Evolution, Structure, and Function of Phenoloxidase in the Pea Aphid *Acyrthosiphon pisum*

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements of the Degree of

Master of Science

By

Eric Edward Van Fleet

January 2012

ABSTRACT

Phenoloxidases (monophenol monooxygenase, EC 1.14.18.1; catechol oxidase, EC 1.10.3.1) are a group of enzymes with copper cofactors that produce reactive quinones and are part of the melanin synthesis pathway, both of which have important roles in immunity. The pea aphid, *Acyrthosiphon pisum*, which according to genome annotation is deficient in many other immune system components, codes for two phenoloxidase proteins that represent putative dimer components and possess the amino acid residues contributing to the active site. Constitutive phenoloxidase activity was detectible in the pea aphid hemolymph. It was activated by both conformational change with methanol and proteolytic cleavage of the propeptide with trypsin. Phenoloxidase activity was not significantly altered by aseptic wounding or infection studies with *Escherichia coli* or *Micrococcus luteus*. Phylogenetic analysis of insect phenoloxidases yielded a topology consistent with a lineage-specific duplication in each order (including Hemiptera). The possibility that the topology could be generated by a duplication, probably in the ancestral insect, followed by coevolution between the two phenoloxidase subunits within each order, was explored but rejected. The three-dimensional structures of the pea aphid phenoloxidases were reconstructed by homology modeling. The models of all three possible dimeric states of phenoloxidase (two homodimers and one heterodimer) did not exhibit conformational change in response to propeptide cleavage and their conformation differed from other modeled insect phenoloxidases. Taken together, these results suggest that the pea aphid has a functional phenoloxidase, but that it may be activated and function in a different way from the phenoloxidases in previously-studied insects.

BIOGRAPHICAL SKETCH

Eric Van Fleet was born on April 9, 1987 in Mifflinburg, Pennsylvania. He graduated from Mifflinburg Area High School in 2005, and enrolled in Ithaca College in Ithaca, New York. He participated in undergraduate research with Dr. Vicki Cameron, studying the effects of the nuclear gene YME1 on the regulation of cytochrome C oxidase and its effects on respiration in *Saccharomyces cervisiae*. He also worked with Dr. Marina Caillaud on several projects, including the sequencing of a putative odorant binding protein in *Manduca sexta* involved in detection of solanaceous plants, and transcriptome analysis of host-differentiated pea aphids. While working with Dr. Caillaud, he contributed to the International Aphid Genomics Consortium and earned an authorship on the genome paper of the pea aphid, *Acyrthosiphon pisum*. He graduated from Ithaca College in 2009 with his Bachelor of Science in Biochemistry, with Honors. His interest in science was fostered by his parents James and Melinda Van Fleet, a science and engineering reference librarian and a biochemist respectively, and a long series of supportive and talented teachers and research advisors.

My success is dedicated to

my parents, who gave me the drive,

my teachers, who gave me the means,

and my wife, who gave me the reason.

ACKNOWLEDGEMENTS

I would like to first thank my advisor, Dr. Angela Douglas, for being my mentor during my graduate studies. She allowed me to explore my research goals, and even when my investigations took unexpected tacks she helped me incorporate them all into the larger scheme of insect immunity and realize their broader impacts. Dr. Douglas has been an important factor in both my professional and personal development over the course of my degree, and without her guidance I cannot conceive accomplishing what I have managed up to this point.

I also thank Dr. Brian Lazzaro, my committee member and director of departmental graduate studies, for his contributions to my thesis and my breadth of knowledge of the sciences. My discussions with him helped refine my ideas, and introduced me to many lines of research that sparked my interests and inspired new directions for my research.

Special thanks to all the members of the Douglas lab; they all served as both friends and colleagues and were always willing to provide a sanity check for my research. Particular thanks to Adam Wong, for putting up with my constant scribbling on the office chalkboard and listening as I debunked my own theories, to Calum Russell, for providing moral support and always having a fascinating side project to talk about, and George Lin, for helping to jump-start my interest in computational biology.

Dr. Cole Gilbert and Dr. Bryan Danforth gave me wonderful opportunities to be a teaching assistant, and with their guidance I developed an appreciation for teaching. No matter where my career takes me, I look forward to one day having the opportunity to inspire children with the wonders of science.

My undergraduate research advisors, Dr. Marina Caillaud and Dr. Vicki Cameron, both provided me excellent opportunities, and without the undergraduate research projects I worked on with them I doubt I would have been able to succeed in my graduate work. Special thanks to Dr. Caillaud for facilitating my participation in the International Aphid Genomics Consortium, where I first met Dr. Douglas.

Finally, I am infinitely grateful for the support of my family. My parents, James and Melinda, were unceasing in their support for my goal of being a scientist, and I've done my best to repay their support with pride. My sister, Liz, despite being my junior, has always been remarkably insightful and has helped me overcome a panalopy of conundrums, and was always willing to lend an ear when times were difficult. And my wife, Amanda, has helped me in ways too innumerable to list in my pursuit of this project.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

The genome sequence of the pea aphid, *Acyrthosiphon pisum*, revealed major apparent deficiencies in its immune system, compared to other insects with sequenced genomes (IAGC 2010; Gerardo et al. 2010). These deficiencies include an incomplete immune deficiency (IMD) pathway, drastic reductions in antimicrobial peptide genes, and no recognizable peptidoglycan recognition proteins. This has lead to the hypothesis that aphids are deficient in their immune response, raising the possibility that their condition is an adaptation to the maintenance their obligate bacterial symbiont, *Buchnera aphidicola*, and secondary defensive symbionts (Burke and Moran 2011). However, other insects that form obligate interactions with bacteria, such as the tsetse fly, do not appear to have similar immunological deficiencies (Roditi and Lehane 2008). In addition, while the phenoloxidase activation pathway in pea aphids is poorly characterized, the phenoloxidase genes appear to be intact according to the genome annotation. Insects generally use phenoloxidase in two capacities: as an antimicrobial defense mechanism, and as the agent of cuticular sclerotization during development, molting, and wound healing (Cerenius and Soderhall, 2004). While aphid phenoloxidases might function in sclerotization, its presence provides a potential mechanism for innate immune defense in the pea aphid, particularly given the presence of phenoloxidase in the bacteriocytes and hemocytes (Poliakov et al. 2011, McLaughlin et al. 2011)

Phenoloxidase is a copper binding oxidative enzyme that mediates the initial steps of melanin production (Figure 1). It is an example of multicopper oxidases, which occur across all organisms, being found in bacteria, fungi, plants, and animals, but having different physiological functions in each (Garcia-Pereja et. al. 1987; Polacheck and Kwon-Chung, 1988; Crawford

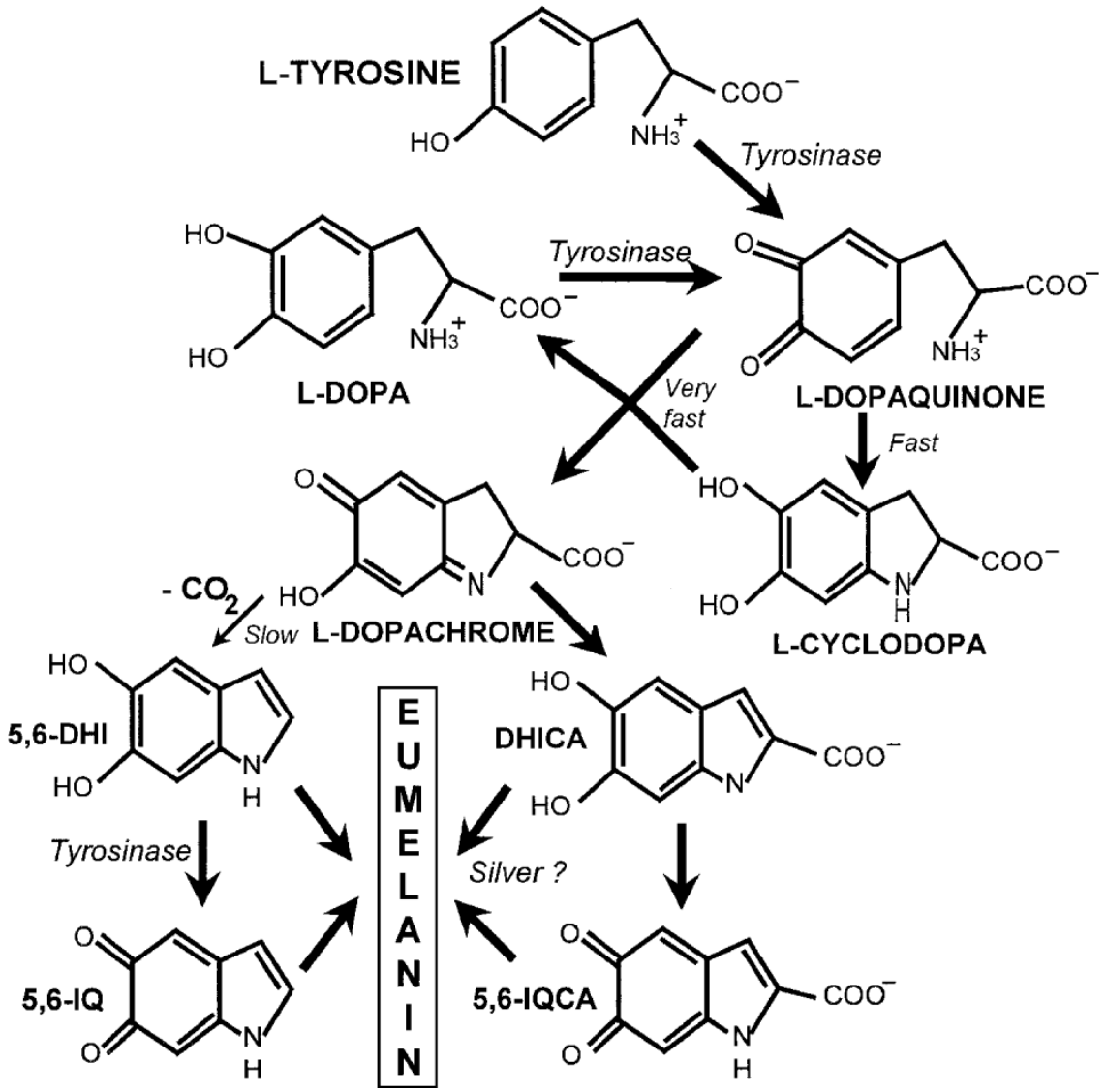


Figure 1. The eumelanin branch of the melanin synthesis pathway. Eumelanin is the form of melanin used for wound healing and encapsulation in insects (Nappi and Sugumaran 1993). L-tyrosine and L-dopa can be converted into L-dopaquinone by tyrosinases, but catechol oxidases only convert L-dopa because of incompatibility with the monophenolic structure of tyrosine. The phenoloxidases in this study are annotated as possessing tyrosinase and catechol oxidase activity. L-dopaquinone rapidly, spontaneously degrades to L-dopachrome. Both tyrosinases and catechol oxidases accept other phenolic substrates, like dopamine, forming dopamine quinone and dopaminochrome (not shown). Image modified from Olivares et al. 2001.

1967). Phenoloxidase converts L-tyrosine (monophenol monooxygenase activity, EC 1.14.18.1) or L-dopa (catechol oxidase activity, EC 1.10.3.1) into L-dopaquinone, a highly reactive oxidizing agent that can function as an antimicrobial agent or spontaneously converted into L-dopachrome, which is a precursor of melanin (Sanchez-Ferrer et. al. 1995). Laccase type phenoloxidases, which only possess the catechol oxidase activity, are the primary agents of cuticular sclerotization in insects and are generally found in the cuticle. Immune-type phenoloxidases, on the other hand, possess both tyrosinase and catechol oxidase activities and are more commonly used in antimicrobial responses. They are primarily stored in the hemocytes or free in the hemolymph depending on the organism, although they can also be found in other tissues. Because of the extreme oxidative properties of dopaquinone, phenoloxidase is stored as an inactive precursor referred to as prophenoloxidase. Activation of phenoloxidase has been shown to be mediated by the Toll pathway *in vivo* in *Drosophila melanogaster* (Ligoxygakis et. al. 2002), but in most arthropods only the direct upstream components, the serine protease cascade that cleaves the propeptide, are characterized (Piao et. al. 2005).

The prophenoloxidase activating system is unusual in the fact that, while it is primarily a serine protease cascade, cleavage of the propeptide is necessary but not sufficient to induce enzymatic activity. The additional activation is mediated by serine protease homologs, which are proteins that contain serine protease domains that are enzymatically non-functional, bind to the phenoloxidase dimer to cause a conformational change (Cerenius and Soderhall 2004). This allows a functional serine protease, identified in several organisms as the prophenoloxidase activating enzyme (PPAE), to access the propeptide cleavage site and remove the propeptide (Wang et al. 2001). The propeptide cleavage is not only important for directly opening up the protein's active site, but it also induces another conformational change which is essential to

enzymatic activation (Hall et al. 1995). This reliance on conformational change for activation is evident in assays for phenoloxidase activity. A large range of compounds has been shown to activate phenoloxidases *in vitro*, including alcohols and detergents (Soderhall 1982, Hall et al. 1995). Mechanistically, these compounds disrupt the electrostatic interaction between the propeptide and the rest of the protein (Li et al. 2009), resulting in conformational change without propeptide cleavage. Generic proteases can also be used as activators for prophenoloxidase, but can have variable effects between organisms (Saul and Sugumaran 1987). Finally, bacterial and fungal cell wall components have been used as activators *in vivo* and *in vitro*, with varying success, to induce the prophenoloxidase activating pathway.

The crystal structure of phenoloxidase from *Manduca sexta* has been analyzed, (Li et al. 2009) and gives insight into the mechanism of phenoloxidase activity *in vivo*. Insect phenoloxidases form a dimer, and while each phenoloxidase subunit is individually active once the propeptide is cleaved (Sanchez-Ferrer et al. 1995), dimerization is required for interaction with the prophenoloxidase activating system to achieve this cleavage *in vivo*. One important output of the analysis of the crystal structure was the demonstration of the involvement of the propeptide region in the dimeric interactions. Propeptide cleavage affects not only the activation of the individual subunits, but also the conformation of the dimer. The underlying mechanism has not been studied in insect phenoloxidases, but dimer conformation has been suggested to play a role in phenoloxidase activation in other systems, including the spiny lobster *Panulirus argus* (Perdomo-Morales et al. 2008). The x-ray crystallography of *Manduca sexta* phenoloxidase was done on the heterodimeric form, and studies have shown that these phenoloxidases show a preference for heterodimer formation (Jiang et al. 1997), with homodimers present but less prevalent. This is not the case in *Drosophila melanogaster,* where

both homodimeric forms of phenoloxidase are prevalent and the heterodimer is rare (Asada and Sezaki 1999). The mechanism of dimer preference is still unknown. However, these studies indicate either that there is a regulatory mechanism towards specific dimer formation, or that protein interacting regions change over evolutionary time, resulting in either homodimers or heterodimers being more efficient in different insects.

Presented here is a characterization of phenoloxidase from the pea aphid, based on enzymatic assays, phylogenetic analysis, and modeling of the three-dimensional structure of the enzyme. Assays of pea aphid phenoloxidase under sterile and septic wounding conditions were performed to determine if phenoloxidase activity was induced either as a wounding response or an immune response to bacterial infection. Subsequently, assays were repeated using prophenoloxidase activators to determine levels of proenzyme available to the organism. Phylogenetic analysis was performed to establish whether pea aphid phenoloxidases resolved into a clade with other, better characterized phenoloxidases that have been shown to be functional immune system components. In the process, an unusual evolutionary pattern of lineage specific duplications was observed among organisms with multiple phenoloxidases within the different insect orders. Further investigations were performed on laccase-type phenoloxidases to detect if the pattern of lineage specific duplication was more widespread. Analysis was also performed on the ecdysone receptor and ultraspiracle system and compared to phenoloxidase, to determine if coevolutionary processes occurring in the former system were similar to those driving the evolutionary history of the phenoloxidases. Finally, the 3D structures of the insect phenoloxidases used in the phylogenetic analysis were modeled, to investigate the effects of protein-protein interactions on the functionality of phenoloxidase dimers. The theoretical dimers of pea aphid phenoloxidase were compared to those of other organisms in terms of surface area of protein-protein

interactions, as well as stabilizing hydrogen bonds and salt bridges, to determine if these dimers operated similarly to other functional phenoloxidases. These comparisons were also done on 3D models with and without propeptides, to discern whether pea aphid phenoloxidase was amenable to activation by a standard prophenoloxidase activating system.

Works Cited

Asada, N. and Sezaki, H. (1999) Properties of phenoloxidases generated from prophenoloxidase with 2-propanol and the natural activator in *Drosophila melanogaster*. Biochemical Genetics 37(5/6):149-158

Burke, G.R. and Moran, N.A. (2011) Responses of the pea aphid transcriptome to infection by facultative symbionts. Insect Molecular Biology 20(3):357-365

Cerenius, L. and Soderhall, K. (2004) The prophenoloxidase-activating system in invertebrates. Immunological Reviews 198:116-126

Crawford, R .M.M. (1967). Phenol oxidase activity and flooding tolerance in higher plants. Nature 214:427-428.

García-Pareja, M.P., Monteoliva-Sanchez, M., García de la Paz, A.M., Corominas, E., Pérez, M.L., and Ramos Cormenzana, A. (1987) Characterization of bacteria with phenol-oxidase activity from soil in the La Laguna area (Spain). Chemosphere **16**:2627–2630.

Gerardo, N.M, Altincicek, B., Anselme, C., Atamian, H., Barribeau, S.M, de Vos, M., Duncan, E.J., Evans, J.D., Gabaldón, T., Ghanim, M., Heddi, A., Kaloshian, I., Latorre, A., Moya, A., Nakabachi, A., Pérez-Brocal, V., Pignatelli, M., Rahbé, Y., Ramsey, J.S., Spragg, C.J., Tamames, J., Tamarit, D., Tamborindeguy, C., Vincent-Monegat, C., and Vilcinskas, A. (2010) Immunity and other defenses in pea aphids, *Acyrthosiphon pisum*. Genome Biology 11(2):R21

Hall, M., Scott, T., Sugumaran, M., Soderhall, K. and Law, J.H. (1995) Proenzyme of *Manduca sexta* phenol oxidase: purification, activation, substrate specificity of the active enzyme, and molecular cloning. PNAS 92:7764-7768

The International Aphid Genomics Consortium (2010) Genome Sequence of the pea aphid *Acyrthosiphon pisum*. PLoS Biology 8(2): e1000313

Jiang, H., Wang, Y., Ma, C. and Kanost, M.R. (1997) Subunit composition of pro-phenol oxidase from *Manduca sexta*: molecular cloning of subunit ProPO-P1. Insect Biochemistry and Molecular Biology 27(10):835-850

Li, Y., Wang, Y., Jiang, H. and Deng, J. (2009) Crystal structure of *Manduca sexta* prophenoloxidase provides insights into the mechanisms of type 3 copper enzymes. PNAS 106(40):17001-17005

Ligoxygakis, P., Pelte, N., Ji, C., Leclerc, V., Duvic, B., Belvin, M., Jiang, H., Hoffmann, J.A. and Reichhart, J.M. (2002b) A serpin mutant links Toll activation to melanization in the host defence of *Drosophila*. EMBO Journal 21:6330-6337

Nappi, A.J. and Sugumaran, M. (1993) Some biochemical aspects of eumelanin formation in insect immunity, in Pathak, J.P.N., ed, Insect Immunity, Springer, New York, p. 132- 145

Olivares, C., Jiménez-Cervantes, C., Lozano, J.A., Solano, F. and García-Borrón, J.C. (2001) The 5,6-dihydroxyindole-2-carboxylic acid (DHICA) oxidase activity of human tyrosinase. Biochemical Journal 354:131-139

Perdomo-Morales, R., Montero-Alejo, V., Perera, E., Pardo-Ruiz, Z. and Alonso Jiménez, E. (2008) Hemocyanin-derived phenoloxidase activity in the spiny lobster *Panulirus argus* (Latreille, 1804). Biochimica et Biophysica Acta 1780:652-658

Piao, S., Song, Y.L., Kim, J.H., Park, S.Y., Park, J.W., Lee, B.L., Oh, B.H. and Ha, N.C. (2005) Crystal structure of a clip-domain serine protease and functional roles of the clip domains. EMBO Journal 24:4404-4414

Polacheck, I. and Kwon-Chung, K.J. (1988) Melanogenesis in Cryptococcus neoformans. Journal of General Microbiology 134:1037-1041

Poliakov, A., Russell, C.W., Ponnala, L., Hoops, H.J., Sun, Q., Douglas, A.E. and van Wijk, K.J. (2011) Large-scale label-free quantitative proteomics of the pea aphid-*Buchnera* symbiosis. Molecular and Cellular Proteomics 10: M10.007039

Roditi, I. and Lehane, M.J. (2008) Interactions between trypanosomes and tsetse flies. Current Opinion in Microbiology 11(4):345-351

Sánchez-Ferrer, A., Rodríguez-López, J.N, García-Cánovas, F, and García-Carmona, F (1995) Tyrosinase: A comprehensive review of its mechanism. Biochimica et Biophysica Acta 1247:1-11

Saul, S.J. and Sugumaran, M. (1987) Protease mediated phenoloxidase activation in the hemolymph of the tobacco hornworm, *Manduca sexta*. Archives of Insect Biochemistry and Physiology 5:1-11

Soderhall, K. (1982) Prophenoloxidase activating system and melanization - a recognition mechanism of arthropods? A review. Developmental and Comparative Immunology 6(4):601-611.

Wang, R., Lee, S.Y., Cerenius, L. and Soderhall, K. (2001) Properties of the prophenoloxiase activating enzyme of the freshwater crayfish, *Pacifastacus leniusculus*. European Journal of Biochemistry 268:895-902

CHAPTER 2: ENZYMATICS OF PEA APHID PHENOLOXIDASES

Introduction

The hallmarks of phenoloxidase as a component of the immune system are easily identifiable. The activation of phenoloxidase is generally strictly spatially restricted, to avoid non-specific damage (Sadd and Siva-Jothy 2006). For this reason, phenoloxidase activity is often measured using the hemolymph or just the hemocytes (Iwama and Ashida 1986) when assaying response to bacterial infection, or hemolymph drawn specifically from the wound site when measuring wound healing (Bidla et al. 2005). However, this poses a problem in the pea aphid, since its size make it difficult to perform these sorts of assays repeatedly and reliably. In addition to hemolymph, whole body homogenates were fractionated with ammonium sulfate, and the fraction of the protein extract containing phenoloxidase was used to perform assays. These assays were initially performed over a 72-hour timecourse after the aphids reached the final larval instar. However, severe mortality in the bacterially infected aphids after 24 hours limited the amount of usable data.

Aphid response to bacterial infection was measured using *Escherichia coli* and *Micrococcus luteus*. These bacteria were selected to represent, respectively, Gram-negative and Gram-positive bacteria as a whole, and specifically bacteria possessing meso-diaminopimelic-acid-type (DAP-type) peptidoglycan or L-lysine-type (Lys-type) peptidoglycan (Swaminathan et al. 2006). Due to the absence of peptidoglycan recognition proteins (PGRPs) in the aphid genome, it is unknown whether or how the aphid can recognize differences in bacterial morphology. Additionally, the absence of key components of the IMD pathway, which has been associated with the immune response to (generally Gram-negative) bacteria with DAP-type

peptidoglycan in *Drosophila melanogaster* (Lemaitre 2004), indicates there may be differential recognition and response between the two classes of bacteria. Our initial hypothesis was that lack of immune activity against Gram-negative bacteria would represent an adaptation to bacterial symbiosis, since *Buchnera aphidicola* and several secondary symbionts are Gram-negative γ-proteobacteria.

Phenoloxidase activity was also measured under the effects of exogenous activating treatments. This latter experiment allowed us to estimate the total available phenoloxidase in the organism. Because of the distinctive activating system of immune-system phenoloxidases as compared to laccase-type phenoloxidases, these treatments would restrict the assay results to reporting the activity of the former.

Methods

Aphid Rearing

Acyrthosiphon pisum clone CWR09-18 was derived from a single parthenogenetic female collected from an alfalfa crop in Freeville, NY, USA in 2009, and was cultured on pre-flowering *Vicia faba* at 20°C and a 16:8 light/dark cycle. I selected this clone for my research because diagnostic PCR assays and microscopical examination revealed that it contained the symbiotic bacterium *Buchnera aphidicola* and no known secondary endosymbiont (CW Russell, pers. com). To obtain aphids lacking *Buchnera*, two-day-old larvae were transferred from plants to the chemically-defined diet of Douglas & Prosser (1992) with 0.5 M sucrose and 0.15 M amino acids supplemented with 50 μg of the antibiotic rifampicin ml^{-1}, with aphids on rifampicin-free diets as controls. Two days later, the aphids were transferred to either rifampicin-free diet or, for hemolymph collection, to plants. All experiments were conducted on 7-day-old insects. Under

all treatments, these insects were final-instar larvae, and the aphids treated with the rifampicin-free diet, but not those derived from rifampicin treatment, contained detectable *Buchnera* as determined by the *Buchnera*-specific PCR assay

Phenoloxidase Assays

The phenoloxidase activity of whole aphids and hemolymph was quantified. Whole aphid samples were processed by homogenizing individual 7-day-old larval aphids in 150 µL ice-cold PBS pH 7.4. After centrifugation at 15,000 RCF at 4°C for 10 minutes, 100 µL supernatant was added to saturated ammonium sulfate, pH 8.0, to form a 50% ammonium sulfate solution. Samples were mixed on ice for 2 minutes, and centrifuged for 30 minutes at 15,000 RCF. The protein pellet was resuspended into 150 µL PBS, pH 7.4. Hemolymph was collected by submerging aphids in water-saturated mineral oil, removing the front pair of legs, and collecting the exuding hemolymph droplets into a graduated microcapillary tube. Hemolymph droplets that contained visible debris were discarded; these were typically from aphids in which the leg was removed too close to the body, and fat body was released along with hemolymph. No visible melanization of the hemolymph droplets occurred within an hour of initial wounding. Hemolymph from up to 30 aphids was pooled to a final volume of 3 µL, which was added to 147 µL PBS, pH 7.4.

Phenoloxidase activity was monitored by change in absorbance at 492 nm as the dopamine substrate was enzymatically converted into dopaminochrome. The experimental sample comprised 50 µL sample and 50 µL 20 mM dopamine (substrate). The negative controls included a heated control for which 50 µL sample was inactivated by incubation at 95^{o}C for 5 minutes prior to phenoloxidase assay; a substrate-negative control with 45 µL sample and 50 µL

PBS pH 7.4; and a sample-negative control with 50 µL PBS pH 7.4 and 50 µL substrate. The protein content of 5 µL hemolymph solution or protein extract was assayed for protein content using the Bradford assay with BSA standard (Bio-rad Laboratories, California, USA). Change in absorbance was converted to dopaminochrome formation per minute per mg protein, using a molar extinction coefficient for dopamine of 3240 $M^{-1}cm^{-1}$. When activating treatments were used, they were added to 100 µL volumes of sample 5 minutes prior to assay and incubated at room temperature. Activating treatments consisted of 50 µL 25 µg trypsin μl^{-1} in 0.9% NaCl solution, or 50 µL 100% methanol, with 50 µL PBS pH 7.4 as control.

Two-way anova for phenoloxidase activity data was performed in SPSS PASW Statistics v.18 (SPSS Inc., Chicago IL).

Bacterial Infections

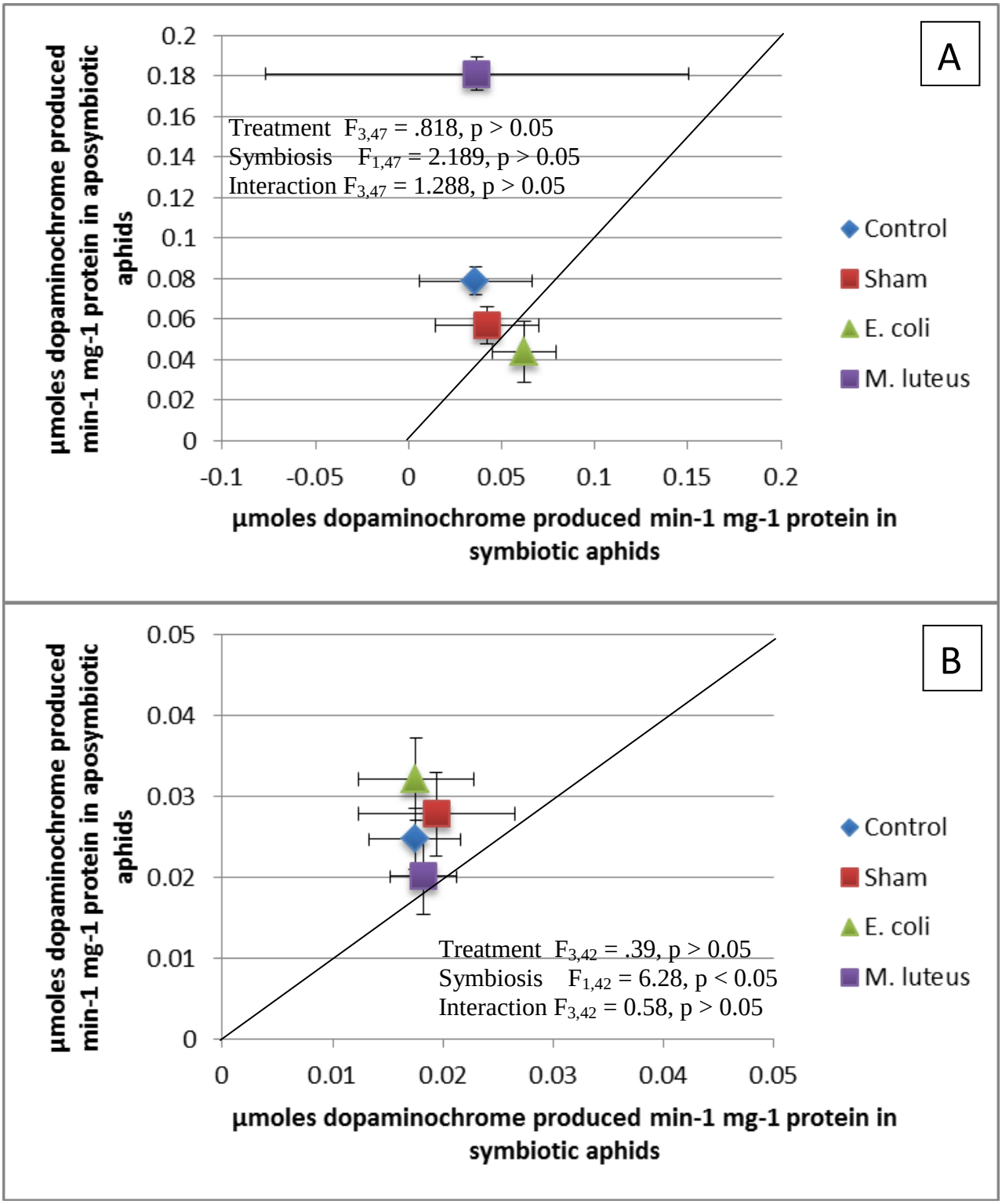
Escherichia coli strain JM109 or *Micrococcus luteus* strain HBN-1 was cultured on LB liquid medium at 37°C. Bacterial cultures were pelleted and resuspended in Carlson's solution (Harada et al. 1997). Bacterial concentration was calculated by cell counts using a hemocytometer, and all solutions were diluted to $1x10^6$ cells/mL. Six-day-old aphids were challenged by stabbing the aphids in the dorsal abdomen with a 0.3 mm insect pin dipped in a suspension of bacteria in Carlson's solution. Sham controls comprised stabbing with a sterile pin dipped in Carlson's solution.

Results

Phenoloxidase Activity Assay

Phenoloxidase activity was assayed at 0, 6, and 24 hours after wounding and/or bacterial challenge in symbiotic and aposymbiotic aphids (Figure 2). There was no significant effect of wounding or bacterial infection on phenoloxidase activity over time ($p > 0.05$). A significant difference between symbiotic and aposymbiotic aphids was observed, with aposymbionts showing a lower level of phenoloxidase activity ($0.01<p < 0.05$). Also, phenoloxidase activity showed a marked reduction over time, with all treatments showing higher levels of activity at the zero-hour time point that drops off by the six-hour time point ($p < 0.001$). These data suggest constitutive phenoloxidase activity regardless of wounding or bacterial infection. This constitutive activity is lower in aposymbiotic aphids than in symbiotic aphids, which links perturbation of *Buchnera* symbiosis to the distribution of phenoloxidase. The overall enzymatic activity also decreased over the course of the assay, corresponding with the aphid's development from the last larval instar into adulthood.

Insect phenoloxidases are generally produced as a proenzyme, and are stored in that form to prevent cellular damage from reactive oxygen species until activated by a serine protease (Saul and Sugumaran, 1987). To determine if this constitutive phenoloxidase activity represented the totality of phenoloxidase enzyme expressed in the pea aphid, phenoloxidase activity was assayed in samples that had been exposed to prophenoloxidase activators (Figure 3). Trypsin was used as a non-specific serine protease, since proteolytic cleavage of the ~55 kDa propeptide is normally required for prophenoloxidase activation. The prophenoloxidase activating serine protease in the pea aphid has not been identified, preventing analysis using proteases more similar to the native enzyme. Methanol was used as a representative of a wide family of alcohols and non-polar substances that have been shown to activate prophenoloxidase (Asada et al. 1993). Both treatments increased phenoloxidase activity, with trypsin inducing a


A
µmoles dopaminochrome produced min-1 mg-1 protein in aposymbiotic aphids
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
-0.1
-0.05
0
0.05
0.1
0.15
0.2
Treatment $F_{3,47}$ = .818, p > 0.05
Symbiosis $F_{1,47}$ = 2.189, p > 0.05
Interaction $F_{3,47}$ = 1.288, p > 0.05
Control
Sham
E. coli
M. luteus
µmoles dopaminochrome produced min-1 mg-1 protein in symbiotic aphids
B
µmoles dopaminochrome produced min-1 mg-1 protein in aposymbiotic aphids
0.05
0.04
0.03
0.02
0.01
0
0
0.01
0.02
0.03
0.04
0.05
Treatment $F_{3,42}$ = .39, p > 0.05
Symbiosis $F_{1,42}$ = 6.28, p < 0.05
Interaction $F_{3,42}$ = 0.58, p > 0.05
Control
Sham
E. coli
M. luteus
µmoles dopaminochrome produced min-1 mg-1 protein in symbiotic aphids

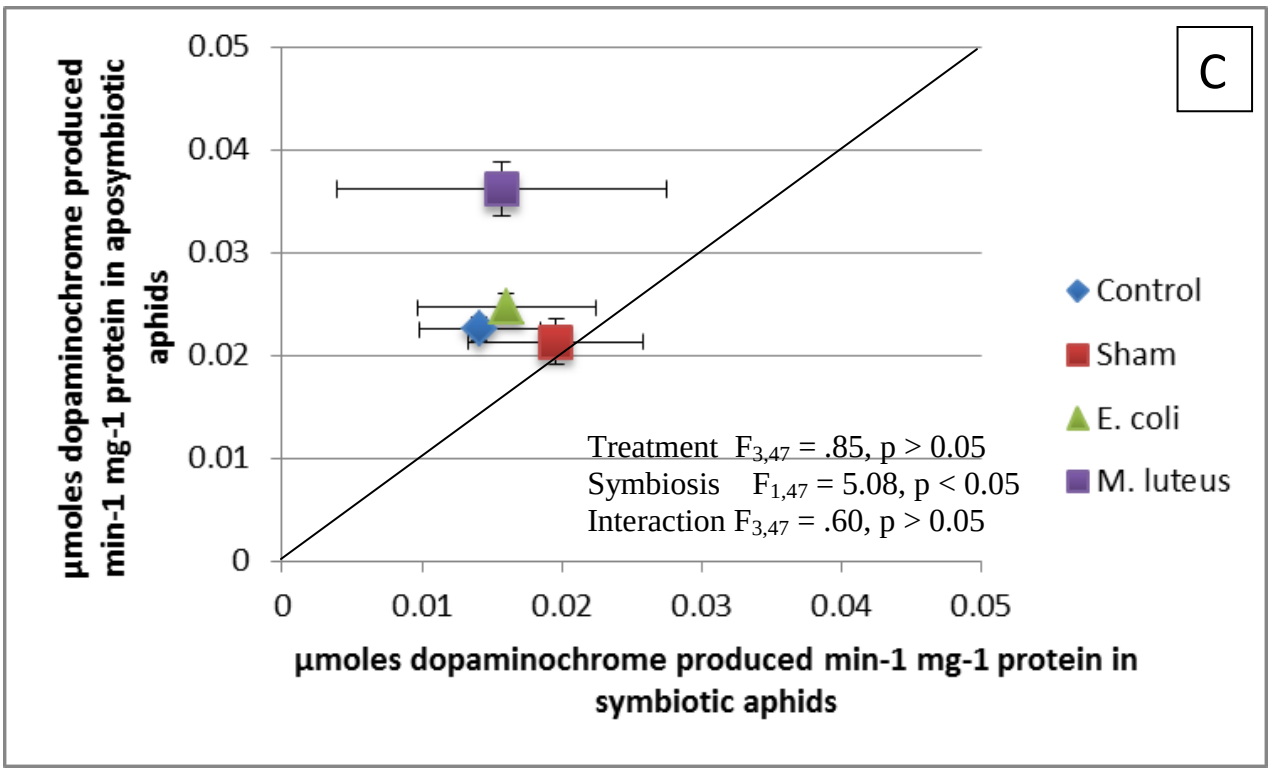


Figure 2. Timecourse assay of phenoloxidase activity in aphids under bacterial infection and wounding conditions. Each datapoint represents five replicates, and each replicate consists of five whole body homogenates. Figures A, B, and C depict phenoloxidase assays at 0, 6, and 24 hours respectively, plotting activity in symbiotic aphids versus aposymbiotic aphids along a line of equivalence. Analysis was performed using ANOVA.

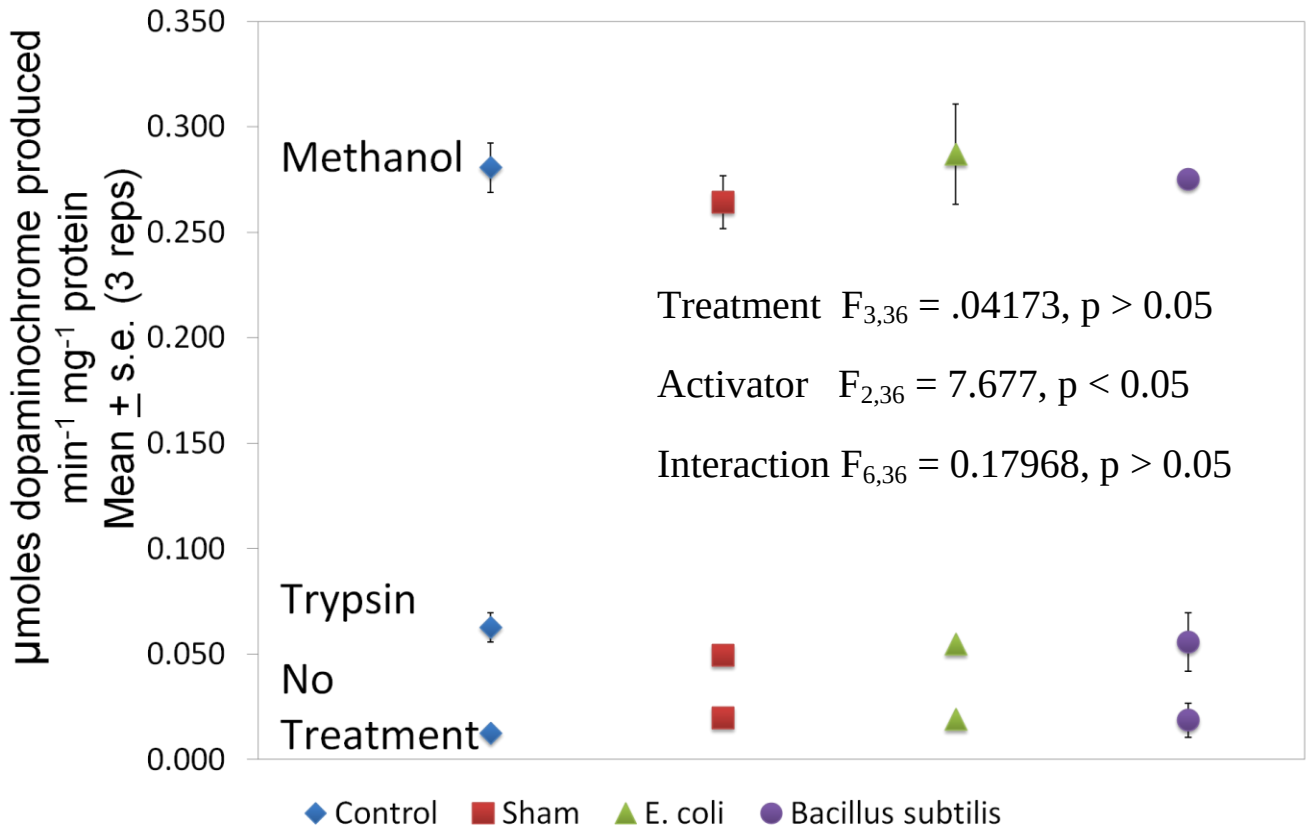


Figure 3. Prophenoloxidase activation in bacterially infected aphids. Each replicate consists of five aphid whole body homogenates. Each biological replicate was subjected to all three activation treatments. Analysis was performed using non-parametric Scheirer-Ray-Hare test.

three-fold increase and methanol inducing a ten-fold increase in enzymatic activity. This effect was independent of bacterial infection treatment, and supports the results of the timecourse assay that there is no inducibility of phenoloxidase by wounding or bacterial treatment. It does, however, demonstrate that in addition to the constitutively active phenoloxidase, there is a large reserve of prophenoloxidase protein in the aphid that is susceptible to exogenous activators.

Discussion

Pea Aphid Phenoloxidase As A Potential Immune System Component

The results of phenoloxidase assays in the pea aphid paint a picture that is very much at odds with other insect immune systems. Phenoloxidase is not activated by either bacterial infection or wounding, and its activity seems to be held at a very low level relative to the amount of prophenoloxidase available in the system. This information, alongside the deficiencies in the immune system mentioned previously, supports the idea that the aphid's immune system, viewed from the vantage of traditional immune pathways, is compromised. However, the lower constitutive phenoloxidase activity in aposymbiotic aphids than symbiotic aphids suggests an association between functional symbiosis with *Buchnera aphidicola* and phenoloxidase. Analysis of the aphid proteome has previously shown that the phenoloxidase protein is expressed, and is enriched in the bacteriocyte relative to other tissues (Poliakov et al., 2011). The phenoloxidase assay data show that there is less phenoloxidase activity per unit protein in the whole body of aposymbiotic aphids compared to symbiotic aphids, raising the possibility that a portion of the activated phenoloxidase may be in the bacteriocyte. However, the proteomic analysis was unable to distinguish between the proenzyme and activated form of phenoloxidase.

It is still unusual that, rather than being used as an immune system component, pea aphid phenoloxidase is localized to the same tissue as its' obligate endosymbiont.

There are several potential explanations for this observation. The first is that phenoloxidase is involved in the regulation of *Buchnera* to prevent overpopulation of the symbiont that could disrupt the beneficial effects of symbiosis for the pea aphid. In this scenario, reactive oxygen species would be produced in the bacteriocyte to limit bacterial growth. This process would very likely destroy the individual bacteriocyte where it was activated, because of nonspecific cell damage caused by phenoloxidase activity. This system, however, could be seen as in contrast with the fragility of the *Buchnera* population. In particular, *Buchnera* has been shown to be particularly sensitive to elevated temperature (Chen et al. 2009), with some populations of aphids going so far as to adopt a secondary symbiont which contributes to heat resistance (Montllor et al. 2002). Activation of even a small amount of phenoloxidase may end up causing widespread damage to the bacteriome, disrupting the symbiosis. Given the effort that the aphid invests in maintaining this symbiosis despite its relative fragility, it is improbable that it would use such a broad and powerful effector for population control of bacteria. However, the potential for widespread damage to the bacteriocyte reveals another possible function of phenoloxidase: the degradation of the bacteriome that occurs in mid-reproductive aphids. Studies have shown a pattern of degradation in *Buchnera* populations that accelerates with aging and proposed some mechanisms for bacterial control (Nishikori et al. 2005). However, not only are the bacteria killed, but the bacteriocytes lysed and undetectable by microscopy at the end of the aphid life (Douglas and Dixon 1987). Phenoloxidase may in fact be localized to the bacteriocyte for a task not related to management of endosymbionts, but to their destruction.

An alternative explanation is that *Buchnera aphidicola* is adapted to an environment rich in reactive oxygen species, and that the role of phenoloxidase is instead to restrict which bacteria have access to the bacteriocyte. Several secondary symbionts have been shown to form symbioses with *Acyrthosiphon pisum* (Fukatsu et al. 2000). When they occur, they occupy bacteriocytes, but these cells are mutually exclusive to the bacteriocytes containing *Buchnera aphidicola* (Fukatsu et al. 2000). In order to manage which bacteria are capable of entering the bacteriocyte, the aphid may have adapted to use phenoloxidase to create an inhospitable environment, allowing only the obligate symbiont and certain beneficial bacteria with a tolerance for oxidizing conditions.

There is some evidence for this based on the genomes of the secondary symbionts and the proteome of *Buchnera aphidicola*. Alkyl hydroperoxide reductase C (AhpC) is the second most abundant protein in the proteome of *Buchnera* (Poliakov et al., 2011). Alkyl hydroperoxide reductase has been shown, in *Escherichia coli*, to be the primary scavenger of hydrogen peroxide (Seaver and Imlay 2001); this is likely the case in *Buchnera* as well, since its genome is a subset of the *Escherichia coli* genome (Prickett et al. 2006). The high proteome ranking of AhpC, despite the reductive evolution that has occurred over the evolutionary history of *Buchnera* (Shigenobu et al. 2000), indicates a strong selection pressure for conservation, which could be explained by a highly oxidizing environment in the bacteriocyte. Homologs to AhpC can also be found via BLAST in the genomes of the aphid secondary symbionts *Regiella insecticola* and *Hamiltonella defensa* (unpublished data). Since hydrogen peroxide is a known byproduct of the enzymatic oxidation of DOPA by phenoloxidase (Komarov et al. 2005), it is possible that over the course of the evolution of symbiosis, endosymbionts of pea aphids have utilized the alkyl hydroperoxide reductase system to maintain symbiosis in the face of a restrictive bacteriocyte

environment created by the pea aphid. This theory is offset by the potential damage to the bacteriocyte cells themselves by phenoloxidase. It is, however, difficult to ignore the persistance of AhpC conservation through the drastic genome reductions that have occurred in the pea aphid's primary and secondary endosymbionts.

Works Cited

Asada, N., Fukumitsu, T., Fujimoto, K. and Masuda, K. (1993) Activation of prophenoloxidase with 2-propanol and other organic compounds in *Drosophila melanogaster*. Insect Biochemistry and Molecular Biology 23(4):515-520

Bidla, G., Lindgren, M., Theopold, U. and Dushay, M.S. (2005) Hemolymph coagulation and phenoloxidase in *Drosophila* larvae. Developmental and Comparative Immunology 29:669-679

Chen, C.Y., Lai, C.Y. and Kuo, M.H. (2009) Temperature effect on the growth of *Buchnera* endosymbiont in *Aphis craccivora*(Hemiptera: Aphididae). Symbiosis 49: 53-59

Douglas, A.E. and Dixon, A.F.G. (1987) The mycetocyte symbiosis in aphids: variation with age and morph in virginoparae of *Megoura viciae* and *Acyrthosiphon pisum*. *Journal of Insect Physiology* 33:109- 113

Douglas, A.E. and Prosser, W.A. (1992) Synthesis of the essential amino acid tryptophan in the pea aphid (*Acyrthosiphon pisum*) symbiosis. *Journal of Insect Physiology* 38:565-568

Fukatsu, T., Nikoh, N., Kawai, R. and Koga, R. (2000) The secondary endosymbiotic bacterium of the pea aphid *Acyrthosiphon pisum* (Insecta: Homoptera). Applied and Environmental Microbiology 66(7):2748-2758

Iwama, R. and Ashida, M. (1986) Biosynthesis of prophenoloxidase in hemocytes of larval hemolymph of the silkworm, *Bombyx mori*. Insect Biochemistry 16(3):547-555

Komarov, D.A., Slepneva, I.A., Glupov, V.V. and Khramtsov, V.V. (2005) Superoxide and hydrogen peroxide formation during enzymatic oxidation of DOPA by phenoloxidase. Free Radical Research 39(8):853-858

Lemaitre, B. (2004) The road to Toll. Nature Reviews Immunology 4:521-527

Montllor, C.B., Maxmen, A. and Purcell, A.H. (2002) Facultative bacterial endosymbionts benefit pea aphids *Acyrthosiphon pisum* under heat stress. Ecological Entomology 27:189-195

Nishikori, K., Morioka, K., Kubo, T. and Morioka, M. (2009) Age- and morph- dependent activation of the lysosomal system and *Buchnera* degradation in aphid endosymbiosis. Journal of Insect Physiology 55:351-357

Poliakov, A., Russell, C.W., Ponnala, L., Hoops, H.J., Sun, Q., Douglas, A.E. and van Wijk, K.J. (2011) Large-scale label-free quantitative proteomics of the pea aphid-*Buchnera* symbiosis. Molecular and Cellular Proteomics 10: M10.007039

Prickett, M.D., Page, M., Douglas, A.E. and Thomas, G.H. (2006) BuchneraBASE: a post-genomic resource for *Buchnera* sp. APS. *Bioinformatics* 22:641-642

Sadd, B.M. and Siva-Jothy, M.T. (2006) Self-harm caused by an insect's innate immunity. Proceedings of the Royal Society B 273(1600):2571-2574

Saul, S.J. and Sugumaran, M. (1987) Protease mediated phenoloxidase activation in the hemolymph of the tobacco hornworm, *Manduca sexta*. Archives of Insect Biochemistry and Physiology 5:1-11

Seaver, L.C. and Imlay, J.A. (2001) Alkyl hydroperoxide reductase is the primary scavenger of endogenous hydrogen peroxide in *Escherichia coli*. Journal of Bacteriology 183(24):7173-7181

Shigenobu, S., Watanabe, H., Hattori, M., Sakaki, Y. and Ishikawa, H. (2000) Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. APS. Nature 407:81-86

Swaminathan, C.P., Brown, P.H., Roychowdhury, A., Wang, Q., Guan, R., Silverman, N., Goldman, W.E., Boons, G.J. and Mariuzza, R.A. (2006) Dual strategies for peptidoglycan discrimination by peptidoglycan recognition proteins (PGRPs). PNAS 103(3):684-689

CHAPTER 3: PHYLOGENETICS OF PEA APHID PHENOLOXIDASE

Introduction

The lack of phenoloxidase activity as an immune response in the pea aphid brought into question how phenoloxidase was utilized in the aphids and whether they functioned similarly to other insect phenoloxidases. The active site of tyrosinase-type phenoloxidase has been thoroughly characterized (Jiang et al. 1997); it is split between two copper binding domains, each of which has three conserved histidine residues that participate in metal binding. The propeptide cleavage site is also highly conserved across insect lineages (Hall et al. 1995). In order to assess if all these conserved domains were present, the two pea aphid phenoloxidase genes identified in the genome annotation were included in a multiple alignment with other, better characterized phenoloxidases.

The aphid sequences were also used in a phylogenetic tree, to establish their relationship to immune system phenoloxidases from other organisms. Insect hexamerins, which are structurally similar to phenoloxidases, but without enzymatic activity (Terwilliger et al. 1999), were included as an outgroup. Because of the unusual evolutionary pattern of apparent lineage specific duplications revealed by the phylogenetic analyses of the present study, a comparison was also done between insect phenoloxidases and insect laccases from the same organisms. In addition to serving as another check for whether pea aphid phenoloxidases clustered with other insect phenoloxidases, the lacccase analysis was performed to see if the evolutionary pattern of the phenoloxidases extended to a family of genes with similar enzymatic activity but a different physiological role.

The interpretation of multiple, nearly identical duplications occurring independently in each insect order was problematic, partially because it did not match up with the lineage specific duplications in the phenoloxidase family that have independently occurred in parts of the Diptera (Christophides et al. 2002). We investigated the possibility of sequence coevolution between phenoloxidase copies from the same organism by a comparative analysis with the ecdysone receptor-ultraspiracle system, which is a heterodimer-forming system which has been shown to undergo coevolution in several insect lineages (Bonneton et al. 2003). While this comparison was unable to fully explain the evolution of insect phenoloxidases, it did suggest methodologies for further investigation.

Methods

Sequence and Phylogenetic Analysis

Phylogenetic analysis was performed on three groups of sequences: insect phenoloxidases (using hexamerins as outgroup), ecdysone receptor and ultraspiracle genes, and insect laccase-type cuticular phenoloxidases. Sequences of phenoloxidase and hexamerin proteins (Table 1), as well as laccase-type phenoloxidases (Table 2) were collected from GenBank. Sequences were selected from insect species with sequenced genomes, or species that have had the molecular properties of their phenoloxidases characterized. Ultraspiracle and ecdysone receptor sequences were taken from previous studies of ecdysone receptor evolution (Bonneton et al. 2003), with the addition of sequences from insect species with sequenced genomes that were not represented in the original study (Table 3).

Sequences were aligned with the ClustalW algorithm using default parameters in MEGA 4 (Tamura 2004). A Bayesian inference (BI) study was conducted for each group of sequences.

Sequence data was compiled and analyzed using MrBayes v.3.1.2 (Huelsenbeck and Ronquist 2001). An initial MrBayes analysis was performed for model finding. The Whelan and Goldman (WAG) fixed amino acid evolution model (Whelan and Goldman, 2001) was used for analysis of the insect phenoloxidases, the Jones model (Jones et al. 1992) was used for the analysis of the ecdysone receptor and ultraspiracle genes, and the BLOSUM model (Henikoff and Henikoff 1992) was used for the analysis of insect laccases. Analyses were initially run for 2,000,000 generations with sampling every 1000 generations. The first 25% of samples were discarded as burn-in. The runs were considered converged when average standard deviation of split frequencies was less than 0.01. Analysis of trees generated in MrBayes was performed using Tracer v.1.5. MrBayes runs with LnL values of less than 200 in Tracer were discarded and rerun with additional generations. Trees were drawn using FigTree v.1.3.1. Bayesian inference models were compared to bootstrapped Maximum Parsimony trees generated in MEGA 4.

Results

Characterization of Pea Aphid Phenoloxidases

Two sequences in the pea aphid genome were identified as phenoloxidase genes by sequence similarity to *Drosophila melanogaster* phenoloxidase subunit A3 (CG8193): LOC100160034 (703 amino acids, 80.65 kDA predicted molecular mass) and LOC100163393 (700 amino acids, 80.11 kDA predicted molecular mass). LOC10060034 is subsequently referred to as *A. pisum* phenoloxidase subunit A and LOC100163393 as *A. pisum* phenoloxidase subunit B. Further searching of the aphid genome annotation using LOC100160034 and

Table 1. Sequences Used in Phylogenetic Analyses of Phenoloxidases

All sequences were annotated phenoloxidases of insects with sequenced genomes, and the crustacean *Daphnia magna*. Four insect hexamerins, also annotated in insects with sequenced genomes, were used as an outgroup to distinguish enzymatically active phenoloxidases from similar protein families that contain the same domains. Sequences from *Manduca sexta* were also included in the analysis for being the most thoroughly annotated phenoloxidases.

Sequence	Species	Accession	Abbreviation
Phenoloxidase subunit 1 Precursor	*Bombyx mori*	NP_001037335	BMPPO1
Phenoloxidase subunit 2	*Bombyx mori*	NP_001037534	BMPPO2
Prophenoloxidase Subunit A1	*Drosophila melanogaster*	NP_476812	DMPPOA1
Prophenoloxidase Subunit A3	*Drosophila melanogaster*	Q9V521	DMPPOA3
Prophenoloxidase 1	*Anopheles gambiae*	XP_312089	AGPPO1
Prophenoloxidase 2	*Anopheles gambiae*	XP_316323	AGPPO2
Prophenoloxidase 3	*Anopheles gambiae*	XP_315073	AGPPO3
Prophenoloxidase 4	*Anopheles gambiae*	XP_315084	AGPPO4
Prophenoloxidase 5	*Anopheles gambiae*	XP_307623	AGPPO5
Prophenoloxidase 6	*Anopheles gambiae*	XP_315075	AGPPO6
Prophenoloxidase 7	*Anopheles gambiae*	CAD31059	AGPPO7
Prophenoloxidase 8	*Anopheles gambiae*	XP_315074	AGPPO8
Prophenoloxidase 9	*Anopheles gambiae*	XP_315076	AGPPO9
Phenoloxidase Subunit A3	*Apis mellifera*	NP_001011627	AMPPO
Prophenoloxidase Subunit 1	*Tribolium castaneum*	NP_001034493	TCPPO1
Prophenoloxidase Subunit 2	*Tribolium castaneum*	NP_001034522	TCPPO2
Prophenoloxidase	*Daphnia magna*		DAPHPPO
Similar to Prophenoloxidase (Subunit A)	*Acyrthosiphon pisum*	XP_001949307.1	APPPOA
Similar to Prophenoloxidase (Subunit B)	*Acyrthosiphon pisum*	XP_001951137.1	APPPOB
Chain A, Crystal Structure of Prophenoloxidase	*Manduca sexta*	3HHS_A	MSPPOA
Chain B, Crystal Structure of Prophenoloxidase	*Manduca sexta*	3HHS_B	MSPPOB
Hemocyanin Subunit F, putative	*Pediculus humanus*	XP_002429710	PHPPO
Hexamerin	*Apis mellifera*	ABR45905	AMHEX

Larval serum protein 1 beta subunit	*Drosophila melanogaster*	AAB58821	DMHEX
Hexamerin 2	*Tribolium castaneum*	NP_001164335	TCHEX
Hexamerin	*Anopheles gambiae*	AAA96405	AGHEX
Prophenoloxidase 1	*Aedes aegypti*	XP_001648968	AAPPO1
Prophenoloxidase 2	*Aedes aegypti*	XP_001661891	AAPPO2
Prophenoloxidase 3	*Aedes aegypti*	XP_001661890	AAPPO3

Table 2. Sequences Used in Phylogenetic Analyses of Laccase-type Phenoloxidases

All sequences were annotated laccases of insects with sequenced genomes, with *Daphnia pulex* as an outgroup (*Daphnia magna* laccases were insufficiently annotated). Laccases from *Nasonia vitripennis* were included since at the time of writing these genes were thoroughly annotated, while the phenoloxidases were not.

Sequence	Species	Accession	Abbreviation
Laccase-1-like Isoform 1	*Acyrthosiphon pisum*	XP_001948070.1	Lac1I1
Laccase-5-like	*Acyrthosiphon pisum*	XP_001950788.1	Lac5
Laccase-7-like	*Acyrthosiphon pisum*	XP_001946224.1	Lac7
Laccase 1	*Manduca sexta*	AAN17506.1	Lac1
Laccase 2	*Manduca sexta*	AAN17507.1	Lac2
Laccase 2	*Apis mellifera*	ACK57559.2	Lac2
Laccase 4	*Apis mellifera*	XP_393845.3	Lac4
Laccase-5-like	*Apis mellifera*	XP_625189.2	Lac5
Laccase-1-like	*Apis mellifera*	XP_001120790.2	Lac1
Laccase 1	*Tribolium castaneum*	NP_001034514.1	Lac1
Laccase 2 Isoform A	*Tribolium castaneum*	NP_001034487.2	Lac2I1
Laccase 2 Isoform B	*Tribolium castaneum*	AAX84203.2	Lac2I2
CG42345 Isoform A	*Drosophila melanogaster*	NP_724412.1	CG42345IA
CG42345 Isoform F	*Drosophila melanogaster*	NP_724413.2	CG42345IF
CG42345 Isoform E	*Drosophila melanogaster*	NP_610170.2	CG42345IE
CG42345 Isoform D	*Drosophila melanogaster*	NP_001137606.1	CG42345ID
CG3759 Isoform A	*Drosophila melanogaster*	NP_609287.3	CG3729IA
RE55660p	*Drosophila melanogaster*	AAO39486.1	RE55660p
CG32557	*Drosophila melanogaster*	NP_573249.1	CG32557I3
Laccase 1	*Anopheles gambiae*	AAN17505	Lac1
Laccase 2 Isoform A	*Anopheles gambiae*	AAX49501	Lac2IA
Laccase 2 Isoform B	*Anopheles gambiae*	AAX49502	Lac2IB
Laccase 3	*Anopheles gambiae*	ABQ95972	Lac3
Hypothetica protein DAPPUDRAFT 317026	*Daphnia pulex*	EFX81873	317026
Hypothetica protein DAPPUDRAFT 49503	*Daphnia pulex*	EFX81872	49503
Laccase-1-like	*Nasonia vitripennis*	XP_001604988	Lac1I1
Laccase-like	*Nasonia vitripennis*	XP_001603789	Lac1I2
Laccase-1-like	*Nasonia vitripennis*	XP_001599997	Lac1I3
Venom laccase	*Nasonia vitripennis*	NP_001155158	Venom
laccase-4-like	*Nasonia vitripennis*	XP_001605369	Lac1I4
L-ascorbate oxidase-like	*Nasonia vitripennis*	XP_001600222	Lac1I5
Laccase-2-like	*Nasonia vitripennis*	XP_001603034	Lac1I6

Laccase-like	*Nasonia vitripennis*	XP_001599970	Lac1I7
Laccase-like	*Nasonia vitripennis*	NP_001155159	Lac
Multicopper Oxidase	*Pediculus humanus*	XP_002423996	Lac1
Multicopper Oxidase	*Pediculus humanus*	XP_002422943.1	Lac2
Multicopper Oxidase	*Pediculus humanus*	XP_002423995.1	Lac3

Table 3. Sequences Used in Phylogenetic Analyses of Ecdysone Receptor and Ultraspiracle

All sequences were taken from Bonneton et al. 2003, or were annotated ultraspiracle or ecdysone receptor sequences from insects with sequenced genomes, with *Daphnia magna* as an outgroup.

Sequence	Species	Accession	Abbreviation
Ecdysone Receptor Isoform A1	*Daphnia magna*	BAF49029	ECRA1
Ecdysone Receptor Isoform A2	*Daphnia magna*	BAF49031	ECRA2
Ecdysone Receptor Isoform B	*Daphnia magna*	BAF49033	ECRB
Ecdysone Receptor	*Pediculus humanus*	XP_002430228	ECR
Ecdysone Receptor	*Locusta migratoria*	AAD19828	ECR
Ecdysone Receptor Isoform B1	*Apis mellifera*	NP_001152827	ECR
Ecdysone Receptor Isoform A	*Apis mellifera*	NP_001091685	ECRA
Ecdysone Receptor	*Acyrthosiphon pisum*	ACR45971	ECR
Ecdysone Receptor Isoform A	*Acyrthosiphon pisum*	NP_001152831	ECRA
Ecdysone Receptor Isoform B1	*Acyrthosiphon pisum*	NP_001152832	ECRB1
Ecdysone Receptor	*Tenebrio molitor*	CAA72296	ECR
Ecdysone Receptor Isoform B	*Drosophila melanogaster*	NP_724460	ECR
Ecdysteroid Receptor	*Aedes albopictus*	AAF19032	ECR
Ecdysone Receptor	*Aedes aegypti*	P49880	ECR
Ecdysone Receptor	*Chironomus tentans*	P49882	ECR
Ecdysone Receptor Isoform A	*Choristoneura fumiferana*	AAC61596	ECR
Ecdysone Receptor Isoform B1	*Bombyx mori*	NP_001166846	ECR
Ecdysone Receptor	*Manduca sexta*	P49883	ECR
Retinoid X Receptor-like protein	*Daphnia magna*	ABF74729	USP
USP Protein	*Tenebrio molitor*	CAB75361	USP
Retinoid X Receptor	*Pediculus humanus*	XP_002424949	RXR
RXR	*Locusta migratoria*	AAF00981	RXR
Ultraspiracle	*Apis mellifera*	NP_001011634	USP
Ultraspiracle	*Acyrthosiphon pisum*	NP_001155140	USP
Ultraspiracle	*Drosophila melanogaster*	NP_476781	USP
Ultraspiracle	*Aedes albopictus*	AAF19033	USP
Ultraspiracle Isoform-A	*Aedes aegypti*	AAG24886	USPA
Ultraspiracle	*Chironomus tentans*	AAC03056	USP
Ultraspiracle Homolog	*Choristoneura fumiferana*	O76202	USP
Utraspiracle Homolog	*Bombyx mori*	NP_001037470	USP
Ultraspiracle Homolog	*Manduca sexta*	P54779	USP

LOC100163393 and other insect phenoloxidase sequences as the query sequence did not yield additional pea aphid sequences with e-value≤0.05. The candidate pea aphid phenoloxidase protein sequences had 83% sequence identity, compared to 47.4% sequence identity between chain A and chain B of *Manduca sexta* prophenoloxidase. Both aphid sequences included three hemocyanin domains (pf00372 [Hemocyanin_M], pf003722 [Hemoncyanin_N], and pf003723 [Hemocyanin_C]) (Figure 4), all of which are structural hallmarks of functional phenoloxidases (Jones et. al. 1990; Fujimoto et. al. 1995).

Phylogenetic Analysis of Insect Phenoloxidases

The evolutionary relationship between *A. pisum* phenoloxidase A, *A. pisum* phenoloxidase B and phenoloxidase genes in other insects was investigated, with the crustacean *Daphnia magna* as an evolutionary outgroup, and several insect hexamerins as a structural outgroup (Figure 5A). Maximum likelihood and Bayesian models agreed on tree topology (Figure 5B), so Bayesian trees were used for the analysis. The phenoloxidases form a coherent clade, which is distinct from the structurally similar hexamerin proteins. With the exception of *Apis mellifera* phenoloxidase, the topology of the tree correlates with the consensus phylogeny of insect diversification. The two pea aphid phenoloxidase proteins are a lineage-specific duplication, as is also evident for *Tribolium castaneum* (Coleoptera). The clustering of phenoloxidase genes in the two Lepidoptera, *Bombyx mori* and *Manduca sexta*, and Diptera, *Drosophila melanogaster* and *Anopheles gambiae*, suggests that the lineage-specific phenoloxidase duplications occurred at the order level in these groups. The closer sequence identity of *A. pisum* phenoloxidase A than to phenoloxidases of other insects compared to *A. pisum* phenoloxidase B (Table 4) indicates that the latter has diverged more rapidly after the putative duplication event.

```
APPPO_LOC100160034  MTDKNNILYL FDRPTEPIFI GKGDDNVSFE VPAEYLTDRY KPLASDIQNR
APPPO_LOC100163393  MADKNNILYL FDRPTEPIFI GKGEENVSFD VPTDYLIDRY KPLASDIQTR
MSPPOA              ADIFDSFELL YDRPGEPMIN TKGEDKVLFE LTEQFLTPEY ANNGLELNNR
MSPPOB              TDAKNNLLYF FDRPNEPCFM QKGEDKVVFE IPDHYYPDKY KSLSNTLSNR

APPPO_LOC100160034  FS----SGKT ISVTKLSSIP DMSFPLQLGR DKAFSLFIPY HSKMAAKLIE
APPPO_LOC100163393  FP----GGKT VPVTRLSSIP DLSIPLGLKR DMPFSLFNQS HGKMAAKLIE
MSPPOA              FGDEEEVSRK IILKNLDKIP EFPKAKQLPN DADFSLFLPS HQEMANEVID
MSPPOB              FGNE--ATKR IPIRNIT-LP NLEVPMQLPY NDQFSLFVPK HRTMAAKLID

APPPO_LOC100160034  IFMASKT--F DELLSLAVYA RDRVNPYMFI YALSVVVTHR PDTRNLELPS
APPPO_LOC100163393  ILMNAKS--Y DELLSLSVYC RDRINPYMFT YALSVALIHR PDTRNLRLPS
MSPPOA              VLMSVTENQL QELLSTCVYA RINLNPQLFN YCYTVAIMHR RDTGKVRVQN
MSPPOB              IFMGMRD--V EDLQSVCSYC QLRINPYMFN YCLSVAILHR PDTKGLSIPT

APPPO_LOC100160034  HVEMFPSLYM DATVFGRARE ESAVVQTGS- RTPIEIPHDY SANDLDFEHR
APPPO_LOC100163393  HSEMFPSLYM DSSVFARARE ESAVVQTGS- RTPIEIPHDY SANNLDAEHR
MSPPOA              YAEIFPAKFL DSQVFTQARE AAAVIPKTIP RTPIIIPRDY TATDLEEEHR
MSPPOB              FAETFPDKFM DSKVFLRARE VSNVVISGS- RMPVNVPINY TANTTEPEQR

APPPO_LOC100160034  ISYFREDIGV NLHHWHWHLV YPFDGP-VDI VNKDRRGELF YYMHEQILAR
APPPO_LOC100163393  ISYFREDIGI NLHHWHWHLV YPFDGP-LNI VNKNRRGELF FYMHQQIIAR
MSPPOA              LAYWREDLGI NLHHWHWHLV YPFSASDEKI VAKDRRGELF FYMHQQIIAR
MSPPOB              VAYFREDIGI NLHHWHWHLV YPFDSADRSI VNKDRRGELF YYMHQQIIGR
                              *    *                                  *

APPPO_LOC100160034  YNMERLSNDM NRVVRLTNWR SPILEGYFPK LDNILANRVW PSRPVNATLS
APPPO_LOC100163393  YNMERLSNNM NRVVRLTNWD QPIAEGYFPK LDNILANRVW PPRPVNAVLQ
MSPPOA              YNCERLCNSL KRVKKFSDWR EPIPEAYYPK LDSLTSARGW PPRQAGMRWQ
MSPPOB              YNVERMCNGL PQVKPFSDFS APIEEGYFPK LDSQVASRTW PPRFAGSVFR

APPPO_LOC100160034  NINREIEQIS FDIEDLERWR DRIFNAIHSG FIINTAGQQV RLTEADGINI
APPPO_LOC100163393  NISREVEQIT FDIEDLVRWR DRIFNAIHSG FIINTAGQQV RLTETDGIDI
MSPPOA              DLKRPVDGLN VTIDDMERYR RNIEEAIATG NVILPD---- KSTKKLDIDM
MSPPOB              NLDRTVDQVK IDVRKLFTWR DQFLEAIQKM AIKMPNGREL PLDEVTGIDM

APPPO_LOC100160034  LGNLIEASIL SLNQNLYGSL HNNGHNAISF IHDPDNRFLE NYGVMGDSAT
APPPO_LOC100163393  LGNIIESSIL SQNPNLYGSL HNNGHNAIAY IHDPDNRFLE NYSVMGDSAT
MSPPOA              LGNMMEASVL SPNRDLYGSI HNNMHSFSAY MHDPEHRYLE SFGVIADEAT
MSPPOB              LGNLMESSII SPNRGYYGDL HNMGHVFAAY THDPDHRHLE QFGVMGDSAT
                                       *          *

APPPO_LOC100160034  AMRDPIFYRW HAYIDDIFQE FKAT---IPS YTIQNLSFDN VRVQSVEISA
APPPO_LOC100163393  AMRDPIFYRW HAYIDDIFQE YKAT---IPS YNVQNLGFDN VSVQSVEVTA
MSPPOA              TMRDPFFYRV HAWVDDIFQS FKEAPHNVRP YSRSQLENPG VQVTSVAVES
MSPPOB              AMRDPFFYRW HRFVDDVFNI YKEK---LTP YTNERLDFPG VRVSSVGIEG
                            *

APPPO_LOC100160034  TGIPRNELAT FWQQSDVDLS RGLDFLPRGS VFARFTHLQH APFNYKITVE
APPPO_LOC100163393  TGLPRNEFAT FWQQSDTDLS RGLDFLPRGS VFARFTHLQH APFNYKIIVE
MSPPOA              AGGQQNVLNT FWMQSDVNLS KGLDFSDRGP VYARFTHLNH RPFRYVIKAN
MSPPOB              ARP--NTLRT LWQQSTVELG RGLDFTPRGS VLARFTHLQH DEFQYVIEVN

APPPO_LOC100160034  NN-GNQRLGT VRIFIAPKYD ERGLPFLFRE QRKLMVELDK FSVTLTRGRN
APPPO_LOC100163393  NN-GNQRIGT VRIFLAPKFD ERGLPFLFRE QRKLFVELDK FSTSLKRGRN
MSPPOA              NT-ASARRTT VRIFIAPKTD ERNLPWALSD QRKMFIEMDR FVVPLSAGEN
MSPPOB              NTTGGNLMGT VRIFMAPKVD DNGQPMSFNK QRRLMIELDK FSQALRPGTN

APPPO_LOC100160034  EITRRSIESS VTIPHEITYR NLDRNRPANN SDAAAAFNFC GCGWPQNMLI
APPPO_LOC100163393  EIVRRSIESS VTIPHEITYR NQGSNRPAAN SDAAPMFNFC GCGWPQNMLI
MSPPOA              TITRQSTESS LTIPFEQTFR DLSIQGSDPR RSELAAFNYC GCGWPQHMLV
MSPPOB              TIRRRSVDSS VTIPYERTFR NQSERPGDPG TAGAAEFDFC GCGWPHHMLI

APPPO_LOC100160034  PKGTAEGFQC QLFVMISNGA NDQVENAQAD GQTCDNASSY CGIRNSRYPD
```

```
APPPO_LOC100163393  AKGSPEGFQC QLFVMVSNGE IDQVANAQGD GQTCDDASSY CGILNSRYPD
MSPPOA               PKGTVGGVAY QLFVMLSNYE LDKIEQPDGR ELSCVEASMF CGLKDKKYPD
MSPPOB               PKGTAQGYPV VLFVMISNWN NDRIEQDLVG --SCNDAASY CGIRDRKYPD

APPPO_LOC100160034  ARSMGYPFDR TPRDGVVTLQ QFLTP-NMVV QDVRIRFSNR TVAPLQNRIG
APPPO_LOC100163393  SRSMGYPFDR TPRDGVVTLQ QFLTT-NMVV QDVRIRFSNR TVAPLQNATA
MSPPOA               ARPMGYPFDR -PSNSATNIE DFSAMSNMGL QDIVIKLSDV TEPNPRNPPA
MSPPOB               KQAMGYPFDR KMANDAATLS DFLRP-NMAV RDCSIQFSDT TVERGQQG--

APPPO_LOC100160034  SQQTSKNPPA KAPGRN
APPPO_LOC100163393  NRNAG---TS NNNKRN
MSPPOA               ---------- ------
MSPPOB               ---------- ------
```

Figure 4. Deduced amino acid sequence of *A. pisum* LOC100160034 and LOC100163393. MEGA 4 was used to calculate sequence identities. The copper-binding domains, contained in the Hemocyanin_N domain (pf003722) are underlined with asterisks marking the conserved histidine residues involved in copper ion binding. The Hemocyanin_N domain (pf003722) contains the serine protease cleavage site (RF) for the propeptide, which is indicated in bold at residues 50 and 51.

Phylogenetic Analysis of Laccase Type Phenoloxidases

To determine if the putative lineage specific duplications of phenoloxidase on the level of order is also obtained for laccases, a phylogenetic analysis was performed on sequences of laccase type phenoloxidases. There was a high level of similarity between the Bayesian and Maximum Likelihood tree topologies (Figure 6A-B). Unlike the phenoloxidases in the previous tree, the laccase type phenoloxidases do not cluster primarily by order, and the lineage specific duplications that are present clearly represent family expansions, as is the case with several of the *Nasonia vitripennis* laccases. When the laccase sequences were analyzed alongside the insect phenoloxidases, they segregated to different branches on the tree in a fashion that suggested an ancestral split between laccases and phenoloxidase before the Daphnia lineage (Figure 7). This shows convincingly that despite a high similarity in enzymatic activity, phenoloxidases and laccases have distinct evolutionary trajectories, presumably due to differences in functional uses in the insects.

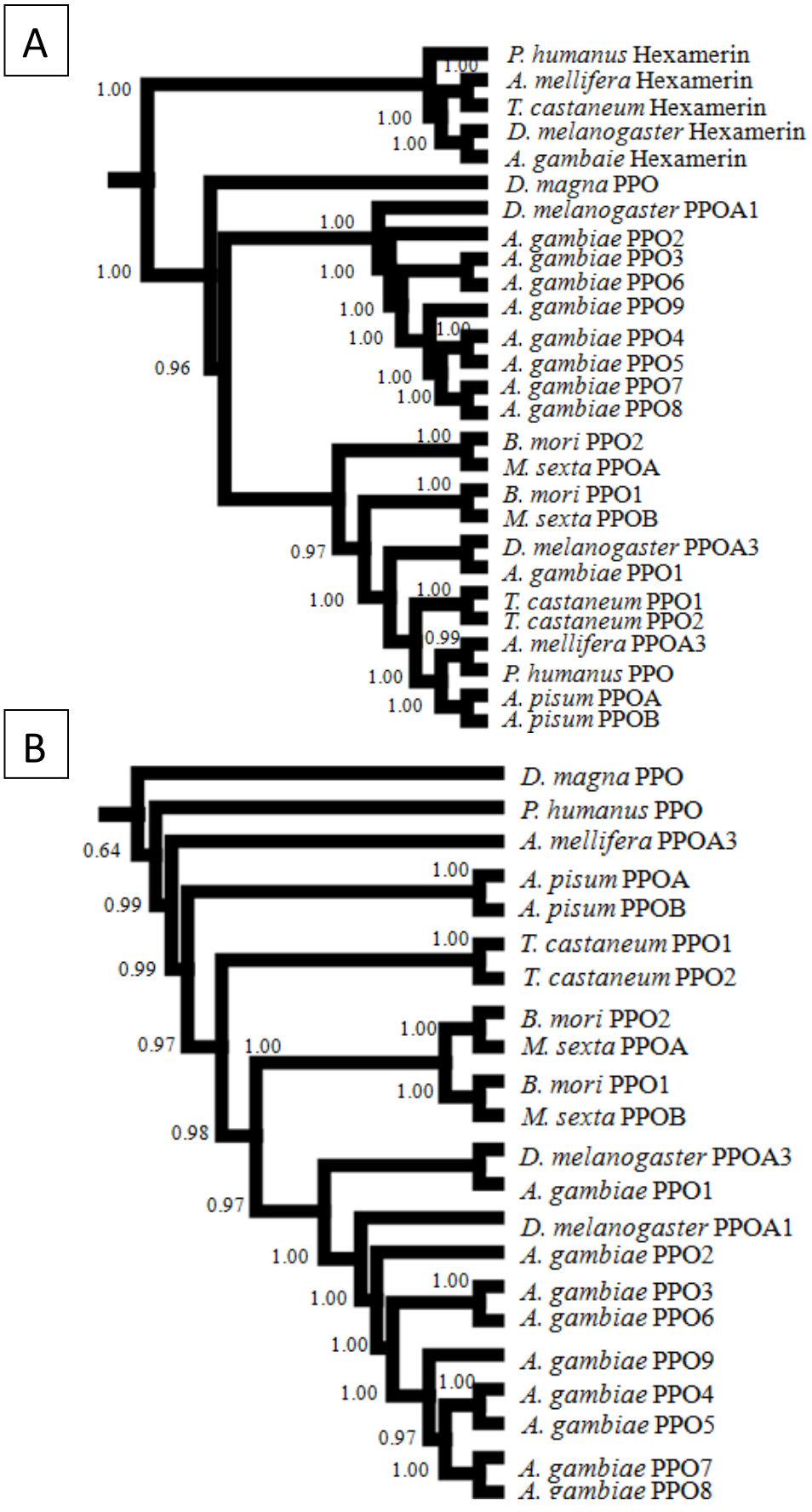
Phylogenetic Analysis of Ecdysone Receptor (EcR) and Ultraspiracle (USP)

Bonneton et al (2003) demonstrated coevolution between the ecdysone receptor and ultraspiracle genes in the Diptera and Lepidoptera lineages. Using additional sequences from the species used in the phenoloxidase studies, the phylogenetic trees for ultraspiracle (Figure 8A-B) and ecdysone receptor (Figure 8C-D) were recreated. When these sequences were combined into a single tree, the topology showed a separation of ultraspiracle and ecdysone receptor sequences from the same organism (Figure 8E). Sequence identity between ultraspiracle and ecdysone receptor is much lower than between phenoloxidase subunits; *Acyrthosiphon pisum* only has a

Table 4. Reciprocal BLAST results between *Acyrthosiphon pisum* putative prophenoloxidases and prophenoloxidases from other organisms

Organism	Top Hit for LOC100163393	BLAST score	Top Hit Against *A. pisum*	BLAST score
D. melanogaster	Phenoloxidase Subunit A3 (CG8193)	817	LOC10060034	830
M. sexta	Phenoloxidase Subunit 1 (O44249.3)	721	LOC10060034	741
T. molitor	Prophenoloxidase (BAA75470.1)	851	LOC10060034	879
A. mellifera	Phenoloxidase Subunit A3 (NP001011627.1)	865	LOC10060034	900
A. gambiae	Prophenoloxidase (XP312089)	796	LOC10060034	826

Organism	Top Hit for LOC10060034	BLAST score
D. melanogaster	Phenoloxidase Subunit A3 (CG8193)	830
M. sexta	Phenoloxidase Subunit 1 (O44249.3)	741
T. molitor	Prophenoloxidase (BAA75470.1)	879
A. mellifera	Phenoloxidase Subunit A3 (NP001011627.1)	900
A. gambiae	Prophenoloxidase (XP312089)	826


A
P. humanus Hexamerin
A. mellifera Hexamerin
T. castaneum Hexamerin
D. melanogaster Hexamerin
A. gambaie Hexamerin
D. magna PPO
D. melanogaster PPOA1
A. gambiae PPO2
A. gambiae PPO3
A. gambiae PPO6
A. gambiae PPO9
A. gambiae PPO4
A. gambiae PPO5
A. gambiae PPO7
A. gambiae PPO8
B. mori PPO2
M. sexta PPOA
B. mori PPO1
M. sexta PPOB
D. melanogaster PPOA3
A. gambiae PPO1
T. castaneum PPO1
T. castaneum PPO2
A. mellifera PPOA3
P. humanus PPO
A. pisum PPOA
A. pisum PPOB
B
D. magna PPO
P. humanus PPO
A. mellifera PPOA3
A. pisum PPOA
A. pisum PPOB
T. castaneum PPO1
T. castaneum PPO2
B. mori PPO2
M. sexta PPOA
B. mori PPO1
M. sexta PPOB
D. melanogaster PPOA3
A. gambiae PPO1
D. melanogaster PPOA1
A. gambiae PPO2
A. gambiae PPO3
A. gambiae PPO6
A. gambiae PPO9
A. gambiae PPO4
A. gambiae PPO5
A. gambiae PPO7
A. gambiae PPO8

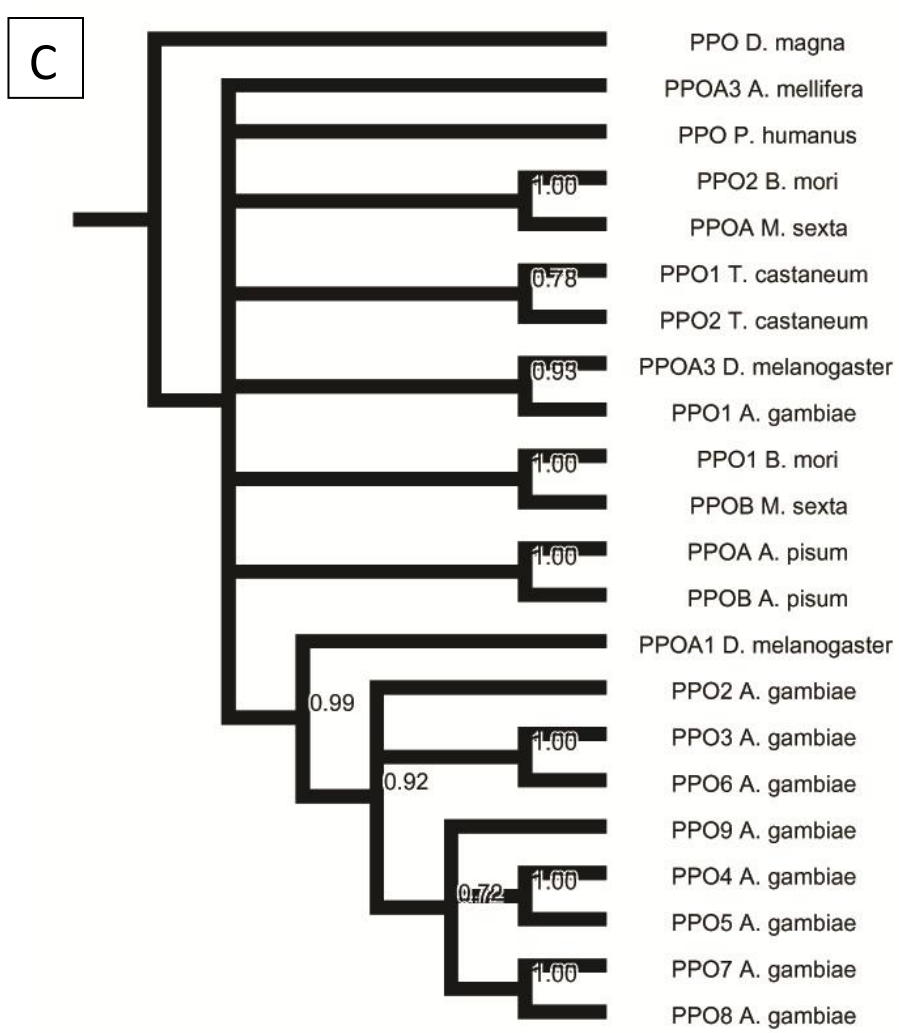
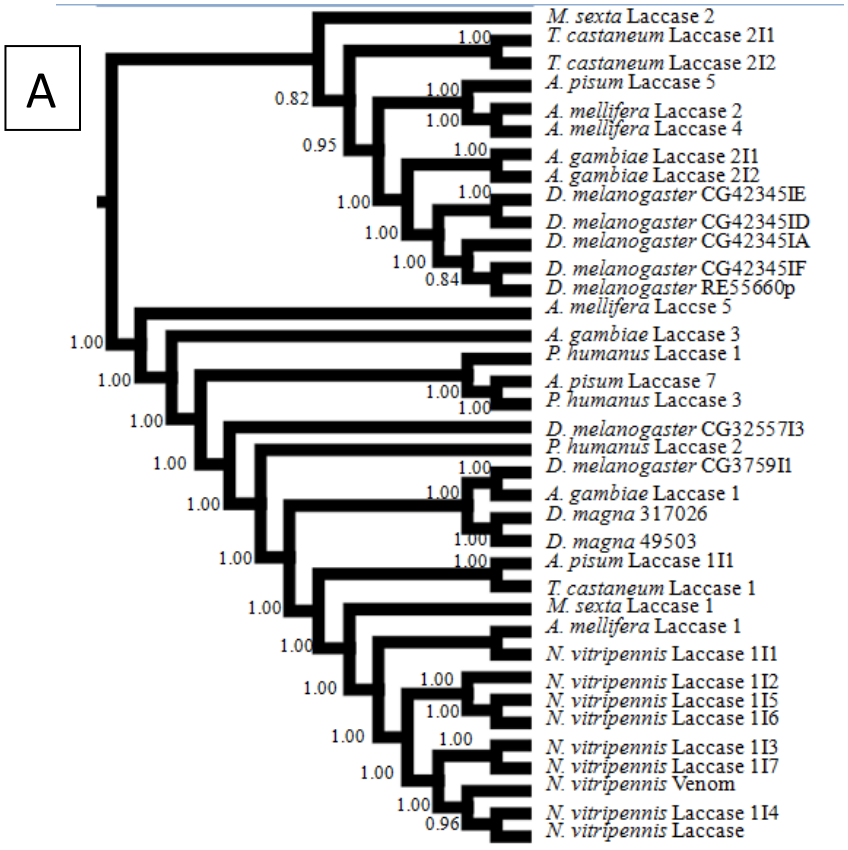


Figure 5. Phylogenetic tree for insect phenoloxidases. Bayesian (A-B) and maximum likelihood (C) trees are presented. For Bayesian trees, branches with a posterior probability of less than 0.80 were collapsed. The Bayesian tree with (A) and without (B) hexamerins is presented. For Maximum Parsimony trees, branches with bootstrap support of less than 0.50 were collapsed. Trees are rooted on the hexamerin clade, when present, and otherwise on *Daphnia magna* prophenoloxidase.


A
M. sexta Laccase 2
T. castaneum Laccase 2I1
T. castaneum Laccase 2I2
A. pisum Laccase 5
A. mellifera Laccase 2
A. mellifera Laccase 4
A. gambiae Laccase 2I1
A. gambiae Laccase 2I2
D. melanogaster CG42345IE
D. melanogaster CG42345ID
D. melanogaster CG42345IA
D. melanogaster CG42345IF
D. melanogaster RE55660p
A. mellifera Laccse 5
A. gambiae Laccase 3
P. humanus Laccase 1
A. pisum Laccase 7
P. humanus Laccase 3
D. melanogaster CG32557I3
P. humanus Laccase 2
D. melanogaster CG3759I1
A. gambiae Laccase 1
D. magna 317026
D. magna 49503
A. pisum Laccase 1I1
T. castaneum Laccase 1
M. sexta Laccase 1
A. mellifera Laccase 1
N. vitripennis Laccase 1I1
N. vitripennis Laccase 1I2
N. vitripennis Laccase 1I5
N. vitripennis Laccase 1I6
N. vitripennis Laccase 1I3
N. vitripennis Laccase 1I7
N. vitripennis Venom
N. vitripennis Laccase 1I4
N. vitripennis Laccase

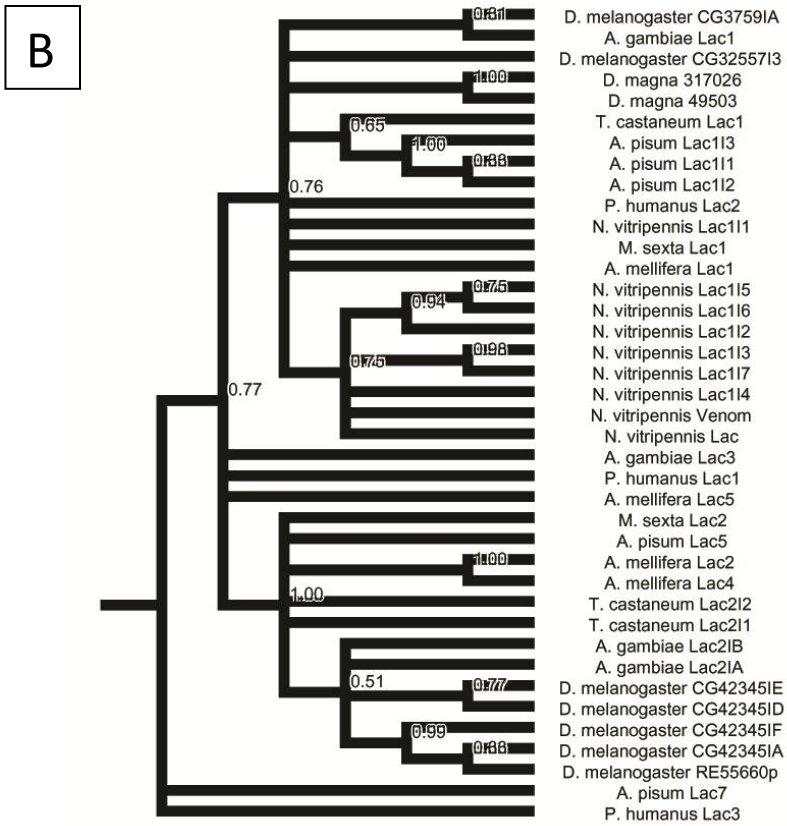

B
D. melanogaster CG3759IA
A. gambiae Lac1
D. melanogaster CG32557I3
D. magna 317026
D. magna 49503
T. castaneum Lac1
A. pisum Lac1I3
A. pisum Lac1I1
A. pisum Lac1I2
P. humanus Lac2
N. vitripennis Lac1I1
M. sexta Lac1
A. mellifera Lac1
N. vitripennis Lac1I5
N. vitripennis Lac1I6
N. vitripennis Lac1I2
N. vitripennis Lac1I3
N. vitripennis Lac1I7
N. vitripennis Lac1I4
N. vitripennis Venom
N. vitripennis Lac
A. gambiae Lac3
P. humanus Lac1
A. mellifera Lac5
M. sexta Lac2
A. pisum Lac5
A. mellifera Lac2
A. mellifera Lac4
T. castaneum Lac2I2
T. castaneum Lac2I1
A. gambiae Lac2IB
A. gambiae Lac2IA
D. melanogaster CG42345IE
D. melanogaster CG42345ID
D. melanogaster CG42345IF
D. melanogaster CG42345IA
D. melanogaster RE55660p
A. pisum Lac7
P. humanus Lac3

Figure 6. Phylogenetic tree for insect laccases. Bayesian (A) and maximum likelihood (B) trees are presented. For Bayesian trees, branches with a posterior probability of less than 0.80 were collapsed. For Maximum Parsimony trees, branches with bootstrap support of less than 0.50 were collapsed. Trees are rooted based on the point of divergence between phenoloxidases and laccases, as shown in Figure 7.

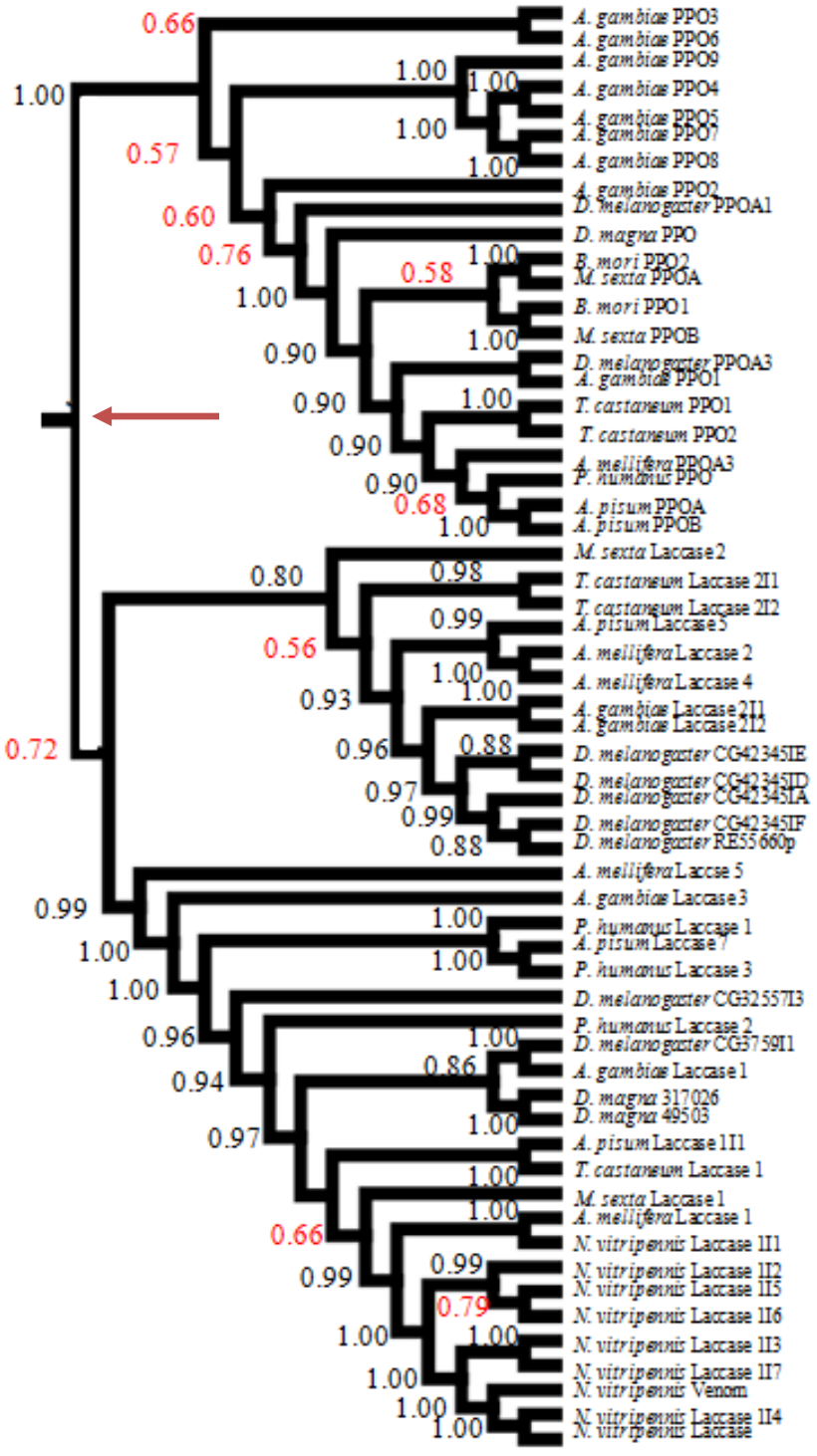
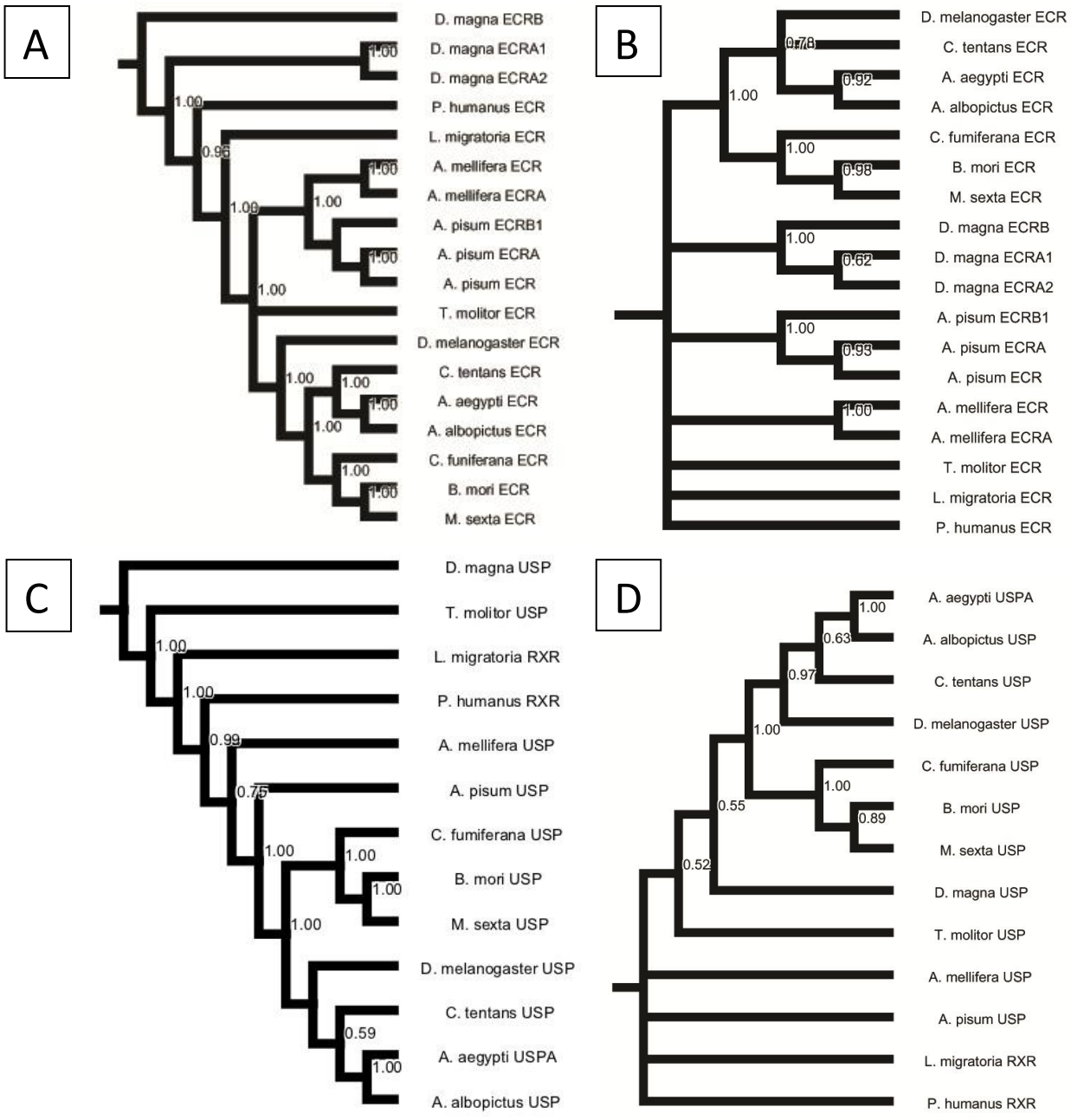


Figure 7. Phylogenetic tree for insect laccases and phenoloxidases. Branches with a posterior probability of less than 0.80 have their probabilities highlighted. The branch indicating the split between the phenoloxidase and laccase lineages is indicated with a red arrow.


A
D. magna ECRB
D. magna ECRA1
D. magna ECRA2
P. humanus ECR
L. migratoria ECR
A. mellifera ECR
A. mellifera ECRA
A. pisum ECRB1
A. pisum ECRA
A. pisum ECR
T. molitor ECR
D. melanogaster ECR
C. tentans ECR
A. aegypti ECR
A. albopictus ECR
C. funiferana ECR
B. mori ECR
M. sexta ECR
B
D. melanogaster ECR
C. tentans ECR
A. aegypti ECR
A. albopictus ECR
C. fumiferana ECR
B. mori ECR
M. sexta ECR
D. magna ECRB
D. magna ECRA1
D. magna ECRA2
A. pisum ECRB1
A. pisum ECRA
A. pisum ECR
A. mellifera ECR
A. mellifera ECRA
T. molitor ECR
L. migratoria ECR
P. humanus ECR
C
D. magna USP
T. molitor USP
L. migratoria RXR
P. humanus RXR
A. mellifera USP
A. pisum USP
C. fumiferana USP
B. mori USP
M. sexta USP
D. melanogaster USP
C. tentans USP
A. aegypti USPA
A. albopictus USP
D
A. aegypti USPA
A. albopictus USP
C. tentans USP
D. melanogaster USP
C. fumiferana USP
B. mori USP
M. sexta USP
D. magna USP
T. molitor USP
A. mellifera USP
A. pisum USP
L. migratoria RXR
P. humanus RXR

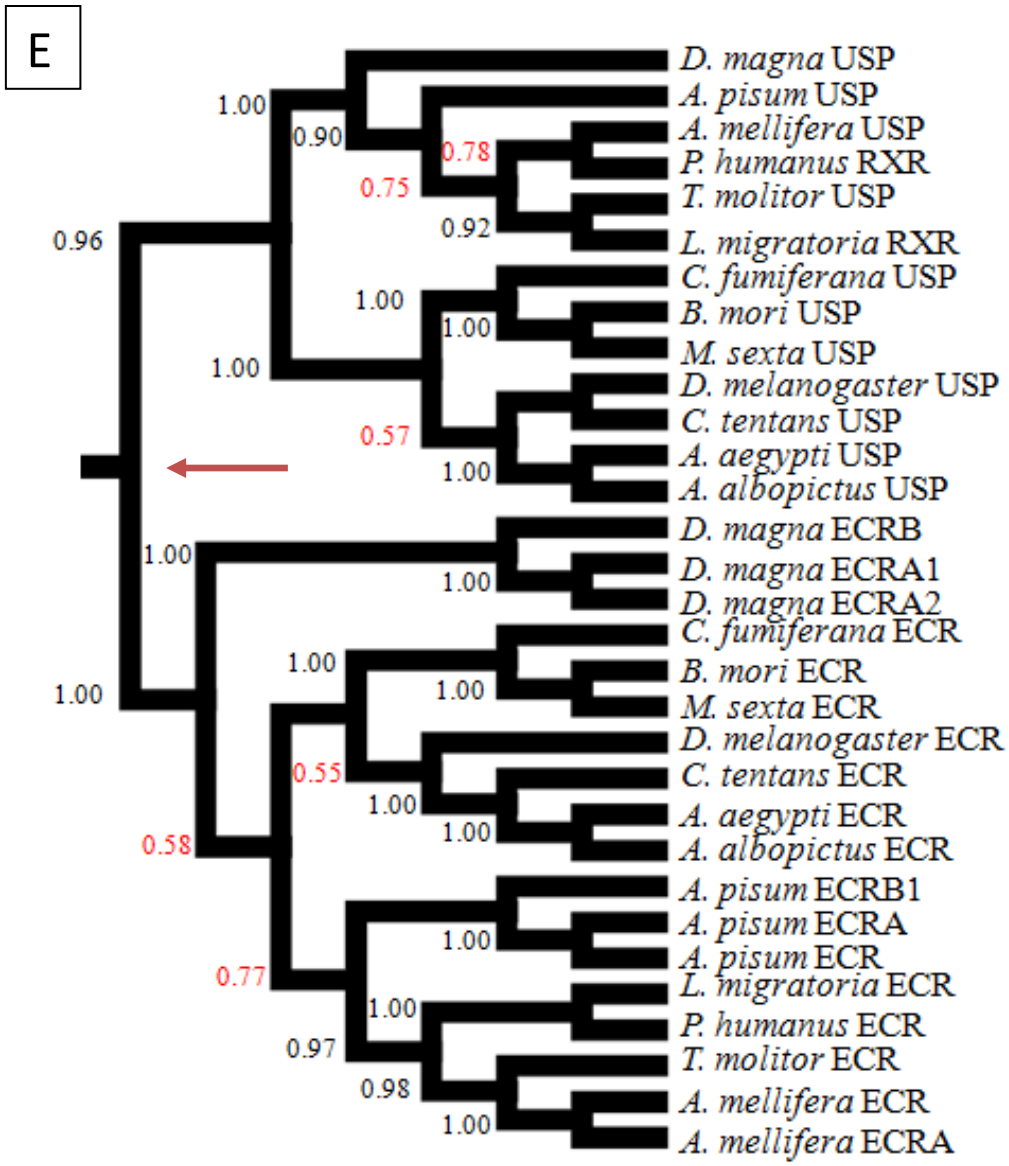


Figure 8. Phylogenetic tree for ultraspiracle (A, B), ecdysone receptor (C, D) and the two together (E). For Bayesian trees (A, C, E), nodes with a posterior probability of less than 0.80 are collapsed, or highlighted in red for the combined ECR-USP tree. For Maximum Parsimony trees (B, D), branches with bootstrap support of less than 0.50 were collapsed. Unlike Bonneton et al, the entire sequence of both ultraspiracle and ecdysone receptor was used for this analysis to allow for simultaneous comparison. An arrow indicates the branch separating ultraspiracle and ecdysone receptor sequences.

31% sequence identity between EcR and USP, while there is 85% sequence identity between the *A. pisum* phenoloxidase subunits. The separation of ultraspiracle and ecdysone receptor sequences to different branches of the tree despite a history of coevolution, as well as the levels of divergence between sequence identities, suggest that coevolutionary processes are not the driving evolutionary force in the phenoloxidase system. As a result, the suggestion of multiple lineage specific duplications seems to be the most valid to explain the diversification of insect phenoloxidases.

Discussion

A striking property of the phylogeny of insect phenoloxidases is that they have a topology consistent with lineage specific duplications. It has been previously demonstrated that phenoloxidase, as an innate immune system component, can undergo rapid evolution in response to natural enemies in *Daphnia magna* (Pauwels et al. 2010). However, the pattern seen in the insects is different than that seen in crustaceans, where an ancestral duplication of phenoloxidase has lead to several classes of enzyme with phenoloxidase activity that are utilized in innate immunity (Terwilliger and Ryan 2006). Since *Daphnia magna* is an evolutionary outgroup both to the insect phenoloxidases and the crustacean phenoloxidases used in the Terwilliger and Ryan analysis, and *Daphnia manga* has only a single phenoloxidase subunit, it is likely that all duplications occurred after the divergence of the crustacean lineage from *D. magna*.

Attempts to find other systems where this repeated lineage specific duplication occurred were unsuccessful, with the partial exception of the ecdysone receptor and ultraspiracle system discussed later in this chapter. The comparison between immune system phenoloxidases and

laccase type phenoloxidases demonstrated that this unusual evolutionary pattern is limited to the former branch of the phenoloxidase gene family. This means that insect immune system phenoloxidases have undergone a duplication, but instead of an ancestral duplication leading to a clade of "A" subunits and "B" subunits, the duplication has occurred several times independently, resulting in systems that are startlingly similar to each other in different insects.

The idea that phenoloxidases duplicated independently in each insect order, however unlikely, is the most parsimonious interpretation of the insect phenoloxidase tree. Evidence from another lineage specific duplication of phenoloxidase, specific to *Anopheles gambiae,* offers insight into when these events may have occurred. Prophenoloxidases 3-9 (PPO3-PPO9) appear to be duplications of prophenoloxidase 2 (PPO2), and are not closer in sequence similarity to *Drosophila melanogaster* prophenoloxidase A1 (PPOA1) than *A. gambiae* PPO2. *A. gambiae* PPO1, though, appears to be most similar to *D. melanogaster* PPOA3, very much like the duplications in the other orders. This suggests that these were two separate duplication events in *Anopheles gambiae,* with the larger expansion occurring after the duplication resulting in PPO1 and PPO2, and with all subsequent phenoloxidases arising from PPO2. This results in the duplication between PPO1 and PPO2 in the Diptera appearing much older than the lineage specific duplications in the other insects, a result that is likely caused by a lack of sequences to properly fill out the tree and determine precise timing of the duplications. However, along with the date from the Lepidoptera it does fairly conclusively identify that these duplications occurred on the level of phylogenetic order

While the lineage specific duplications are nearly ubiquitous throughout the phylogeny, there are several lineages which only possess a single phenoloxidase. *Pediculus humanus* is the most ancestrally diverged organism on the tree other than *Daphnia magna,* and because of its

position it is tempting to say that whatever conditions promoted the duplications in other lineages may not have applied to the Phthiraptera because of its basal position in the evolution of insects. However, the single phenoloxidase subunit in *Apis mellifera* is much more difficult to explain. Despite being annotated as prophenoloxidase A3, there does not appear to be an A1 subunit, since the name was based off of BLAST similarity to *Drosophila melanogaster* prophenoloxidase A3 (Lourenco et al. 2005). The same paper confirms that it is the only prophenoloxidase in *Apis mellifera*. The phylogenetic analysis shows it clustering closely to *P. humanus*, and both sequences being close to the *D. magna* root. This indicates that both *A. mellifera* and *P. humanus* phenoloxidases, and perhaps all organisms with a single copy of phenoloxidase, may have a drastically different evolutionary history than insect phenoloxidases from orders that have undergone lineage specific duplictions.

One potential alternative explanation for the lineage specific duplications is that there was actually an ancestral duplication that is somehow being masked by the evolutionary processes involved. Concerted evolution could be acting to keep within-order sequence identity high, leading to apparent lineage specific duplication. However, this is unlikely, since it is not suggested by the positioning of the positioning of the genes on their respective scaffolds in *A. pisum*, and insect phenoloxidases lack of repetitive elements in their sequence. Another intriguing possibility is that the necessity for dimer formation to achieve phenoloxidase activity in insects has resulted in coevolution, and that this somehow has contributed to the evolutionary pattern witnessed in these studies. Of particular note is the matter of dimer preference in insect phenoloxidase. For example, *Manduca sexta* phenoloxidase has been shown to preferentially form a heterodimer, while *Drosophila melanogaster* instead forms either of its homodimers. While no difference in activity has yet been demonstrated within *M. sexta* or *D. melanogaster*

between the heterodimer and either homodimer, the physiological constraint of creating only a specific dimeric form suggests that an organism's dimer preference may be related to the evolution of its phenoloxidases. A possibility which would explain both this preference in dimer form and the unusual evolutionary pattern would be a pattern of coevolution between phenoloxidase subunits from the same organism. Mutations affecting the residues at the protein-protein interacting regions could drive dimer stability, and thus constrain sequence diversification. If, however, phenoloxidases dimers from different organisms, or even different dimers from the same organism, have different protein-protein interacting regions, the dimer form that is stabilized by this process may vary from organism to organism, and may lead to strong sequence identity for a given region within an organism compared to random diversification in all other organisms.

To investigate this hypothesis, we sought evidence from another dimer-forming system which has shown evidence of coevolution: the ecdysone receptor-ultraspiracle system. Bonneton et al. (2003) demonstrated that ligand binding domain of ecdysone receptor had a significant increase in substitution rate in the Diptera and Lepidoptera lineages. This is a domain that is involved in heterodimer formation, and thus would have evolved rapidly to maintain interactions with ultraspiracle, whose copies in the aforementioned lineages are highly divergent. When modeled together on a phlyogenetic tree, ecdysone receptor and ultraspiracle segregate to their own clades, and do not recapitulate the pattern observed in the insect phenoloxidases. Therefore, despite being theoretically similar to phenoloxidase in terms of potential for dimer coevolution, the pattern of lineage specific duplications is not recapitulated, meaning that coevolution is unlikely to be driving the evolution of insect phenoloxidases. An analysis of the substitution rate of the protein-protein interacting regions of the phenoloxidase sequences may be able to shed

further light on any coevolutionary relationship that may be occurring, but until more insect sequences become available it is impossible to further investigate the unusual lineage specific duplications this study revealed

Works Cited

Bonneton, F., Zelus, D., Iwema, T., Robinson-Rechavi, M. and Laudet, V. (2003) Rapid divergence of the ecdysone receptor in Diptera and Lepidoptera suggests coevolution between ECR and USP-RXR. Molecular Biology and Evolution 20(4):541-553

Christophides, G.K., Zdobnov, E., Barillas-Mury, C., Birney, E., Blandin, S., Blass, C., Brey, P.T., Collins, F.H., Danielli, A., Dimopoulos, G., Hetru, C., Hoa, N.T., Hoffmann, J.A., Kanzok, S.M., Letunic, I., Levashina, E.A., Loukeris, T.G., Lycett, G., Meister, S., Michel, K., Moita, L.F., Muller, H.M., Osta, M.A., Paskewitz, S.M., Reichhart, J.M., Rzhetsky, A., Troxler, L., Vernick, K.D., Vlachou, D., Volz, J., von Mering, C., Xu, J., Zheng, L., Bork, P. and Kafatos, F.C. (2002) Immunity-related genes and gene families in *Anopheles gambiae*. Science 298(5591):159-165

Fujimoto, K., Okino, N., Kawabata, S., Iwanaga, S. and Ohnishi, E. (1995) Nucleotide sequence of the cDNA encoding the proenzyme of phenol oxidase A1 of *Drosophila melanogaster*. PNAS 92:7769-7773

Hall, M., Scott, T., Sugumaran, M., Soderhall, K. and Law, J.H. (1995) Proenzyme of *Manduca sexta* phenol oxidase: purification, activation, substrate specificity of the active enzyme, and molecular cloning. PNAS 92:7764-7768

Henikoff, S. and Henikoff, J.G (1992) Amino acid substitution matrices from protein blocks. PNAS 89:10915-10919

Huelsenbeck, H.P. and Ronquist, F. (2001) MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17(8):754-755

Jiang, H., Wang, Y., Ma, C. and Kanost, M.R. (1997) Subunit composition of pro-phenol oxidase from *Manduca sexta*: molecular cloning of subunit ProPO-P1. Insect Biochemistry and Molecular Biology 27(10):835-850

Jones, G., Brown, N., Manczak, M., Hiremath, S., and Kafatos, F.C. (1990) Molecular cloning, regulation, and complete sequence of a hemocyanin related, juvenile hormone-suppressible protein from insect hemolymph. Journal of Biological Chemistry 265:8596-8602.

Jones DT, Taylor WR & Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. Computer Applications in the Biosciences 8:275-282

Lourenço, A.P., Zufelato, M.S., Bitondi, M.M.G. and Simões, Z.L.P. (2005) Molecular characterization of a cDNA encoding prophenoloxidase and its expression in *Apis mellifera*. Insect Biochemistry and Molecular Biology 35(6):541-552

Pauwels, K., De Meester, L., Put, S., Decaestecker, E. and Stoks, R. (2010) Rapid evolution of phenoloxidase expression, a component of innate immune function, in a natural population of *Daphnia magna*. Limnology and Oceanography 55(3):1408-1413

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011) MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution 28:2731-2739

Terwilliger, N.B., Dangott, L. and Ryan, M. (1999) Cryptocyanin, a crustacean molting protein: Evolutionary link with arthropod hemocyanins and insect hexamerins. PNAS 96(5):2013-2018

Terwilliger, and Ryan, M. (2006) Functional and phylogenetic analysis of phenoloxidase from Brachyuran (*Cancer magister*) and Brachiopod (*Artemia franciscana, Triops longicaudatus*) crustaceans. Biological Bulletin 210:38-50

Whelan, S. and Goldman, N. (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Molecular Biology and Evolution 18(5):691-699

CHAPTER 4: THREE-DIMENSIONAL STRUCTURE MODELING OF PEA APHID PHENOLOXIDASE

Introduction

As various studies have demonstrated, dimer formation is critically important to the function of insect phenoloxidases, and the crystal structure of *Manduca sexta* phenoloxidase allows a deeper study into the dynamics of dimer formation (Li et al. 2009). 3D structures provide a wealth of information about the protein-protein interacting regions and the stabilizing bonds that work alongside the electrostatic bonding to enhance the interactions. This information can be used to map important residues, determine the organization of the active site, and deduce how a protein interacts with its upstream activators. Computer modeling of these 3D structures *de novo* has proven difficult and computationally expensive (Woodley and Catlow 2008), but advances have been made in both homology based modeling and protein-protein docking. Since the crystal structure of an insect phenoloxidase is available, modeling of the 3D structure of related genes presents itself as a way to more deeply examine the mechanisms of phenoloxidase activity in the pea aphid, as well as studying the role of dimer preference on the structure of active phenoloxidases.

Insect phenoloxidases were modeled against the 3D crystal structure of *Manduca sexta*, and subsequently submitted to a protein-protein docking software which was able to generate the predicted heterodimeric and homodimeric forms of these models. Models were created for the sequences both with and without their propeptide, resulting in six dimer models per pair of sequences. These models were analyzed with software that reported the sequence and polarity of the protein-protein interacting region, calculated the solvent inaccessible surface area that

mediated protein-protein interactions, and made predictive calculations about the number and location of salt bridges and hydrogen bonds within the interaction. These data allowed identification of trends in conformational change induced by propeptide cleavage and dimer preference, and served as an indicator of dimer stability.

Methods

Phenoloxidase Dimer Modeling and Analysis

Putative 3D structures of insect phenoloxidases were generated using the SWISS-MODEL protein structure homology modeling server (Arnold et al. 2006; Kiefer et al. 2009; Peitsch 1995). Models were generated based on both the full amino acid sequence and the sequence after removal of the propeptide according to its conserved location within the multiple alignment. All models, with the exception of *Drosophila melanogaster* prophenoloxidase A1, were successfully modeled against *Manduca sexta* prophenoloxidase subunit B (PDBid 3HHSB; Li et al. 2009). 3D structure models were then submitted to the GRAMM-X protein docking server (Tovchigrechko and Vasker 2005; Tovchigrechko and Vasker 2006) for modeling of dimer interactions. Insect lineages with two phenoloxidase genes had both homodimer structures and the heterodimer structure modeled, both with and without propeptides. Lineages with a single phenoloxidase gene had only the homodimer with and without propeptide modeled. 3D structure models were visualized in UCSF Chimera version 1.4.1 (Pettersen et al. 2004) or PyMOL version 1.3 (The PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC). Surface area, stabilizing bonds, and electrostatic interactions were calculated by the PROTORP web server (Reynolds et al., 2009).

Results

Analysis of Phenoloxidase Dimers

Every phenoloxidase subunit, with the exception of *Drosophila melanogaster* prophenoloxidase A1, was successfully modeled against *Manduca sexta* prophenoloxidase subunit B (3HHSB), both with and without the propeptide. Prophenoloxidase A1 from *Drosophila* was instead modeled against an oxygenated type 2 hemocyanin subunit (PDBID: 1NOLA). Based on the molecular characterization, a conserved propeptide cleavage site for *Daphnia magna* prophenoloxidase could not be identified, so only a homodimer model of the complete protein was created. A comparison between dimers with and without propeptides shows that propeptide cleavage reduces the surface area of protein-protein interaction by up to 53.8%, in the case of *Pediculus humanus,* and by an average of 15.9% (Table 5). This average is skewed by several dimers which do not appear to have their interacting regions altered by propeptide reduction, including all three *Acyrthosiphon pisum* models, the *Apis mellifera* homodimer, and the prophenoloxidase subunit B homodimer of *Manduca sexta*. The *Drosophila melanogaster* dimer models also exhibit this property, but this may be related to the alternative model for *D. melanogaster* prophenoloxidase A1. Propeptide cleavage appears to affect stability of the interacting region, with a significant reduction in number of salt bridges in dimer models that have reduced potein-protein interacting regions (Figure 9A, $T_{1,19}$=.768, $p > 0.05$). However, hydrogen bonds show no trend in relation to protein-protein interacting regions (Figure 9D, $T_{1,19}$=8.12, $p < 0.001$). The number of both hydrogen and salt bridges show a linear regression only slightly different from the line of equivalence in models with low protein-protein interaction area differences between propeptide and non-propeptide models (Figure 9B,9E). Among models with high protein-protein interaction area differences, the difference between the linear regressions and the lines of equivalence for salt bridges and hydrogen bonds are highly different,

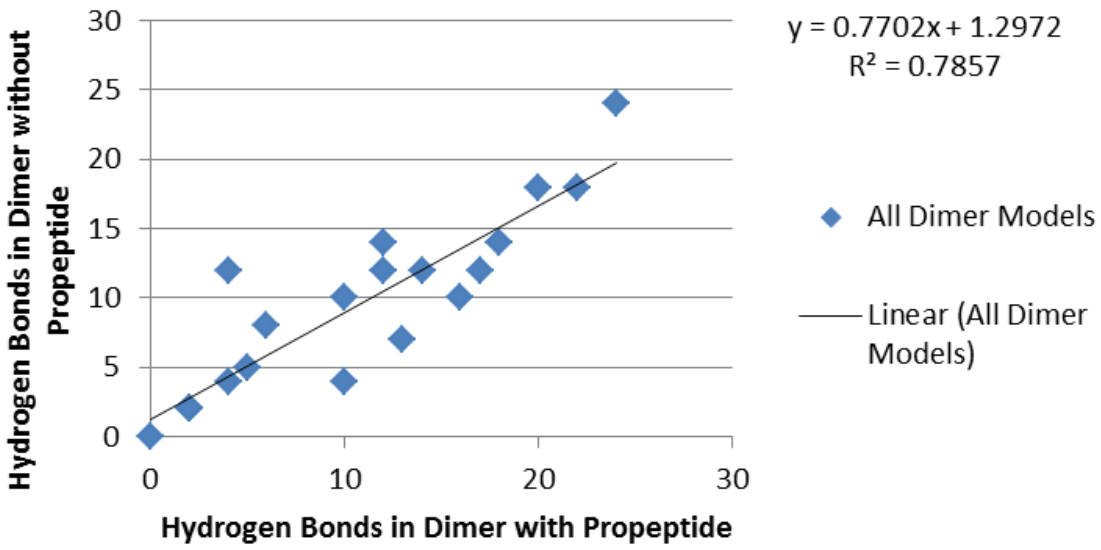
and the slopes are less than 1 (Figure 9C, 9F). This indicates that while propeptide removal has little effect on stabilizing bonds in models that do not have reduction of their interacting region under non-propeptide conditions, in models that have a greater than 15.9% change in interacting region hydrogen bonds and salt bridges are eliminated by propeptide removal.

Acyrthosiphon pisum phenoloxidase dimers are unusual among the models analyzed here in that they are not only nearly indistinguishable from each other among all factors measured, but they also show no alterations from cleavage of the propeptide (Table 5). Examinations of the 3D models using PYMOL show that while hetero- and homo-dimers from the same organism can have different binding site orientations (Figure 10), and that removal of the propeptide can result in conformational changes of the dimer (Figure 11), neither of these is the case for *Acyrthosiphon pisum* (Figure 12). *A. pisum* phenoloxidase dimers have the lowest protein-protein interaction surface area of all the modeled dimers. They also seem to have a very "open" conformation both with and without the propetide, a condition that is only seen in the original *M. sexta* model after propeptide cleavage. These 3D structure models not only demonstrate the unusual evolutionary history of phenoloxidase resulting in dramatically different dimer conformations of what is presumably an active enzyme, but also specifically highlight *A. pisum* phenoloxidases, which varies in nearly every respect from the most thoroughly characterized phenoloxidases of *M. sexta*.

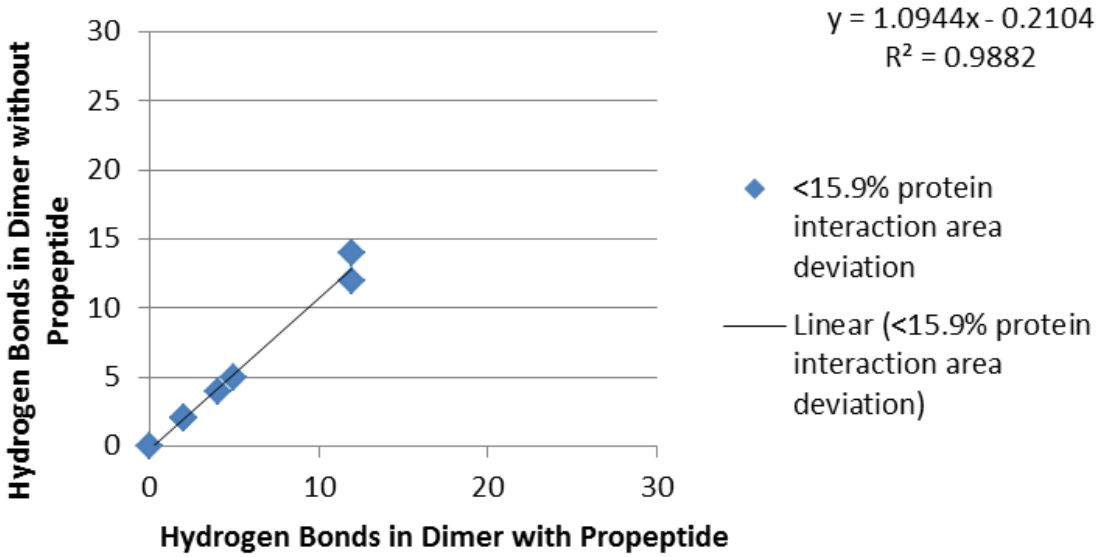
Table 5. Protein-protein interaction surface area for phenoloxidase dimer models. The non-solvent accessible surface area of each phenoloxidase dimer model was calculated in Angstroms squared ($Å^2$). Change between propeptide and non-propeptide versions of the same dimer model were calculated as a percentage, with a positive percentage indicating a reduction in non-solvent accessible surface area in the non-propeptide model. The *Anopheles gambiae* PPO2 homodimer was discluded from the average percent change calculations as a major outlier.

Model	Propeptide Interacting Region ($Å^2$)	Non-propeptide Interacting Region ($Å^2$)	% Change
A. gambiae PPO1 Homodimer	2380.54	1717.29	27.9
A. gambiae PPO2 Homodimer	1492.01	2519.11	-68.8
A. gambiae Heterodimer	1682.1	1340.22	20.3
A. mellifera Homodimer	1681.72	1694.24	-0.7
***A. pisum* PPOA Homodimer**	**1198.35**	**1198.23**	**0.0**
***A. pisum* PPOB Homodimer**	**1216.31**	**1216.85**	**0.0**
***A. pisum* Heterodimer**	**1196.06**	**1196.33**	**0.0**
B. mori PPO1 Homodimer	2681.25	1571.07	41.4
B. mori PPO2 Homodimer	2250.21	1718.77	23.6
B. mori Heterodimer	2423.4	1454.32	40.0
D. magna Homodimer	2090.13	----------	----------
D. melanogaster PPOA1 Homodimer	1366.87	1374.8	-0.6
D. melanogaster PPOA3 Homodimer	1728.9	1589.63	8.1
D. melanogaster Heterodimer	1418.31	1487.6	-4.9
M. sexta PPOA Homodimer	2314.51	1759.45	24.0
M. sexta PPOB Homodimer	1800.63	1879.07	-4.35
M. sexta Heterodimer	2454.76	1849.84	24.64
P. humanus Homodimer	2399.45	1108.58	53.8
T. castaneum PPO1 Homodimer	2409.25	1787.75	25.8
T. castaneum PPO2 Homodimer	2394.49	1754.77	26.7
T. castaneum Heterodimer	2442.89	1768.07	27.6
	Average % Change	Average % Change (Above Average Total Change)	Average % Change (Below Average Total Change)
	15.9	30.5	-0.2

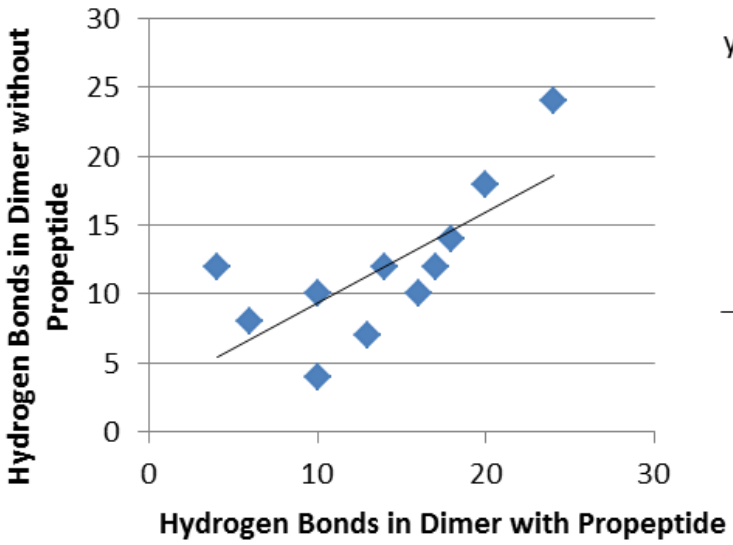
A


Number of Hydrogen Bonds
Hydrogen Bonds in Dimer without Propeptide
Hydrogen Bonds in Dimer with Propeptide
0
5
10
15
20
25
30
0
10
20
30
y = 0.7702x + 1.2972
R² = 0.7857
All Dimer Models
Linear (All Dimer Models)

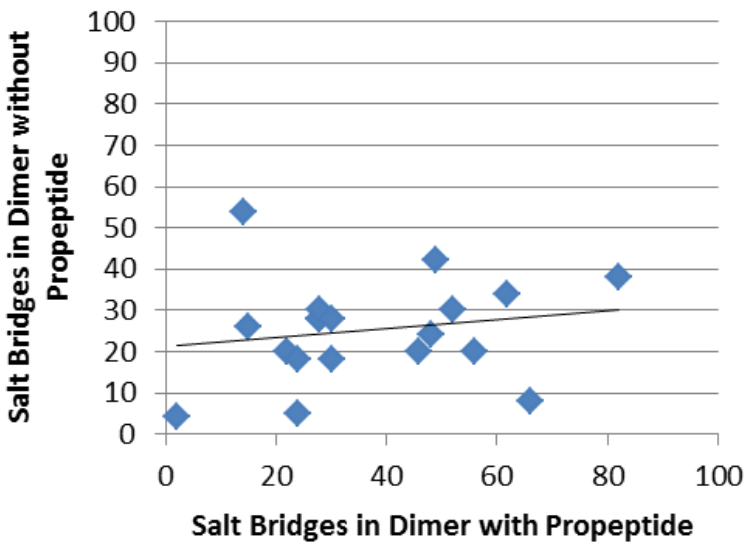
B


Number of Hydrogen Bonds
Hydrogen Bonds in Dimer without Propeptide
Hydrogen Bonds in Dimer with Propeptide
0
5
10
15
20
25
30
0
10
20
30
y = 1.0944x - 0.2104
R² = 0.9882
<15.9% protein interaction area deviation
Linear (<15.9% protein interaction area deviation)

C


Number of Hydrogen Bonds
Hydrogen Bonds in Dimer without Propeptide
Hydrogen Bonds in Dimer with Propeptide
0
5
10
15
20
25
30
0
10
20
30
y = 0.6549x + 2.8602
R² = 0.5223
>15.9% protein interaction area deviation
Linear (>15.9% protein interaction area deviation)

D


Number of Salt Bridges
Salt Bridges in Dimer without Propeptide
Salt Bridges in Dimer with Propeptide
0
10
20
30
40
50
60
70
80
90
100
0
20
40
60
80
100
y = 0.1073x + 21.214
R² = 0.0317
All Dimer Models
Linear (All Dimer Models)

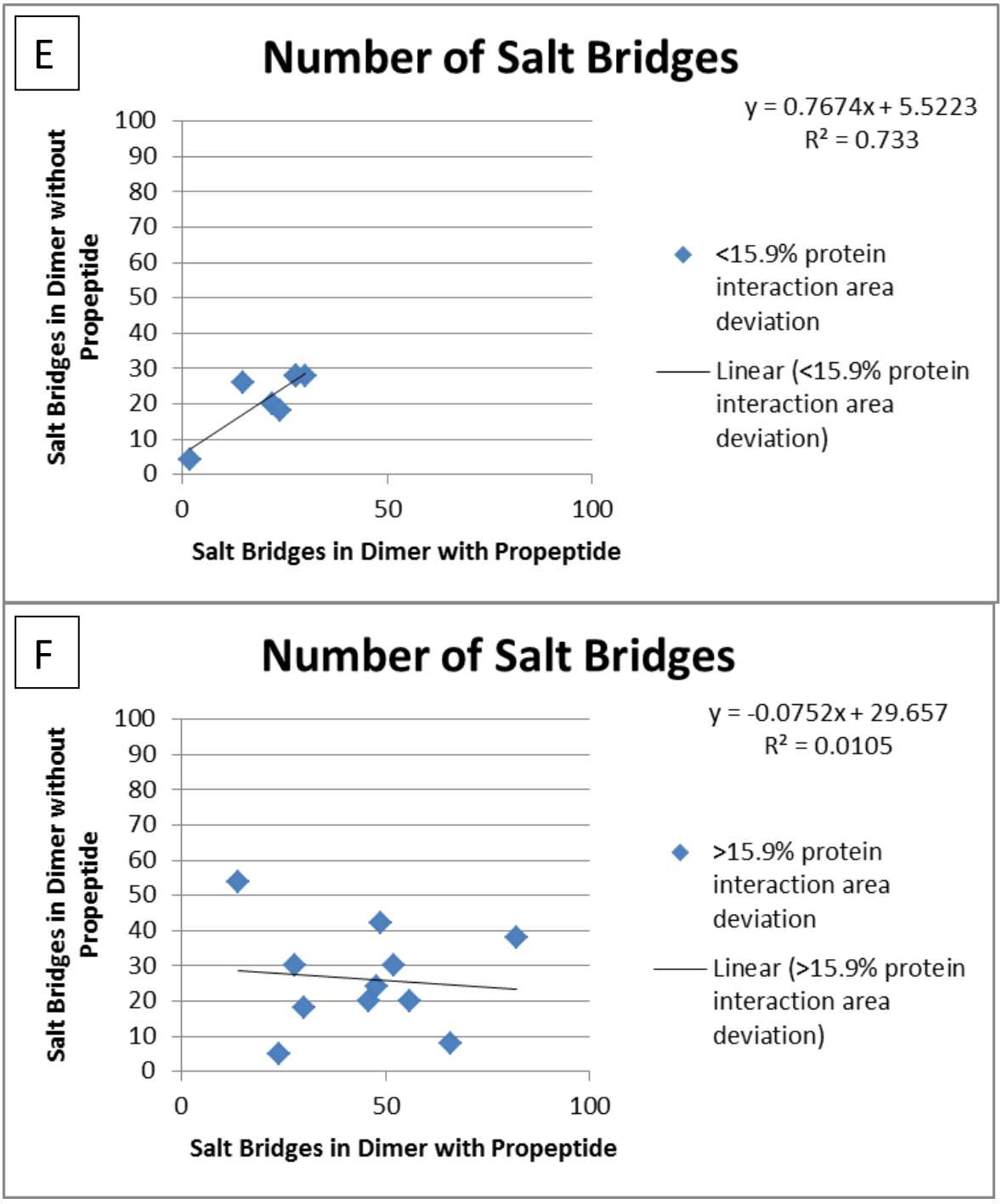


Figure 9. Comparisons of stabilizing bonds between propeptide and non-propeptide phenoloxidase dimer models. Each data point represents the ratio of salt bridges or hydrogen bonds between a predicted 3D model with and without its propeptide. Salt bridges (A-C) and hydrogen bonds (D-F) were calculated by PROTORP. PROTORP was unable to calculate

disulfide bonds or water bridges based on the models generated by GRAMM-X. Comparisons are done between the full dataset (A,D), and also split between models with less than the average percent change (15.9%) (B,E) or greater than 15.9% change (C,F) between their propeptide and non-propeptide forms.

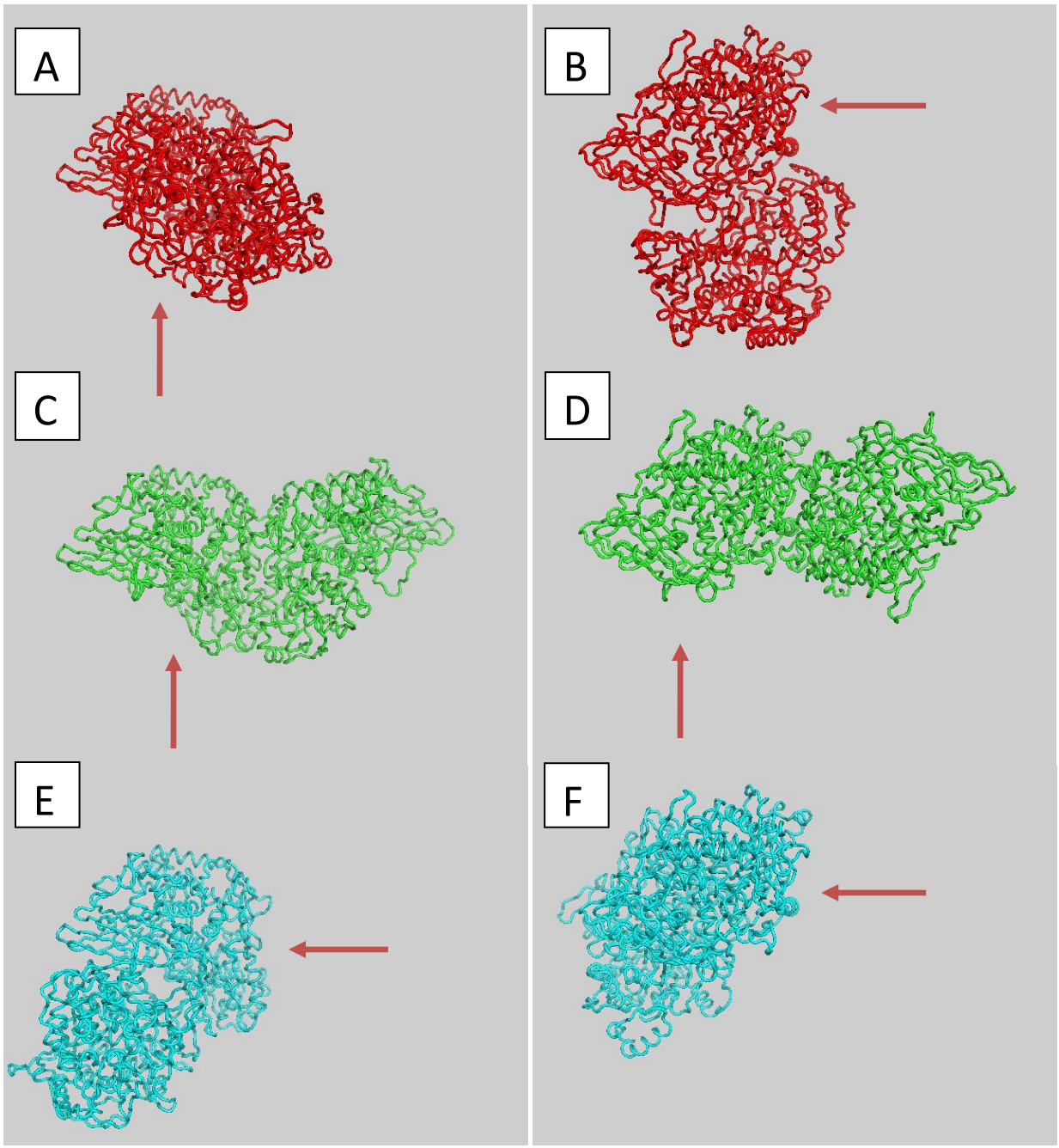


Figure 10. Alternate protein-protein interaction sites between homodimer and heterodimer models within species. Dimer models of *Anopheles gambiae* shown at the model's default orientation (A, C, E) and rotated ~90° around the x-axis (B, D, F). The models shown are the heterodimer of *A. gambiae* PPO1 and *A. gambiae* PPO2 (A-B), the homodimer of *A. gambiae* PPO1 (C-D), and the homodimer of *A. gambiae* PPO2 (E-F). All PROTORP models are created

so that one subunit, indicated by the red arrow in this figure and referred to henceforth as the A subunit, would overlay Manduca sexta PPOA (3HHSA) if the two models were loaded into the same 3D space. Therefore, this figure represents the relative position of the B subunit between models with overlapping A subunits.

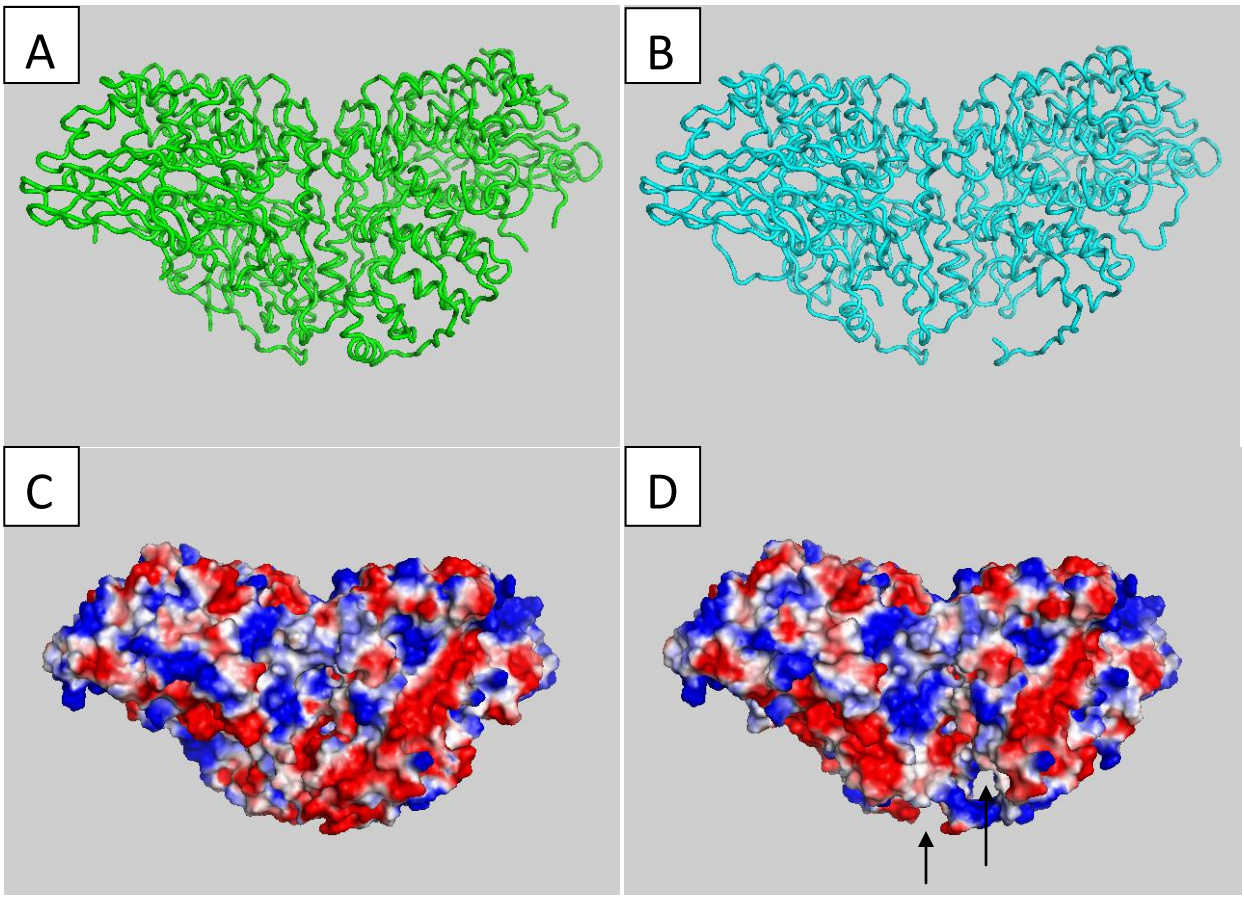


Figure 11. Conformational change of phenoloxidase dimer as a result of propeptide removal. The heterodimer of *M. sexta* PPOA and PPOB is shown both with (A,C) and without (B,D) propeptide, using either a projection of the secondary and tertiary structure (A-B) or the surface area electrostatics (C-D). Arrows indicate gaps in the protein-protein interaction area induced by propeptide cleavage, and represent a channel to the active site of the phenoloxidase dimer.

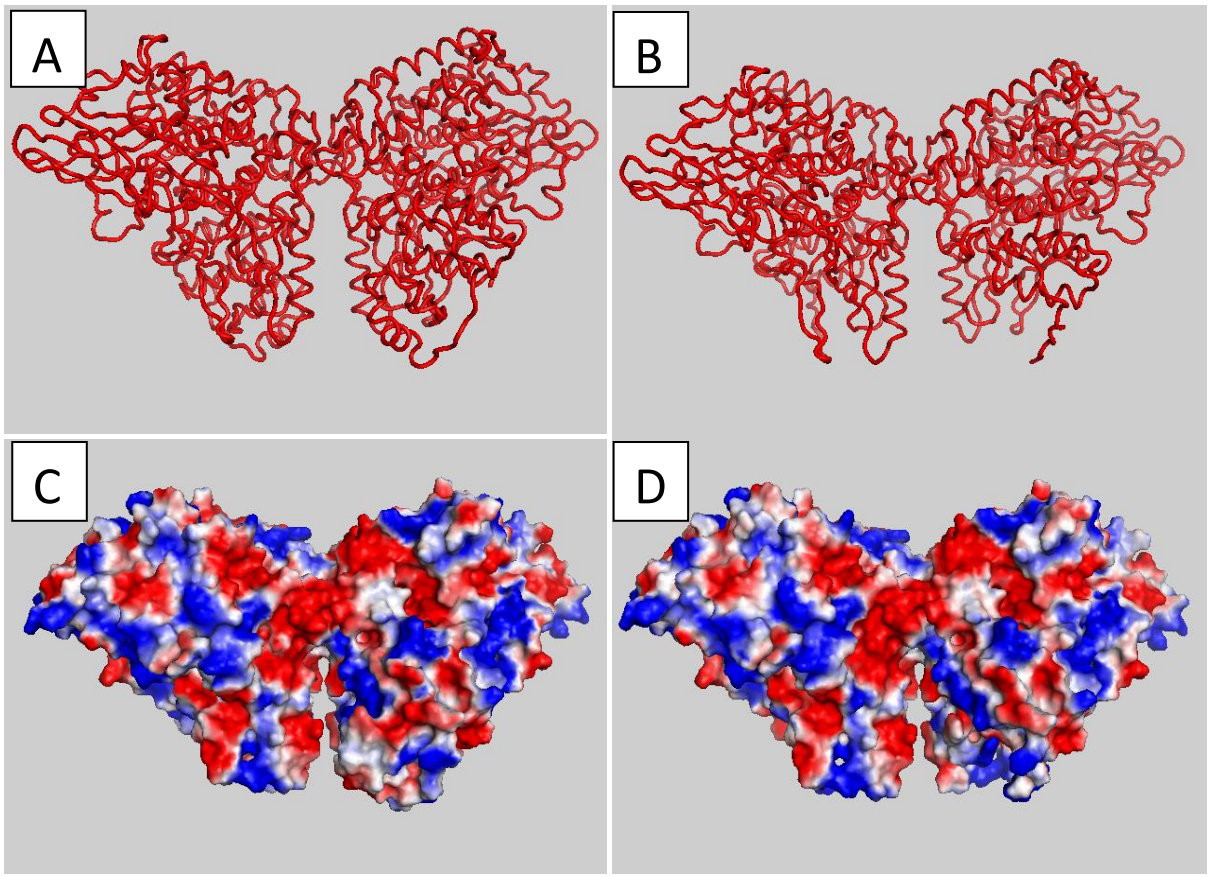


Figure 12. 3D structure models of *Acyrthosiphon pisum* phenoloxidase dimers. The heterodimer of *A. pisum* PPOA and PPOB is shown both with (A,C) and without (B,D) propeptide, using either a projection of the secondary and tertiary structure (A-B) or the surface area electrostatics (C-D). Both models appear to have a constitutively open channel leading to the active site.

Discussion

Analysis of the 3D structure models for the phenoloxidase dimers was undertaken to explore the protein-protein interactions which may underlie the function of insect phenoloxidases. The identification of two groups of dimer models based on the change in area of the protein-protein interacting region between models with and without the phenoloxidase propeptide demonstrates quite clearly that some of these proposed models represent enzymatically functional proteins, while others are not. The fact that nearly half of the dimer models produced from this analysis did not exhibit conformational change in response to propeptide cleavage underscores that while the mechanism of dimer preference is not thoroughly understood, it potentially has profound impact on dimer stability. The variability of protein-protein interacting regions between different dimer models from the same organism is also worth noting; if a region involved in the interactions undergoes random mutation, it does not necessarily affect all the potential dimers.

One goal for the 3D structure models was to use information from phenoloxidases whose preferred dimer state is known as a predictive model for dimer preference of the other sequences. As mentioned previously, a preference for heterodimer formation has been demonstrated in *Manduca sexta* phenoloxidase (Jiang et al. 1997), as well as *Bombyx mori* (Ashida and Yamazaki 1990). The heterodimer model for *M. sexta* phenoloxidase has the largest protein-protein interface, as well as the highest number of hydrogen bonds and salt bridges both with and without propeptide cleavage, in comparison to the homodimeric models. *D. melanogaster* PPOA3 homodimer seems to meet these criteria as well, but given the modelling issues encountered with PPOA1, it is difficult to determine if it is actually the most stable *D. melanogaster* dimer. Also, other organisms in the analysis do not have consistent indicators of

dimer stability (such as *Tribolium castaneum*). Given the lack of a consistent trend, without additional data about the preferred dimeric state *in vivo*, it is impossible to determine the preferred state using the *in silico* data. Further studies could be done with polypeptide antibodies or recombinant proteins to identify dimer preference, and determine if there is a pattern in dimer formation more complex than this study can explore.

There is only single organism, *Acyrthosiphon pisum,* whose predicted models do not appear to exhibit conformational change from propeptide cleavage in either its heterodimer or homidimer states. One important observation is that the lack of change in protein-protein interacting region with propeptide cleavage indicates that the propeptide does not contribute to these interactions. This shows that the pea aphid phenoloxidase 3D structure is naturally in the "open" conformation that is found only after propeptide cleavage in other organisms like *M. sexta*. An "open" structure suggests, unusually, that pea aphid phenoloxidase in a dimeric state is constitutively active. This is paradoxical to the assay data, which shows very little endogenous phenoloxidase activity but a huge increase due to exogenous activators. It is possible that the generated models are incorrect, but since all the phenoloxidase subunits were modeled successfully using *Manduca sexta* prophenoloxidase as a base such investigations would have to be done without homology modeling. Furthermore, the docking software was able to accurately recreate the *M. sexta* heterodimer, resulting in a model nearly identical to the crystal structure produced by Li et al. (2009), which lends support to the accuracy of the methodology for insect phenoloxidases.

One explanation for the discrepancy between the enzymatic and 3D structure data is that while the conformation of the dimer is open, the propeptides of the individual phenoloxidase subunits are still blocking the active sites. Lack of phenoloxidase activation under conditions of

wounding or infection could thus be explained by an incompatibility of the unusual dimeric form with the prophenoloxidase activating system. However, further studies into the interior pockets of the active sites of the 3D models, and whether they were blocked in the dimer, would have to be done to make this assertion with any confidence. It is also unlikely that this mechanism would be as strictly regulated as the typical phenoloxidase activating pathway, since the open conformation of the dimer may allow propeptide cleavage by nonspecific serine proteases, which could lead to uncontrolled cell damage.

Another possibility is that the dimer is, in fact, constitutively active, but the limiting rate of prophenoloxidase activation is formation of the dimer. Phenoloxidase dimers are the weakest of those modeled in terms of the surface area of their protein-protein interactions, and have a very low number of stabilizing hydrogen and salt bridge bonds. Aphid phenoloxidases may be unable to form dimers without some yet-unidentified stabilizing factor, or its interactions can only be maintained under certain physiological conditions. This latter explanation seems more plausible, considering that the pea aphid dimers have the highest percentage of nonpolar residues in their interacting regions of all the dimers modeled, and these interactions may be affected by cellular conditions. Various methods, including HPLC analysis (Jiang et al. 1997) and gel-filtration chromatography (Fujimoto et al. 1993) have previously been used to identify whether phenoloxidase is monomeric or dimeric *in vivo*, and similar methodologies could be used to determine under what conditions phenoloxidase dimers form in the pea aphid. A corollary of the previous concept is that phenoloxidase activity may derive from monomers instead of dimers. However, based on characterization of the prophenoloxidase activating system in other insects, there is no known mechanism for activation of monomeric phenoloxidase. Further study of the prophenoloxidase activating system in the pea aphid would be necessary for conclusive proof.

The final possibility to explain the conflict between enzymatic and modeling data is that the phenoloxidase activity that is being detected is not, in fact, coming from the phenoloxidase genes characterized in this study. While LOC100160034 and LOC100163393 are the only two pea aphid sequences which have the domains associated with tyrosinase-type phenoloxidases, the observed activity could be coming from laccases, or perhaps even another unidentified phenoloxidase sequence using different copper-binding domains. This could be tested by determining the substrate specificity of the phenoloxidase whose activity has been observed; evidence of tyrosinase activity would implicate that the two phenoloxidases focused on in this study are, in fact, the source of phenoloxidase activity in the pea aphid. While the 3D structure of pea aphid phenoloxidase has raised more questions than it answered, it does reveal how atypical phenoloxidase in the pea aphid appears to be.

Works Cited

Arnold, K., Bordoli, L., Kopp, J. and Schwede, T. (2006). The SWISS-MODEL workspace: A web-based environment for protein structure homology modelling. Bioinformatics 22:195-201

Ashida, M. and Yamazaki, I. H. (1990). Biochemistry of the phenoloxidase system in insects: with special reference to its activation. In Ohnishi, E. and Ishizaki, H. (ed.), Molting and Metamorphosis, Japan Scientific Societies Press, Tokyo, p. 239.

Fujimoto, K., Okino, N., Kawabata, S., Iwanaga, S. and Ohnishi, E. (1995) Nucleotide sequence of the cDNA encoding the proenzyme of phenol oxidase A1 of *Drosophila melanogaster*. PNAS 92:7769-7773

Jiang, H., Wang, Y., Ma, C. and Kanost, M.R. (1997) Subunit composition of pro-phenol oxidase from *Manduca sexta*: molecular cloning of subunit ProPO-P1. Insect Biochemistry and Molecular Biology 27(10):835-850

Kiefer, F., Arnold, K., Künzli, M., Bordoli, L. and Schwede, T. (2009). The SWISS-MODEL repository and associated resources. Nucleic Acids Research 37:D387-D392.

Li, Y., Wang, Y., Jiang, H. and Deng, J. (2009) Crystal structure of *Manduca sexta* prophenoloxidase provides insights into the mechanisms of type 3 copper enzymes. PNAS 106(40):17001-17005

Peitsch, M. C. (1995) Protein modeling by e-mail. Bio/Technology 13:658-660.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004) UCSF Chimera: A visualization system for exploratory research and analysis. Journal of Computational Chemistry 25(13):1605-1612

Tovchigrechko, A. and Vakser, I.A. (2005) Development and testing of an automated approach to protein docking. Proteins 60(2):296-301.

Tovchigrechko, A. and Vakser, I.A. (2006) GRAMM-X public web server for protein-protein docking. Nucleic Acids Research 34:W310-W314.

Reynolds, C., Damerell, D. and Jones, S. (2009) Protorp: a protein-protein interaction analysis tool. Bioinformatics 25(3):413-414

Woodley, S.M. and Catlow, R. (2008) Crystal structure prediction from first principles. Nature Materials 7:937-946

CHAPTER 5: CONCLUSION

The analysis of *Acyrthosiphon pisum* phenoloxidases has produced some surprising results about its form and function. While it appears to be enzymatically active, it is not utilized in response to wounding or bacterial infection. It instead seems to be localized to the same tissue as the aphids' symbiotic bacteria, where it could be serving any of a number of purposes, including management of symbiotic bacteria or degradation of the bacteriocyte over the aphid lifespan. The phylogenetic analysis revealed that, while the identified sequences were in fact immune-system phenoloxidases and not laccases, they exhibited a pattern of lineage specific duplications, along with the other insect phenoloxidases. Searching for an explanation of this pattern using the three-dimensional models of phenoloxidase dimers only made more apparent how unusual pea aphid phenoloxidase was, showing a dimer that appears to be in a state of constitutive activation, and raising the question of whether there is a fundamental difference in the mechanisms of phenoloxidase regulation in the pea aphid compared to other insects. While it is impossible with the available data to examine how the evolutionary history of phenoloxidase syncs with the incorporation of *Buchnera* as an obligate endosymbiont, it seems clear that phenoloxidase in the pea aphid has been subject to far different physiological contraints because of bacterial symbiosis. Because of the importance of phenoloxidase as a component of the innate immune system in other insects, it is difficult to imagine how it became adapted to its current use, but it seems to be an integral component of the bacteriocyte. Even if the role of phenoloxidase in the pea aphid is that of the sword of Damocles, it has a bearing on the fate of the aphid-bacterial symbiosis, and further investigations into the immune system of the pea aphid will likely reveal the extent of the adaptations the aphid has undergone to maintain and protect both its symbionts and itself.