acidifies the normally basic interior of the vesicle. The luciferase goes through a conformational activation and the luciferin is released from a protective luciferase binding protein (LBP) once the pH is acidic. A multitude of other mechanisms exist, but none, at first glance, are as simple as the bacterial method (3). They have been found to be able to luminesce in a continuous manner, instead of short and controlled bursts, and the level of light emitted depends on the level of expression of the components of the luciferase system. It is not, however, as simple as it appears.

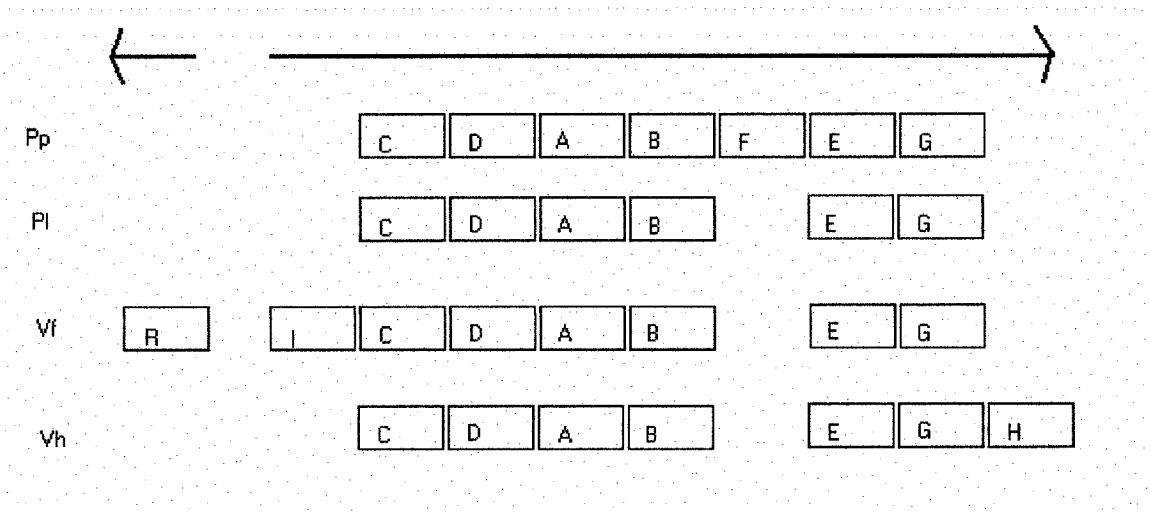
1.2–Bacterial Bioluminescence

The organization of the genes involved in bacterial bioluminescence has been found to be highly conserved in many luminous bacteria, and it is of primary importance to the regulatory mechanisms that control its induction (11). The genes involved in bacterial luminescence, *lux* genes, or their gene products, *lux* proteins, can be grouped into two general categories, the genes responsible for the structural elements of the luminescence system and those involved in the regulatory mechanisms. The first category groups the genes whose products catalyze the light-emitting reaction and who are responsible for the synthesis of the necessary substrates. The latter group encompasses those which are involved, either directly or indirectly, in the regulation of expression of the structural genes.

The *lux* structural genes are grouped, in luminescent bacteria, in a single operon, termed the *lux* operon (7). This group contains the genes that code for the luciferase α and β subunits, LuxA and LuxB, respectively, as well as those encoding for the acyl-reductase, acyl-synthetase and acyl-transferase polypeptides, LuxC, LuxE and LuxD, respectively, responsible for the synthesis of the long-chain aliphatic aldehyde substrate necessary for the luciferase-catalyzed reaction (8,11). A comparison of the *lux* operons from *P. phosphoreum, P. leiognathi, V. fischeri* and *V. harveyi* (Figure 1.3) shows that there is a conserved order of luxCDABE in these organisms, with variability being introduced as additional genes become part of the operon. The luxF gene has been found in certain strains of the *Photobacterium* genus, between *lux* B and E (11), and codes for a

27 kDa non-fluorescent flavoprotein, whose function is yet to be identified. LuxG is also a 27 kDa protein (11,13,14), and its gene has been found immediately downstream of the *lux*E gene in marine, but not terrestrial, bacteria. Sequence alignment has shown that it has a 40% homology to the *E.coli fre* protein, which has flavin reductase activity. In *V. harveyi*, an additional gene has been found downstream of *lux*G, called *lux*H, who shares a 64% identity with the *E.coli HtrP* protein, and has been shown to code for a 3,4-dihydroxy-2-butanone 4-phosphate synthase, important in riboflavin synthesis (11,15,54). In *V. fischeri*, a gene termed *lux*I has been found upstream of *lux*C, and has been found to encode for a 22kDa homoserine lactone synthase(11,12,15), whose importance will be discussed later in this section. Additional proteins have been found to affect the luminescence reaction, such as LuxY in *V. fischeri* and LumP in *P. phosphoreum* and *leiognathi* (3,11), and though they are not part of the operon, they are able to change kinetic properties of the luciferase as well as change the emission spectrum.

Figure 1.3- Comparison of the organization of the *lux* operons of *P. phosphoreum* (Pp), *P. leiognathi* (Pl), *V. fischeri* (Vf) and *V. harveyi* (Vh), where the solid arrows represent the direction of transcription.


Pp
C
D
A
B
F
E
G
Pl
C
D
A
B
E
G
Vf
R
I
C
D
A
B
E
G
Vh
C
D
A
B
E
G
H

The regulatory proteins that control luminescence involve far more complex interactions than the structural genes, and much is yet unknown. It has been recently found that the method by which luminescence is regulated, at least in certain bacteria, involves a phenomenon known as quorum sensing (16). This is a process by which bacteria are able to respond to signals emitted by themselves and other bacteria in the local environment, communicating the level of bacterial growth and providing information on the nutritional levels in the environment in which they are growing, through the production of small, freely-diffusible molecules now referred to as autoinducers. The discovery that chemically similar molecules are produced by plant and animal pathogens that affect virulence, among other properties, has generated considerable interest in the lux system of regulation as a possible model for other organisms(17,18,27,28,29,30). Two systems have so far been investigated in light-emitting bacteria, that of *V. harveyi and V. fischeri*, and though they are not fully understood, will be presented here to give an example of this phenomenon and to show the diversity and complexity that can exist in the control of luminescence, even within a single genus.

The *V. fischeri* system of luminescence regulation involves two proteins, LuxI, which has been previously described, and LuxR. The gene encoding for LuxR (19), seen in Figure 1.3, is found in close proximity to the *lux* operon, separated by a 219 base-pair nucleotide sequence, and is transcribed in the opposite direction. This protein of 29 kDa has been found to be similar to other members of the prokaryotic σ-family of transcription regulators, having a carboxy-terminal domain with a DNA-binding motif,

but differing in its amino terminus. In most transcriptional regulators, there is usually a phosphorylation site at the N-terminus, which has been replaced, in *V. fischeri* LuxR, by a ligand-binding domain. The method of regulation then occurs as follows. LuxI, the homoserine lactone synthase, is produced at low levels during cell growth, and thus produces a constant, though small, amount of 3-oxo-hexanoyl-homoserine lactone, the *V. fischeri* autoinducer. During growth, this molecule, due to its diffusible nature, is able to accumulate in the local environment. Once a critical concentration is reached, usually at the mid- to late-logarithmic stages of growth, it is able to interact with LuxR, causing an increase in the transcriptional level of both *lux*R and the *lux* operon, which includes *lux*I, thus forming a positive-feedback auto-inductory loop that increases luminescence in a drastic fashion over a short period of time. One of the possible functions of this, and a relatively obvious one, is population dispersal, for as the bacteria are multiplying there is a constant consumption of nutrients and an accumulation of waste products, and when the population density is high enough, luminescence can attract predators and encourage predatory actions that would lead to population dispersal, allowing the bacteria to propagate and continue to multiply.

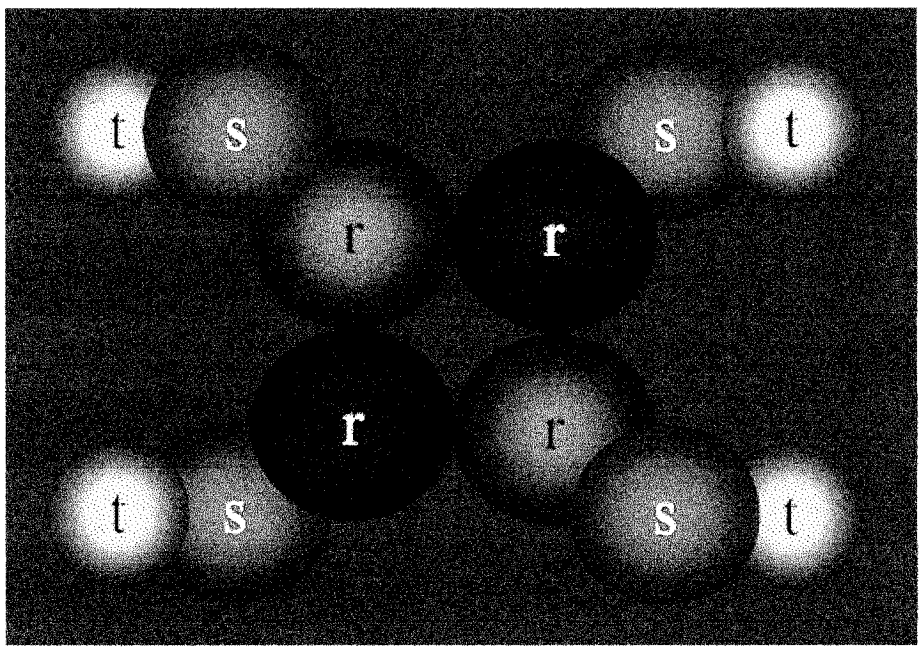
In contrast to the relative simplicity in regulation of the *V. fischeri lux* system, that of *V. harveyi* differs considerably and is far more complex, much of which is still unknown to researchers. The regulatory mechanism involves a far greater number of components, counting among them Lux M,N,O,P,Q,R,S and U and some as of yet unidentified components (20,21). The LuxM and LuxS proteins have been found to be synthases producing a 3-hydroxybutanoyl homoserine lactone and a furanosyl-borate

ester, respectively, which function as autoinducers in much the same way as the autoinducers in *V. fischeri,* and which will be referred to as AI-1 and AI-2, respectively, for simplicity. As in *V. fischeri*, these are small, freely-diffusible molecules that accumulate in the local environment, but they do not, however, directly bind to a transcriptional regulator. Instead, AI-1 and AI-2 have been found to interact with the LuxN and LuxP/Q proteins, respectively, which act as receptors on the cell membrane and are the first part of a phosphorelay system. These receptors have the ability to be phosphorylated by ATP at specific histidine residues, which then causes a subsequent transphosphorylation of a downstream aspartate residue. This phosphorylation leads to the recruitement and phosphorylation of LuxU, yet another component of the relay system. In its phosphorylated form, LuxU acts as a kinase, phosphorylating LuxO, currently thought to be the central luminescence response regulator in *V. harveyi.* LuxO bears similarity to the prokaryotic σ-^{54} family of transcription regulators, and in its phosphorylated form has been found to repress the expression of LuxR, which is unrelated to the LuxR protein of *V. fischeri*, through the expression of an as of yet unidentified factor X. At a specific concentration, again occurring at the mid- to late-logarithmic stages of cell growth, the autoinducers are able to interact with their cognate receptor, which causes them to change from a kinase to a phosphatase activity. This causes the de-phosphorylation of LuxU, and subsequently LuxO. Once LuxO is no longer able to repress the expression of LuxR, also a transcriptional regulator, LuxR production rises and so does the expression of the *lux* operon, leading to light production (21).

1.3- Fatty Acid Reductase complex

In order to produce light, bacterial luciferases require three substrates, a long-chain aliphatic aldehyde, reduced flavin mononucleotide ($FMNH_2$) and oxygen (7, 8, 9). While the latter two substrates are readily available, the long chain aldehyde is not since high levels of this reactive compound can be very toxic to the cell. As such, the synthesis of these aldehydes is a process which must be tightly controlled. As previously seen, the expression of the genes responsible for aldehyde synthesis is highly regulated, occurring concurrently with that of the aldehyde-consuming luciferase. Though these genes are present in most luminous bacteria, the activities of the LuxC, D and E polypeptides have so far only been isolated and purified from *P. phosphoreum* (22). These three proteins have been found to associate into a large multimeric complex, termed the fatty acid reductase (FAR) complex (22,23), composed of four subunits of each component, and has been found to be responsible for aldehyde biosynthesis. LuxC has been found to associate as a tetramer, which acts as the core of the complex, and to which LuxE has been found to bind. Following this, each LuxE is able to interact with a LuxD protein, though weakly, forming a dodecamer of over 500 kDa in molecular mass and depicted, in cartoon form, in Figure 1.4.

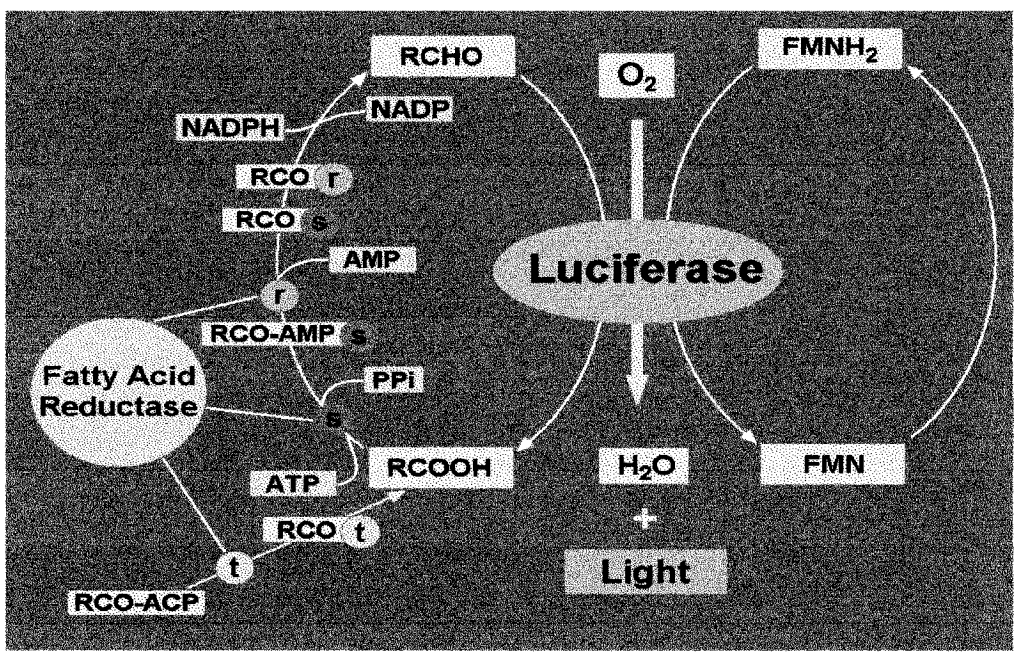
Figure 1.4- Schematic representation of the fatty-acid reductase complex from *P. phosphoreum*, where r is the acyl-reductase (LuxC), s is the acyl-synthetase (LuxE) and t is the acyl-transferase (LuxD).


t
s
s
t
r
r
r
r
t
s
s
t

The mechanism by which aldehyde production is thought to occur is depicted on the left side of Figure 1.5 (9,11,24,25). The fatty acyl group must first be diverted from another biosynthetic pathway, usually from an acyl-acyl carrier protein (ACP) or from an acyl-CoA. The acyl-transferase, the **t** subunit in the Figure (LuxD), is able to accomplish this through its esterase function, cleaving the bond from the acyl group to its carrier through the action of a serine residue, producing an acyl-acyl-transferase adduct that can be cleaved and release free acid. It has been curiously observed that the rate at which this cleavage occurs, that of the acyl group from the serine residue, is increased in the presence of the **s** subunit in the Figure, the acyl-synthetase (LuxE), possibly arising from a conformational change that occurs during their interaction.

Following release of the fatty acid, the synthetase subunit is able to catalyze an ATP-dependent activation which results in an acyl-AMP adduct, and pyrophosphate formation, and though this activated acyl group is not covalently bound to the synthetase, it remains firmly associated to it. A cysteine residue on the synthetase is then able to attack the acyl-AMP adduct, release AMP and form an acyl-acyl-synthetase intermediate. Lux E has the ability to release the acyl group, though this occurs at a very low level. Again, the rate of transfer of the acyl group to the synthetase has been found to increase in the presence of the acyl-reductase, LuxC, even if the reductase has been previously inactivated, again possibly owing to a conformational change that occurs during the interaction of the two polypeptides. The final step in the reaction, once the acyl group has been transferred to a cysteinyl residue on the reductase subunit, is the NADPH-dependent reduction of the acyl group and release of free aldehyde.

Figure 1.5- Diagrammatic representation of the production of the reactive aldehyde substrate (left) and the luciferase reaction (right) that occurs during the luminescence reaction.


RCHO
O_2
$FMNH_2$
NADPH
NADP
RCO r
RCO s
AMP
Luciferase
r
RCO-AMP s
Fatty Acid Reductase
PPi
s
RCOOH
H_2O
FMN
ATP
RCO t
+
t
Light
RCO-ACP

Due to the increase in activity that is observed for the LuxD and LuxE polypeptides when the binding partner for the next step in the reaction is present, it has been suggested that there is a channeling mechanism that is present in the complex (26). A model has been proposed where the cysteinyl residue of the synthetase subunit comes in close proximity to the reductase cysteine residue, possibly allowing for a direct transfer of the acyl moiety. This would of course be advantageous to the catalytic efficiency of the enzyme complex, but would also confer an advantage to the cell as the potentially toxic intermediates produced during the reaction would not be able to damage the cell.

During the course of investigation of the properties of LuxC, it was noticed that they bore some resemblance to those aldehyde dehydrogenases (ALDH's) which are studied in our lab. Both ALDH's and LuxC reductases are of a comparable size, usually 50-55 kDa, and the cysteine residue which is found to be catalytically essential for ALDH's and LuxC reductases is found in a similar position in their amino acid sequences. Despite the fact that the direction of the catalytic reaction is opposite (oxidation vs. reduction), their similar substrates have sparked an interest in the possible catalytic and structural links between the two enzymes.

1.4- *Vibrio harveyi* aldehyde dehydrogenase

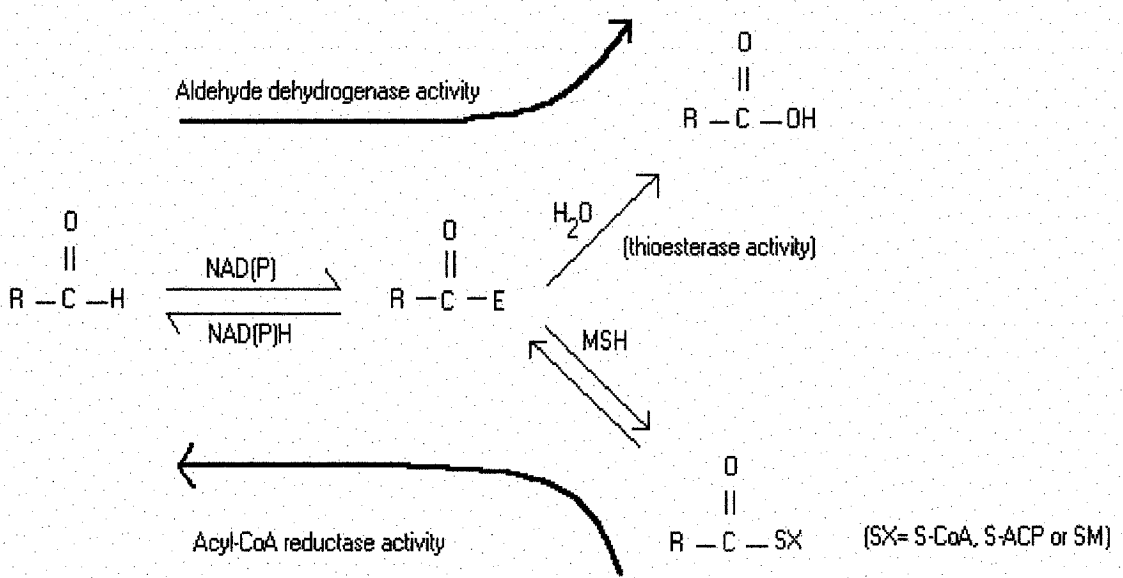
The aldehyde dehydrogenase superfamily of enzymes is one that is quite large, and to date counts well over 500 members from many different organisms and tissues, separated into different classes (31,32). Class 1 aldehyde dehydrogenases are usually cytosolic and are ubiquitously distributed in a variety of tissues. Class 2 enzymes are mitochondrial and they are found to be strongly expressed in liver tissue. These 2 classes show a nearly 70% identity in sequence, are usually functional as tetramers and commonly use NAD as a cofactor to carry out their reactions. The third class of aldehyde dehydrogenases are usually cytosolic, though they differ substantially from the two first classes, as they are usually homodimers, are able to use both NAD and NADP as cofactors and show a low level of sequence identity (<30%) with the ALDH's of the first two classes. It has also been noticed that the *N*-terminal domain of the class 1/2 enzymes is longer, and that the class 3 family has an extended *C*-terminal domain, which according to structural analyses affects the oligomerization state. Sequence alignment of 145 aldehyde dehydrogenases has revealed the existence of 10 conserved sequence motifs, in which residues thought to be catalytically important are located, and which have been found in most ALDH's, irrespective of class or origin (32).

The *V. harveyi* aldehyde dehydrogenase, an enzyme with polypeptides of 55 kDa, naturally associates as a homodimer, as do the polypeptides of class 3 ALDH's (33). Though the sequence homology to other dehydrogenases is low (<25%), it has been seen that of the 23 residues which are at least 80% conserved in other dehydrogenases, 19 are

present in the *V. harveyi* aldehyde dehydrogenase. This enzyme, though, shows an unusual preference for NADP over NAD. This type of specificity has only been seen in one other ALDH whose structure has been characterized, that isolated from *Streptococcus mutans* (34). Structural characterization of these two enzymes has implicated certain residues for NADP specificity, which will be discussed in Chapter 4.

Alignment of the previously mentioned sequence motifs in aldehyde dehydrogenases with the sequence of PpLuxC has shown that some of these motifs are also present in LuxC, and through this resemblance allows certain residues to be recognized that may play an important role in substrate specificity and in catalysis. The reactions carried out by both ALDH's and LuxC reductases also show some commonalities, and though the initial substrates are not the same, parts of the reaction pathway are very similar, as seen in Figure 1.6. Through these similarities, and the experiments performed on potentially common catalytic resides, it is our hope to show that the *P. phosphoreum* LuxC is part of the aldehyde dehydrogenase extended family and that it may be the first member of this family to preferentially catalyze the reductase reaction as opposed to the dehydrogenase reaction.

Figure 1.6- Reaction scheme of the mechanism of the dehydrogenase and reductase activities, where CoA is co-enzyme A, ACP is acyl-carrier protein and M is a methyl group


Aldehyde dehydrogenase activity
O
||
R — C — OH
O
||
R — C — H
NAD(P)
NAD(P)H
O
||
R — C — E
H2O
(thioesterase activity)
MSH
Acyl-CoA reductase activity
O
||
R — C — SX
(SX= S-CoA, S-ACP or SM)

Chapter 2 - Materials and Methods-

Expression and purification of *P. phosphoreum* LuxC

The gene containing the native PpLuxC was inserted into the PET-23a+ vector which was then used to transform competent *E.coli* BL21 cells. Transformed cells were then grown at 30°C in Luria-Burtani medium containing 75 μg/mL of ampicillin to an OD_{600} of between 1.0 and 1.2, after which they were induced for protein production with IPTG, to a final concentration of 0.6 mM, and allowed to incubate for one additional hour at 30°C. The cells were then harvested by centrifugation and the resultant cell pellet was stored at -70°C overnight. The pellet was then sonicated on ice in a buffer of 50 mM Na/K phosphate buffer, pH= 7, with 20 mM β-mercaptoethanol and the lysate was cleared by centrifugation. The resulting solution was loaded onto a DEAE-sepharose chromatography column and eluted with a Na/K PO_4 gradient of 0.05-1.0 M, with 20 mM β-mercaptoethanol. Fractions containing the protein of interest, as determined by SDS-PAGE analysis, were then dialyzed against a buffer of 50 mM Na/K phosphate buffer, pH= 7, with 20 mM β-mercaptoethanol and then loaded on a 2'-5'-ADP-sepharose affinity chromatography column (23,35) and the protein was eluted using a 1 mM NADP-containing 50 mM Na/K phosphate buffer, pH= 7, with 0.2 M NaCl and 20 mM β-mercaptoethanol. The resultant pure protein, as determined by SDS-PAGE, was dialyzed into a solution of 30% glycerol, 50 mM Na/K phosphate, pH= 7, and 20 mM β-mercaptoethanol and stored at -20°C.

Site-directed mutagenesis of *P. phosphoreum* LuxC

Site-directed mutagenesis was carried out on PpLuxC cloned into the M13mp19 vector using the Muta-Gene M13 *in vitro* mutagenesis kit (Bio-Rad). The procedures were followed as per the instruction manual and mutants were generated through the use of mutagenic primers designed to create the N152D, K174Q and S177E substitutions.

Expression and purification of the mutant forms of *P. phosphoreum* LuxC

Expression and purification of the N152D, K174Q and S177E mutants of PpLuxC was done just like the native recombinant enzyme, except that the mutant enzymes would not bind the 2'-5'-ADP-Sepharose affinity chromatography column and were thus passed only through the DEAE-Sepharose column, with the exception of N152D which was passed twice through this column. The resultant proteins were of sufficient purity (>80%), as judged by SDS-PAGE, and final enzyme preparations were dialyzed into the same storage buffer as the native recombinant enzyme and stored at -20°C.

Expression and purification of the native and mutant forms of *V.harveyi* ALDH

The native Vh-ALDH, and its K172Q, T175E and E253A mutants were prepared just as the wild-type PpLuxC until the purification steps, with the exception that induction in the presence of IPTG was carried out for two hours instead of one. The native enzyme, along with its E253A mutant, were purified in a single affinity chromatography step using a 2'-5'-ADP-Sepharose column (40,61) by washing with 50

mM Na/K phosphate buffer, pH=7, 20 mM β-mercaptoethanol before and eluting in the same buffer containing 100 μM NADP. The T175E mutant of ALDH was purified in a single step as well, using a β-NAD-Sepharose column, using the same buffers as the previous enzymes for washing and elution, except that the elution buffer was modified to contain 1 mM NAD instead of 100 μM NADP. The K172Q mutant of ALDH would not bind either of the affinity chromatography columns, and was purified by being passed through a DEAE-Sepharose column twice, using a Na/K phosphate buffer gradient from 0.05 to 1.0 M, with 20 mM β-mercaptoethanol, for the first column and the same buffer, with the addition of 0.5 M NaCl, for the second column. All enzymes were assayed for purity by SDS-PAGE and, once found to be of sufficient quality (>80%), were stored in the same buffer and conditions as the native recombinant PpLuxC.

Aldehyde dehydrogenase activity

Aldehyde dehydrogenase activity was measured for all proteins using freshly made NAD or NADP in a Na/K phosphate buffer at pH= 8, and activity was measured at a saturating concentration of dodecanal (1% v/v of a 0.1% stock in isopropanol). Enzyme concentrations used for each assay varied between mutants and wild-type enzymes, and was reflective of their ability to catalyze the reaction. NAD(P)H production was measured fluorometrically, using a Hitachi F-3010 fluorometer, with an excitation wavelength of 340 nm and using the emission of the excited reduced NAD(P)H at 460 nm. The wavelength dispersion used for both excitation and emission wavelengths was 10 nm. The Michaelis constants and k_{cat}s were calculated from Lineweaver-Burk

plots of activity versus NAD(P) concentration, and statistics were derived from linear correlation coefficients output by Excel™.

Acyl-CoA reductase activity

Acyl-CoA reductase activity was measured for all enzymes using a luciferase-coupled assay and measured using a photomultiplier tube where 1 light unit is equal to 9 X 10^9 quanta/sec (55). Freshly made NADPH or NADH, of varying concentrations, was incubated with 5 μM myristoyl-CoA and a small amount of the enzyme in a 50 mM Na/K phosphate buffer at pH=7, containing 20 mM β-mercaptoethanol and 10 mM Mg^{2+}, and allowed to react for between 1 to 4 minutes, after which 5 μg of native *P. phosphoreum* luciferase was added. The sample was then placed in the photomultiplier tube and 1.0 mL of 50 μM chemically-reduced flavin mononucleotide ($FMNH_2$) injected. The light produced by the luciferase is reflective of the amount of aldehyde produced, which can then be used to calculate reductase activity. The kinetic constants, again, were calculated from Lineweaver-Burk plots of activity versus NAD(P)H concentrations and statistics derived from linear correlation coefficients output by Excel™ .

Crystallization of native *P. phosphoreum* LuxC

Native *P. phosphoreum* LuxC was prepared as described above and was dialyzed into a buffer of 20 mM HEPES, pH= 7, 10% glycerol, 0.2 M NaCl and 10 mM DTT for crystallization purposes, with a final concentration of 6.8 mg/ml. Prior to being setup for crystallization, the protein was incubated for 30 minutes on ice with 15 mM NADPH and

25 mM DTT, so as to allow formation of a binary enzyme·NADPH complex. Using a method of hanging-drop vapor-diffusion with a mother liquor of 10% PEG-3350, 3%PEG-400, 10% glycerol, 400 mM NH_4Cl and 100 mM PIPES, pH=6.5, thin plate-like crystals were obtained after 2-3 days of incubation at 17°C. To improve crystal quality, the plate-like crystals were crushed by vortexing with Teflon™ balls in mother liquor and the resulting solution was used as a seeding solution for crystal growth (62). The mother liquor was also modified to improve crystal growth, increasing the concentration of PEG-3350 from 10% to 13%, changing nothing else. A drop containing 1 μL of modified mother liquor and 1 μL of protein solution was allowed to equilibrate, after which a cat whisker dipped in seed-solution was streaked through the drop. The streak-seeding technique improved quality of the crystals, increasing the size and thickness of the crystals and resulting in crystals of a hexagonal morphology and useable size to appear after 5-7 days.

Expression, purification and crystallization of selenomethionyl-PpLuxC

The selenomethionyl-derivative of acyl-CoA reductase was produced just as the native enzyme, with the exception that all media used during growth contained selenomethionine instead of methionine, as described for LeMaster medium (46) (contains all amino acids except methionine, and is supplemented with thiamine) and using the DL41 *E. coli* Met^- strain of cells. Purification and crystallization occurred just as the native enzyme, except that no plate-like crystals were initially obtained, so seeding was done using the seeds obtained for the native enzyme, and incubation at 298 K for 5-7

days was sufficient to produce crystals resembling the native protein crystals in size and appearance.

Data collection and preliminary crystallographic analysis

A dataset was collected for the native protein at the Stanford Synchrotron Radiation Laboratories (SSRL, Stanford, California) using a cryostream device to keep the crystals at 100 K and diffraction was measured using a MAR345 detector. Crystals were looped in nylon loops of an appropriate size and were frozen directly in the cryostream, without additional cryoprotectant. For the selenium derivative, three datasets were collected at the tunable beamline 5.0.2 of the Advanced Light Source (ALS, Berkeley, California), collecting data at the peak, inflection and high-energy remote wavelengths for selenium on an ADSC Quantum210 detector. The selenium-derivative crystals were looped in nylon loops and frozen in the form of propane "popsicles" in liquid nitrogen. These "popsicles" were directly mounted on the goniostat and allowed to thaw under a stream of nitrogen at 100 K.

For the native dataset, a 180° oscillation range was used, in slices of 1 degree, and a full dataset was collected on a single crystal. The selenomethionine-derivative datasets were collected over a 200° oscillation range, collecting 100° of data on two opposing quadrants, in wedges of 30 degrees, using a 1 degree oscillation per diffraction image, and all three datasets were obtained on a single crystal.

The data collected was processed and scaled using DENZO/SCALEPACK of the HKL suite of software (43,44). To obtain an initial solution, the peak and inflection datasets were used in a direct methods program, SnB (Shake'n'Bake) (48). This allowed the identification of 20 of the 22 selenium positions. These positions were further refined using the program MLPHARE of the CCP4 suite of software(48). DM, a program which performs density modification, also of the CCP4 suite of software, was used to improve the quality of the initial electron density map. Using MAPMAP, of the Uppsala software factory (USF, Uppsala, Sweden) (49), the initial electron density map was collapsed and the resulting bones model was used to build the initial model, along with the map itself, in the program O (50). Using the selenium sites obtained, a rotation matrix was also calculated and from the monomer that was modeled, the full tetrameric structure of LuxC was calculated.

Acylation and De-acylation experiments

The wild type and N152D mutant of *P. phosphoreum* LuxC , as well as the E253A mutant of *V. harveyi* aldehyde dehydrogenase were first dialyzed into a buffer of 10% glycerol, 50 mM Na/K phosphate, pH= 7, 0.2 M NaCl and 0.25 mM DTT overnight. The freshly dialyzed enzymes were then incubated with 10 μM [^{3}H]myristoyl-CoA, at 2 Ci/mmol, over a period of time to allow acylation to occur (63). To measure the level of acylation, the reaction was stopped at various time points using a 1:1 sample buffer for SDS-PAGE (70 mM Tris, pH= 6.8, 15% glycerol, 1.4% sodium dodecyl sulphate) that contained 5mM N-ethylmaleimide (NEM). The acylated proteins were then allowed to

react for one hour at 37 °C with NEM, after which they were separated by SDS-PAGE. The gels were stained, destained and then soaked for 30 minutes in a solution of Amplify™, stained using Coomassie R-250 brilliant blue and destained in a solution of 10% acetic acid, 40% methanol and 50% water. They were then dried over Whatman filter paper and exposed to Kodak XAR film for 24-72 hours at -70 °C. The rate of acylation/deacylation was calculated by scintillation counting of the bands from the SDS gel, by first soaking the gel/filter paper in 200 µL of water, removing the filter paper, then solubilizing the gel slice by adding 400 µL of a 90% hyamine HCl solution and shaking at 37 °C overnight and then adding 10 mL of Cytoscint™ scintillation cocktail and measuring counts in a Beckman LS-3801 liquid scintillation counter. To measure de-acylation, NADPH was added to the acylated samples and allowed to incubate for up to 20 minutes, stopping the reaction at various time points using the same NEM-containing sample buffer as above, after which the procedure is identical.

Chapter 3 - Kinetic and mutational analysis of *Photobacterium phosphoreum* LuxC and *Vibrio harveyi* ALDH.

The *P. phosphoreum* LuxC is known to be responsible, in part, for the synthesis of the long-chain aliphatic aldehyde substrate necessary to the bioluminescence reaction (35). It has been shown that this is accomplished through the formation of an acyl-LuxC intermediate that is followed by the NAD(P)H-dependent reduction of the acyl chain. As seen in Figure 1.6 (Chapter 1), the mechanism by which LuxC operates is, in part, similar to the mechanism by which the *V. harveyi* aldehyde dehydrogenase, among other aldehyde dehydrogenases, catalyses the oxidation of aldehydes, both of them performing redox chemistry on their attached acyl groups. This observation, coupled to the fact that (a)-the polypeptides are of a similar size, (b)-that the tetrameric structure of LuxC is similar to the dimeric and tetrameric forms seen in ALDH's (36) and that (c)-sequence alignment has shown some homology between PpLuxC and ALDH's, has prompted further investigations into the possible link between PpLuxC and Vh-ALDH.

As mentioned in Chapter 1, the aldehyde dehydrogenase superfamily of enzymes counts well over 500 members from a variety of tissues and organisms. Though these enzymes are able to utilize a number of different substrates, the basic reaction chemistry remains the same. Multiple sequence alignments have been performed and have revealed the existence of several motifs which are highly conserved among these enzymes, shown

in Figure 3.1 and listed in Table 3.1 (32). Though these motifs are spread throughout the primary sequence, the three-dimensional structures have shown that these motifs cluster in the vicinity of the active-site, showing that certain elements are necessary for the redox chemistry that is carried out and the enzymatic mechanism. These motifs have been show to contain highly conserved amino acids residues, some of which serve mechanistically important roles. From the alignment of 145 different aldehyde dehydrogenases, 16 residues have been identified to be conserved in 90% of cases, and of these, 12 have been shown to cluster in 7 of the 10 conserved sequence motifs (shown in Table 3.1 in bold). Sequence alignment of PpLuxC with the Vh-ALDH sequence and a consensus sequence for ALDH's (Figure 3.2) has shown that there is a low level of identity that is present (<15%), and that 8 out of the 12 residues which are found in the conserved sequence motifs are present in PpLuxC in similar positions.

Figure 3.1- Sequence alignment of 145 aldehyde dehydrogenases with the highly conserved motifs highlighted in color (obtained at http://www.psc.edu/biomed/pages/research/Col_HBN_ALDH.html).

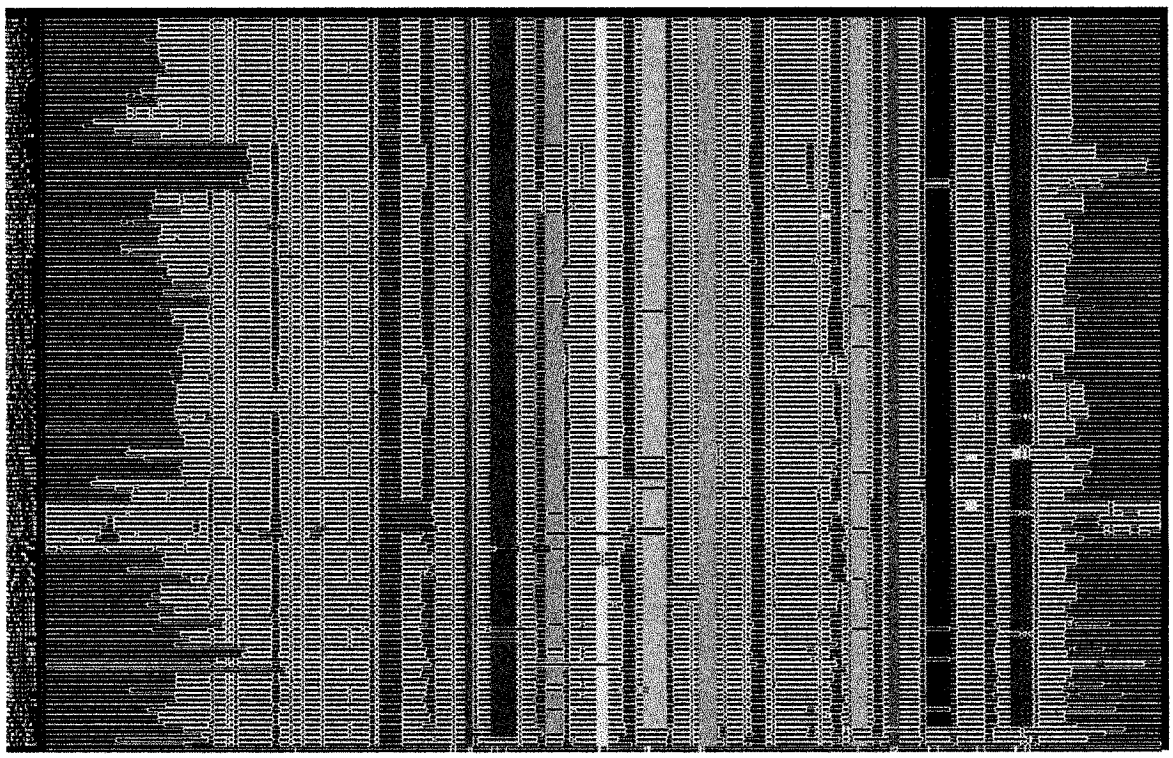


Table 3.1- Ten most conserved motifs in aldehyde dehydrogenases obtained through multiple sequence alignments(32).

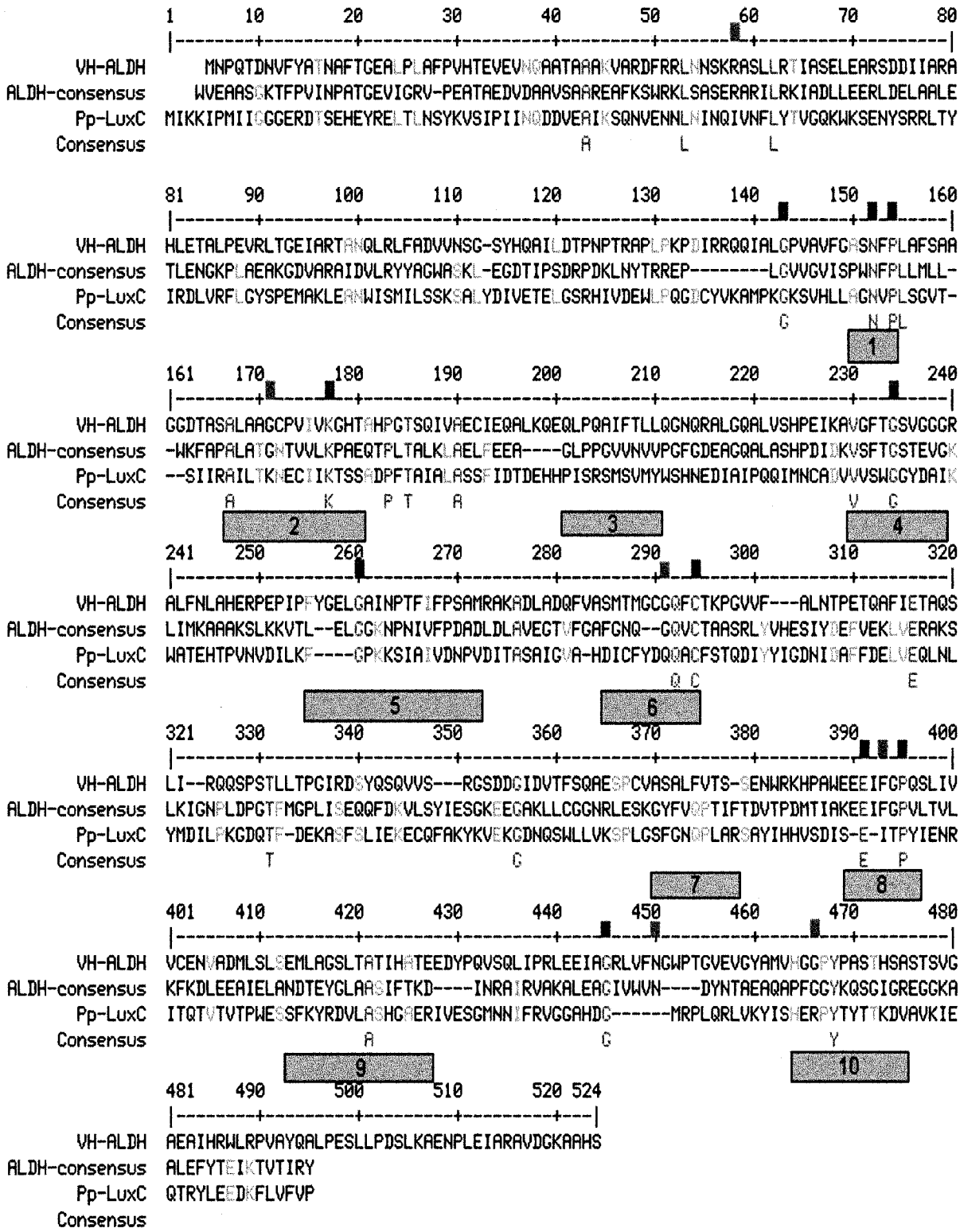
Motif Number	Length	Sequence[1]	Indices[2]
1	5	P-W/F-**N**-F/Y-**P**	279-283
2	14	A-L-A-A-**G**-N-T-V-V-L-**K**-P-A/S-E	296-309
3	10	G-F/L-P-P-G-V-V/L-N-V/I-I/V	327-341
4	10	I/V-S/A-F-T-**G**-S-T-E/A-V/T-G	364-373
5	16	L-E-L-**G**-G-K-S/N-P-X-I-V-F-D-D-A-D	397-416
6	8	F-F-N-Q/A-**G**-Q-X-**C**	430-437
7	9	G-Y-F/Y-I/V-Q-P-T-V/I-F/L	533-542
8	7	E-**E**-I-**F**-G-**P**-V	560-566
9	15	N-D-T/S-E-Y-G-L-A-A-A/G-V/I-F-T/S-K/R-D/N	586-600
10	12	P-F-G-**G**-F-K-X-S-G-I-G-R	641-654

1-Letters in bold are conserved in at least 90% of the sequences used in the alignment and X represent residues with no clear consensus.

2-Indices are numbered consecutively in order of appearance in ALDH sequences

Of particular importance are the identification of a cysteine and an asparagine residues potentially involved in the catalytic mechanism of PpLuxC. The cysteine residue in ALDHs, found in motif 6, which has been determined to be the catalytic active-site nucleophile (36,56), and the asparagine residue in motif 1, which has been implicated in the stabilization of the thiohemiacetal intermediate (32,37,56), are the two most highly conserved residues in ALDHs and have been found to be present in PpLuxC at similar positions in the sequence (Cys^{286} and Asn^{152}).

Figure 3.2- Sequence alignment of the consensus sequence of aldehyde dehydrogenases, the sequence of *V. harveyi* aldehyde dehydrogenase and the sequence of *P. phosphoreum* LuxC performed with hierarchal clustering outlined in Multalin (53). The positions of the ten conserved motifs in the aldehyde dehydrogenase consensus sequence are shown by numbered yellow bars. The highly conserved residues (>90%) in aldehyde dehydrogenases that are present in PpLuxC are highlighted with blue bars and those which are highly conserved and not found in PpLuxC are highlighted with red bars. Residues in orange show pair-wise identities of PpLuxC with either the *V. harveyi* aldehyde dehydrogenase or the consensus sequence for aldehyde dehydrogenases. Residues in red show positions which are conserved in all three sequences. The consensus sequence for aldehyde dehydrogenases was obtained through multiple alignment of available aldehyde dehydrogenase sequences. Numbering starts at the first position of the *P. phosphoreum* LuxC sequence and includes gaps.


VH-ALDH
ALDH-consensus
Pp-LuxC
Consensus
1
2
3
4
5
6
7
8
9
10

Motif 2 has shown itself to be important in ALDHs in influencing cofactor specificities and affinities, and two positions have shown themselves to be specifically important (38,39). A lysine residue in motif 2 (residue 177 of Figure 3.2), has been implicated in interacting with the nucleotide cofactor, along with a residue three amino acids downstream. All three sequences show the presence of the conserved lysine residue (K172 in Vh-ALDH and K174 in PpLuxC). In NAD-specific enzymes, the amino acid residue three amino acids downstream is often found to be a glutamate residue, whereas NADP-preferring enzymes have shown this residue to usually be a serine or threonine. The ALDH consensus sequence in Figure 3.2 shows a glutamate in this position, as the majority of ALDH's are NAD-specific. When comparing the sequences of PpLuxC and the NADP-specific Vh-ALDH, however, we see that a threonine residue is found in the Vh-ALDH at this position (T175) and that PpLuxC has a serine residue (S177), which is consistent with their reported cofactor specificities.

Important differences in the sequence have also been found between Vh-ALDH and PpLuxC. These differences may provide further functional information and insight into how a single amino acid residue may affect the function of an enzyme. Two glutamic acid residues, E209/268 and E333/399 in Class 1&2 and Class 3 ALDHs have long been the object of study and their functions are still debated (40,57,58,59,60). The first of these glutamates has been suggested to be involved in aiding in the base-activation of the active-site cysteine nucleophile and/or the coordination of a water molecule, so as to increase its nucleophilic character, and encourage its attack on, and the release of, the oxidized acyl group (the enzyme's thioesterase activity). The second glutamate has also

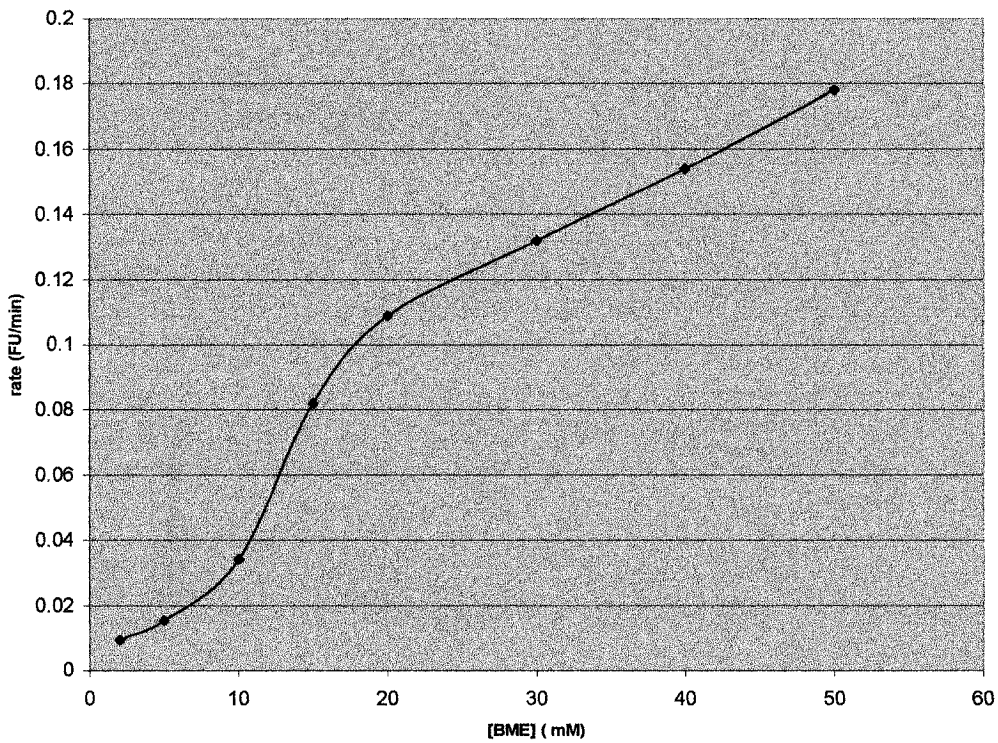
been implicated in these functions, and has additionally been suggested to help in the correct placement of the niacin ring of NAD(P) to allow efficient hydride transfer. Structural characterization has not been able to clarify this question, as both residues are in close proximity to the active-site cysteinyl residue and could be in a favorable position to interact, depending on the specific ALDH. The ambiguity here lies in the fact that the cysteine residue in the active-site has much higher temperature factors than its surrounding protein environment, indicative of thermal motion which could alter its position to favor either of the glutamate residues. Functional analysis, though, has provided some information. Methyl-malonyl semi-aldehyde dehydrogenases (MMSALDH's), are part of the ADLH superfamily and are known to be lacking the first, but not the second, glutamate residue (32), and it has been shown that these particular ALDH's are not able to cleave and release their final product as a free acid but rather as a CoA ester, implying the necessity for a molecule with a higher nucleophilicity than water to allow product release. This may indicate that the first, and not the second, glutamate is necessary for water activation and thus thioesterase activity. Additionally, it has been seen that mutation of the first glutamate in human liver mitochondrial ALDH causes a decrease in the turnover rate, as well as a change in the rate-limiting step from hydride transfer to product release (64), and that mutation of the second glutamate also causes a decrease in the catalytic activity of the Vh-ALDH, but that no change to the rate-limiting step has occurred (40). Together, this may provide evidence that it is indeed the first glutamate residue which is involved in the thioesterase activity, and that the second residue influences hydride transfer, but the base-activation (deprotonation) of the active site cysteine remains in question. The absence of this first glutamate in PpLuxC would

encourage its implication in thioesterase activity, as this activity is not required in the acyl-CoA reductase reaction and is thus consistent with its primary role as a reductase.

3.1 – Dehydrogenase and acyl-CoA reductase activities of PpLuxC and Vh-ALDH.

The *V. harveyi* aldehyde dehydrogenase has been previously shown to have a higher affinity for NADP over NAD during its catalysis of long-chain aliphatic aldehyde oxidation, and has equally been shown to be able to catalyze the reduction of activated fatty acids to aldehydes, again with a marked preference for NADPH over NADH (33). Previous studies have also demonstrated that PpLuxC is able to catalyze the reduction of activated fatty acids with a higher affinity for NADPH over NADH (35), but cannot catalyze the dehydrogenase reaction. During the course of this investigation, it was found that LuxC could indeed function as a dehydrogenase by adding a high concentration of exogenous thiol (i.e. β-mercaptoethanol) (Figure 3.3), which presumably acts as a nucleophile to cause product release, much in the same way as MMSALDHs require CoA.

Figure 3.3- Dehydrogenase activity of the native *P. phosphoreum* LuxC, with saturating dodecanal concentration and 1 mM $NADP^+$, at neutral pH with varying β-mercaptoethanol (BME) concentrations. The rate is measured as a fluorescence (FU) change per minute.


0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
rate (FU/min)
0
10
20
30
40
50
60
[BME] (mM)

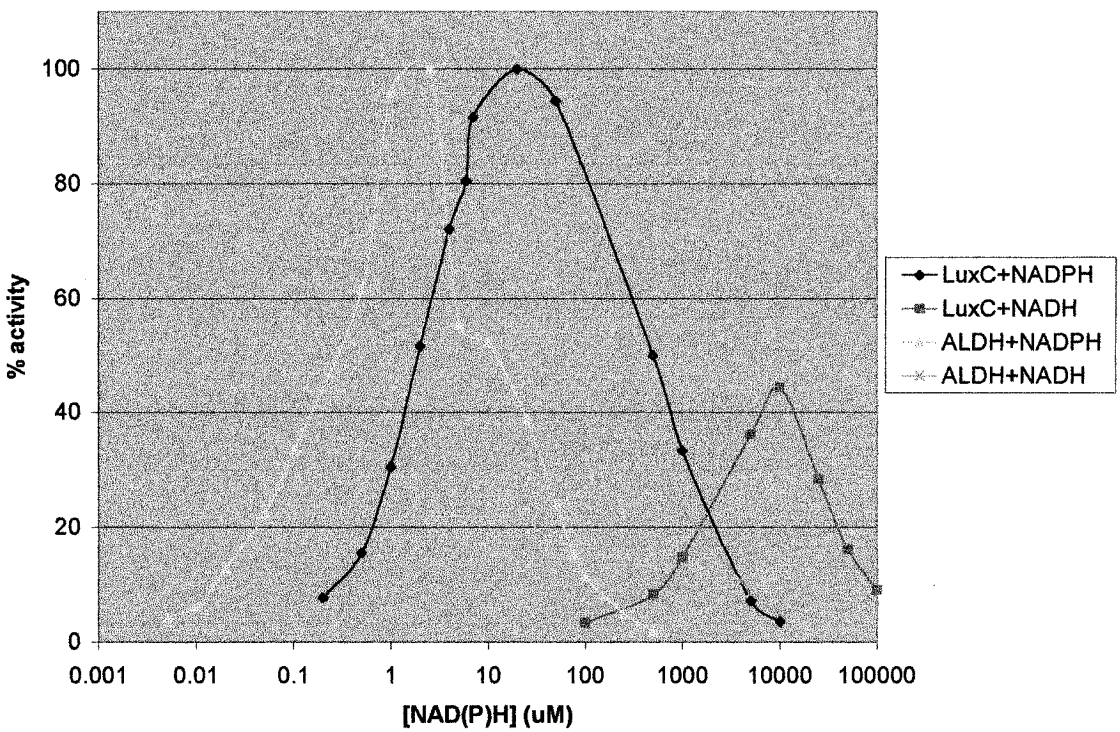
The kinetic parameters of the dehydrogenase and acyl-CoA reductase activities are listed in Table 3.2 for both the native PpLuxC and Vh-ALDH. Though the K_m values and catalytic rates (k_{cat}) with the various cofactors differ between the two enzymes, the differences remain proportional. The catalytic rates of all the PpLuxC-catalyzed reactions are approximately 10% of the same reactions catalyzed by Vh-ALDH, and the K_m values for the nucleotide cofactors are ten to fifty times higher for PpLuxC than for Vh-ALDH, and though they differ from one another, they follow the same trends. Both enzymes show the lowest K_m value for NADPH, followed by NADP and then NADH. These similarities alone are not strong evidence that the two enzymes are related but they create a basis for further comparisons.

Table 3.2- Kinetic parameters of the dehydrogenase and reductase reactions for the native *P. phosphoreum* LuxC and *V. harveyi* aldehyde dehydrogenase. Values were extrapolated using double-reciprocal plots of all data.

Activity	K_m (μM)		k_{cat} (min^{-1})	
	PpLuxC	Vh-ALDH	PpLuxC	Vh-ALDH
Reductase (NADH)	1700 ± 65	35 ± 3.0	4.4 ± 0.4	36 ± 0.9
Reductase (NADPH)	2.5 ± 0.2	0.2 ± 0.1	5.7 ± 1.0	55 ± 4.2
Dehydrogenase (NAD)	-	300 ± 35	-	3500 ± 91
Dehydrogenase (NADP)	61 ± 8	1.9 ± 0.4	53 ± 1.4	565 ± 31

When comparing the rates of the reductase reaction with both enzymes over a range of NAD(P)H concentrations (Figure 3.4), we can see that a maximal activity is reached with NADPH that is higher than with NADH, and this at a lower concentration. From Table 3.2, we can also see that the ratio of k_{cat} of NADH/NADPH is the same for both enzymes and that the ration of K_m of NADH/NADPH only shows a four-fold difference between the two. At high cofactor concentrations (Figure 3.4), we can notice that both enzymes seem to suffer from substrate inhibition. This effect had previously been seen in Vh-ALDH (33), and it was proposed to occur through congestion of the active-site by competition for entry between the cofactor and the activated acid (41). Since both the Vh-ALDH and PpLuxC have been shown to be able to catalyze the same reactions, it is possible that the substrate inhibition that is seen in PpLuxC occurs through a similar congestion, as their active-sites geometries can be proposed to be similar due to substrate requirements.

Figure 3.4- Reductase activities of both the *V. harveyi* aldehyde dehydrogenase and the *P. phosphoreum* LuxC, with changing cofactor concentrations. Values were normalized to the maximum rate observed for each enzyme.


100
80
60
40
20
0
% activity
0.001
0.01
0.1
1
10
100
1000
10000
100000
[NAD(P)H] (uM)
LuxC+NADPH
LuxC+NADH
ALDH+NADPH
ALDH+NADH

Since the similarities between these enzymes continued to grow, and the rates of the dehydrogenase and reductase reactions are measurable for both wild-type enzymes (Table 3.2), analysis was undertaken on residues whose functions have been identified in Vh-ALDH and who can be found to have counterparts in PpLuxC through sequence alignment. These residues were thus mutated in both enzymes and their effects on the kinetic properties of the enzymes were investigated, to see if the changes, as compared to the wild-type values, would be correlated in both enzymes. This would serve to both clarify the catalytic mechanism of PpLuxC and could provide additional evidence toward the mechanistic similarities between PpLuxC and Vh-ALDH.

3.2 – Mutational effects

Through the sequence alignment of PpLuxC and Vh-ALDH, two residues which have been implicated in cofactor binding and affinity in Vh-ALDH (K172 and T175) were found to have counterparts in PpLuxC (K174 and S177, respectively). The K172 residue in the *V. harveyi* enzyme is one that is found in many ALDH's and is involved in nucleotide binding through two interactions: one with the 3'-hydroxyl group of the ribose ring of NAD(P) and the other with the hydroxyl group of the 2'-phosphate moiety of NADP (39). The other residue, T175 (or S177 in PpLuxC), is one which is found in NADP-specific enzymes and is proposed to interact with the NADP cofactor through an additional hydrogen bond with the hydroxyl of the 2'-phosphate group (39). Since both of the enzymes in question are NADP(H) specific, and the proper lysine and serine residues are found in PpLuxC, it was proposed that these residues would interact with the cofactor in much the same way as in Vh-ALDH. Consequently, the lysine residues were mutated to glutamines and the serine/threonine residues were changed to glutamates, and their effects on the kinetic parameters of the enzymes were examined (Table 3.3).

Table 3.3- Acyl-CoA reductase kinetic parameters with NADH and NADPH for *P. phosphoreum* LuxC and *V. harveyi* aldehyde dehydrogenase, as well as their respective mutants.

Enzyme	**Cofactor**		**K_m (μM)**	
LuxC		Wild-type	K174Q	S177E
	NADH	1700 ± 65	11 800 ± 400	1900 ± 82
	NADPH	2.5 ± 0.2	7 200 ± 360	31 ± 5.0
Vh-ALDH		Wild-type	K172Q	T175E
	NADH	35 ± 3.0	600 ± 44	28 ± 2.1
	NADPH	0.2 ± 0.1	500 ± 68	47 ± 6.4

Enzyme	**Cofactor**		**k_{cat} (min^{-1})**	
LuxC		Wild-type	K174Q	S177E
	NADH	4.4 ± 0.4	0.24 ± 0.05	0.11 ± 0.02
	NADPH	5.7 ± 1.0	0.96 ± 0.06	0.16 ± 0.04
Vh-ALDH		Wild-type	K172Q	T175E
	NADH	36 ± 0.9	5.8 ± 0.9	6.3 ± 0.2
	NADPH	55 ± 4.2	4.4 ± 0.3	8.1 ± 1.1

We can see, in Table 3.3, that the apparent affinity for NADH on mutating the potentially equivalent lysine residue of both enzymes decreased, resulting in a 7-fold increase in the K_m of the K174Q mutant of PpLuxC and a 17-fold increase in the K_m of the K172Q mutant of Vh-ALDH. In the presence of NADPH, this effect on K_m was even more pronounced, with the PpLuxC K174Q mutant demonstrating a nearly 3000-fold increase in K_m and the Vh-ALDH K172Q mutant displaying a 2500-fold increase, when comparing to their wild-type K_ms. This effect was expected as the lysine residue interacts with NADPH at two separate points and at only one with NADH. The serine and threonine residues are proposed, however, to interact only with the NADPH, and not the NADH, cofactor. As shown in Table 3.3, the K_m for NADH of the S177E mutant of LuxC and of the T175E mutant of Vh-ALDH were the same as the K_ms for their respective wild-type enzymes, whereas their K_ms for NADPH increased 12 fold for the S177E mutant and nearly 250-fold for the T175E mutant, as compared to those of their respective wild-type enzymes.

The dehydrogenase assay(data not shown) also showed that the T175E mutant of Vh-ALDH had a six-fold lower K_m for NAD and a 1500-fold higher K_m for NADP, which is consistent in its influence on NADP, and not NAD, binding, though a decrease in K_m for NAD may be indicative that another interaction, favorable to NAD, may have come into effect. Since no activity could be detected for the native PpLuxC in the presence of NAD, such a comparison is not readily available, but it was seen that the S177E mutant of PpLuxC was able to catalyze the dehydrogenase reaction in the presence of NAD, possibly owing to a similar improvement that resulted in an decrease

of the K_m for NAD in the Vh-ALDH T175E mutant, and it was also seen that the K_m for NADP increased 20-fold in the S177E mutant, as compared to the wild-type PpLuxC.

The catalytic rates of the reductase reaction were, however, found to be reduced in the mutant enzymes, with their activities being between 3-18% of the wild-type values (Table 3.3). Although these residues are not implicated in the catalytic steps of the mechanism, it is possible that the effect is exerted through a less "ideal" placement of the cofactor in the active-site for catalysis, resulting in higher energies of activation for the reaction and thus causing a reduction in the catalytic rates.

One of the residues which is very highly conserved among aldehyde dehydrogenases is an asparagine residue in motif 1. This residue is proposed to stabilize the oxyanion intermediate that is formed during the nucleophilic attack of the active-site cysteine residue's thiol group on the carbonyl carbon of the acyl moiety (thiohemiacetal formation) (56). An asparagine residue has been identified in PpLuxC at position 152, and is proposed to act in a similar capacity to that observed for ALDH's. Accordingly, this residue was mutated to an aspartate (N152D) and its effects on the kinetics of the reaction were determined, as well as its effects on acylation due to its role in stabilizing acyl-enzyme formation.

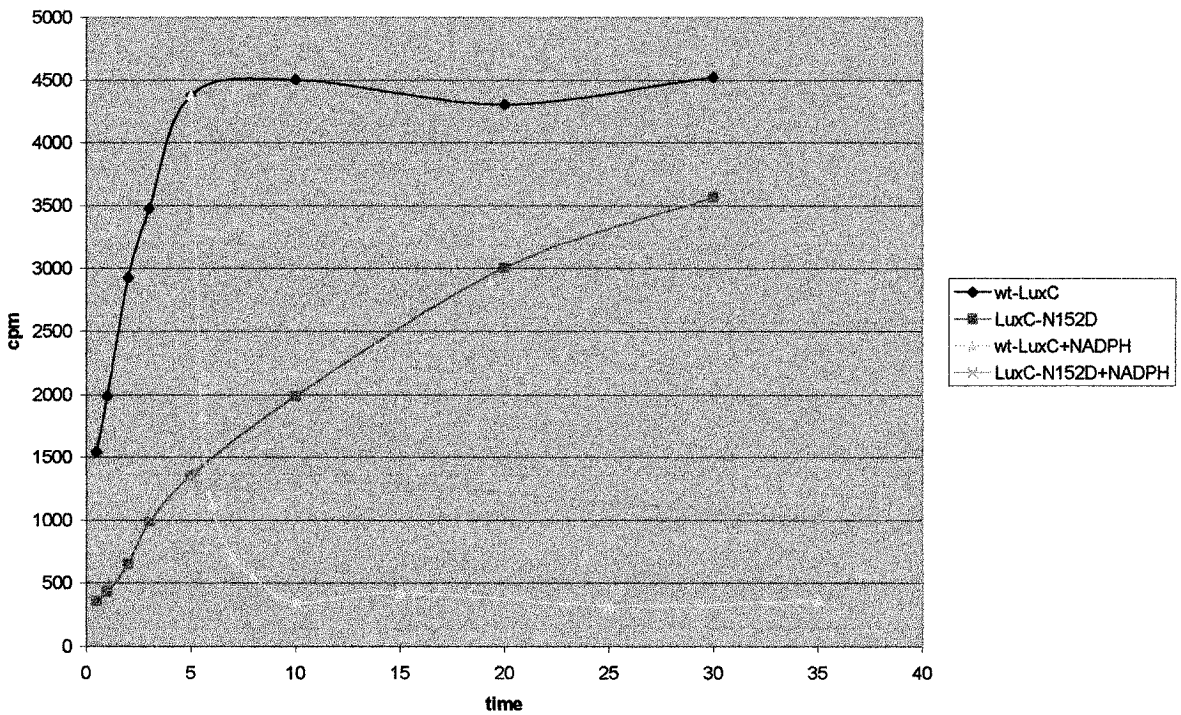
The kinetic results showed that the reductase activity (k_{cat}) of the N152D mutant was reduced to 1% of the native enzyme activity, with both the NADH and NADPH cofactors, and that the apparent K_m of this mutant for NADH was increased almost 20-

fold, and almost 40-fold for NADPH. While N152 is not known to interact with the nucleotide cofactor, it is possible that the introduction of a charge in the N152D mutant would affect cofactor binding since it is located in the vicinity of the active site. Also, the drop in activity that is seen, which occurs irrespective of the cofactor, correlates well with its proposed role in intermediate stabilization.

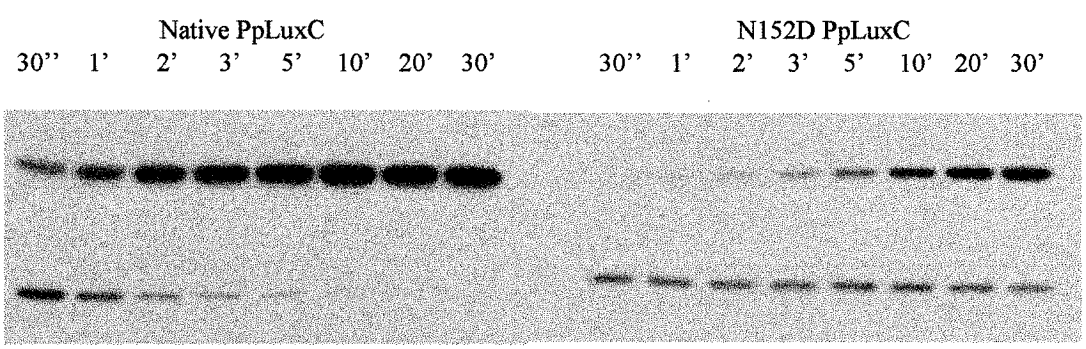
The asparagine residue's role was also investigated through acylation/deacylation assays, since it is implicated in acyl-enzyme formation. When comparing the results from acylation of the native LuxC and its N152D mutant, we can see that the overall level of acyl-enzyme that is observable is greater in the native enzyme, and that the rate of formation of the acylated intermediate is lower in the asparagine mutant enzyme. Moreover, when NADPH was added, which causes deacylation, the level of acylation remaining was found to be higher in the mutant form of the enzyme, and that it decreased at a very low rate, showing that reduction in the presence of NADPH occurred more slowly.

Figure 3.5- Panel A: Acylation and de-acylation profiles of the native *P. phosphoreum* LuxC and its N152D mutant. The blue and magenta lines indicate acylation of the native and N152D mutant of LuxC, respectively, and the yellow and cyan lines indicate de-acylation, again of the native and N152D mutant of LuxC, respectively, upon addition of NADPH. Panel B: Autoradiograph of the [^{3}H]myristoyl-CoA-labeled protein during acylation, with the top lines being acylation (native LuxC upper left and N152D LuxC upper right) and the lower lines showing de-acylation with NADPH (native LuxC lower left and N152D LuxC lower right) over a time period of 0.5 to 30 minutes.

A


5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
cpm
0
5
10
15
20
25
30
35
40
time
wt-LuxC
LuxC-N152D
wt-LuxC+NADPH
LuxC-N152D+NADPH

B


Native PpLuxC
30'' 1' 2' 3' 5' 10' 20' 30'
N152D PpLuxC
30'' 1' 2' 3' 5' 10' 20' 30'

A residue of marked importance in Vh-ALDH, the glutamate residue in motif 5 (residue 258 in Figure 3.2), was the final object of study during this investigation. It has been previously shown that the dehydrogenase activity of an E253A mutant is nearly obliterated, with a decrease of nearly 25000-fold in turnover rate at high substrate concentrations (40). This residue has been implicated in activation of the cysteine residue in the active site and/or coordination of a water molecule for an attack on the acyl-enzyme intermediate, causing fatty acid release (the enzyme's thioesterase activity). Since an equivalent residue is missing in PpLuxC and the thioesterase activity is only required for the dehydrogenase mechanism, the effects of mutating this residue on the reductase function of Vh-ALDH were investigated.

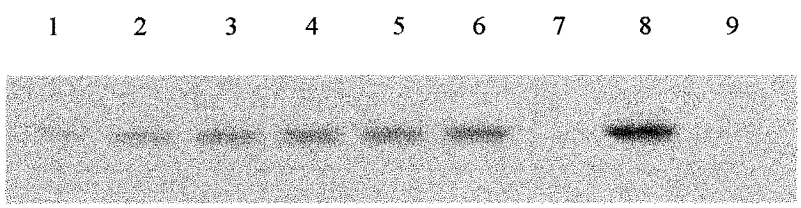
In contrast to the drastic effects on dehydrogenase activity, the effects of mutating E253 on reductase activity were very small (Table 3.4). At neutral pH, pH= 7, the E253A mutant of Vh-ALDH retained 15% of the wild-type reductase activity. An increase of the pH, to pH=8, caused an increase in the reductase activity of the mutant while the wild-type reductase activity was significantly decreased. This led to the E253A mutant having nearly twice the activity of the wild-type enzyme at pH= 8. This result clearly shows that the glutamate residue, which is essential for dehydrogenase activity, is not required for reductase activity. Consequently, this would be consistent with the primary role of E253 in Vh-ALDH being to increase the nucleophilicity of water, a function needed only for dehydrogenase activity and not reductase activity.

Previous attempts to detect an acylated intermediate of Vh-ALDH upon reaction with [^{3}H]myristoyl-CoA have not been successful, presumably due to the high thioesterase activity. Since the E253 residue appears to be primarily responsible for thioesterase activity, it would be predicted that this mutant could now be acylated. Figure 3.6 indeed shows that the E253A mutant can be readily acylated, indicating that the thioesterase activity is indeed blocked in this mutant and that the cysteine residue can still function as a nucleophile. Moreover, addition of NADPH but not β-mercaptoethanol caused the rapid removal of the myristoyl group, consistent with the retention of reductase activity. This result on deacylation of the E253A mutant, coupled with its acylation and kinetic properties, indicates that the role of the E253 residue may almost exclusively lie in increasing the nucleophilic character of water and at most indirectly affecting the nucleophilicity of the active site cysteine residue.

Table 3.4- Turnover rate (k_{cat}) of the reductase activity of the native and E253A mutant of Vh-ALDH at pH= 7 and pH= 8.

pH	Enzyme	k_{cat} (min^{-1})
7		
	Vh-ALDH wild-type	55 ± 4.2
	Vh-ALDH E253A	8.7 ± 1.1
8		
	Vh-ALDH wild-type	8.8 ± 0.6
	Vh-ALDH E253A	14.4 ± 1.4

Figure 3.6- Autoradiograph of the [^{3}H]myristoyl-CoA-labeled Vh-ALDH E253A protein with time. Lanes 1-6 show incubation of the E253A mutant for 0.5, 1, 3, 5, 10 and 15 minutes, respectively, with 5 μM [^{3}H]myristoyl-CoA. Lane 7 shows an additional 5 minutes of incubation with NADPH from the 10 minute mark; lane 8 shows an additional 5 minutes of incubation with 100mM β-mercaptoethanol from the 10 minute mark; lane 9 shows an additional 5 minutes of incubation with both NADPH and β-mercaptoethanol from the 10 minute mark.


1
2
3
4
5
6
7
8
9

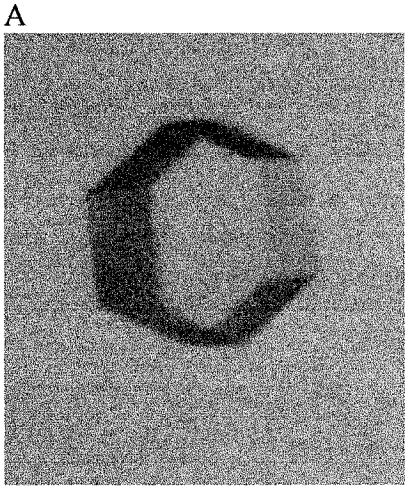
Chapter 4 - Crystallization and Preliminary X-Ray Diffraction Studies

4.1-Crystallization

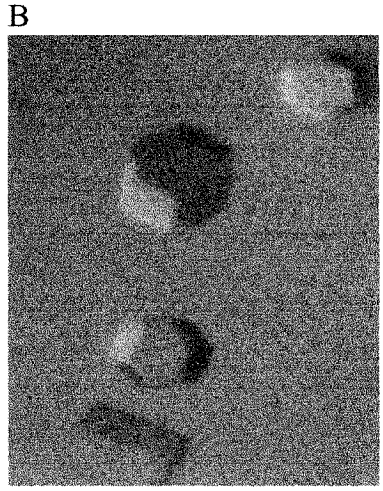
The native and selenomethionyl derivative of *P. phosphoreum* acyl-CoA reductase were purified, as mentioned in Chapter 2, and were crystallized in a solution containing NADPH to obtain crystals of the binary enzyme·NADPH complex. Clear crystals were obtained that had a hexagonal morphology (Figure 4.1) with typical dimensions of 0.1 x 0.1 x 0.08 mm and that diffracted to a minimum d-spacing of 2.7 Å. Indexing of the diffraction pattern showed that they belonged to the trigonal spacegroup $P3_221$, with cell dimensions of a=b=95.4 Å, c=202.1 Å, $\alpha=\beta=90°$, $\gamma=120°$. Using these cell dimensions and two monomers of 58 kDa per asymmetric unit, a Matthews coefficient (42) of 2.29 $Å^3$/Da and solvent content of 46% were calculated.

Figure 4.1- Crystals of native (panel A) and selenomethionyl (panel B) LuxC, after 5 days of incubation at 298 K.

A



B



4.2- Other Crystallization Experiments

Prior to obtaining the hexagonal crystals that were used in the diffraction experiment, two other crystals forms were obtained from the same mother liquor. The first of these crystals forms had a tetragonal bypiramidal morphology and had a very clear appearance, but unfortunately did not diffract beyond 4 Å. The second form that was observed resembled hollow-rods with a solid base. Though these crystals diffracted well, with a maximum resolution of 2.2 Å, they were found to be merohedrally twinned, with an approximate twinning fraction of 0.46, and all attempts to detwin the data were unsuccessful. Interestingly, the type of crystal obtained depended on the incubation temperature during crystal growth. At 277 K, the bypiramidal and hollow-rod crystals predominated; as the incubation temperature was increased, the hexagonal crystals would form, up to 298 K where they were the only crystals obtained. At 303 K, no crystals were seen, replaced instead by a precipitate of a crystalline appearance, like freshly fallen snow, which was of no use to structural determination.

4.3-Initial Phase Determination

Once we had obtained crystals of PpLuxC and found that they were able to diffract to a useful resolution, data was collected on both the native and the selenomethionyl derivative of the protein. In order to collect a sufficient amount of data, a dataset covering a 180° oscillation range, in increments of one degree, was collected on

the native crystal. Three datasets covering a 200° oscillation range each were collected at the peak, inflection and remote wavelengths for selenium, again in increments of one degree and this time using an inverse beam geometry. The data statistics for all datasets are presented in Table 4.1. Since they all showed a high degree of completeness and redundancy and since initial diagnostics revealed no signs of twinning, they were deemed satisfactory and analysis was continued.

Table 4.1- Data collection statistics from the native *P. phosphoreum* LuxC crystal dataset and the three datasets collected for the selenomethionyl-derivative peak, inflection and remote wavelengths. I/σI is the signal-to-noise ratio and R_{merge} is an indicator of agreement between datasets.

	Native	Peak	Inflection	Remote
Resolution	2.67 Å	3.0 Å	3.0 Å	3.0 Å
Space Group	$P3_221$	$P3_221$	$P3_221$	$P3_221$
Wavelength of radiation	1.0800 Å	0.9793 Å	0.9795 Å	0.9184 Å
Number of reflections	147425	262 892	258433	261678
Number of unique reflections	29907	40950	41 121	40800
R_{merge} (%)	4.8 (68.9)	8.1 (50.8)	7.0 (42.7)	6.0 (55.1)
Redundancy	4.9	6.4	6.4	6.5
Completeness*	95.9 % (76.5%)	99.9 % (100.0 %)	99.8 % (99.9%)	99.9 % (99.7%)
I/σI *	24.8 (1.2)	15.7 (3.1)	17.9 (3.4)	26.5 (3.0)

* Data in parentheses indicates result for highest resolution shell

After all data was processed and scaled (43,44), the first step in structural determination involves finding a solution to the phase portion of the structure factor equation, one which will allow for the calculation of an interpretable electron density map. The method employed for this determination in PpLuxC was that of multiple-wavelength anomalous diffraction, or MAD, where the anomalous contributions of the selenium atoms to the diffraction pattern are used to get an initial phase approximation (46, 47). Using a direct-methods approach and the program SnB (48), or Shake'n'Bake as it is affectionately known, a solution which defined the positions of 20 of the 22 expected selenium atoms in the asymmetric unit was found using either the peak or inflection wavelength data. This positional information contained phase information which was then applied to all non-hydrogen atoms, and an initial electron density map was calculated that showed protein/solvent boundaries (Figure 4.2) and thus an approximately correct solution. The positions of the *pseudo* heavy-atoms were then refined with a maximum-likelihood based method in MLPHARE (45), and these refined phases produced a map with density that revealed not only clear protein/solvent boundaries, but also clear elements of secondary structure in the correct handedness (Figure 4.3), showing that the initial solution was correct and that it was improving, as well as confirming the identity of the point group. The density was then improved through histogram matching and solvent flattening, both density modification techniques (45), and a map of sufficient quality to start model building was calculated.

Figure 4.2- Low resolution electron density map of LuxC, using the initial phase information contained in the 20 selenium atoms located.

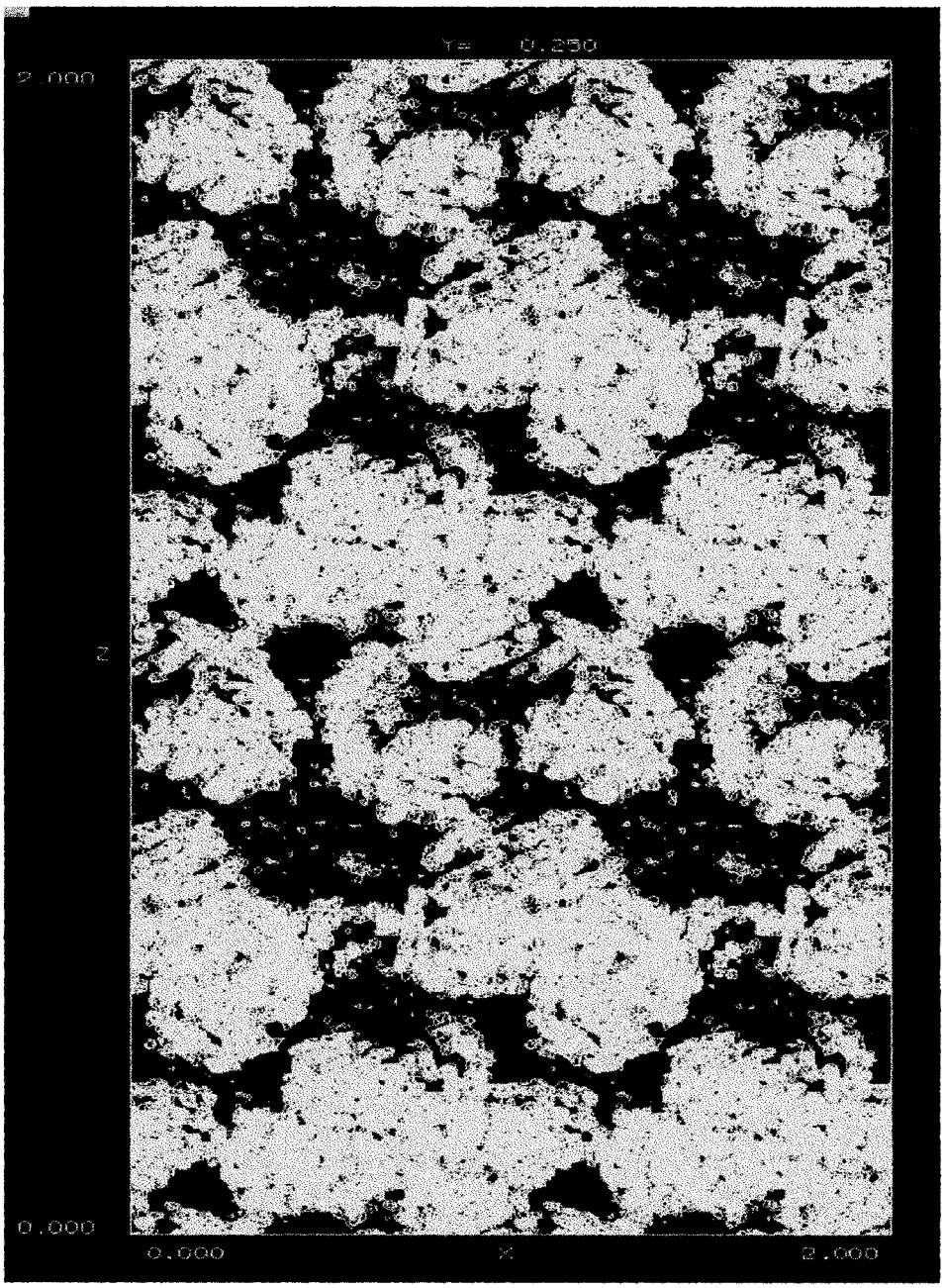
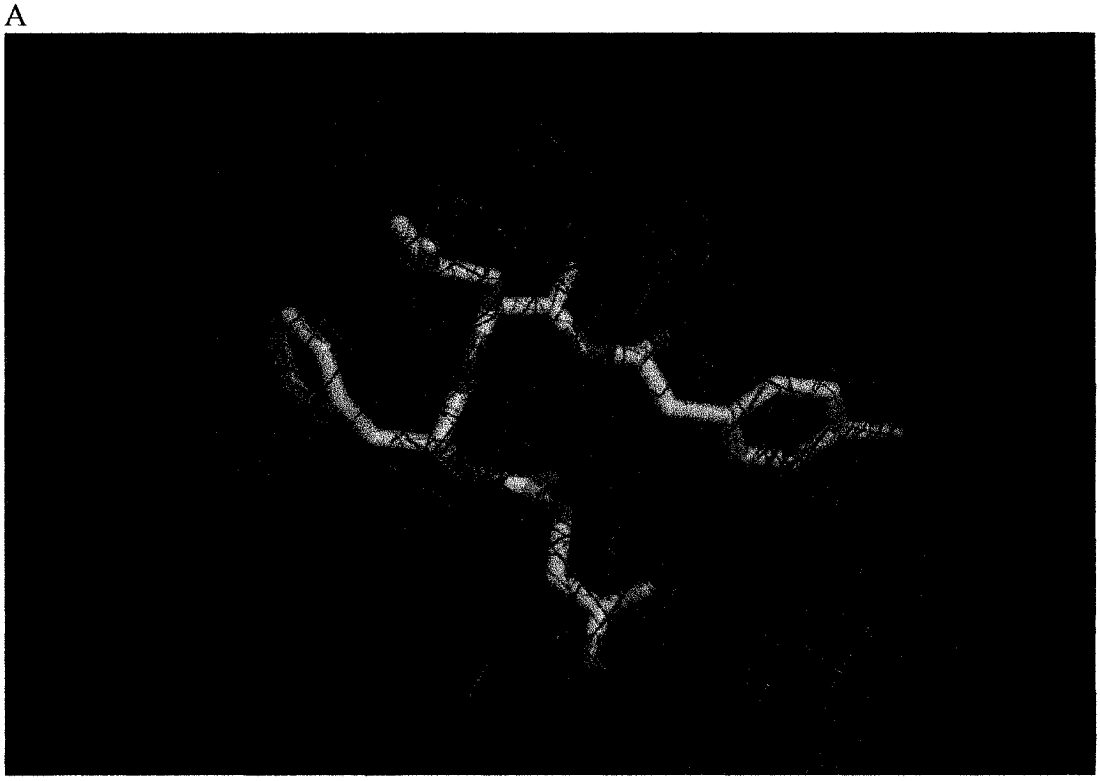
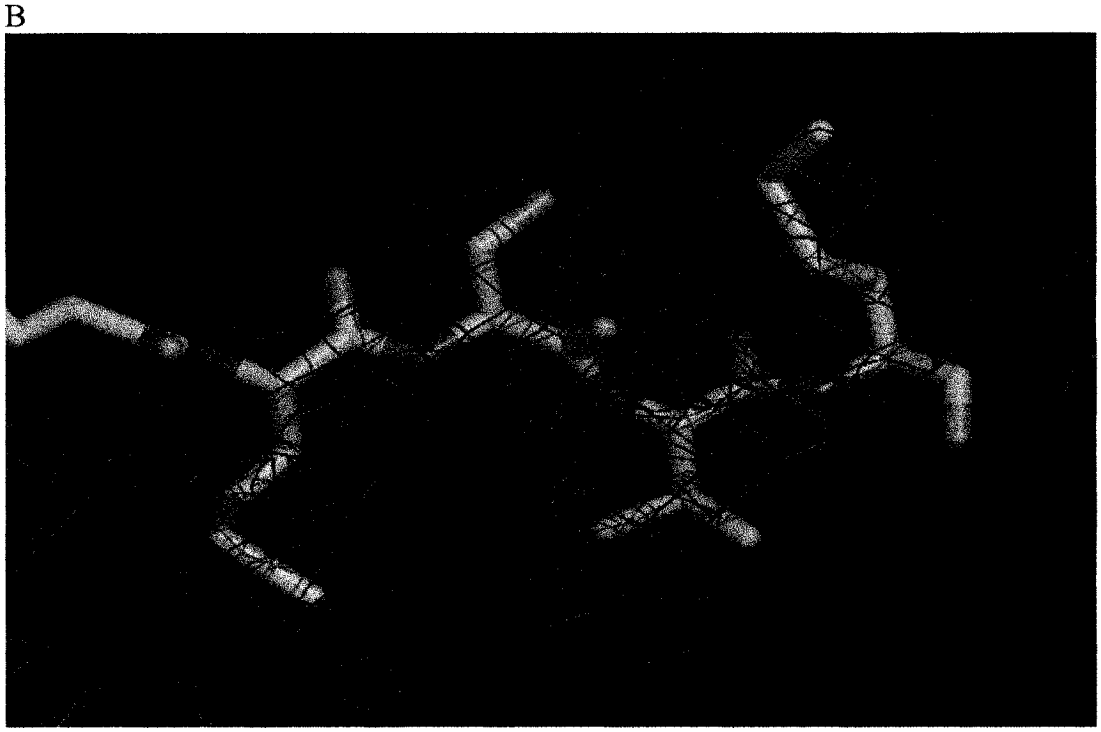

Y= 0.250
2.000
Z
0.000
0.000
X
2.000

Figure 4.3- Electron density map showing, in panel A, density for an alpha-helix fragment, containing a tetrapeptide NFLY, and in panel B, density for a beta-pleated sheet fragment, containing a tetrapeptide MSVM, where the atoms in green are carbons, those in red are oxygens, those in blue are nitrogens and those in orange are sulphurs.

A



B



4.4- Model Building

Using the electron density map that was generated after density modification, and the program MAPMAN of the RAVE suite of software (49), model construction was started. MAPMAN was first used to create a bones map of the electron density, or a map where the density is collapsed to its center and appears as a skeleton of the original map. By then overlaying the density on the bones, a structure begins to emerge, but it is one with no structural or chemical sense as it is not continuous and has many branch points. This model is visualized using "O" (50), and, with the electron density map, a trace of the α-carbon backbone is manually done (Figure 4.4). Since the selenium positions correspond to methionine residues, these were initially used to anchor the amino acid sequence and allow the threading of the primary structure through the backbone trace. Large aromatic side-chains also have a characteristic appearance in the electron density maps, and were used as anchors in the domains between methionine residues to ensure no "slipping", or misplacement, of the primary sequence. Then using conserved parameters of hydrogen bonding distances, bond lengths, bond and dihedral angles and side chain orientations present in all protein structures, a model containing 375 of the 479 amino acids was built.

Moreover, since the selenium atoms present in the asymmetric unit were found and since only two protomers are contained in the asymmetric unit, a rotation matrix could be approximated between the two polypeptide chains by manually assigning the

selenium atoms to one of the two chains (Figure 4.5). This matrix, along with the operators that are specific to the space group, allowed the calculation of both the dimeric and tetrameric structures, as seen in Figures 4.6 and 4.7, and surface representations (space-filling models) in Figure 4.8.

4.5- Structural Refinement

Structural refinement of the preliminary model was undertaken using CNS (51) to correct for dihedral angles and improper protein geometries. Refinement initially showed very poor statistics, using either the monomer or the dimer, with R and Rfree values of nearly 46%. Closer investigation of the dimeric model with the electron density revealed that a simple rotation of the model was inadequate to allow refinement. The domain that links the catalytic and nucleotide binding domains positions the two domains in a slightly different orientation in each monomer, with position varying by as much as 3.0 Å between the monomers. This difference proved to be too much for the refinement programs to correct, as simulated annealing resulted in warped geometries that had moved out of regions of density. To correct this, a manual rebuild of the second monomer would be necessary, but this was not permitted by restraints on time. Structural elements of the preliminary model, though not refined, are, overall, correct and those that are of significant importance will be discussed next.

Figure 4.4- Trace of the alpha-carbon backbone of a single polypeptide chain of PpLuxC.

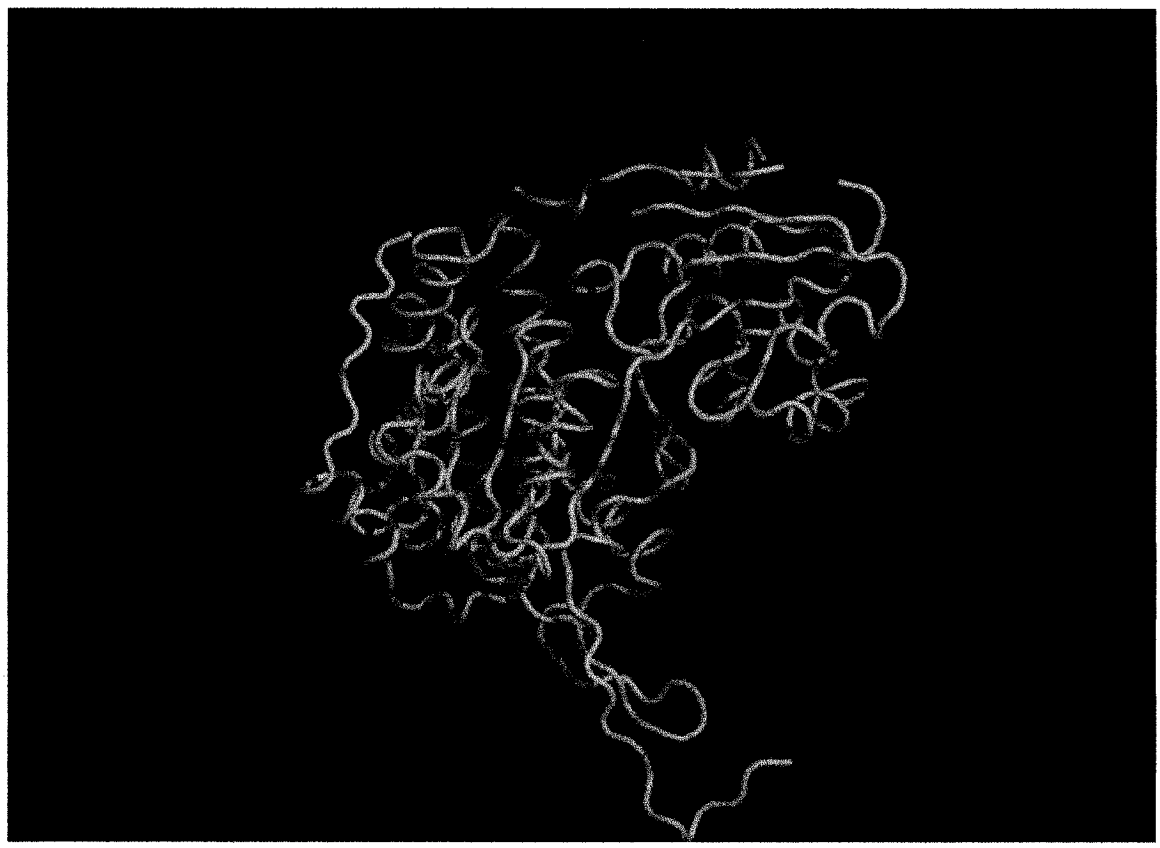


Figure 4.5- Positions of the 20 selenium atoms contained in a single asymmetric unit, with the ones belonging to different polypeptide chains colored in red and cyan, and with a polypeptide chain positioned over one set of these selenium atoms.

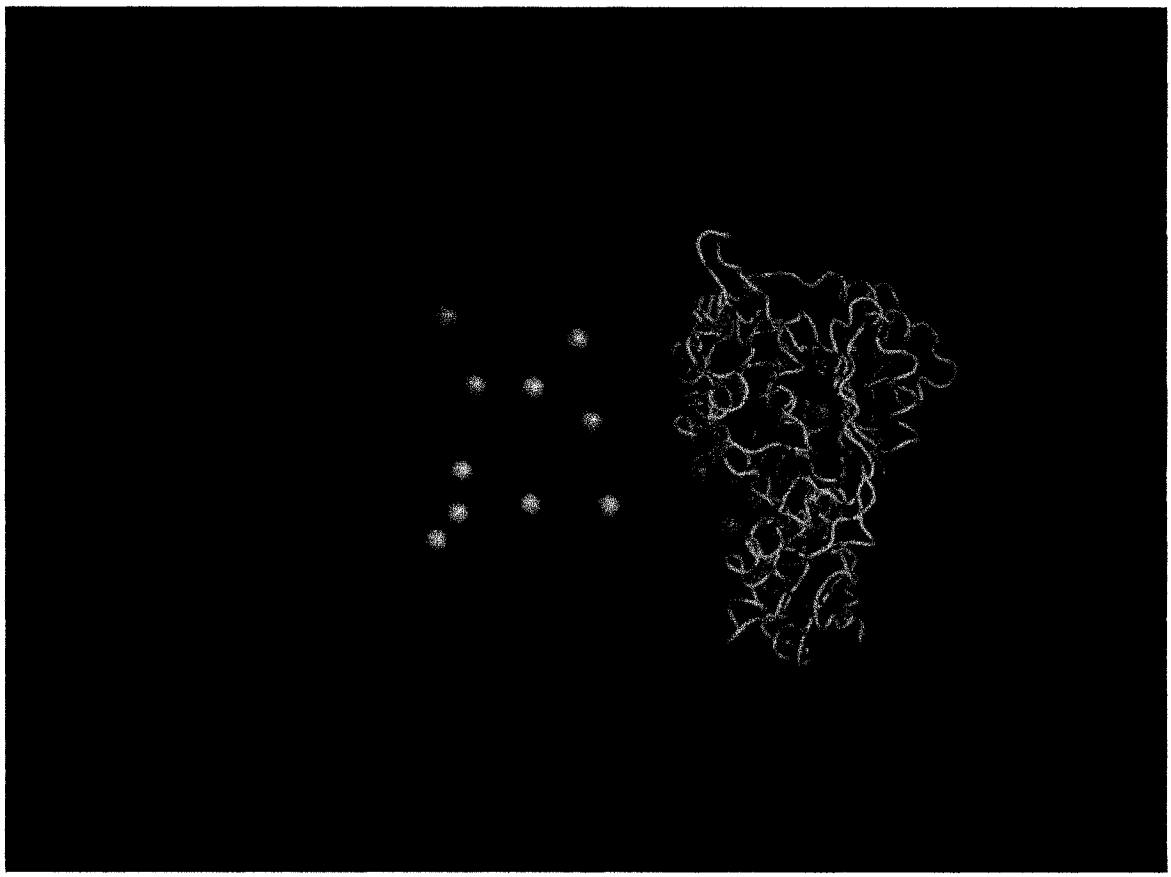


Figure 4.6- Trace of the alpha-carbon backbone for the symmetry-generated tetramer of PpLuxC.

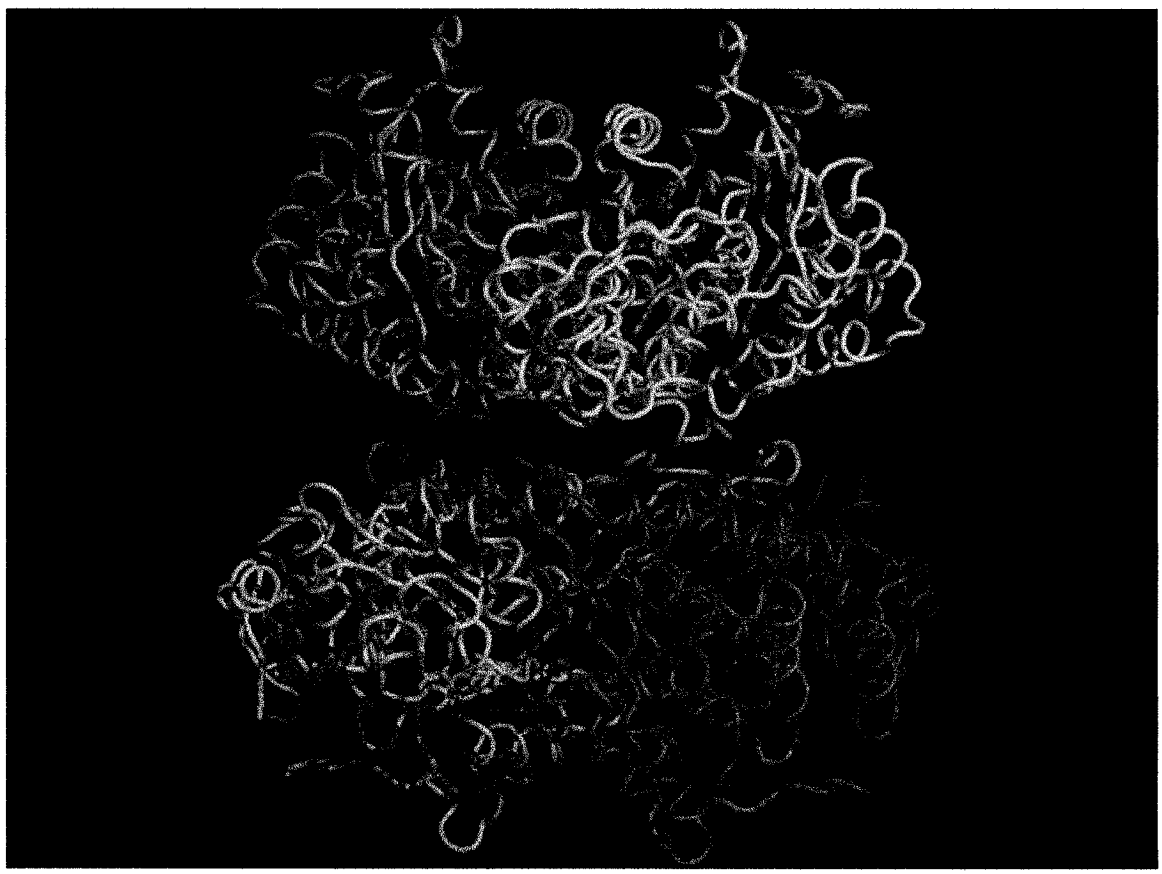


Figure 4.7- Trace of the alpha-carbon backbone for non-crystallographically symmetry-related dimer of PpLuxC.

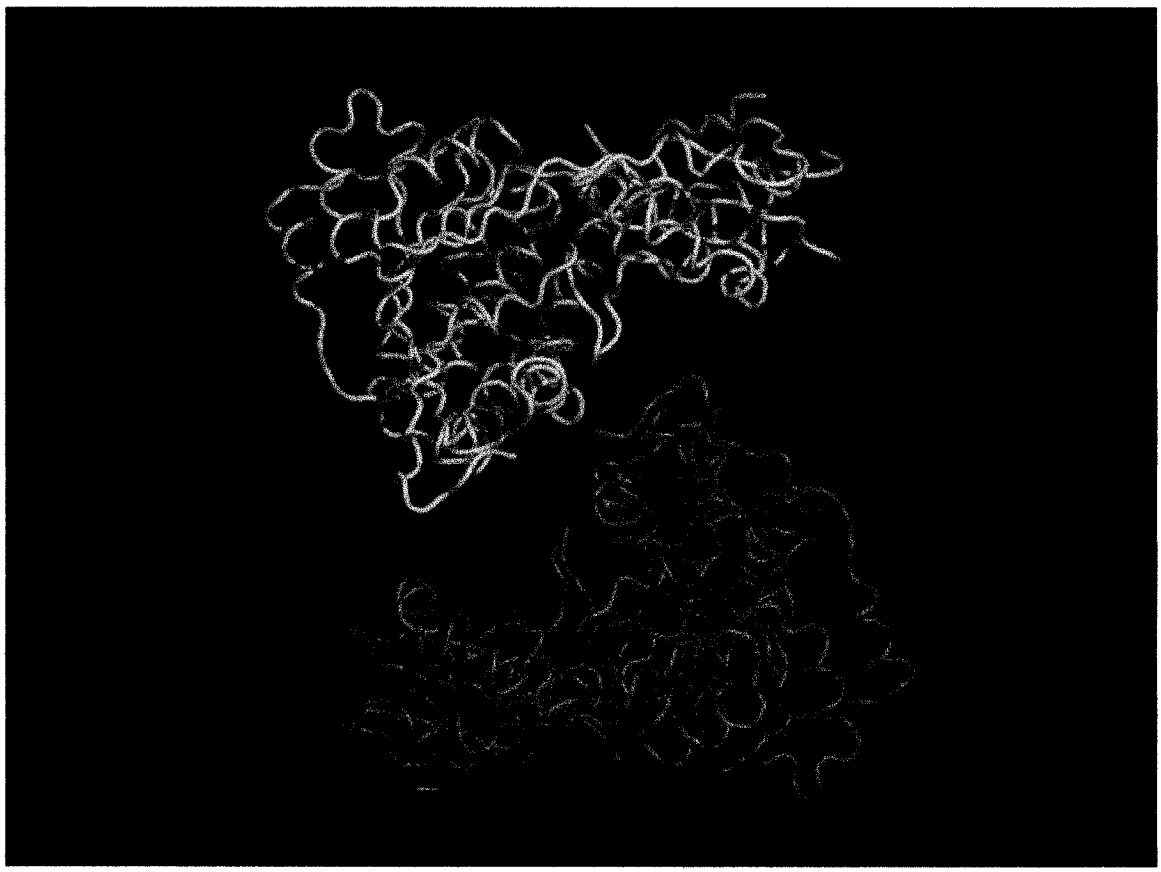
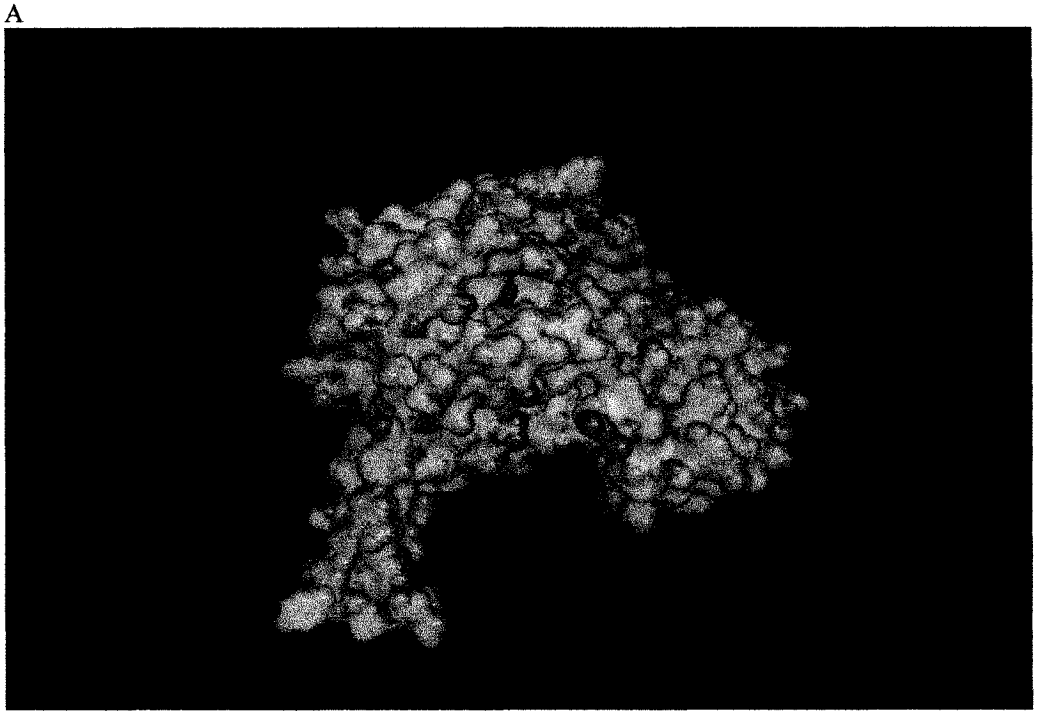
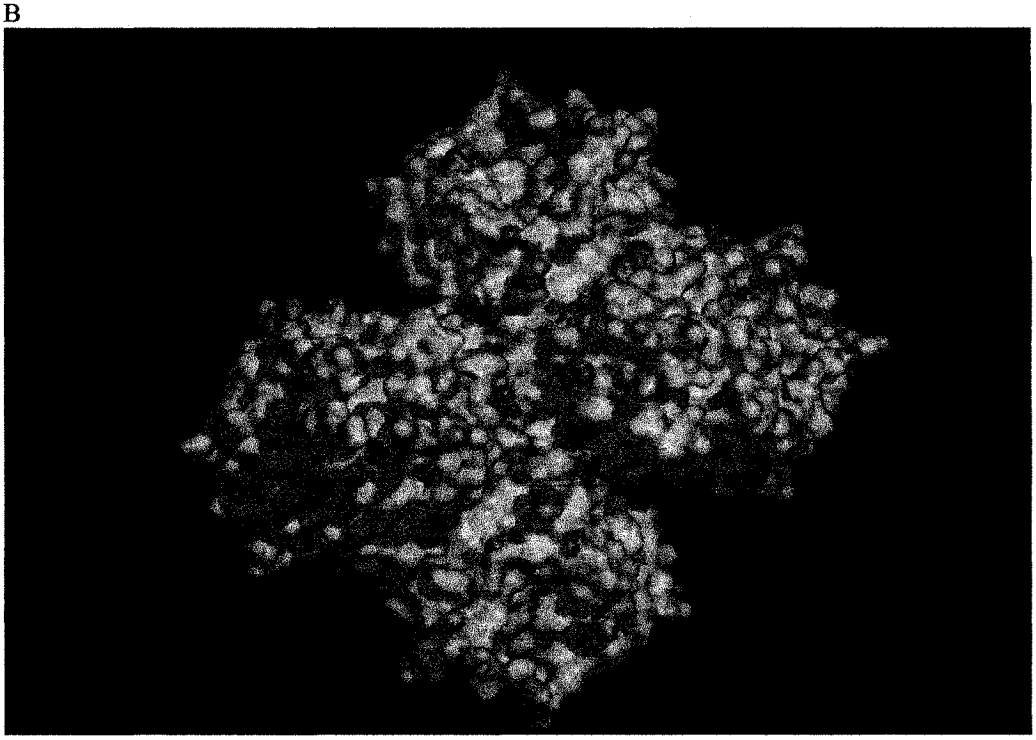


Figure 4.8- Surface representation of one polypeptide chain of the PpLuxC model in panel A and a surface representation of the tetrameric model in panel B.

A



B



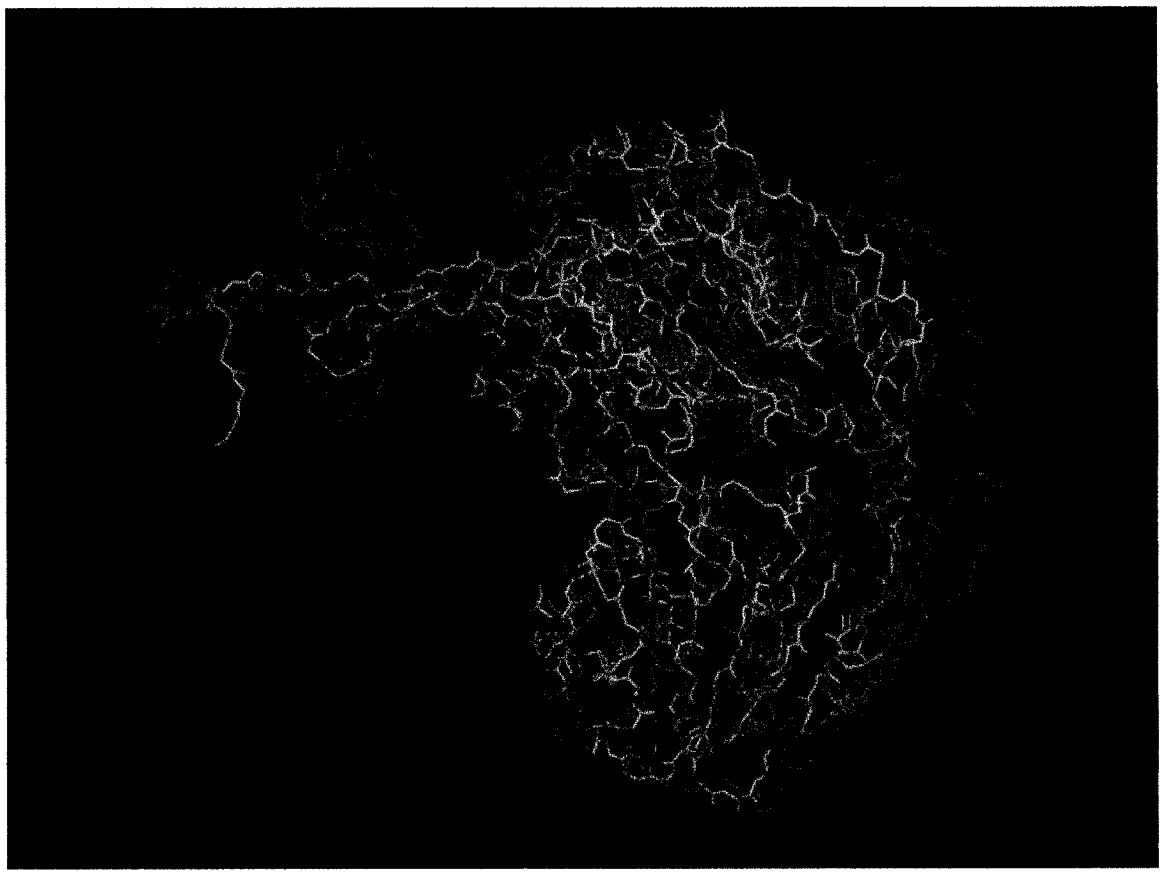
4.6 – Preliminary structural analysis

The structure of the *P. phosphoreum* acyl-CoA reductase has been solved, and though it is not in its final refined form, a number of structural characteristics have emerged that show a very strong similarity to the structures of aldehyde dehydrogenases, as shown in Figure 4.9, where the structure of the Vh-ALDH is superimposed with that determined for PpLuxC. The preliminary structure contains 375 of the 479 amino acids, and shows three principal domains. A nucleotide binding domain (NBD) has been found to extend from residues 29 to 127, 140 to 253 and 429 to 443. The first few residues of the amino terminus, in the model, are relatively disordered and show no secondary structure, but these lead into a series of four α-helices that surround the NBD. Residues 128 to 139 and 444 to 456 make up the dimerization, or oligomerization, domain (OD) which is a 3-stranded antiparallel β-sheet. The third and final domain is the catalytic domain (CD) which is found to extend from amino acid 254 to 429.

Closer investigation has revealed a number of interesting features. The nucleotide binding domain contains a 5-stranded parallel β-sheet, which corresponds to the Rossman fold, or the dinucleotide binding motif. There is some residual electron density within the vicinity of the proposed Rossman fold that is unoccupied, and since the enzyme was crystallized in the presence of NADPH, we have proposed that this density corresponds to part of the cofactor, more precisely the adenosine moiety of NADPH. This density lies at an angle to the binding motif, as is seen in the crystal structure of the binary NADP·Vh-ALDH complex (39), and numerous possible points of contact can be

proposed. The side chain of serine 177 is within hydrogen bonding distance to the proposed location of the hydroxyl group of the 2'-phosphate of NADPH, and the side chain of lysine 174 is in a position to form contacts with both the hydroxyl group of the 2'-phosphate moiety of NADPH and 3'-hydroxyl group of the adenosyl ribose ring of NADPH. A histidine residue, His 210, has also been noticed to be in a curious position, lying co-planar with the adenosine ring and being within the distance for some hydrogen-bonding or electrostatic interactions, and possibly acting to restrain the movement of the adenosine ring and correctly position it for hydride transfer to occur. Main-chain contacts are also possible as two glycine residues, which lie in the "fingerprint" region of the dinucleotide binding motif, Gly 231 and 232, are positioned in a tight turn at the end of the β_4 strand of the binding motif and whose main-chain nitrogens are within hydrogen-bonding distance to the adenosine ring. The main-chain atoms of the β_1 strand of the fold are also within hydrogen-bonding distance to the adenosine moiety, at residues 149, 150 and 152. These findings, along with the fact that the proposed position of the adenosine ring is approximately 6 Å from the active-site cysteine residue, have aided in the proposal that this density is in fact part of the NADPH cofactor. The nicotinamide portion is not visible, which has been noticed in a number of other NADP(H)-dependent enzyme structures (57), and is proposed to be caused by flexibility in the active-site which allows multiple conformations of the cofactor to be present, and thus has the effect of diminishing the amount of electron density which can be observed.

Figure 4.9 – Alignment of the structure of Vh-ALDH, in red, obtained from the Protein Data Bank (PDB accession number 1EZ0) and of the structure determined for PpLuxC, in yellow.



The active-site cysteine residue, which has been shown to form the acylated-enzyme intermediate, is located in a hydrophobic pocket which isolates the active site from the rest of the enzyme. Four phenylalanine residues, F252, F280, F287 and F426, as well as tryptophan 230, are located within 5 to 8.6 Å from the cysteine, and three of these, F280, F287 and F426, appear to form a "wall" with the three aromatic rings forming a flat hydrophobic patch on one side of the active site.

The cysteine residue's sulfhydryl group must be in its thiolate form to be able to cause acyl-group transfer from the acyl-LuxE complex to LuxC, and two residues have been found to be within hydrogen bonding distance to aid in the base-activation of the sulfhydryl group. The carbonyl oxygens of Asp433 are located within 4 Å of the cysteine group, and the main chain oxygen of Gly253 is only 4.5 Å away, putting them all in a position to catalyze the deprotonation of the cysteine's sulfhydryl group. Since aldehyde dehydrogenases have shown that there are significantly higher temperature factors for the cysteinyl residues in their active sites, which is most likely due to an added conformational flexibility, and since there is a large resemblance between the PpLuxC and various aldehyde dehydrogenases, it is possible that this residue is also very flexible, and that the distances to the aspartate and glycine residues could change, placing it closer to the latter residues or placing it within distance to some other residues which are not apparent. We must consider, though, that this may not be the case, and since temperature factor refinement was not carried out, this point is still unclear.

The oligomerization domain of LuxC is composed of, so far, 3 β-strands, and the final 20 residues of the carboxyl-terminus of the protein are expected to form a fourth and final strand, as has been seen in both the dimeric and tetrameric structures of aldehyde dehydrogenases. When the dimeric structure of the enzyme is generated, through rotation symmetries proper to the space-group and the rotation matrix initially established with the selenium atoms, it can be seen that the oligomerization domain β-sheet acts to extend the sheet of the catalytic domain, to form a 9- or 10- stranded extended β-sheet. The ambiguity in the number of final strands is caused by poor density in a region of the CD, which has prevented complete modeling of this domain, though a 6^{th} and final strand is proposed to exist. This would promote a strong binding and stabilization of the dimeric structure. The contacts that would form and stabilize the tetrameric structure have not yet been defined, as the model is incomplete and unrefined, but it is known to exist and to be active in this form, and structural refinement should help in identifying these interactions.

Chapter 5 – Discussion

Sequence alignment of the *P. phosphoreum* LuxC with the *V. harveyi* aldehyde dehydrogenase and a consensus sequence for aldehyde dehydrogenases has shown that despite a low level of overall identity (<15%) certain important elements are conserved. The substrate inhibition and nucleotide specificity for both the PpLuxC and Vh-ALDH enzymes serves to indicate that their active-site geometries may be very similar. This conclusion can be further supported by the fact that they are able to catalyze the same, or very similar, reactions using the same substrates. Two residues known to be critical for nucleotide binding in Vh-ALDH (38,39) were found to be present in PpLuxC at the same location in the primary sequence, and mutations carried out on these residues, K174 and S177 in PpLuxC, and K172 and T175 in Vh-ALDH, resulted in changes in catalytic activity and cofactor affinity of similar magnitude and in the same direction for both enzymes. Additionally, an asparagine residue in PpLuxC, Asn152, matched the position of the asparagine residue in ALDHs which is proposed to stabilize thiohemiacetal formation (56). Mutation of this residue in PpLuxC had the effect of reducing the enzymes' activity, an effect which is seen when this residue is mutated in ALDHs. Moreover, acylation studies of this mutant showed that both the rate of acylation and deacylation were decreased, confirming that its effect on activity was exerted through disruption of a stabilizing interaction (e.g. in thiohemiacetal formation) involved in (de)acylation and showing that this residue accomplishes a similar function to its counterpart in ALDHs.

A glutamic acid residue, critical to the dehydrogenase reaction in ALDH's, but absent in PpLuxC, was also investigated. Mutation of E253 to alanine in *V. harveyi* ALDH causes a 25000-fold decrease in the enzyme's dehydrogenase activity (40). In contrast, the reductase activity of the E253A mutant of Vh-ALDH had 15% of the reductase activity of the wild-type enzyme under standard assay conditions (pH= 7). Moreover, increasing the pH of the assay to pH=8 resulted in the reductase activity of the E253A mutant being higher than that of the wild-type enzyme, while the dehydrogenase activity of the mutant could still not be detected. This implicates the glutamate residue directly in activation of a water molecule to increase its nucleophilic character and not in the activation of the cysteine residue. The absence of this residue in PpLuxC can explain why this enzyme preferentially catalyzes the reductase and not the dehydrogenase reaction.

The structural characterization of PpLuxC provided further evidence towards the similarities that may exist with ALDHs. The overall structure of the preliminary model of PpLuxC is very similar to that of Vh-ALDH, showing the same patterns of connectivity and arrangement of α-helices and β-strands, and showing that the overall fold and arrangement of different domains is very similar. The oligomerization properties of PpLuxC are known to be similar to those of ALDH's (36), and structural analysis has shown that the β-sheet extension responsible for the strong dimerization seen in Vh-ALDH is present in PpLuxC. Also, evidence was found in the PpLuxC electron density map that allowed for partial placement of the cofactor in the structure, which showed that the binding mode of the adenosine portion of NADP is similar in PpLuxC and Vh-ALDH

(39), and is further supported by the similar positioning of residues involved in cofactor binding in both enzymes, residues whose roles were evaluated through the kinetic studies mentioned earlier.

When these pieces of information are put together, they strongly indicate that LuxC is part of the aldehyde dehydrogenase extended family, and that it may be the first member of this family to preferentially catalyze the reduction of fatty acids rather than the oxidation of fatty aldehydes. Since mutation of the E253 residue in Vh-ALDH caused a loss of dehydrogenase activity without a comparable loss in reductase activity, and this residue is absent from PpLuxC, it appears that this residue may be critical in determining the direction of the reaction in the ALDH family.

References

1- Harvey, E.N. (1957) A History of Bioluminescence, American Philosophical Society, Philadelphia.

2- Boyle, R (1668) Philosophical transactions of the Royal Society of London, volume 2, p381.

3- Wilson, T. and Hastings, J.W. (1998) Bioluminescence, Annu. Rev. Cell Dev. Biol., no14, pp197-230.

4- Ghiradella, H (1977) Fine structure of the tracheoles of the lantern of a photurid firefly, J.Morphol., no153, pp187-204.

5- Hastings, J.W. and Morin, J.G. (1991) Bioluminescence, neural and integrative animal physiology, Wiley Interscience, New York, NY, pp131-170.

6- Hastings, J.W. (1996) Chemistries and colors of bioluminescent reactions: A review, Gene no 173, pp5-11.

7- Meighen, E.A. (1992) Bioluminescence, bacterial. Encyclopedia of Microbiology, volume 7.

8- Meighen, E.A. (1991) Molecular biology of bacterial bioluminescence, Microbiological reviews, p123-142.

9- Meighen, E.A. and Dunlap, P.V. (1993) Physiological, biochemical and genetic control of bacterial bioluminescence, Advances in microbial physiology, v34, pp 1-67.

10- Hastings, J.W. and Dunlap, J.C. (1986) Cell-free components in dinoflagellate bioluminescence. The particulate activity: scintillons; the soluble components:luciferase, luciferin and luciferin binding protein. Meth Enzymol. No 133, pp307-323.

11- Meighen, E.A. (1994) genetics of bacterial bioluminescence, Annu. Rev. Genet. v. 28, pp117-139.

12- Martin, M., Showalter, R. and Silverman, M. (1989) Identification of a locus controlling expression of luminescence genes in *Vibrio harveyi*. J. Bact., pp 2406-2414.

13- Swartzman, E., Kapoor, S., Graham, A.F. and Meighen, E.A. (1990) A new *Vibrio fischeri lux* gene precedes a bidirectional termination site for the lux operon, J. Bact., v172, no 12, pp6797-6802.

14- Lee, C.Y., Szittner, R.B. and Meighen, E.A. (1991) The lux genes of the luminous bacterial symbiont, *P. leiognathi*, of the ponyfish, Eur. J. Biochem., v201, pp 161-167.

15- Swartzman, E., Miyamoto, C.M., Graham, A.F. and Meighen, E.A. (1990) Delineation of the transcriptional boundaries of the lux operon of *Vibrio harveyi* demonstrates the presence of two new lux genes, J.Biol. Chem., v 265, no. 6, pp3513-3517.

16- Mok, K.C., Wingreen, N.S. and Bassler, B.L. (2003) *Vibrio harveyi* quorum sensing : a coincidence detector for two autoinducers control gene expression. EMBO J., v 22, no. 4, pp870-881.

17- Pesci, E.C. and Iglewski, B.H. (1999) Quorum sensing in *Pseudomonas aeruginosa*, Cell-cell signaling in bacteria, American Society for Microbiology, Washington, D.C., pp 147-155.

18- Whiteley, M., Lee, K.M. and Greenberg, E.G. (1999) Identification of genes controlled by quorum sensing in *Pseudomonas aeruginosa.* P.N.A.S., U.S.A., v96, pp13904-13909.

19- Callahan, S.M. and Dunlap, P.V. (2000) LuxR- and Acyl-homoserine-lactone-controlled non-lux genes define a quorum sensing regulon in *Vibrio fischeri*, J. Bact., v 182, no.10, pp2811-2822.

20- Miyamoto, C.M., Dunlap, P.V., Ruby, E.G. and Meighen, E.A. (2003) LuxO controls LuxR expression in *Vibrio harveyi*: evidence for a common regulatory mechanism in *Vibrio*, Mol. Micro., v 48, no. 2, pp537-548.

21- Swartzman, E., Silverman, M. and Meighen, E.A. (1992) The luxR gene-product of *Vibrio harveyi* is a transcriptional activator of the lux promoter, J. Bact., v 174, no. 22, pp7490-7493

22- Wall, L. and Meighen, E.A. (1986) Subunit structure of the fatty acid reductase complex from *Photobacterium phosphoreum*, Biochemistry, v 25, pp 4315-4321.

23- Riendeau, D., Rodrigues, A. and Meighen, E.A. (1982) Resolution of the fatty acid reductase from *Photobacterium phosphoreum* into acyl-protein synthetase and acyl-CoA reductase activities. J. Biol. Chem., v 257, no. 12, pp6908-6915.

24- Rodrigues, A., Wall, L., Riendeau, D. and Meighen, E.A. (1983) Fatty acid acylation of proteins in bioluminescent bacteria. Biochemistry, v 22, no. 24, pp5604-5611.

25- Wall, L., Rodriguez, A. and Meighen, E.A. (1986) Intersubunit transfer of fatty acyl groups during fatty acid reduction. J. Biol. Chem., v261, no. 34, pp15981-15988.

26- Soly, R.R. and Meighen, E.A. (1991) Identification of the acyl-transfer site of fatty acyl protein synthetase from bioluminescent bacteria. J. Mol. Biol., v219, pp69-77.

27- Hammer, B.K. and Bassler, B.L. (2003) Quorum sensing controls biofilm formation in *Vibrio Cholerae.* Mol. Microbiol.,v50(1), pp101-104

28- Cvitkovitch, D.G., Li, Y.H. and Ellen, R.P. (2003) Quorum sensing and biofilm formation in Streptococcal infections. J.Clin. Invest., v112(11)pp1626-1632

29- Smith, R.S. and Iglewski, B.H. (2003) *Pseudomonas aeruginosa* quorum sensing as a potential antimicrobial target. J.Clin. Invest, v112(10), pp1460-1465.

30- Novick, R.P. and Jiang, D. (2003) The Staphylococcal saeRS system coordinates environmental signals with agr quorum sensing. Microbiology, v149(Pt 10), pp2709-2717.

31- Perozich, J., Kuo, I., Wang, B.C., Boesch, J.S., Lindahl, R. and Hempel, J. (2000) Shifting the NAD/NADP preference in class 3 aldehyde dehydrogenases. Eur. J. Biochem., v 267, pp6197-6203.

32- Perozich, J., Nicholas, H., Wang, B.C., Lindahl, R. and Hempel, J. (1999) Relationships within the aldehyde dehydrogenase extended family. Protein Science, v 8, pp137-146.

33- Byers, D. and Meighen, E.A. (1984) *Vibrio harveyi* aldehyde dehydrogenase. J. Biol. Chem., v259, no. 11, pp7109-7114.

34- Cobessi, D., Tete-Favier, F., Marchal, S., Azza, S., Branlant, G. and Aubry, A. (1999) Apo and holo crystal structures of an NADP-dependent aldehyde dehydrogenase from *Streptococcus mutans.* J. Mol. Biol. V 290, pp161-173.

35- Rodriguez, A., Riendeau, D and Meighen, E.A. (1983) Purification of the acyl coenzyme A reductase component from a complex responsible for the reduction of fatty acids in bioluminescent bacteria. J. Biol. Chem., v 258, no. 8, pp5233-5237.

36- Hempel, J., Nicholas, H. and Lindahl, R. (1993) Aldehyde dehydrogenases: Widespread structural and functional diversity within a shared framework. Protein Science, v 2, pp1890-1900.

37- Zhang, L., Ahvazi, B., Szittner, R., Vrielink, A. and Meighen, E.A. (2001) Differences in nucleotide specificity and catalytic mechanism between *Vibrio harveyi* aldehyde dehydrogenase and other members of the aldehyde dehydrogenase superfamily. Chemico-Biological Interactions, pp29-38

38- Sheikh, S., Ni, L., Hurley, T.D. and Weiner, H. (1997) The potential roles of the conserved amino acids in human liver mitochondrial aldehyde dehydrogenase. J. Biol. Chem., v272, no. 30, pp18817-18822.

39- Ahvazi, B., Coulombe, R., Delarge, M., Vedadi, M., Zhang, L., Meighen, E.A. and Vrielink, A. (2000) Crystal structure of the NADP-dependent aldehyde dehydrogenase from *Vibrio harveyi*: Structural implications for cofactor specificity and affinity. Biochem J., v 349, pp853-861.

40- Vedadi, M. and Meighen, E.A. (1997) Critical glutamic acid residues affecting the mechanism and nucleotide specificity of *Vibrio harveyi* aldehyde dehydrogenase. Eur. J. Biochem., v 246, pp698-704.

41- Bognar, A. and Meighen, E.A. (1983) Mechanism of a long-chain fatty aldehyde dehydrogenase induced during the development of bioluminescence in Beneckea harveyi. Can J Biochem Cell Biol. v 61, no. 5, pp301-306.

42- Matthews BW. (1968) Solvent content of protein crystals. J Mol Biol **33:**491-497.

43- Otwinowski, Z. (1993) in proceedings of the CCP4 suite study weekend, Data Collection and Processing (Sawyer, L., Isaacs, N. and Bailey, S., eds.)pp56-62, SERC Daresbury Laboratory, Warrington, U.K.

44- Minor, W. (1993) XDISPLAYF program, Purdue University.

45- Collaborative Computational Project, No. 4 (1994) The CCP4 suite: programs for protein crystallography, Acta Crystallogr. D., v50, pp760-763.

46- Hendrickson, W.A., Horton, J.R. and LeMaster, D.M. (1990) Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structures. EMBO J., v 9, pp1665-1672.

47- Ramakrishnan, V. and Biou, V. (1997) Treatment of multiwavelength anomalous diffraction as a special case of multiple isomorphous replacement. Methods Enzymol., v276, pp538-557.

48- Weeks, C.M., DeTitta, G.T., Hauptman, H.A., Thuman, P. and Miller, R. (1994). Structure solution by minimal function phase refinement and Fourier filtering: II. implementation and applications, Acta Crystallogr. A., v 50, pp210-220.

49- Kleywegt, G.J. and Jones, T.A. (1996). xdlMAPMAN and xdlDATAMAN - programs for reformatting, analysis and manipulation of biomacromolecular electron-density maps and reflection data sets. Acta Crystallogr. D., v 52, pp826-828.

50- Jones, T.A., Zou, J.Y., Cowan, S.W. and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A. v 47, pp110-119.

51- Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T. and Warren, G.L. (1998) Crystalography and NMR system:A new software suite for macromolecular structure determination. Acta Crystallogr. D., v 54, pp905-921.

52- Kantardjieff K.A., Rupp B. (2003). Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. Protein Science v 12, pp 1865-1871.

53- Corpet, F. (1988) Multiple sequence alignment with hierarchical clustering, Nucl. Acids Res., v16, no. 22, pp10881-10890.

54- Richter, G., Volk,R., Krieger, C., Lahm, H.W., Röthlisberger, U. and Bacher, A. (1992) Biosynthesis of riboflavin: cloning, sequencing, and expression of the gene coding for 3,4-dihydroxy-2-butanone 4-phosphate synthase of Escherichia coli., J. Bacteriol. V174, no. 12, pp 4050-4056.

55- Hastings, J.W. and Weber, G (1963) J. Opt. Soc. Am. 53, pp 1410-1415.

56- Cobessi, D., Tete-Favier, F., Marchal, S., Branlant, G., and Aubry, A. (2000) Structural and biochemical investigations of the catalytic mechanism of an NADP-dependent aldehyde dehydrogenase from Streptococcus mutans. J Mol Biol. v 300, no. 1, pp 141-152.

57- Hammen, PK., Allali-Hassani, A., Hallenga, K., Hurley, T.D., Weiner, H. (2002) Multiple conformations of NAD and NADH when bound to human cytosolic and mitochondrial aldehyde dehydrogenase. Biochemistry, v 41, no. 22, pp 7156-7168.

58- Wei, B., Mays, D.C., Lipsky, J.J., Weiner, H. (2003) Chemical modifications to study mutations that affect the ability of the general base (E268) to function in human liver mitochondrial aldehyde dehydrogenase. Chem. Biol. Interact. pp85-91.

59- Wang, X. and Weiner, H. (1995) Involvement of glutamate 268 in the active site of human liver mitochondrial (class 2) aldehyde dehydrogenase as probed by site-directed mutagenesis. Biochemistry, v 10, no.34(1), pp 237-243.

60- Marchal, S., Cobessi, D., Rahuel-Clermont, S., Tete-Favier, F., Aubry, A. and Branlant, G. (2001) Chemical mechanism and substrate binding sites of NADP-dependent aldehyde dehydrogenase from Streptococcus mutans. Chem. Biol. Interact. Pp15-28.

61- Vedadi, M., Szittner, R., Smillie, L. and Meighen, E.A. (1995) Involvement of cysteine 289 in the catalytic activityof an NADP-specific fatty aldehyde dehydrogenase from *Vibrio harveyi*, Biochem., v 34, pp16725-16732.

62- Stura, E (1999) Protein Crystallization, Techniques, Strategies, and Tips edited by Bergfors, T, International University Line, La Jolla, pp141-153.

63- Lee, C.Y. and Meighen, E.A. (1997) Cysteine-286 as the site of acylation of the Lux-specific fatty acyl-CoA reductase, B.B.A., no 1338, pp215-222.\

64- Wei, B., Mays, D.C., Lipsky, J.J. and Weiner, H. (2003) Chemical modifications to study mutations that affect the ability of the general base (E268) to function in human liver motichondrial aldehyde dehydrogenase. Chemico-Biol. Int., pp85-91.