STRUCTURAL STUDIES OF PROTEIN-SMALL MOLECULE INTERACTIONS: *BOMBYX MORI* PHEROMONE BINDING PROTEIN SPECIFICITY FOR BOMBYKOL, AND THE SUBSTRATE SPECIFICITY OF PANTOCIN A BIOSYNTHETIC PROTEINS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Catherine Lynn Lautenschlager

January 2005

STRUCTURAL STUDIES OF PROTEIN-SMALL MOLECULE INTERACTIONS: *BOMBYX MORI* PHEROMONE BINDING PROTEIN SPECIFICITY FOR BOMBYKOL, AND THE SUBSTRATE SPECIFICITY OF PANTOCIN A BIOSYNTHETIC PROTEINS

Catherine Lynn Lautenschlager, Ph. D.
Cornell University 2005

Interactions between small molecules and proteins mediate the biological processes of life. A knowledge of how proteins specifically interact with small molecules provides us with the tools to probe complex biology, but also allows us to develop potent and specific drugs to manipulate these processes to prevent, manage and cure disease.

In the first part of this thesis, the interactions between non-pheromone small molecules and the pheromone binding protein from *Bombyx mori* were explored by using X-ray crystallography to solve structures of these complexes. Pheromone signal transduction in moths provides a relatively simple model of the complex biology of neurological processing.

The second part of this thesis aimed to discover the functional form of a peptide substrate that interacts with biosynthetic enzymes to make pantocin A. This antibiotic is effective against *Erwinia amylovora,* the pathogen that causes fire blight disease. The functional form of the substrate was explored through genetic manipulation of the biosynthetic pathway that produces pantocin A.

BIOGRAPHICAL SKETCH

Catherine Lynn Lautenschlager was born on May 7, 1978 in Charlotte, North Carolina. Her interest in science was cultivated by her father through star-gazing and philosophy sessions, by her mother who provided a stimulating and well-rounded childhood, and by her sister, who let Cathy scribble all over her science project with red crayon in her youth. The author graduated from Oak Ridge High School in Orlando, Florida in 1996 and earned her B.A. in Microbiology from the University of Florida in 1999. Her interest in biochemistry further developed through the support of her parents, sister, teachers, and fellow labmates.

To my parents

ACKNOWLEDGMENTS

This work would not have been possible without the support of my advisors, collaborators, family and friends.

I would like to thank Dr. Jon Clardy, my advisor, for his exceptional encouragement, patience, and support throughout my graduate study, and for making a move mid-way through my study that ended up benefiting his graduate students in many ways. Special thanks to Sean Brady; without his intellectual guidance and friendship, I would probably not have made it through my first two years or my last year. To Cristy, Cindy, Jessie and Katherine, for their friendship. To Ryan and Muthu, for their advice as post-docs. To my collaborator, Dr. Walter Leal, for his advice and input.

Many thanks to Dr. Steven Ealick and Dr. Jerrold Meinwald for serving on my committee for the past five years.

Very special thanks to my parents for their love and support throughout my life; to my sister who has always gotten me through hard times; and to Jeremy, who reminds me that life does not revolve around the lab.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

PBP: pheromone binding protein

Ea273: *Erwinia amylovora* strain 273

CHAPTER 1

Introduction

Small molecules have the power to wreak havoc on our world. They can alter gene expression (1), agonize or antagonize cellular processes, trigger behaviors ranging from fighting to mating (2,3), and kill cells or organisms (4). Likewise, small molecule drugs have the power to fix catastrophic biological problems, suppressing cancer (5-7), repairing genetic disorders (8), and inhibiting the growth of pathogens (9) and the amplification of viruses (10) in the body. The language of small molecules is mediated by proteins. A knowledge of how specific proteins interact with small molecules can provide insights into complex biological processes and hasten the design potent and specific drugs to combat disease.

Chemosensory transduction provides an example of a complex biological process that is sensitive, specific and rapid. In chemosensory perception, an external chemical stimulus is received and interpreted by neurological processes. Examples of chemosensory transduction include the neurological processing of a small molecule odorant into a perception of smell, or the transmission of an external chemical signal into a complex behavior such as colonization or mating. In the silkworm moth, *Bombyx mori*, a female releases the sex pheromone bombykol into the air, a signal that is received by a male *B. mori* moth. Processing of this chemical signal induces a response in which the male moth flies to the female to mate from distances up to miles. Sex pheromone processing in insects provides a relatively simple model for the study of complex neurological processing. In addition, sex pheromone reception depicts an interesting example of the ability to distinguish self from

non-self. Male moths will only respond to sex pheromone signals of females from their own species, though moth sex pheromones can differ in as subtle ways as chain length, positioning of double bonds, or functional group (Fig. 1.1). Some species even use the same set of pheromone compounds that differ only in their composition of the blend.

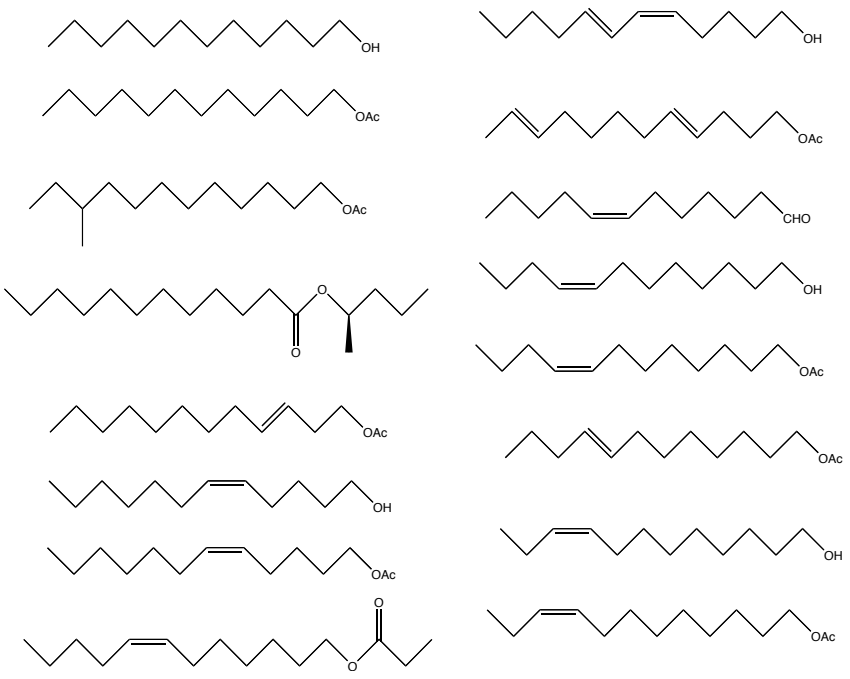


Figure 1.1. Sex pheromones from various moth species.

Small molecule antibiotics inhibit the growth of microbes. Most antibiotics are derived from mold or bacteria that have developed pathways to synthesize these compounds. Certain strains of the bacterium *Erwinia hericola* encode antibiotic-producing biosynthetic pathways that produce compounds effective in combating the bacterium *Erwinia amylovora*, a pathogenic bacterium that causes fire blight disease (11). Fire blight is a disease of rosaceous plants and has devastated crops in the Hudson Valley region as

well as other regions throughout the world. One of these antibiotics, pantocin A, is completely synthesized by only two enzymes from a peptide substrate encoded in a biosynthetic pathway. The target of this antibiotic is an enzyme involved in the biosynthesis of histidine, L-histidinol phosphate aminotransferase.

When first discovered, antibiotics were termed "wonder drugs". Antibiotic resistance has now risen to the point where it is necessary to find discover new classes of drugs effective against resistant strains of rising pathogens. New small molecule antibiotics such as pantocin A not only provide us with molecular scaffolds for the design of more potent and specific drugs, but can serve as probes to investigate bioprocesses (*e.g.* the biosynthesis of L-histidine) in addition to providing clues as to what systems in microbes serve as good targets for drug discovery.

References

1. Kuruvilla, F.G., Shamji, A.F., Sternson, S.M., Hergenrother, P.J., Schreiber, S.L. (2002) *Nature* **416**, 653-7.

2. Hartgens, F., Kuipers, H. (2004) *Sports Med.* **34**, 513-554.

3. Rasmussen, L.E., Schulte, B.A. (1998) *Anim.. Reprod. Sci.* **53**, 19-34.

4. Olsnes, S. (2004) *Toxicon* **44**, 361-368.

5. Lin, N.U., Winer, E.B. (2004) *Breast Cancer Res.* **6**, 204-210.

6. Piekarz, R., Bates, S. (2004) *Curr. Pharm. Des.* **10**, 2289-2298.

7. Cho, L.C., Choy, H. (2004) *Oncology* **18**, 29-39.

8. Saad, S.T., Lajolo, C., Gilli, S., Marques, J.F. Jr., Lima, C.S., Costa, F.F., Arruda, U.R. (2004) *Am. J. Hematol.* **77**, 45-49.

9. Meyerhoff, A., Albrecht, R., Meyer, J.M., Dionne, P., Higgins, K., Murphy, D. (2004) *Clin. Infect. Dis.* **39**, 303-308.

10. Chrusciel, R.A., Strohbach, J.W. (2004) *Curr. Top. Med. Chem.* **4**, 1097-1114.

11. Jin, M., Liu, L, Wright, S.A., Beer, S.V., Clardy, J. (2003) *Agnew Chem. Int. Ed. Engl.* **42**, 2898-2901.

CHAPTER TWO

X-ray crystallographic structure of apo B-form of Bombyx mori pheromone binding protein

Introduction

To attract a mate, a female moth releases pheromone, a volatile small molecule that is detected by the male moth. The pheromone signal is transduced in the antennae of male moths by an incompletely understood process. The hydrophobic pheromone reaches a G-protein coupled receptor after being transported through the antennal lymph by a pheromone binding protein (PBP). Once pheromone is detected, mating behavior is initiated and the male moth follows the pheromone plume, zig-zagging through the air to reach its mate (1).

Pheromone reception is highly dynamic. Signals must be rapidly detected, transduced, and degraded for the male to follow the trail of pheromone and locate a mate from a great distance. There is a need for sensitivity to discern the low concentration of the pheromone and selectivity to discriminate pheromone from other volatile molecules. Pheromone reception has been shown to be highly discriminatory, and even a slight alteration in the chemical structure of bombykol causes a several order of magnitude loss of potency as judged by electroantennograms and loss of activity as a sexual attractant (2). Electroantennograms have also shown an individual pheromone receptor can generate an electrical signal to the presence of only a single molecule of pheromone (3). A quick temporal response is also needed to stay on the pheromone trail. From wind tunnel experiments, it is estimated that ~200 molecules are sufficient to elicit a

behavioral response, and the projected time for this process is on the order of 100 msec (4). The rate-limiting step for pheromone detection may well be the journey of the hydrophobic pheromone through the aqueous lymph to reach the receptor.

In *Bombyx mori,* the BmPBP facilitates the rapid transfer of bombykol pheromone to the antennal receptor. In their most well-documented role, PBPs mediate the delivery of the hydrophobic sex pheromone to its receptor in the dendritic membrane. It has also been suggested PBPs play a role in deactivation and degradation of the pheromone (5). The lymph of pheromone-sensitive olfactory hairs in lepidopteran antennae contains 160 mg/mL PBP (10-20 mM). This extraordinary protein concentration ensures that each molecule of pheromone finds a PBP to convey it to the receptor. The high concentration of PBP may also reflect multiple PBP roles, such as degradation of pheromone signal or of non-pheromone small molecules.

The first structure of an invertebrate PBP was previously reported by Sandler et al. (6, Fig. 2.1a) for BmPBP with a bound bombykol molecule. The X-ray structure of the BmPBP•bombykol complex at 1.8 Å resolution revealed a completely helical protein with multiple disulfide bridges that completely encased the pheromone in a hydrophobic pocket of the protein. BmPBP is small (142 residues) and highly soluble, with six □ helices that are held together tightly by three disulfide bridges. The converging ends of four helices form one end of the pheromone binding pocket; the other end is capped by helix □3. Bombykol is completely engulfed in a hydrophobic cavity. Ser56 forms a hydrogen bond with the alcohol group of the pheromone, and Phe12 and Phe118 sandwich one set of conjugated double bonds in the molecule; other interactions aren't specific. The binding pocket

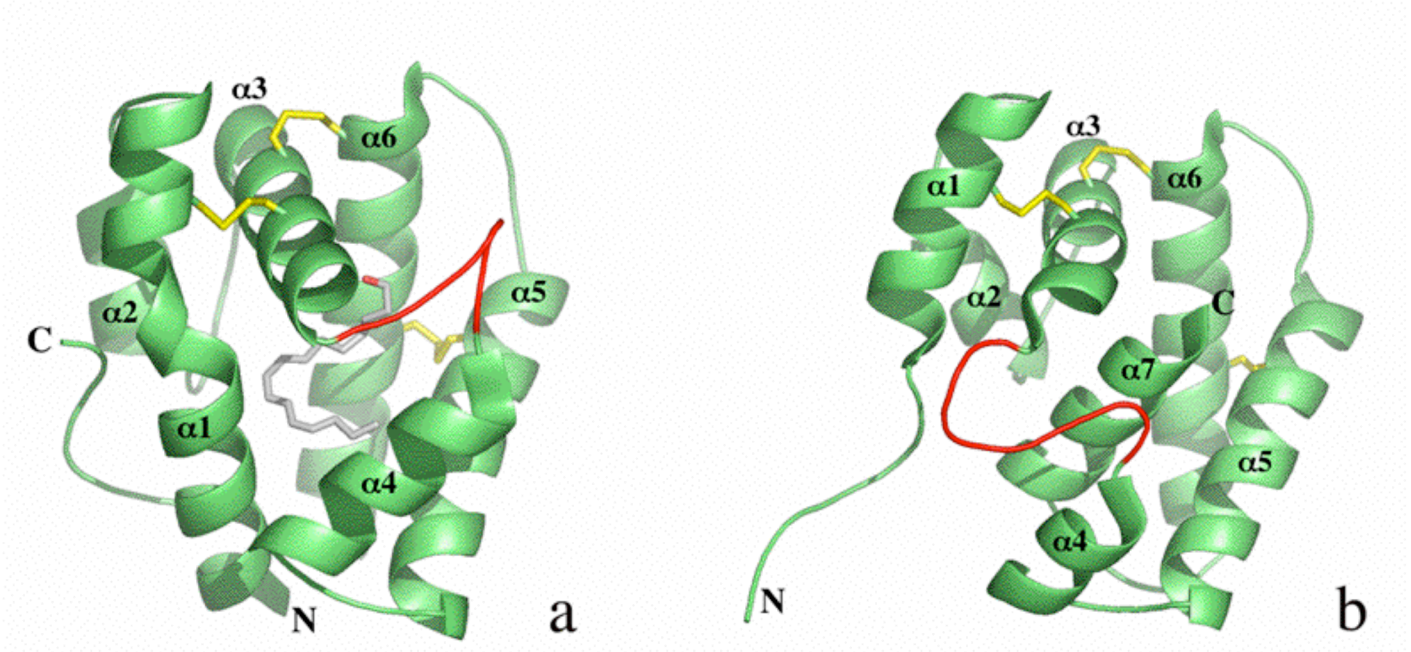


Figure 2.1. Structures of bombykol-bound and A-form *Bombyx mori* PBP. BmPBP is shown in green ribbons, bombykol as sticks with red alcohol oxygen atom, disulfide bridges in yellow, loop region between □3 and □4 in red. **N** is the N-terminus, **C** is the C-terminus, helices □1 - □7 as labeled. *a*, X-ray crystal structure of the BmPBP•bombyol complex at 1.8 Å resolution (Protein Data Base: 1DQE). *b*, NMR structure of A-form BmPBP (Protein Data Base: 1GM0). Figure created with PyMOL.

residues, Met5, Phe12, Phe36, Trp37, Ile52, Ser56, Phe76, Val94, Glu98, Ala115, and Phe118, are highly conserved across lepidopteran OBPs (6).

Both circular dichroism and tryptophan fluorescence studies showed that *Bombyx mori* PBP existed in two major conformations in a pH-dependent fashion (7). The low pH form, the A-form, was characterized by NMR (8, Fig. 2.1b). Above pH 5.5, the B-form PBP is in its "closed" conformation. Below pH 5.5, the A-form PBP is in its "open" conformation and does not bind pheromone. These structures suggested a model for pheromone release at the receptor (8). As the pheromone-bound protein approaches the receptor cell membrane a local drop in pH (9) could trigger the conformational change in the protein, ejecting the pheromone to its receptor. The crystal structure of BmPBP•bombykol first identified a loop region between α3 and α4 that could allow ligand access to the binding pocket (6). Within this loop, three conserved histidine residues (His69, His70, and His95) are present, one or more of which are likely to become protonated below pH 5.5. Protonation could cause charge repulsion and a conformational change in the loop, thus allowing bombykol to exit from the binding pocket.

The NMR structure of A-form BmPBP illustrates several major differences between the two protein conformations (Fig. 2.1). In the bombykol-BmPBP structure, the C-terminus of the protein is a disordered tail. In A-form PBP, it is the N-terminal region of the protein that is disordered, and the C-terminal region forms helix α7 that fills the binding pocket of the protein. The kink in α3 of the protein may form to accommodate α7 in the cavity (Fig. 2.1).

Though BmPBP is one of the best-characterized proteins in chemoreception, the story is incomplete. A detailed structural mechanism for

pheromone egress or entrance does not exist, and the role, if any, that BmPBP, or PBPs in general, play in discriminating between pheromones and structural analogs is not clear. The present study was designed to answer a structural question: what is the structure of the unbound cavity of BmPBP at neutral pH? The current theory is that unbound PBP at neutral pH has an empty binding cavity that is primed to bind pheromone in its cavity, and that a conformational change only occurs when the chemical environment passes a pH of approximately 5.5.

Materials and Methods

Expression and Purification of B. mori PBP. Plasmids for *B.mori* PBP were provided by Dr. Walter Leal at UC Davis. The protein was expressed recombinantly in *E. coli* as described (7).

Removal of ligand. PBP is a member of the family of odorant binding proteins, which are general binding proteins of hydrophobic molecules. Since BmPBP tends to pick up suitably shaped hydrophobic molecules during recombinant expression in *E. coli* (10), a method verified by mass spectrometry experiments to remove unwanted ligand was used in protein preparation (11). Following purification by ion exchange, hydroxyapatite colulmn, and gel filtration, the protein sample was dialyzed in 0.2 M sodium citrate pH 4.5 and incubated for 30 minutes at 37 °C with hydroxyalkoxypropyl dextran resin (Sigma), a hydrophobic binding resin, to open the protein and remove any endogenous ligand. Protein was eluted from the resin in the low pH sodium citrate buffer, dialyzed against 10 mM Tris pH 8.0, and concentrated to 20 mg/mL for crystallization.

Crystallization of apo B-form BmPBP. Crystals were obtained by vapor diffusion hanging drop method at 4 °C. A two microliter sample of protein was mixed with an equal volume of well solution containing 0.2 M sodium fluoride and 20% PEG 3350. The pH of the well solution and drops containing crystals was verified to be 7.5 (±0.5) by a litmus test. 2R,3R-butanediol (8%) was used as a cryoprotectant.

Data collection, phasing, and data processing. Complete data sets at 2.3 Å resolution were collected on A1 at CHESS (Cornell University, Ithaca, NY) using 1 ° oscillation steps. The data were integrated using DENZO of the HKL suite (12) and scaled using SCALA of the CCP4 suite (13). Data set statistics are shown in Table 2.1. Further data reduction was performed with CCP4i (13).

Molecular replacement with both the bombykol-bound and A-form BmPBP structures (Protein Data Bank entries 1DQE and 1GM0, respectively) was attempted using MOLREP (14). No replacement solution was found with the ligand-bound protein; a successful molecular replacement solution was found using the first model of the NMR A-form structure.

Model building and refinement. After building an initial model in O (15), the C-terminal helix of the model was omitted. Electron density was recalculated and the C-terminal residues were built into the electron density map manually to decrease potential model bias. Refinement was performed by iterative cycles of model building and simulated annealing using O and CNS (16). The first seven residues were omitted in the final refinement of the model.

Table 2.1. Crystallographic data for *Bombyx mori* apo PBP, pH 7.5

Data collection	
Source	CHESS F1
Wavelength (Å)	0.950
Space group	$P2_12_12_1$
Unit cell (Å)	
a	54.24
b	70.79
c	75.49
Resolution (Å)	51.3 – 2.10
Completeness, %	91.1
Observed reflections	103903
Unique reflections	12733
R_{sym}, (%)[a]	5.4
Refinement statistics	
Highest resolution (Å)	2.30
R factor (%)[b,c]	
Working	22.0
Free	26.9
rms deviation[d]	
Bond lengths (Å)	0.048
Angles, °	2.53
Average B-factor (Å)	26.6

[a] $R_{sym} = \Sigma|I - \langle I \rangle| / \Sigma I$
[b] $R = \Sigma_{hkl}||F_o| - |F_c|| / \Sigma_{hkl}|F_o|$
[c] R_{work} calculated from set of reflections in which 5% of the total reflections were randomly omitted from the refinement and used to calculate R_{free}.
[d] rms = root mean square

Results

Crystallographic results and comparison with other B. mori PBP structures. The X-ray structure of apo B-form BmPBP (Fig. 2.2a) most closely resembles the A-form BmPBP structure, which is also unliganded. The N-terminal region of the protein is disordered as in the A-form structure, and the first seven residues were omitted in the final cycle of refinement. The kink in helix α3 of apo B-form BmPBP is more apparent than in previous structures, and the loop region between α3 and α4 is longer through early termination of α3. This change in the loop may reflect the its conformational flexibility or repositioning of the residues in this region. The C-terminal tail of apo B-form BmPBP forms an ordered helix occupying the binding pocket. The presence of helix α7 in the binding pocket of the protein confirms that no ligand remains bound to the protein.

Helix 7 in the binding pocket. The difference in configuration of the C-terminus in the crystal structures of apo B-form BmPBP and BmPBP•bombykol is striking (Figs. 2.1a, 2.2a). Both structures were from crystals prepared under high pH B-form conditions, however, the C-terminal tail of the protein is a disordered coil in the bombykol-bound structure and forms an ordered alpha helix in the apo structure. This demonstrates that two conformations of the protein can exist at neutral pH, and that the conformation is sensitive to the presence or absence of ligand, along with pH.

Helix α7 makes many stabilizing contacts in the binding pocket, interacting with residues Leu8, Ser9, Leu10, Ser30, Trp37, Ser56, His70, Gly71, Ile91, His95, Ala115, and Phe118 (Figs. 2.3, 2.4). Many of these residues compose the binding pocket in the bombykol BmPBP structure. Hydrogen

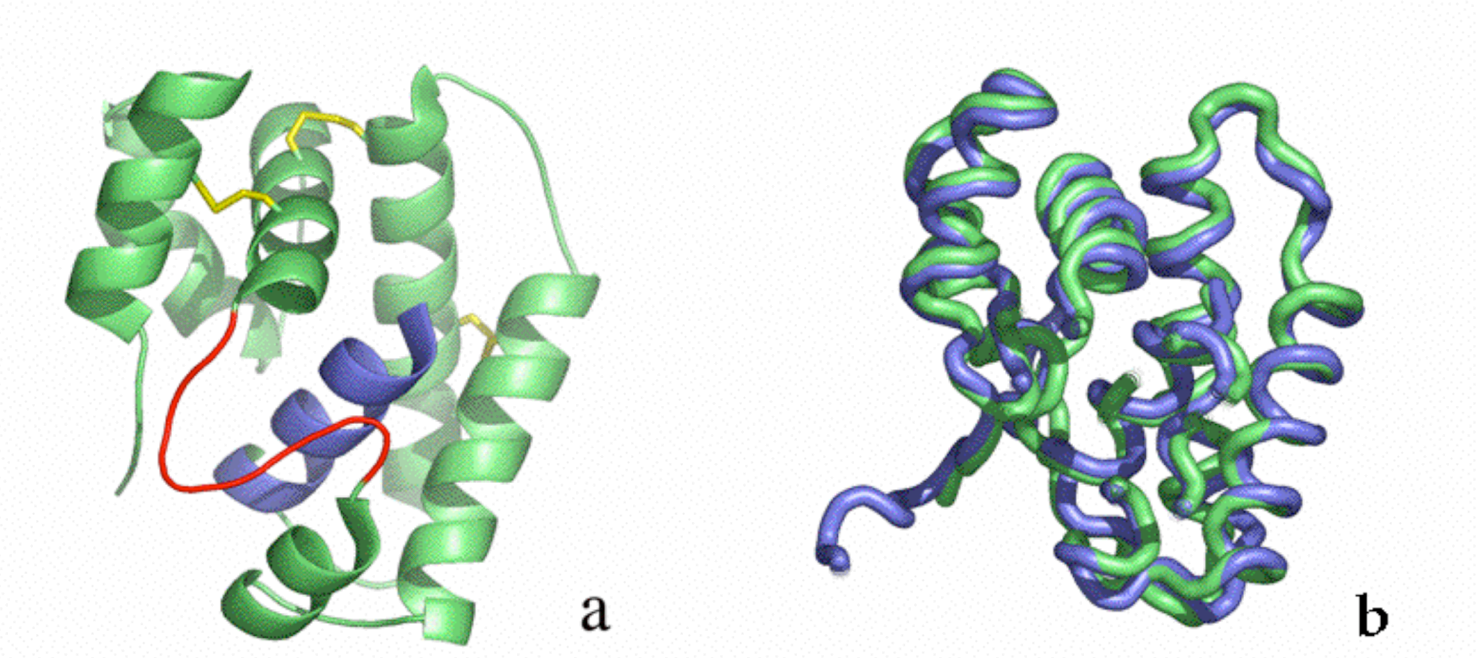
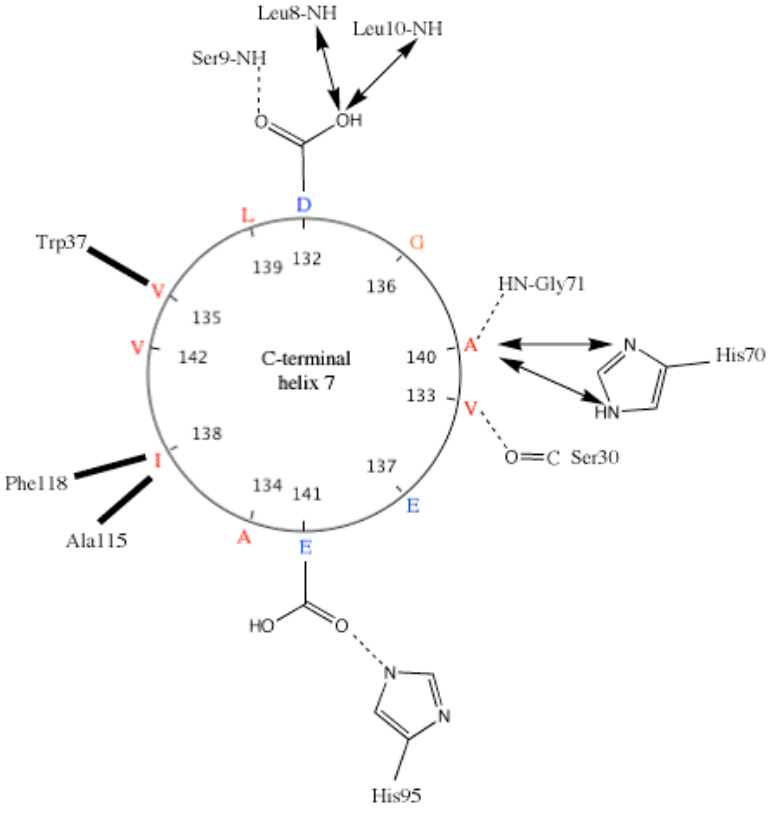


Figure 2.2. Structure of apo *B. mori* PBP, pH 8.0 at 2.3 Å resolution and superimposition with A-form BmPBP. *a*, Crystal structure of B-form apo BmPBP; helix □7 shown in blue. BmPBP is shown in green ribbons, disulfide bridges in yellow, loop region between □3 and □4 in red. *b*, C□ superimposition of apo structures of BmPBP; B-form (green), A-form (blue). Backbone of two superimposed structures depicted as tubes. Pictures and superimposition done in PyMOL.

Figure 2.3. Helix 7 and its interactions. Helix wheel of α7 (residues 132-142) from *B. mori* PBP was created with the Helix Wheel program (Turcotte). Acidic residues D132, E137, and E142 lie on one face of the helix and are indicated in blue; nonpolar residues are indicated in red. The helix exhibits amphipathic characteristics; the presence of glycine and perhaps alanine on the hydrophilic face would likely not be disfavored. Interactions with binding pocket residues are indicated: hydrogen-bonding depicted as dashed line (- - - - -), electrostatic interactions as double-headed arrow (↔), and hydrophobic interactions as broad line (▮). Hydrogen bonding to carbonyl oxygen or amide nitrogen of residues 132-142 as appropriate unless a side chain H-bond is depicted. When a potential proton donor and acceptor are more than 3.5 Å apart, the interaction was identified as electrostatic. Figure from ChemDraw.


Leu8-NH
Leu10-NH
Ser9-NH
O
OH
D
L
G
Trp37
V
139
132
136
HN-Gly71
135
V
142
C-terminal
helix 7
140
A
N
His70
133
HN
V
O=C Ser30
138
I
137
Phe118
134
141
E
Ala115
A
E
HO
O
N
N
His95

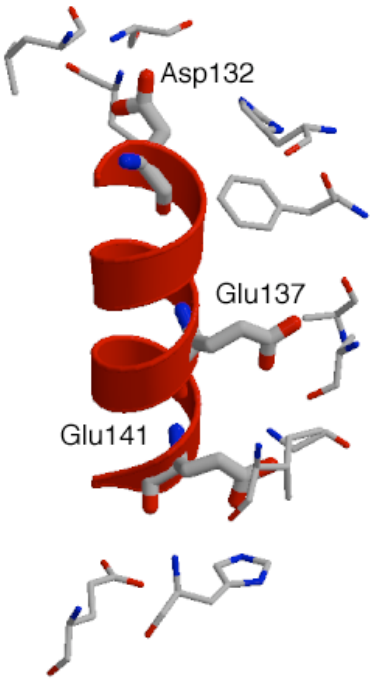


Figure 2.4. Location of acidic residues in helix 7. C-terminal tail adopts an amphipathic helical conformation in binding pocket of PBP. The positioning of three acidic residues in the C-terminal tail is completely conserved in lepidopteran PBPs. Picture created with MOLSCRIPT and RASTER3D.

bonding between carbonyl oxygens and amide hydrogens occurs between residue pairs Ser30/Val133 and Ala140/Gly71; H-bonding also occurs through side chain interactions between residue pairs His95/Asp141, and through the carboxylate group of Asp132 and the amide group of Ser9. Electrostatic interactions occur between residue pairs Asp132/Leu8, Asp132/Leu10, and Ala140/His70; hydrophobic interactions occur between Val135/Trp37 and Ile138/Phe118.

The interaction of helix 7 with His70 and His95 is interesting as these residues are believed to play a role in loop destabilization and pheromone

access to the binding pocket. Side chain atom N atom of His95 is 3.028 Å from an O of Glu141 and is in good orientation for hydrogen bonding. Hydrogen bond donor and acceptor atom distances of 2.5 – 3.2 Å fall into the category of "moderate, mostly electrostatic" H-bonding (17). Side chain atoms N and N of His70 are within 3.9 and 3.7 Å, respectively, of the carbonyl oxygen of Ala140. These distances would be rather large for H-bonding, and at high pH one of these ring nitrogens should not be protonated and therefore cannot act as a hydrogen bond donor. Hydrogen bonding distances of 3.2 – 4.0 Å fall into the category of "weak, electrostatic" bonding; H-bonds are not considered energetically significant when the distance between donor and acceptor atoms exceeds 3.5 Å (17). The interactions between His70 and Ala140 are weakly electrostatic; H-bonding may occur between His95 and Glu141. Helix 7 interactions with His70 and His95 may explain the early truncation of helix 3 and the longer looping region that follows.

Helix 7 is composed of mostly nonpolar amino acids with the exception of three acidic residues: D132, E137, and E141. These three residues lie on the same face of the helix, giving it a pronounced amphipathic character (Fig. 2.4). The amphipathic nature of helix 7 may play a role in the conformational change associated with release of pheromone at the receptor as discussed in the next section.

Discussion

Amphipathic helix 7. There are a number of examples of peptides and regions of proteins that switch from disordered to an ordered, amphipathic helix upon association with a membrane. Proteins and peptides exhibiting such a coil helix transition include viral fusion proteins, signal presequences, and

antibacterial and hemolytic peptides (18,19,20,21). Among known lepidopteran PBP sequences, the positioning of the three acidic residues in the C-terminal tail is maintained (Fig. 2.5). The structures of BmPBP indicate the tail of *B. mori* PBP undergoes a coil helix transition, perhaps through helix formation at the receptor membrane. Such a transformation may help to drive the overall conformational change associated with the protein upon membrane interaction, ejecting the ligand. It is known that the response of the receptor cell requires both the pheromone and PBP (22). The PBP is most likely important in solubilizing the pheromone in the lymph, but a more intimate interaction with the receptor cannot be ruled out. The three conserved acidic residues in the C-terminal tail might be necessary for pheromone release in the direct vicinity of the receptor; if amphipathic helix formation is critical to drive the conformational change that ejects the ligand, mutation of these residues may slow or abolish pheromone release.

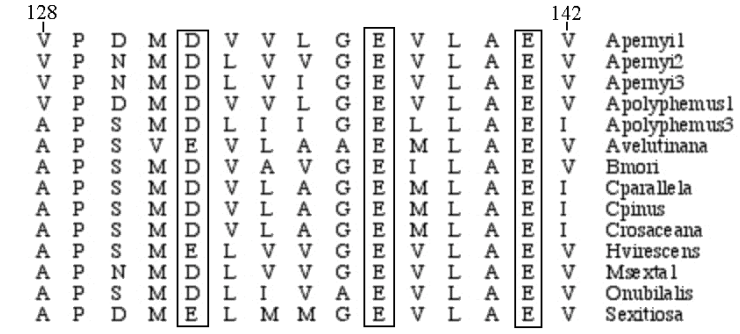


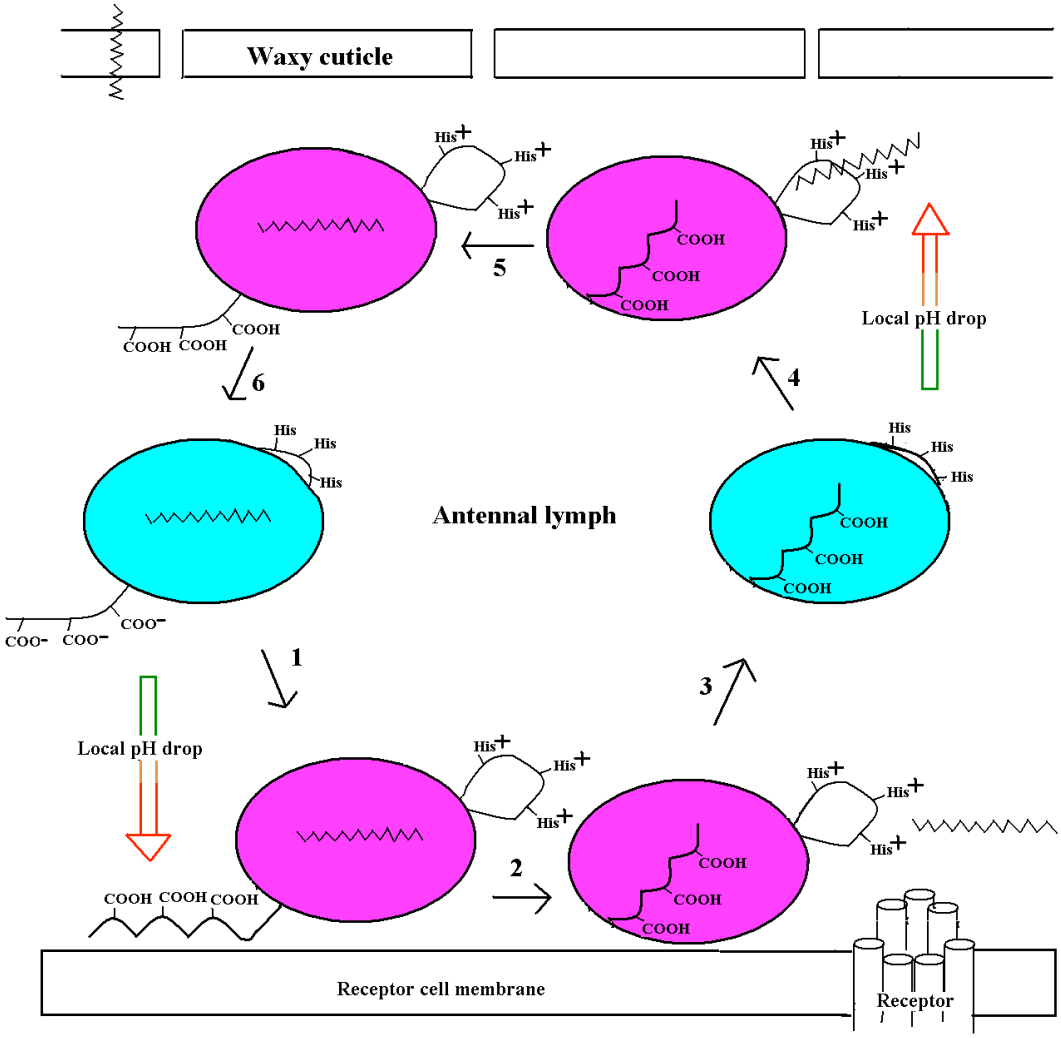
Figure 2.5. Sequence alignment of moth PBP C-terminal residues. Acidic residues 132, 137, and 141 are highlighted in boxes (using *B. mori* PBP residue numbers; marked on top ruler). Protein sequences from the GenBank; alignment created using MEGALIGN from Lasergene (DNAStar) and Adobe Photoshop.

Amphipathic helix formation dependent on the presence of a hydrophobic surface insinuates helix 7 formation and subsequent ejection of pheromone only occurs at a membrane surface.

The acidic residues of the tail may play an even more complex role in conformational change at the receptor membrane. A recent paper on cytidylyltransferase (CTP), a protein involved in phosphatidylcholine synthesis, analyzes one domain of the protein that undergoes a coil helix transition at anionic membranes (23). Domain M of CTP contains three glutamates that were shown to participate in the selectivity for anionic (over zwitterionic) surfaces. Mutating the three glutamates to glutamines in a synthetic peptide that mimicked the domain abolished its preference for anionic membranes. It is suggested the three glutamates become protonated in the low pH environment of anionic membranes; protonation would decrease the charge and increase the hydrophobicity of the domain, allowing the helix to insert into the membrane.

A similar phenomenon may be exhibited by B. mori PBP (Fig. 2.6). A local drop in pH is associated with the receptor cell membrane; upon approach of the pheromone-bound PBP to the receptor, an amphipathic helix may form and protonation of the three conserved acidic residues in the free C-terminal tail could neutralize the charge and increase the hydrophobicity of the region, allowing it to be more easily tucked into the hydrophobic pocket of the protein. This would imply that specifically at the negatively-charged receptor membrane, a shift in the conformation of the PBP occurs that favors ejection of the ligand to the receptor. This would also imply that mutating the three acidic residues to basic residues would undermine conformational change at the negatively-charged membrane. It has been shown that *B. mori*

Figure 2.6. Hypothetical model for ligand release at pheromone receptor. 1) As pheromone-bound protein approaches the membrane, C-terminal acidic residues are protonated and the C-terminus forms an ordered amphipathic helix. 2) Helix formation initiates conformational change in protein; protonation of histidine residues in loop destabilizes region and allows phermone to be ejected as helix α7 pushes into binding pocket. 3) Unliganded protein diffuses away from the membrane into higher pH region; histidine residues are deprotonated. 4) pH drop at cuticle protonates histidines in loop; loop is destabilized and ligand can enter binding pocket. 5) C-terminal tail competes with ligand for binding pocket. 6) pH increases as PBP moves away from cuticle, ionizing C-terminal acidic residues; C-terminus is no longer favored in hydrophobic binding pocket as it is displaced by pheromone. The blue oval represents B-form PBP; pink oval represents A-form PBP; pheromone is depicted as a jagged line. Histidines are indicated on looping region between helices α3 and α4; conserved acidic residues are indicated by carboxylate groups at C-terminus.


Waxy cuticle
His+
COOH
Local pH drop
5
6
4
His
COO-
Antennal lymph
1
3
Local pH drop
2
Receptor cell membrane
Receptor

PBP exhibits its conformational transition upon interaction with anionic, but not neutral, phospholipids vesicles (7).

To continue the cycle, as the pH drops at the receptor cell membrane the three conserved histidine residues in the looping region are also protonated. The resulting ionization of this loop destabilizes α3, switching the PBP to an open conformation and allowing pheromone to exit as helix α7 pushes into the binding pocket. The unbound protein now diffuses away from the receptor membrane, the histidine residues are deprotonated as the pH increases. As the PBP approaches the negatively-charged dendritic membrane at the pore cuticle (9), there is another local drop in pH, the histidine residues are again protonated and the protein reverts to its open conformation, allowing incoming pheromone access to the binding pocket. An equilibrium between pheromone or C-terminal tail occupation of the binding pocket may ensue – bombykol may be most favored by the binding pocket, with uncharged C-terminal tail less favored, and charged C-terminal tail unfavored. Ionization of the C-terminal tail as the PBP diffuses away from the cuticle and the pH rises pushes the equilibrium towards binding pheromone in the pocket. At this point, the protein is again in its closed conformation, protecting pheromone from degradation, and the cycle can begin again.

During protein preparation, BmPBP was cleared of any ligand it might have acquired during recombinant expression in *E. coli* with the use of a hydrophobic binding resin, Lipidex, in citrate buffer at pH 4.5 (see Materials and Methods), making the apo B-form X-ray crystal structure consistent with the proposed model for conformational change and pheromone release described above.

Structure of the binding pocket. There are three potential fates for a vacant protein cavity: it can collapse, be filled with something other than the ligand, or remain open and empty, though this last option is not likely in the case of BmPBP because of its large cavity size. Empty protein cavities are thought to be energetically unfavorable due to loss of van der Waals contacts (24). A study comparing 121 protein structures in the PDB database with single cavities ranging from 10 – 204 $Å^3$ in volume classified some general tendencies (25). Cavities were grouped into three categories: solvated, empty, and unknown. Solvated cavities were defined by the presence of water molecules in the cavity, empty cavities contained solvent information with the PDB file but had no water molecules indicated in the pocket (but could contain poorly-ordered water), and unknown was used to describe cavities associated with PDB structures with no solvent information. Larger cavities were found to more likely be solvated and delineated by polar residues, and empty cavities were favored when the pocket was formed by nonpolar residues. The hydrophobic cavity of BmPBP would likely fall in the last category. Empty cavities rarely exceeded 50 $Å^3$ in size; the size of the binding pocket in the bombykol B-form structure is calculated to be 167 $Å^3$,more than three times that size, by VOIDOO (26). The calculated penalty of having an empty cavity in a protein is 24 cal $mol^{-1}Å^{-3}$ (27), so the nominal energy penalty for the empty cavity would be ~ 4 kcal mol^{-1}. The C-terminal tail of BmPBP can fill the binding pocket of the protein as seen in apo A-form NMR and crystal apo B-form structures, and the protein has been shown to bind fatty acid derivatives from *E. coli* when recombinantly expressed; an occupied binding pocket is likely more energetically favorable than an open, empty cavity, making it more favorable for BmPBP to bind either ligand or helix □7 in its pocket.

Ligand-dependent conformational change. The crystal structure of B-form apo-BmPBP suggests that in addition to the well-characterized pH-dependent conformational change associated with the protein, there may also be a ligand-dependent conformational change. Such a change has been demonstrated by circular dichroism and UV-spectroscopy for the three PBPs from *Antheraea polyphemus* (21). An equilibrium of at least two major conformations of BmPBP may exist at higher pH, favoring a disordered C-terminal tail when ligand is present, and an ordered C-terminal helix when ligand is absent.

References

1. Marsh, D., Kennedy, J.S., Ludlow, A.R. (1978) *Physiol. Entomol*., 3, 221–240.

2. Steinbrecht, R.A., Kasang, G. (1972) *Olfaction and Taste IV*, 193-199.

3. Vickers, N.J., Baker, T.C. (1994) *Proc. Natl. Acad. Sci. USA*. 91:5756-5760.

4. Kaissling, K.E. (1998) *Ann. N Y Acad. Sci.* **30**, 320-2.

5. Kaissling, K.E., Priesner, E. (1970) *Naturweissenschaften* **57**, 23-28.

6. Sandler, B.H., Nikonova, L., Leal, W.S., Clardy, J. (2000) *Chem. & Biol.* **7**, 143-51.

7. Wojtasek, H., and Leal, W.S. (1999) *J. Biol. Chem.* **274**, 30950-30956.

8. Horst, R., Damberger, F., Luginbuhl, P., Guntert, P., Peng, G., Nikonova, L., Leal, W.S., Wuthrich, K. (2001) *Proc. Natl. Acad. Sci.* **98**, 14374-79.

9. Kaissling, K. E. & Thorson, J. (1980) in *Receptors for Neurotransmitters, Hormones and Pheromones in Insects*, eds. Sattelle, D. B., Hall, L. M. & Hildebrand, J. G. (Elsevier, Amsterdam), pp. 261-282.

10. Oldham, N.J., Krieger, J., Breer, H., Fischedick, A., Hoskovec, M., Svatos, A. (2000) *Agnew. Chem. Int. Ed.* **39**, 4341-43.

11. Oldham, N.J., Krieger, J., Breer, H., Svatos, A. (2001) *Chem. Senses* **26**, 529-531.

12. Otwinowski, Z., Minor, W. (1997) *Methods Enzymol.* **276**, 307-326.

13. Collaborative Computation Project, Number 4 (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50**, 760-763.

14. Vagin, A., Teplyakov, A. (2000) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **56**, 1622-1624.

15. Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard. (1991) *Acta Crystallogr. Sect. A* **47,** 110-119.

16. Brunger, A.T., Adams, P.D., Clore, M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M, Simonson, T., Warren, G.L. 1998. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54**, 905-921.

17. Jeffrey, G.A. "An introduction to hydrogen bonding" Oxford U. Press, 1997.

18. Lampio, A., Kilpelainen, I., Pesonen, S., Karhi, K., Auvinen, P., Somerharju, P., Kaariainen, L. (2000) *J. Biol. Chem.* **275**, 37853-9.

19. Hammen, P.K., Gorenstein, D.G., Weiner, H. (1996) *Biochemistry* **35**, 3772-81.

20. Wieprecht, T., Apostolov, O., Beyermann, M., Seelig, J. (1999) *J. Mol. Biol.* **294**, 785-94.

21. Nicol, F., Nir, S., Szoka, F.C. Jr. (1996) *Biophys. J.* **71**, 3288-301.

22. Pophof, B. (2002) *Naturwissenschaften* **89**, 515-8.

23. Johnson J.E., Xie M., Singh L.M., Edge R., Cornell R.B. (2003) *J. Biol. Chem.* **278**, 514-22.

24. Pakula, A.A., Sauer, R.T. (1989) *Ann. Rev. Genetics* **23**, 289-310.

25. Hubbard, S.J., Gross, K.H., Argos, P. (1994) *Protein Eng.* **7**, 613-26.

26. Kleywegt, G.J., Jones, T.A. (1993) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50,** 178-185.

27. Eriksson, A.E., Baase, W.A., Zhang, X.J., Heinz, D.W., Blaber, M., Baldwin, E.P., Matthews, B.W. (1992) *Science* **255**, 178-83.

28. Mohl, C., Breer, H., Krieger, J. (2002) *Invert. Neurosci.* **4**, 165-74.

CHAPTER THREE

X-ray crystallographic structures of *Bombyx mori* pheromone binding protein (PBP) in complex with non-pheromone ligands

Introduction

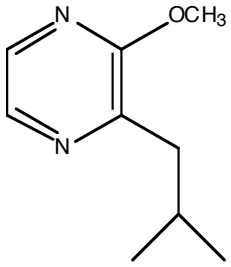
Small molecule sex pheromones act as chemical messages to elicit sexual behavior in insects. The sex pheromone of *Bombyx mori*, bombykol (Fig. 3.1), is released by the female moth, diffuses through the air, and is detected in the highly branched antennae of male moths. The male follows the pheromone plume to its mate. Pheromone is detected in male moth olfactory hairs, or sensilla, which cover the antennae. The pheromone adsorbs to the olfactory hair, diffuses through pore tubules into the aqueous sensillar lymph where it is bound to the pheromone binding protein, and is transported to the receptor on the neural cell.

The lymph of pheromone-sensitive olfactory hairs in *Bombyx mori* antennae contains a high concentration of pheromone-binding proteins (PBPs), 10-20 mM or about 160 mg/mL. PBPs are members of the odorant-binding protein family, a group of proteins that act to solubilize hydrophobic compounds in an aqueous environment. This subclass of odorant binding proteins (OBPs) mediates the delivery of the sex pheromone to its receptor in the dendritic membrane. Lepidopteran PBPs differ structurally from other OBPs.

Pheromone reception in male moths is highly sensitive; a male moth needs to identify and respond only to the pheromone of its species. This specificity is particularly impressive given the degree of similarity of moth sex pheromones. Sex pheromones produced by female Lepidopterae are typically

Bombykol

Iodohexadecane



2-isobutyl-3-methoxypyrazine
Bell pepper odorant

Figure 3.1. Ligands for *B. mori* PBP complexes

twelve- to eighteen-carbon partially unsaturated aliphatic chains. These molecules differ only by carbon chain length, placement of double bonds, and functional group (aldehyde, alcohol, or acetate). Some species use a single molecule as their pheromone, while others use a blend of molecules with a well-defined composition. A question of considerable debate is whether PBPs play a role in discriminating among molecules to augment the specificity of the signaling process.

There is evidence to both support and contradict a role for PBPs in pheromone recognition. Analysis of the primary structure of PBPs from

different moth species shows limited diversity among the proteins. The sequences are about 70% identical and 85% similar to each other (BLOSUM 62 matrix in MEGALIN), and analyzing the amino acid sequences for residues that might be involved in specificity has yielded nothing conclusive, except for a serine residue (Ser56 in *B. mori*) that is believed to interact with the alcohol group in 14-16 carbon chain pheromones. However, several species of moth have more than one PBP, coinciding in instances where the species pheromone signal consists of a blend of molecules. It has been hypothesized that in species with pheromonal blends, each PBP binds a different component of the blend.

Ligand affinity of pheromone binding proteins has been hard to study through binding experiments due to the hydrophobicity and insolubility of the pheromones. Free ligand is difficult to separate from PBP-bound ligand because the hydrophobic molecules adhere to the surfaces of the vials and plastic pipettes used in experiments. To explore the specificity of *B. mori* PBP we crystallized the protein with non-pheromone ligands (Fig. 3.1) and solved the structures of these complexes by X-ray crystallography. To investigate limitations that would be imposed by stringent specificity we complexed the PBP with iodohexadecane. This ligand has a chain length that is longer than that of bombykol, lacks the restraint provided by the two double bonds in the pheromone, and replaces the alcohol functional group with an iodine atom. The overall shape of iodohexadecane is somewhat similar to moth pheromones, being comprised of a long hydrocarbon chain, though the molecule isn't locked into the hook-shaped geometry exhibited by bombykol due to its *cis* double bond between carbon atoms 12 and 13. The large iodine atom would also presumably be easy to identify in an electron density map.

We also wanted to explore a potential ligand with quite different geometry than that of a typical moth sex pheromone, and complexed the protein with bell pepper odorant (2-isobutyl-3-methoxypyrazine). This aromatic molecule lacks the long chain of bombykol but retains its hydrophobic nature. Bell pepper odorant is a volatile molecule that might be encountered by a male moth, and as pheromone binding protein is a member of the odorant binding protein family of proteins, we were curious to see whether the odorant would be bound. Presence of electron density fitting both of these candidate ligands in the binding pocket of *B. mori* PBP provides solid experimental information that can be used to structurally characterize the degree of specialization of the PBP binding pocket for its natural ligand, bombykol.

Materials and Methods

Protein preparation. Plasmids for *B. mori* PBP were obtained from Dr. Walter Leal (UC Davis, CA). Recombinant protein was expressed in *E. coli* BL21 (DE3) cells and purified as previously described (2). A slight modification of the purification procedure was used to "empty" the protein of any fatty acid or other ligand the protein might have picked up during expression (3). Recombinant PBP was incubated with a Lipidex resin at pH 4.5, 37 °C to bind any hydrophobic ligand acquired during recombinant expression. Ligand-free protein was eluted with 0.2 M sodium phosphate pH 4.5.

Ligand addition. Ligands were dissolved in methanol to make them more soluble for protein complex formation and added at pH 4.5 prior to dialyzing the protein back to pH 8.0. Protein concentration was determined by Bradford Assay using bovine serum albumin standards, and ligand was added to the

protein in a 10X molar excess. The protein-ligand solution was agitated to disperse the ligand, and the protein complex was dialyzed against 10 mM Tris buffer pH 8.0. Protein complexes were concentrated to 20 mg/ml, and unbound ligand removed by centrifugation. As a control, protein purified as described above with Lipidex treatment, omitting the ligand addition step, was dialyzed against 10 mM Tris pH 8.0 and crystallized at this pH; no electron density was discovered in the binding pocket of PBP using this protocol, ensuring the only molecules present for binding are those we add to the protein. The structure of *B. mori* PBP complexed with bombykol was solved again following the above procedure to ensure the validity of the method.

Crystallization and structure determination. Crystals were obtained by vapor diffusion hanging drop method at 22 °C. A two microliter sample of protein was mixed with two microliters of well solution containing 40% polyethylene glycol (MW 4,000), 100 mM Tris pH 8.0, and 50 mM $MgCl_2$. Complete data sets at 1.9 Å resolution for ihd-BmPBP and 2.0 Å resolution for the bpo-BmPBP were collected on F1 at CHESS (Cornell University, Ithaca, NY). The data were integrated using DENZO of the HKL suite (4) followed by scaling with SCALA of the CCP4 suite (5). Data set statistics are shown in Table 3.1. Further data reduction was performed with CCP4i (5).

The structures were solved by molecular replacement using the CCP4 version of MOLREP (6). The model used was the *B. mori* PBP pheromone complex, chain A (Protein Data Bank entry 1DQE). The bombykol ligand was omitted from this model and only one monomer was used for molecular replacement.

Table 3.1. Crystallographic data for *Bombyx mori* PBP complexes

	ihd-BmPBP	bpo-BmPBP
Ligand	iodohexadecane	bell pepper odorant
Data collection		
Source	CHESS F1	CHESS F1
Wavelength (Å)	0.950	0.950
Space group	$I4_1$	$I4_1$
Unit cell (Å)		
a,b	83.79	83.85
c	34.88	34.82
Resolution (Å)	29.62 – 1.70	32.11 – 2.00
Completeness, %	96.6	94.5
Observed reflections	121796	109033
Unique reflections	12785	8240
R_{sym}, (%)[a]	9.0	11.5
Refinement statistics		
Highest resolution (Å)	1.91	2.00
R factor (%)[b,c]		
Working	22.4	20.9
Free	26.4	24.5
rms deviation[d]		
Bond lengths (Å)	0.037	0.043
Angles, [o]	2.77	2.71
Average B-factor (Å)	29.9	32.6

[a] $R_{sym} = \sum|I - \langle I \rangle| / \sum I$
[b] $R = \sum_{hkl}||F_o| - |F_c|| / \sum_{hkl}|F_o|$
[c] R_{work} calculated from set of reflections in which 5% of the total reflections were randomly omitted from the refinement and used to calculate R_{free}.
[d] rms = root mean square

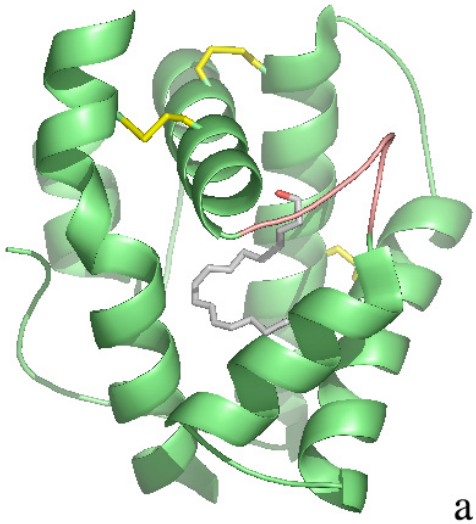
For each complex, only one molecule was found in the asymmetric unit. Model building was done with O (7). The structure was completed and refined by iterative cycles of model building and simulated annealing using O and CNS v.1.1 (8). The eight C-terminal residues were omitted in the final refinement of each complex due to structural disorder.

Results

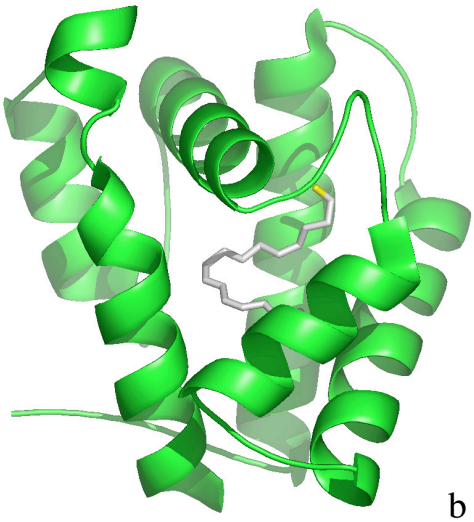
Overall structures. All structures were solved by molecular replacement using the crystal structure of *B. mori* PBP bound to bombykol at pH 8 (Fig. 3.2, ref. 9). The ligand was removed from the model for molecular replacement. In the bombykol complex, the protein is a dimer of identical alpha helical monomers. Four of six alpha helices converge to form a hydrophobic binding pocket for bombykol, and three disulfide bonds stabilize the structure. A loop region between helix 3 and 4 is believed to provide entrance for the ligand by becoming destabilized upon protonation of one or all of three histidine residues at low pH (9). The bound bombykol has a roughly planar, hook-shaped conformation within the binding pocket. The hydroxyl group of bombykol forms a hydrogen bond with the sidechain of a Ser56, and one set of double bonds in bombykol is sandwiched between the Phe12 and Phe118 aromatic rings.

Both the ihd-BmPBP and bpo-BmPBP structures (Fig. 3.2) show an overall protein conformation similar to that of the bombykol-bound structure. The C-terminal region of the protein exists as a disordered loop, and the N-terminal region is an ordered helix. All three disulfide bonds are present. The electron density in the pockets of these complexes clearly fits the modeled ligands. Backbone alignments of the proteins were performed using

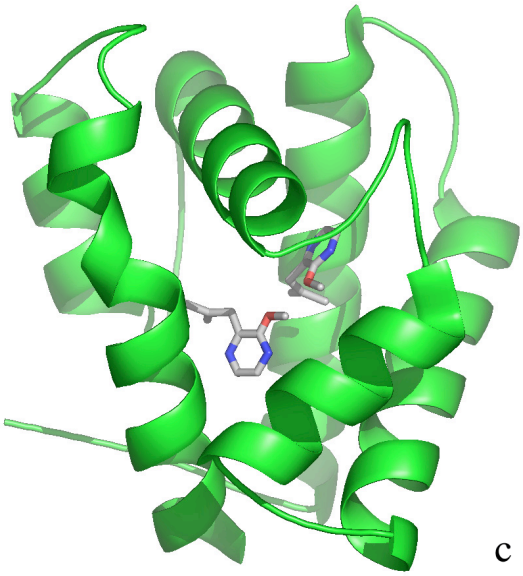
Figure 3.2. Crystal structures of *Bombyx mori* PBP in complex with ligands. Secondary structures are depicted in *ribbon* representations. Ligands are depicted in ball-and-stick figures with CPK colors. *a*, previously described bombykol complex, bom-BmPBP, at 1.8 Å resolution. *b,* iodohexadecane complex, ihd-BmPBP, at 1.9 Å resolution. *c,* bell pepper odorant complex, bpo-BmPBP, at 2.1 Å resolution. Pictures generated using PyMol.



a



b



c

LSQMAN (10). The rms deviation between mainchain C_ atoms of the iodohexadecane- and bombykol-bound proteins was 0.757 Å, and between the bell pepper odorant- and bombykol- bound proteins was 0.775 Å, indicating nearly identical backbone conformations among the complexed PBPs (Fig. 3.3). The only region where mainchain atoms do not align is in the loop between helices □3 and □4. This loop contains the three completely conserved histidines believed to cap the binding pocket and allow pH-dependent entrance and egress of the ligand. Loops are naturally floppy, and the structural differences in loop positioning could be an artifact of differences in crystal packing or represent the motion of the protein. Alternatively, slight structural or energetic changes may occur when the PBP is bound to non-pheromone ligand that act to prime the "incorrect" molecule for ejection from the pocket long before the protein arrives at the receptor. The volumes of the binding pocket cavities were calculated using VOIDOO (11) and were defined as 166.9 $Å^3$, 200.0 $Å^3$, and 207.9 $Å^3$ for the bombykol, iodohexadecane, and bell pepper odorant complexes, respectively. The deviation in the calculated volume of the bombykol complex binding pocket could be due to slight changes in the orientation of side chains of residues surrounding the pocket, defining the borders of the cavity in a different manner. This could depict some shifting in the orientation of side chains in the binding pocket of the protein, perhaps indicating that bombykol is more tightly enclosed within the cavity. Structural alignments of the C□ backbones of the complexes do not indicate any obvious differences in positioning of the residues that form the binding pocket (Fig. 3.4). In this figure, binding pocket residues are grouped by their degree of conservation among lepidoteran pheromone- and odorant-binding proteins. Figures 3.4a and 3.4d display residues that are highly

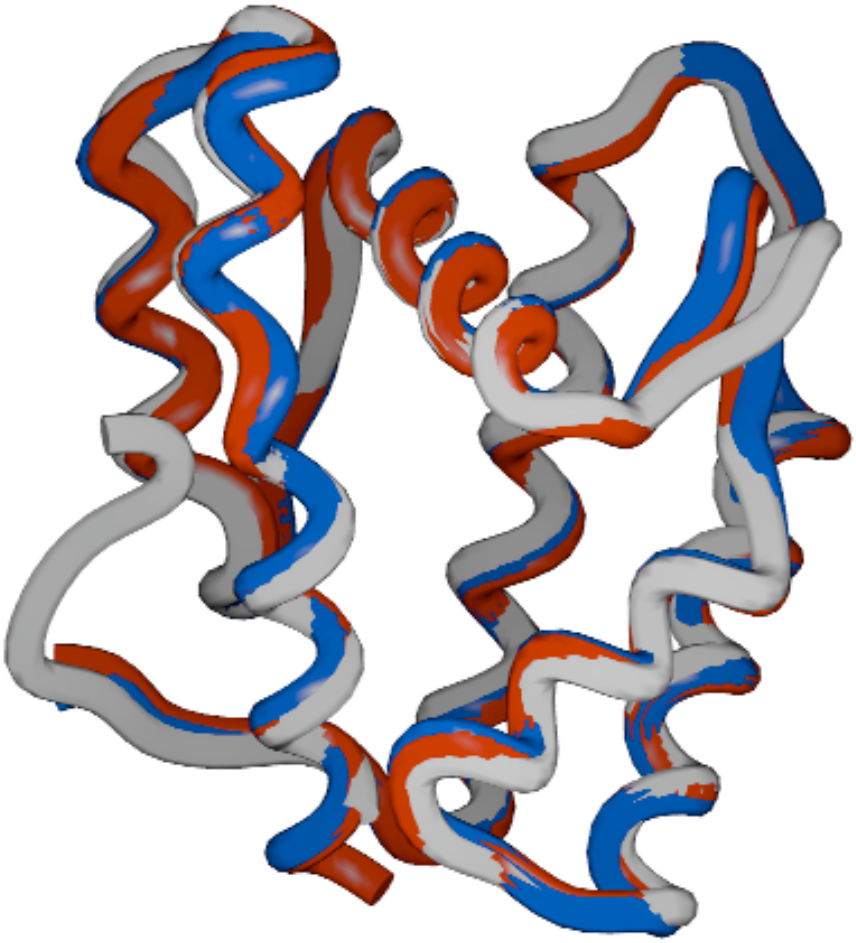
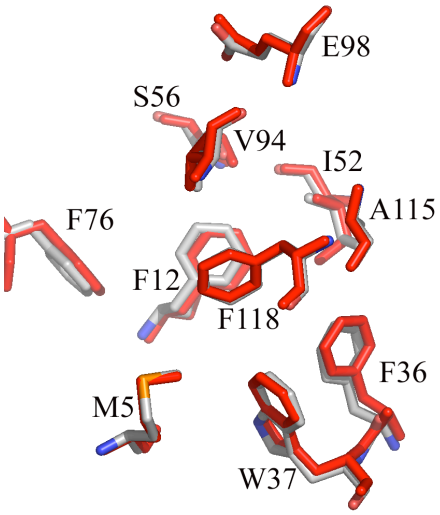
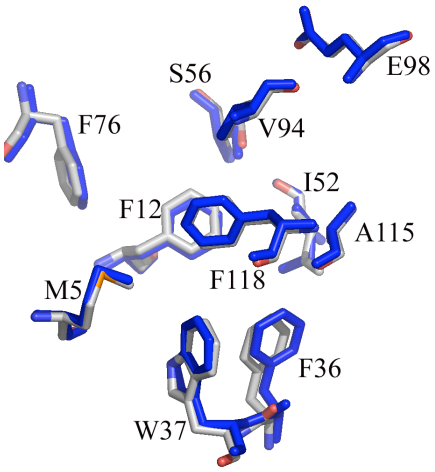


Figure 3.3. Overlay of PBP structures generated by superpositioning with LSQMAN. Backbones depicted in worm. Protein complex bom-BmPBP in gray, ihd-BmPBP in red, and bpo-BmPBP in blue. Root mean square deviation between ihd-BmPBP and bom-BmPBP mainchain Cα atoms is 0.757 Å; root mean square deviation between bpo-BmPBP and bom-BmPBP mainchain Cα atoms is 0.775 Å. Picture generated in SPOCK.

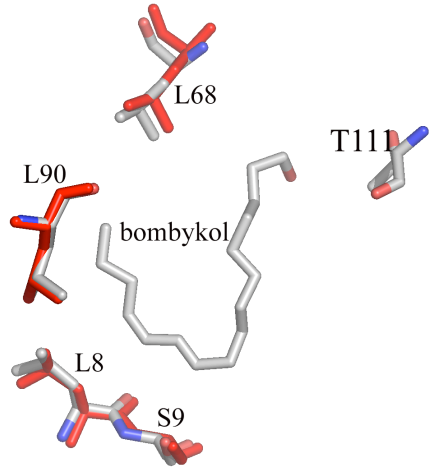
Figure 3.4. Structural alignments binding pocket residues of the *B. mori* PBP complexes. Bombykol-bound PBP is displayed in CPK colors, iodohexadecane-bound PBP in red, and bell pepper odorant-bound PBP in blue. Figures ***a*** and ***d*** highlight residues of the binding pocket that are highly conserved among lepidopteran PBP and OBP families; ***b*** and ***e*** show residues that are highly conserved only among lepidopteran PBPs; ***c*** and ***f*** portray residues that are least conserved in the binding pocket of moth PBPs. Pictures generated with PyMol.


E98
S56
V94
I52
A115
F76
F12
F118
F36
M5
W37

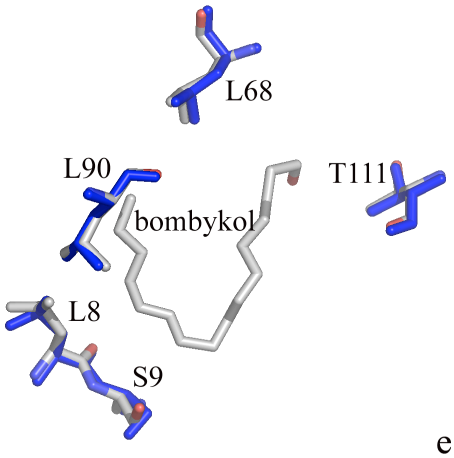
a


S56
E98
F76
V94
I52
F12
A115
M5
F118
F36
W37

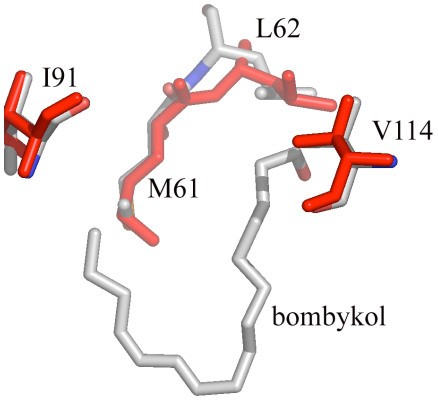
d


L68
T111
L90
bombykol
L8
S9

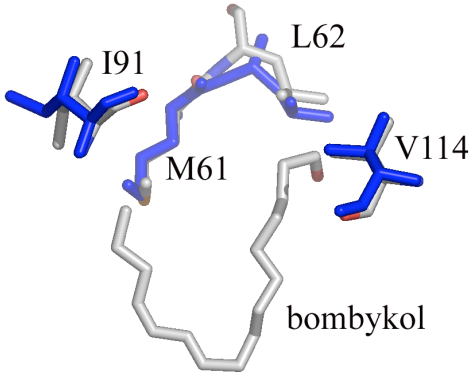
b


L68
L90
T111
bombykol
L8
S9

e


L62
I91
V114
M61
bombykol

c


L62
I91
M61
V114
bombykol

f

conserved among the moth OBP family of proteins. No discernible movement is displayed in these residues in the bombykol, iodohexadecane, and bell pepper odorant bound complexes. Figures 3.4b and 3.4e depict residues that are highly conserved among pheromone-binding proteins, but not among the odorant binding protein family. Very slight changes in the positioning of residues Leu8, Ser9, and Leu68 are seen among the complexes. Figures 3.4c and 3.4f depict residues of the binding pocket that are least conserved among lepidopteran pheromone binding proteins. Larger, but still moderate, shifting is observed for residues Leu62 and Ile91. These shifts do not appear to correlate highly with differences in ligand geometry in the binding pocket, though they could indicate more subtle repulsion between polar and hydrophobic groups. Surprisingly, there appears to be some correlation between the degree of conservation of binding pocket residues and their apparent rigidity when the PBP is bound to various ligands. Highly conserved residues seemingly do not shift to fit different ligands in the binding pocket, whereas the least conserved residues in the binding pocket show a wider range of positioning when the protein is bound to various molecules. This could indicate these semi-conserved residues provide a broader range of flexibility in the binding pocket, and that these residues may be crucial in establishing the affinity of a given moth PBP for its pheromone. Perhaps more puzzling is a closer look at the cavities of each complex as defined by VOIDOO (Fig. 3.5). The outlined bombykol cavity appears completely inaccessible to solvent in the bom-BmPBP structure. However, the binding pockets for both the iodohexadecane-bound and bell pepper odorant-bound PBPs appear to be solvent accessible, as calculated and drawn by VOIDOO. The channel leading out of the cavity is directly opposite the

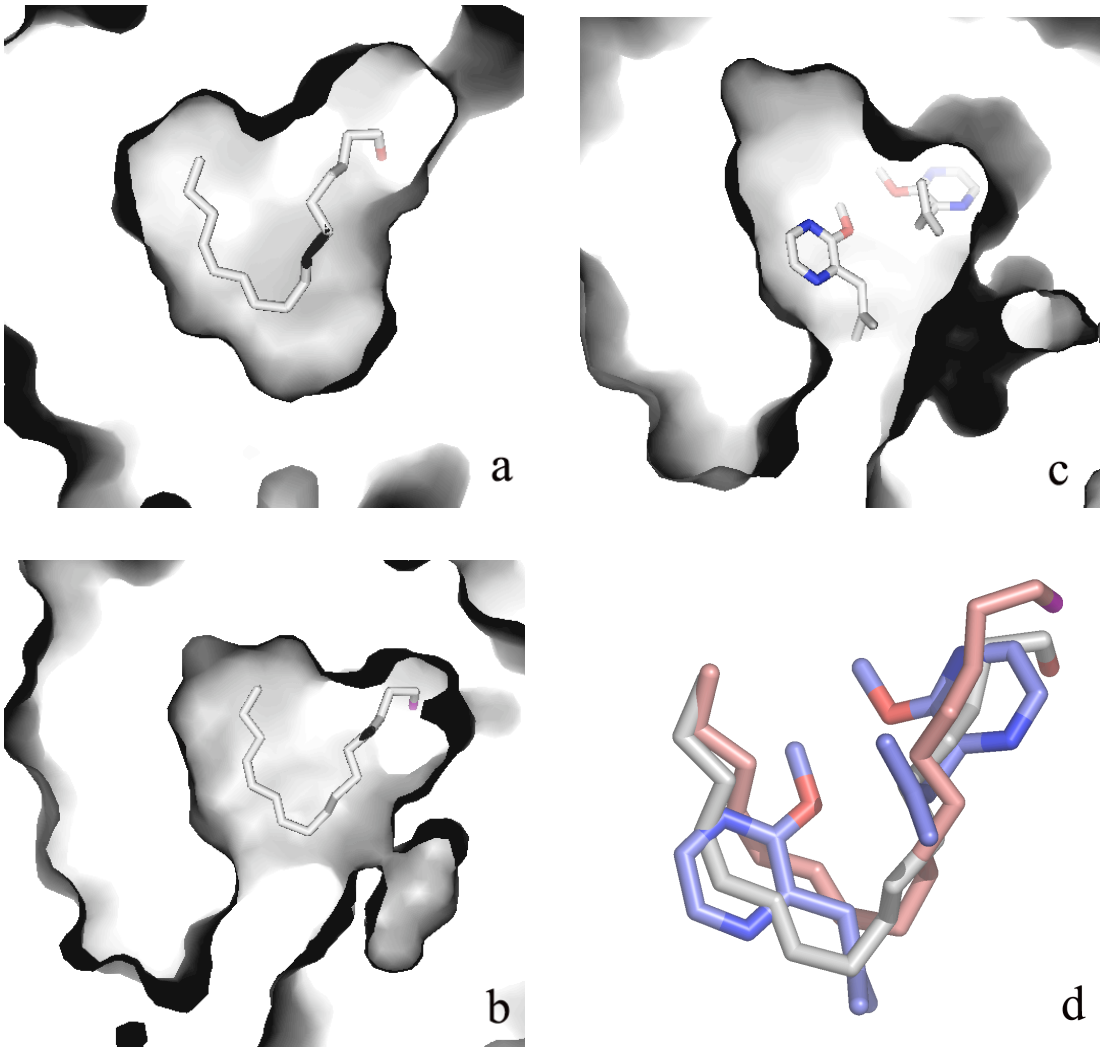


Figure 3.5. Ligand geometries in the BmPBP binding pocket. ***a***, bombykol; ***b***, iodohexadecane; ***c***, bell pepper odorant in the binding cavity defined by bombykol. Tunnels exiting the binding cavity can be seen for the iodohexadecane and bell pepper odorant complexes. ***d***, overlay of all three ligands in aligned BmPBP structures. Figures from PyMol.

histidine loop predicted to play a role in pH-dependent release of bombykol to its receptor. The presence of a channel exiting the binding cavity when the PBP is bound to "incorrect" ligand (*i.e.,* not bombykol) could reveal an interesting way PBPs thwart non-pheromone ligand mobility in the antennal lymph. If "incorrect" ligands create a channel in the protein that makes it favorable for the PBP to release the ligand during transport in the lymph, these molecules are less likely to make it to the receptor. This could speed up pheromone processing at the receptor, as the majority of the molecules that are actually carried to the receptor cell would be bombykol. Very slight structural changes in the binding pocket when pheromone is bound (as compared to other ligands) coupled with higher affinity for bombykol (greater binding energy) could change the conformation of the protein enough to trigger closure of the channel. Or perhaps a greater motion of non-pheromone ligands in the binding pocket contributes to the appearance of a channel. It is also possible that the channels were arbitrarily defined by the VOIDOO program and do not represent novel paths out of the binding pocket of the PBP.

Iodohexadecane complex. The electron density in the binding pocket of the ihd-BmPBP structure is similar to that of bombykol in the PBP. The electron density of the ligand is continuous (Fig. 3.6), and a strong intensity signal presumably indicating the iodine atom at the end of the ligand was easily identified. The occupancy of all iodohexadecane atoms as defined by a B factor was less than 50 except for that of the iodine atom, which had a B factor of 65. This indicates a fairly specific fit of the ligand within the binding pocket. Although the geometry of iodohexadecane is not restrained by double

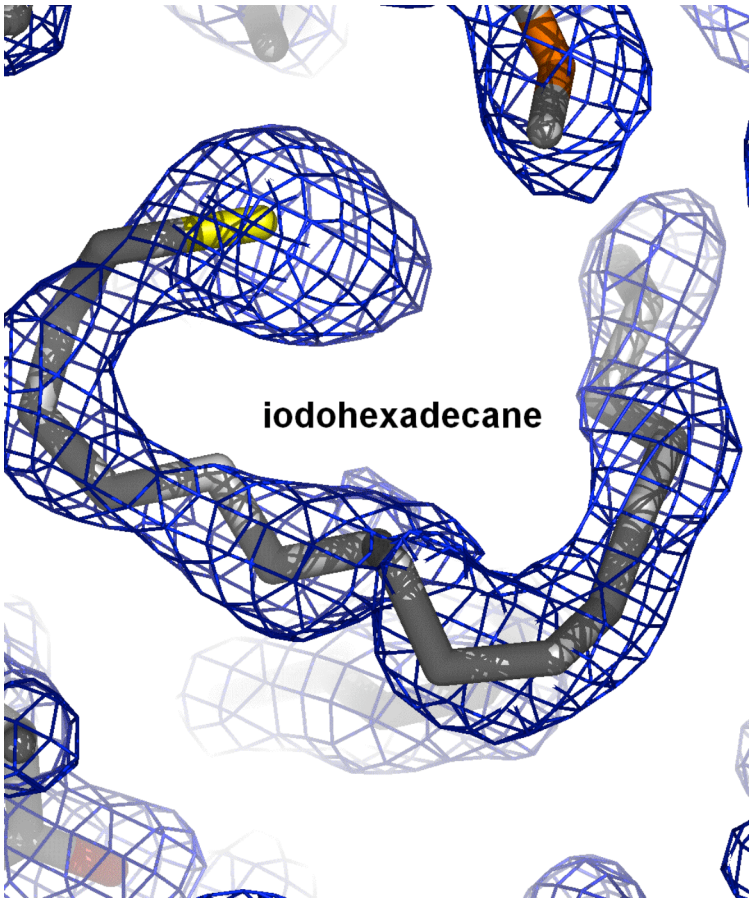


Figure 3.6. $2F_o$–F_c electron density map of iodohexadecane in binding pocket of *B. mori* PBP. Some residues were omitted for picture clarity. Figure from O.

bonds, the ligand adopted a configuration similar to that of bombykol in the binding pocket. As bombykol is locked into this hook-shaped geometry by its double bonds, this may indicate the protein has more affinity for bombykol. This structure clearly shows the iodine atom of the ligand interacting with Ser56, mimicking the interaction of the alcohol group of bombykol with the Ser56 in the pheromone-PBP structure. As modeled, the iodine is 3.7 Å from the sidechain oxygen of Ser56. Other residues located within 4.0 Å of iodohexadecane are Phe12, Leu 62, Leu68, Phe76, Thr111, Val114, Ala115, and

Phe118. All are residues that also compose the binding pocket in the bombykol complex.

Bell pepper odorant complex. Initially, one molecule of bell pepper odorant was modeled into the structure of bpo-BmPBP. The methoxy oxygen of bell pepper odorant lies 3.8 Å from the sidechain oxygen of Ser56. The ligand is within at least 4.0 Å of residues Ile52, Ser56, Leu62, Leu68, Val94, Thr111, Val114, Ala115, and Phe118. An additional comparable mass of electron density was located in the binding pocket, and a second molecule of bell pepper odorant was modeled into the binding pocket of the protein (Fig. 3.7). Addition of this second molecule further improved statistics, indicating *B. mori* PBP was accommodating more than one molecule of this ligand in the binding pocket. The curving of the ring in the second bell pepper odorant molecule is reflected in the electron density. As modeled, this second molecule is within at least 4.0 Å of residues: Leu8, Ser9, Phe12, and Phe36. The closest atom to the methoxy oxygen of this second bell pepper odorant molecule is Phe12, which is 3.6 Å away. The two bell pepper odorant molecules are separated by a distance of 2.9 Å. All residues located close to these two ligand molecules were also involved bombykol-PBP interactions. Two water molecules are modeled into the binding pocket. The occupancy of all bell pepper odorant atoms as defined by a B factor was between 45 and 60. The average B-factor of the bell pepper odorant atoms was 53.9; this may reflect more flexibility in the positioning of the bell pepper odorant molecules in the binding pocket. The high B-factor could also indicate a lower ligand occupancy. This may indicate that bell pepper odorant is bound less tightly by the protein than is iodohexadecane.

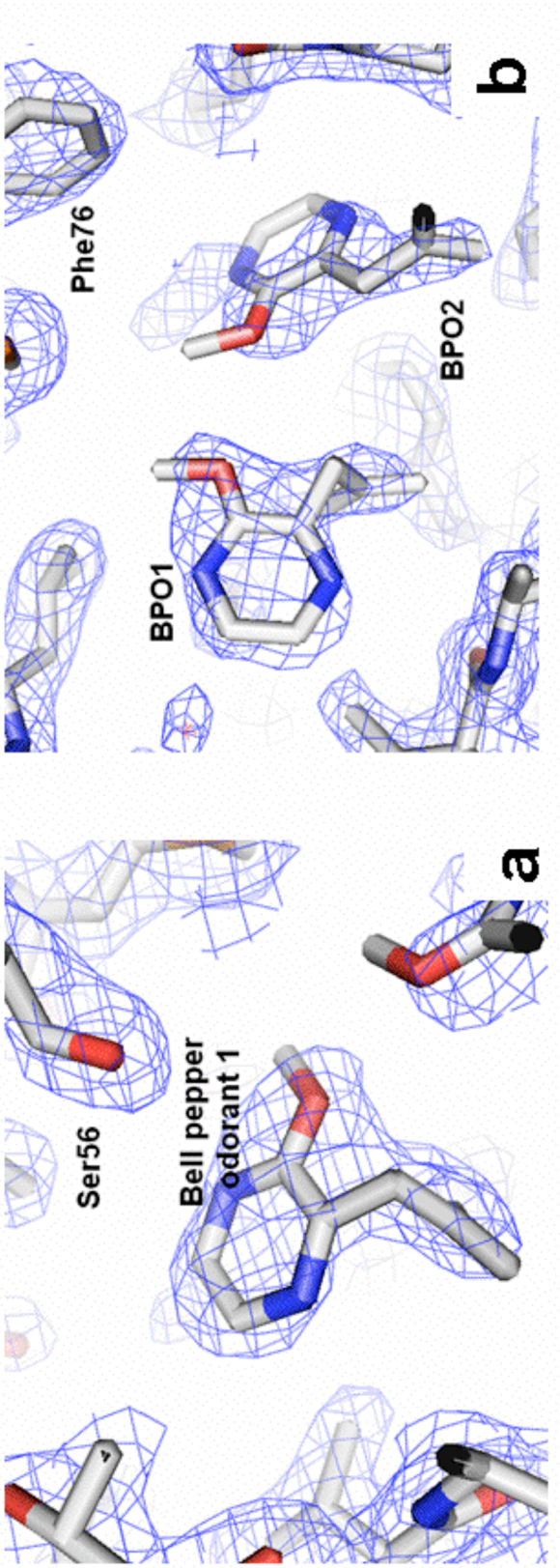


Figure 3.7. $2F_o$–F_c electron density maps of bell pepper odorant in PBP binding pocket. *a*, first bell pepper odorant molecule modeled in. *b*, both molecules in the final structure. BPO is bell pepper odorant. Pictures from O.

Discussion

PBP promiscuity. From these two complexes, we believe *B. mori* PBP can accommodate molecules a variety of hydrophobic molecules within its binding pocket. Polar groups can interact with Ser56, as seen through interactions with the methoxy group of bell pepper odorant. Two molecules of bell pepper odorant can be fit into the electron density in the binding pocket of the protein. It is interesting that bell pepper odorant can be bound by the protein as it contains more polar character than bombykol, has a ring rather than a chain, and has different polar groups to be accommodated.

The positioning of the ligands within the binding cavity and overlay of these ligands are shown in Fig. 3.4. Bombykol, iodohexadecane, and bell pepper odorant essentially fill the same space within the binding pocket. The structures of these two new complexes show ligands with very different geometries can fit into the cavity of the PBP, from straight chain carbon compounds to aromatic molecules. However, the hook structure adopted by iodohexadecane (similar to the shape of bombykol in the binding pocket) might indicate that appropriate positioning of trans and cis double bonds is influential in binding long chain ligands with restrained geometries. This would suggest a compound like bombykol, the sex pheromone, might be bound preferentially over other ligands due to favorable geometry and protein-ligand interactions. As described in results, the calculated volume of the binding pocket in the bombykol-bound structure was somewhat smaller than for the non-pheromone complexes, further suggesting bombykol is more tightly bound by the PBP. Additionally, a channel may appear when PBP is bound to a molecule other than pheromone that facilitates egress of these "incorrect" ligands from the binding pocket before the protein reaches the

receptor. The discovery that the binding pocket of *B. mori* PBP can accommodate non-pheromone ligands is not so surprising considering that in a form of the protein, the C-terminal tail of the protein occupies the cavity (12).

Comparison of B. mori PBP complexes to other insect antennal binding proteins. Other small antennal binding proteins have demonstrated binding to more than one kind of compound. In the cockroach PBP structure (13) a fluorescence reporter, amino-naphthalene sulfonate, was bound to the protein. This molecule was found to be displaced by two out of four components of the species pheromone blend: 3-hydroxy-butan-2-one, and butane-2,3-diol. The apo structure of this protein contains a glycerol molecule from protein preparation in its binding cavity. PBPs from *Antheraea polyphemus* and *Antheraea pernyi* were found to bind both pheromone and non-pheromone components as revealed by native gel electrophoresis (14). Each protein binds a particular pheromonal component preferentially. The structure of a chemosensory protein (CSP) from *Mamestra brassicae* (15) revealed three molecules of 12-bromo-dodecanol in the binding pocket of the protein, reminiscent of our observation of two bell pepper odorant molecules in the *B. mori* PBP complex. This provides evidence that other antennal binding proteins have somewhat adaptable binding pockets, though certain ligands may be bound with higher affinity than others.

It is important to note our experiments are not affinity experiments, and that a 10-fold excess of ligand was incubated with the protein. While iodohexadecane and bell pepper odorant fit in the binding pocket of *B. mori* PBP, the equilibrium of ligand binding and dissociation was not explored.

References

1. Kaissling, K.E., Priesner, E. (1970) *Naturweissenschaften* **57**, 23-28.

2. Wojtaset, H., Leal, W.S. (1999) *J. Biol. Chem.* **274**, 30950-30956.

3. Oldham, N.J., Krieger, J., Breer, H., Svatos, A. (2001) *Chem. Senses* **26**, 529-531.

4. Otwinowski, Z., Minor, W. (1997) *Methods Enzymol.* **276**, 307-326.

5. Collaborative Computation Project, Number 4 (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50**, 760-763.

6. Vagin, A., Teplyakov, A. (2000) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **56**, 1622-1624.

7. Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard. (1991) *Acta Crystallogr. Sect. A* **47,** 110-119.

8. Brunger, A.T., Adams, P.D., Clore, M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M, Simonson, T., Warren, G.L. 1998. *Acta Crystallogr. Sect. D Bio.l Crystallogr.* **54**, 905-921.

9. Sandler, B.H., Nikonova, L., Leal, W.S., Clardy, J. (2000) *Chem. & Biol.* **7**, 143-51.

10. Kleywegt, G.J. and Jones, T.A. (1994) CCP4/ESF-EACBM Newsletter on Protein Crystallography **31,** 9-14.

11. Kleywegt, G.J., Jones, T.A. (1993) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50,** 178-185.

12. Horst, R., Damberger, F., Luginbuhl, P., Guntert, P., Peng, G., Nikonova, L., Leal, W.S., Wuthrich, K. (2001) *PNAS* **98**, 14374-9.

13. Lartigue, A., Gruez, A., Spinelli, S., Riviere, S., Brossut, R., Tegoni, M., Cambillau, C. (2003) *J. Biol. Chem.* **278**, 30213-30218.

14. Maida, R., Ziegelberger, G., Kaissling, K.E. (2003) *J. Comp. Physiol. B* **173**, 563-73.

15. Campanacci, V., Lartigue, A., Hallberg, B.M., Jones, T.A., Giudici-Orticoni, M., Tegoni, M., Cambillau, C. (2003) *PNAS* **100**, 5069-5074.

16. Kraulis, P.J. (1991) *J. Appl. Crystallogr.* **24**, 946-950.

17. Merritt, E.A., and Bacon, D.J. (1997) *Methods Enzymol.* **277**, 505.

CHAPTER FOUR

Elucidation of the structure of the functional PaaP substrate in the biosynthesis of the pantocin A antibiotic

Introduction

Fire blight is a bacterial disease of rosaceous plants that can ruin entire orchards and is the most devastating bacterial disease of apples and pears worldwide (1). The disease was first discovered in the Hudson Valley region in the 1800s, and has since spread to Europe, the Mediterranean, the Middle East, New Zealand, and Australia. The sporadic nature of fire blight outbreaks and their destructive potential make the disease difficult and costly to control.

Fire blight gets its name from the scorched appearance of infected plants. Losses range from delay of fruit bearing and loss of tree limbs to death of individual trees to loss of entire crops. The effect was particularly devastating on the pear industry in Midwest North America; pear was once a popular crop in the Midwest, comparable to the apple industry in the Northeast. Fire blight largely eliminated the possibility of commercial pear production in the region (2).

The disease is caused by the bacterium *Erwinia amylovora* (3). The pathogen is harbored on cankers formed on trees, particularly in the winter when hailstorms are prominent in infected regions and cause physical damage to trees. In the following spring, a combination of wet weather and warmer temperatures allows the bacterium to take hold and spread; blossoms wilt and die, fruit turns brown, and bacteria seeps from blossoms. Insects, particularly bees, spread the bacteria from tree to tree. The bacterium can work its way to

the rootstock to spread the disease to an entire orchard, ruining entire crops of apple, pear, and other rosaceous tree crops.

Management of fire blight is largely preventative, involving selection of resistant cultivar trees and sanitizing cankers in winter. Spraying streptomycin and copper compounds on infected blossoms to stop further spread and pruning diseased trees are also implemented treatment techniques. Nature, however, has developed its own way of controlling the disease. *Erwinia herbicola*, a related bacterium, is an effective biocontrol agent of fire blight (4). In the competition for resources, bacteria have developed antibiotic-producing pathways to defend against competing microbes. Many strains of *Erwinia herbicola* have been found to produce antibiotics effective against the *Erwinia amylovora* pathogen. Work done previously by our group identified a variety of antibiotics produced by *E. herbicola* effective against *E. amylovora*, the molecular basis for antibiotic activity, and the target of these antibiotics (5,6).

Genomic libraries of biologically relevant strains of *E. herbicola* led to the identification of antibiotic-producing clones of *Escherichia coli*. Bioassay-guided fractionation of large-scale fermentations of these clones through chemical chromatography led to the isolation of pantocin A, one of a set of compounds found that exhibits antibiotic activity against *E. amylovora*. The structure of the compound was elucidated through mass spectrometry and labeled NMR studies (Fig. 4.1).

This heterologous expression approach to natural product isolation has the added benefit of access to the DNA-encoding biosynthetic machinery responsible for pantocin A production. Gene cluster sequencing led to the discovery of three major ORFs in the biosynthetic pathway, a finding that was

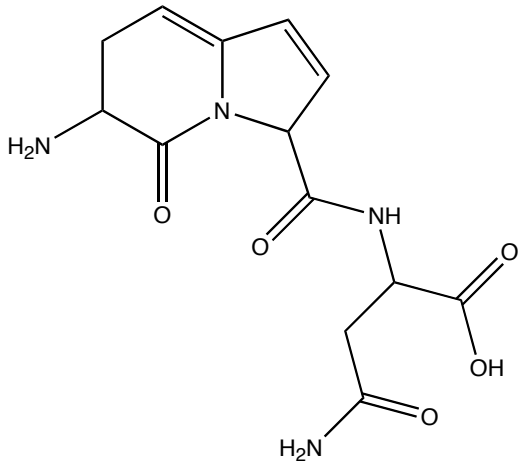


Figure 4.1. Structure of pantocin A

further confirmed by transposon mutagenesis. An additional small ORF was found to be essential for pantocin A activity. The genes comprising the pathway, *paaA, paaB, paaC,* and *paaP*, were cloned and transformed into *E. coli* and found to confer antibiotic activity against *E. amylovora* (5,6,7).

From sequence similarity searches, it was deduced that *paaA* and *paaB* provided the machinery to synthesize pantocin A. PaaA is a 42 kDa soluble protein showing sequence similarity to members of the bacterial ThiF/MoeB/HesA family (37-44% identity). Residues 131 to 270 appear to form a conserved ATP-binding domain, as visualized in the MoeB crystal structure (1JWB). All critical ATP-binding residues are conserved in PaaA. The function of this family of proteins is to activate a carboxy-terminal glycine carboxylate from a partner protein via adenylation. PaaB is a 26 kDa protein with proposed similarity to the Pfam 2OG-Fe(II) oxygenase superfamily as indicated through motif scanning. This family of proteins performs oxidation reactions such as hydroxylation, oxidative ring closure, and desaturation. Isopenicillin N synthase belongs to this family of enzymes, and is responsible for forming the beta-lactam and thiazolidine rings of isopenicillin N antibiotic.

The *paaC* gene was predicted to encode a transmembrane protein and was suspected to act as a transporter pump. It was later found that a plasmid carrying *paaC* confers resistance to extracellular pantocin A in *E. coli* (6).

The scaffold of pantocin A suggested it had a peptide origin. The biosynthesis of peptide antibiotics usually occurs using one of two methods: nonribosomal peptide synthetase production, or posttranslational modification of peptide precursors. The size and organization of the *paaPABC* gene cluster made posttranslational modification of a precursor the more likely candidate. The *paaP* gene encodes a peptide of 30 amino acid, is essential for antibiotic production, and was therefore considered a candidate for the peptide precursor. The peptide contains two glutamate residues at positions 16 and 17, likely candidates for pantocin A's bicyclic ring moeity. An asparagine located at position 18 is also found in the structure of the antibiotic. Isotope-feeding experiments confirmed the presence of these glutamate residues in the pantocin A structure. The discovery of a minor product (PA2) in fermentation cultures showed the subsequent alanine (at position 19 in PaaP) attached C-terminal to the aspargine moiety of the major product, further confirming PaaP as the substrate of the antibiotic biosynthetic pathway. The *paaP* gene encodes a peptide of 30 amino acids. The two glutamates hypothesized to participate in bicyclic ring formation are mid-sequence, at positions 16 and 17. The final pantocin A structure consists of the bicylic ring with an attached asparagine group, presumably derived from the asparagine residue at position 18 of the PaaP peptide (5,6).

From the predicted functions of each ORF, a mechanism was proposed for the biosynthesis of pantocin A (7, Fig. 4.2). All steps are proposed to be carried out by PaaA and PaaB.

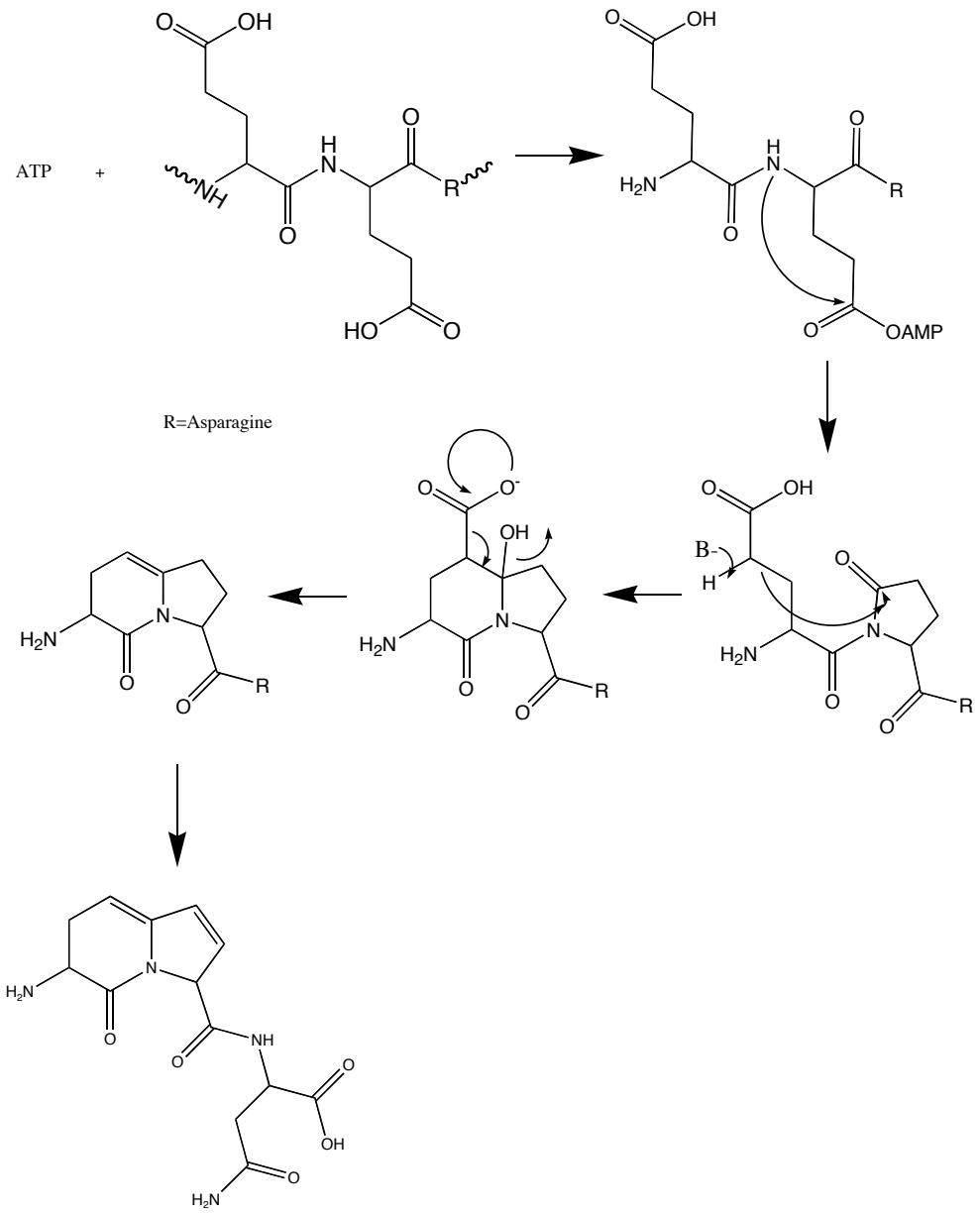
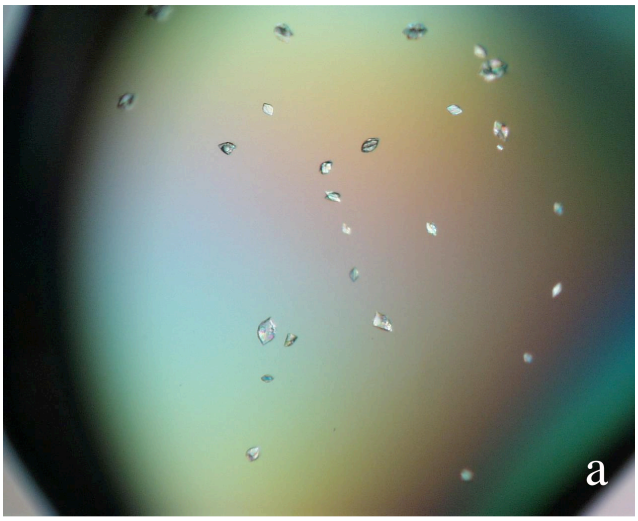


Figure 4.2. Proposed mechanism for pantocin A biosynthesis. R=Asn

Next, the biological function of pantocin A was explored. Pantocin A exhibits activity against *E. amylovora* at an IC_{50} of 200 nM. Activity could be abolished by adding histidine into the test minimal media, providing evidence that the target of pantocin A was an enzyme of the *E. amylovora* histidine biosynthetic pathway (7). *Erwinia amylovora* is closely related to *Escherichia coli.* Competition assays using *E. coli* histidine biosynthetic pathway enzymes in the presence and absence of pantocin A found the antibiotic inhibited the enzyme L-histidinol phosphate aminotransferase. Competition experiments with the tripeptide Ala-Gly-Gly provided evidence pantocin A entered cells through a tripeptide transporter (7).

Initial attempts to crystallize PaaA failed. Crystallization conditions were screened using kits from Hampton and Emerald Biosciences, and gave no protein crystals. Further screening of wells exhibiting good precipitation led to identification of a condition that gave crystals that diffracted poorly on an in-house source. These crystals appeared under the conditions 3% PEG 20,000, 0.1 M Tris pH 8.4 at 4 °C (Fig. 4.3a). Addition of $ZnCl_2$ additive resulted in larger crystals that also diffracted poorly (Fig. 4.3b). Dissolved crystals ran as a band on a protein gel at a molecular weight of ~45 kDa, corresponding to that of the tagged PaaA construct. Addition of ATP did not improve crystal quality. Addition of a truncated PaaP peptide TEENA, which contains the three residues that form the pantocin A antibiotic, caused the protein to form a dense particulate reminiscent of denatured protein. A slight variation of this peptide, TEELA, also failed to produce crystals. Dehydration, seeding, crystal annealing, glutaraldehyde crosslinking, and room temperature data collection in capillary tubes did not improve diffraction quality. At this point, radiolabelled pyrophosphate competition assays,



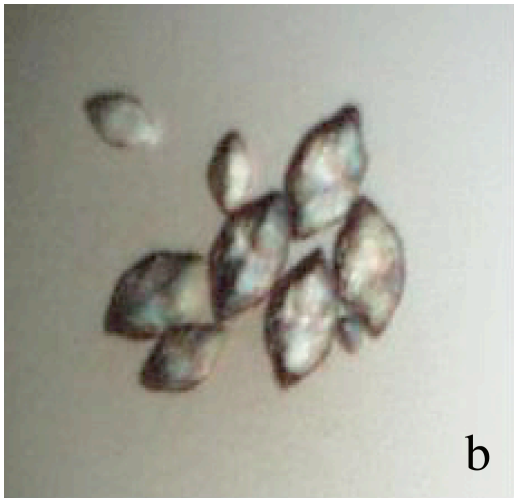


Figure 4.3. PaaA crystals. a) in well containing 3% PEG 20,000, 0.1 M Tris pH 8.4, b) in well containing 3% PEG 20,000, 0.1 M Tris pH 8.4, and 0.01 M $ZnCl_2$.

ATP/NADH coupled ATPase assays, and mass spectroscopy of *in vitro* PaaA reactions using the TEENA peptide were performed to determine whether TEENA is an effective substrate. Results were inconclusive. In view of the importance of knowing the structure of the functional substrate of PaaA for both protein crystallization and elucidation of enzyme mechanism, a comprehensive genetic assay was developed to determine the structure of the functional substrate of PaaA.

There are three obvious possibilities for the structure of the functional PaaP-derived substrate: the complete 30mer PaaP peptide; a shortened version of the peptide derived by autocatalytic processing, or a processed version of the peptide derived by proteolysis *via* a cellular protease. In the first case, PaaA and/or PaaB would play a role in cleaving the functional form of the substrate at positions 15 and 19 or 20. PaaA is a 42 kDa enzyme, large enough to have two domains. One region of the enzyme, formed by residues 131-275, is believed to contain an ATP-binding cassette that is predicted to adenylate the side chain of the glutamate at position 17 of PaaP, leading to cyclization through nucleophilic attach by the backbone amide nitrogen. Residues 131-275 may be one of two domains of this enzyme; a second domain could be responsible for cleaving the 30mer peptide prior or subsequent to this cyclization, though the potential presence of two domains in PaaA and their function has not been studied. PaaB is a 26 kDa enzyme and could also consist of two domains, with one domain cleaving the substrate, though this is also speculative. Alternatively, hydrolysis of the PaaP peptide could be the consequence of ring cyclization or some other enzymatic reaction. In total, at least two cleavage reactions would need to occur to excise the central portion of PaaP found in the final pantocin A product. PaaA or PaaB could be

involved in one or both of these cleavages. However, the *paaPABC* gene cluster contains only two enzymes responsible for the complete biosynthesis of pantocin A. These two enzymes are responsible for a predicted five-step mechanistic process, not including peptide cleavages, suggesting PaaA and PaaB may not be not responsible for cleaving the 30mer substrate. Cleavage by cellular proteases or peptide autocleavage are also potential ways for a peptide to be processed prior to use as substrate. Our goal was to establish which residues of PaaP are necessary for pantocin A production.

Further elucidation of the mechanism of pantocin A biosynthesis using *in vitro* and X-ray crystallographic studies will require identification of the active substrate. A genetic assay was used to detect essential residues of PaaP in an effort to identify the functional form of PaaP substrate in pantocin A biosynthesis. Addition of substrate to enzyme in cocrystallization or soaking experiments can sometimes greatly improve the quality of crystal diffraction, and this may help in the case of PaaA.

Materials and Methods

Expression and purification of PaaA. DNA for the *paaA* gene was inserted into a pET30b vector. Cultures of *Escherichia coli* expressing the construct were grown in LB for two hours at 37 °C, then induced with 1 mM IPTG and 0.2% arabinose. Growth continued 2-4 additional hours at 37 °C. Cells were harvested by centrifugation and protein recovered by sonication. Protein was purified by nickel-affinity column by step elution (Qiagen) followed by gel filtration on Superdex 200 (Amersham Pharmacia) at 4 °C. PaaA was concentrated to concentrations of 10, 15 and 20 mg/mL for crystallization trials.

Crystallization of PaaA. Crystallization of PaaA was attempted by hanging and sitting drop diffusion at 4 °C. 2 uL of protein was mixed with 2 uL of well solution, suspended over a well containing 500 uL of solution. Trays were incubated for periods of up to one year.

Crystal dehydration. Two methods of dehydration were tested. In the first, poorly-diffracting crystals from wells containing 3% PEG 20,000 and 0.1 M Tris pH 8.4 were transferred sequentially to wells containg 4%, 5%, 6%, 7%, 8%, 9%, and 10% PEG 20,000 in 0.1 M Tris pH 8.4 for 48 hours each. In the second method, cover slips from hanging drops containing poorly-diffracting crystals were exposed to air for 10 to 30 minutes, then 5 uL of fresh solution containing 5%, 8% or 10% PEG 20,000 was added to the drop and the wells recapped.

Glutaraldehyde crosslinking. Drops containing poorly-diffracting crystals in 3% PEG 20,000, 0.1 M Tris pH 8.4 with and without $ZnCl_2$, $YtCl_2$, and $StCl_2$ additive were crosslinked with 5% v/v glutaraldehyde after 1 week, 2 weeks, and 1 month of growth at 4 °C.

paaPABC construct. DNA corresponding to the *paaPABC* pathway including 500 bases upstream of the start site (to include regulatory and promoter elements) was inserted in a pUC19 vector (NEB) and transformed into DH5 cells for expression.

Cassette design for rapid mutation. Two restriction sites (*BclI* and *Bpu10I*, Fermentas) unique to the *pUC19/paaPABC* plasmid construct mapping to the

5′ and 3′ ends of the *paaP* gene were found using MacVector. Neither of these sites occurs in the pUC19 vector nor in the remainder of the *paaPABC* cluster. The *BclI* restriction site (-TGATCA-) contains the second two bases encoding *paaP* Met1 and all three bases encoding Ile2 (the 6th base of the restriction site is removed during digestion). The *Bpu10I* restriction site (-CCTTAGC-) contains the last base encoding Val28, all three bases encoding Leu29, and the first base encoding Ser30 (the last two bases of the restriction site is removed during digestion). Two sequential digests of *paaPABC*/*pUC19* construct with *BclI* followed by *Bpu10I* resulted in a "cassette" that essentially removes the *paaP* gene. This cassette can be easily loaded with DNA insert allowing easy removal, quick mutation or complete replacement of *paaP* without affecting the positioning of upstream regulatory elements or *paaA*, *paaB*, and *paaC* genes. The prepared cassette vector was dephosphorylated to further optimize cassette loading by ligation.

5mer, 12mer, 30mer paaP mutants. To determine if a minimal sequence of paaP centered around the residues forming the bicyclic ring of pantocin A is sufficient for antibiotic production, DNA coding portions of the paaP gene centered around residues 16-18 were inserted into the cassette for bioassay. Inserts encoding the 5mer (TEENA), 12mer (SAITEENAMYTK), and full-length 30mer (MIKFSTLSQRISAITEENAMYTKGQVIVLS) PaaP peptide were inserted into the pUC19 cassette and transformed into DH5 cells. All plasmids were sequenced prior to use in a crude bioassay (see below) to determine whether the mutant PaaP genes permitted antibiotic production. All cassette constructs maintained residues M1, I2, L29 and S30 due to the sequence of the restriction sites.

N-terminal truncation mutants. The pUC19/*paaPABC* construct was digested with BclI (Fermentas), followed by removal of sticky ends by incubation with Mung Bean nuclease (Epicentre) at 37 °C for 5 minutes. The DNA was gel extracted (Qiagen gel extraction kit) and digested with Bpu10I (Fermentas). The cassette, which at this point has one blunt end at the N-terminus of the *paaP* gene and one sticky end at the other terminus, was dephosphorylated and ligated to annealed DNA primers (Integrated DNA Technologies protocol) encoding the start methionine followed by the *paaP* gene truncated at residues 4, 8, 11, and 15.

Alanine/glycine scanning. Site-directed mutagenesis of the active *paaPABC* clone in a pUC19 vector, expressed in *E. coli* was performed using the GeneTailor Mutagenesis kit from Invitrogen. In brief, the plasmid containing the biosynthetic pathway is methylated, then site-directed mutants are created using mutagenic primers, amplifying the complete 2.5 kb plasmid with a high-fidelity polymerase. PCR product is then transformed into DH5α-T1^{R} cells, which permit replication of the unmethylated mutagenized product. Isolated DNA from transformed clones was sequenced for each PaaP mutant to confirm the alanine mutation (intrinsic alanines were mutated to glycine). Plasmids were then transformed into DH5α cells for bioassays.

Crude bioassay. Bioassays were performed using indicator strain *E. amylovora* strain 273 (Ea273) in a growth challenge assay. A 10-mL culture of Ea273 was grown in LB at 30 °C, shaking at 250 rpm until cells reached an optical density (A^{600nm}) of 0.8. The culture was spun for 10 minutes at 800 x *g*, and cells were resuspended in 1 mL 5mM sodium phosphate pH 6.8. Cells were mixed with

100 mL of 0.7% agar and 27 mL 5X GAT salts (20 g. glucose, 0.3 g L-asparagine, 0.05 g nicotinic acid, 11.5 g K_2HPO_4, 4.5 g. KH_2PO_4, 0.12 g. $MgSO_4$, 100 ug thiamine per 200 mL) for a top agar overlay. 4 mL of the overlay was poured onto 10 mL (100 mm) plates of GAT agar. DH5 clones expressing mutants were spotted onto the hardened overlay and plates were incubated for 40 hours at 30 °C. Zones of inhibition indicating pantocin A antibiotic activity could be seen after 20 hours of incubation. Alanine mutants that affected the ability of PaaP to be used as a substrate resulted in decreased or absence of inhibitory growth zone.

Quantitative bioassay. 1-mL cultures of *E. coli* expressing *paaPABC* with alanine or glycine substitutions in *paaP* (site-directed mutagenesis) were grown overnight at 37 °C. A 10 mL aliquot of GAT minimal medium (20 g. glucose, 0.3 g. L-asn, 0.05 g. nicotinic acid, 0.12 g. $MgSO_4$, 11.5 g. K_2HPO_4, 4.5 g. KH_2PO_4, 100 ug thiamine per liter culture) with 0.6 ug/mL ampicillin was added to each culture. Cultures were incubated for 24 hours at 37 °C to chew up all antibiotic, cells were spun at 4,000 x *g* at 4 °C, and the supernatant was reserved in 50-mL falcon tubes. Supernatant from each mutant was incubated at 90 °C for 1 hour to kill any remaining cells, then diluted 1:10 in fresh GAT media (no antibiotic). 0.1 mL of an overnight culture (30 °C) of *Erwinia amylovora* Ea273 was added to 10 mL of each diluted supernatant, and cultures were grown from 12-16 hours at 30 °C. Growth inhibition of the Ea273 pathogen was monitored by absorbance using a spectrophotometer. The absorbance of 1 mL samples of culture was measured and these measurements standardized to the growth of a control (supernatant from *E. coli* expressing the pUC19 vector incubated with 0.1mL Ea273).

Supernatant bioassay. 10 uL of heat-inactivated supernatant from *E. coli* clones expressing *paaPABC* with *paaP* alanine or glycine substitution mutants (see Quantitative bioassay) was applied to a GAT plate containing a top agar overlay of Ea273 (see Crude bioassay) and incubated overnight at 30 °C.

Results

5mer, 12mer, 30mer. The 30mer clone produced a zone of inhibition similar to that of wildtype *paaPABC*, confirming the cassette system is effectual for rapid mutation of the *paaP* gene. Neither the 5mer nor the 12mer *paaP* constructs produced a zone of inhibition in the crude assay (data not shown), indicating the need for residues distant from the –EEN- residues (the residues that comprise the pantocin A product) for antibiotic production.

Alanine/glycine scanning – crude bioassay. (Fig. 4.4) Alanine substitution of residues Lys3, Phe4, Ile11, Glu16, Glu17, and Phe21 resulted in a loss of Ea273 growth inhibition (no zone). Diminished zones were observed for R10A, S12A, N18A, and Q25A mutations. Substituting leucine at position Ile11 resulted in a zone of inhibition, indicating antibiotic activity. Substituting phenylalanine at position Tyr21 resulted in a zone of inhibition, also indicating antibiotic activity (Fig. 4.5). A reversion of the F21A mutation using site-directed mutagenesis resulted in restoration of antibiotic activity, which implies that this technique was successfully used to change single residues in the 6.5 kB construct.

N-terminal truncations. None of the PaaP N-terminal truncation mutants exhibited antibiotic activity in the crude assay (data not shown).

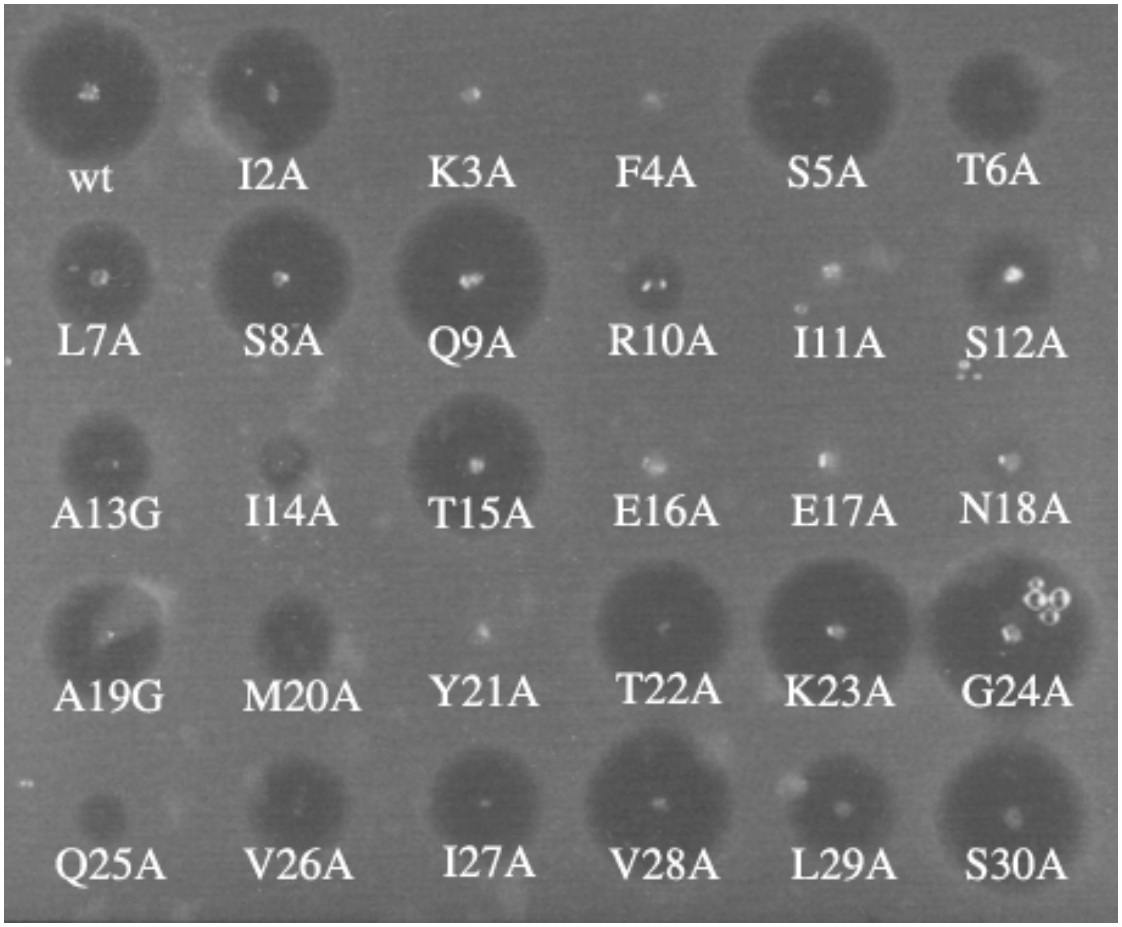


Figure 4.4. Crude bioassay of PaaP mutants. Clones were applied to a Ea273 top agar overlay on a GAT plate and tray was incubated at 30 °C overnight. Zones of inhibition indicate antibiotic activity.

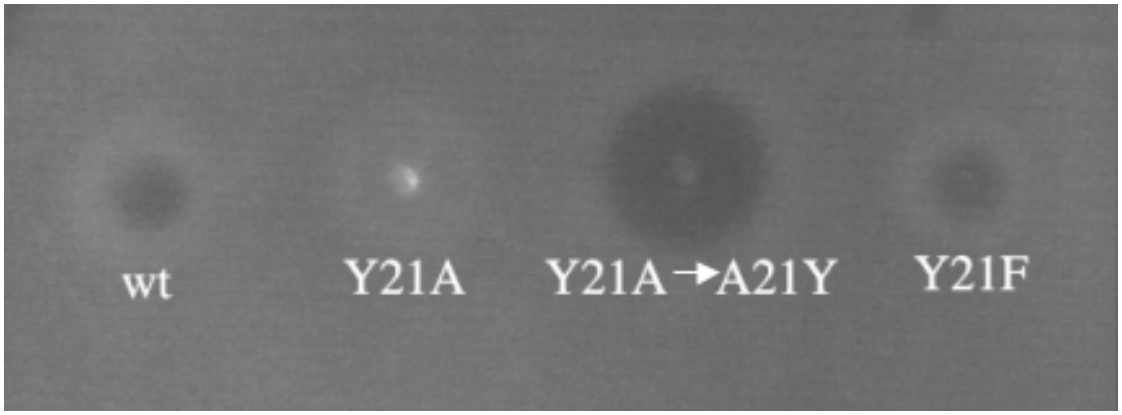


Figure 4.5. Crude bioassay of PaaP mutant with substitutions at residue 21. Wildtype, alanine substitution, reversion from alanine substitution, and conservative phenylalanine mutation. Clones were applied to Ea273 top agar overlay on a GAT plate, and incubated at 30 °C overnight. Zones of pathogen growth inhibition indicate antibiotic activity.

C-terminal early terminations. None of the mutants containing stop codon substitutions at positions 21, 24, and 27 displayed antibiotic activity in the crude assay (data not shown) and were not used in the quantitative assay.

Alanine/glycine scanning – quantitative bioassay. (Fig. 4.6) Five individual experiments were performed with each set of mutants. Supernatants from *paaP* mutants corresponding to amino acid changes K3A, F4A, I11A, E17A, E18A, N19A, and Y21A displayed no apparent antibiotic effect on the Ea273 pathogen. Mutant S12A and Y21F supernatants showed an intermediate level of Ea273 growth inhibition. Supernatants from all other mutations displayed antibiotic activity similar to that of the wildtype *paaPABC* clone, within statistical error.

Alanine/glycine scanning – supernatant bioassay. Results from the supernatant bioassay corresponded with those of the quantitative bioassay (data not shown). Supernatants from mutants K3A, F4A, L11A, E16A, E17A, and Y21A failed to produce zones of inhibition on the Ea273 top agar. Supernatant from mutants S12A and N18A produced diminished zones of inhibition.

Discussion

The inability of the 5mer and 12mer PaaP constructs centered around the bicyclic ring-forming unit of the substrate to display detectable antibiotic activity suggests that PaaP residues distant from the site of enzymatic activity have an essential function in either substrate processing or substrate recognition. As expected, mutating any of the three residues that comprise the pantocin A antibiotic results in a lack of antibiotic activity. The inability of

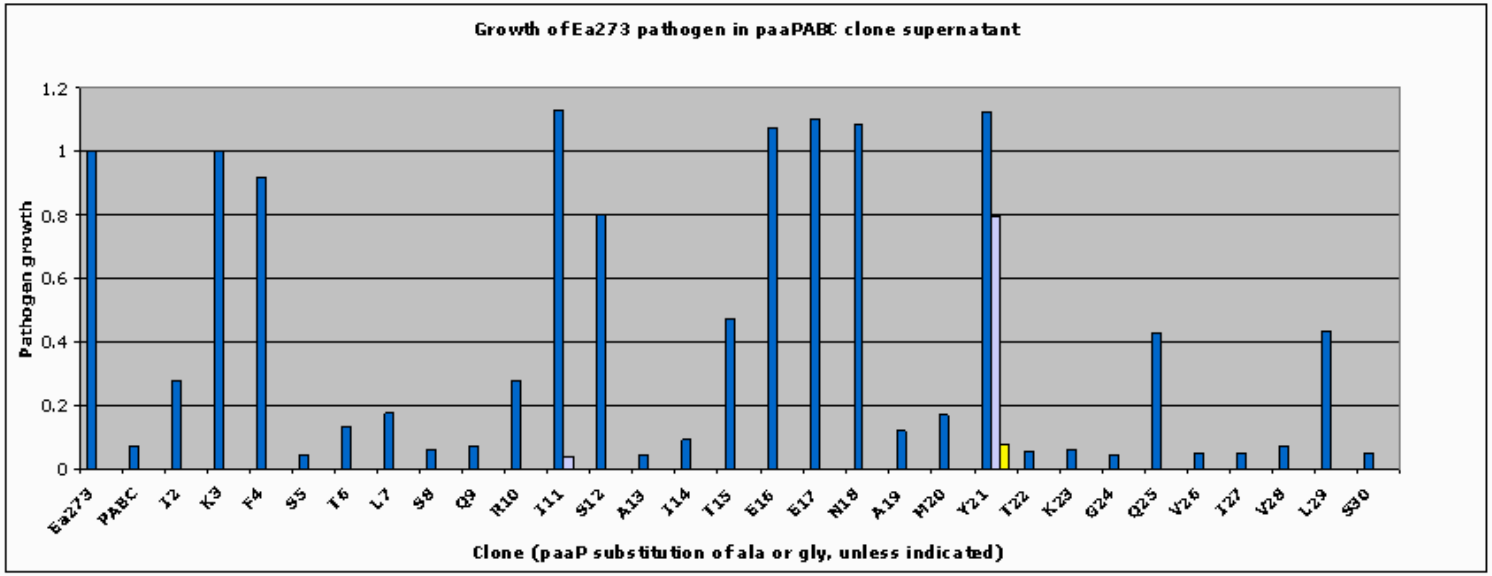


Figure 4.6. Results from quantitative bioassay. Results are standardized to the growth of Ea273 in absence of antibiotic-producing clone. Conservative mutations IllL and Y21F in light blue. The yellow box represents the site-directed mutagenic reversion from the Y21A mutant. A total of five experiments were done with each set of clones.

mutants K3A, F4A, I11A, and Y21A to produce zones of inhibition in the crude, quantitative, and supernatant bioassays indicates these four residues play some specific role in PaaP conversion to pantocin A product. This is somewhat unexpected as these residues are distant from the residues that form the antibiotic. One might expect that residues near the –EEN- reaction center would be important for specificity or recognition as they might have some interaction with an enzyme, such as a protease, PaaA, or PaaB. Surprisingly, this is not the case. T15A, A19G, and A19V substitutions had little to no detectable effect on antibiotic activity. More surprisingly, I11A and Y21A substitutions, mutations in residues distant from the substrate reaction center, had dramatic effects and seemingly abolished pantocin A activity. The hydrophobic nature of Ile11 and Tyr21 and their identification as essential residues suggests autocatalytic processing does not occur. Autoproteolysis of peptides is frequently observed in post-translational modification of peptides and activation of proteins, but is usually promoted through nucleophilic attack by a serine or cysteine side chain (8-12). At a length of 30 amino acids, PaaP would be rather small in comparison to known peptides that exhibit autoprocessing. Isoleucine and tyrosine are bulky, aliphatic residues; such residues are common specificity determinants in the substrate binding of proteases, and they could also be specificity determinants for other enzymes as well. Conservative mutation of Ile11 to a leucine resulted in complete antibiotic activity in both the crude and quantitative assays. Conservative mutation of Tyr21 to a phenylalanine resulted in antibiotic activity in the crude assay, but little activity in the quantitative assay. Residues Lys3 and Phe4 also seem to be important for the production of pantocin A. These four residues may be important recognition sites for PaaA or PaaB.

Alternatively, these residues may form a kind of grip or anchor that positions residues 16-19 in some way that is favorable for catalytic reaction. Calmodulin has been shown to bind to a 26-residue peptide, anchoring it by two residues near the termini of the peptide (13). A tryptophan residue and a phenylalanine residue located at positions 4 and 17, respectively, anchor the peptide that forms a helix and hairpin turn. An analysis showed peptides that bind with high affinity to calmodulin contain aromatic or long-chain hydrophobic residues separated by a stretch of twelve residues. The PaaP peptide could similarly be anchored by L11 and Y21, allowing proper configuration of residues that participate in the enzymatic reaction.

The cassette rapid-mutation system provided a quick and efficient way to replace, remove, and mutate the *paaP* gene. Insertion of DNA encoding the full 30mer PaaP peptide resulted in full antibiotic activity.

References

1. Online publication, http://www.rbgsyd.gov.au/information_about_plants/pests_diseases/fact_sheets/fire_blight

2. Online publication, http://www.ipm.ucdavis.edu/PMG/PESTNOTES/pn7414.html

3. Bonn, W.G., and van der Zwet, T. (2000) Distribution and economic importance of Fire Blight. In *Fire Blight: the Disease and its Causative Agent Erwinia Amylovora.* Vanneste, J.L., (ed.). New York: CABI Publishing, pp. 37-53.

4. Chatterjee, A.K., Gibbins, L.N., Carpenter, J.A. (1969) *Can. J. Microbiol.* **15,** 640-642.

5. Jin, M., Liu, L., Wright, S.A., Beer, S.V., Clardy, J. (2003) *Angew Chem. Int. Ed. Engl.* **42,** 2898-2901.

6. Jin, M., Wright, S.A., Beer, S.V., Clardy, J. (2003) *Angew Chem. Int. Ed. Engl.* **42,** 2902-2905.

7. Jin, M. (2003) *Antibiotics from Erwinia herbicola.* Doctoral dissertation, Cornell University, Ithaca, NY.

8. Ghosh, K., Pal, A., Chattopadhyaya, R. (2004) *Biochem. J.* **379,** 325-330.

9. Rozanov, D.V., Strongin, A.Y. (2003) *J. Biol. Chem.* **278,** 8257-8260.

10. Zhou, Y., Lindberg, I. (1994) *J. Biol. Chem.* **269,** 18408-18413.

11. Shinde, U., Fu, X., Inouye, M. (1999) *J. Biol. Chem.*, **274**, 15615-15621.

12. Doran, J.D., Nomizu, M., Takebe, S., Menard, R., Griffith, D., Ziomek, E. (1999) *Eur. J. Biochem.* **263**,145-151.

13. Osawa, M., Tokumitsu, H., Swindells, M.B., Kurihara, H., Orita, M., Shibanuma, T., Furuya, T., Ikura, M. (1999) *Nat. Struct. Biol.* **6**, 819-824.