# STRUCTURE AND FUNCTION OF PHEROMONE-BINDING PROTEINS FROM THE GYPSY MOTH, *LYMANTRIA DISPAR*

by

Nicolette Sharon Honson
Bachelor of Science, Simon Fraser University 2001

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

In the
Department
of
Chemistry



SIMON FRASER UNIVERSITY

Summer 2006

# APPROVAL

**Name:** Nicolette Sharon Honson

**Degree:** Doctor of Philosophy

**Title of Thesis:** Structure and Function of Pheromone-Binding Proteins from the Gypsy Moth, *Lymantria dispar*

**Examining Committee:**

**Chair:** Dr. D.J. Vocadlo (Assistant Professor)

Dr. E. Plettner (Associate Professor)
Senior Supervisor

Dr. A.J. Bennet (Professor)
Committee Member

Dr. J.L. Thewalt (Associate Professor)
Committee Member

Dr. L. Craig (Assistant Professor)
Internal Examiner

Dr. P.L. Davies (Professor)
External Examiner
Biochemistry Department
Queen's University

**Date Approved:** August 10, 2006

# ABSTRACT

Many species of moth use pheromones to attract mates. Pheromones are detected by olfactory neurons housed in specialized sensory hairs on the antennae. Pheromone-binding proteins (PBPs) are thought to transport the hydrophobic pheromone through the antennal lymph to the olfactory neuron.

The antennal lymph contains a variety of potential ligands for PBPs, none of which have been identified. I have determined the concentrations of male gypsy moth (*Lymantria dispar*) PBP1 and PBP2 *in vivo*. I have also identified four fatty acids and cholesterol as compounds which are found in the sensillar lymph and which associate with either *L. dispar* PBP1 or PBP2. The concentrations of these compounds in the lymph are in the high millimolar range. Experimental evidence supports fatty acid formation into vesicles *in vitro*. I have also determined that PBPs selectively associate with low millimolar concentrations of these vesicles. The existence of vesicles or multilamellar structures in the lymph may be a mechanism of odorant transport. PBPs may not even be necessary for odorant transport, but instead may function in long-term scavenging of odorants. Such scavenging would prevent saturation of odorant receptors, thereby preventing prolonged depolarization of the nerve cell.

Previously, *L. dispar* PBPs have not been structurally characterized. PBPs are highly α-helical proteins containing three disulfide bonds. These disulfide linkages are thought to contribute to the folding and stability of these proteins. The disulfide connectivities were determined, and the second disulfide linkage was found to exhibit

unique chemical and structural properties. To further investigate the role of this disulfide bridge, this bond was chemically modified by cyanylation of PBP1 (PBP1.CN) to lock the protein in this singly reduced form. To determine if this bond plays a role in PBP1 conformational stability, biochemical properties of PBP1.CN were compared with unmodified protein. Modified PBP1 lacking this second disulfide bond was more susceptible to heat denaturation than unmodified protein. To determine if lack of this disulfide bond affects ligand binding, a GC-binding assay was employed to compare pheromone-binding affinities of PBP1.CN with unmodified protein. It was found that reduction of the C2-C5 bond plays a role in ligand discrimination.

**Keywords:** gypsy moth; pheromone-binding protein; fatty acids; disulfide bond; insect

*To Jeremy Pettigrew*

*and my parents, Alice and Duncan.*

"Discovery consists of seeing what everybody has seen and thinking what nobody has thought."

- Albert von Szent-Gyorgy (1962)

# ACKNOWLEDGEMENTS

I would like to thank my supervisor and mentor, Dr. Erika Plettner, for her support, guidance, and many good ideas during the course of my doctoral studies. She taught me all of my biochemistry skills from the ground up, and for that I am extremely grateful. I would also like to thank the members of my supervisory committee, Dr. Andrew Bennet and Dr. Jenifer Thewalt, for their time and helpful suggestions.

My past and present laboratory colleagues, the "Plettner Babes" Ms. Adina Rojubally, Ms. Yongmei Gong, Ms. Taraneh Lajevardi, as well as Ms. Angelica Kowcun, Ms. Violet Hung, Mr. James Inkster, Ms. Ewa Sokolowski, Ms. Sheila Smith, Dr. Srinivas Nagabandi, Dr. Ranjeet Nair, Ms. Johanna Griscti, Ms. Ivy Ling, Mr. Adam Ludlow, and Ms. Mianwei Wong, are all thanked for their friendship and support, especially at "coffee time".

Mr. David Lee, Dr. Philip Patty, and Dr. Barbara Frisken from the Physics Department are thanked for their help with performing the dynamic light scattering experiments. I would like to thank Dr. Margaret Johnson for her expertise in protein modeling. Ms. Regine Gries is thanked for her help with rearing of the gypsy moths. Dr. Mark Paetzel from the Department of Molecular Biology and Biochemistry is also thanked for his guidance in the techniques of protein crystallography and for his time spent obtaining X-ray diffraction patterns of my protein crystals.

Finally, I would like to thank NSERC for an NSERC doctoral postgraduate scholarship and Simon Fraser University for financial support.

# TABLE OF CONTENTS

Approval ..... ii
Abstract ..... iii
Dedication ..... v
Quotation ..... vi
Acknowledgements ..... vii
Table of Contents ..... viii
List of Figures ..... xi
List of Schemes ..... xiii
List of Tables ..... xiv
List of Abbreviations ..... xv

**Chapter 1: Introduction ..... 1**

1.1 Lepidopteran Pheromone Communication ..... 1
1.2 Detection of Pheromones and General Odorants ..... 4
1.2.1 An Overview of Insect Olfaction and Gustation ..... 4
1.2.2 Perireceptor Events in Moth Odorant and Pheromone Detection ..... 7
1.2.3 Receptor Events in Moth Odorant and Pheromone Detection ..... 8
1.3 Odorant-Binding Proteins ..... 10
1.3.1 Discovery of OBPs and Biological Properties ..... 10
1.3.2 Primary and Secondary Structural Characteristics ..... 14
1.3.3 Tertiary Structure Determination ..... 15
1.4 Binding Studies of PBPs ..... 18
1.4.1 The PBP Binding Cavity ..... 18
1.4.2 Ligand Binding Experiments ..... 18
1.5 *L. dispar* PBP1 and PBP2 ..... 23
1.5.1 *L. dispar* Pheromone Composition ..... 23
1.5.2 Effect of pH and Ion Concentration on Pheromone Binding ..... 23
1.5.3 Synergistic Binding Effects ..... 28
1.5.4 PBP1 and PBP2 Exhibit Subtle Ligand Discrimination ..... 29
1.6 PBPs and their Role in Olfaction ..... 30
1.6.1 The Objectives of this Work ..... 31

**Chapter 2: Materials and Methods ..... 32**

2.1 General Methods ..... 32
2.1.1 Purchased/Donated Chemicals ..... 32
2.1.2 Expression, Isolation, and Purification of Recombinant PBPs ..... 33
2.1.3 Analytical Methods ..... 36
2.2 Identification of Endogenous Ligands ..... 39
2.2.1 Extraction of Endogenous Compounds from the Aqueous Whole Antennal Homogenate ..... 39

2.2.2 Extraction of Endogenous Compounds from the Cuticle and from Whole Antennal Homogenates in Ethyl Acetate ..... 40
2.2.3 Separation of PBP1 and PBP2 from the Aqueous Whole Antennal Homogenate ..... 41
2.2.4 Identification of Endogenous Ligands ..... 42
2.2.5 Determination of PBP1 and PBP2 Concentrations ..... 44
2.2.6 Determination of Fatty Acid Vesicle Size Using Dynamic Light Scattering ..... 45
2.2.7 Isothermal Titration Calorimetry ..... 47
2.3 Structural Characterization of PBPs ..... 48
2.3.1 *In silico* Measurements and Comparisons ..... 48
2.3.2 Peptide Mapping and Disulfide Properties ..... 49
2.3.3 Circular Dichroism Experiments ..... 51
2.3.4 GC Binding Assay ..... 52
2.3.5 Crystallography ..... 54

**Chapter 3: Endogenous Ligands Bind *L. dispar* PBP1 and PBP2 *in vivo* ..... 56**

3.1 Fatty Acids And Cholesterol are Present in the Extract of the Aqueous Whole Antennal Homogenate ..... 56
3.1.1 Extraction of the Aqueous Whole Antennal Homogenate ..... 56
3.1.2 Cuticular and Whole Antennal Extractions into Ethyl Acetate ..... 64
3.1.3 Whole Antennal Perfusion Experiments ..... 67
3.2 PBP1 Binds C18 Fatty Acids Endogenously ..... 69
3.3 PBP2 Binds C16 and C18 Fatty Acids Endogenously ..... 73
3.4 PBP1 and PBP2 Concentrations in the Whole Antennal Tissue Homogenate ..... 75
3.5 Endogenous Fatty Acids Associate with PBPs as Vesicles ..... 79
3.6 Binding Constant Determination of Endogenous Ligands Using Isothermal Titration Calorimetry ..... 85
3.7 Discussion ..... 88
3.7.1 Compounds Present in the Sensillar Lymph ..... 88
3.7.2 Reanalysis of PBP Concentrations *in vivo* ..... 90
3.7.3 PBPs Selectively Associate with Fatty Acid Vesicles ..... 90
3.7.4 Possible Roles of Endogenous Ligands and Vesicles, and a New Role for PBPs ..... 93

**Chapter 4: Structural Characterization of *L. dispar* PBP1 and PBP2 ..... 97**

4.1 Peptide Mapping of *L. dispar* PBP1 and PBP2 ..... 97
4.1.1 Peptide Mapping by Partial Reduction and Cyanylation Reactions ..... 98
4.1.2 Peptide Mapping by CNBr Reactions ..... 103
4.2 Examining the C2-C5 Disulfide Bond ..... 107
4.2.1 Torsion Angle Comparisons with Known PBP Structures ..... 107
4.2.2 Distances Between Helix 3 and Helix 6 ..... 114
4.2.3 Sterics and Electronics of the C2-C5 Disulfide Bond ..... 118
4.3 Conformational Analysis of Cyanylated C2-C5 PBP1 ..... 118
4.3.1 MALDI-TOF MS ..... 118
4.3.2 PAGE Gel and Western Blot ..... 119
4.3.3 CD Spectra ..... 120

4.3.4 Ligand Binding Determinations by the GC Assay ..... 125
4.4 Crystallization Trials of PBP1 and PBP2 ..... 126
4.4.1 Crystallization Conditions ..... 126
4.4.2 Diffraction of PBP1 Crystals ..... 127
4.5 Discussion ..... 130
4.5.1 Disulfide Connectivities of *L. dispar* PBP1 and PBP2 ..... 130
4.5.2 Properties of the C2-C5 Disulfide Bond: Effect on Helix Orientation and Torsion Angle Comparisons with Known PBP Structures ..... 130
4.5.3 Biological Function of the C2-C5 Disulfide Bond ..... 140

**Appendices ..... 143**

Appendix 1 Determination of Sensillar Lymph Volume ..... 143
Appendix 2 LC Traces of PBP1 and PBP2 ..... 144
Appendix 3 Detailed CNBr Digest Assignments of all PBP1 and PBP2 Fragments (Excluding Fragments Containing a Disulfide Bond) ..... 145
Appendix 4 Confirmation of Disulfide Containing Fragments by Identification of Intramolecular Fragmentation Peptides for *L. dispar* PBPs ..... 149
Appendix 5 Disulfide Bond Torsion Angles for C1-C3 and C4-C6 ..... 151

**Literature Cited ..... 154**

# LIST OF FIGURES


Figure 1.1 Structures of pheromones mentioned throughout Chapter 1 ..... 3
Figure 1.2 The pheromone-sensing system of *L. dispar* adult males ..... 6
Figure 1.3 Cladogram of known OBPs ..... 12
Figure 1.4 Overall fold of two OBPs ..... 16
Figure 1.5 Effect of pH on pheromone binding to PBP1 and PBP2 at constant ionic strength ..... 26
Figure 1.6 The binding pocket of PBP1 with the possible hydrogen bonding network illustrated ..... 27
Figure 3.1 Compounds identified in the extract of the aqueous whole antennal homogenate of male *L. dispar* ..... 58
Figure 3.2 GC-MS traces of trimethylsilylated compounds found in the extract of the aqueous whole antennal homogenate ..... 60
Figure 3.3 GC-MS fatty acid calibration ..... 63
Figure 3.4 GC-MS cholesterol calibration ..... 64
Figure 3.5 Bar graph of the amounts of fatty acids and cholesterol found in cuticular and antennal extractions directly into ethyl acetate ..... 66
Figure 3.6 PBP1 and PBP2 separation from the aqueous whole antennal homogenate ..... 71
Figure 3.7 GC traces of the PBP1 extract ..... 72
Figure 3.8 GC trace of the PBP2 extract ..... 75
Figure 3.9 Determination of PBP concentrations *in vivo* by immunoblotting ..... 78
Figure 3.10 Determination of PBP concentrations *in vivo* by Coomassie staining ..... 79
Figure 3.11 DLS measurements of PBP1 associated fatty acids ..... 84
Figure 3.12 DLS measurements of PBP2 associated fatty acids ..... 85
Figure 3.13 ITC titration of linoleic acid into a PBP2 solution ..... 87
Figure 4.1 Mass spectra of *L. dispar* PBPs produced from LC-ESI-MS peptide separation and analysis ..... 105
Figure 4.2 C2-C5 disulfide bond torsion angle ranges of insect PBPs ..... 112
Figure 4.3 Histograms illustrating the disulfide bond torsion angle ranges of insect PBPs and the hemolymph protein TmolTHP12 ..... 113
Figure 4.4 Sequence alignments of the insect-binding proteins and the C2-C5 disulfide bond ..... 116

Figure 4.5 Orientation of helices 3 and 6 of various insect PBPs and TmolTHP12 ..... 117
Figure 4.6 PAGE gel and immunoblot of PBP1 and cyanylated PBP1 ..... 120
Figure 4.7 CD spectra of PBP1 and cyanylated PBP1 at various temperatures ..... 122
Figure 4.8 Thermal stability of PBP1 and cyanylated PBP1 ..... 124
Figure 4.9 Crystals of PBP1 co-crystallized with **(+)-1** ..... 129
Figure 4.10 Location of the disulfide bonds in LdisPBP1 and TmolTHP12 ..... 132
Figure 4.11 Comparison of $d_5$ and the angle measured between helix 3 and helix 6 of insect PBPs ..... 134
Figure 4.12 Orientation of the C2-C5 disulfide bond in insect PBPs ..... 140

# LIST OF SCHEMES

Scheme 1.1 *L. dispar* PBP chemical attachment of a fluorescent dansyl moiety.....21
Scheme 4.1 Reduction and cyanylation reaction mechanism of *L. dispar* PBPs .....100
Scheme 4.2 Fragments of *L. dispar* PBPs produced after cyanylation and subsequent cleavage under basic conditions .....102

# LIST OF TABLES


Table 3.1 Endogenous compounds found in the extract of the aqueous whole antennal homogenate ..... 58

Table 3.2 Amounts of compounds found in cuticular and antennal ethyl acetate extracts ..... 65

Table 3.3 Amounts of compounds found in perfusion experiments in DMSO ..... 68

Table 3.4 Endogenous compounds found bound to PBP1 ..... 73

Table 3.5 Endogenous compounds found bound to PBP2 ..... 74

Table 3.6 DLS results of fatty acid vesicles associated with PBP1 and PBP2 ..... 83

Table 3.7 Thermodynamic binding parameters for PBP1 and PBP2 ..... 86

Table 4.1 Singly reduced and cyanylated isomers of PBP1 and PBP2 produced via cyanylation cleavage reaction ..... 98

Table 4.2 CNBr chemical cleavage peptide fragments of partially reduced PBP1 and PBP2. ..... 103

Table 4.3 Disulfide bond properties and distances between helix 3 and 6 ..... 109

Table 4.4 Molar ellipticity values of modified and unmodified PBP1 ..... 123

Table 4.5 Estimated secondary structure content of various OBPs ..... 125

Table 4.6 Dissociation constant determination of **1** bound to either unmodified or modified PBP1 ..... 126

Table 4.7 PBP1 crystallization trials ..... 127

# LIST OF ABBREVIATIONS

| Abbreviation | Meaning |
|---|---|
| ABP | antennal-binding protein |
| ABSF | 4-(2-aminoethyl)benzenesulfonyl fluoride |
| ACN | acetonitrile |
| AL | antennal lobe |
| AMA | 1-aminoanthracene |
| amu | arbitrary mass units |
| ANS | 8-anilino-1-sulphonic acid |
| AP | alkaline phosphatase |
| Asp, D | aspartic acid |
| AU | arbitrary units |
| BSTFA | *bis*(trimethylsilyl)-trifluoroacetamide |
| C16 | fatty acid with 16 carbons |
| C18 | fatty acid with 18 carbons |
| cAMP | cyclic adenosine monophosphate |
| CD | circular dichroism |
| CDAP | 1-cyano-4-dimethylamino-pyridinium |
| cGMP | cyclic guanosine monophosphate |
| CHAPS | 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate |
| CMC | critical micellar concentration |
| CNS | central nervous system |
| Cys, C | cysteine |
| DAG | diacylglycerol |
| $ddH_2O$ | doubly distilled water |
| DLS | dynamic light scattering |
| DMSO | dimethylsulfoxide |
| DTT | dithiothreitol |

| Abbreviation | Meaning |
|---|---|
| EDTA | ethylenediaminetetracetetic acid |
| EI | electron impact |
| ESI | electrospray ionization |
| FA | fatty acid |
| FPLC | fast protein liquid chromatography |
| $\Delta G$ | change in Gibbs free energy |
| G | general program for the GC-MS |
| GC | gas chromatography |
| Gly, G | glycine |
| GOBP | general odorant-binding protein |
| GPCR | G-protein coupled receptor |
| GR | gustatory receptor |
| GRN | gustatory receptor neuron |
| h | hour |
| HEPES | N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid |
| His, H | histidine |
| HPLC | high performance liquid chromatography |
| $IP_3$ | (1,4,5)-triphosphate |
| IPTG | isopropyl-β-D-thiogalactopyranoside |
| ITC | isothermal titration calorimetry |
| ITC residue | isothiocyanate residue |
| $K_d$ | dissociation constant |
| LC | liquid chromatography |
| MALDI-TOF | matrix assisted laser desorption ionization time-of-flight |
| $M_{ave}$ | average molecular weight |
| min | minute |
| $M_{monoiso}$ | monoisotopic molecular weight |
| MS | mass spectrometry |
| MWCO | molecular weight cut-off |
| n | stoichiometry |

| Abbreviation | Meaning |
|---|---|
| nc | not calculated |
| nd | not determined |
| NMR | nuclear magnetic resonance |
| NPN | N-phenyl-1-napthylamine |
| OBP | odorant-binding protein |
| OR | olfactory receptor |
| ORN | olfactory receptor neuron |
| PAGE | polyacrylamide gel electrophoresis |
| PBP | pheromone-binding protein |
| PDB | protein data bank |
| PEG 4000 | polyethylene glycol with a molecular weight of 4000 |
| Phe, F | phenylalanine |
| $PIP_2$ | phosphatidylinostitol bisphosphate |
| PLC | phospholipase C |
| PMSF | phenylmethylsulfonyl fluoride |
| PVDF | polyvinylidene fluoride |
| R | universal gas constant |
| R. I. | retention index |
| $\Delta S$ | change in entropy |
| S | steroid program for the GC-MS |
| s | second |
| *s.* | sensilla |
| SDS | sodium dodecyl sulfate |
| SE | standard error |
| Ser, S | serine |
| SHT | steroid high temperature program for the GC-MS |
| SNMP | sensory neuron membrane protein |
| $\tau$ | torsion angle |
| TCA | trichloroacetic acid |
| TCEP | *tris*(carboxyethyl)phosphine |
| TFA | trifluoroacetic acid |

| Abbreviation | Meaning |
|---|---|
| Thr, T | threonine |
| TLC | thin layer chromatography |
| TMS | trimethylsilyl |
| Tris | *tris*(hydroxymethyl)aminomethane |
| UV | ultraviolet |

# CHAPTER 1: INTRODUCTION

## 1.1 Lepidopteran Pheromone Communication

Insects communicate *via* semiochemicals, chemical messages produced and sensed by insects (Nation 2002). Semiochemicals employed by insects are small, organic molecules. Two types of semiochemicals exist, allelochemicals and pheromones. Allelochemicals are chemical signals produced to allow communication between members of different species; pheromones are chemical signals produced to allow communication among members of the same species. Pheromones may elicit physiological (primer) or behavioral (releaser) responses (Borden 1993).

Pheromones play an important role in reproduction of many insect species. The first sex pheromone identified was from the order Lepidoptera, which includes insects such as moths and butterflies (Butenandt *et al.* 1959). Many species of moth use sex pheromones to attract mates. In moths, the females produce the pheromone and the males locate the "calling" females upwind. The sex attractant pheromone not only ensures male/female communication over long distances but also contributes to reproductive isolation of related species. Closely related species of moth often share one or more sex pheromone component(s), yet males will only fly towards their own species-specific blend. For example, the nun moth (*Lymantria monacha)* and the gypsy moth (*Lymantria dispar*) share the same sex pheromone attractant, (+)-disparlure ((+)-**1**) (Figure 1.1). However, 90% of the *L. monacha* pheromone blend consists of an antagonist, (-)-disparlure ((-)-**1**). Male *L. dispar* moths are able to detect both

enantiomers, yet male *L. monacha* moths cannot detect **(-)-1** (Hansen 1984). This is an important attribute since **(-)-1** acts to cancel the response of *L. dispar* males to **(+)-1**, helping to prevent cross-species mating. In addition to **(+)-1** and **(-)-1**, compounds **2**, **3**, **(+)-4**, **(-)-4**, **(+)-5**, and **(-)-5** have been identified as part of the *L. monacha* pheromone blend, while compound **6** has been identified as a second pheromone component in the *L. dispar* blend (Gries *et al.* 1996). Addition of **3** and **(+)-4** to **(+)-1** in a 1:10:10 ratio was 10 times more attractive to *L. monacha* males than **(+)-1** alone, while addition of **6** to **(+)-1** resulted in attraction of *L. dispar* but not *L. monacha* males (Gries *et al.* 1996). The disparlure precursor **3** is found in both species, however, addition of **3** to **(+)-1** or racemic disparlure results in a reduced attractiveness towards male *L. dispar* moths (Grant *et al.* 1996). These additional compounds enhance the specificity of the pheromone signal by acting either synergistically or antagonistically. Thus male moths are exquisitely selective to the pheromones released. Such specificity is a result of the diverse chemical structures and composition of pheromone blends. The olfactory system of male moths has specialized to distinguish components in this pheromone blend.

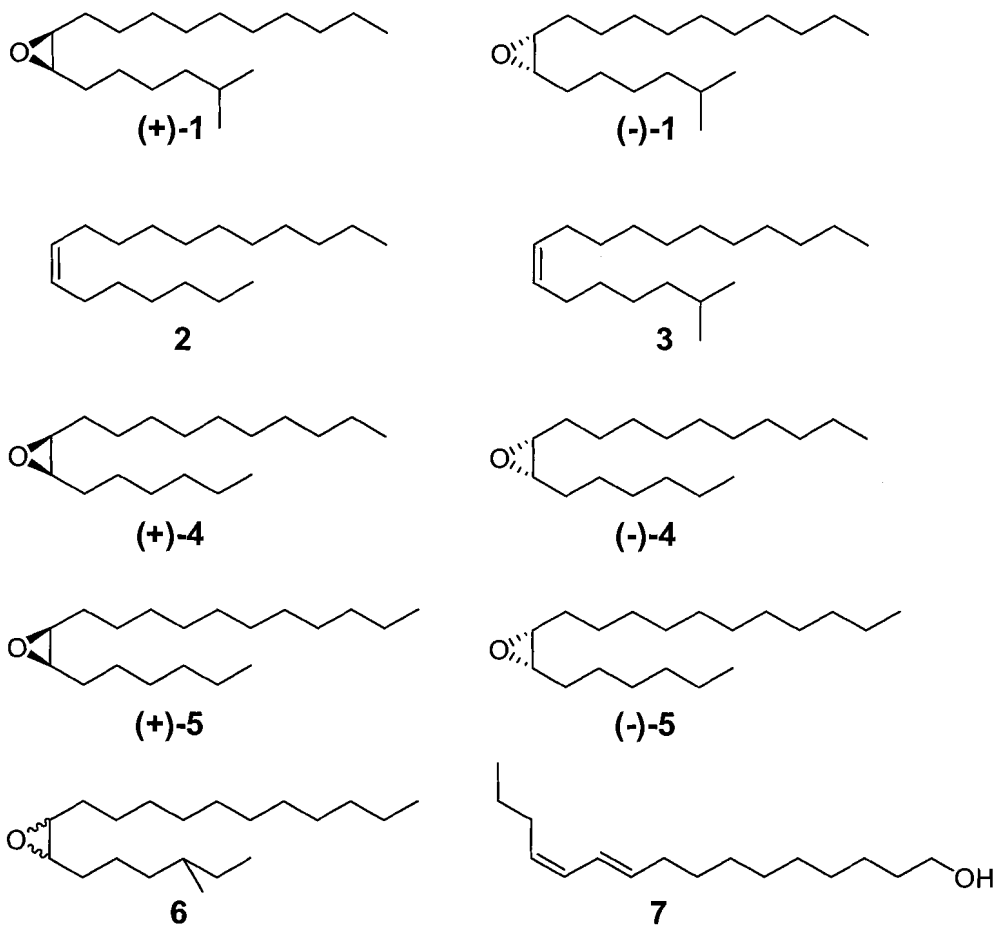



**Figure 1.1 Structures of pheromones mentioned throughout Chapter 1.**

Male moths are also extremely sensitive to the chemical signal released by females. For example, male moths are capable of detecting picogram quantities of pheromone many hundreds of meters downwind (Bierl *et al.* 1970, Kaissling and Priesner 1970). The lowest detectable dose of the sex pheromone **7** that results in wing-flapping motion has been measured for the silkworm moth, *Bombyx mori*, and was found to be 10,000 molecules per mL of air for a 1 second duration, or 1 molecule present on 25% of the total number of sensilla (see section 1.2.1 for definition of sensillum) (Schneider *et al.* 1968). The maximum detectable dose was found to be $3\times10^{11}$ molecules per mL of air for a 4 second duration, or approximately 3,000,000 molecules per sensillum (Kaissling

1977). Thus, to induce a nerve impulse requires a very low concentration of molecules, and very high doses are required to cause adaptation.

## 1.2 Detection of Pheromones and General Odorants

### 1.2.1 An Overview of Insect Olfaction and Gustation

Insects detect odorants and pheromones through hair-like structures called sensilla, which can be found on many parts of the body. Olfactory sensilla are mainly found on the head, specifically on the antennae (Figure 1.2 A, B) or the maxillary palps. Gustatory sensilla can be found on the legs, wings, mouthparts, and the ovipositor in females (Dahanukar 2005).

A sensillum is a hollow cuticular structure, covered by a waxy layer. Pores extend through the cuticle and protrude into the lumen of the hair (Figure 1.2 C). The lumen of the hair is filled with an aqueous solution known as the sensillar lymph. The pores are thought to allow passage of pheromones and odorants from the air to the lymph. Receptor neurons extend their dendrites into the lymph filled cavity of each hair, while their axons connect to the antennal lobe (AL) in the brain. Odorants and pheromones are recognized by olfactory receptor neurons (ORNs), while tastants are detected by gustatory receptor neurons (GRNs) in taste sensilla. The membranes of these dendrites house several olfactory receptor (OR) or gustatory receptor (GR) proteins which belong to the seven transmembrane domain superfamily of G-protein-coupled receptors (GPCRs). Interaction of the olfactory receptor with the odorant or pheromone results in the opening of ion channels leading to a voltage change across the cell membrane (Stengl *et al.* 1999). This change is relayed to the AL of the central nervous system (CNS)

(Hildebrand 1996). Here, the nerve signals are processed and passed on to the protocerebrum through projection neurons. Once the olfactory stimulus is received, the motor system is induced and the result is a nerve impulse such as wing flapping in the case of pheromone detection or feeding if a food source is located.

Each odorant or pheromone molecule stimulates one or more ORs (Hallem *et al.* 2004). ORNs expressing the same type of OR converge to one glomerulus in the AL (Vosshall *et al.* 2000). In fact, imaging studies have revealed that individual odorants give rise to activation of unique patterns of glomerular activation. This same pattern is retained once the glomeruli synapse with projection neurons, suggesting this odor map is relayed to higher brain centers (Wang *et al.* 2003).

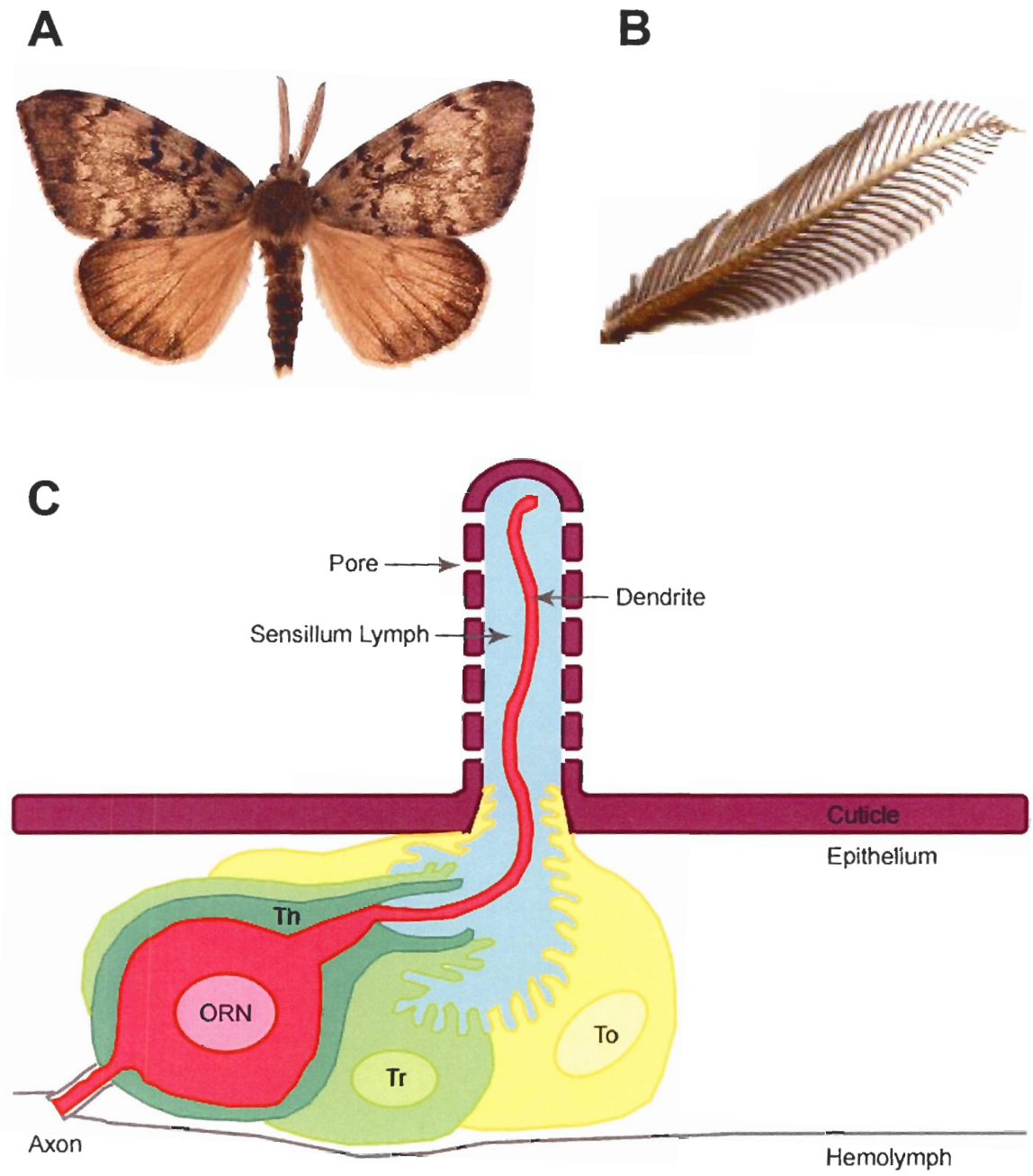



**Figure 1.2** **The pheromone-sensing system of *L. dispar* adult males. (A) Photograph of an *L. dispar* adult male. (B) One antenna from an *L. dispar* adult male. (C) Diagram of the olfactory *sensillum trichodeum*. ORN is the olfactory receptor neuron (shown in red); Th, Tr, and To are the auxillary cells thecogen (shown in dark green), trichogen (shown in light green), and tormogen (shown in yellow), respectively. The illustration is not to scale.**

### 1.2.2 Perireceptor Events in Moth Odorant and Pheromone Detection

To attract a mate, it is usually the female moth which releases the pheromone from glands located in her posterior abdomen segment. Males frequently have long, feathery antennae, allowing for effective interception of the mating signal. Olfactory sensilla are non-randomly distributed over the antennae in large numbers (Kaissling 1987). For example, a single antenna of the male Tobacco hornworm (*Manduca sexta*) contains about 100,000 sensilla, while the female has 100,000-150,000 sensilla (Lee and Strausfeld 1990). Although this may suggest the female to be better equipped at detecting pheromone attractants, the opposite is true. Sensilla are composed of two types: pheromone-sensitive *s. trichodea* (long hairs, between 30 to 600 μm in length), and general odorant-tuned *s. basiconicae* (medium-length hairs, between 10 and 80 μm) (Keil 1999). The male antennae of the tobacco hornworm contains 43,000 pheromone-specific sensilla and 55,000 general odorant sensilla, while the female antenna contains only sensilla of the general odorant type (Lee and Strausfeld 1990).

The sensillum is filled with an aqueous sensillum lymph which bathes the dendrites. This lymph is completely separate from the hemolymph. The sensillar lymph has been estimated to contain 200 mM $K^+$ and 40 mM $Na^+$ (Kaissling 2004). Such an ion composition produces a transepithelial potential of +40 mV (Kaissling 2004). The lymph also contains soluble proteins known as odorant-binding proteins (OBPs). The true function of these proteins is still unknown and is highly debated. Because of their location in such sensory organs and their ability to bind various odorants and pheromones, the function of these small, hydrophilic proteins has been thought by many in the field to be binding and transport of hydrophobic odorants across the aqueous layer

to the surface of the dendritic membrane. Others believe the protein may function as scavengers of excess pheromone. Once near the surface of the dendritic membrane, the OBP may interact with the receptor to either present or deliver the odorant or pheromone to the nerve cell. The mechanism of interaction with the OR is unknown. Degrading enzymes are also present in the lymph and may deactivate the odorant or pheromone which have already stimulated the nerve cell and prevent the cell from over-stimulation (Vogt *et al.* 1985). The components and mechanisms involved in deactivation are unknown.

### 1.2.3 Receptor Events in Moth Odorant and Pheromone Detection

Many *Or* and *Gr* genes have been isolated from the fruit fly *Drosophila melanogaster* (Clyne *et al.* 1999, Vosshall *et al.* 1999, Clyne *et al.* 2000, Stortkuhl and Kettler 2001, Wetzel *et al.* 2001, Robertson *et al.* 2003). Candidate ORs from the malaria vector mosquito *Anopheles gambiae* (Fox *et al.* 2001, Fox *et al.* 2002, Hill *et al.* 2002), the Tobacco Budworm moth *Heliothis virescens* (Krieger *et al.* 2002, Gohl and Krieger 2006), and the silkworm moth *Bombyx mori* (Sakurai *et al.* 2004) were subsequently identified.

Once an odorant or pheromone is recognized by the receptor, a cascade of events involving production of a second messenger is initiated. OR stimulation results in reorganization of G proteins that leads to activation of a key enzyme. In invertebrates this enzyme is phospholipase C (PLC) which hydrolyzes membrane bound phosphatidylinositol bisphosphate ($PIP_2$) to produce the second messengers inositol (1,4,5)-triphosphate ($IP_3$) and diacylglycerol (DAG) (Krieger and Breer 1999). Pheromone stimulation has resulted in production of $IP_3$ in the antennae of the silkworm

moth (Breer *et al.* 1990). $IP_3$ then can activate cation channels in the plasma membrane, producing an electrical potential (change in voltage). The cyclic adenosine monophosphate (cAMP) or cyclic guanosine monophosphate (cGMP) pathway may also play a role in signal transduction. The secondary messenger cGMP has been found in the antennae of the silkworm moth upon pheromone stimulation (Ziegelberger *et al.* 1990). Also, cyclic nucleotide activated channels have been found in insect antennae (Baumann *et al.* 1994, Krieger *et al.* 1999).

It is still unclear which ion channels are involved in the olfactory signalling process. Odorant or pheromone stimulation of cultured receptor cells of *M. sexta* resulted in the opening of $IP_3$-dependent $Ca^{2+}$ channels. The entering third messenger $Ca^{2+}$ ions activate an unknown cation channel which leads to an influx of $K^+$ ions resulting in an electrochemical potential across the cell (Stengl *et al.* 1992, Stengl 1994). $IP_3$-dependent $Ca^{2+}$ channels were also identified in the dendrites of OR cells of the silkmoths *B. mori*, *Anthereae pernyi*, and *Anthereae polyphemus* by immunocytochemical experiments (Laue and Steinbrecht 1997). The calcium-binding protein calmodulin was found to be expressed with high labelling densities in the dendrites of olfactory receptor cells of the silkmoths, thus supporting the role of calcium as a third messenger (Laue and Steinbrecht 1997).

A novel antennal membrane-bound protein named the sensory neuron membrane protein (SNMP) was identified first in *A. polyphemus* (Rogers *et al.* 1997), and later in *B. mori*, *H. virescens*, and *M. sexta* (Rogers *et al.* 2001). Unlike the putative ORs previously detected, this novel antennal protein does not belong to the GPCR family of proteins. SNMP is a 67 kDa protein expressed on the dentritic membranes of olfactory

neurons in adult antennae. The protein is homologous to the two-transmembrane domain receptor proteins which includes CD36, a family of proteins which are involved in protein-protein interactions. Immunological studies have determined SNMP expression in both receptor cells of female and male *A. polyphemus* moths tuned to pheromone and general odorants, although most of the expression is in pheromone-specific cells, thus implicating a role in the olfactory process (Rogers *et al.* 2001). Furthermore, SNMP expression parallels the expression of OBPs which begins during adult development and increases significantly towards the end of adult development and into the adult stage (Rogers *et al.* 1997). Possible functions of the SNMP may be to mediate cell-cell adhesion or cytoskeletal-membrane interactions. Alternatively, the SNMP may interact with odorant- or pheromone-bound OBPs by stabilizing the complex at the membrane surface, thus assisting delivery of odors or pheromones to the OR (Rogers *et al.* 1997).

## 1.3 Odorant-Binding Proteins

### 1.3.1 Discovery of OBPs and Biological Properties

OBPs constitute a large superfamily of insect-binding proteins. The first OBP discovered was a pheromone-binding protein (PBP) isolated from the antennae of *A. polyphemus* (Vogt and Riddiford 1981). When a radiolabeled analog of the major pheromone component of *A. polyphemus* was incubated with an antennal homogenate, and the proteins from the homogenate were separated by sodium dodecyl sulfonate (SDS) polyacrylamide gel electrophoresis (PAGE), a radioactive protein band was observed with an approximate molecular weight of 15 kDa. This protein was unique to the antennae of male moths (it was not seen in other body tissues or in the hemolymph) and total PBP concentrations were initially estimated to be about 10 mM in male *A.*

*polyphemus* antennae (Klein 1987). Later, the PBP concentration from the gypsy moth, *L. dispar*, was estimated at 13.4 mM (Vogt *et al.* 1989). However, these protein concentrations may be inaccurate as they were approximated from non-calibrated gels. Thus the true PBP concentrations *in vivo* have never been determined.

To date, OBPs have been identified in a number of insect orders including Lepidoptera, Coleoptera, Diptera, Hymenoptera, and Hemiptera (Figure 1.3) (Vogt *et al.* 1999). Ten years after the discovery of PBPs, general OBPs (GOBPs), which bind a variety of general odorant molecules associated with food and the environment, were identified in the antennae of both sexes of Lepidoptera (Vogt *et al.* 1991, Vogt *et al.* 1991). They were separated into two classes based on their relative electrophoretic mobility using native PAGE: GOBP1 (migrates slower) and GOBP2 (migrates faster) (Vogt *et al.* 1991). PBPs have about 30% sequence identity to GOBPs, while the GOBPs have about 50% sequence identity to each other.

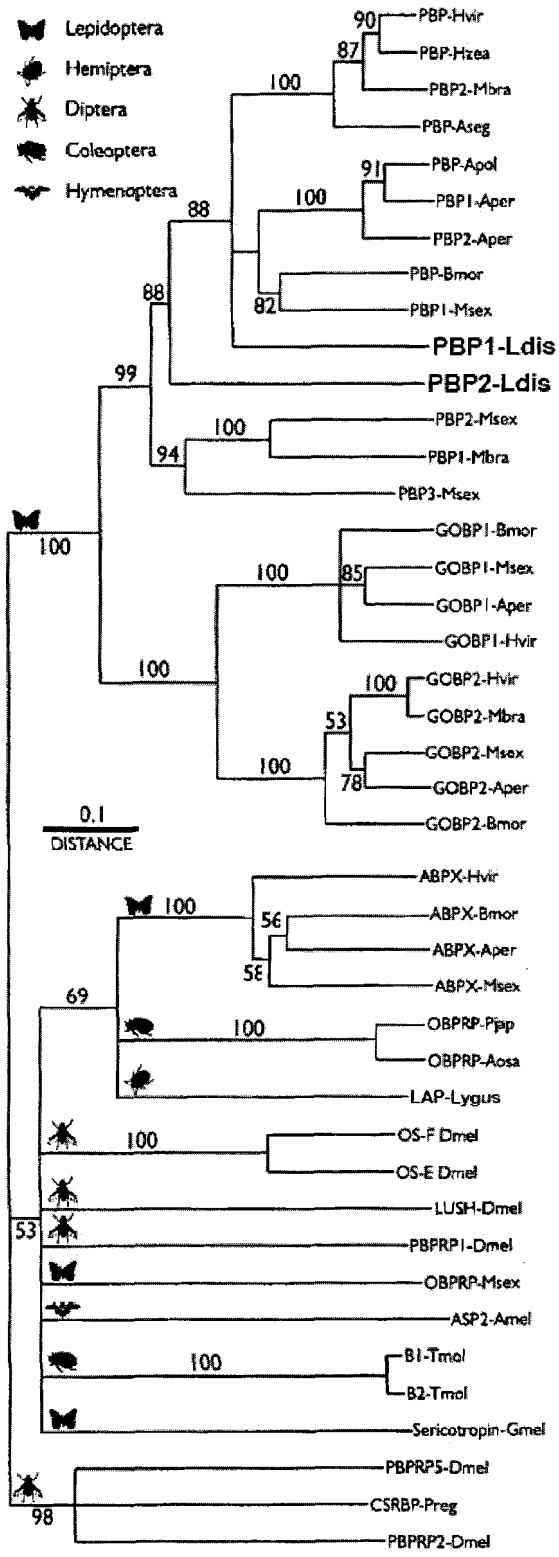



**Figure 1.3** **Cladogram of known OBPs. *L. dispar* PBP1 and PBP2 are shown in bold. The branch lengths are proportional and numbers indicate scale distance. Adapted from (Vogt *et al.* 1999), by permission of author and publisher © Oxford University Press, 1999.**

Finally, antennal-binding proteins (ABPs) were first isolated in *D. melanogaster* (McKenna *et al.* 1994, Pikielny *et al.* 1994) and were then discovered in the antennae of male *B. mori* (Krieger *et al.* 1996). Primary structure characterization of these proteins reveal similarities to PBPs and GOBPs, including the same hydrophobic and hydrophilic domains and six conserved cysteine residues. However, ABPs seem to have a low degree of sequence identity (11-26%) when compared with other PBPs or GOBPs from *B. mori* or *D. melanogaster* (Krieger *et al.* 1996).

The above proteins belong to the insect OBP superfamily of proteins. OBPs are characterized as small, hydrophilic proteins found in the sensory sensilla of insects and having the capability of binding low molecular weight compounds. They have a large sequence identity range spanning 6 to 99% (Tegoni *et al.* 2004).

OBPs are expressed during specific life stages, in specific tissues and specific individuals, depending on their function. For example, in the fire ant *Solenopsis invicta*, workers express an OBP that appears to mediate a queen pheromone activity (Krieger and Ross 2002). In *D. melanogaster*, only the T1 and T2 populations of sensilla express OBP76a (LUSH) (Kim *et al.* 1998). Moth PBPs tend to be expressed most highly in adult male antennae. Indeed, expression of PBPs in Lepidoptera begins 35-40 hours before emergence of the adult moth from the pupa (Vogt *et al.* 1989). OBPs are expressed in epithelial cells and secreted into extracellular spaces. In the case of Lepidopteran PBPs, supporting cells at the base of the sensillum (Figure 1.2 C) secrete the protein into the sensillar lymph.

Expression of multiple OBPs is common in many insects. The presence of two PBPs encoded by two different genes was first detected in the male antennae of *L. dispar*

(Vogt *et al.* 1989). It has been shown that the two PBPs, named PBP1 and PBP2 (sometimes referred to in this work as LdisPBP1 and LdisPBP2, respectively), preferentially bind different enantiomers of the pheromone (see section 1.5.1), suggesting multiple PBPs may be essential for proper pheromonal communication. More recently, studies of male *M. sexta* antennae have revealed expression of two ABPs, two GOBPs, and three PBPs in this species to localize in specific sensilla and annular regions with some overlap, while females lacked this distinct zonal expression (Nardi *et al.* 2003). The presence of multiple OBPs suggest they may be involved in specificity of the olfactory process.

### 1.3.2 Primary and Secondary Structural Characteristics

Lepidopteran OBPs are small, hydrophilic, acidic (pI ~ 5) proteins with an approximate molecular weight of 13-17 kDa (110-150 amino acids in length). OBPs share similar sequence identity (50-60%) within the same Lepidopteran species, but little similarity with OBPs from *D. melanogaster* and other non-lepidopteran species. Circular dichroism (CD) experiments reveal a tightly folded protein with mainly α-helical structure. The disulfide bridge pairings of OBPs were first elucidated by peptide mapping for the *B. mori* PBP (BmorPBP) (Leal *et al.* 1999, Scaloni *et al.* 1999), and later for the honey bee *Apis mellifera* L. PBP1 (Briand *et al.* 2001) and *L. dispar* PBP1 and PBP2 (see section 4.1) (Honson and Plettner 2006). A combination of enzymatic and chemical cleavage techniques were employed to determine the disulfide linkages of these tightly folded PBPs. Most OBPs have six conserved cysteine residues which form three interlocking disulfide bridges. There are a few OBPs which have extra cysteines located at the N-terminus such as *A. pernyi* PBPs Aper-1 and Aper-2 (Du and Prestwich 1995).

Other non-antennal OBPs possess only four of the six conserved cysteine residues (they lack the second and fifth cysteine residues). An example of such a four-cysteine protein is the hemolymph protein THP12 from the mealworm *Tenebrio molitor* (TmolTHP12) (Rothemund *et al.* 1997, Rothemund *et al.* 1999).

### 1.3.3 Tertiary Structure Determination

The first available three-dimensional structure of a member of the insect OBP family was the nuclear magnetic resonance (NMR) structure of the four-cysteine hemolymph protein TmolTHP12 (Figure 1.4) (Rothemund *et al.* 1997, Rothemund *et al.* 1999). The protein was found to consist of six α-helices, separated by loops, in a glove-like arrangement. Helices 1 and 3, as well as helices 5 and 6, are held close together by disulfide bonds Cys14-Cys45 and Cys85-Cys102, respectively (Rothemund *et al.* 1999). The first solved tertiary structure of a six-cysteine insect OBP was the crystal structure of the BmorPBP in complex with the species-specific pheromone **7** (Figure 1.4) (Sandler *et al.* 2000). NMR structures of liganded and unliganded forms of BmorPBP were solved later (Horst *et al.* 2001, Lee *et al.* 2002). Since then, four more OBPs have had their structures solved by either X-ray diffraction or NMR studies. These include the LUSH protein (OBP76a) in *D. melanogaster* (Kruse *et al.* 2003), PBP3 from *A. polyphemus* (ApolPBP) (Mohanty *et al.* 2003, Mohanty *et al.* 2004, Zubkov *et al.* 2005), a PBP (formally classified as an antennal-specific protein) from *A. mellifera* L. called ASP1 (AmelASP1) (Lartigue *et al.* 2003, Lartigue *et al.* 2004), and a PBP from the cockroach *Leucophaea maderae* (LmaPBP) (Lartigue *et al.* 2003, Lartigue *et al.* 2003). The tertiary structure of all the 6-cysteine OBPs reveals six compact α-helices connected by extended stretches with three interlocking disulfide bridges, confirming peptide mapping results.

For BmorPBP, the disulfide linkages are Cys19-Cys54 which connects α-helices 1 and 3, Cys50-Cys108 which connects helices 3 and 6, and Cys97-Cys117 which links helices 5 and 6 (Sandler *et al.* 2000).

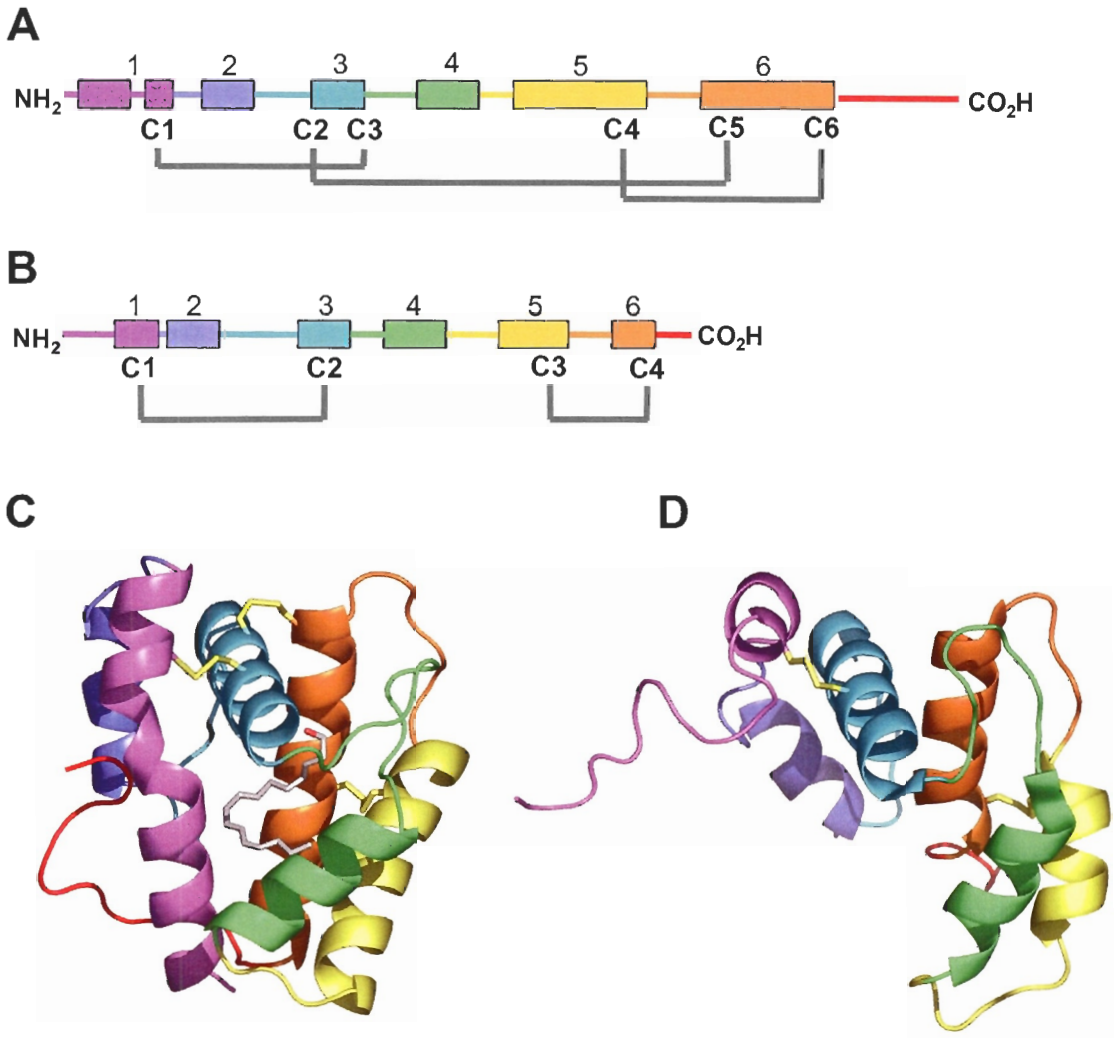



**Figure 1.4** **Overall fold of two OBPs. Cartoon representation of the disulfide pairings of BmorPBP (A) and TmolTHP12 (B). The solved three-dimensional structures of BmorPBP by X-ray diffraction (bound pheromone 7 is shown in grey) (Sandler *et al.* 2000) (C) and TmolTHP12 by NMR (Rothemund *et al.* 1999) (D) were prepared using PyMOL (Delano, W.L. The PyMOL Molecular Graphics System. (2004) Delano Scientific, San Carlos, CA, USA. http://www.pymol.org).**

The solved structures of the BmorPBP have exhibited large structural changes when the pH of the buffer is varied. The NMR structure of the unliganded protein at

neutral pH reveals helix 1 to be α-helical while the C-terminal tail is disordered, existing in varying positions on the outer surface of the protein (Lee *et al.* 2002). At acidic pH, the N-terminus becomes disordered, while residues 131-142 of the C-terminus adopt an α-helical conformation to create a seventh helix. This new helix is found inserted into the binding cavity normally occupied by the silkworm moth pheromone **7** (Horst *et al.* 2001). It is unclear how the ligand, which is entirely surrounded when bound in most OBPs studied, enters or exits the binding cavity. One possible entry/exit point is a flexible loop region (approximately residues 55-75) within the N-terminal segment. Vast conformational changes of the C-terminus have been proposed to be the mechanism of ligand expulsion near the dendritic membrane, thereby releasing the pheromone to the receptor (Horst *et al.* 2001).

However, the impressive difference in the C-terminal region of BmorPBP is not a structural feature reflected among the other solved OBPs. For example, the X-ray structure of the *D. melanogaster* OBP LUSH, was insensitive to pH in the range 4.6 to 6.5 (Kruse *et al.* 2003). LmaPBP has a stunted C-terminus (Lartigue *et al.* 2003), while the C-terminal region of AmelASP1 and *D. melanogaster* OBP forms one of the walls of the binding cavity (Kruse *et al.* 2003, Lartigue *et al.* 2004). Recently, the NMR structure of unliganded ApolPBP at acidic pH revealed an unstructured C-terminal α-helix which does not insert itself into the binding cavity (Zubkov *et al.* 2005). Thus the proposed ligand expulsion mechanism by C-terminal helix insertion cannot be universal to OBPs.

## 1.4 Binding Studies of PBPs

### 1.4.1 The PBP Binding Cavity

The binding pocket of PBPs is heart-shaped and is formed by four anti-parallel α-helices comprised of helices 1, 4, 5, and 6. Indeed, many lepidopteran pheromones are extended molecules with a bend mid-length. The ligand is housed within this relatively large, hydrophobic cavity (Figure 1.4 C). For example, **7** forms non-covalent interactions with the aliphatic side chains of the BmorPBP binding cavity wall. These interactions include the two conserved phenylalanine residues (Phe12 and Phe118) which stack on both sides of the ligand π system, while the hydroxyl group is involved in a hydrogen bond with a serine (Ser56) near the cavity entrance (Klusak *et al.* 2003). The opening to the binding pocket is narrow, and it has been proposed that movement of a loop (residues 60-69) may widen the opening, allowing the pheromone to enter and exit the pocket (Sandler *et al.* 2000). These movements, required for access to an enclosed binding site, do not appear to be necessary for all OBPs: the OBPs of *L. maderae*, *A. mellifera* L., and *D. melanogaster* have larger, more solvent accessible binding sites than BmorPBP.

### 1.4.2 Ligand Binding Experiments

A number of binding assays have been utilized to assess the binding of various pheromones and ligands with PBPs. The first binding assay developed was a native PAGE experiment where OBP is incubated with different radiolabeled ligands. The sample is then loaded onto a native PAGE gel and electrophoresed. The gel is then either quickly stained and destained (Vogt *et al.* 1989) or blotted onto a membrane (Bohbot *et al.* 1998, Plettner *et al.* 2000). The PBP band is then excised and placed in a scintillation counter to measure radioactivity, or the blot is scanned with a thin layer chromatography

(TLC) plate scanner. This screen allows for only qualitative detection of pheromone binding.

Dissociation constants for pheromone or ligand binding could be determined by a vial adsorption binding assay developed by Du and Prestwich (Du and Prestwich 1995). Various concentrations of radiolabeled pheromone are incubated in a plastic vial coated with either 1-decanol or 1-dodecanol (to reduce nonspecific adsorption to the plastic surface). PBP-ligand complexes adsorb to the vial surface while unbound pheromone remains in solution. A sample of the unbound pheromone is then counted for radioactivity. In subsequent experiments, a filtration step employing size exclusion gel matrix was added before measuring radioactivity. In this manner, PBP-ligand complexes were collected as the filtrate while unbound ligand was trapped in the column, allowing for direct measurement of bound ligand (Plettner *et al.* 2000). The problem with these assays was the adsorption of hydrophobic pheromone or ligand onto the vial, even after coating vials with a long chain alcohol in an effort to prevent adsorption of the hydrophobic pheromone. Also, there is a limited number of radiolabeled ligands available for testing.

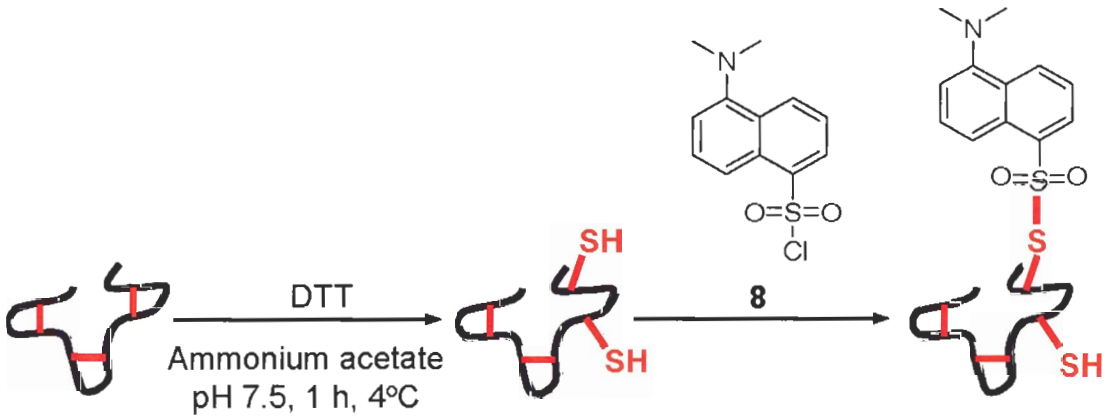
To test the binding strengths of non-radiolabeled ligands, I developed a competitive radioassay during my undergraduate research semester where non-radiolabeled ligands are incubated with PBP that had been pre-incubated with a radiolabeled reporter molecule that is known to bind the protein. The solution is then passed through a filter and counted for radioactivity as explained above. Addition of the cold ligands was expected to result in displacement of the reporter molecule. Thus, the less radioactivity detected bound to protein, the stronger the binding of the cold ligand

(Honson *et al.* 2003). This is an indirect way to measure dissociation constants that are not affected by synergistic or antagonistic effects between ligands (Honson *et al.* 2003). To my surprise, I found that the radiolabeled reporter ligand bound more strongly to the protein in the presence of some cold ligands than in the absence. This was one indication that there are synergistic interactions between ligands with respect to PBP binding (Honson *et al.* 2003). Earlier, an antagonistic effect was observed between the enantiomers of **1**: racemic **1** binds to both *L. dispar* PBPs more weakly than does either (+)-**1** or (-)-**1** alone (Plettner *et al.* 2000).

To directly measure binding of non-radiolabeled ligands, fluorescence binding assays were developed. The first was a fluorescence displacement assay which involved binding of a fluorescent reporter molecule such as 1-aminoanthracene (AMA) with *M. brassicae* or *A. polyphemus* PBPs (Campanacci *et al.* 2001). PBPs from these species are capable of binding AMA. When the protein solution is irradiated at a specific wavelength, bound AMA fluoresces, resulting in a blue shift and an increase in fluorescence intensity due to the hydrophobic environment of the probe. When the PBP-AMA complex is titrated with a competitor, AMA is displaced into the aqueous solution and the fluorescent probe is quenched, resulting in a red shift in the emission maximum and a decrease in fluorescence emitted (Campanacci *et al.* 2001). However, AMA turned out not to be a general probe for all PBPs since only some are capable of AMA binding (Campanacci *et al.* 2001, Inkster *et al.* 2005). Since this study, other probes have been utilized in fluorescence displacement assays. These include 8-anilino-1-sulphonic acid (ANS) to study ligand binding with *L. maderae* PBP (Riviere *et al.* 2003), and N-phenyl-1-naphthylamine (1-NPN) to study ligand binding with the *D. melanogaster* OBP LUSH

(Zhou *et al.* 2004), the migratory locust *Locusta migratoria* OBP1 (Ban *et al.* 2003), the paper wasp *Polistes dominulus* OBP-1 (Calvello *et al.* 2003), and *L. dispar* PBP1 (Inkster *et al.* 2005). The two major problems with this assay include: 1) indirect measurements of dissociation constants, 2) the probe may or may not have an effect on competitor ligand binding.

In order to avoid the above problems I have developed a dansyl fluorescence assay where *L. dispar* PBPs are covalently modified with **8** by selectively reducing one of the three disulfide bonds and attaching the dansyl moiety as a thiosulfonate (Scheme 1.1) (Honson *et al.* 2003). Upon ligand binding, the fluorescence emission of the protein increases. However, the dansylated protein is short-lived, probably because of re-formation of the disulfide bond or hydrolysis of the dansyl moiety (Honson *et al.* 2003).




**Scheme 1.1** ***L. dispar* PBP chemical attachment of a fluorescent dansyl moiety. The above scheme is described in reference (Honson *et al.* 2003).**

A number of non-invasive ligand binding assays have been developed. Endogenous tryptophan fluorescence has been utilized to explore ligand binding of

ApolPBP (Bette *et al.* 2002) and *L. dispar* PBP1 and PBP2 (Honson *et al.* 2003). Some ligands quench the tryptophan fluorescence, which results in a decrease in fluorescence as the protein is titrated with ligand. The main problem with this assay is the relatively small change in fluorescence, which results in great errors when determining dissociation constants. Isothermal titration calorimetry (ITC) is another direct way of measuring the binding affinity of a ligand for a protein. In this method, the change in enthalpy ($\Delta H$) is recorded as ligand is titrated into the protein solution. Thermodynamic parameters such as the Gibbs free energy change ($\Delta G$) and the change in entropy ($\Delta S$), as well as the stoichiometry (n) and the dissociation constant ($K_d$), can all be determined by measuring the heat released during binding in one ITC run (Wiseman *et al.* 1989). A mammalian lipocalin, the mouse urinary protein-I (MUP-I), has been studied by ITC (Sharrow *et al.* 2003). The thermodynamics of the mouse pheromone 2-*sec*-butyl-4,5-dihydrothiazole and analogs binding to MUP-I were characterized. Binding of the pheromone with MUP-I (dissociation constants were measured) resulted in a favorable enthalpy change, but an unfavorable entropy change because of binding interactions with the pheromone alkyl chain. Finally, a gas chromatography (GC) based assay has been developed (Danty *et al.* 1999) (Y. Gong and E. Plettner, unpublished). Similar to the radioassay described above, protein is incubated with ligand and bound protein is separated from unbound ligand by gel filtration. Bound ligand is then removed from the protein by extraction into organic solvent and the concentration of the ligand is determined by GC. This assay is currently being utilized in determining ligand binding kinetics and dissociation constants for *L. dispar* PBPs.

## 1.5 *L. dispar* PBP1 and PBP2

### 1.5.1 *L. dispar* Pheromone Composition

The gypsy moth, *L. dispar*, is a forest pest in North America, Europe and Asia. Male moths are extremely sensitive to the chemical signals released by females. As mentioned in section 1.1, the female sex pheromone is 7*R*,8*S*-*cis*-2-methyl-epoxyoctadecane, commonly known as (+)-disparlure ((+)-**1**) (Bierl *et al.* 1970). Synthetic (+)-**1** is active in the laboratory and field tests in amounts as little as 2 pg (Bierl *et al.* 1970). Male moths are also exquisitely selective to the pheromone components. For example, racemic **1** is not attractive to the moths, even at high doses. The pheromone produced by female *L. dispar* moths has been estimated to contain almost 100% (+)-**1**, with trace amounts of the enantiomer, (-)-**1** (Hansen 1984) and **6** (Gries *et al.* 1996). Field testing experiments revealed synthetic (+)-**1** to be an extremely potent attractant, while traps baited with (-)-**1** as well as with the alkene precursor **3** were slightly repellent (Miller *et al.* 1977). The PBPs of *L. dispar* reflect this enantiomeric preference as was proven using the radioassay with gel filtration as previously described (Plettner *et al.* 2000). It was found that PBP2 binds (+)-**1** more strongly, while PBP1 has a greater affinity for the (-) enantiomer. The racemic mix seemed to bind less strongly than either of the preferred enantiomers (Plettner *et al.* 2000).

### 1.5.2 Effect of pH and Ion Concentration on Pheromone Binding

I have been involved in publishing three papers describing binding assays performed on PBP1 and PBP2. The first paper addressed the much debated question of whether pH affects pheromone binding (Kowcun *et al.* 2001). *In vitro*, pheromone binding to *L. dispar* PBPs decreases when the pH is lowered while keeping the ionic

strength constant. However, in the sensillum, pH and ionic strength likely change concurrently. We estimated the concurrent pH and salt concentrations and measured PBP/pheromone binding at several combinations of pH and ionic strength. We have found that an increase in salt concentration cancels this pH effect, and there is no net change in binding of PBP1 or PBP2 with its preferred enantiomer (Kowcun *et al.* 2001).

We also detected critical residues involved in binding by obtaining $pK_a$ values when measuring pheromone binding at various pHs. By following the pH dependence of pheromone binding at constant ionic strength, we can detect ionizations in the PBP-pheromone complex that are relevant to pheromone binding. Four pH profiles were constructed with PBP1 and PBP2 using radiolabeled **(+)-1** and **(-)-1** (Figure 1.5). The patterns observed for PBP1 and PBP2 with **(+)-1** or **(-)-1** differed significantly. In all cases, the data fit best to a set of 3 $pK_a$ values suggesting that 3 amino acid residues which titrate in that pH range are involved in pheromone binding. With increasing pH, as one traverses the first two $pK_a$s, an increase in pheromone binding is seen with **(+)-1**. However, with **(-)-1**, a decrease in ligand binding is seen when traversing the second $pK_a$ (Figure 1.5). The first two ionizations in each of the four profiles are in the range consistent with histidine residues. PBPs have five highly conserved histidine residues. Titration of the acidic histidine ($pK_a$ approximately 5.8 and 6.3 for **(+)-1** and **(-)-1**, respectively) leads to an increase binding of both enantiomers to both PBPs. The subtle but highly reproducible dependence of this acidic $pK_a$ on the ligand suggests that this $pK_a$ is dependent on the protein/ligand interaction. Titration of the second histidine ($pK_a$ approximately 7.8 for **(+)-1** and **(-)-1**) leads to increase binding with **(+)-1**, but a decrease or no change with **(-)-1**, suggesting that the second histidine residue is important in ligand

recognition, perhaps *via* a hydrogen bonding network with the epoxide oxygen moiety of **1**. The third, most basic residue observed has a $pK_a$ of about 9, suggesting titration of either a lysine or tyrosine (Kowcun *et al.* 2001).

A possible interpretation of these results is shown in Figure 1.6. The homology models of PBP1 and PBP2 using the crystal structure of the silkworm moth (PDB accession code 1DQE) as a template with **(-)-1** or **(+)-1** superimposed indicates a direct interaction with the epoxide and residue 9. When PBP1 binds its preferred enantiomer, **(-)-1**, Thr9 acts as a hydrogen bond doner and the epoxide of **(-)-1** is the hydrogen bond acceptor, allowing extension of the hydrogen bonding network and increasing binding affinity. Titration of His124 results in termination of this extended hydrogen bonding network and a large drop in binding affinity is observed. When PBP1 binds **(+)-1**, the epoxide of **(+)-1** can no longer act as a hydrogen bond acceptor, thus there is not such a significant change in binding affinity when His124 is titrated. Residue 9 of PBP2 is an alanine, so there is no such direct hydrogen bond with either enantiomer of **1** and little change is observed in the pH profile as the more basic histidine is titrated.

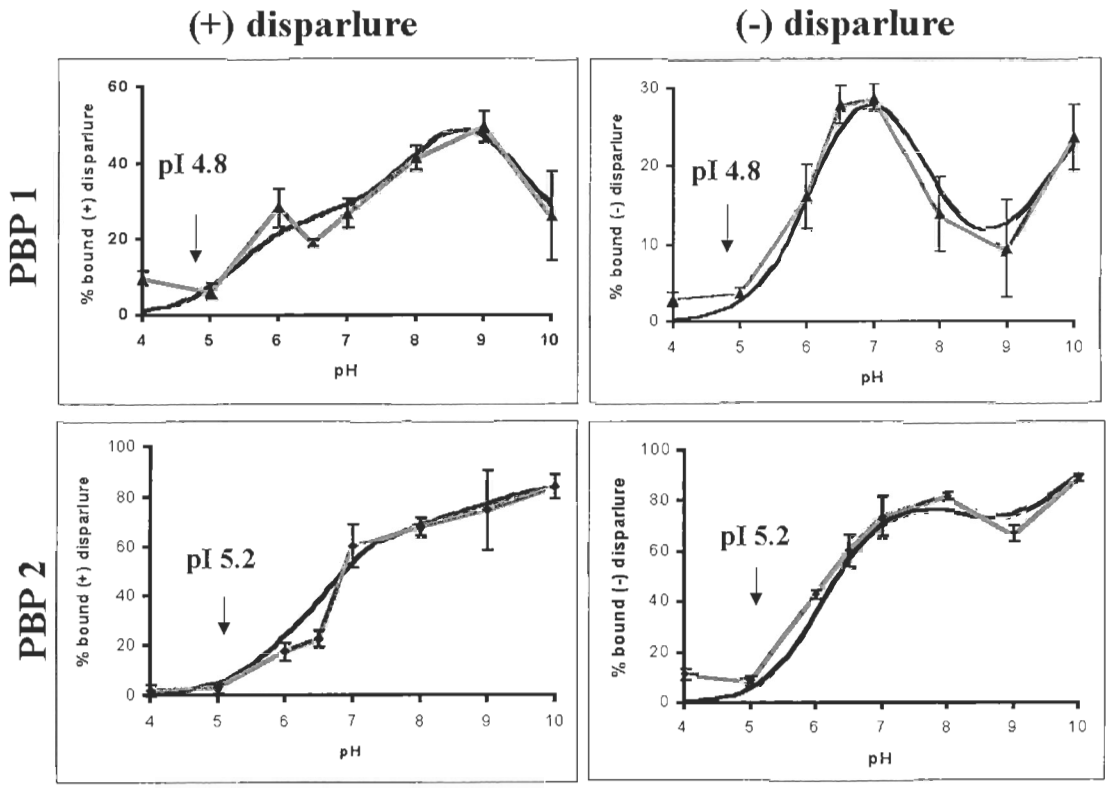



**Figure 1.5** **Effect of pH on pheromone binding to PBP1 and PBP2 at constant ionic strength. The amount of radiolabeled (+)-1 or (-)-1 bound to either PBP1 or PBP2 was measured. Each point is the average of three replicates with bars indicating standard error. The dark curves represent calculated pH profiles, obtained using the three $pK_a$ values detected in each experiment. This figure is from reference (Kowcun *et al.* 2001), by permission of the publisher ©American Society for Biochemistry and Molecular Biology.**

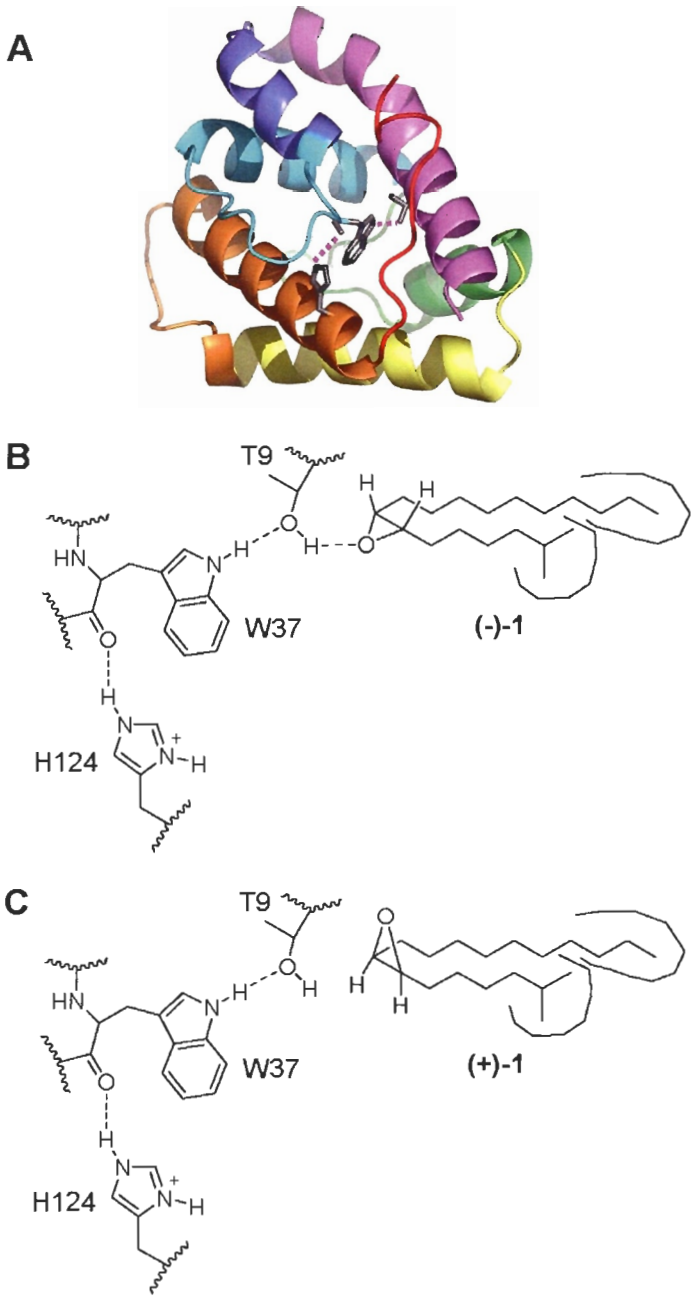



**Figure 1.6** The binding pocket of PBP1 with the possible hydrogen bonding network illustrated. (A) Illustrated is the homology model of PBP1 and the location of residues T9, W37, and H124 (with hydrogen bonds shown in pink). This image was prepared using PyMOL (Delano, W.L. The PyMOL Molecular Graphics System. (2004) Delano Scientific, San Carlos, CA, USA. http://www.pymol.org). (B) When PBP1 binds its preferred enantiomer, (-)-1, T9 acts as a hydrogen bond doner and the epoxide of (-)-1 is the hydrogen bond acceptor, allowing extension of the hydrogen bonding network and increasing binding affinity. (C) When PBP1 binds (+)-1, the epoxide of (+)-1 can no longer act as a hydrogen bond acceptor, resulting in decreased binding affinity.

### 1.5.3 Synergistic Binding Effects

The second paper describes the competition assay, the dansyl fluorescence assay, and the endogenous tryptophan fluorescence assay (Honson *et al.* 2003). In the competition assay, the PBP1 or PBP2 preferred enantiomer of **1** was used as the radiolabeled reporter ligand and a variety of cold analogs were titrated into the bound protein solution. Interestingly, a number of compounds were observed to enhance the binding of the reporter ligand instead of competing for the same binding site, as expected. The greatest synergistic effect was seen with the alkene precursor of **1**. To explain the enhancement of pheromone binding in the presence of another ligand, we hypothesized that PBPs must work as dimers or higher order multimers in order to recognize two different molecules. The BmorPBP crystallized as a head-to-tail dimer, with one pheromone bound per monomeric unit (Sandler *et al.* 2000). Interestingly, the BmorPBP structure reveals His69 of one PBP molecule to be only 5.6 Å away from His123 of the second PBP in the dimer (Sandler *et al.* 2000). In the first experiment exploring the binding of pheromone at various pHs, we have found that histidine residues may be important to binding. Perhaps titration of the first histidine may cause a conformational change resulting in an equilibrium shift between PBP in the monomer form and PBP as a dimer. The ligand binding itself may lead to a conformational change, perturbing this monomer/dimer equilibrium. To obtain $K_d$ values of **(+)-1** and analogs of **1** such as the alkene precursor **3** as well as an aziridine analog, the dansyl fluorescence assay was utilized where one of the three disulfide bonds of the proteins were selectively reduced followed by covalent attachment of the fluorescent dansyl moiety. The fluorescence change was monitored as the analogs were titrated into solution. Dissociation constants obtained by this novel fluorescence assay were compared with $K_d$ values obtained using

the endogenous tryptophan fluorescence. The results suggest PBPs bind analogs of the pheromone with some discrimination, but the greatest difference in binding is seen at differing ligand:PBP ratios. We hypothesized that these ratios reflect PBP multimerization. We and others have observed that PBPs equilibrate between monomeric, dimeric and higher-order multimeric forms (Leal 2000, Newcomb *et al.* 2002, Ban *et al.* 2003, Honson *et al.* 2003). These equilibria should be noticeable at higher PBP concentrations. Dansyl fluorimetric titrations were performed using 1-3 μM dansylated PBP1. Higher PBP concentrations resulted in a 2-fold increase in binding. Radioassays were performed at low ligand:PBP ratios, and the result was much weaker binding than fluorescent titrations performed at high ligand-PBP ratios. We hypothesized that at low pheromone concentrations, the PBP acts as a transporter. At high pheromone concentrations, multimers of PBP-ligand complexes form. This effect may function to sequester excess pheromone at high stimulus doses, thereby preventing an overloading of receptors. The PBP may provide an automatic stimulus attenuation: at low doses the protein acts as a transporter and at high doses the protein can act as a carrier as well as a scavenger (Honson *et al.* 2003).

### 1.5.4 PBP1 and PBP2 Exhibit Subtle Ligand Discrimination

The third paper briefly explores binding of a variety of compounds with PBP1, both related and unrelated structurally to **1**, using the fluorescence displacement assay with 1-NPN as a reporter (Inkster *et al.* 2005). Values of the test ligand concentration necessary to displace 50% of the reporter ($IC_{50}$) were determined for various ligands. $K_d$ values cannot be determined with these competition assays because ligands dissociate slowly from PBPs (Y. Gong and E. Plettner, unpublished results), and there is variable

kinetic competition between the reporter ligand and the test ligands. Thus, the displacement assay only enabled us to rank the relative kinetic affinity of PBP1 for various ligands. Compounds similar to **(-)-1** revealed strongest affinity, while compounds varying in shape and size were capable of weak binding with PBP1 (Inkster *et al.* 2005).

## 1.6 PBPs and their Role in Olfaction

PBPs have been implicated in olfaction because of their localization in olfactory organs and their ability to bind odorants. This is how the suggestion that PBPs may be odorant transporters arose. However, the role of PBPs must extend beyond being a simple transporter. There are cases where the pheromone or odorant is water soluble and thus requires no transporter. For example, the male cockroach *L. maderae* emits the water soluble compound 3-hydroxy-butan-2-one (Lartigue *et al.* 2003), while *D. melanogaster* can detect ethanol as an odorant. Interestingly, the *D. melanogaster* OBP LUSH is still required for detection of the alcohol (Kim *et al.* 1998). Flies in which LUSH was knocked out could not detect ethanol or vaccenyl acetate (an aggregation pheromone). The sensilla, where LUSH is normally expressed, showed abnormal neuronal activity which could only be corrected by injection of recombinant LUSH protein. A recombinant lepidopteran PBP could not rescue the LUSH mutation (Xu *et al.* 2005). Thus, although PBPs and OBPs are essential for olfaction, their biological role and mode of action are still unclear.

### 1.6.1 The Objectives of this Work

The first step in pheromone binding is thought to involve desorption of the hydrophobic pheromone from the waxy cuticle by PBPs. Here the question arises whether PBPs have an empty binding site, when no pheromone has entered the lymph. It is likely that the antennal lymph contains a variety of potential ligands for PBPs, none of which have been identified. Binding of these endogenous ligands to PBPs would result in the pheromone needing to first displace any endogenously bound ligand. If this were so, a change is necessary in the way researchers study binding kinetics of insect PBPs. In addition, the presence of an endogenous ligand may result in the PBP adopting a certain conformation, perhaps exposing residues involved in pheromone-specific interactions. Lastly, endogenously bound ligands may act to synergize or antagonize pheromone binding. The first objective of this thesis is to isolate and identify any endogenous ligands of PBP1 and PBP2 which may be present in the antennal sensillum lymph of *L. dispar* male moths, and to determine their binding specificities.

The second objective of this thesis is to structurally characterize *L. dispar* PBP1 and PBP2. PBPs are known to be highly α-helical proteins containing three disulfide bonds. These disulfide linkages are thought to contribute to the folding and stability of these proteins. We have found that reduction of one of the three disulfide linkages using a chemical reducing agent results in both PBP1 and PBP2 retaining binding ability with their preferred enantiomer of **1** *in vitro* (Kowcun *et al.* 2001). I wish to determine the disulfide connectivities of *L. dispar* PBP1 and PBP2, as well as the role these disulfide bonds play in PBP conformational stability and their effect on ligand binding.

# CHAPTER 2: MATERIALS AND METHODS

## 2.1 General Methods

### 2.1.1 Purchased/Donated Chemicals

The reagents purchased for protein expression are as follows: isopropyl-β-D-thiogalactopyranoside (IPTG) and DNAse/RNAse mix from Invitrogen (Burlington, ON), *tris*(hydroxymethyl)aminomethane (Tris), ampicillin, phenylmethylsulfonyl fluoride (PMSF), 4-(2-aminoethyl)benzenesulfonyl fluoride (ABSF), Triton-X-100, β-mercaptoethanol, cysteine, and cystine from Sigma (St. Louis, MO), ethylenediaminetetraacetic acid (EDTA) from BDH Inc. (Toronto, ON), and glycine from Bioshop (Burlington, ON).

All reagents used for the endogenous ligand experiments were of the highest purity available. Ethyl acetate was distilled prior to use. Palmitic, linoleic, and oleic acid were from Sigma as were 3-[(3-cholamidopropyl)dimethylammonio]-1-propane sulfonate (CHAPS), and *bis*-(trimethylsilyl)trifluoroacetamide (BSTFA). Stearic acid was obtained from Matheson Coleman & Bell (Cincinnati, OH), cholesterol from Fisher Scientific (Fair Lawn, NJ), potassium hydroxide from BDH Inc., sodium chloride from Caledon (Georgetown, ON), **(+)-1** and dimethyldichlorosilane were obtained from Aldrich (St. Louis, MO). **(-)-1** was a gift from R. Gries (Simon Fraser University). Isopropanol was from Anachemia (Montreal, QC), and postassium chloride was from EM Science (Gibbstown, NJ).

Reagents for peptide mapping and cyanylation of proteins are as follows: Trichloroacetic acid (TCA) was purchased from Aldrich. Dithiothreitol (DTT), *tris*(carboxyethyl)phosphine (TCEP), 1-cyano-4-dimethylamino-pyridinium $BF_4^-$ (CDAP), CNBr, sinapinic acid, and trifluoroacetic acid (TFA) were purchased from Sigma. Ammonium acetate was purchased from BDH Inc., urea from Bioshop Canada Inc., ammonium bicarbonate and sodium dihydrogen phosphate from Fisher Scientific, and formic acid from Anachemia.

For protein crystallography, sodium and potassium dihydrogen phosphate were purchased from Fisher Scientific, N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid (HEPES) from Bioshop, polyethylene glycol (PEG) 4000 from Sigma, and $MgCl_2$ from BDH Inc.

### 2.1.2 Expression, Isolation, and Purification of Recombinant PBPs

PBP1 for the peptide mapping experiment and PBP2 for all experiments described in this chapter were expressed in *Escherichia coli* (with the pHN1+ vector containing ampicillin resistance, supplied by L. Chen and G.L. Verdine, Harvard, via G.D. Prestwich), isolated, and purified as detailed in the literature (Prestwich 1993, Plettner *et al.* 2000). Clones containing the PBP gene were transformed into JM 109 *E. coli* cells by electroporation and grown in LB medium at 37°C. A 10 mL solution of LB containing ampicillin (50 μg/mL) was inoculated and allowed to grow for 8 h. All 10 mL was then used as inocula for 500 mL cultures containing ampicillin (50 μg/mL) and allowed to grow overnight. The 500 mL cultures were then centrifuged at 7000 × g for 10 min, and the pellet resuspended in fresh 2 L of LB medium containing ampicillin (50 μg/mL).

When the optical density at 590 nm reached 0.5-0.8, the culture was induced with 1 mM of IPTG, and the culture was grown for an additional 4.5 h at 27°C. The culture was then centrifuged at 7000 × g for 20 min, and the pellets were kept either at 4°C if used the next day, or flash frozen in liquid $N_2$ and kept at -36°C until needed.

Pellets were washed three times with 100 mM pH 7.0 Tris-HCl buffer and then resuspended in 300 mL lysis buffer (80 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA, 4% glycerol, pH 7.2). Lysozyme (0.1 mg) was added to the resuspension solution and allowed to stir at 4°C for 2 h. Protease inhibitors PMSF and ABSF (16 mg and 0.2 mg, respectively), as well as a DNAse/RNAse mix (100 μg) was added and the solution allowed to incubate for an additional 30 min. The cells were lysed by sonication in a Branson Sonifier 450 (Branson Ultrasonics Corporation, Danbury, CT, settings were duty 50, output 9, 30 minutes), and then centrifuged at 10,000 × g for 10 min to harvest inclusion bodies.

The inclusion body pellet was resuspended in 150 mL of Triton/Tris wash buffer (0.2% Triton-X-100 in 50 mM Tris-HCl pH 6.8), passed through a tissue homogenizer, and then centrifuged at 10,000 × g for 10 min. This process was repeated twice more.

Inclusion bodies were then subjected to a denaturation/renaturation scheme. The pellet was resuspended in 45 mL of 8 N guanidinium HCl and passed through a tissue homogenizer. A reducing solution (10 mM DTT, 19 mM β-mercaptoethanol in 200 mM Tris-HCl pH 8.0) was then added and the solution stirred at 4°C for 2 h under argon. The reduced protein was then diluted into a renaturation buffer (5 mM cysteine in 100 mM Tris-HCl pH 8.0) and stirred at 4°C for 30 min. An oxidizing solution (200 mM cystine in 0.5 N NaOH) was then added and the solution stirred an additional 1 h at 4°C . The

solution was then centrifuged at 7,000 × g for 20 min and the supernatant was concentrated to 45 mL in an Amicon Cell from Millipore (Billerica, MA) using a 10,000 nominal molecular weight cut-off (MWCO) UltraSep Disc from Osmonics Inc. (Oakville, ON).

After concentration, the renatured protein was dialysed against 4 L of 80 mM Tris-HCl pH 8.0 and centrifuged at 10,000 × g for 20 min. The supernatant was again concentrated in the Amicon Cell to about 15 mL. All 15 mL were loaded onto 2 large 12% native PAGE gels, and electrophoresis was allowed to proceed at 160 V using a large electrophoresis chamber (Bio-Rad). Gels were then eluted at 200 mA into a Tris/glycine running buffer using the Whole Gel Eluter (Bio-Rad). Fractions from the eluter were analyzed by 16% native and SDS PAGE. The PBPs had the correct mass by matrix assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) and were positively identified by Western blot. These PBPs were found to be active through binding experiments as described elsewhere (Plettner *et al.* 2000, Kowcun *et al.* 2001, Honson *et al.* 2003).

PBP1 for the binding assays, vesicle constitution, and structural characterization were expressed in *E. coli* using the periplasmic expression system (the pET-22b vector containing ampicillin resistance, purchased from Invitrogen (Burlington, ON)). Vectors containing the PBP1 gene were transformed by electroporation into competent *E. coli* BL21 (DE3) cells. LB (10 mL) containing ampicillin (50 μg/mL) was inoculated and grown for 8 h at 37°C. All 10 mL was then used as inocula for 500 mL cultures containing ampicillin (50 μg/mL) and allowed to grow overnight. The 500 mL cultures were then centrifuged at 7,000 × g for 10 min, and the pellet resuspended in fresh 2 L of

LB medium containing ampicillin (50 μg/mL). When the optical density at 590 nm reached 0.5-0.8, IPTG was added to give a final concentration of 0.1 mM, and the culture was grown for an additional 20 h at 27°C. The culture was then centrifuged at 7,000 × g for 20 min, and the pellet were kept either at 4°C if used the next day, or was flash frozen in liquid $N_2$ and kept at -36°C until needed.

The cell pellet was resuspended in 5 mL resuspension buffer (80 mM Tris-HCl pH 8.0) and subjected to five rounds of freezing in liquid $N_2$ and thawing in a warm water bath (Johnson and Hecht 1994). After the last thaw, 5 mL of resuspension buffer was added and the solution stirred slowly at 4°C for 1 h. The solution was then centrifuged at 10,000 × g for 20 min, and the supernatant containing released protein was then purified by large gel electrophoresis and isolated using the Whole Gel Eluter as described above.

PBP1 and PBP2 used for binding studies were delipidated by incubation with charcoal as follows. A 4 μL solution of charcoal (0.05 g/mL) was added to 1 mL of protein. The solution is incubated at 4°C for 1 h with stirring. The charcoal was removed by centrifugation at 20,000 × g for 5 min., and the supernatant filtered to remove remaining charcoal in a 300 K MWCO Nanosep® centrifugal device (Pall Life Sciences, Ann Arbor, MI) at 5,000 × g. The filtrate was collected and flash frozen in liquid nitrogen.

### 2.1.3 Analytical Methods

Protein concentration was determined spectrophotometrically by measuring the ultraviolet (UV)-absorbance at 280 nm. The concentrations of PBP1 and PBP2 were

calculated using extinction coefficients (13020 $M^{-1}$ $cm^{-1}$ and 14300 $M^{-1}$ $cm^{-1}$, respectively) as determined by their amino acid compositions.

PBP1 and PBP2 for the calibration Western blots were expressed in *Escherichia coli* (with the pHN1+ vector containing ampicillin resistance, supplied by L. Chen and G.L. Verdine, Harvard, via G.D. Prestwich) and expressed, isolated, and purified as from reference (Prestwich 1993, Plettner *et al.* 2000).

Polyacrylamide gel electophoresis (PAGE) was performed on 12% and 16% gels (49:1 bis, pH 8.8) with 4% stacking gels (pH 6.8). Western blotting was performed by transferring proteins onto PVDF membranes (BioRad) using a 25 mM Tris, 230 mM glycine buffer at 200 mA for 1 h. After blotting, gels were stained, and complete transfer of protein from the gel to the membrane was detected. For dot blots, PVDF membranes were pre-soaked in methanol and allowed to air dry. Just prior to the membrane becoming completely dry, 0.4 pmol of each PBP and 1 μL of a single homogenized antenna was applied to the membrane. Polyclonal antibodies were generated previously against the pure, recombinant proteins in rabbits (Plettner *et al.* 2000). These antiPBP1 or antiPBP2 antibodies were used as the primary antibody in Western blots. Goat anti rabbit antibodies, conjugated with alkaline phosphatase, were used as the secondary antibody. The bands were visualized by using the alkaline phosphatase substrate (AP Conjugate Substrate Kit) from Bio-Rad.

Samples for MALDI-TOF MS were prepared by applying 1 μL of a 1:1 mix of protein solution (about 30 μM) and sinnapinic acid (10 mg/mL) in acetonitrile, 3% TFA, and doubly deionized water (3:1:6) to the gold plate. Samples were allowed to air dry, and 1 μL of sinnapinic acid was spotted over top of the dried sample and mixed using the

pipette tip. Again, the sample was allowed to air dry before loaded into the MALDI-TOF MS. Mass spectra were obtained on a MALDI-TOF MS (model PerSeptive Voyager-DE STR from PE Applied Biosystems) with a nitrogen laser (337 nm). Spectra were obtained in linear mode with delayed extraction at an acceleration voltage of 25 kV. Mass spectra obtained were the average of 64 individual spectra, each obtained from one laser pulse.

For GC-MS analysis, mass spectra were obtained on a Varian CP 3800 gas chromatograph equipped with an SPB-5 column (Supelco, 30 m, 0.25 mm internal diameter, 0.25 μm filter), and interfaced with a Saturn 2000 ion trap mass spectrometer. The chromatograph was operated in splitless mode and programmed as follows. For the general program: injector (220°C), column oven (140°C (1.00 min), 20°C/min to 205°C (11.00 min), 40°C /min to 260°C (19.37min) (total 36.00 min)). For the steroids program: injector (280°C), column oven (220°C (3.00 min), 20°C/min to 260°C (25.00 min) (total 30.00 min)). For the steroids high temperature program: injector (280°C), column oven (260°C (3.00 min), 20°C /min to 280°C (30.00 min), 30°C /min to 310°C (5.00 min) (total 40.00 min)). Mass spectra were recorded from 5-17 min for the general program, and 5-30 min for the steroids program, and 5-40 min for the steroids high temperature program. All mass spectra were recorded in the electron-impact (EI) mode with automatic gain control, from 50-500 m/z for the general program, and 50-650 m/z for both steroids programs (scan time 0.65 s/scan for the general program, 0.76 s/scan for both steroids programs, emission current 30 μamps, prescan 100 μs, background 49 m/z). Both treatments and buffer control extracts were run on the corresponding program, and

the control gas chromatographs and mass spectra were subtracted from the treatment gas chromatographs and mass spectra.

For the GC binding assay, chromatographs were obtained on a Varian 3400 gas chromatograph. The chromatograph was operated in splitless mode and programmed as follows: injector (280°C), column oven (100°C (5.00 min), 10°C/min to 200°C (14.00 min), 50°C /min to 250°C (11.00min) (total 30.00 min)).

## 2.2 Identification of Endogenous Ligands

### 2.2.1 Extraction of Endogenous Compounds from the Aqueous Whole Antennal Homogenate

For extractions of antennal homogenate, silanized glass vials were used to prevent absorption of any endogenous ligands that may be present in the homogenate. Glass vials were silanized by filling the vials with 10% dimethyldichlorosilane in toluene allowing the solution to react for 1 h. The vials were drained, rinsed with methanol three times, then with dd$H_2O$ five times and dried overnight in an oven (120°C).

To obtain whole antennal homogenates in aqueous buffer, antennae from newly emerged *L. dispar* males were cut and stored frozen at -36°C. A total of 492 antennae were used in the extraction. Antennae were first thawed and suspended in 7.38 mL of 100 mM Tris-HCl, pH 7.5. The solution was sonicated three times for 2 min with stirring between each sonication period in a Branson Sonifier 450 using the micro tip (settings were duty 30, output control 4). The solution was then passed through a tissue homogenizer. The homogenate was clarified by centrifugation at 10,000 × g for 30 min in a Hermle Model Z383 bench top centrifuge (Hermle Labor Technik, Wehingen, Germany). To remove any remaining membrane structures, the supernatant was

ultracentrifuged at 100,000 × g for 1 h in a Beckman Optima L-90K Ultracentrifuge (Beckman Coulter Inc., Fullerton, CA) and the resulting supernatant was collected (6 mL), flash frozen (in liquid $N_2$), and stored at -36°C.

A quarter of the antennal supernatant (1.5 mL) was used for extraction of whole antennal homogenates in aqueous buffer. The solution was divided into two silanized glass vials and 0.324 mL of freshly distilled ethyl acetate was added. The vials were vortexed for 2 min and then centrifuged for 45 min using a physicians compact centrifuge to break any emulsions. The top organic layer from each vial was then transferred to a clean glass ampoule. The procedure was repeated once more and the solvent collected was concentrated down under vacuum to 25 μL for GC-MS analysis.

### 2.2.2 Extraction of Endogenous Compounds from the Cuticle and from Whole Antennal Homogenates in Ethyl Acetate

For cuticular extracts, antennae from 10 newly emerged *L. dispar* males were extracted. Each pair of antennae were dipped into a glass vial containing 200 μL ethyl acetate for 10 s (antennae were not excised from the moth or damaged in any way). The ethyl acetate was then pooled and centrifuged in a physician's compact centrifuge for 5 min to remove any particles. The organic solvent was then transferred to a clean glass ampoule and the solvent was concentrated down to 25 μL for GC-MS analysis. The whole procedure was repeated two more times for a total of three replicates (30 moths in total).

For whole antennal homogenates in ethyl acetate (a solvent used to dissolve non-polar to weakly polar compounds), the same 20 antennae used for the cuticular extraction were homogenized by resuspending the antennae in 2 mL ethyl acetate. The antennae

were sonicated for 1 min (duty 30, output 3), and the sonication was repeated twice more. The sonicated solution was ground in a tissue homogenizer, and antennal debris was then removed by centrifuging in a physician's centrifuge for 10 min. The supernatant was rotor evaporated to about 0.5 mL, and then concentrated further by vacuum to about 10 μL. Ethyl acetate was added to this solution for a final volume of 25 μL for GC-MS analysis.

For the DMSO perfusion experiments, 2 μL of DMSO were allowed to perfuse into one pair of antennae for about 4 h. About 100 μL of distilled ethyl acetate was then added and the antennae were extracted for GC-MS analysis.

### 2.2.3 Separation of PBP1 and PBP2 from the Aqueous Whole Antennal Homogenate

Three quarters of the antennal supernatant (4.5 mL) was used for isolation of PBP1 and PBP2. The solution was concentrated in 3 K MWCO Nanosep® centrifugal devices and then passed through a 300 K MWCO Nanosep® centrifugal device to filter the solution to yield a final volume of 50 μL. The solution was purified *via* FPLC by injection into the BioLogic Duo Flow Chromatography System (Bio-Rad Laboratories, Hercules, CA) fixed with a UNO Q-1 anion exchange column (Bio-Rad Laboratories, with a 1.3 mL column bed volume) and eluted using an increasing salt gradient consisting of two buffers: Buffer A is 20 mM Tris-HCl pH 7.4, and Buffer B is 20 mM Tris-HCl, 1.0 M NaCl. The following program was employed at a constant flow rate of 2 mL/min: isocratic flow of 100% Buffer A for 6.5 mL, linear gradient of 65% Buffer A and 35% Buffer B for 47.5 mL, linear gradient of 0% Buffer A and 100% Buffer B for 6 mL, isocratic flow of 0% Buffer A and 100% Buffer B for 5.2 mL, isocratic flow of 100%

Buffer A and 0% Buffer B for 6.5 mL. Fractions were collected as follows: 3 mL fractions were collected from 0-11 mL, 0.5 mL fractions were collected from 11-49 mL, and 3 mL fractions were collected from 49-71.7 mL.

Fractions were analyzed by 16% native PAGE and were Coomassie and silver stained, as well as Western blotted and detected with either PBP1 or PBP2 antibodies. The corresponding PBP1 (1.5 mL) and PBP2 (2 mL) fractions were pooled in silanized glass vials. An equivalent amount of freshly distilled ethyl acetate was added to the protein solution, vortexed for 2 min and allowed to stand for 10 min before the organic layer was removed. The procedure was repeated once more, and the combined organic solvent was concentrated by rotary evaporation to about 0.25 mL. The solution was then transferred to a glass ampoule and evaporated under vacuum to 25 μL for GC-MS analysis. Buffer controls were prepared by subjecting buffer to the same procedures, including passage through the BioLogic system.

### 2.2.4 Identification of Endogenous Ligands

To confirm the identity of the acids in the extracts, the extract was silylated using BSTFA, which reacts with hydroxyl groups to form trimethylsilyl derivatives. BSTFA samples were prepared by taking 3-5 μL of the sample into a glass ampoule and evaporating to dryness under a vacuum. The sample was then resuspended in 5 μL of neat BSTFA, sealed, and allowed to react for 1 h. The silylated product was then diluted with 45 μL of ethyl acetate.

To determine the acid content in the extracts of whole antennal homogenates in aqueous buffer, a calibration of trimethylsilylated acid standards consisting of 5, 2, 1, 0.5,

and 0.1 ng of each compound was run on the GC-MS in the general program. The linear equations obtained were used to calculate the amount of acid present in a diluted BSTFA sample. For BSTFA treated whole antennal extracts, a solution of 3.9 antenna equivalents/μL was used. A 1 μL aliquot was derivatized with BSTFA and diluted to 50 μL, resulting in a 0.079 antenna equivalents/μL solution. A fifty times dilution was then injected into the GC-MS, resulting in 0.0016 antenna equivalents being injected. For example, the amount of palmitic acid was determined from the calibration curve to be 2.3 ng injected on the GC-MS, resulting in a total of 1.46 μg/antenna.

To determine the acid content in the PBP1 antennal extract, the linear equations obtained from the calibration of acid standards were used to extrapolate the amount of acid present in a diluted BSTFA sample. For BSTFA treated PBP1 antennal extracts, a solution of 14.76 antenna equivalents/μL was used. A 3 μL aliquot was derivatized with BSTFA and diluted to 50 μL, resulting in a 0.8856 antenna equivalents/μL solution. A volume of 2 μL of this diluted BSTFA solution was then injected into the GC-MS.

To determine the acid content in the PBP2 antennal extract, the linear equations obtained from the calibration of acid standards were used to extrapolate the amount of acid present in a diluted BSTFA sample. For BSTFA treated PBP2 antennal extracts, a solution of 14.76 antenna equivalents/μL was used. A 3 μL aliquot was derivatized with BSTFA and diluted to 50 μL, resulting in a 0.89 antenna equivalents/μL solution. A volume of 2 μL of this diluted BSTFA solution was then injected into the GC-MS.

Retention indices (R. I.) were determined as follows: R. I. = [(log t(A) – log t(N))/(log t(N+1)- log t(N))]*100 + N*100 (Miller and Bruno 2003), where t(A) is the

retention time of the compound which the R. I. is to be determined, t(N) is the retention time of a hydrocarbon standard which elutes just before A, t(N+1) is the retention time of a hydrocarbon standard which elutes just after A, and N is the number of carbons of the hydrocarbon standard which elutes just before A.

### 2.2.5 Determination of PBP1 and PBP2 Concentrations

The concentrations of PBP1 and PBP2 from the aqueous whole antennal homogenate were determined from absorbance readings at 280 nm of the aqueous whole antennal homogenate FPLC trace. The graphing software Origin 6.0 was used to calculate the areas under the corresponding peaks. For PBP1, the area was calculated from 665-746 s, resulting in a peak width of 81 s. The maximum wavelength occurred at 712 s with an absorbance of 0.01166 AU. For PBP2, the area was calculated from 746-843 s, with a corresponding peak width of 97 s. The maximum wavelength occurred at 799.2 s with an absorbance of 0.01027 AU.

For PBP1 and PBP2 concentration determinations from calibrated Western blots, one antenna was homogenized in 20 μL of 20 mM Tris-HCl pH 7.4, centrifuged at 20,000 × g, and the pellet was discarded. This process was repeated so that a total of 10 antennae for each PBP could be loaded onto a 12% native PAGE gel. Gels were electrophoresed at 160 V, transferred to a PVDF membrane by Western blotting, and protein bands were detected using AP (see section 2.1.3). Gels and blots were scanned using a flatbed scanner and images were saved as ".tiff" files at a resolution of 1200 dots per inch.

The amount of protein present in each band was measured by the Quantity One Basic software (Bio-Rad). Volume (in units of intensity × $mm^2$) of each protein band were determined by selecting the area around the protein band. The volume of a representative background area was also selected. The adjusted volume was calculated by subtracting the background volume from the protein band volume. The amount of protein in each band was then calculated from a calibrated protein standard curve of adjusted volume (in arbitrary units (AU)) versus the amount of protein (in μg). The calibration curve was generated by fitting to a logarithmic regression curve.

Calibrated native PAGE was performed on 16% gels using varying concentrations of recombinant PBP1 and PBP2. A total of 6 antennae were homogenized individually as described above and loaded onto the same gels (1 antenna/lane). Gels were electrophoresed at 120-160 V and stained by diluting 0.6 mL of a 10% (w/v) ethanol stock of Coomassie® Brilliant Blue R (Sigma) in 50 mL of fixative solution (12.5% methanol, 7% glacial acetic acid, in ddH2O). Gels were stained for 5 h, and then placed in destaining solution (7% glacial acetic acid in ddH2O) for 10 h. The destained gels were then scanned and protein quantified as described above.

### 2.2.6 Determination of Fatty Acid Vesicle Size Using Dynamic Light Scattering

Fatty acid vesicles were produced with and without PBP1 or PBP2 at concentrations reflecting those found *in vivo*. All fatty acid vesicles were prepared from 10 mg/mL palmitic, linoleic, oleic, or stearic acid stocks in isopropanol. In a glass vial, fatty acid stocks were added and vortexed for 1 min. A 100 mM Tris-HCl, 200 mM KCl, and 40 mM NaCl pH 7.4 buffer (to mimic the sensillar lymph environment (Kaissling 2004)) containing either PBP1 or PBP2 was then added to the fatty acid mix and vortexed

for 2 min. The buffer was passed through a 0.22 μm sterile filter before addition to protein or fatty acids.

For PBP1 associated fatty acid vesicles, the concentrations of each fatty acid used in the dynamic light scattering (DLS) measurements are similar to the concentrations found *in vivo* and are as follows: 0.346 mM of oleic acid and 0.118 mM of stearic acid. This is approximately 2.7 times the amounts found *in vivo*. For PBP2 associated fatty acid vesicles, the concentrations of each fatty acid was a 370 × dilution of what was observed *in vivo* (due to the formation of large aggregates which were inaccessible to light scattering) and are as follows: 18.4 μM of palmitic acid, 20.0 μM of linoleic acid, 4.17 μM of oleic acid, and 5.99 μM of stearic acid. Samples containing PBP1 or PBP2 were prepared by adding a 5 × dilution of the proteins found *in vivo* (because of the large volumes of protein needed to fill the light scattering sample cell) and are as follows: 54.7 μM PBP1 and 52.0 μM PBP2.

Dynamic light scattering measurements were made using an *ALV-Laser Vertriebsgesellschaft m.b.H* ALV-5000 Multiple Tau Digital Correlator. A JDS Uniphase 1145P 23 mW He-Ne (632.8 nm) laser was used to provide a vertically polarized laser light beam. The scattered intensity signal was measured over an angular range from 12° to 150° by a Thorn EMI Electron Tube followed by an ALV Photomultiplier Tube. The intensity autocorrelation function was measured and fitted using Cumulant Data Analysis. DLS measurements of fatty acid samples with and without protein were taken at two different times to demonstrate fatty acid aggregation. The first and second measurements were made within 32 min to 61 min, and 158 min to 178 min, respectively, of when the sample was prepared.

### 2.2.7 Isothermal Titration Calorimetry

Calorimetry experiments were performed on a VP-ITC MicroCalorimeter (MicroCal Incorporated, Northhampton, MA) at Simon Fraser University. The sample cell contained 20 μM of delipidated PBP1 or PBP2 in 100 mM Tris-HCl pH 7.4, 250 mM KOH for palmitic acid titrations, or 100 mM Tris-HCl pH 7.4, 2% MeOH, 80 μM CHAPS for all other titrations. Protein solutions were titrated with 2 mM ligand in 100 mM Tris-HCl pH 7.4, 250 mM KOH for palmitic acid titrations, or 100 mM Tris-HCl pH 7.4, 2% MeOH, 80 μM CHAPS for all other titrations. Titrations consisted of one 1 μL aliquot and then thirty 2 μL aliquots of the test ligand. Injections were of a 20 s duration with 210 s intervals between injections, and was collected with a 2 s filter period at 25°C. In addition, control solutions without protein were titrated with each test ligand. The reference cell contained 100 mM Tris-HCl pH 7.4, 250 mM KOH for palmitic acid titrations, or 100 mM Tris-HCl pH 7.4, 2% MeOH, 80 μM CHAPS for all other titrations. Raw data was processed using the Origin software supplied by the manufacturer. Heats of dilution and mechanical mixing from control runs without protein were subtracted from the heat of binding of solutions with protein. The first 1 μL injection has been omitted from the data analysis. $K_d$ calculations were performed by fitting the data to a non-linear regression model assuming one binding site.

The change in enthalpy (ΔH) for compounds which exhibited slow and noisy heat release were determined by integration of the sample run from 2.2 min to 109.5 min to obtain the heat released during this time. Heat release of control runs were measured in the same way. The change in enthalpy could then be calculated by subtraction of the heat released from the control run from the heat released from the sample run.

The change in entropy (ΔS) was determined as follows:

$$\Delta G_d = - RT \ln K_d \quad (1)$$

$$\Delta G_d = \Delta H - T\Delta S \quad (2)$$

$$\Delta S = (\Delta G_d - \Delta H) / (-T) \quad (3)$$

where $\Delta G_d$ is the change in the Gibbs free energy of dissociation, R is the universal gas constant (8.314 J $K^{-1}$ $mol^{-1}$), and T is the temperature (298 K). We obtain ΔS as follows ((1) into (3)):

$$\Delta S = ((- RT \ln K_d) - \Delta H)) / (-T) \quad (4)$$

## 2.3 Structural Characterization of PBPs

### 2.3.1 *In silico* Measurements and Comparisons

Models of *L. dispar* PBP1 and PBP2 were obtained using the crystal structure of the BmorPBP (Sandler *et al.* 2000) as a template. BmorPBP was used to generate the homology model of *L. dispar* PBP1 and PBP2, respectively. Homology modeling was performed as described in (Honson *et al.* 2003). The 143 amino acid sequence of PBP1 and the 145 amino acid sequence of PBP2 were submitted to the fully automated SWISS-MODEL server using the 'Iterative magic fit' algorithm available within the Swiss-PDB Viewer interface (http://www.expasy.ch/spdbv/) (Guex and Peitsch 1997). Comparative protein modeling was used to generate a protein model since both proteins exhibit significant sequence identity with BmorPBP (61% and 48% with PBP1 and PBP2, respectively). Models of the complexes of pheromones **(+)-1** and **(-)-1** bound to PBP1 and PBP2 were constructed as follows. The ligands were built using Insight/Discover

software (Accelerys, Inc.) implemented on a Silicon Graphics O2 platform. Ligands were constructed assuming all-*trans* conformations for the alkyl chains and were subjected to brief energy minimizations to obtain reasonable structures before docking. Energy minimizations were performed using the CVFF force field {Dauber-Osguthorpe, 1988 #153}. Ligands were docked by manual positioning within the binding site, at a similar location to the position of bombykol within the binding site of BmorPBP (Sandler *et al.* 2000). These complexes were then optimized by energy minimizations (200 steps, using the conjugate gradient algorithm) before examination of the intermolecular interactions (hydrogen bonds, hydrophobic interactions and cation–$\pi$ interactions).

Disulfide bond angles, bond lengths, and the distance between two points of the bridged helices (helix 3 and 6) were also determined using the Swiss-PDB Viewer interface. The angle measured is defined as follows: $\tau$ is the torsion angle with $C_\alpha$ of the lower numbered cysteine residue nearest the viewer (the Newman projection along the S-S bond). The bond length measured is defined as follows: $b_1$ is the S-S bond length of a disulfide bond. The distance between $C_\beta$ of the cysteines comprising the disulfide bridge is $d_1$. Distances between two points of the bridged helices were also measured and are labeled as $d_2$, $d_3$, and $d_4$. Finally, distances were measured between sulfur atoms of the fifth cysteine residue of PBPs and the oxygen atom of a highly conserved aspartic acid amino acid two residues upstream and are labeled as $d_5$.

### 2.3.2 Peptide Mapping and Disulfide Properties

Peptide mapping of PBP1 and PBP2 are described in (Honson and Plettner 2006). PBP1 or PBP2 was precipitated with 1/9 volume of 100% TCA for a final TCA concentration of 10% and incubated on ice for 30 min. The protein solution was then

centrifuged for 5 min @ 13,600 × g, the pellet washed with 50 μL cold acetone and air dried overnight. The final dry weight of protein used in the following reactions was approximately 0.5-1 mg.

Cyanylation reactions for peptide mapping were performed following the procedure of Briand *et al.* (Briand *et al.* 2001) with some modifications (Wu and Watson 1997). The TCEP reduced and CDAP cyanylated protein solution was fractionated by reversed-phase HPLC which contains a Supelco preparative column (25 cm × 4.6 mm, packed with 5 μM C18 Supelcosil™ LC-318 column).

CNBr reactions for peptide mapping were performed following a general procedure (Smith 2002). CNBr chemical cleavage of partially reduced PBP1 and PBP2 was accomplished by dissolving about 1 mg of the protein pellet in water to give a 5 mg/mL solution. After chemical cleavage, samples were then fractionated and masses obtained by LC-ESI-MS at the University of British Columbia's Michael Smith Laboratory/Laboratory of Molecular Biophysics Proteomics Core Facility. Analysis was performed on an API Q STAR PULSARi Hybrid LC-MS-MS which contains an LC Packings column (75 μm I.D. × 15 cm, packed with 3 μm C18 100A pepMap column). *L. dispar* PBPs were eluted by the following linear gradient at a flow rate of 200 nL/min (solution A is 98% dd$H_2O$, 2% acetonitrile (ACN), 0.1% formic acid; solution B is 85% ACN, 15% dd$H_2O$, 0.1% formic acid): 5% solution B to 40% solution B in 60 minutes.

Cyanylation of PBP1 at the C2-C5 disulfide bond for structural characterization was performed as follows: 580 nmol of PBP1 in 50 mM Tris-HCl pH 7.0 was stirred under argon for 1 h at room temperature with a 15 × moles excess of TCEP. A 5 × moles excess of CDAP was then added and the solution incubated at room temperature for 30 h

in the dark with stirring under argon. The reaction was monitored by MALDI-TOF MS. Once two cyanylated groups were detected attached to the protein (addition of 52 Da), the reaction was immediately passed through a C-18 Sep Pak column from Waters (Millford, MA) with a linear gradient starting at 5% ACN and 95% 20 mM Tris-HCl pH 7.4 to 80% ACN at a flow rate of 3 mL/min. Fractions containing protein were detected using a mini-Bradford assay (Bradford reagent from Sigma) in 96 well plates (2:5:11 ratio of sample:Bradford reagent:dd$H_2O$). Visual inspection of the plate revealed which fractions contain protein. Doubly cyanylated protein fractions (PBP1 was reduced at one disulfide bond and cyanylated, so it carries two S-CN moieties) were identified by MALDI-TOF MS. These fractions were pooled and rotor evaporated to remove ACN. The protein solution was then concentrated in 3K Nanoseps to a final volume of 50 μL. The concentrated solution was injected into the FPLC containing the UNO-Q1 anion exchange column and unmodified and modified protein was separated using the method described above. The FPLC purified doubly cyanylated protein was then used for the experiments described below.

### 2.3.3 Circular Dichroism Experiments

Circular dichroism (CD) spectra were recorded at 0.25 mg/mL for the unmodified and modified PBP1. A 20 mM sodium phosphate buffer, pH 7.2, was used for all measurements on a JASCO-810 CD spectropolarimeter. Measurements were taken by averaging five scans in a thermostatted cell from 25°C to 95°C in 5°C increments. The cell path length was 1 mm, and spectra were recorded from 180 nm to 260 nm in 0.5 nm increments.

Secondary structure content was determined using the analysis software SELCON3 (the *self-consistent* method for protein CD analysis, version 3) (Sreerama *et al.* 2000), available online through the CDPro software package at http://lamar.colostate.edu/~sreeram/CDPro. The reference set SMP56 was used, as it contained the largest representation of proteins (43 soluble proteins and 13 membrane proteins).

### 2.3.4 GC Binding Assay

The following GC binding assay was developed by Y. Gong and E. Plettner (unpublished). To 650 μL of a 2 μM PBP1 solution, cyanylated PBP1 (PBP1.CN) or PBP2 solution in 20 mM Tris-HCl pH 7.4 buffer is incubated for 1 h at room temperature with 10 μM of either **(+)-1** or **(-)-1** (delivered from a 1 mM stock solution in ethanol). This solution (100 μL) is then passed through a small column containing 0.061-0.065 g of P2 size exclusion column material (Bio-Rad) which has been swelled with 20 mM Tris-HCl pH 7.4 buffer. The column is then rinsed with the same buffer and the filtrate collected into a silanized glass vial. A 50 μL volume of a 1:1 mix of distilled hexane:ethyl acetate is added to the filtrate and the vial is vortexed on high for 2 min. The organic layer is carefully removed and measured by using a 100 μL Hamilton syringe and transferred to a clean glass vial containing $Na_2SO_4$ as a drying agent. The organic extraction is repeated once more and the extract pooled. The extract is allowed to dry for 1 h before injection into the GC. As a control to determine the total amount of **(+)-1** or **(-)-1** in solution (to account for pheromone absorption onto the glass vial surface), 100 μL of the incubated protein solution is transferred to a clean glass vial and extracted with 1:1 hexane:ethyl acetate as described above. A control was also

performed to determine the amount of column bleed, where **(+)-1** or **(-)-1** without protein was incubated and extracted as described above.

To determine the amount of **(+)-1** or **(-)-1** detected by the GC, a calibration was performed by Y. Gong. Briefly, 1.6 μL of a 12 ng/uL stock solution of **(+)-1** or **(-)-1** in (1:1) hexane:ethyl acetate was injected 3 times. The average of the three injections had a total area count of 39684.0. From this calibration, it was determined that 0.0005 ng/count of **(+)-1** or **(-)-1** was detected.

The dissociation constant was determined as follows:

$$PBP + L_{sol} \leftrightarrows PBP.L \quad (5)$$

$$K_d = [PBP]\,[L_{sol}] / [PBP.L] \quad (6)$$

where PBP is the free PBP and PBP.L is the PBP-pheromone ligand complex; $L_{sol}$ is the free ligand in solution (see below).

The PBP.L elutes from the mini P2 gel column, so the ligand extracted from the eluent allows determination of [PBP.L].

The ligand extracted from a non-filtered aliquot allows us to determine the total ligand present in solution:

$$[L_{tot}] = [L_{sol}] + [PBP.L] \quad (7)$$

This measurement is necessary, because some of the free ligand adsorbs on the vial wall:

$$L_{sol} \Leftrightarrow L_{ad} \quad (8)$$

where $L_{sol}$ is the ligand in solution and $L_{ad}$ is the ligand adsorbed. The total ligand added was 10 μM, thus:

$$10\ \mu M = [L_{sol}] + [L_{ad}] + [PBP.L] \quad (9)$$

The total PBP added to the system was 2 μM, thus:

$$[PBP_{tot}] = [PBP] + [PBP.L] = 2\ \mu M \quad (10)$$

Having measured $[PBP_{tot}]$, $[L_{tot}]$, and [PBP.L], we obtain $K_d$ as follows ((7) and (10) into (6)):

$$K_d = ([PBP_{tot}] - [PBP.L])\ ([L_{tot}] - [PBP.L]) / [PBP.L] \quad (11)$$

### 2.3.5 Crystallography

Crystal Screen 1™, VDX™ plates with sealant and siliconized plastic cover slips were all purchased from Hampton Research (Aliso Viejo, CA). PBP1 pHN1+ crystals were grown using protein purified as described above (see 2.1.2). PBP1 pET crystals were grown using protein purified as above, with an extra purification step added. The protein was further purified using the FPLC containing the anion exchange column and eluted using the method described above for antennal PBP separations (see 2.2.3). All stock protein solutions were in 20 mM Tris-HCl pH 7.4. PBP1 co-crystallized with (+)-1 or (-)-1 were incubated overnight at 4°C with 5-10 times excess pheromone. All reservoir buffers were passed through a 0.22 μM sterile filter before use.

The hanging drop vapor diffusion method was used to grow crystals. In one well, 500 μL of reservoir buffer was added. Protein solution (1 μL) was spotted onto one

plastic cover slip, and 1 μL of reservoir buffer was spotted overtop (without mixing). Plates were stored in a room kept at 18°C.

# *CHAPTER 3:* ENDOGENOUS LIGANDS BIND *L. DISPAR* PBP1 AND PBP2 *IN VIVO*

It is likely that the antennal lymph contains a variety of potential ligands for PBPs, none of which have been identified. These endogenous ligands may participate in positive or negative pheromone blend effects. This first experimental chapter describes the isolation and identification of compounds present in the sensillar lymph. PBP1 and PBP2 were then separated from the lymph and extracted to identify endogenous ligands of these proteins. Also, the concentrations of PBP1 and PBP2 were analyzed and compared with the concentrations of endogenous ligands which they bind.

## 3.1 Fatty Acids and Cholesterol are Present in the Extract of the Aqueous Whole Antennal Homogenate

### 3.1.1 Extraction of the Aqueous Whole Antennal Homogenate

Analysis of the extract of the aqueous whole antennal homogenate by GC-MS (see section 2.2.4) revealed several fatty acids, hydrocarbons and at least one steroid (Table 3.1). Compounds were positively identified by mass spectral fragmentation patterns, as well as comparison with standards when available. Retention indices (R. I.) are included when describing peak retention times, as the R. I. remains virtually unchanged across different programs (see experimental section 2.1.3) and gives a more universal indication of retention on a particular column (Miller and Bruno 2003) (see section 2.2.4). The general program revealed two major peaks. The fastest eluting major peak (8.49 min, R. I. = 1972) was palmitic acid. The second-eluting peak was large and

broad (13.12 min, R. I. = 2173), and was identified as linoleic acid. A third smaller hydrocarbon peak was observed (16.42 min, R. I. = 2303). With the steroids program, which was designed to detect molecules of higher molecular weight, linoleic acid was seen again (5.82 min, R. I. = 2177) as well as a large, broad peak running late in the program (27.49 min, R. I. = 3161). Analysis of the MS fragmentation pattern suggested a steroid with an $M^{\cdot+}$ peak of 386. The compound was identified as cholesterol by comparison with a standard. This result was confirmed using a higher temperature "steroids" program, SHT (see 2.2.4), with which cholesterol eluted (15.03 min, R. I. = 3165).

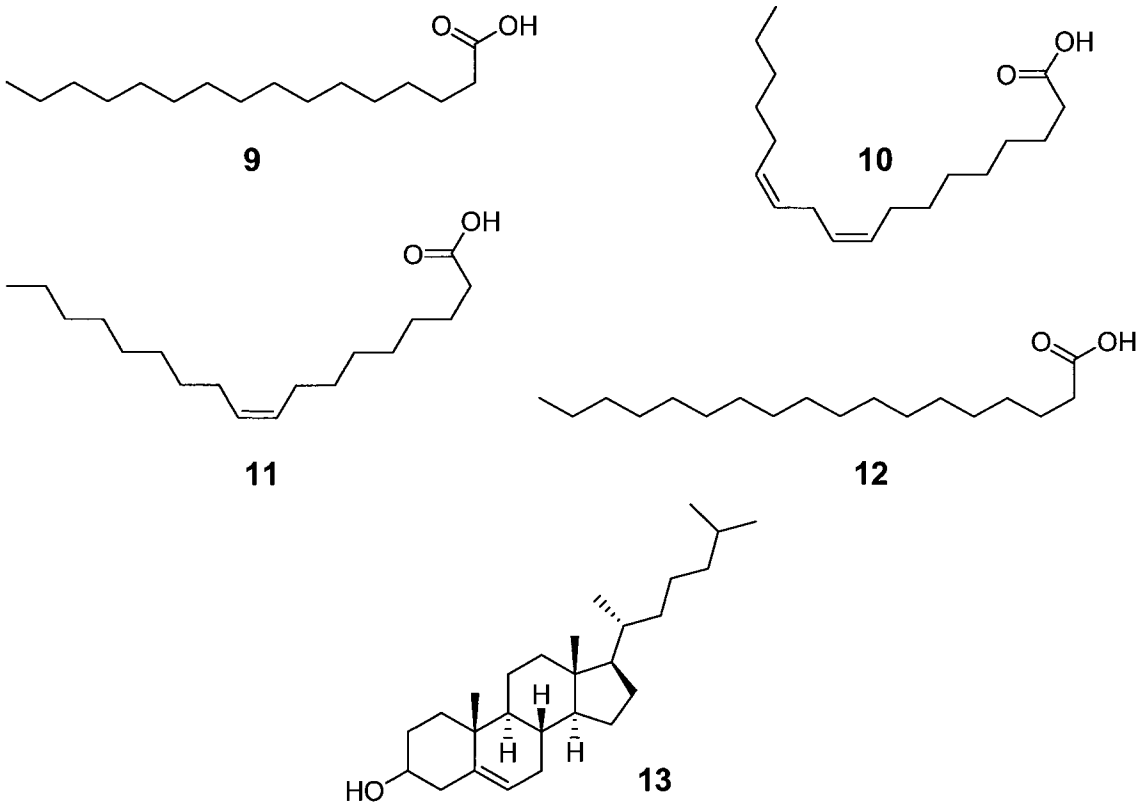



**Figure 3.1** **Compounds identified in the extract of the aqueous whole antennal homogenate of male *L. dispar*. Structures are palmitic acid (9), linoleic acid (10), oleic acid (11), stearic acid (12), and cholesterol (13), all of which were identified in whole antennal extracts of male *L. dispar*.**

**Table 3.1** **Endogenous compounds found in the extract of the aqueous whole antennal homogenate.**

| Compound[a] | Mol. weight (Da) | Prog.[b] | Compound retention time (min) | Standard retention time (min) | Compound R. I. | Amount[d] (μg/ antenna) | Conc.[e] (M) |
|---|---|---|---|---|---|---|---|
| palmitic acid | 256.4 | G | 8.49 | 8.30 | 1972 | nd | nd |
| trimethylsilyl palmitate[a] | 328.4 | G | 9.85 | 9.85 | 2045 | 1.46 | 0.190 |
| linoleic acid | 280.5 | G | 13.12 | nd | 2173 | nd | nd |
| | | S | 5.82 | nd | 2177 | nd | nd |
| trimethylsilyl linoleate[a] | 352.5 | G | 14.43 | 14.45 | 2221 | 11.05 | 1.31 |
| oleic acid | 282.5 | G | not observed | 12.15 | nd | nd | nd |
| trimethylsilyl oleate[a] | 354.5 | G | 14.64 | 14.65 | 2230 | 9.34 | 1.10 |
| | | S | 6.31 | nd | 2237 | nd | nd |
| stearic acid | 284.5 | nd[c] | nd | nd | nd | nd | nd |

| Compound[a] | Mol. weight (Da) | Prog.[b] | Compound retention time (min) | Standard retention time (min) | Compound R. I. | Amount[d] (μg/ antenna) | Conc.[e] (M) |
|---|---|---|---|---|---|---|---|
| trimethylsilyl stearate[a] | 356.5 | G | 15.52 | 15.52 | 2268 | 0.76 | 0.089 |
| | | S | 6.49 | nd | 2259 | nd | nd |
| cholesterol | 386.7 | S | 27.49 | 27.32 | 3161 | nd | nd |
| | | SHT | 15.03 | 14.80 | 3165 | nd | nd |
| trimethylsilyl cholesterol[a] | 458.7 | S | 29.64 | nd | 3224 | nd | nd |
| | | SHT | 15.56 | 15.65 | 3184 | 4.15 | 0.358 |

[a] Trimethylsilyl esters or ethers were generated on a small scale using BSTFA.
[b] Program abbreviations are as follows: G for general, S for steroids, and SHT for steroids high temperature.
[c] Abbreviation "nd" = not determined.
[d] Since detection of trimethylsilylated compounds was more efficient (see section 2.2.4), these data were used to estimate the amount of each compound present in the antenna.
[e] The total amount estimated was used to estimate the average concentration of the detected compounds in the antenna (see section 2.2.4).

To confirm the identity of the acids the extract was silylated using BSTFA (Table 3.1) (see section 2.2.4) to give the trimethylsilyl (TMS) derivatives. The general program revealed four sharp peaks corresponding to palmitic acid (9.85 min, R. I. = 2045), linoleic acid (14.43 min, R. I. = 2221), oleic acid (14.64 min, R. I. = 2230), and stearic acid (15.52 min, R. I. = 2268) (Figure 3.2 A). Compounds were positively identified by mass spectral fragmentation patterns as well as comparison with fragmentation patterns from standards. The higher weight derivatized compounds oleic acid (6.31 min, R. I. = 2237) and stearic acid (6.49 min, R. I. = 2259), as well as cholesterol (29.64 min, R. I. = 3224), were seen in the steroids program and the identities of the acids were confirmed against standards run on the same program (Figure 3.2 B). Finally, only silylated cholesterol was seen in the steroids high temperature program (15.56 min, R. I. = 3184) (Figure 3.2 C) with its mass spectrum matching the standard. Oleic acid elutes very near linoleic acid and, without trimethylsilylation, the two compounds could not be separated.

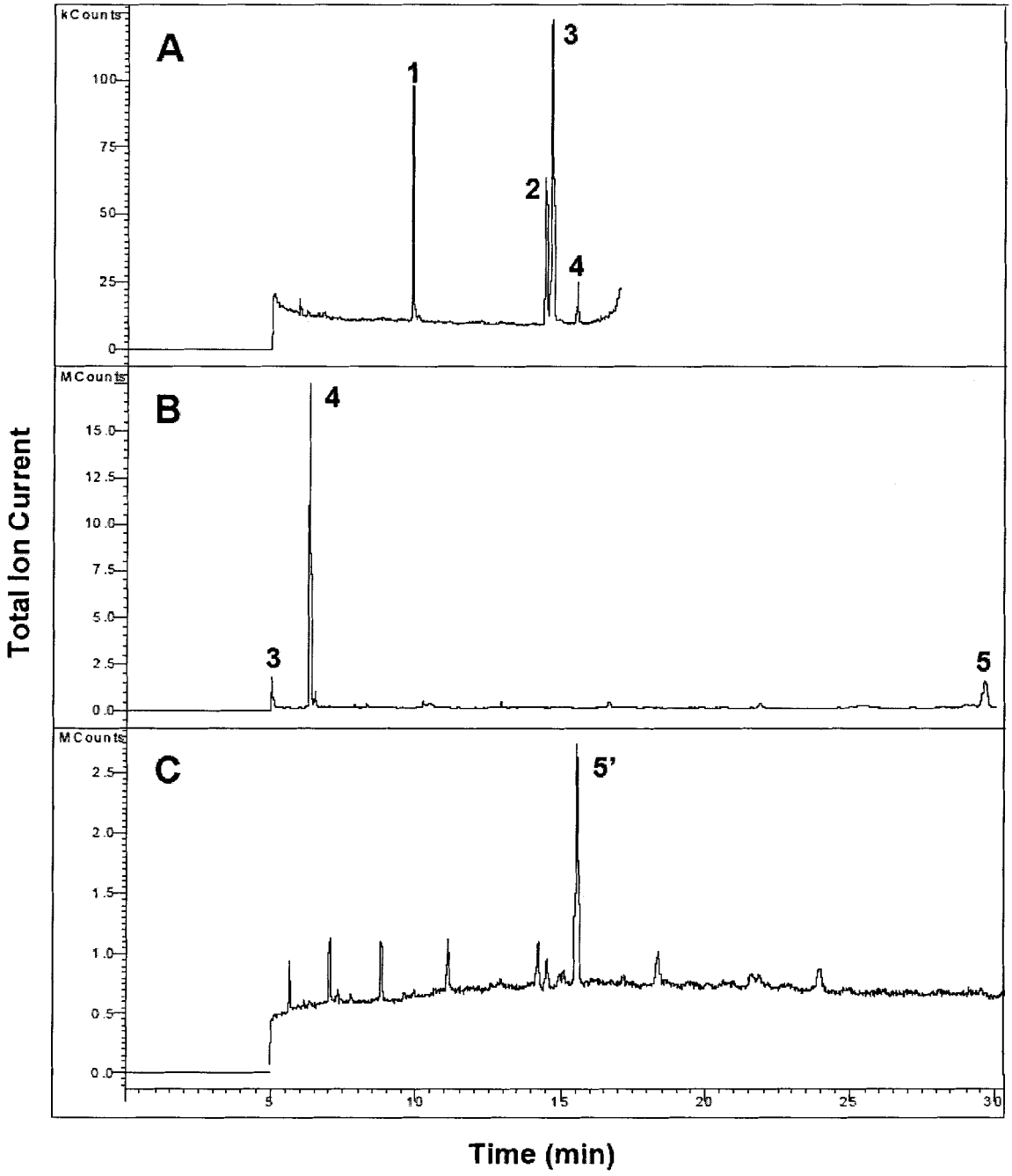



**Figure 3.2** **GC-MS traces of trimethylsilylated compounds found in the extract of the aqueous whole antennal homogenate. (A) The GC trace using the general program reveals sharp peaks of TMS palmitate (1), TMS linoleate (2), TMS oleate (3), and TMS stearate (4). (B) The steroids program enabled the detection of derivitized TMS oleate (3) and TMS stearate (4) as well as TMS cholesterol (5). (C) The high temperature steroids program revealed only silylated cholesterol (5'). The other peaks in this table are from the BSTFA reagent.**

To determine the acid content in the extract of the aqueous whole antennal homogenate, a calibration of trimethylsilylated acid standards consisting of 5, 2, 1, 0.5, and 0.1 ng of each compound was run on the GC-MS in the general program (Figure 3.3 A). The linear equations obtained were used to calculate the amount of acid present in a diluted BSTFA sample (Figure 3.3 B) (see section 2.2.4). From the calibration curve, the amount of palmitic acid was determined to be 2.3 ng, resulting in a total of 1.46 μg/antenna. The acid content for linoleic, oleic, and stearic acid was determined to be 11.05, 9.34, and 0.76 μg/antenna, respectively (Table 3.1).

The sensillar lymph, contained in the sensory hairs on the antenna (see section 1.2.1) is separate from the hemolymph (an insect's equivalent of blood) circulation. Hemolymph, when collected from a cockroach circulatory system, is known to contain 0.45 μg of total free fatty acids (Oguri and Steele 2003). This free fatty acid content in hemolymph should be similar for other insects. Thus, only about 2% of free fatty acids I detected in the antenna could come from hemolymph (total fatty acids detected in lymph is 22.6 μg/antenna). Furthermore, I determined that the cuticle of the antenna only contains 1.3%-15.3% of the fatty acids detected in the extract of the aqueous whole antennal homogenate (see section 3.1.2). Therefore, it is reasonable to assume that most of the free fatty acids and cholesterol detected in the extract of the aqueous whole antennal homogenate came from the combined sensillar lymph from all the sensory hairs on the antenna. A detailed survey of the types of sensilla present on *L. dispar* male moth antennae by light microscopy has revealed 22,035 *s. trichodea* (long length hairs), and 768 *s. trichodea* (medium length hairs) (Scheffler 1975). There are also 3222 *s. basiconica*, 497 *s. coeloconica*, 309 *s. chaetica*, and 85 *s. styloconica* (Scheffler 1975).

Altogether, there were 26,916 sensilla of various types present on one male *L. dispar* antenna surveyed. The volume of a sensillum is estimated to be about 1 pL/hair (see Appendix 1). Since there are about 30,000 hairs of all types (Scheffler 1975), the combined sensillar lymph volume can be estimated to be 0.03 μL (1pL/hair × 30,000 hairs). Therefore, the average concentration in the sensillar lymph of palmitic, linoleic, oleic, and stearic acid was determined to be 0.190, 1.31, 1.10, and 0.0890 M, respectively (see Appendix 1 for a sample calculation of how the palmitic acid concentration was determined).

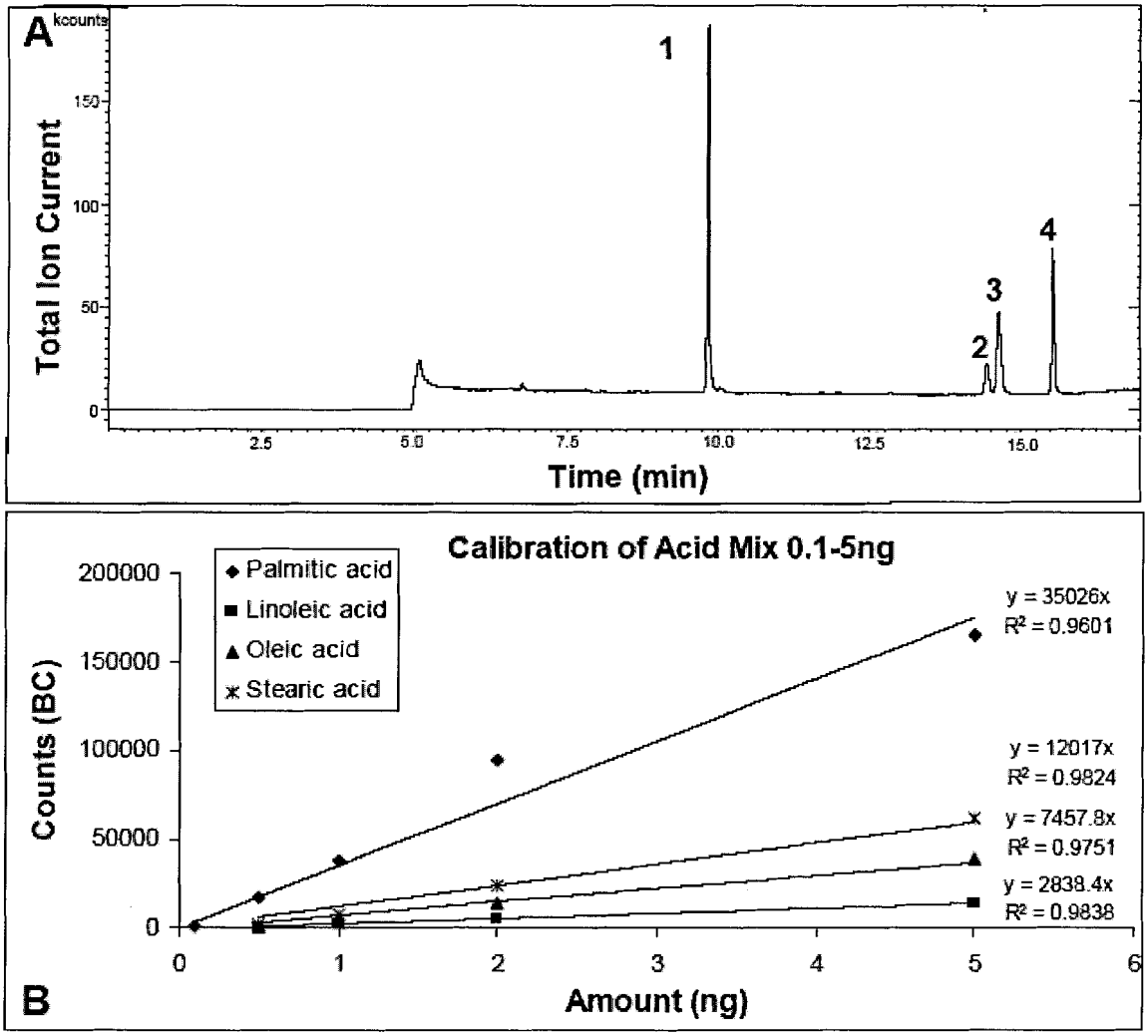



**Figure 3.3** **GC-MS fatty acid calibration. A palmitic (1), linoleic (2), oleic (3), and stearic acid (4) mix consisting of 5, 2, 1, 0.5, and 0.1 ng after derivatization with BSTFA was run on the GC-MS in the general program. (A) GC trace of the 5 ng fatty acid standard mix. (B) The linear equations obtained were used to extrapolate the amount of acid present in a diluted BSTFA sample.**

To determine the cholesterol content in the extract of the aqueous whole antennal homogenate, a calibration of trimethylsilylated acid standards consisting of 50, 20, 10, 5, and 2 ng of each compound was run on the GC-MS using the steroids high temperature program (Figure 3.4). The linear equations obtained were used to calculate the amount of cholesterol present in a diluted BSTFA sample (see section 2.2.4). From the calibration

curve, the amount of cholesterol was determined to be 4.15 μg/antenna, resulting in a concentration of 0.358 M (Table 3.1).

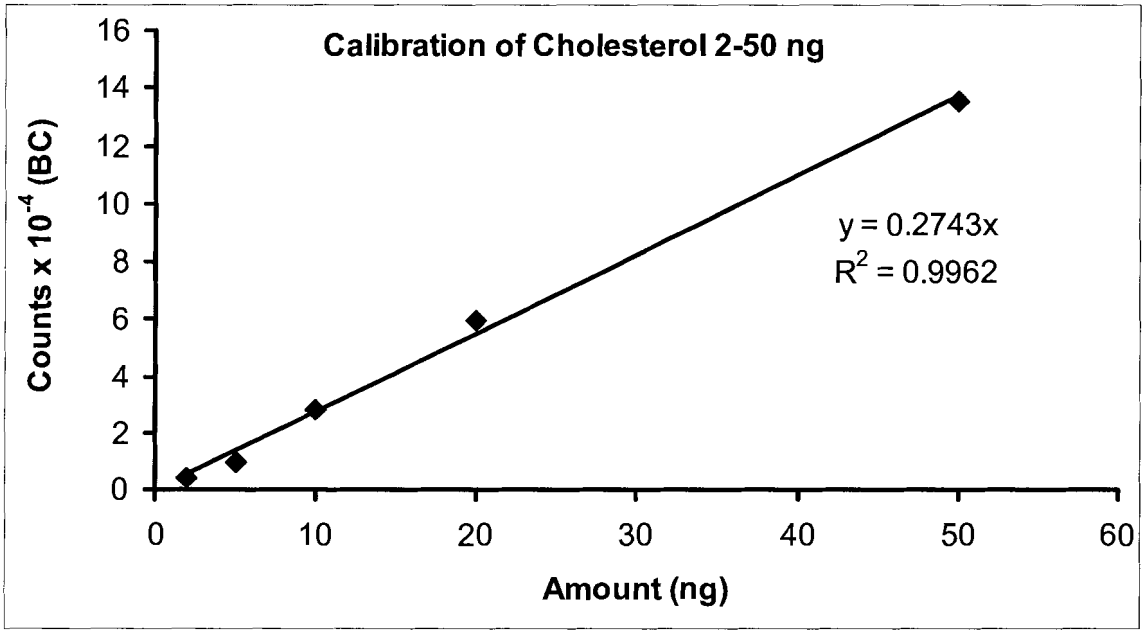



**Figure 3.4** **GC-MS cholesterol calibration. Cholesterol consisting of 50, 20, 10, 5, and 2 ng after derivitization with BSTFA was run on the GC-MS in the steroids high temperature program. The linear equations obtained were used to extrapolate the amount of cholesterol present in a diluted BSTFA sample.**

### 3.1.2 Cuticular and Whole Antennal Extractions into Ethyl Acetate

The fatty acids detected in the extract of the aqueous whole antennal homogenate could come from four sources: 1) the cuticle, 2) the lymph of sensory hairs, 3) the hemolymph or 4) lipase hydrolysis of phospholipids. The amount of free fatty acids and cholesterol that are present in the cuticle of male *L. dispar* antennae were determined by dipping intact antennae into ethyl acetate. All extracts were derivitized with BSTFA. Only derivitized palmitic acid and cholesterol were observed in all three replicates. The amount of palmitic acid was found to be 0.02-0.22 μg/antenna, or 1.3%-15.3% of the amounts determined in the total aqueous extract of the whole antennal homogenate

(Table 3.2) (Figure 3.5). The amount of cholesterol was found to be 0.04-0.13 μg/antenna, or 1.1%-3.1% of the amounts determined in the total extract of the aqueous whole antennal homogenate (Table 3.2) (Figure 3.5). These amounts are relatively minor considering the large amounts of fatty acids present in the aqueous whole antennal homogenate.

**Table 3.2 Amounts of compounds found in cuticular and antennal ethyl acetate extracts. Amounts of compounds found in antennal extracts into aqueous buffer are also shown for comparison.**

| Compound | Amount in aq. extract (μg/antenna) | Cuticular extract into ethyl acetate | | Whole antennal extract into ethyl acetate | |
|---|---|---|---|---|---|
| | | Avg. amount (μg/antenna)[a] | Avg. % of aq. extract | Avg. amount (μg/antenna)[a] | Avg. % of aq. extract |
| palmitic acid | 1.46 | 0.13 ± 0.07 | 8.5 ± 5.0 | 1.06 ± 0.06 | 72.8 ± 3.9 |
| linoleic acid | 11.05 | 0 | 0 | 1.12 ± 0.37 | 10.1 ± 3.3 |
| oleic acid | 9.34 | 0 | 0 | 0.59 ± 0.20 | 6.29 ± 2.1 |
| stearic acid | 0.76 | 0 | 0 | 0.19 ± 0.02 | 24.8 ± 3.0 |
| cholesterol | 4.15 | 0.09 ± 0.03 | 2.2 ± 0.7 | 17.7 ± 8.7 | 425 ± 209 |

[a] Values are the average of three replicates. Standard error is indicated after each value.

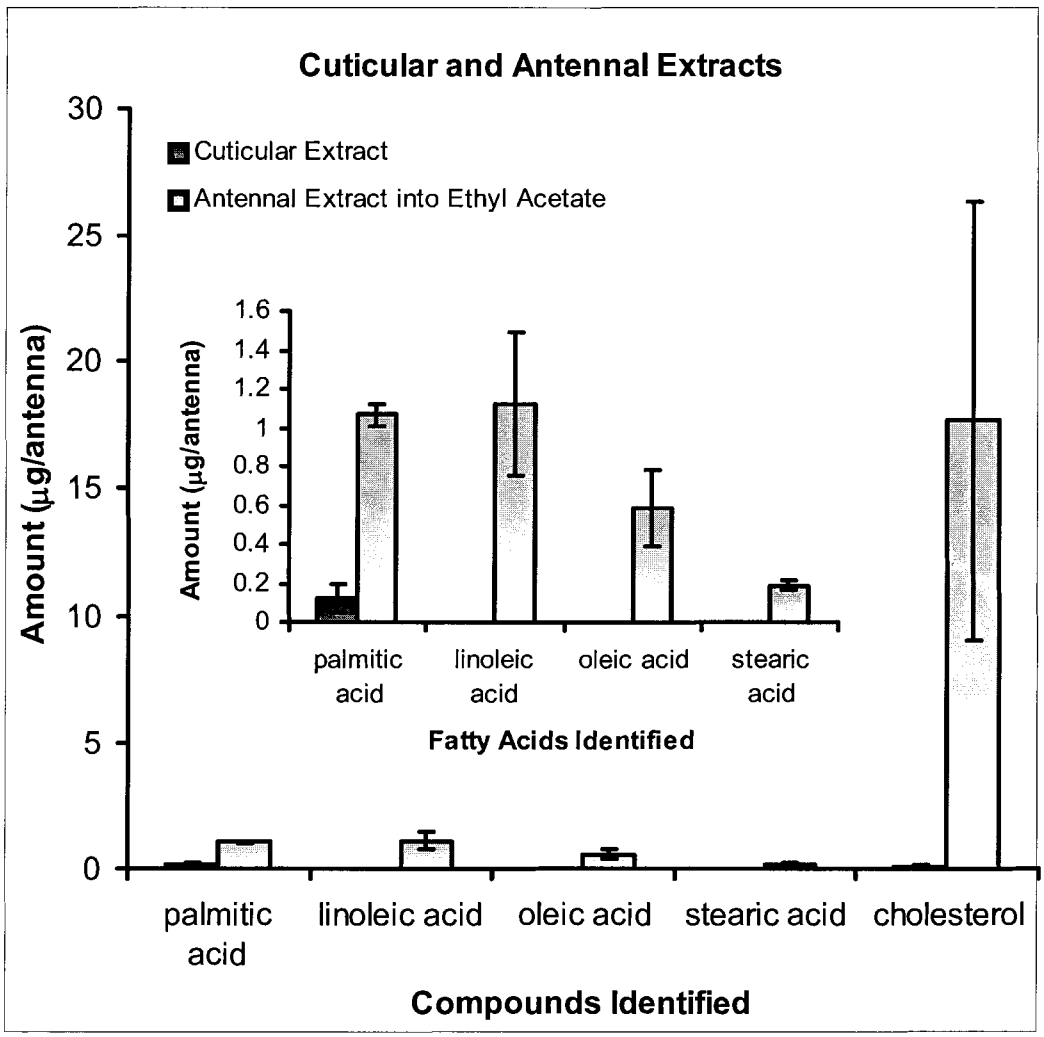



**Figure 3.5** **Bar graph of the amounts of fatty acids and cholesterol found in cuticular and antennal extractions directly into ethyl acetate. Standard errors are indicated by thin bars, and the inset is the same graph but with cholesterol removed.**

To ensure the free fatty acids were not produced from lipases, which may be released during homogenization, whole antennal extractions into ethyl acetate were performed. Extraction directly into organic solvent would denature any lipases released, thus preventing their activity. The extract was derivitized before GC-MS analysis. Again, all four fatty acids and cholesterol were identified. The amounts of palmitic, linoleic, oleic, and stearic acid were determined to range from 1.00-1.15, 0.78-1.72, 0.36-

0.9, and 0.17-0.23 μg/antenna, respectively. When compared to the whole antennal extracts into aqueous buffer, the percentage of each acid ranged from 68.5%-79.1%, 7.1%-15.5%, 3.9%-9.6%, and 22.1%-29.7%, for palmitic, linoleic, oleic, and stearic acid, respectively (Table 3.2) (Figure 3.5). The amounts of these fatty acids found when extracted directly into ethyl acetate may reflect the solubility of these fatty acids. A higher percentage of palmitic and stearic acid were extracted. Both these acids are very water insoluble, and may solubilize into organic solvent much more readily than linoleic and oleic acid. PBPs and other membranous components may also help partition fatty acids into aqueous buffer by non-specific association, but they may retain these acids in the presence of organic solvent. In either case, the amount of free fatty acids extracted directly into ethyl acetate are still in the hundreds of millimolar range.

The amount of cholesterol was determined to range from 4.34-28.42 μg/antenna, or 105%-685% of the whole antennal extract into aqueous buffer (Table 3.2) (Figure 3.5). The increase in the amount of cholesterol may again be due to increased solubility of cholesterol in ethyl acetate. The variability of the amounts of fatty acids and cholesterol determined in different replicates suggest differences exist between pupae reared at different times. Perhaps environmental conditions such as food source or the age of the moth has an affect on fatty acid and cholesterol composition. In either case, high amounts (in the molar or hundreds of millimolar range, see Table 3.1) of free fatty acids and cholesterol are present in *L. dispar* male antennae.

### 3.1.3 Whole Antennal Perfusion Experiments

Antennae were also subjected to perfusion experiments where a small amount of DMSO was incubated with antennae for more than 4 h (data courtesy of E. Plettner). The

perfused antennae were then extracted with ethyl acetate, derivatized with BSTFA, and the amounts of fatty acids were determined by GC-MS analysis. Perfusion allows for extraction of fatty acids without homogenization. Hence, fatty acids detected are from the lymph and not from disruption of lipid membranes. Also, perfusion in such a small amount of DMSO would allow lipases to remain active, allowing us to compare these results with the results from the aqueous whole antennal homogenate (where lipases could also have been active), and the ethyl acetate whole antennal homogenates (lipases are not active). Once again, trimethylsilylated palmitic, linoleic, oleic, and stearic acid were observed (Table 3.3). The amounts of these fatty acids were in the same order as amounts found in homogenized whole antennal extractions in ethyl acetate (see section 3.1.2). Cholesterol could not be detected since these extracts were only run on the general program. Since the amounts of fatty acids and cholesterol are similar with the amounts found in the aqueous and ethyl acetate whole antennal homogenates, we can conclude that lipases are not likely major contributors to the free fatty acid pool detected in the antennal extracts.

**Table 3.3** **Amounts of compounds found in perfusion experiments in DMSO. Amounts of compounds found in the extract of the aqueous whole antennal homogenate are also shown for comparison.**

| Compound | Amount in aq. extract (μg/antenna) | Perfusion in DMSO | |
|---|---|---|---|
| | | Avg. amount (μg/antenna)[a] | Avg. % of aq. extract |
| palmitic acid | 1.46 | 0.4 ± 0.10 | 26.1 ± 6.9 |
| linoleic acid | 11.05 | 3.2 ± 1.37 | 28.5 ± 10.2 |
| oleic acid | 9.34 | 1.3 ± 0.54 | 14.9 ± 5.1 |
| stearic acid | 0.76 | 0.6 ± 0.10 | 71.9 ± 15.5 |

## 3.2 PBP1 Binds C18 Fatty Acids Endogenously

Since large amounts of fatty acids and cholesterol were seen in the whole antennal extracts, I isolated PBP1 and PBP2 from the antenna to observe if they bind any of these endogenous compounds. Three-quarters of the aqueous whole antennal tissue homogenate was purified on a strong anion exchange column by FPLC to separate PBP1 and PBP2 from other antennal proteins (Figure 3.6 A). Fractions containing PBP1 and PBP2 were identified by native PAGE (Figure 3.6 B) and Western blotting (Figure 3.6 C). After pooling and extracting the corresponding fractions with ethyl acetate, 1-2 μL of the ethyl acetate extract was run on the GC-MS to identify any ligands that may be associated with these proteins. To account for any contaminating compounds that may have eluted from the FPLC, control samples of buffer extracts and blank FPLC runs were also run on the GC-MS (data not shown).

The general program did not reveal any significant peaks. When the extract was run with the steroids program, two large broad peaks were observed late in the program (28.26 min, R. I. = 3190, and 29.49 min, R. I. = 3221) (Figure 3.7 B) (Table 3.4). The broad peaks both contain a 386 m/z fragment, suggesting they both contain a sterol backbone and are similar to cholesterol whose $M^{\cdot+}$ is 386 m/z. These two compounds are referred to as cholesterol derivative1 (28.26 min, R. I. = 3190) and cholesterol derivative2 (29.49 min, R. I. = 3221). Finally, when the PBP1 extract was run using the "steroids high temperature" program, the two cholesterol derivatives were observed as expected and consisted of large, broad peaks (10.99 min, R. I. = 2986, and 13.38 min, R. I. = 3101) (Figure 3.7 C). Interestingly, two other large, broad peaks were also seen later in the program (18.81 min, R. I. = 3287 and 20.56 min, R. I. = 3334) (Figure 3.7 C).

These two compounds could not be identified by the fragmentation pattern seen in the mass spectra.

To confirm the identities of the compounds seen in the PBP1 extract, BSTFA was used to silylate the PBP1 extract. The gas chromatogram of the general program revealed that PBP1 associates with the saturated and singly unsaturated C18 fatty acids detected in the extract of the aqueous whole antennal homogenate; namely, small oleic and stearic acid peaks (14.76 min, R. I. = 2235, and 15.59 min, R. I. = 2271, respectively) were observed (Figure 3.7 A) (Table 3.4). Only buffer contaminants were observed in the steroids program. Two small, broad peaks of unknown compounds were observed in the gas chromatogram of the steroids high temperature program (17.92 min, R. I. = 3261, and 19.98 min, R. I. = 3319), as well as a small unidentified peak (22.71 min, R. I. = 3385).

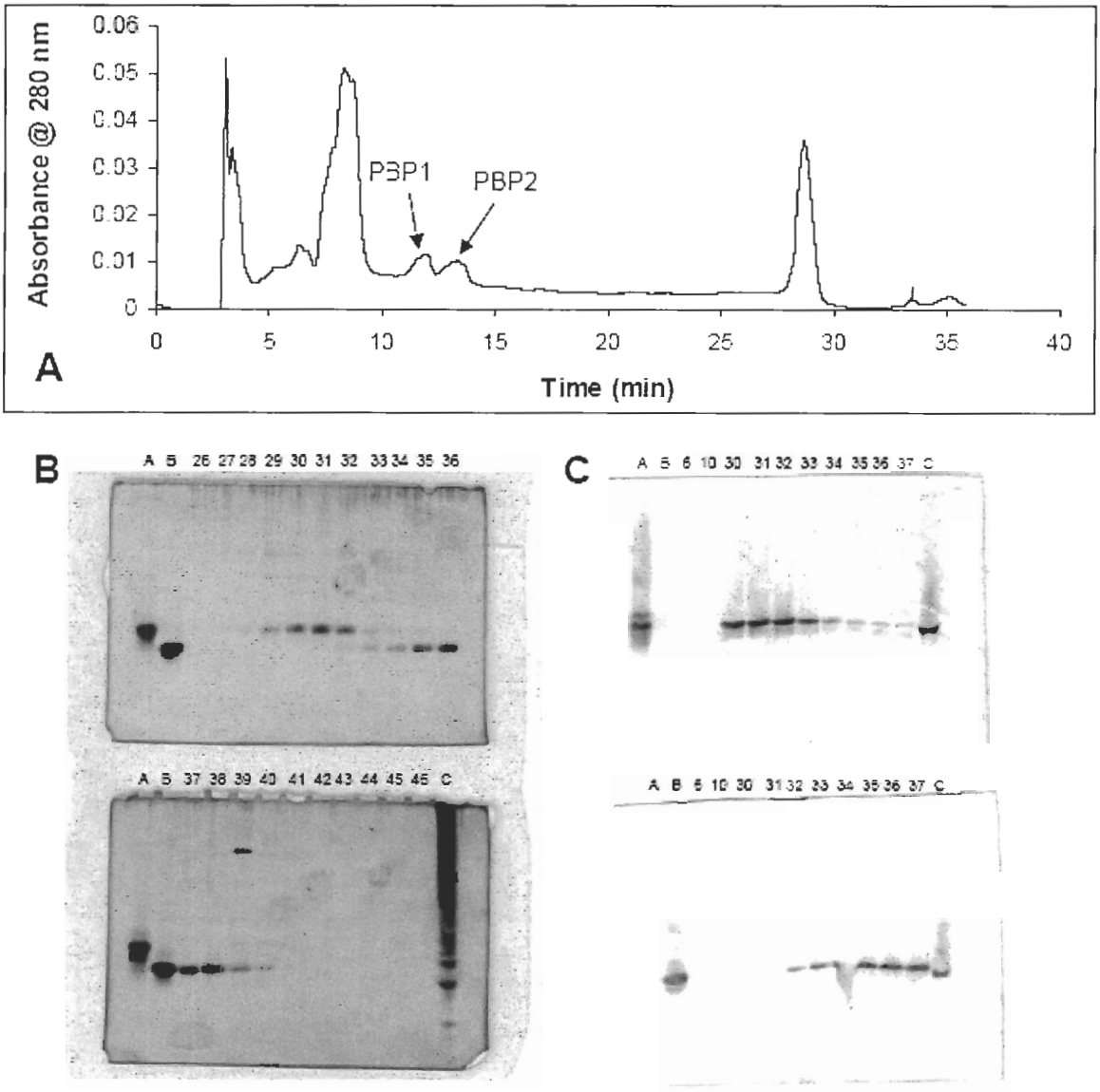



**Figure 3.6** **PBP1 and PBP2 separation from the aqueous whole antennal homogenate. (A) FPLC trace of the whole antennal extract with PBP1 and PBP2 peaks indicated. (B) 16% native PAGE of fractions containing PBP1 and PBP2. (C) Western blotting confirmation of PBP1 and PBP2. Top blot treated with PBP1 anti-serum; bottom blot treated with PBP2 anti-serum. Lane A is a recombinant PBP1 standard, lane B is a recombinant PBP2 standard, lane C is the clarified whole antennal homogenate, and numbered lanes correspond to the FPLC eluted fractions.**

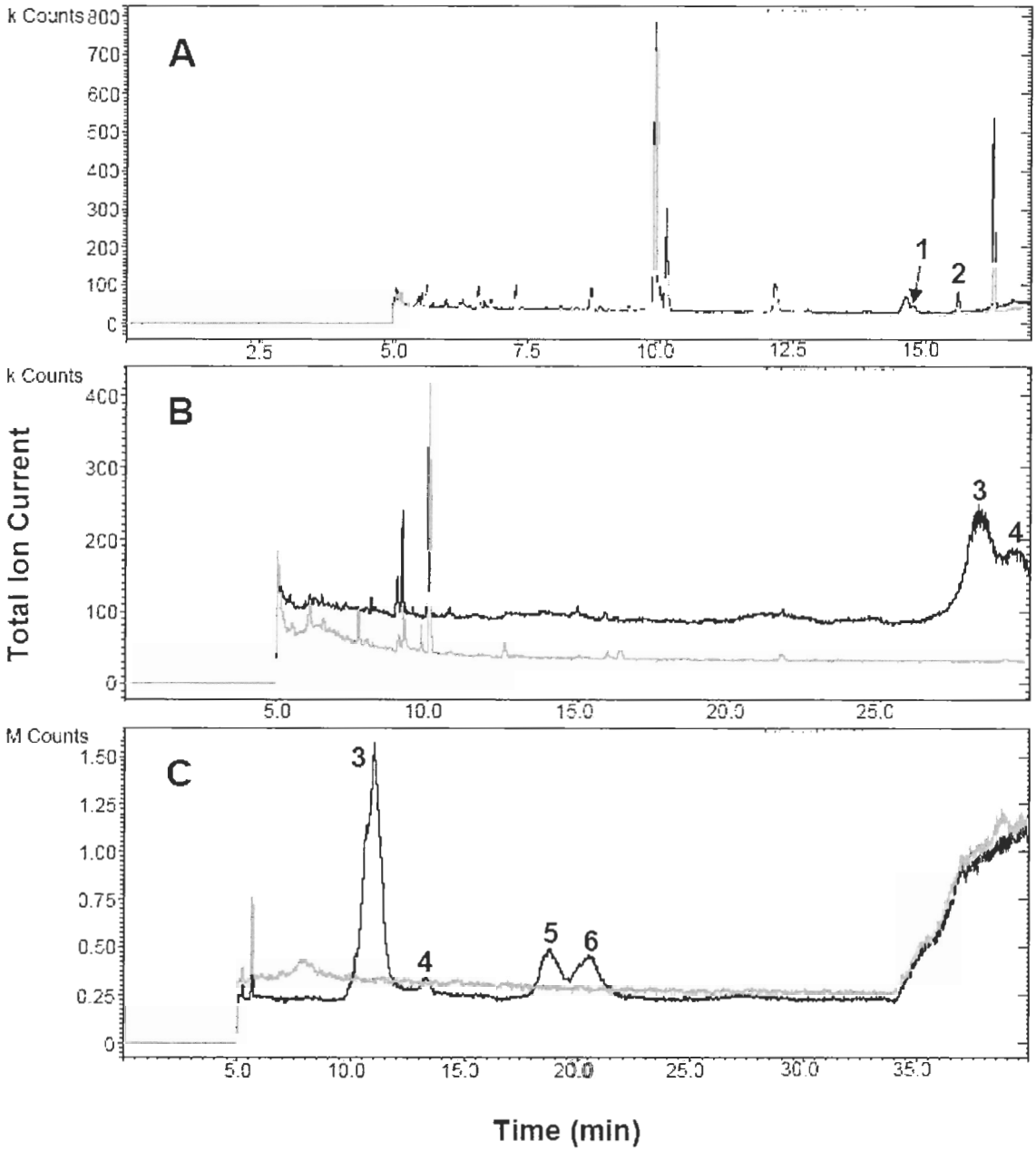



**Figure 3.7** **GC traces of the PBP1 extract. GC traces of the PBP1 extract are shown in black while the buffer extract controls are shown in grey (spectra are overlaid). PBP1 seems capable of associating with only two of the four fatty acids identified in the extract of the aqueous whole antennal homogenate. (A) The general program shows BSTFA derivatized oleic (1) and stearic acid (2) peaks. (B) Cholesterol derivative peaks (3) and (4) were detected with the steroids program (no BSTFA was added). (C) Cholesterol derivative peaks (3) and (4) as well as two unidentified compounds (5, 6) were detected with the steroids high temperature program (no BSTFA was added).**

**Table 3.4 Endogenous compounds found bound to PBP1.**

| Compound[a] | Mol. weight (Da) | Prog.[b] | Compound retention time (min) | Standard retention time (min) | Compound R. I. | Amount[d] (ng/ antenna) | Conc.[e] (μM) |
|---|---|---|---|---|---|---|---|
| oleic acid | 282.5 | G | not observed | 12.15 | nd | nd | nd |
| trimethylsilyl oleate[a] | 354.5 | G | 14.76 | 14.65 | 2235 | 1.08 | 127 |
| stearic acid | 284.5 | nd[c] | nd | nd | nd | nd | nd |
| trimethylsilyl stearate[a] | 356.5 | G | 15.59 | 15.52 | 2271 | 0.37 | 43.4 |
| cholesterol derivative1 | nd | S | 28.26 | nd | 3190 | nd | nd |
| | nd | SHT | 10.99 | nd | 2986 | nd | nd |
| cholesterol derivative2 | nd | S | 29.49 | nd | 3221 | nd | nd |
| | nd | SHT | 13.38 | nd | 3101 | nd | nd |

[a] Trimethylsilyl esters or ethers were generated on a small scale using *bis*-trimethylsilyl trifluoroacetamide (BSTFA).
[b] Program abbreviations are as follows: G for general, S for steroids, and SHT for steroids high temperature.
[c] Abbreviation "nd" = not determined.
[d] Since detection of trimethylislylated compounds was more efficient (see section 2.2.4), these data were used to estimate the amount of each compound present in the antenna.
[e] The total amount estimated was used to estimate the average concentration of the detected compounds in the antenna (see section 2.2.4).

From the calibration curve, the amount of oleic and stearic acid (with a BSTFA treated buffer control subtracted) was determined to be 1.08, and 0.37 ng/antenna, respectively (Table 3.4). The concentrations of oleic and stearic acid were determined to be 127 and 43.4 μM, respectively.

## 3.3 PBP2 Binds C16 and C18 Fatty Acids Endogenously

The general program revealed two peaks with mass spectral fragmentation patterns similar to hydrocarbons, yet no saturated hydrocarbon standard we ran matched the sample spectra (6.34 min, R. I. = 1851, and 14.77 min, R. I. = 2236). Endogenous ligands were not observed using either steroids program.

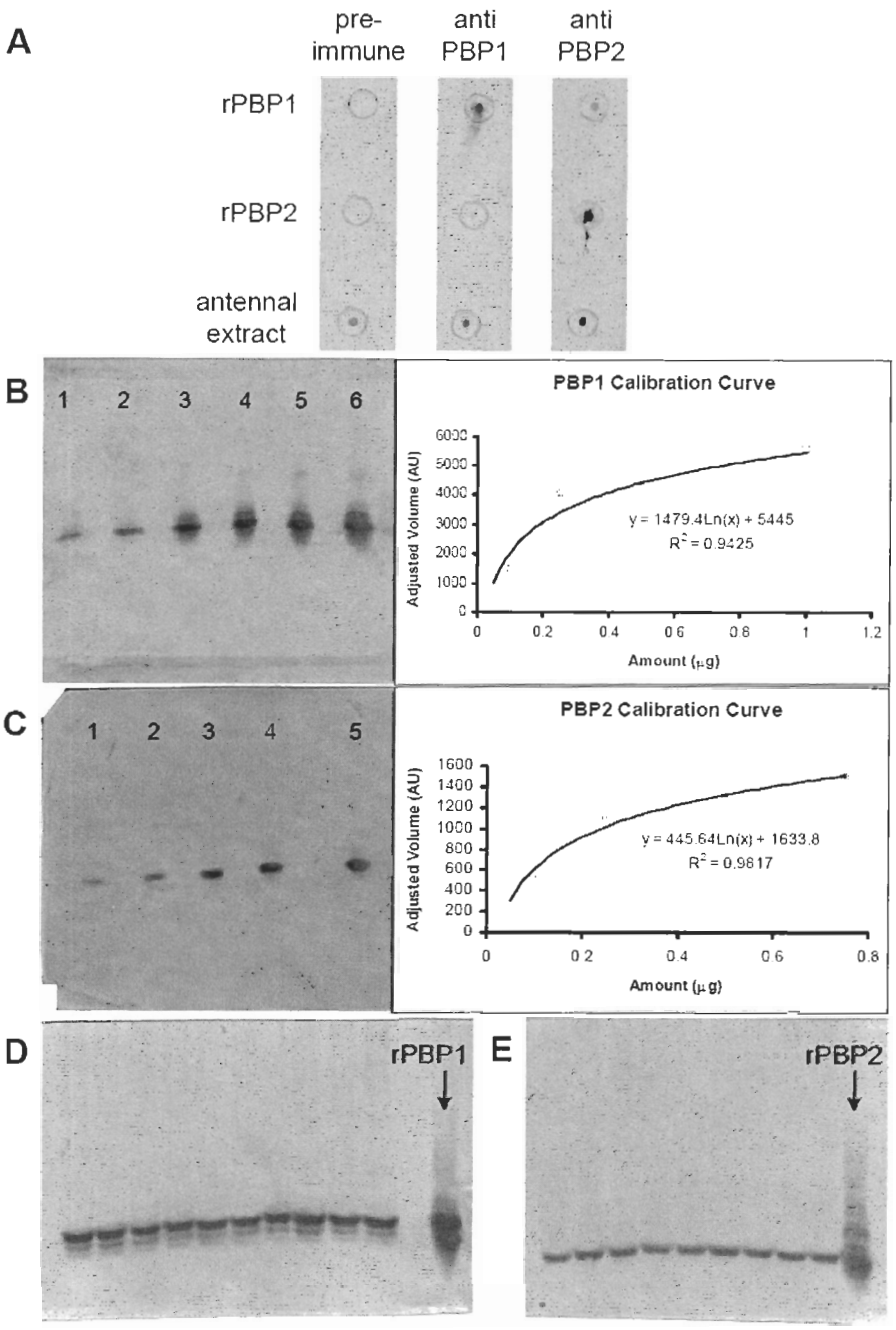



**Figure 3.9** **Determination of PBP concentrations *in vivo* by immunoblotting. (A) Dot blots on PVDF membrane. Note the pre-immune bleed does not show any staining against recombinant PBP1 or PBP2. Also, there is little cross-reactivity of the PBP1 antibody against recombinant PBP2, and little cross-reactivity of the PBP2 antibody against recombinant PBP1. (B) The PBP1 calibration immunoblot (left), and calibration curve (right). (C) The PBP2 calibration immunoblot (left) and calibration curve (right). Lanes 1 through 6 are PBP concentrations as follows: 0.05 µg, 0.1 µg, 0.25 µg, 0.5 µg, 0.75 µg, and 1 µg. (D) PBP1 immunoblot with recombinant PBP1 standard (rPBP1) indicated. (E) PBP2 immunoblot with recombinant PBP2 standard (rPBP2) indicated. Each lane contains the supernatant of one homogenized antenna. Adjusted volumes were calculated for all lanes.**

| Compound[a] | Mol. weight (Da) | Prog.[b] | Compound retention time (min) | Standard retention time (min) | Compound R. I. | Amount[d] (ng/ antenna) | Conc.[e] (mM) |
|---|---|---|---|---|---|---|---|
| trimethylsilyl oleate[a] | 354.5 | G | 14.83 | 14.65 | 2238 | 13.0 | 1.53 |
| stearic acid | 284.5 | nd[c] | nd | nd | nd | nd | nd |
| trimethylsilyl stearate[a] | 356.5 | G | 15.62 | 15.52 | 2272 | 18.8 | 2.20 |

[a] Trimethylsilyl esters or ethers were generated on a small scale using *bis*-trimethylsilyl trifluoroacetamide (BSTFA).
[b] Program abbreviations are as follows: G for general and S for steroids.
[c] Abbreviation "nd" = not determined.
[d] Since detection of trimethylislylated compounds was more efficient (see section 2.2.4), these data were used to estimate the amount of each compound present in the antenna.
[e] The total amount estimated was used to estimate the average concentration of the detected compounds in the antenna (see section 2.2.4).

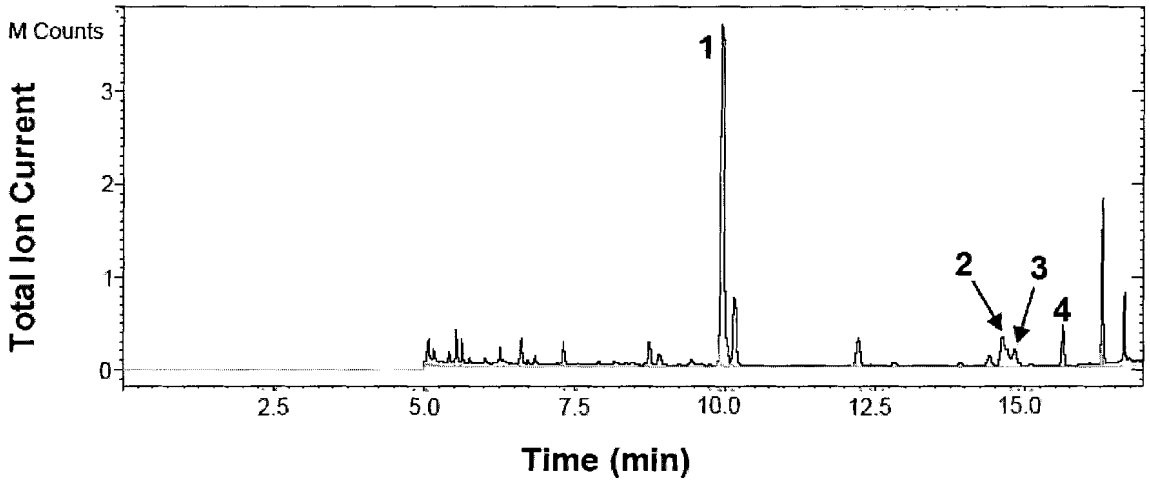



**Figure 3.8** **GC trace of the PBP2 extract. GC trace of the PBP2 extract is shown in black while the buffer extract control is shown in grey (traces are overlaid). PBP2 seems to be capable of associating with all the fatty acids identified in the whole antennal extract. The general program trace revealed derivatized palmitic (1), linoleic (2), oleic (3), and stearic acid (4).**

## 3.4 PBP1 and PBP2 Concentrations in the Whole Antennal Tissue Homogenate

The concentrations of PBP1 and PBP2 isolated *in vivo* from male *L. dispar* antennae were calculated from the FPLC trace of the whole antennal tissue homogenate, from calibrated Western blots, and from calibrated Coomassie-stained gels. From the

FPLC trace we could determine the concentration of each protein from the area under the corresponding peak (Figure 3.6 A). For PBP1, the concentration was determined as follows: 0.78844 abs × s (the area under the PBP1 peak) / 13,020 $M^{-1}$/abs (the PBP1 extinction coefficient) / 81 s (the length of time PBP1 was eluted from the column) = 0.748 μM. The buffer volume containing PBP1 was $3.33 \times 10^{-5}$ L/s (column flow rate) × 81 s = 2.7 mL. Therefore, 0.748 μM × 2.7 mL / 369 antennae (total number of antennae homogenate injected into the FPLC) = $5.47 \times 10^{-12}$ mol/antenna. The volume of lymph in the *s. trichodea* of one antenna can be estimated. From previous calculations, the volume of one hair is 1 pL (see Appendix 1). The number of *s. trichodea* was found to be 22,000 (Scheffler 1975). Thus, the volume of lymph in each *s. trichodea* is 1 pL/hair × 22,000 hairs = 0.022 μL (PBPs reside only in *s. trichodea*, thus this number of hairs is appropriate for estimating PBP concentrations). Since the molecular weight of PBP1 is 16,139.8 Da, the total concentration of PBP1 was estimated to be 0.25 mM in the *s. trichodea* lymph, or 0.088 μg/antenna.

For PBP2, the concentration was determined as follows: 0.82398 abs × s (the area under the PBP2 peak) / 14,300 $M^{-1}$/abs (the PBP2 extinction coefficient) / 97 s (the length of time PBP2 was eluted from the column) = 0.594 μM. The buffer volume containing PBP2 was $3.33 \times 10^{-5}$ L/s (column flow rate) × 97 s = 3.2 mL. Therefore, 0.594 μM × 3.2 mL / 369 antennae = $5.20 \times 10^{-12}$ mol/antenna. Since the molecular weight of PBP2 is 16,141.6 Da, the total concentration of PBP2 was estimated to be 0.24 mM in the *s. trichodea* lymph, or 0.084 μg/antenna.

PBP concentrations were confirmed by Western blots. Dot blots on PVDF membrane were performed to ensure that the pre-immune bleed does not show any

staining against recombinant PBP1 or PBP2. Also, there is little cross-reactivity of the PBP1 antibody against recombinant PBP2, and little cross-reactivity of the PBP2 antibody against recombinant PBP1 (Figure 3.9 A). Calibrated immunoblots of PBP1 and PBP2 were used to obtain a logarithmic regression curve from which an equation relating the adjusted volume of the band (measured using the Quantity One Basic software from Bio-Rad) with the amount of protein loaded (Figure 3.9 B, C). An immunoblot with repeating lanes each containing supernatant from one homogenized antenna was analyzed and the adjusted volume of the bands measured (Figure 3.9 D, E) (immunoblotting of homogenized antennal supernatant was performed by Ivy Ling). The average amounts of PBP1 and PBP2 were determined to be 0.112 ± 0.006 μg/antenna and 0.163 ± 0.013 μg/antenna, respectively. Thus, the concentrations of PBP1 and PBP2 are 0.34 ± 0.02 mM and 0.46 ± 0.04 mM, respectively.

I also analyzed Coomassie stained gels of antennal extracts by densitometry to obtain the amount of PBP detected against calibration gels with recombinant PBP1 (Figure 3.10 A) and PBP2 (Figure 3.10 B). The total amounts/antenna determined by this method were 0.635 ± 0.039 μg PBP1 and 0.733 ± 0.028 μg PBP2. This translated to concentrations of 1.79 mM PBP1 and 2.06 mM PBP2. This method provides an over-estimate of the PBP concentration, because the staining does not distinguish between the PBPs of interest and other OBPs electrophoresing in the same region.

The three PBP concentration estimates are similar with each other, but are lower than published estimates (see section 3.7.2). PBPs interact with other components of the lymph (see section 3.5) and, for that reason, it was important to obtain good estimates of PBP concentrations.

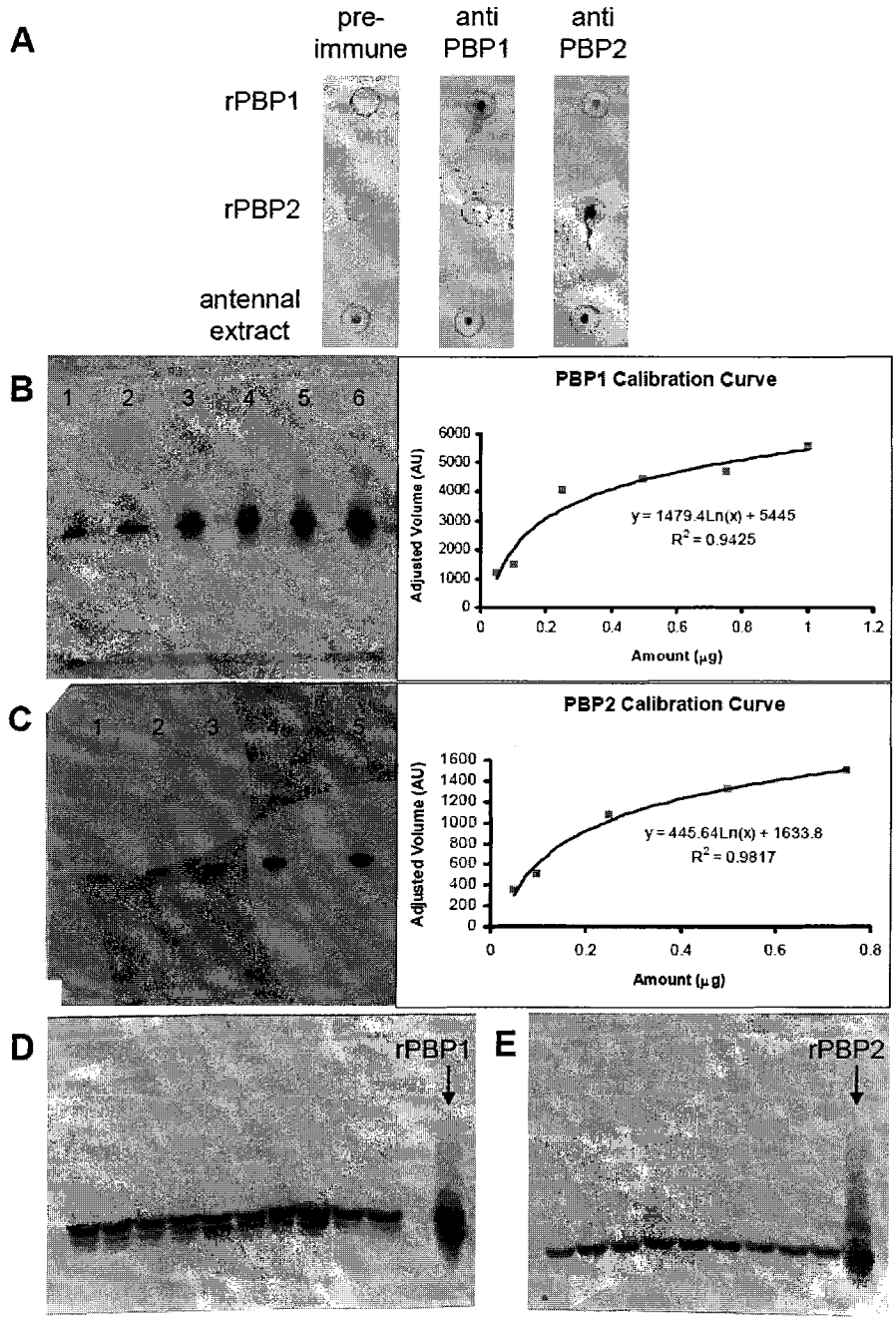



**Figure 3.9** **Determination of PBP concentrations *in vivo* by immunoblotting. (A) Dot blots on PVDF membrane. Note the pre-immune bleed does not show any staining against recombinant PBP1 or PBP2. Also, there is little cross-reactivity of the PBP1 antibody against recombinant PBP2, and little cross-reactivity of the PBP2 antibody against recombinant PBP1. (B) The PBP1 calibration immunoblot (left), and calibration curve (right). (C) The PBP2 calibration immunoblot (left) and calibration curve (right). Lanes 1 through 6 are PBP concentrations as follows: 0.05 µg, 0.1 µg, 0.25 µg, 0.5 µg, 0.75 µg, and 1 µg. (D) PBP1 immunoblot with recombinant PBP1 standard (rPBP1) indicated. (E) PBP2 immunoblot with recombinant PBP2 standard (rPBP2) indicated. Each lane contains the supernatant of one homogenized antenna. Adjusted volumes were calculated for all lanes.**

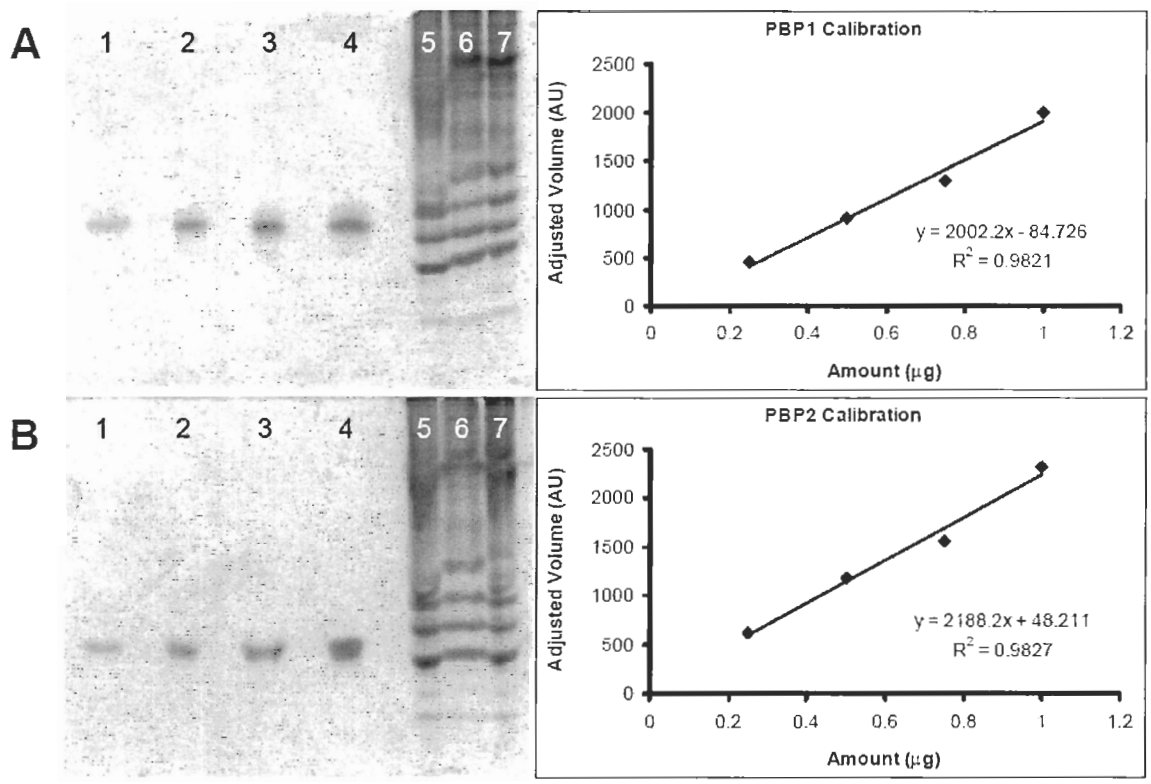



**Figure 3.10 Determination of PBP concentrations *in vivo* by Coomassie staining. (A) The calibrated 16% native PAGE gel of PBP1 (left) and calibration line (right). (B) The calibrated 16% native PAGE gel of PBP2 (left) and calibration line (right). Lanes 1 through 4 are PBP concentrations as follows: 0.25 µg, 0.5 µg, 0.75 µg, and 1 µg. Lanes 5, 6, and 7 each contain the supernatant of one homogenized antenna. Adjusted volumes were calculated for all lanes.**

## 3.5 Endogenous Fatty Acids Associate with PBPs as Vesicles

The above calculations of the fatty acid content present in whole antennal extracts, as well as fatty acids bound to PBP1 or PBP2, suggest that there is a large amount of fatty acids present in the sensillar lymph. These concentrations well exceed the CMCs of the individual acid components: 1.3 µM for palmitic acid (Hoyrup *et al.* 2001), 10 µM for linoleic acid (Lopez-Nicolas *et al.* 1995), 10-15 µM for oleic acid (Gietzen *et al.* 1982), and 393 µM for stearic acid (Chirita *et al.* 2003). Only PBP1 bound stearic acid (43.4 µM/antenna bound) was present in amounts less than its

estimated CMC. Such high concentration of acid in the lymph suggests aggregates, such as micelles or vesicles, are likely present. From extractions of the isolated PBPs, there appeared to be an excess of fatty acid bound to each PBP. Oleic and stearic acid were 0.5 and 0.2 times the amount of PBP1 while palmitic, linoleic, oleic, and stearic acid were 28, 29, 6, and 9 times in excess of PBP2, respectively. These lipid aggregates were associating with the proteins non-covalently, yet were associating strongly enough to be carried through the strong anion exchange column.

In order to determine the aggregation number of the fatty acid aggregates, DLS experiments were performed. The results from these experiments indicate that the main particles were fatty acid vesicles and higher weight aggregates, while only trace amounts of micelles were observed (Table 3.6). Figure 3.11 shows the results from DLS measurements of PBP1 and PBP1 associated fatty acids. PBP1-associated fatty acids oleic and steric acid form vesicles after 32 min in the absence of PBP1 which start at 86 nm in radius and increase to larger aggregates. After 158 min, the smaller vesicles were no longer observed and only large structures were detected. When recombinant PBP1 is present in the fatty acid mixture, we still observed the 86 nm radius vesicles after 60 min, along with higher order aggregates. Even after 178 min, these smaller vesicles could still be detected when PBP1 was present. This suggests PBP1 is capable of preventing vesicle aggregation. DLS measurements of PBP1 alone reveal a small particle with radius of 4.4 nm in size. The solved crystal structure of BmorPBP with pheromone bound reveals the BmorPBP to have dimensions of 4.0 × 3.5 × 3.0 nm (Sandler *et al.* 2000). PBP1 is about 61% homologous with BmorPBP, so our measurement suggests that PBP1 aggregates into dimeric or multimeric structures with two PBP-lengths as the diameter. This

suggests apo PBP1 forms a dimeric structure. In the presence of fatty acids, this small protein particle was absent, and a particle with a 10 nm radius was observed, suggesting PBP1 may form a tetramer or a micelle structure with PBP1 and fatty acids. It is interesting to note that higher order aggregates of PBP1 were observed even in the absence of fatty acids. However, the peak contributions from these higher order aggregates of pure PBP1 do not contribute to the higher order aggregates seen in the fatty acid samples because of their smaller amplitudes. After a 178 min incubation of fatty acids and PBP1, a particle with a radius of 24 nm appears. This may indicate fatty acid micelles, or perhaps higher order multimers of PBP1.

Figure 3.12 shows the results from DLS measurements of PBP2 and PBP2 associated fatty acids. Micelles (10 nm radius) and vesicles (86 nm radius) of PBP2 associated fatty acids palmitic, linoleic, oleic, and stearic acid formed after 61 min in the absence of PBP2. These smaller particles disappeared after 161 min to form larger vesicles (145 and 688 nm radius). When recombinant PBP2 is present in the acid mix, we observed very small particles with a radius of 1.3 nm, as well as particles that have a radius of 47 nm. Larger aggregates appeared at 59 min. After 178 min, similar-sized particles were seen (radius 59 nm). This suggests PBP2 stabilizes some smaller micelles and vesicles. Interestingly, particles with a 3 nm radius were detected with only PBP2. This may correspond to the PBP2 monomer. This monomer peak disappeared in the presence of fatty acids, yet particles (radius 47 nm and 59 nm at 59 min and 178 min, respectively) smaller than the fatty acid vesicles (radius 86 nm) were observed. Perhaps PBP2 associates with fatty acid vesicles to create smaller, more stable vesicles than those vesicles seen with only fatty acids. As with PBP1, higher order aggregates of PBP2 were

observed even in the absence of fatty acids. Again, pure PBP2 aggregates have small amplitudes, which suggests that they probably do not contribute to the higher order aggregates seen in the fatty acid samples. The largest particles ($> 1 \times 10^4$ nm) in the "protein only" samples may arise from protein sample preparation (i.e. concentration in spin columns may introduce unwanted particles such as dust).

**Table 3.6 DLS results of fatty acid vesicles associated with PBP1 and PBP2. Fatty acid samples were measured at two different time points to demonstrate their aggregative nature.**

| Sample[a] | Peak | Radius (nm) | Sample | Peak | Radius (nm) |
|---|---|---|---|---|---|
| PBP1-associated fatty acids @ 32 min | 1 | 86 | PBP2-associated fatty acids @ 61 min | 1 | 10 |
| | 2 | 380 | | 2 | 86 |
| | 3 | $4.7 \times 10^3$ | | 3 | 441 |
| | 4 | $1.9 \times 10^5$ | | 4 | $1.8 \times 10^3$ |
| PBP1-associated fatty acids @ 158 min | 1 | $1.9 \times 10^3$ | | 5 | $2.1 \times 10^{5\,b}$ |
| | 2 | $8.6 \times 10^{3\,b}$ | PBP2-associated fatty acids @ 161 min | 1 | 145 |
| | 3 | $3.8 \times 10^{5\,b}$ | | 2 | 688 |
| PBP1 | 1 | 4.4 | | 3 | $1.8 \times 10^3$ |
| | 2 | 80 | | 4 | $1.0 \times 10^{5\,b}$ |
| | 3 | $6.9 \times 10^3$ | | 5 | $7.4 \times 10^{6\,b}$ |
| | 4 | $1.6 \times 10^{6\,b}$ | PBP2 | 1 | 3.0 |
| PBP1 with fatty acids @ 60 min | 1 | 10 | | 2 | 80 |
| | 2 | 86 | | 3 | $1.9 \times 10^3$ |
| | 3 | 698 | | 4 | $2.6 \times 10^{5\,b}$ |
| | 4 | $5.9 \times 10^3$ | PBP2 with fatty acids @ 59 min | 1 | 1.3 |
| | 5 | $1.6 \times 10^{5\,b}$ | | 2 | 47 |
| PBP1 with fatty acids @ 178 min | 1 | 24 | | 3 | $1.9 \times 10^3$ |
| | 2 | 80 | | 4 | $6.4 \times 10^{4\,b}$ |
| | 3 | $1.8 \times 10^3$ | | 5 | $3.0 \times 10^{6\,b}$ |
| | 4 | $9.3 \times 10^3$ | PBP2 with fatty acids @ 178 min | 1 | 59 |
| | 5 | $1.9 \times 10^{5\,b}$ | | 2 | $1.9 \times 10^3$ |
| | | | | 3 | $4.4 \times 10^{5\,b}$ |

[a] For PBP1 the fatty acids are oleic and stearic acid, and for PBP2 the fatty acids are palmitic, linoleic, oleic, and stearic acid. Amplitudes of the peaks are shown in Figure 3.11 and Figure 3.12 for PBP1 and PBP2, respectively.

[b] The resolution of the DLS instrument allows accurate measurements for radii of $1.0 \times 10^4$ nm or below. Radii reported above this length are only estimates.

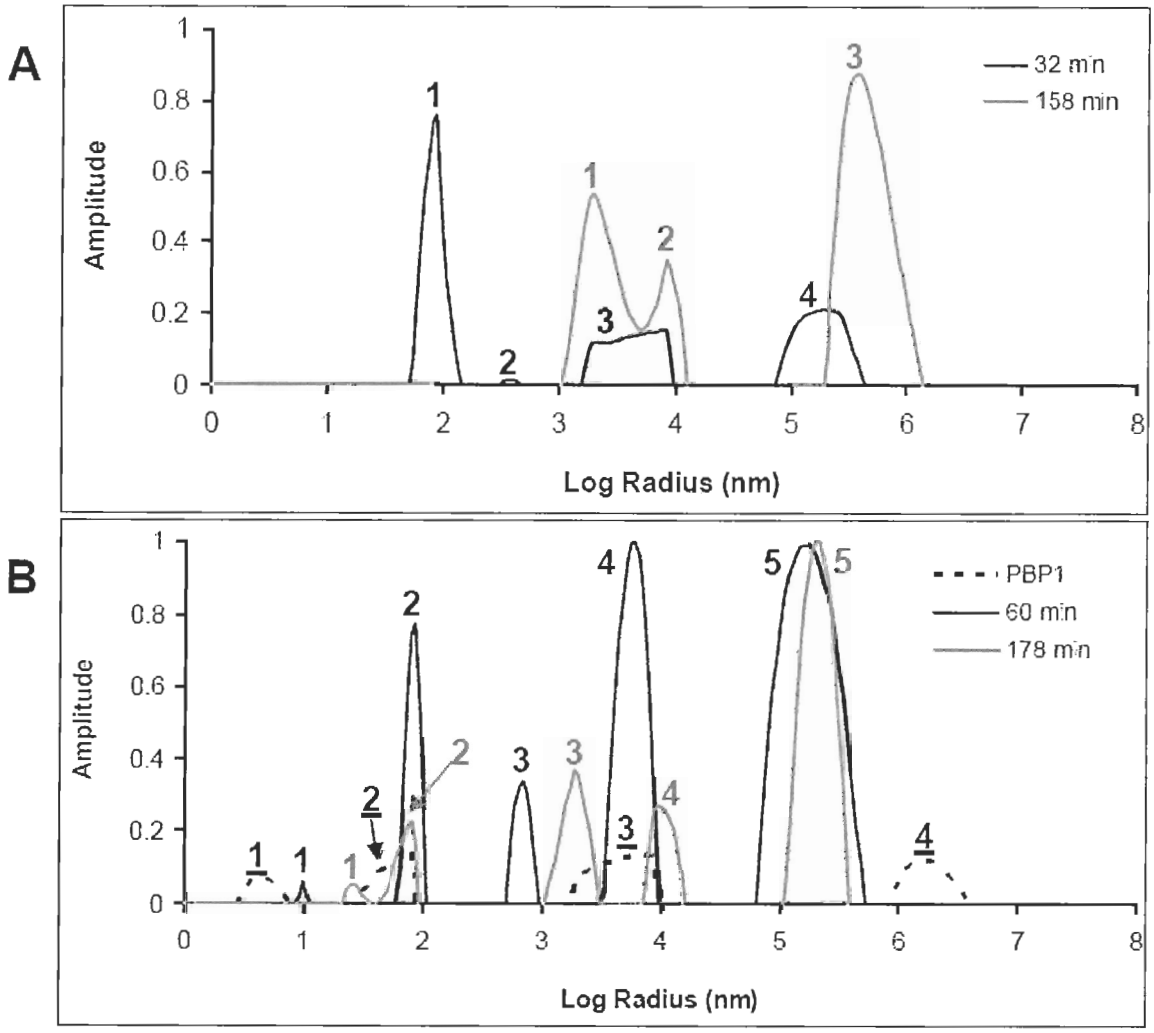



**Figure 3.11** **DLS measurements of PBP1 associated fatty acids. (A) A mixture of oleic and stearic acid prepared at a 2.7 times excess of concentrations observed associated with PBP1 *in vivo* were measured. (B) PBP1 as well as PBP1 in the presence of its associated fatty acids were measured (PBP1 was present in a 5-fold dilution of its concentration observed *in vivo*).**

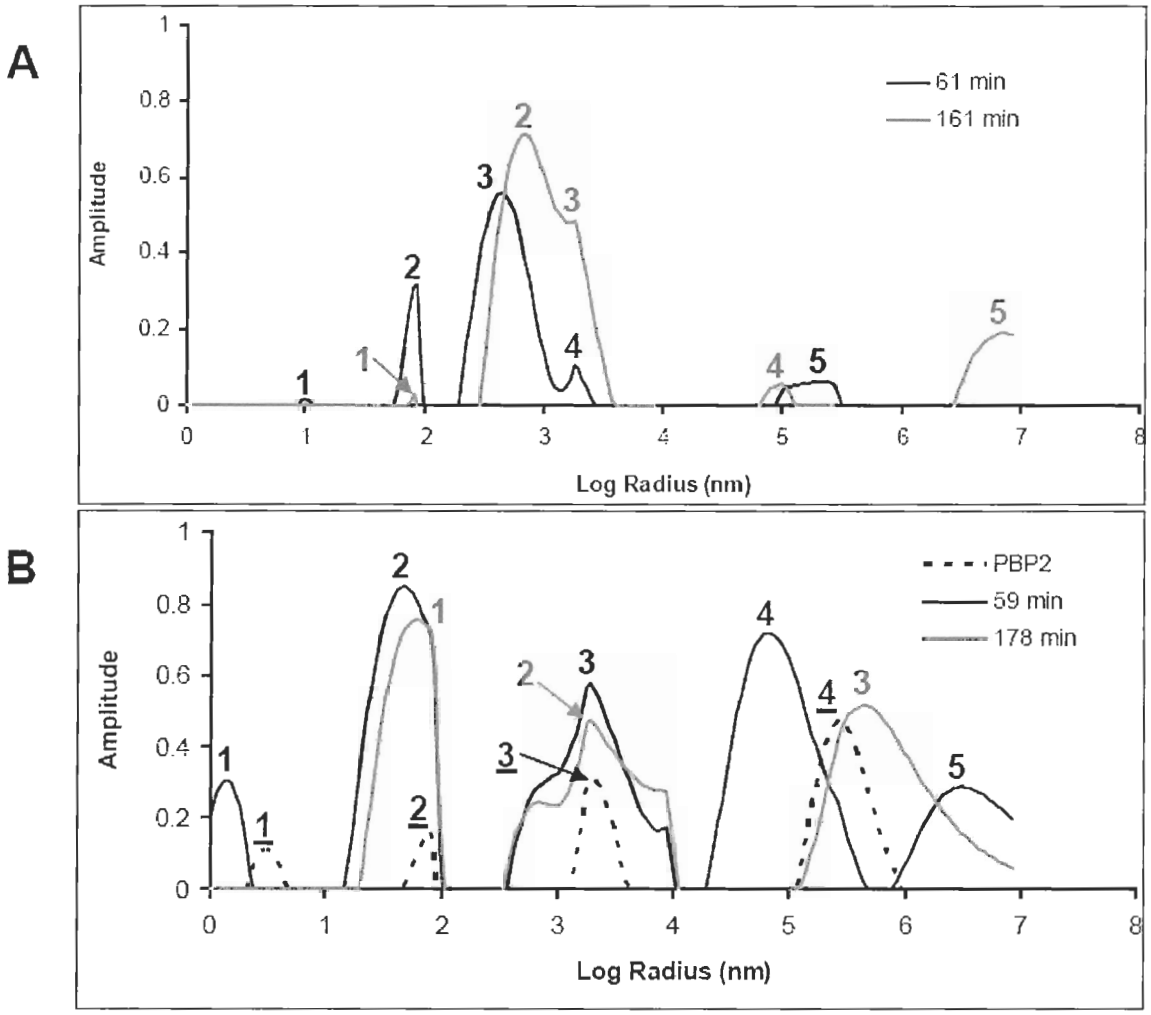



**Figure 3.12** **DLS measurements of PBP2 associated fatty acids. (A) A mixture of palmitic, linoleic, oleic and stearic acid prepared at a 370-fold dilution of the concentrations observed associated with PBP1 *in vivo* was measured. (B) PBP2 as well as PBP2 in the presence of its associated fatty acids were measured (PBP2 was present in a 5-fold dilution of its concentration observed *in vivo*).**

## 3.6 Binding Constant Determination of Endogenous Ligands Using Isothermal Titration Calorimetry

The endogenous compounds positively identified in the extract of the aqueous whole antennal homogenate were assayed to determine whether they bind PBP1, PBP2, or both, and to obtain thermodynamic parameters such as $\Delta H$, $\Delta S$, binding stoichiometry, and the dissociation constant of ligand binding (Table 3.7). Endogenous ligands must

strongly associate with the proteins in order to remain associated after being passed through a strong anion exchanger. Thermodynamic binding parameters for palmitic acid and stearic acid could not be determined because of the low solubility of these fatty acids in the ITC buffer. In addition to solubility problems, many of the compounds I tested gave very slow and noisy heat release (Figure 3.13). Thus, $K_d$ values could only be determined for a few compounds. However, enthalpy contributions could still be calculated for all those ligands which bind PBPs (even those with slow binding kinetics), by calculating the sum of the injection peak areas. Heat release or absorption are indicative of ligand-protein interaction processes. Both the dissociation constant values and the change in enthalpy are consistent with the ligand binding preferences of each PBP.

**Table 3.7 Thermodynamic binding parameters for PBP1 and PBP2.**

| Compound | $K_d$ (μM) | | stoichiometry[a] (n) | | ΔH (kcal/mol) | | ΔS (e.u.) | |
|---|---|---|---|---|---|---|---|---|
| | PBP1 | PBP2 | PBP1 | PBP2 | PBP1 | PBP2 | PBP1 | PBP2 |
| oleic acid | 2.04 | 1.41 | 1.18 | 0.82 | -2.14 | -2.94 | 18.85 | 16.90 |
| linoleic acid | nc[b] | 5.07 | nc | 1.60 | -2.50[d] | -2.85 | nc | 14.67 |
| **(+)-1** | 11.10[c] | 2.35[c] | nc | nc | 0.24[d] | -0.58[d] | 23.47[e] | 23.81[e] |
| cholesterol | nc | nc | nc | nc | 0.66[d] | -0.10[d] | nc | nc |

[a] This stoichiometry arises from the data being fitted to a non-linear regression model assuming one binding site.
[b] Abbreviation "nc" = not calculated because of very slow kinetics of heat release.
[c] $K_d$ determined using the GC binding assay (see section 4.3.4).
[d] The change in enthalpy was calculated by manual integration of peak areas (see section 2.2.7).
[e] The change in entropy was calculated using values of ΔH and $K_d$ (see section 2.2.7).

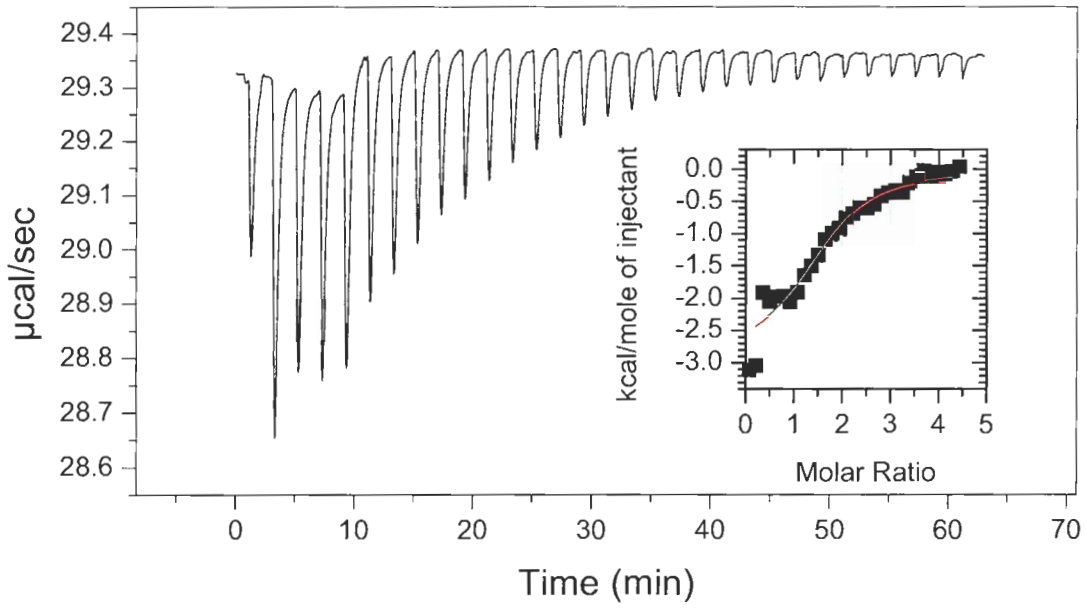



**Figure 3.13 ITC titration of linoleic acid into a PBP2 solution. Linoleic acid was titrated into a 20 µM solution of PBP2. The inset graph is the processed raw titration data (with a control titration without PBP2 subtracted) generated by the Origin software supplied by the manufacturer. The data was fitted to a non-linear regression model assuming one binding site.**

Ligand binding to PBP2 seems to be an exothermic process, while PBP1 binds compounds both exothermically and endothermically. Although both PBP1 and PBP2 bind oleic acid, PBP2 binds oleic acid more strongly as reflected by the smaller dissociation constant and larger enthalpy change, respectively (Table 3.7). These results are also reflected in the amounts seen bound in the antenna: a 6:1 oleic acid:PBP2 ratio was detected, while only a 0.5:1 oleic acid to PBP1 ratio was observed. Thus, in one sensillum containing equal amounts of PBP1 and PBP2, there may be a greater driving force for oleic acid to more strongly associate with PBP2 than with PBP1. A similar observation is made with **(+)-1**. ITC results are consistent with previous binding studies, suggesting PBP2 binds **(+)-1** more strongly than does PBP1 (Plettner *et al.* 2000). In this study, this preference manifested itself in a greater absolute heat change upon binding

(−0.58 kcal/mol) when compared with PBP1 (0.24 kcal/mol). It is also interesting to note that PBP1 binding of pheromone **(+)-1** is an endothermic process. Similarly, cholesterol binding to PBP1 is also endothermic (0.66 kcal/mol) when compared to the heat released upon binding with PBP2 (-0.10 kcal/mol). PBP1 binding to cholesterol may be an entropy driven process. The observation that only a small amount of heat is released when cholesterol is titrated into a solution with PBP2 may indicate weak binding with PBP2. This is consistent with the observations made in the FPLC separations of antennal proteins: only PBP1 bound steroids (see section 3.2).

## 3.7 Discussion

### 3.7.1 Compounds Present in the Sensillar Lymph

High concentrations of fatty acids and cholesterol were found in the sensillar lymph both when *L. dispar* male antennae were homogenized in aqueous buffer as well as in organic solvents such as ethyl acetate or DMSO. Homogenization into aqueous buffer allowed us to isolate native, functional, PBP1 and PBP2. Homogenizations into organic solvent allowed us to confirm that the fatty acids are indeed present in the sensillar lymph and were not generated post-mortem by lipases released during sonication. Cuticular extractions also ruled out the possibility that the fatty acids were from the cuticle.

It is interesting to note the relative amounts of fatty acids found from homogenized whole antennal extracts into aqueous buffer differed from homogenized whole antennal extracts into organic solvent. All fatty acids extracted into organic solvent were less than amounts extracted into aqueous buffer, with the amount of palmitic

acid being closest to the amount extracted into aqueous buffer (73%). Solubility likely plays a role in the amounts of fatty acids extracted. Palmitic acid, which is the most difficult of all the fatty acids to dissolve into aqueous buffer, would be expected to partition better into ethyl acetate than do linoleic, oleic, and stearic acid. Why more fatty acid was found in aqueous buffer extractions may lie in the fact that PBPs are capable of binding or associating with these acids, and extraction from the aqueous phase into ethyl acetate may be facilitated by PBPs (i.e. PBPs may allow fatty acids to dissociate with any membranous structures released during sonication). In either case, the concentrations of fatty acids found in the whole antennal extract are well into the millimolar range. High concentrations of fatty acids have recently been shown to come from the diet of the insect (Wang *et al.* 2006). In the butterfly *Morpho peleides*, palmitic, linoleic, oleic, and stearic acids were found in homogenized whole body extracts at the following amounts, respectively: $4.1 \pm 2.8$, $2.2 \pm 1.5$, $2.8 \pm 1.2$, $1.1 \pm 0.5$ μg/mg body weight. (Wang *et al.* 2006).

Cholesterol was also estimated to be present in high concentrations in the sensillar lymph. More cholesterol was found when antennae were homogenized directly into ethyl acetate than when antennae were homogenized into aqueous buffer. This may again be due to the increased solubility of cholesterol in ethyl acetate. Cholesterol is required for the production of insect hormones (ecdysteroids), which are important sterols that regulate a number of developmental stages. Insects cannot synthesize cholesterol *de novo*, they must ingest it from their food source (Belles *et al.* 2005).

### 3.7.2 Reanalysis of PBP Concentrations *in vivo*

The amounts of PBP1 and PBP2 found from FPLC, immunoblots, and Coomassie stained gels differed from previously established values. The amounts largely quoted are 6.7 mM for both PBP1 and PBP2 (Vogt *et al.* 1989), 27 to 28 times the amounts found by FPLC (0.25 mM and 0.24 mM for PBP1 and PBP2, respectively), and 15 to 20 times the amounts found by calibrated immunoblots (0.34 mM and 0.46 mM for PBP1 and PBP2, respectively). The quoted amounts are 3 to 4 times the amounts found by Coomassie stained gels (1.79 mM and 2.06 mM for PBP1 and PBP2, respectively). Coomassie stained gels, however, provides an over-estimate of the PBP concentration, because the staining does not distinguish between the PBPs of interest and other OBPs electrophoresing in the same region. In addition, electron micrographs showing PBPs stained with gold particles by immunocytochemical localization reveal sparse staining of PBPs, consistent with concentrations more likely within our estimates (Steinbrecht *et al.* 1992, Zhang *et al.* 2001). This lower expression of PBPs supports the results of kinetic experiments performed within our research group (Y. Gong and E. Plettner, unpublished results). This reanalysis of PBP concentrations is invaluable to researchers trying to mimic PBP concentrations *in vivo*.

### 3.7.3 PBPs Selectively Associate with Fatty Acid Vesicles

Extraction of PBP1 and PBP2 revealed that these proteins associate with fatty acids and cholesterol endogenously. The concentrations associated with each protein were again in the millimolar range. These concentrations are well above the CMC of each fatty acid, suggesting these acids are micellar in solution (see section 3.5). However, DLS revealed structures of greater than 200 nM in diameter, suggesting

vesicles are formed in aqueous solution at neutral pH. The structures observed are not micelles because micelles are much smaller. For example, oleic acid micelles have been determined to have a radius of about 1.3 nM when formed in aqueous buffer at pH 8.5 (Chen and Szostak 2004). The particles in the fatty acid suspensions, with and without PBP, were also found to aggregate over time, suggesting the vesicles may be associating to form bigger vesicles, or multilamellar stacks. Single chain amphiphiles are capable of forming bilayers (Gebicki and Hicks 1973) or vesicles when the pH of the solution approaches the $pK_a$ of the fatty acid (Hargreaves and Deamer 1978). The $pK_a$ values of palmitic, linoleic, oleic, and stearic acid in a spread monolayer have been determined to be 8.7 (Kanicky *et al.* 2000), 9.2, 9.9, and 10.2, respectively (Kanicky and Shah 2002). The $pK_a$ value of oleic acid vesicles in water has been estimated to be between 8.0-8.5 (Cistola *et al.* 1988). Short-chain carboxylic acids which remain in solution, such as acetic acid, have a $pK_a$ of about 4.8. Indeed, the $pK_a$ values of carboxylic acids increase with increasing carbon chain length. The basis for this trend is thought to be increased van der Waals interactions between chains of adjacent molecules in concentrated fatty acid solutions. This results in shielding of the hydrogen atom between oxygen atoms, leading to a higher apparent $pK_a$ of the fatty acids in micelles or vesicles (Kanicky and Shah 2002). Vesicles have also been observed to form from fatty acid micelle precursors under aqueous conditions (Chen and Szostak 2004), and formation of these vesicles have even been found to generate and maintain a pH gradient (Chen and Szostak 2004).

The concentration of fatty acids associated endogenously with PBP1 is about 56% of the total protein concentration *in vivo*. In addition, two unidentified cholesterol derivatives also bind PBP1. The amounts of these derivatives could not be determined

because no standard could be run to produce a calibration curve. However, it is possible that the 44% of PBP1 not binding fatty acids may be binding these cholesterol derivatives. This suggests PBP1 may not only associate with fatty acid vesicles, but may selectively bind only one C18 fatty acid molecule per PBP. Perhaps the C18 acid mimics **1** in the PBP binding site. The molar excess of fatty acids associated with PBP2 are in excess by 74-fold, thus PBP2 may bind some of the fatty acids within the pheromone-binding pocket, but the rest must associate with the protein elsewhere (i.e. on its surface). Oleic acid was the only fatty acid soluble enough in aqueous buffer to test its thermodynamic binding parameters with both PBP1 and PBP2, and both proteins were capable of binding this fatty acid.

DLS experiments were attempted at the concentrations of PBPs and fatty acids found *in vivo*. However, in the case of PBP2, the particles formed under those conditions far exceeded the sizes the instrument could detect. Also, the DLS required a large sample volume of protein solution. For this reason, PBP1 and its associated fatty acids were studied as a 5-fold dilution of PBP1. PBP2 and its associated fatty acids were studied at a 5-fold dilution of PBP2 and a 370-fold dilution of fatty acids. DLS experiments revealed that PBP1 may stabilize fatty acid vesicle size. However, PBP2 may destabilize fatty acid vesicle size to create significantly smaller vesicles. This suggests that PBP1 passively associates with fatty acid vesicles that have been preformed, while PBP2 interacts with preformed vesicles to alter their size. Structural characterization of related PBPs (see section 1.3.3) reveal an unstructured C-terminal tail. The homology model of PBP1 which was constructed using the solved structure of BmorPBP, contains a C-terminal tail comprised of 11 amino acid residues (DVAVGELLADT), 4 of which are

hydrophobic. Homology modeled PBP2 contains an 18-residue C-terminal tail (NWAPDVELLVADFLAESQ), 8 of which are hydrophobic. The hydrophobic C-terminal tail may be capable of insertion into fatty acid vesicles. This is more likely for PBP2, which is predicted to have a tail of similar length to the fatty acids present in the sensillar lymph. For example, a protein contributing to the formation of amyloid fibrils in neurodegenerative diseases is capable of forming a disordered structure consisting of 15 protein units with a hydrophobic C-terminal tail (Mukhopadhyay *et al.* 2006). In addition, PBP2 was found to associate with all four fatty acids, while PBP1 only associated with the C18 fatty acids. This again supports the idea that PBP2 associates with fatty acid vesicles, while PBP1 binds fatty acids within its pheromone-binding pocket. The specific ratios of fatty acids found associated with PBP2 may be due to the nature of vesicle formation *in vivo*, and not due to the selectivity of the protein.

### 3.7.4 Possible Roles of Endogenous Ligands and Vesicles, and a New Role for PBPs

Binding of ligands to OBPs have been found to induce conformational changes in the protein (Lee *et al.* 2002, Zhou *et al.* 2004). Endogenous ligands likely also induce conformational changes, thus one role of endogenous ligands such as the fatty acids may be to hold the protein in an appropriate conformation for odorant reception. Fatty acids as ligands may also help induce selectivity towards the proteins as an odorant entering the lymph must be a strong enough binder to displace the endogenous molecule. The existence of vesicles or multilamellar structures in the lymph may also be a mechanism of odorant transport. Instead of PBPs partitioning hydrophobic odorants into the aqueous lymph, vesicles may sequester and transport odorants to the receptor. Alternatively,

odorants may pass through channels within a complex multilamellar network towards the receptor. PBPs themselves associate with vesicles, perhaps *via* insertion and van der Waals interactions of their hydrophobic tails with the carbon chains of these fatty acids. They may still be capable of binding odorant when associated with vesicles.

The presence of fatty acids in the sensillar lymph may also provide defense against bacteria and pathogens that may enter the antenna through the many pores covering its surface. Bacterium have sizes ranging from 0.02 to 400 $\mu m^3$ (Koch 1996), while the diameter of a pore tubule located in *s. trichodea* of *A. polyphemus* is about 0.03 μm (Keil 1999). Thus, it is possible that a bacterium may enter the lymph through a pore tubule. However, bacteria of this size are usually dead. It is possible, however, that bacteria can enter the lymph directly if the hair is damaged. Fatty acids are known antibacterial and antiviral agents, although their mechanism of action is relatively unknown. Recently, linoleic and oleic acid have been found to inhibit the bacterial enoyl-acyl carrier protein reductase (FabI), which is essential for bacterial fatty acid synthesis, while stearic acid (the saturated form of linoleic acid) did not exhibit any FabI inhibition (Zheng *et al.* 2005). Free fatty acids such as palmitic, linoleic, oleic, and stearic, at concentrations $\geq$ 75 μM, 250 μM, 225 μM, and 150 μM, respectively, have recently been shown to be toxic to a macrophage cell line (J774) (Lima *et al.* 2006). The fatty acid concentrations determined in the lymph are about 600-5000 times in excess of these concentrations. How such high concentrations of free fatty acids in the lymph are not toxic to the neuron remains a question. Perhaps incorporation of these fatty acids into bilayers reduces their toxicity to the moth and the neuronal cell itself. Or perhaps one of

the highly abundant olfaction-specific proteins of unknown function (such as SNMP) provides protection to the neuron against free fatty acid toxicity.

PBPs may not even be necessary for odorant transport, but may have an entirely different function. Instead of odorant transport, they may play a role in preventing the desensitization of odorant receptors by clearing away odorants already bound to receptors, thereby preventing prolonged hyperpolarization of the nerve cell. Such a function would explain why PBPs are necessary to elicit a successful olfactory response, even for water-soluble compounds like ethanol, as was demonstrated with *D. melanogaster* mutants which lacked the OBP LUSH (see section 1.6) (Xu *et al.* 2005).

### 3.7.5 Future Work

Cholesterol and two cholesterol derivatives were found present in lymph extracts. The cholesterol derivatives seem to contain a large number of hydroxyl groups as derivatization with BSTFA resulted in a compound too massive to detect, and detection of underivatized sample is inconsistent. We hope to identify the cholesterol derivatives by direct injection of the extract as well as the standard into the MS ion trap. In this way we hope to see a sharper, more consistent peak.

The presence of endogenous fatty acids changes the way the sensillar lymph is perceived. Instead of being an aqueous solution, it appears to be a soapy solution consisting of fatty acid vesicles. To explore whether **1** partitions into fatty acid vesicles, a ligand transport experiment can be performed. In such an experiment, a mock *s. trichodea* would be created in a glass vial. The system would contain 4 different components. The first would mimic the dendritic membrane by creating phospholipid

vesicles containing sucrose (to allow this layer to sink and stay at the bottom of the vial) and a dye such as bromophenol blue. The second would mimic the sensillar lymph and contain an aqueous solution with or without PBPs and fatty acids. The third component would be a thin wax layer to represent the cuticle. Finally, a cotton plug containing **1** could be placed in the lid of the vial (not in contact with the wax layer). The pH and salt composition of the lymph component and the phospholipid vesicle component would be adjusted to mimic *in vivo* conditions. The system would be left to equilibriate for 1-3 days, and each component could then be extracted with organic solvent to yield the amount of **1** present in each layer. From this experiment we can determine the location of **1,** as well as whether PBPs, fatty acid vesicles, or both, are capable of transporting **1** to the phospholipid vesicle layer.

# CHAPTER 4: STRUCTURAL CHARACTERIZATION OF *L. DISPAR* PBP1 AND PBP2

PBPs are known to be highly α-helical proteins containing three disulfide bonds. These disulfide linkages are thought to contribute to the folding and stability of these proteins. This second experimental chapter describes the determination of the disulfide bond linkages of *L. dispar* PBP1 and PBP2. Two different peptide mapping methods were employed. The first uses TCEP to partially reduce the protein, cyanylation with CDAP, and cleavage under basic conditions followed by reduction of the remaining two disulfide bonds using a higher concentration of TCEP. The second method uses CNBr to chemically cleave the protein. Mapping using the cyanylation procedure resulted in our discovery that the C2-C5 disulfide bond is the most easily reduced disulfide linkage. To further investigate the role of this disulfide bridge, this bond was chemically modified by cyanylation of PBP1 (PBP1.CN) to lock the protein in this singly reduced form. The biochemical properties of this modified protein were then explored, first using far-UV circular dichroism to obtain secondary structure information, and second using a GC-binding assay to determine the pheromone-binding affinities of PBP1.CN.

## 4.1 Peptide Mapping of *L. dispar* PBP1 and PBP2

The results of this section describing the peptide mapping of *L. dispar* PBP1 and PBP2 have been published in *Naturwissenschaften* (Honson and Plettner 2006).

### 4.1.1 Peptide Mapping by Partial Reduction and Cyanylation Reactions

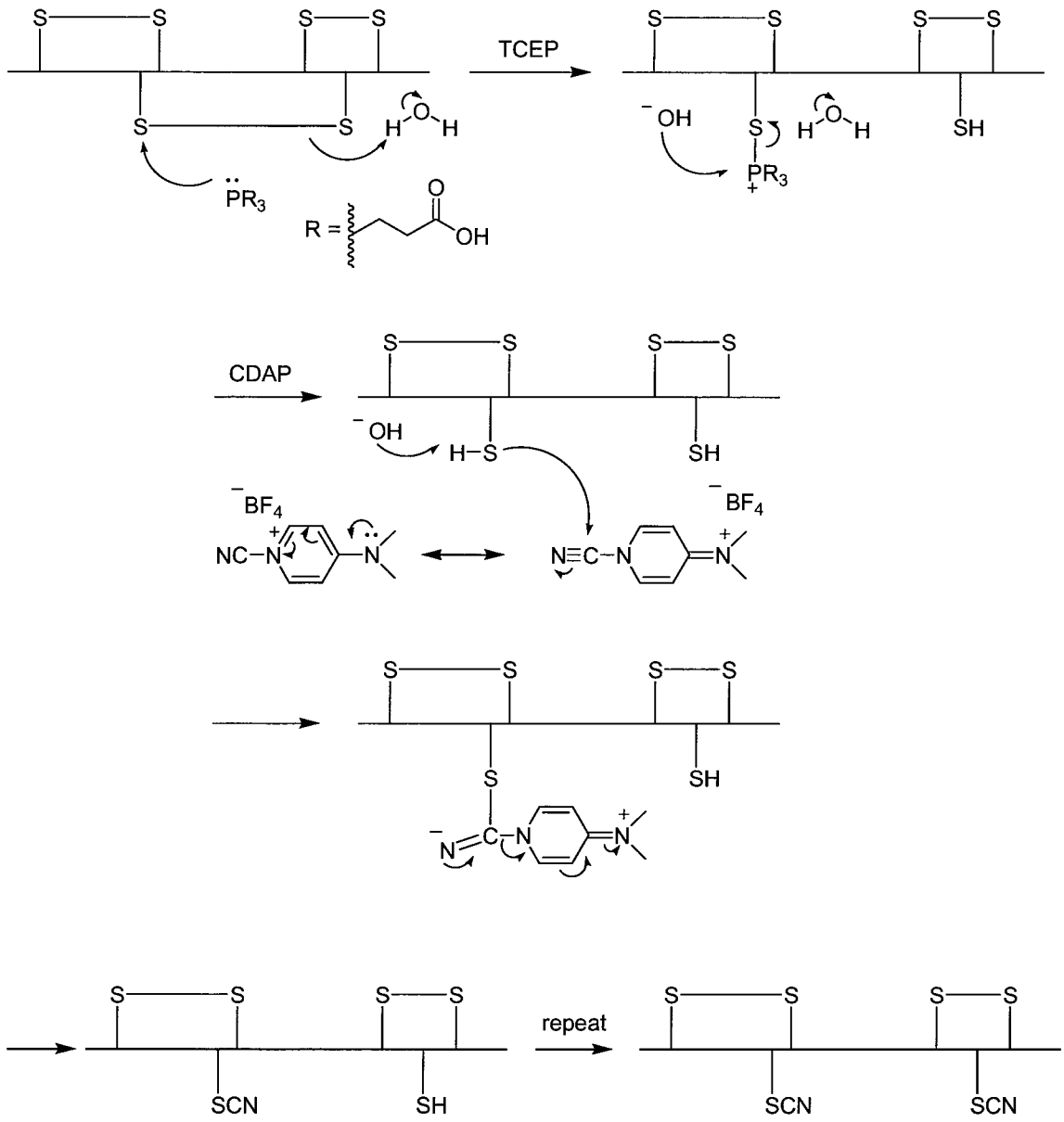
All fragments resulting from the cyanylation reaction are listed in Table 4.1. Partial reduction and cyanylation of both PBP1 and PBP2 led to only one cyanylated product (Scheme 4.1). Two protein peaks were present in the HPLC trace, those of the intact protein and the cyanylated product as identified by MALDI-TOF MS (Table 4.1). There are no free cysteines in PBP1 or PBP2 under non-reducing conditions (Plettner *et al.* 2000). Under reducing conditions, with three-fold TCEP over protein, only one disulfide bond has been reduced after incubation for 30 min at room temperature. We are able to detect this single reduction by cyanylation, followed by MS analysis. Other reductants, namely 2-mercaptoethanol and DTT, gave the same result. The molecular weight of the singly reduced PBP1 isomer was calculated from the oxidized native PBP1 weight of 16139.8 Da with the mass of two hydrogen atoms added to give the average molecular weight ($M_{ave}$) of 16141.8 Da. The observed weight by MALDI-TOF MS, however, was 16207.1 Da. The difference in weight is due to addition of one $Na^+$ and one $K^+$ ion. Association of monovalent cations is common with peptides and proteins analyzed using MALDI-TOF MS (Leite *et al.* 2004). The cyanylated product also contains one $Na^+$ and one $K^+$ ion (Table 4.1).

**Table 4.1 Singly reduced and cyanylated isomers of PBP1 and PBP2 produced via cyanylation cleavage reaction. Table is adapted from reference (Honson and Plettner 2006).**

| PBP | Reaction | $M_{ave}$ (Da) expected | $M_{ave}$ (Da) observed (by MALDI-MS) | Error (%) | Peptide location | Peptide sequence |
|---|---|---|---|---|---|---|
| PBP1 | reduction | 16202.9[a] | 16207.1 | 0.02 | n/a | n/a |
| PBP1 | cyanylation | 16254.9[b] | 16268.6 | 0.08 | n/a | n/a |
| PBP1 | $NH_4OH$ cleavage | 3887.9 | 3889.5 | 0.04 | 109-143 | ITC-QTMLNLAMCFKAEIHKLD WAPTLDVAVGELLADT |

| PBP | Reaction | $M_{ave}$ (Da) expected | $M_{ave}$ (Da) observed (by MALDI-MS) | Error (%) | Peptide location | Peptide sequence |
|---|---|---|---|---|---|---|
| PBP1 | $NH_4OH$ cleavage | 5745.0 | 5739.0 | 0.10 | 1-49 | SKEVMKQMTINFAKPMEA CKQELNVPDAVMQDFFNF WKEGYQITNREAG |
| PBP1 | $NH_4OH$ cleavage | 6600.1 | 6611.6 | 0.17 | 50-108 | ITC-VILCLAKKLELLDQDMNLH |
| PBP1 | $NH_4OH$ cleavage | 6600.1 | 6578.4[c] | 0.33 | 50-108 | HGKAMEFAMKHGADEAM AKQLLDIKHSCEKVITIVAD DP |
| PBP2 | reduction | 16143.6 | 16144.6 | 0.01 | n/a | n/a |
| PBP2 | cyanylation | 16195.7 | 16194.3 | 0.01 | n/a | n/a |
| PBP2 | $NH_4OH$ cleavage | 4080.1 | 4070.8 | 0.23 | 110-145 | ITC-QRAVNVAMCFKAHVHKLN WAPDVELLVADFLAESQ |
| PBP2 | $NH_4OH$ cleavage | 5662.9 | 5654.3 | 0.15 | 1-49 | SKDVMHQMALKFGKPIKL CQQELGADDSVVKEFLDF WKDGYVMKDRQTG |
| PBP2 | $NH_4OH$ cleavage | 6699.0[d] | 6705.6[e] | 0.10 | 50-109 | ITC-MLICMAMKLELLDSAMEIH HGSTFAFAKAHGADEAMA QQIIDIVHGCTTTYPAAETN DP |

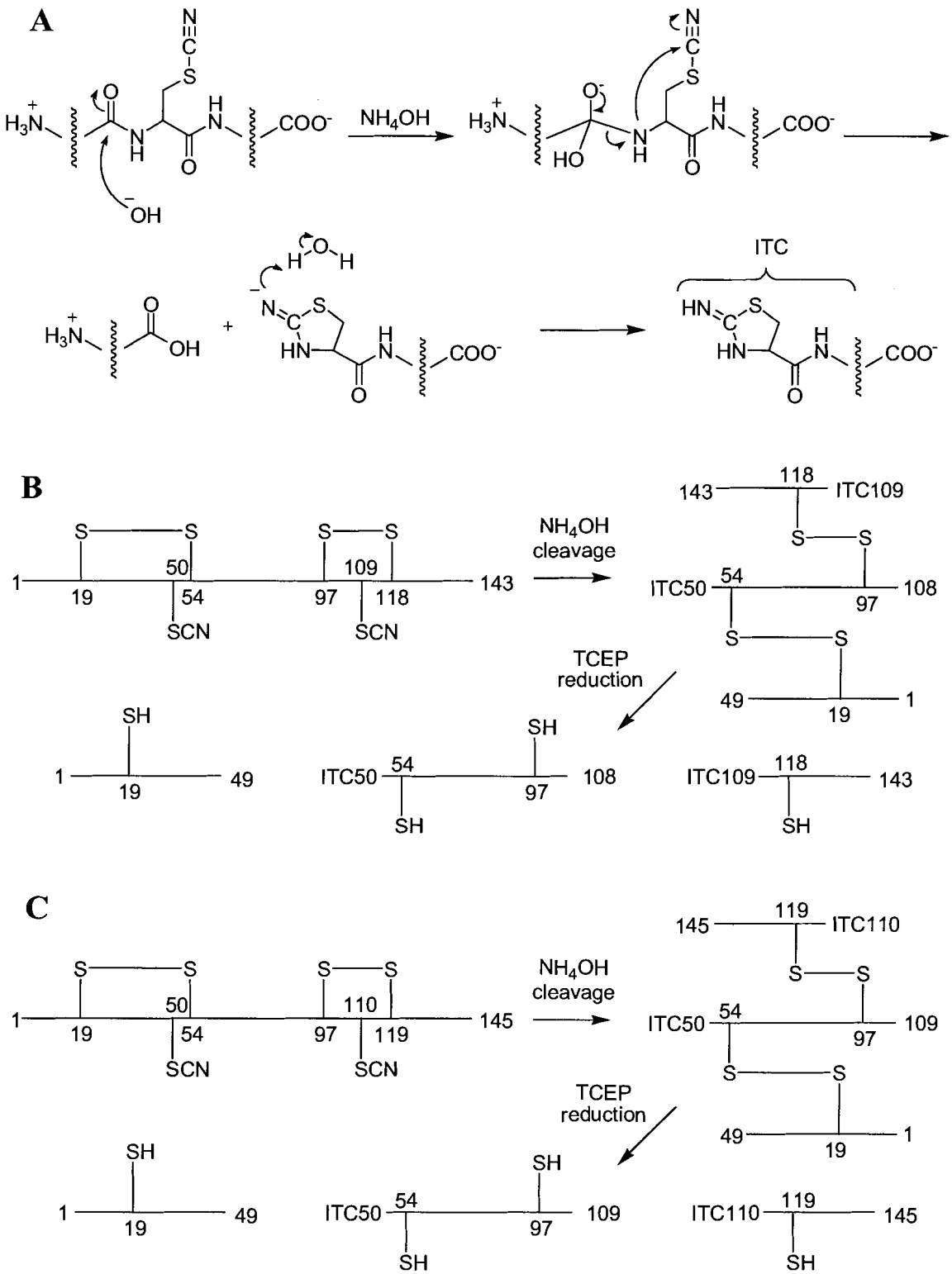
[a] Calculated from $M_{ave}$ of singly reduced PBP1 (16141.8 Da) with $M_{ave}$ of $Na^+$ and $K^+$ added, and $H^+$ subtracted (since singly charged ions are observed).
[b] Calculated from $M_{ave}$ of PBP1 (16141.8 Da) with $M_{ave}$ of $Na^+$, $K^+$ and two CN groups added, and $H^+$ subtracted (since singly charged ions are observed).
[c] Calculated from addition of peaks appearing at 2882.3 Da and 3696.1 Da.
[d] Calculated as the sinapinic acid adduct product (6491.8 Da with addition of sinapinic adduct with $M_{ave}$ of 207.2 Da) via attack of the sinapinic acid carbonyl by a side-chain amine present within the fragment.
[e] Calculated from addition of peaks appearing at 3084.3 Da and 3621.3 Da.




**Scheme 4.1 Reduction and cyanylation reaction mechanism of *L. dispar* PBPs.**

Cleavage of the doubly cyanylated product with $NH_4OH$ led to the appearance of an isothiocyanate (ITC) residue at the N-terminus of each fragment (except for the first residue of the protein) (Scheme 4.2). Subsequent total reduction employing 0.1 M TCEP and MALDI-TOF MS analysis of the solution revealed five peaks for the reaction with PBP1, and four peaks for the reaction with PBP2 (Table 4.1). From the resulting

fragments, the reduced disulfide was found to be Cys50-Cys109 and Cys50-Cys110 for PBP1 and PBP2, respectively (Scheme 4.2), and can be explained as follows. For PBP1, the observed peaks at 6611.6 Da and 3889.5 Da corresponds to the ITC residue fragments ITC50-108 (expected mass of 6600.1) and ITC109-143 (expected mass of 3887.9), respectively, while fragment 1-49 (expected mass of 5745.0 Da) had an observed mass of 5739.0 Da. Two other peaks present on the MALDI spectra were located at 2882.3 Da and 3696.1 Da. Addition of these two peaks gave a mass of 6578.4 Da, suggesting part of the 50-108 fragment had cleaved during exposure to the MALDI laser. Basic cleavage of PBP2 resulted in two $NH_4OH$ cleavage products at 4070.8 Da (expected mass of 4080.1 Da) and 5654.3 Da (expected mass of 5662.9 Da) which corresponds to fragments ITC110-145 and 1-49, respectively. Fragment ITC50-109 has an expected mass of 6491.8 Da, however, no peak of this mass was detected. Peaks located at 3084.3 Da and 3621.3 Da suggests a possible cleavage of ITC50-109 into two smaller fragments which can result during exposure to the MALDI laser (Hoffmann and Stroobant 2001). Addition of these peaks yields an observed mass of 6705.6 Da which is the peptide fragment ITC50-109 (expected mass of ITC50-109 is 6491.8 Da) detected as the sinapinic adduct formed *via* attack of the sinapinic acid carbonyl by a side-chain amine present in the fragment (expected mass of ITC50-109 as the sinapinic adduct is 6699.0 Da). Matrix adduct formation with peptides and proteins has long been observed with MALDI (Loboda and Chernushevich 2005).




**Scheme 4.2** **Fragments of *L. dispar* PBPs produced after cyanylation and subsequent cleavage under basic conditions. (A) Mechanism of base cleavage. (B) PBP1 cleavage products. (C) PBP2 cleavage products. Scheme is adapted from reference (Honson and Plettner 2006).**

### 4.1.2 Peptide Mapping by CNBr Reactions

All fragments resulting from the CNBr reactions are listed in Table 4.2 and have been identified as peaks "a" to "i" for both PBP1 and PBP2. Cleavage of intact proteins was accomplished with CNBr reactions using partially reduced proteins when reacted with excess reagent for 24 h in the dark at room temperature. An initial MALDI-TOF MS of the cleavage reaction mixture taken did not show all the fragments as resolved peaks. Peptides were identified upon separation and analysis by LC-ESI-MS (Table 4.2, Figure 4.1).

**Table 4.2 CNBr chemical cleavage peptide fragments of partially reduced PBP1 and PBP2. Table is adapted from reference (Honson and Plettner 2006).**

| PBP | Peak label | Product | $M_{monoiso}$ (Da) expected | $M_{monoiso}$ (Da) observed (by ESI-MS) | Error (%) | Peptide location | Peptide sequence |
|---|---|---|---|---|---|---|---|
| PBP1 | a | $[M+H]^+$ | 545.30 | 546.50 | 0.22 | 1-5 | SKEVM |
| | b | $[M+H]^+$ | 873.48 | 872.94 | 0.06 | 9-16 | TINFAKPM |
| | c | $[M+2H]^{2+}$ | 1699.38 | 1699.52 | 0.01 | 113-143 | LNLAMCFKAEIHKLD WAPTLDVAVGELLA DT |
| | c' | oxid. Met $[M+H]^+$ | 3002.50 | 3004.83 | 0.08 | 117-143 | MCFKAEIHKLDWAPT LDVAVGELLADT |
| | c'' | formylated c' | 3030.49 | 3032.31 | 0.06 | 117-143 | |
| | c''' | formylated c'' | 3058.49 | 3059.54 | 0.03 | 117-143 | |
| | | | | 3060.69 | 0.07 | | |
| | d | $[M+H]^+$ | 3858.96 | 3860.19 | 0.03 | 34-66 | FNFWKEGYQITNRE AGCVILCLAKKLELLD QDM |
| | d' | $[M+H]^+$ | 2932.53 | 2932.29 | 0.01 | 41-66 | YQITNREAGCVILCL |
| | d'' | formylation of d' | 2960.52 | 2960.23 | 0.01 | 41-66 | AKKLELLDQDM |
| | e | $[M+3H]^{3+}$ | 1571.11 | 1574.58 | 0.26 | 75-117 | EFAMKHGADEAMAK |
| | e' | 2 × formylated $[M+3H]^{3+}$ | 1589.77 | 1592.66 | 0.18 | 75-117 | QLLDIKHSCEKVITIV ADDPCQTMLNLAM |
| | e'' | 3 × formylated $[M+3H]^{3+}$ | 1599.10 | 1602.62 | 0.22 | 75-117 | |
| | f | $[M+H]^+$ | 2473.31 | 2475.19 | 0.08 | 53-74 | LCLAKKLELLDQDMN |
| | f' | homoseryl of f | 2491.32 | 2493.79 | 0.10 | 53-74 | LHHGKAM |
| | f'' | cysteic acid of f | 2521.30 | 2521.75 | 0.02 | 53-74 | |
| | g | $[M+2H]^{2+}$ | 2800.38 | 2801.84 | 0.05 | 67-117 | NLHHGKAMEFAMKH GADEAMAKQLLDIKH SCEKVITIVADDPCQ TMLNLAM |

| PBP | Peak label | Product | $M_{monoiso}$ (Da) expected | observed (by ESI-MS) | Error (%) | Peptide location | Peptide sequence |
|---|---|---|---|---|---|---|---|
| | h | $[M+2H]^{2+}$ | 2852.46 | 2851.11 | 0.05 | (87-112) + (118-143)[a] | AKQLLDIKHSCEKVIT IVADDPCQTM + CFKAEIHKLDWAPTL DVAVGELLADT |
| | i | $[M+2H]^{2+}$ | 2878.90 | 2879.10 | 0.01 | 17-66[a] | EACKQELNVPDAVM QDFFNFWKEGYQIT NREAGCVILCLAKKL ELLDQDM |
| PBP2 | a | $[M+H]^{+}$ | 971.54 | 971.11 | 0.04 | 58-66 | KLELLDSAM |
| | a' | formylated $[M+H]^{+}$ | 999.54 | 999.26 | 0.03 | 58-66 | |
| | b | $[M+H]^{+}$ | 1173.62 | 1171.46 | 0.18 | 56-66 | AMKLELLDSAM |
| | b' | homoseryl $[M+H]^{+}$ | 1191.63 | 1189.48 | 0.18 | 56-66 | |
| | c | $[M+H]^{+}$ | 2078.97 | 2079.48 | 0.02 | 67-86 | EIHHGSTFAFAKAHG |
| | c' | $[M+2H]^{2+}$ | 1040.49 | 1039.13 | 0.13 | 67-86 | ADEAM |
| | c'' | formylated $[M+H]^{+}$ | 2106.96 | 2107.52 | 0.03 | 67-86 | |
| | c''' | homoseryl of c'' | 2124.97 | 2125.52 | 0.03 | 67-86 | |
| | d | $[M+H]^{+}$ | 3079.49 | 3078.85 | 0.02 | 58-86 | KLELLDSAMEIHHGS |
| | d' | homoseryl $[M+H]^{+}$ | 3097.50 | 3096.93 | 0.02 | 58-86 | TFAFAKAHGADEAM |
| | e | $[M+H]^{+}$ | 3382.62 | 3381.84 | 0.02 | 87-118 | AQQIIDIVHGCTTTYP AAETNDPCQRAVNV AM |
| | f | $[M+H]^{+}$ | 4398.25 | 4400.68 | 0.06 | (9-43) + (52-55)[a] | ALKFGKPIKLCQQEL GADDSVVKEFLDFW KDGYVM + LICM |
| | g | $[M+2H]^{2+}$ | 2464.24 | 2466.00 | 0.07 | 1-43 | SKDVMHQMALKFGK PIKLCQQELGADDSV VKEFLDFWKDGYVM |
| | h | formylated $[M+2H]^{2+}$ | 3591.70 | 3592.44 | 0.02 | 52-118 | LICMAMKLELLDSAM EIHHGSTFAFAKAHG |
| | h' | h – 3×$H_2O$ $[M+2H]^{2+}$ | 3564.18 | 3564.37 | 0.06 | 52-118 | ADEAMAQQIIDIVHG CTTTYPAAETNDPC QRAVNVAM |
| | h'' | $[M+2H]^{2+}$ | 3534.65 | 3535.80 | 0.03 | 53-118 | ICMAMKLELLDSAME IHHGSTFAFAKAHGA DEAMAQQIIDIVHGC TTTYPAAETNDPCQ RAVNVAM |
| | i | $[M+3H]^{3+}$ | 3267.91 | 3268.95 3268.99 | 0.03 | 56-145[a] | AMKLELLDSAMEIHH GSTFAFAKAHGADE AMAQQIIDIVHGCTTT YPAAETNDPCQRAV NVAMCFKAHVHKLN WAPDVELLVADFLAE SQ |

[a] Fragments which contain disulfide bond linkages.

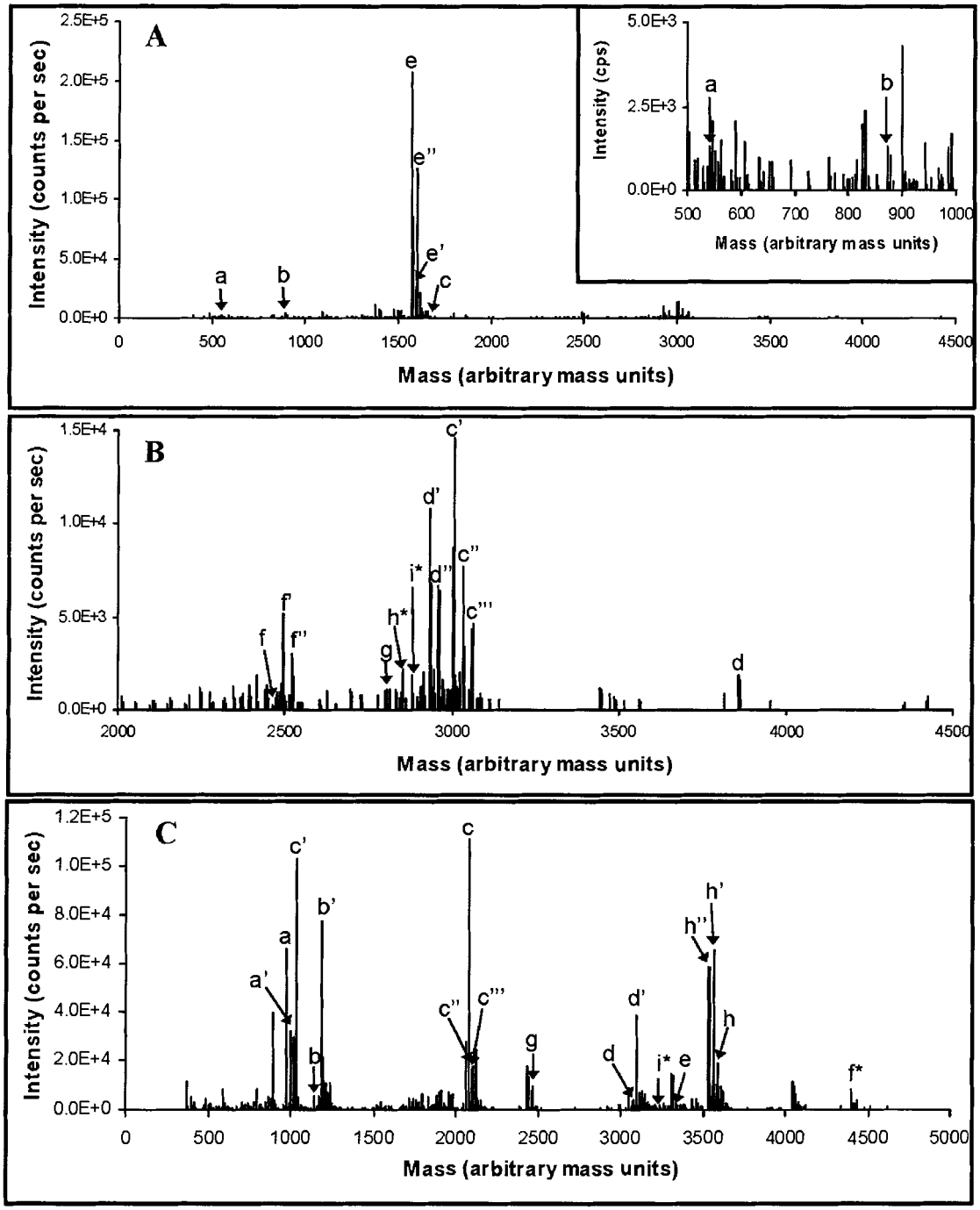



**Figure 4.1** **Mass spectra of *L. dispar* PBPs produced from LC-ESI-MS peptide separation and analysis. (A) PBP1 mass spectra from 0 to 4500 amu. The inset shows peaks from 500 to 1000 amu. (B) PBP1 mass spectrum revealing peaks in the region of 2000-4500 amu. (C) PBP2 mass spectrum from 0 to 5000 amu. Fragments containing disulfide linkages are indicated by '*'. This figure is from reference (Honson and Plettner 2006), by permission of the journal ©Naturwissenschaften.**

Elution of PBP1 peptide fragments by chromatographic separation began after 35 min and was stopped after 100 min (Appendix 2). Masses in the range of 400-4500 Da were determined by mass reconstruction (Figure 4.1 A and B). Many of the larger fragments were observed as $[M+2H]^{2+}$ or $[M+3H]^{3+}$ ions. From these fragments the following two disulfide bridges were mapped: Cys19-Cys54 and Cys97-Cys118. The most easily reduced disulfide as determined from the above cyanylation reactions, Cys50-Cys109, did not withstand the strong reducing conditions of 2-mercaptoethanol. The major peaks shown at 1574.48, 1592.66, and 1602.62 Da were much larger than all other peaks, with the largest peak at 1574.48 Da being seven times greater than the next largest peak cluster (see below) (Figure 4.1 A). The disulfide linkage Cys97-Cys118 was detected within fragment residues 87-112 and 118-143 at 2851.11 Da as determined from the $[M+2H]^{2+}$ ion (1.5% of the base peak, expected mass of 2852.46 Da) (Figure 4.1 B). The disulfide bond Cys19-Cys54 was located on the fragment containing residues 17-66. This fragment was observed at 2879.10 Da, as determined from the $[M+2H]^{2+}$ ion (1.5% of the base peak, calculated mass of 2878.90 Da). This fragment contains one missed cleavage site at Met30. These two peaks were very small compared to the base peak, but detection of fragmentation peptides from these two original peptides supports their assignment (Appendix 3). Assignment of all other fragments can be found within Appendix 4

Elution of PBP2 peptide fragments obtained by CNBr cleavage *via* chromatographic separation began after 23 min and was stopped after 80 min (Appendix 2). Masses in the range of 370-4650 Da were determined by mass reconstruction (Figure 4.1 C). The following two disulfide bridges were mapped: Cys19-Cys54 and Cys97-

Cys119. Again, the most easily reduced disulfide as determined from the above cyanylation reactions, Cys50-Cys110, did not withstand the reducing conditions of 2-mercaptoethanol. The base peak was determined to be the $[M+H]^+$ peak from the fragment containing residues 67-86. The Cys19-Cys54 disulfide linkage was found within the fragment containing residues 9-43 and 52-55 as the $[M+H]^+$ ion. The peak appears at 4400.68 Da (9.8% of the base peak, expected mass of 4398.25 Da). The Cys97-Cys119 disulfide linkage was found within residues 56-145 and includes four missed cleavages at methionines 57, 66, 86 and 118. The peak observed is the $[M+3H]^{3+}$ ion peak at 3268.95 Da and 3268.99 Da (1.9% and 3.3% of the base peak, expected mass of 3267.91 Da). Intramolecular peptide cleavage from both the N and C-terminus of the disulfide bond containing fragments were observed, confirming these assignments (Appendix 3). Assignment of all other fragments can be found within Appendix 4.

## 4.2 Examining the C2-C5 Disulfide Bond

### 4.2.1 Torsion Angle Comparisons with Known PBP Structures

To further explore reasons why the interlocking disulfide bridge (Cys50-Cys109 in PBP1 and Cys50-Cys110 in PBP2) is the most easily reduced disulfide bond of the three present and what its role might be, we measured the C2-C5 disulfide bond angles of all PBPs which have published structures (Table 4.3). The first angle measured was $\tau$, the torsion angle between the $C_\beta$ groups of the cysteine residues along the S-S bond measured with the lower numbered cysteine residue at the front. The measured torsion angles reveal that the cysteines which comprise the C2-C5 exist only in certain conformational ranges (Figure 4.2, Figure 4.3 B). BmorPBP seems to span three different conformational regions: -84.39° to -38.27°, 44.76° to 56.03°, and 94.50° to

110.41° (Figure 4.2). PDB identification code 1DQE (Sandler *et al.* 2000) is the solved crystal structure and thus only two torsion angles were available (one from each subunit of the dimer). One molecule of **7** binds each subunit. PDB identification codes 1GM0 (Horst *et al.* 2001) and 1LS8 (Lee *et al.* 2002) are NMR structures and represent the maximum and minimum torsion angle values of each conformational range from the twenty available solutions. LmaPBP (crystallized as a dimer) only exhibits one τ region: 101.14° to 111.85°. All three structures were solved using X-ray diffraction, and each protein is bound to a small molecule: 1ORG to glycerol, 1OW4 to ANS, and 1P28 to racemic 3-hydroxy-butan-2-one (Lartigue *et al.* 2003). The NMR structure of ApolPBP reveals two disulfide conformational regions, −151.10°-144.80° and 129.78°-163.13° (PDB identification code 1QWV(Mohanty *et al.* 2004)). Finally, AmelASP1 (this protein is a PBP, see section 1.3.3) is shown to bind *n*-butyl-benzene-sulfonamide in the crystal structure (PDB identification code 1R5R (Lartigue *et al.* 2004)) and only one τ angle of 101.91° (Figure 4.2) could be measured from the monomer. It is interesting to note that proteins binding a ligand exhibit the same disulfide conformational range of approximately 100°. This may, however, be a result of the crystal packing since only proteins which where solved by X-ray diffraction show one τ region.

The torsion angles of the first and third S-S bonds in insect PBPs (C1-C3 and C4-C6) as well as the two disulfide bridges in the hemolymph protein TmolTHP12 (C1-C2 and C3-C4) were also measured to see if the C2-C5 torsion angles of PBPs are unique (Figure 4.3 A, C). The first and third disulfide bonds and the two disulfides within TmolTHP12 exhibit torsion angles within the frequently seen ranges, approximately 100° and -80°.

The S-S bond length of the C2-C5 bond was also measured to see if there was any correlation with τ. The disulfide bond lengths ($b_1$) were similar among the above PBPs, with values ranging from 2.02 Å to 2.16 Å (Table 4.3). Finally, the distance between the $C_\beta$ groups of each cysteine was measured ($d_1$). The larger the absolute value of τ, the larger the distance. ApolPBP had the largest distance between cysteines, where $d_1$ >4.3 Å in all 20 NMR structures. The distances $d_2$, $d_3$, and $d_4$ (Table 4.3) are discussed in section 4.2.2. The distance $d_5$ (Table 4.3) is discussed in section 4.2.3.

**Table 4.3 Disulfide bond properties and distances between helix 3 and 6. Table is adapted from reference (Honson and Plettner 2006).**

| Protein[a] | Ligand | PDB ID | Structure no. | τ (°) | $b_1$ (Å) | $d_1$ (Å) | $d_2$ (Å) | $d_3$ (Å) | $d_4$ (Å) | $d_5$ (Å) |
|---|---|---|---|---|---|---|---|---|---|---|
| BmorPBP | **7** | 1DQE | A | 108.75 | 2.11 | 4.14 | 11.24 | 13.90 | 12.94 | 4.53 |
| | | | B | 110.41 | 2.11 | 4.18 | 11.32 | 14.28 | 13.12 | 4.21 |
| | | | **Avg.**[b] | | **2.11** | **4.16** | **11.28** | **14.09** | **13.03** | **4.37** |
| | | | **S.E.**[c] | | **0.00** | **0.02** | **0.04** | **0.19** | **0.09** | **0.16** |
| BmorPBP | n/a | 1GM0 | 1 | 105.32 | 2.03 | 4.02 | 11.32 | 14.46 | 13.24 | 3.82 |
| | | | 2 | 98.26 | 2.03 | 3.91 | 11.20 | 14.07 | 12.98 | 3.46 |
| | | | 3 | 101.70 | 2.03 | 3.91 | 11.18 | 14.47 | 13.39 | 4.09 |
| | | | 4 | 104.68 | 2.03 | 4.00 | 11.17 | 14.33 | 13.50 | 3.64 |
| | | | 5 | 44.76 | 2.03 | 3.40 | 11.19 | 14.41 | 13.31 | 3.98 |
| | | | 6 | 102.03 | 2.04 | 4.01 | 11.21 | 14.36 | 13.04 | 3.88 |
| | | | 7 | -47.05 | 2.02 | 3.42 | 11.06 | 14.40 | 13.17 | 3.20 |
| | | | 8 | -82.45 | 2.04 | 3.78 | 11.21 | 14.28 | 13.12 | 3.86 |
| | | | 9 | 109.32 | 2.05 | 4.10 | 11.06 | 13.93 | 12.84 | 3.88 |
| | | | 10 | 102.56 | 2.04 | 4.00 | 10.89 | 14.25 | 13.01 | 3.98 |
| | | | 11 | -78.84 | 2.04 | 3.74 | 11.15 | 14.49 | 13.19 | 4.60 |
| | | | 12 | -70.83 | 2.04 | 3.73 | 10.98 | 14.26 | 13.00 | 3.80 |
| | | | 13 | -79.57 | 2.04 | 3.76 | 10.96 | 14.01 | 12.89 | 3.65 |
| | | | 14 | -80.51 | 2.03 | 3.75 | 11.00 | 14.40 | 13.06 | 3.89 |
| | | | 15 | 51.65 | 2.03 | 3.45 | 11.13 | 14.55 | 13.38 | 3.40 |
| | | | 16 | -38.27 | 2.04 | 3.44 | 11.18 | 14.32 | 13.19 | 3.38 |
| | | | 17 | 105.62 | 2.04 | 4.01 | 10.94 | 14.24 | 13.18 | 3.98 |
| | | | 18 | -61.98 | 2.02 | 3.58 | 11.22 | 14.32 | 13.12 | 3.17 |
| | | | 19 | -82.06 | 2.05 | 3.86 | 11.03 | 14.27 | 13.16 | 4.83 |
| | | | 20 | 101.84 | 2.03 | 3.93 | 11.31 | 14.06 | 13.12 | 3.84 |
| | | | **Avg.** | | **2.04** | **3.79** | **11.12** | **14.29** | **13.14** | **3.82** |
| | | | **S.E.** | | **0.00** | **0.05** | **0.03** | **0.04** | **0.04** | **0.09** |
| BmorPBP | n/a | 1LS8 | 1 | 105.93 | 2.03 | 4.03 | 11.58 | 14.71 | 13.76 | 3.23 |
| | | | 2 | 96.79 | 2.04 | 4.01 | 12.24 | 14.58 | 13.35 | 3.29 |
| | | | 3 | -78.85 | 2.03 | 3.80 | 11.41 | 14.63 | 13.42 | 3.32 |
| | | | 4 | 94.50 | 2.04 | 3.95 | 12.06 | 14.44 | 13.67 | 3.60 |
| | | | 5 | -76.77 | 2.03 | 3.65 | 11.64 | 15.17 | 13.64 | 3.46 |

| Protein[a] | Ligand | PDB ID | Structure no. | τ (°) | $b_1$ (Å) | $d_1$ (Å) | $d_2$ (Å) | $d_3$ (Å) | $d_4$ (Å) | $d_5$ (Å) |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | 6 | -78.74 | 2.04 | 3.80 | 11.77 | 14.43 | 13.25 | 3.54 |
| | | | 7 | -76.47 | 2.04 | 3.79 | 11.59 | 14.57 | 13.32 | 3.53 |
| | | | 8 | 102.74 | 2.03 | 3.95 | 12.04 | 14.70 | 13.31 | 3.84 |
| | | | 9 | 102.33 | 2.05 | 4.06 | 11.79 | 14.71 | 13.76 | 3.34 |
| | | | 10 | 102.42 | 2.04 | 4.05 | 11.88 | 14.87 | 13.49 | 3.75 |
| | | | 11 | 107.05 | 2.05 | 4.16 | 12.22 | 14.23 | 13.31 | 3.50 |
| | | | 12 | -83.07 | 2.04 | 3.97 | 11.78 | 14.48 | 13.22 | 3.48 |
| | | | 13 | -78.23 | 2.03 | 3.80 | 11.27 | 15.12 | 13.67 | 3.16 |
| | | | 14 | -81.88 | 2.04 | 3.87 | 11.54 | 14.60 | 13.30 | 3.36 |
| | | | 15 | 56.03 | 2.03 | 3.46 | 11.83 | 14.98 | 13.57 | 3.67 |
| | | | 16 | 101.72 | 2.05 | 4.18 | 12.25 | 14.35 | 13.38 | 3.25 |
| | | | 17 | -75.77 | 2.03 | 4.17 | 11.03 | 14.80 | 13.34 | 3.28 |
| | | | 18 | -76.66 | 2.03 | 3.66 | 11.36 | 14.76 | 13.53 | 3.33 |
| | | | 19 | -84.39 | 2.03 | 3.83 | 11.13 | 14.57 | 13.17 | 3.78 |
| | | | 20 | 100.57 | 2.04 | 4.04 | 11.90 | 14.90 | 13.48 | 3.91 |
| | | | **Avg.** | | **2.04** | **3.91** | **11.72** | **14.68** | **13.45** | **3.48** |
| | | | **S.E.** | | **0.00** | **0.04** | **0.08** | **0.05** | **0.04** | **0.05** |
| LmaPBP | gly-cerol | 1ORG | A | 109.79 | 2.09 | 4.22 | 12.86 | 14.09 | 14.01 | 3.97 |
| | | | B | 101.14 | 2.07 | 4.16 | 13.41 | 14.22 | 14.11 | 3.70 |
| | | | **Avg.** | | **2.08** | **4.19** | **13.14** | **14.16** | **14.06** | **3.84** |
| | | | **S.E.** | | **0.01** | **0.03** | **0.28** | **0.07** | **0.05** | **0.14** |
| LmaPBP | ANS | 1OW4 | A | 109.13 | 2.08 | 4.31 | 13.38 | 14.09 | 14.10 | 4.04 |
| | | | B | 109.80 | 2.06 | 4.23 | 13.38 | 14.19 | 14.09 | 3.82 |
| | | | **Avg.** | | **2.07** | **4.27** | **13.38** | **14.14** | **14.10** | **3.93** |
| | | | **S.E.** | | **0.01** | **0.04** | **0.00** | **0.05** | **0.00** | **0.11** |
| LmaPBP | H3B2 | 1P28 | A | 109.26 | 2.05 | 4.07 | 13.38 | 14.17 | 14.10 | 3.45 |
| | | | B | 111.85 | 2.03 | 4.15 | 13.22 | 14.03 | 13.93 | 3.94 |
| | | | **Avg.** | | **2.04** | **4.11** | **13.30** | **14.10** | **14.02** | **3.70** |
| | | | **S.E.** | | **0.01** | **0.04** | **0.08** | **0.07** | **0.08** | **0.24** |
| ApoIPBP | n/a | 1QWV | 1 | 140.21 | 2.03 | 4.42 | 11.93 | 21.06 | 16.25 | 5.73 |
| | | | 2 | 139.87 | 2.03 | 4.44 | 11.68 | 20.61 | 16.56 | 7.90 |
| | | | 3 | 152.06 | 2.03 | 4.53 | 12.02 | 21.64 | 16.41 | 5.63 |
| | | | 4 | 163.13 | 2.03 | 4.54 | 12.54 | 21.17 | 16.30 | 5.82 |
| | | | 5 | 141.43 | 2.03 | 4.48 | 12.34 | 21.73 | 16.79 | 5.37 |
| | | | 6 | 137.39 | 2.03 | 4.41 | 11.71 | 21.77 | 16.51 | 7.67 |
| | | | 7 | 142.67 | 2.04 | 4.47 | 12.02 | 21.01 | 16.55 | 7.88 |
| | | | 8 | 153.40 | 2.03 | 4.53 | 12.50 | 20.95 | 15.66 | 5.73 |
| | | | 9 | 144.52 | 2.03 | 4.47 | 12.32 | 21.94 | 17.05 | 6.05 |
| | | | 10 | 134.02 | 2.03 | 4.40 | 11.88 | 21.24 | 16.58 | 5.50 |
| | | | 11 | 143.03 | 2.03 | 4.45 | 12.08 | 21.95 | 17.49 | 8.34 |
| | | | 12 | 140.32 | 2.04 | 4.47 | 11.48 | 22.17 | 16.51 | 5.62 |
| | | | 13 | 149.81 | 2.03 | 4.50 | 12.69 | 21.59 | 16.31 | 6.27 |
| | | | 14 | 129.78 | 2.03 | 4.34 | 11.88 | 21.38 | 16.02 | 6.13 |
| | | | 15 | 147.80 | 2.03 | 4.49 | 11.67 | 20.86 | 17.00 | 5.09 |
| | | | 16 | -144.80 | 2.03 | 4.46 | 10.77 | 22.10 | 16.14 | 6.87 |
| | | | 17 | -151.10 | 2.03 | 4.50 | 11.18 | 21.75 | 16.46 | 6.02 |
| | | | 18 | 132.64 | 2.03 | 4.36 | 11.55 | 20.82 | 16.67 | 6.15 |
| | | | 19 | 133.40 | 2.03 | 4.43 | 12.13 | 20.87 | 16.85 | 6.21 |
| | | | 20 | 132.76 | 2.03 | 4.38 | 12.44 | 22.08 | 16.51 | 9.03 |
| | | | **Avg.** | | **2.03** | **4.45** | **11.94** | **21.43** | **16.53** | **6.45** |
| | | | **S.E.** | | **0.00** | **0.01** | **0.11** | **0.11** | **0.09** | **0.25** |
| Amel-ASP1 | NBBS | 1R5R | A | 101.91 | 2.16 | 4.16 | 13.93 | 14.23 | 13.64 | 4.28 |

| Protein[a] | Ligand | PDB ID | Structure no. | τ (°) | $b_1$ (Å) | $d_1$ (Å) | $d_2$ (Å) | $d_3$ (Å) | $d_4$ (Å) | $d_5$ (Å) |
|---|---|---|---|---|---|---|---|---|---|---|
| Tmol-THP12 | n/a | 1C3Y | 1 | n/a | n/a | n/a | 13.98 | 11.24 | 12.80 | n/a |
| | | | 2 | n/a | n/a | n/a | 14.03 | 11.73 | 12.99 | n/a |
| | | | 3 | n/a | n/a | n/a | 13.50 | 10.91 | 11.57 | n/a |
| | | | 4 | n/a | n/a | n/a | 13.28 | 11.92 | 12.15 | n/a |
| | | | 5 | n/a | n/a | n/a | 13.49 | 10.99 | 12.03 | n/a |
| | | | 6 | n/a | n/a | n/a | 13.59 | 11.83 | 13.00 | n/a |
| | | | 7 | n/a | n/a | n/a | 13.99 | 11.43 | 12.41 | n/a |
| | | | 8 | n/a | n/a | n/a | 13.65 | 11.66 | 13.02 | n/a |
| | | | 9 | n/a | n/a | n/a | 13.81 | 11.01 | 12.54 | n/a |
| | | | 10 | n/a | n/a | n/a | 13.34 | 11.94 | 12.75 | n/a |
| | | | 11 | n/a | n/a | n/a | 13.58 | 11.64 | 13.09 | n/a |
| | | | 12 | n/a | n/a | n/a | 13.47 | 11.90 | 12.19 | n/a |
| | | | 13 | n/a | n/a | n/a | 13.26 | 12.09 | 12.41 | n/a |
| | | | 14 | n/a | n/a | n/a | 13.66 | 11.45 | 12.49 | n/a |
| | | | 15 | n/a | n/a | n/a | 13.93 | 11.43 | 12.43 | n/a |
| | | | 16 | n/a | n/a | n/a | 13.55 | 11.61 | 11.88 | n/a |
| | | | 17 | n/a | n/a | n/a | 13.82 | 11.42 | 12.19 | n/a |
| | | | 18 | n/a | n/a | n/a | 14.19 | 11.28 | 13.13 | n/a |
| | | | 19 | n/a | n/a | n/a | 13.84 | 11.13 | 11.74 | n/a |
| | | | 20 | n/a | n/a | n/a | 13.94 | 11.56 | 13.01 | n/a |
| | | | 21 | n/a | n/a | n/a | 13.90 | 11.70 | 12.70 | n/a |
| | | | 22 | n/a | n/a | n/a | 13.58 | 11.80 | 12.49 | n/a |
| | | | 23 | n/a | n/a | n/a | 13.68 | 11.82 | 12.65 | n/a |
| | | | **Avg.** | | **n/a** | **n/a** | **13.70** | **11.54** | **12.51** | **n/a** |
| | | | **S.E.** | | **n/a** | **n/a** | **0.05** | **0.07** | **0.09** | **n/a** |

[a] Column headings are as follows: τ is the angle between the $C_\beta$ groups of the cysteine residues down the S-S bond measured with the lower numbered cysteine residue at the front, $b_1$ is the S-S bond distance, $d_1$ is the distance between the $C_\beta$ groups of the cysteine residues (1DQE(Sandler *et al.* 2000), 1GM0(Horst *et al.* 2001), 1LS8 (Lee *et al.* 2002): C50-C108; 1ORG, 1OW4, 1P28 (Lartigue *et al.* 2003): C43-C104; 1QWV (Mohanty *et al.* 2004): C50-C108; 1R5R (Lartigue *et al.* 2004): C47-C98), $d_2$, $d_3$, and $d_4$ are the $C_\alpha$ distances between pairs of residues on helices 3 and 6 (1DQE, 1GM0, 1LS8: E47-W110, L55-V114, A51-V114; 1ORG, 1OW4, 1P28: T40-L106, A48-F110, L44-F110; 1QWV: L47-K110, L55-V114, A51-V114; 1R5R: S44-K100, L52-L104, Y48-L104; 1C3Y (Rothemund *et al.* 1999): K38-S95, I46-F99, H42-F99, respectively), and $d_5$ is the distance between the sulfur atom of the fifth cysteine residue and the oxygen of the highly conserved aspartic acid residue two amino acids upstream (1DQE, 1GM0, 1LS8: D106-C108; 1ORG, 1OW4, 1P28: D102-C104, 1QWV: D106-C108; 1R5R: D96-C98).

[b] Average values of distances $b_1$, $d_1$, $d_2$, $d_3$, $d_4$, and $d_5$ are provided when possible.

[c] S.E. is the ± standard error of distances $b_1$, $d_1$, $d_2$, $d_3$, $d_4$, and $d_5$, calculated from all the available structures analyzed.

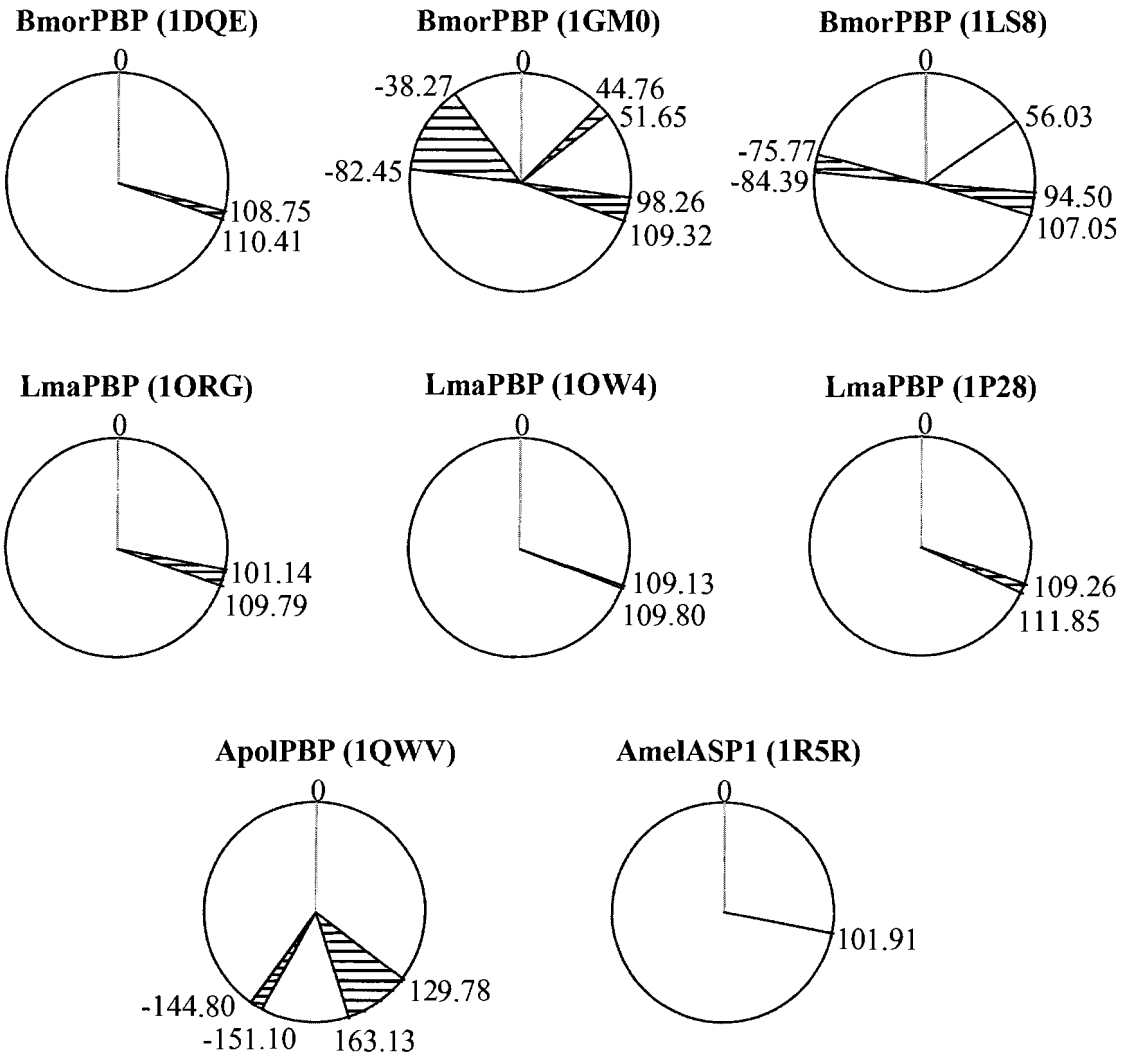



**Figure 4.2** **C2-C5 disulfide bond torsion angle ranges of insect PBPs. The corresponding PDB identification codes are shown in parenthesis. Torsion angle ranges were measured for solved crystal structures of BmorPBP (1DQE) (Sandler *et al.* 2000), LmaPBP (1ORG), (1OW4), and (1P28) (Lartigue *et al.* 2003), and AmelASP1 (1R5R) (Lartigue *et al.* 2004). All crystal structures were co-crystallized with a ligand bound. Conformational ranges for solved NMR structures are BmorPBP (1GM0) (Horst *et al.* 2001) and (1LS8) (Lee *et al.* 2002), and ApolPBP (1QWV) (Mohanty *et al.* 2004). All NMR structures were solved as apo-proteins. BmorPBP (1GM0), however, has a structured C-terminus tail occupying the binding cavity.**

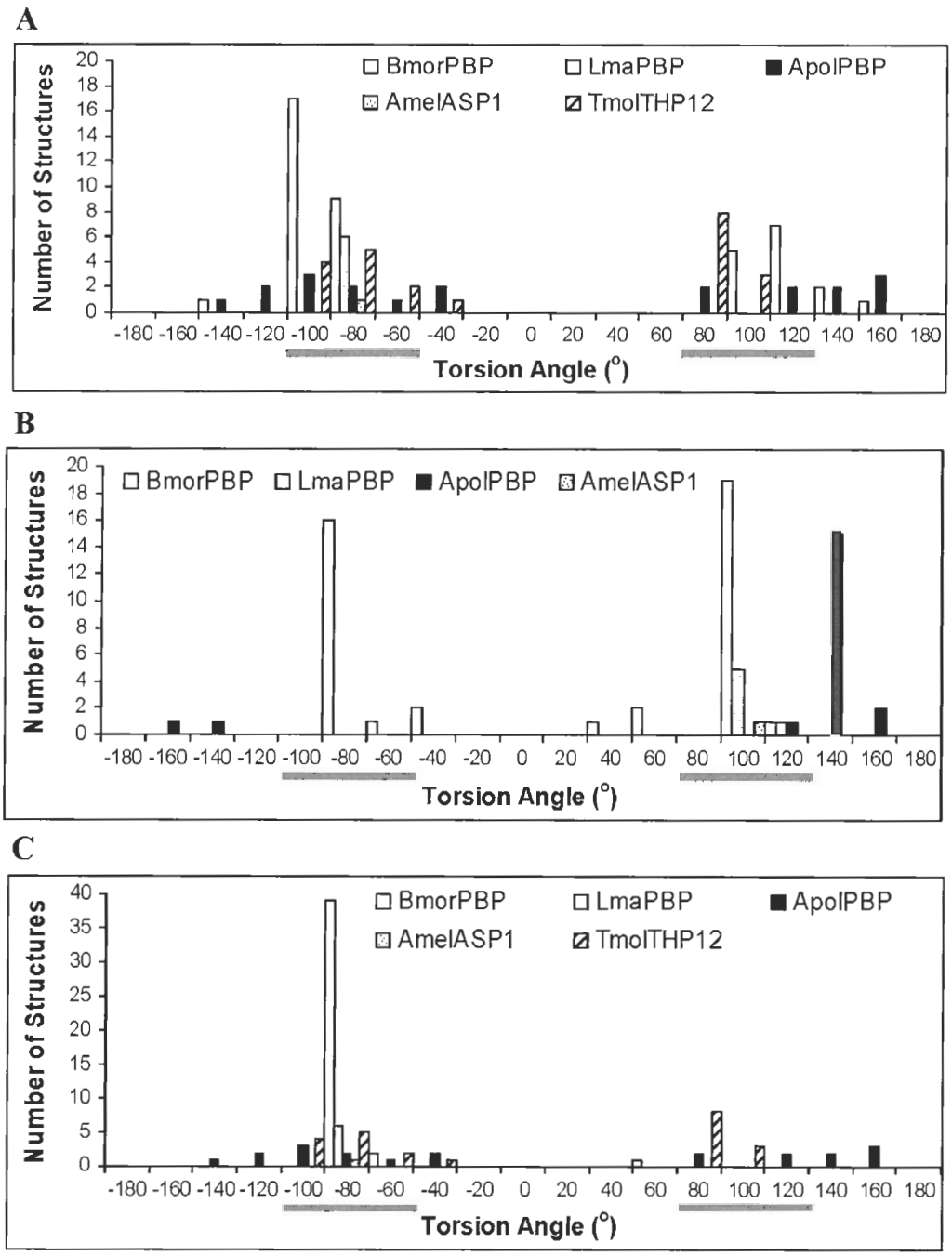



**Figure 4.3** **Histograms illustrating the disulfide bond torsion angle ranges of insect PBPs and the hemolymph protein TmolTHP12. (A) C1-C3 torsion angle ranges observed for BmorPBP, LmaPBP, ApolPBP, AmelASP1 and TmolTHP12. (B) C2-C5 torsion angle ranges (same data was used to generate Figure 4.2) observed for BmorPBP, LmaPBP, ApolPBP, and AmelASP1. Note most τ ranges for ApolPBP are greater than the commonly observed ranges. (C) C4-C6 torsion angle ranges observed for BmorPBP, LmaPBP, ApolPBP, AmelASP1 and TmolTHP12. Torsion angles underlined in grey indicate ranges observed for 351 disulfide bridges among 131 monomeric protein structures (Petersen *et al.* 1999). Torsion angle ranges were measured for solved crystal structures as described in Figure 4.2.**

### 4.2.2 Distances Between Helix 3 and Helix 6

We also investigated the orientation of the two helices connected by the most easily reducible disulfide linkage. As a comparative control, we included measurements of a hemolymph protein from the mealworm beetle *T. molitor*. This 12 kDa hydrophilic hemolymph protein (TmolTHP12, PDB accession number 1C3Y (Rothemund *et al.* 1999)) has four highly conserved cysteine residues consisting of six α-helices and possesses sequence homology with odorant and pheromone-binding proteins (Rothemund *et al.* 1997, Rothemund *et al.* 1999, Graham *et al.* 2001). TmolTHP12 is capable of binding small hydrophobic ligands within a hydrophobic groove (Rothemund *et al.* 1999). Sequence alignments of TmolTHP12 with BmorPBP, LdisPBP1, and LdisPBP2 show the two disulfide bridges of TmolTHP12 corresponding to the first and third bridge of the PBPs (Figure 4.4 A). The most easily reduced disulfide bridge of *L. dispar* PBPs connects helices 3 and 6 and is missing in TmolTHP12. We measured the intermolecular distances between helices 3 and 6 within PBPs that have solved structures using pairs of amino acids (one residue on each helix). Three distances were determined from three pairs of amino acids: $d_2$ (residues closest to the N-terminus of the α-helix), $d_3$ and $d_4$ (residues closest to the C-terminus) (Table 4.3). The measurements reflect the orientation of helices 3 and 6 for PBPs being closest at the N-terminus of each helix (where the most easily reduced disulfide resides) and widening towards the C-terminus of the helix. The average values of $d_2$, $d_3$ and $d_4$ provide information about the relative orientation of helices 3 and 6. The standard errors of the averages provide an indication of the degree of flexibility. For example, in ApolPBP, the distance $d_2$ is significantly larger than in BmorPBP. The greater standard error for ApolPBP also suggests greater mobility than in

BmorPBP. Distances were similar between all PBPs except for ApolPBP which displayed $d_3$ distances approximately 40% larger and $d_4$ distances about 20% larger than the other PBPs. TmolTHP12, however, had a larger inter-helix distance near the N-terminus ($d_2$ is about 20% larger than PBPs excluding ApolPBP), and a slightly smaller distance towards the C-terminus ($d_3$ is about 20% smaller and $d_4$ is approximately 10% smaller than PBPs excluding ApolPBP). The varying orientation of helices 3 and 6 can be observed in a superimposition of the PBPs and TmolTHP12 (Figure 4.5). The exceptionally large displacement of ApolPBP helix 6 suggests a correlation between the disulfide bond angle $\tau$ and the orientation of the helices which it interconnects. TmolTHP12 lacks this interlocking disulfide bridge, with helix 6 oriented close to helix 3, while helix 6 of BmorPBP, AmelASP1 and LmaPBP lie in an intermediate position.

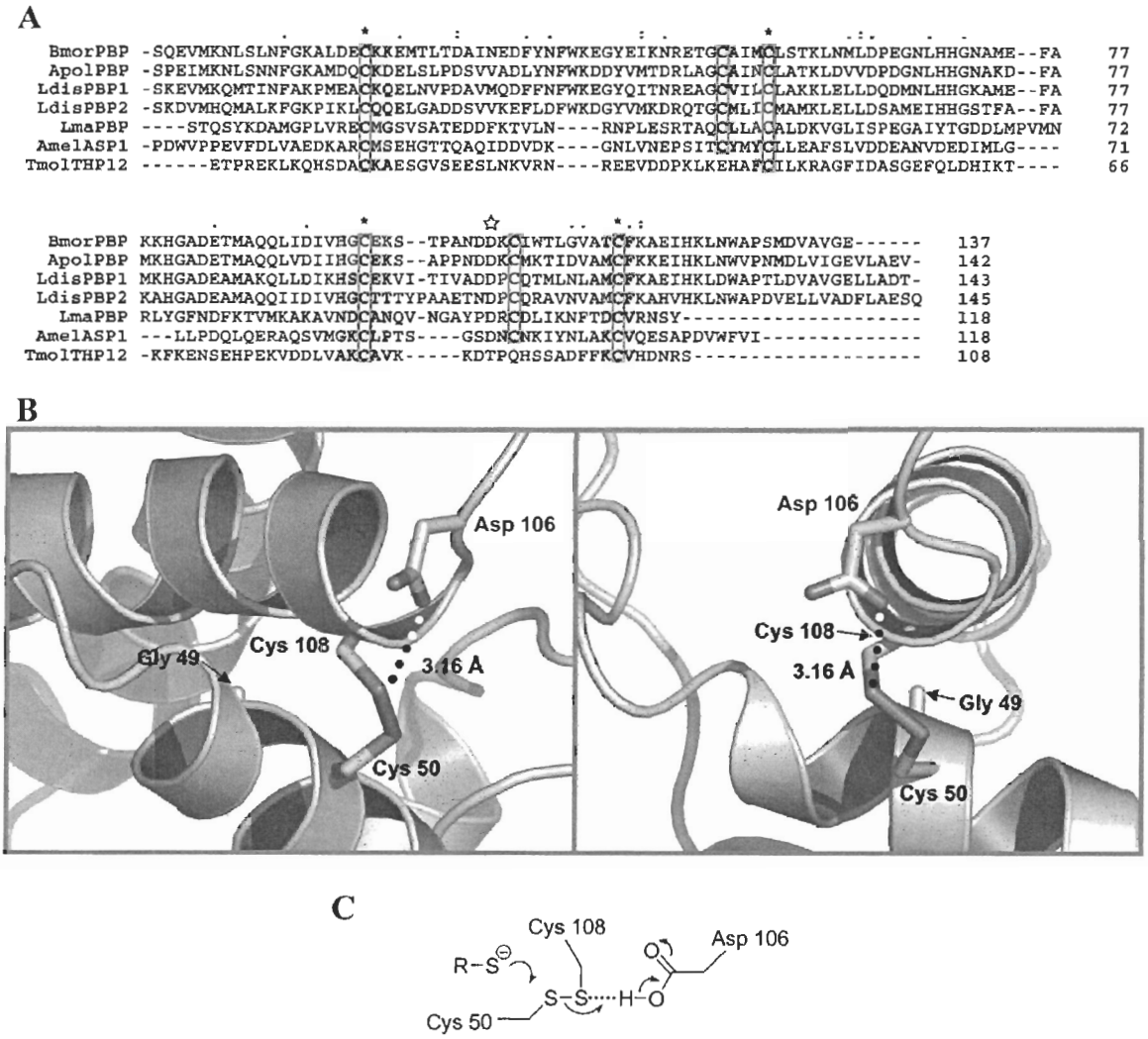



**Figure 4.4** **Sequence alignments of the insect-binding proteins and the C2-C5 disulfide bond. (A) Sequence alignments of the insect-binding proteins BmorPBP, ApolPBP, LdisPBP1, LdisPBP2, LmaPBP, AmelASP1, and the insect hemolymph protein TmolTHP12. Conserved cysteine residues are outlined in grey. Notice C2 and C5 are missing in TmolTHP12, yet the other four cysteines are conserved. The symbol '*' indicates a fully conserved residue, ':' specifies strongly conserved residues (score > 0.5), while '.' indicates a weakly conserved residue (score =< 0.5). The star indicates the highly conserved aspartic acid (Asp106 in BmorPBP) two residues upstream from C5. Sequence alignments were produced using the Clustal X (v. 1.83) multiple sequence alignment program. (B) Front and side views of the possible hydrogen bonding interactions between the sulfur acceptor of Cys108 (C5) and the oxygen donor of Asp106 from the NMR structure of BmorPBP (1LS8) (Lee *et al.* 2002). A glycine residue next to Cys50 (C2) results in an accessible region for nucleophilic attack to occur. Surrounding amino acid side chains are oriented away from the C2-C5 disulfide linkage (not shown), allowing the disulfide to be highly solvent accessible. Images were prepared using PyMOL (Delano, W.L. The PyMOL Molecular Graphics System. (2004) DeLano Scientific, San Carlos, CA, USA. http://www.pymol.org). (C) Scheme of nucleophilic attack at C2 and subsequent proton transfer from Asp106 to the sulfur of C5. This figure is from reference (Honson and Plettner 2006), by permission of the journal ©Naturwissenschaften.**

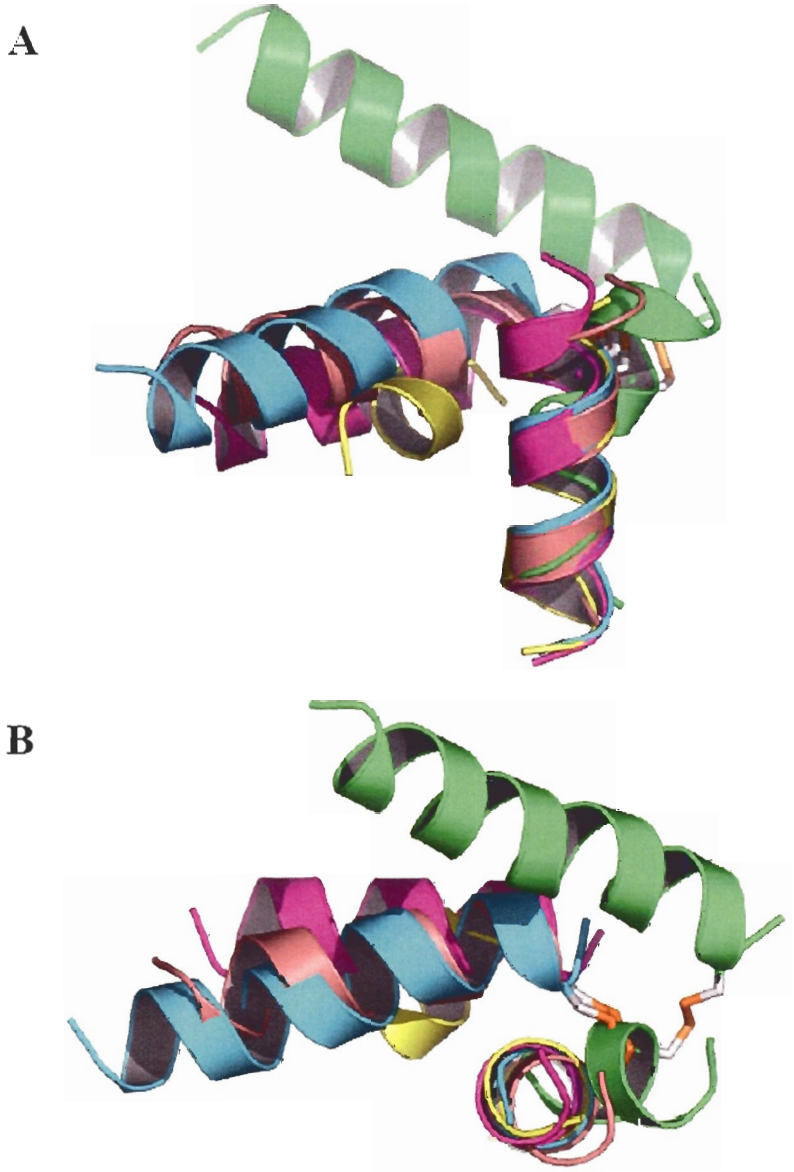



**Figure 4.5** **Orientation of helices 3 and 6 of various insect PBPs and TmolTHP12. The C2-C5 disulfide linkage is shown. A superimposition of helix 3 from all PBPs with solved structures and TmolTHP12 was generated by fitting the backbone $C_\alpha$ atoms of the third cysteine from the PBPs and the second cysteine from THP12, as well as two neighboring amino acid residues (see Table 4.3). (A) Side view with helix 3 oriented vertically. (B) Top view looking down helix 3. Helices are colored as follows: ApolPBP (1QWV) (Mohanty *et al.* 2004) - green, BmorPBP (1GM0) (Horst *et al.* 2001) – blue, AmelASP1 (1R5R) (Lartigue *et al.* 2004) – pink, LmaPBP (1ORG) (Lartigue *et al.* 2003) – purple, TmolTHP12 (1C3Y) (Rothemund *et al.* 1999) – yellow. Note the C-terminus of helix 6 is oriented closely to helix 3 in TmolTHP12, while in ApolPBP the C-terminus of helix 6 is splayed further from helix 3. Images were prepared using PyMOL (Delano, W.L. 2004).**

### 4.2.3 Sterics and Electronics of the C2-C5 Disulfide Bond

To explore reasons why the interlocking disulfide bridge is easily reduced, we examined the sequence alignments of the insect PBPs and found a highly conserved aspartic acid residue two amino acids upstream from the fifth cysteine residue (Figure 4.4 A). This acidic residue is close enough in some PBPs to form a hydrogen bond between itself and the fifth cysteine (Figure 4.4 B). This nearby acid may assist reduction of the interlocking disulfide bridge by protonating the leaving group during disulfide reduction (Figure 4.4 C). The distance, $d_5$, between the fifth cysteine residue and the aspartic acid residue two amino acids upstream was measured for the insect PBPs (Table 4.3). The $d_5$ values reveal that all PBPs measured are within hydrogen bonding distance for an O-H···S bond (3.19 Å – 3.51 Å (Steiner 2002)) except for ApolPBP and AmelASP1, both of which have a $d_5$ distance greater than 4.2 Å. AmelASP1, however, only has one solved crystal structure, and there may be conformational flexibility when the protein is in solution.

## 4.3 Conformational Analysis of Cyanylated C2-C5 PBP1

### 4.3.1 MALDI-TOF MS

The modified PBP1 (PBP1.CN) was confirmed by MALDI-TOF MS. Doubly cyanylated PBP1 (PBP1 was reduced at one disulfide bond and cyanylated, so it carries two S-CN moieties and is referred to as PBP1.CN) was identified by an increase in mass of 52 Da. Attemps to doubly cyanylate PBP2 were unsuccessful when the reaction was performed at neutral pH. The C2-C5 disulfide bond of PBP2 reacts faster than that of PBP1, resulting in quadruply cyanylated PBP2 at pH 7.0 (two disulfide bonds were reduced and the four resulting sulfhydryl groups were cyanylated). This may be due to

PBP2 existing in different conformations at neutral and acidic pH. Cyanylation of protein for structural characterization was performed at neutral pH to ensure properly folded protein, whereas peptide mapping was performed at acidic pH.

### 4.3.2 PAGE Gel and Western Blot

To determine whether or not reduction of the C2-C5 bond results in a conformational change, I ran a 16% native PAGE gel and performed a Western blot, followed by detection using alkaline phosphatase (AP) (Figure 4.6). The native PAGE gel revealed that PBP1 and PBP1.CN both have the same electrophoretic mobilities, suggesting that reduction of the C2-C5 disulfide bond does not result in a large conformational change within the protein.

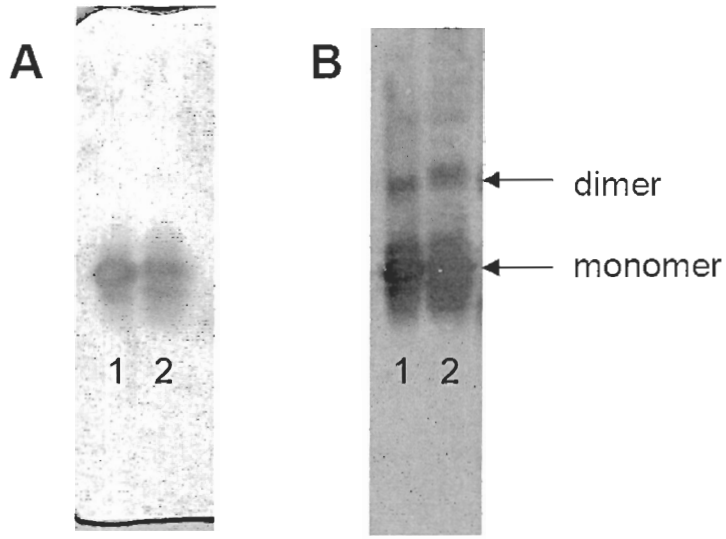



**Figure 4.6** **PAGE gel and immunoblot of PBP1 and cyanylated PBP1. (A) Coomassie stained 16% native PAGE gel. Lane 1 is recombinant PBP1, lane 2 is cyanylated PBP1. (B) Western blot of the native PAGE gel (A). Lane 1 is recombinant PBP1, lane 2 is cyanylated PBP1. Dimer is visible in the immunoblot (and not in the Coomassie stained gel) due to the increased sensitivity afforded by antibody detection.**

### 4.3.3 CD Spectra

To further analyze the conformations brought about by reduction of the C2-C5 disulfide bond, far-UV CD spectra were obtained to compare the secondary structure of modified and unmodified PBP1. The CD spectra of both modified and unmodified PBP1 at 25°C were similar to results obtained from other insect OBPs (Figure 4.7) (Rothemund *et al.* 1999, Wojtasek and Leal 1999, Mohl *et al.* 2002). A curve characteristic for mainly α-helical proteins was obtained, with a maximum centered around 195 nm indicating the $\pi-\pi^*_{\perp}$ absorbance, and two minima at approximately 210 nm and 218 nm, indicating the $\pi-\pi^*_{\parallel}$ and $n-\pi^*$ absorbances, respectively (Table 4.4). Slight thermodenaturation was observed as the temperature was increased in 5°C increments, from 25°C to 95°C. Recooling of the sample to 25°C resulted in a curve almost identical to the unheated sample, suggesting that denaturation is reversible (Figure 4.7 C, D). At 95°C, cyanylated

PBP1 showed slightly less α-helical content than the unmodified protein (Figure 4.7 C, D). The most dramatic difference between modified and unmodified PBP1 occurred around the 195 nm region. At 25°C, the maximum wavelength of both proteins are centered at about 195 nm with similar molar ellipticities (195 nm and $4.6 \times 10^6$ for PBP1, and 194 nm and $5.0 \times 10^6$ for cyanylated PBP1). At 95°C, the molar ellipticity of PBP1 decreases by $1.9 \times 10^6$ and a 2 nm red shift is observed. The red shift suggests that the environment around the chromophores is more polar on average. This could be due to increased interactions of the protein with water. Interestingly, the molar ellipticity of cyanylated PBP1 decreases by $3 \times 10^6$, a 63% larger change than the unmodified protein, and this is accompanied by a larger red shift of 4 nm (Table 4.4). This suggests cyanylated PBP1, which lacks the C2-C5 disulfide bridge, is more susceptible to denaturation at high temperatures than unmodified PBP1. The melting temperature ($T_m$) could not be calculated at each transition because there was no levelling off of data points (Figure 4.8). Perhaps the temperature required for fully denatured protein samples was never reached. This supports the notion that PBPs are extremely stable proteins, even at high temperatures.

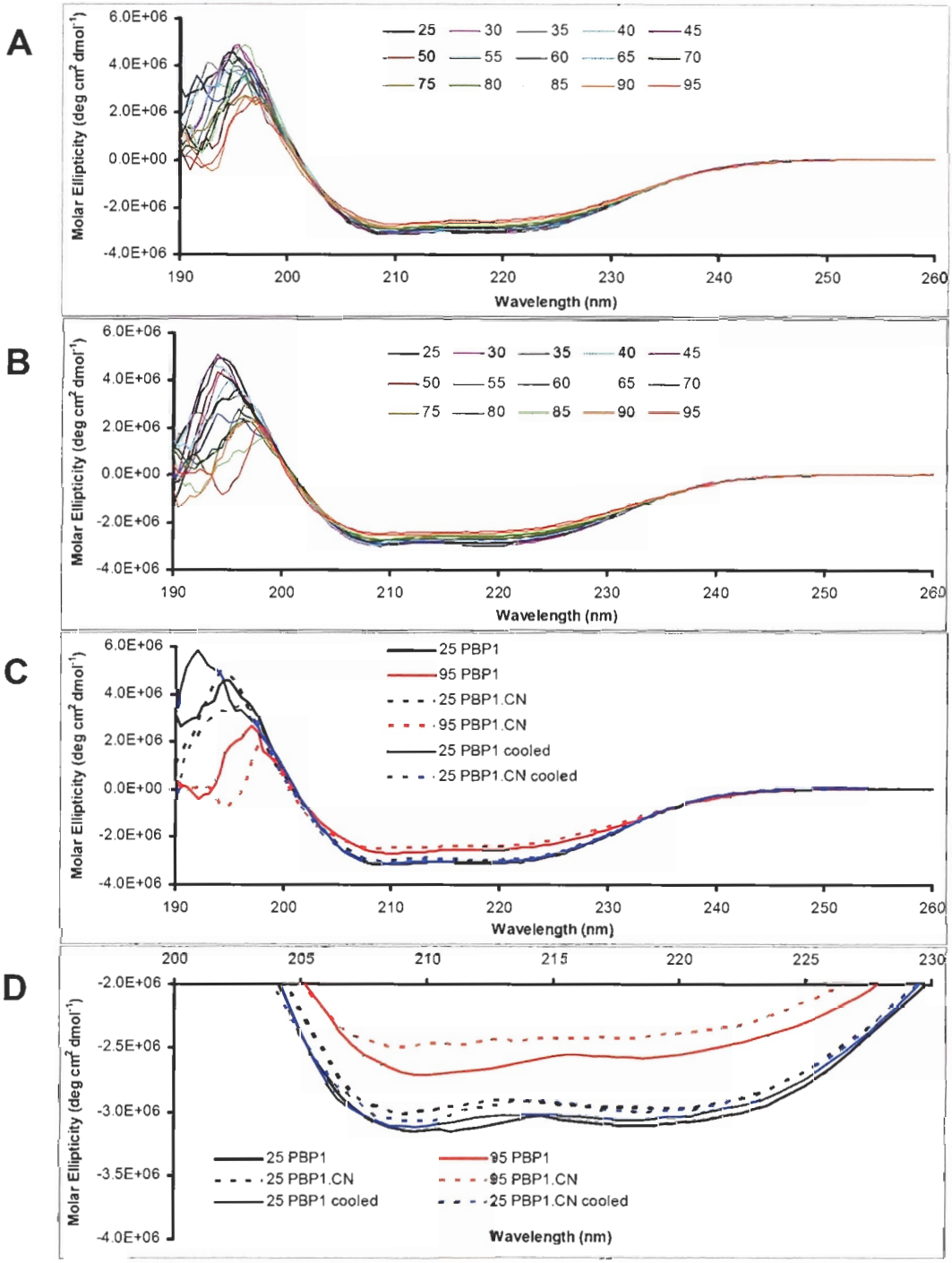



**Figure 4.7** **CD spectra of PBP1 and cyanylated PBP1 at various temperatures. (A) CD spectra of unmodified PBP1. (B) CD spectra of cyanylated PBP1. (C) A comparison of the CD spectra of PBP1 and cyanylated PBP1 (PBP1.CN) at 25°C and 95°C. (D) A close up of the $\pi$–$\pi^*_{\parallel}$ and n–$\pi^*$ absorbancees from (C). "Cooled" indicates the protein was heated from 25°C to 95°C (in 5°C intervals) and then allowed to cool back to 25°C before the CD spectra was recorded.**

**Table 4.4 Molar ellipticity values of modified and unmodified PBP1.**

| Protein | Temp. | $\pi$–$\pi^*_{\perp}$ | | $\pi$–$\pi^*_{\parallel}$ | | n–$\pi^*$ | |
|---|---|---|---|---|---|---|---|
| | (°C) | λ (nm) | [Θ] | λ (nm) | [Θ] | λ (nm) | [Θ] |
| PBP1 | 25 | 195 | $4.6 \times 10^6$ | 211 | $-3.2 \times 10^6$ | 218 | $-3.1 \times 10^6$ |
| | 95 | 197 | $2.7 \times 10^6$ | 211 | $-2.7 \times 10^6$ | 217 | $-2.6 \times 10^6$ |
| PBP1.CN | 25 | 194 | $5.0 \times 10^6$ | 209 | $-3.0 \times 10^6$ | 220 | $-3.0 \times 10^6$ |
| | 95 | 198 | $2.0 \times 10^6$ | 209 | $-2.5 \times 10^6$ | 218 | $-2.4 \times 10^6$ |

The percentage of α-helical content was determined using the software SELCON3 (Table 4.5). The total helical content was found to be 55.7% for both modified and unmodified PBP1 at all temperatures (estimated by adding the percentages of random coil helix and disordered helix). The change in molar ellipticity values may have been too small for the SELCON3 algorithm to differentiate. However, the CD spectra clearly show a cooperative process, characterized by loss of helical content at higher temperatures (Figure 4.7 and Figure 4.8). The estimated α-helical content for both modified and unmodified PBP1 at 25°C were similar to α-helical content determined from the solved three-dimensional structures from other insect OBPs (Table 4.5).

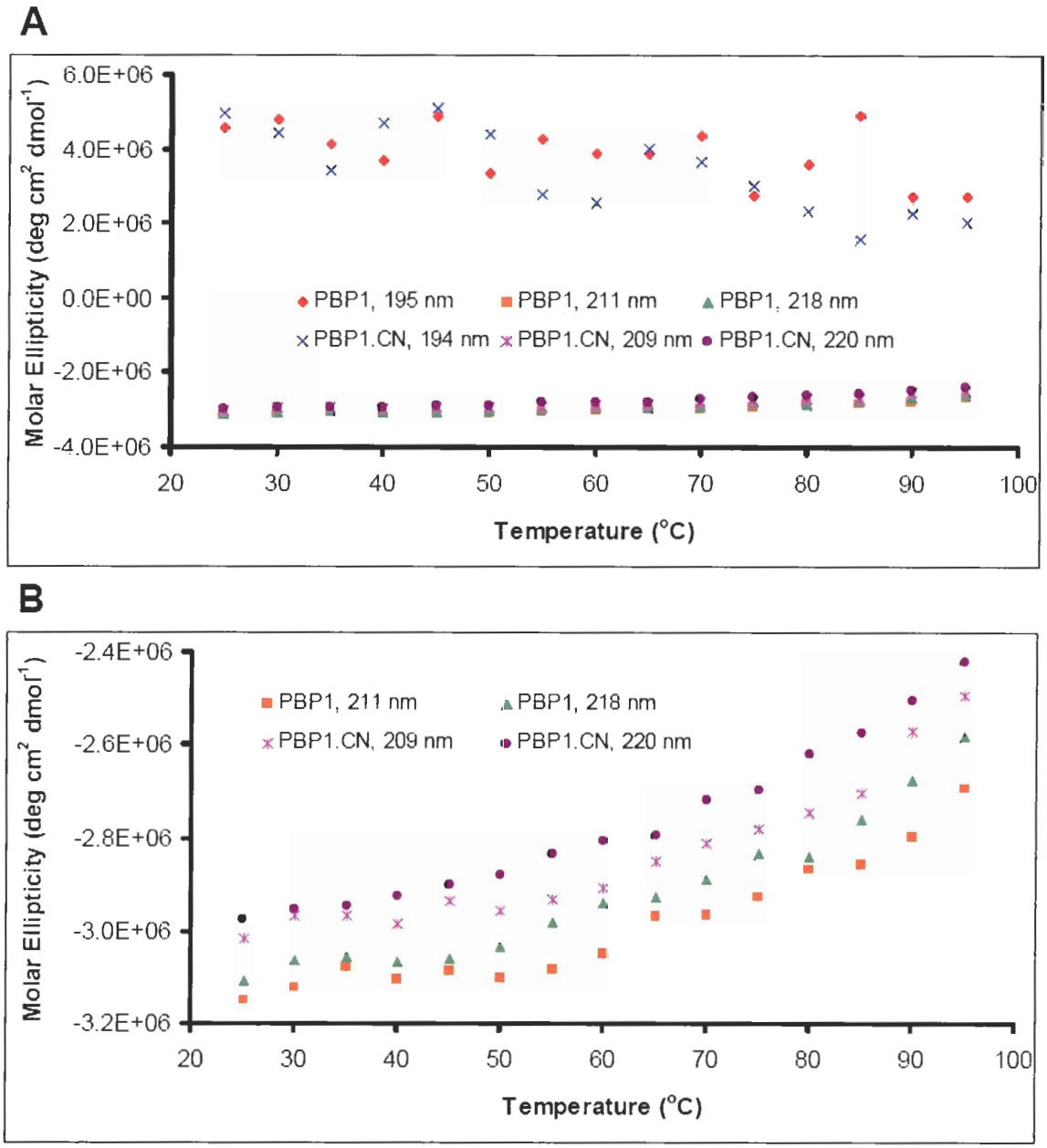



**Figure 4.8** **Thermal stability of PBP1 and cyanylated PBP1. Molar ellipticity values at $\lambda_{max}$ and $\lambda_{min}$ were used to generate each data point. The melting temperature ($T_m$) of each data set could not be calculated because the full extent of denaturation was not reached at 95°C. (A) Data plotted for molar ellipticities at all three $\lambda_{max}$ and $\lambda_{min}$. (B) Data plotted only for molar ellipticities at $\pi$–$\pi^*$ and n–$\pi^*$ transitions.**

**Table 4.5 Estimated secondary structure content of various OBPs.**

| Protein | Temp. (°C) | H(r)[a] | H(d) | S(r) | S(d) | Turns | Random coil |
|---|---|---|---|---|---|---|---|
| PBP1[b] | 25 | 36.3% | 19.4% | 2.6% | 2.9% | 16.8% | 22.4% |
| | 95 | 36.3% | 19.4% | 2.6% | 2.9% | 16.8% | 22.4% |
| PBP1.CN[b] | 25 | 36.3% | 19.4% | 2.6% | 3.0% | 16.9% | 22.4% |
| | 95 | 36.3% | 19.4% | 2.6% | 3.0% | 16.9% | 22.4% |
| LmaPBP[c] | 25 | 65.3% | | n.c.[d] | n.c. | n.c. | n.c. |
| BmorPBP[c] | 25 | 62.8% | | n.c. | n.c. | n.c. | n.c. |
| AmelASP1[c] | 25 | 54.6% | | n.c. | n.c. | n.c. | n.c. |
| ApolPBP[c] | 25 | 51.4% | | n.c. | n.c. | n.c. | n.c. |
| TmolTHP12[c] | 25 | 43.5% | | n.c. | n.c. | n.c. | n.c. |

[a] Abbreviations are as follows: H(r) is for regular α-helix, H(d) is for disordered α-helix, S(r) is for regular β-sheet, S(d) is for disordered β-sheet.
[b] Estimated secondary structure content by SELCON3.
[c] Calculated from the crystal structures described in Figure 4.5. The number of helical residues were divided by the number of residues solved in the three-dimensional structure and multiplied by 100 to obtain a percentage.
[d] n.c. = not calculated.

### 4.3.4 Ligand Binding Determinations by the GC Assay

To determine if the C2-C5 disulfide bond played a role in ligand binding, the dissociation constants were determined for 10 μM of either (+) or **(-)-1** with 2 μM of either modified or unmodified PBP1 or PBP2 by the GC binding assay (developed by Y. Gong and E. Plettner, unpublished). (Table 4.6). The $K_d$ values determined by the GC binding assay for PBP1 and PBP2 are similar to values obtained using the radioassay (Plettner *et al.* 2000). The results indicate the $K_d$ values determined for both modified and unmodified PBP1 with **(-)-1** to be almost identical. Thus the C2-C5 disulfide bond does not affect the binding of the preferred enantiomer with PBP1. However, there was very weak binding of PBP1.CN with the non-preferred enantiomer, **(+)-1**, when

compared with the $K_d$ value of unmodified PBP1. This suggests the C2-C5 disulfide bond plays a role in ligand discrimination.

**Table 4.6 Dissociation constant determination of 1 bound to either unmodified or modified PBP1.**

| Protein | Ligand | Average $K_d$[a] (μM) |
|---|---|---|
| PBP1 | **(-)-1** | 2.2 ± 1.9 |
| | **(+)-1** | 11 ± 10 |
| PBP1.CN[b] | **(-)-1** | 2.7 ± 2.5 |
| | **(+)-1** | 145 ± 100 |
| PBP2 | **(-)-1** | 12 ± 10 |
| | **(+)-1** | 2.3 ± 1.1 |

[a] $K_d$ values are reported as an average ± S.E. of 5 replicates. Control assays performed included determination of the concentration of ligand eluted from the column without protein as well as determination of the concentration of ligand adsorbed onto the glass vials.
[b] This PBP was reduced at one disulfide bond and cyanylated, so it carries two S-CN moieties (see section 2.3.2).

## 4.4 Crystallization Trials of PBP1 and PBP2

### 4.4.1 Crystallization Conditions

A variety of buffer, salt, and precipitant conditions were utilized to screen crystal growing conditions for both PBP1 and PBP2 with and without ligand. All trials were performed using the hanging drop vapor diffusion method. Crystal Screen 1™ (Hampton Research) was used in initial screenings. PBP1 was then chosen as the protein to develop the crystallization procedure since only PBP1 crystals were observed. Once PBP1 crystals were obtained, optimization was performed by varying the pH of the buffer and precipitant concentrations. Table 4.7 summarizes the conditions which yielded crystals suitable for X-ray diffraction. It is interesting to note that in conditions using PEG 4000,

$MgCl_2$ was necessary to produce crystals. Thus crystallization plates using various concentrations of $MgCl_2$ or changing the salt should be set up for future crystallization trials.

**Table 4.7 PBP1 crystallization trials.**

| Ligand | Expression vector | Protein conc. (mg/mL) | Ligand equiv. | Reservoir buffer | Growth time (days) |
|---|---|---|---|---|---|
| **(+)-1** | pET | 40.5 | 5 | 1.4 M Na, K $H_2PO_4$ in 0.1 M Na Hepes pH 7.0 | 21 |
| | pHN1+ | 54.6 | 10 | 30% PEG 4000, 0.1 M Tris-HCl pH 7.5 – 9.0, 0.2 M $MgCl_2$ | 4 |
| | | 30, 35, 40, 45, 50 | | 29-30% PEG 4000, 0.1 M Tris-HCl pH 7.5 – 9.0, 0.2 M $MgCl_2$ | 2 |
| | | 10, 20, 30 | | 28.75% PEG 4000, 0.1 M Tris-HCl pH 8.0, 0.2 M $MgCl_2$ | 4 |

### 4.4.2 Diffraction of PBP1 Crystals

A number of diffraction quality crystals were obtained. Crystals were mounted and exposed for X-ray diffraction by Dr. Mark Paetzel (Department of Molecular Biology and Biochemistry, Simon Fraser University). In most cases, a cryprotectant was used consisting of the same reservoir buffer concentrations with 15% (v/v) glycerol. Most crystals diffracted, however, the patterns obtained were too diffuse due to the crystals being mosaic. Although the crystals appeared to be single, it may be that two or more crystal layers are present. Figure 4.9 shows crystals of liganded PBP1 obtained using two different conditions. Because crystals obtained were consistently layered, in

future trials ligand will be added also to the reservoir buffer. This is to ensure saturation of the protein with ligand.

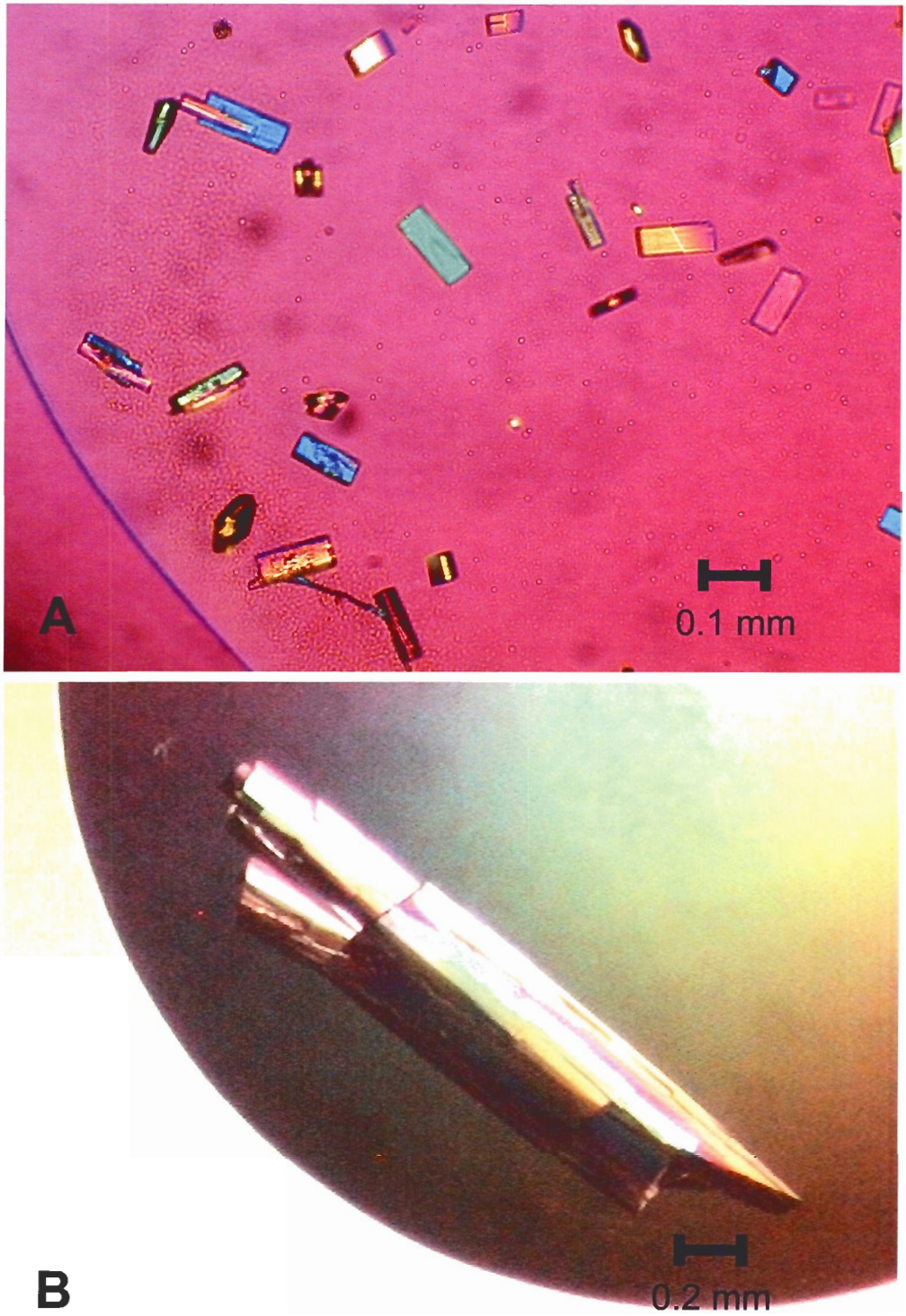



**Figure 4.9** Crystals of PBP1 co-crystallized with (+)-1. (A) PBP1 pET (40.5 mg/mL) with 5 times excess of (+)-1. The reservoir buffer contains 1.4 M Na, K dihydrogen phosphate in 0.1 M Na Hepes pH 7.0. (B) PBP1 pET (54.6 mg/mL) with 10 times excess of (+)-1. The reservoir buffer contains 30% PEG 4000, 0.1 M Tris-HCl pH 8.5, 0.2 M $MgCl_2$.

## 4.5 Discussion

### 4.5.1 Disulfide Connectivities of *L. dispar* PBP1 and PBP2

In the cyanylation experiments used to determine the disulfide connectivities of PBP1 and PBP2, only one disulfide could be reduced and cyanylated under the peptide mapping conditions described in section 2.3.2. To obtain peptides, CNBr was employed after partial reduction. The information obtained from these two peptide mapping techniques revealed *L. dispar* PBPs to contain three interlocking disulfide bridges. The disulfide linkages are cysteines 19-54, 50-109, and 97-118 in PBP1, and cysteines 19-54, 50-110, and 97-119 in PBP2. These results are similar to other mapped and solved structures of insect PBPs such as BmorPBP (Wojtasek and Leal 1999, Sandler *et al.* 2000, Horst *et al.* 2001, Lee *et al.* 2002), LmaPBP (Lartigue *et al.* 2003), ApolPBP (Mohanty *et al.* 2004, Zubkov *et al.* 2005), and AmelASP1 (Briand *et al.* 2001, Lartigue *et al.* 2004). Interestingly, this second disulfide linkage (C2-C5) was found to be the most easily reduced of all three bridges as determined by cyanylation and subsequent digestion.

### 4.5.2 Properties of the C2-C5 Disulfide Bond: Effect on Helix Orientation and Torsion Angle Comparisons with Known PBP Structures

Our results from the cyanylation reaction reveal that the C2-C5 disulfide is the most rapidly reduced disulfide with various reductants: TCEP (shown above), 2-mercaptoethanol, and DTT (see below). Disulfide cleavage by these reductants proceeds *via* nucleophilic attack on the oxidized disulfide (Shaked *et al.* 1980). Homology modeling of PBP1 and PBP2 using the crystal structure of BmorPBP as a template reveals that the most easily reduced disulfide bridge is also the most exposed and

accessible disulfide (Figure 4.10). This disulfide has C5 within hydrogen bonding distance of an aspartic acid residue. This acidic residue may protonate the sulfhydryl group of C5 during reduction (Figure 4.4 B, C). Furthermore, C2 has an adjacent glycine residue on its N-terminal side, making it accessible to reducing thiolate and phosphine nucleophiles (Figure 4.4 B). This glycine is conserved in lepidopteran PBPs (Figure 4.4 A). In a previous paper, we had developed a new fluorescence binding assay where one disulfide bridge of *L. dispar* PBPs was first reduced with DTT followed by reaction of the reduced protein with the fluorescent probe dansyl chloride (Honson *et al.* 2003). The identity of this modified disulfide bridge as the most easily reducible disulfide bond was determined in this work (Cys50-Cys109 and Cys50-Cys110 in PBP1 and PBP2, respectively).

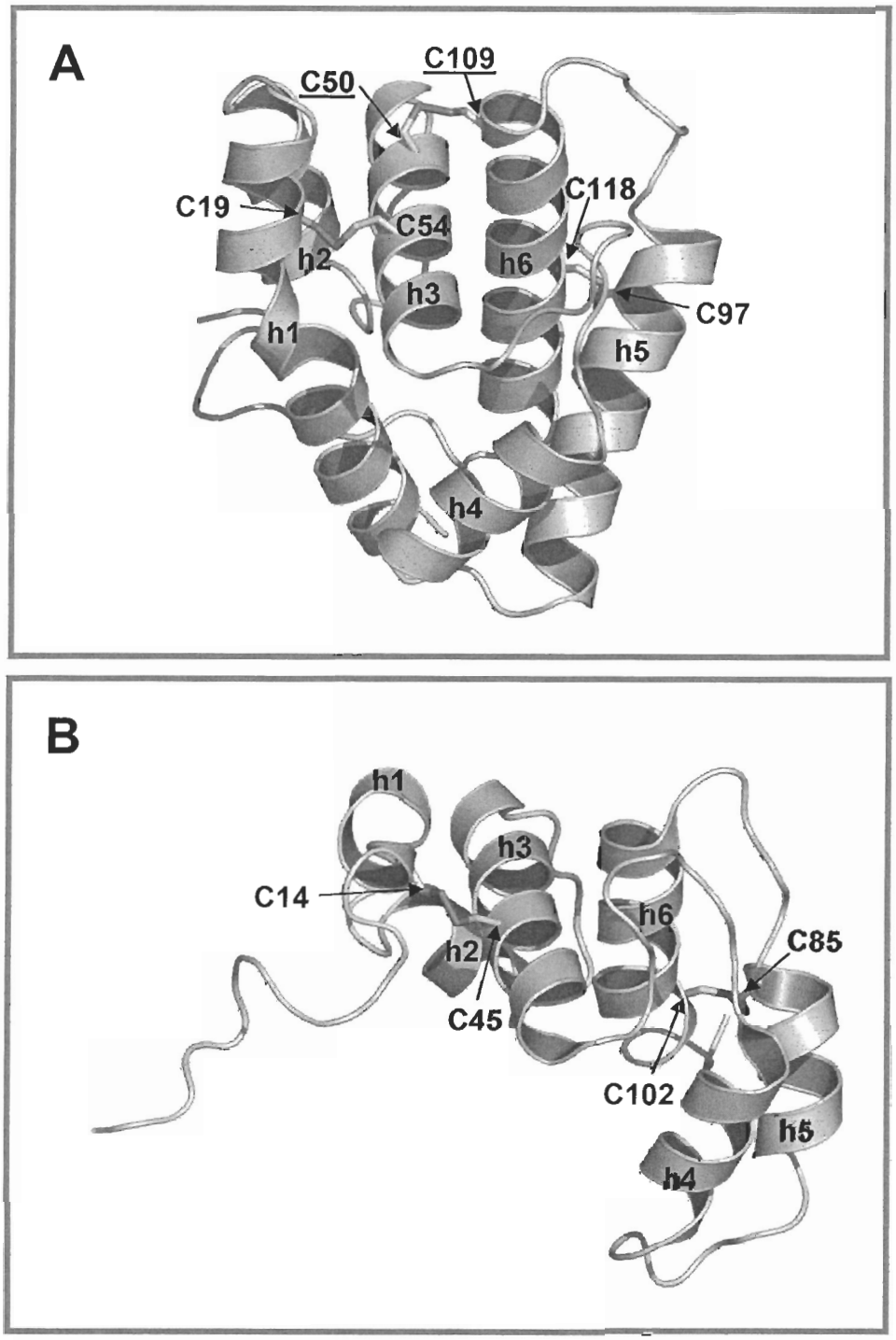



**Figure 4.10** **Location of the disulfide bonds in LdisPBP1 and TmolTHP12. (A) Homology model of PBP1 using the crystal structure of BmorPBP (1DQE) (Sandler *et al.* 2000) as a template. The Cys50-Cys109 bond is underlined. (B) NMR solved structure of TmolTHP12. Note the missing disulfide bond corresponds to the second disulfide linkage in insect PBPs. Images were prepared using PyMOL (Delano, W.L. 2004). Figure is adapted from reference (Honson and Plettner 2006), by permission of the journal ©Naturwissenschaften.**

The reaction kinetics of the C2-C5 disulfide bond reduction was observed while monitoring the progress of the cyanylation reactions by MALDI-TOF MS. In reactions to produce large quantities of cyanylated protein for structural characterization, the PBP1 C2-C5 disulfide bond was first reduced and cyanylated after 3 h, while the PBP2 C2-C5 disulfide bond was reduced and cyanylated in less than 1 h. Interestingly, the $d_5$ distances of PBP1 and PBP2 (measured from the homology model created using BmorPBP as a template (1DQE), with preferred enantiomers of **1** manually docked and minimized), were 4.53 Å (from the O of Asp107 to the S of Cys109) and 4.32 Å (from the O of Asp108 to the S of Cys110), respectively. This supports the hypothesis that hydrogen bonding interactions, between the sulfur acceptor of C5 and the oxygen donor of the aspartic acid two residues upstream, may assist reduction of the interlocking disulfide bridge (Figure 4.4). In the homology models, this PBP2 aspartic acid 108 oxygen is closer to C5 than it is in PBP1, and I have observed that PBP2 reduces and cyanylates faster than PBP1. Thus, I can then predict that ApolPBP will undergo reduction and cyanylation even more slowly than PBP1 because it has a larger average $d_5$ distance (6.45 Å) than PBP1 (Table 4.3). Interestingly, distance $d_5$ of the proteins illustrated in Figure 4.5 correlates with the angle between helix 3 and helix 6 when looking down helix 3 (Figure 4.5 B, Figure 4.11). The C2-C5 disulfide bond of BmorPBP has the shortest $d_5$ and BmorPBP helix 6 is closest to helix 3, while the C2-C5 disulfide bond of ApolPBP has the longest $d_5$ and ApolPBP helix 6 is farthest from helix 3.

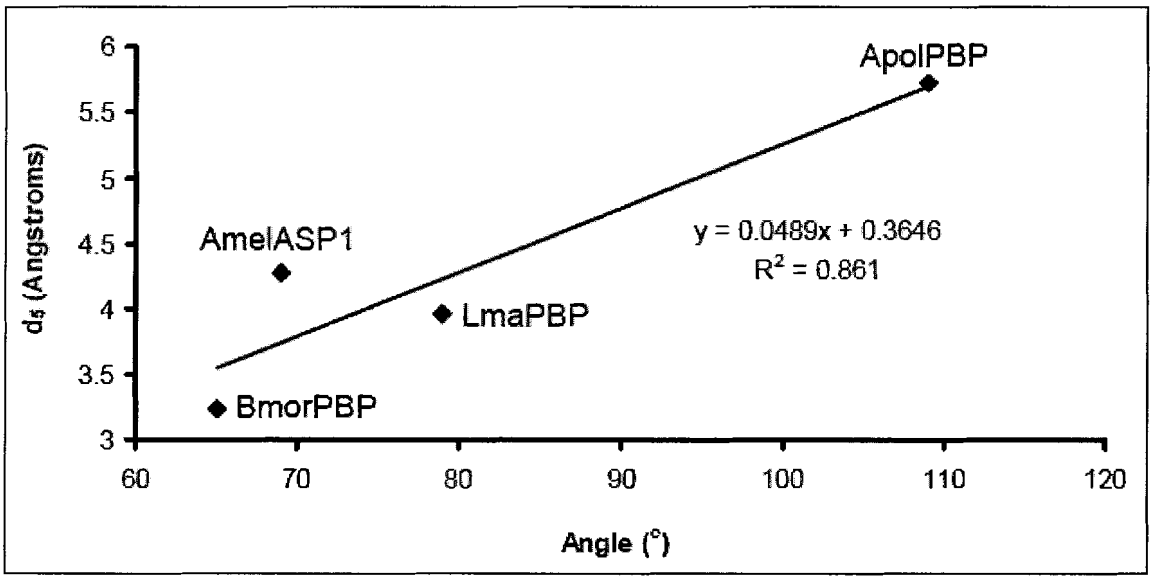



**Figure 4.11 Comparison of $d_5$ and the angle measured between helix 3 and helix 6 of insect PBPs. The same solved PBPs as in Figure 4.5 B (the top view looking down helix 3) were used. Angles measured were between a vertical line drawn down the center of helix 3 and from the line drawn through the center of helix 6.**

All members of the PBP family have three conserved disulfide bridges. Other insect non-antennal proteins have been discovered that have sequence similarity to insect OBPs but contain only two disulfide bridges, such as the *T. molitor* hemolymph protein (TmolTHP12) (Rothemund *et al.* 1997). The missing disulfide bridge in TmolTHP12 corresponds to the second interlocking disulfide bridge (C2-C5) in insect PBPs (Figure 4.4 A). TmolTHP12 has essentially the same folded structure as PBPs except for subtle differences. For example, helices 3 and 6 have a smaller interhelix angle (Rothemund *et al.* 1997, Rothemund *et al.* 1999, Graham *et al.* 2001) (Figure 4.10 and Figure 4.11).

We performed detailed measurements to determine reasons why the C2-C5 disulfide bridge is the most easily reducible of the three disulfide linkages. The torsion angles measured reveal that the S-S bond exists only within certain conformational ranges. The most frequent range among the PBPs studied was from about 95° to 112°

(for BmorPBP, LmaPBP, and AmelASP1) (Figure 4.2). All structures solved by X-ray diffraction had only this one conformational range which may be indicative of the crystal packing properties. All protein structures solved by X-ray diffraction were also co-crystalized with a ligand bound. Perhaps ligand binding results in an overall structural change to the protein leading to the preferred (frequent) conformational state. This conformational state correlates with a particular rotamer range about the interlocking disulfide bond.

The NMR structures of BmorPBP had two additional conformational ranges from 45° to 56° and -84° to -38°, suggesting that the second disulfide bond of the protein exists in one of three conformations when there is no ligand present. The conformational range of the NMR solved structure of ApolPBP are unlike the other PBPs studied, with C2-C5 torsion angle values ranging from 130° to 163°, and -151° to -145° (Figure 4.2, Figure 4.3 B). It is interesting to note that even though ApolPBP had high sequence similarity with the other PBPs measured, the C2-C5 bond prefers to be oriented very differently. All NMR structures reported so far have been solved without ligand.

The NMR-solved BmorPBP structure with PDB identification code 1GM0 is distinct from the others, because the C-terminal tail is inserted into the binding pocket. The disulfide bond in this structure has three torsion angle conformational ranges (Horst *et al.* 2001). Perhaps specific ligand interactions are needed to trigger rotation of the disulfide bond to a preferred conformer.

The various conformational ranges of the C2-C5 disulfide do not seem to affect the overall protein fold. However, this second disulfide linkage does affect the

orientation of helix 6 relative to helix 3, especially for the unique torsion angles exhibited by ApolPBP (see below).

An earlier study comparing the conformational characteristics of disulfide bridges in 22 different proteins found that all 72 disulfide bridges exhibit torsion angles of 90° ± 20° in both right and left handed disulfide conformations (Srinivasan *et al.* 1990). Since then, a more extensive investigation has been performed which analysed the conformations of cysteines for 351 disulfide bridges in 131 non-homologous monomeric protein structures determined by X-ray diffraction or NMR experiments (Petersen *et al.* 1999). This study included measurements of S-S bond lengths and torsion angles. The S-S bond length was found to be approximately 2.02 Å, similar to the bond lengths measured for insect PBPs, and the S-S bond torsion angle ranges in these structures to be either -110° to -50° (maxima at -80°) or 70° to 130° (maxima at 100°), with an equal distribution of each (Petersen *et al.* 1999). These studies suggest most of the C1-C3, C2-C5, and C4-C6 S-S bond torsion angles measured in insect PBPs as well as the C1-C2 and C3-C4 S-S bond torsion angles measured in the hemolymph protein TmolTHP12 are similar to other single-chain protein structures (Figure 4.3 A-C, Appendix 5). However, the C2-C5 disulfide bond in ApolPBP possesses unique conformational ranges, unlike any insect PBP and unlike any monomeric protein structure analyzed. Even some of the C1-C3 and C4-C6 S-S bonds in ApolPBP exhibit $\tau$ values outside of the expected ranges (Figure 4.3 A-C).

We compared the orientation of the helices which contain the second and fifth cysteine residues in PBPs with TmolTHP12 to see if the lack of a disulfide bond in TmolTHP12 has an effect on the protein structure. By superimposing the third cysteine

of helix 3 which is conserved among BmorPBP, LmaPBP, AmelASP1, ApolPBP, and TmolTHP12, we can see the position of helix 6 (Figure 4.5). To quantify the position of helix 6 relative to helix 3, we measured the distances between three different pairs of amino acids (Table 4.3). The N-termini of the helices are positioned close to one another and are almost perpendicular. However, the position of helix 6 in TmolTHP12 and ApolPBP differ dramatically from the other proteins studied. The C-terminus of helix 6 in TmolTHP12 is oriented closest to helix 3, while the C-terminus of helix 6 in ApolPBP is farthest from the third helix. The main structural displacement observed when TmolTHP12 was superimposed with the PBPs occurs in helix 6. Smaller displacements were also seen with helices 1 (all PBPs studied), 4 (LmaPBP) and 5 (BmorPBP, LmaPBP, and ApolPBP) (data not shown). These smaller displacements may simply be the result of the shift in helix 6. The large displacement of TmolTHP12 and ApolPBP helix 6 suggests a correlation between the C2-C5 disulfide bond torsion angle and the orientation of the helices which it interconnects. This missing disulfide bond in TmolTHP12 which connects helices 3 and 6 in ApolPBP may splay the two helices apart, orienting the C-terminus of helix 6 further from helix 3 and the N-terminus slightly closer. ApolPBP exhibits unique $\tau$ conformational ranges, resulting in its extreme displacement of helix 6 compared to the other PBPs. Mohanty and researchers have suggested the acetate containing pheromone of ApolPBP may enter the hydrophobic binding cavity through an opening located between a loop created by helix 2 and 3, helix 1, and the C-terminal end of helix 6 (Mohanty *et al.* 2004). Perhaps the large distance between helix 6 and helix 3 allows entry of the pheromone.

It is clear that the second disulfide linkage plays a role in the orientation of helix 6. We propose that formation of this disulfide bond guides the protein into its preferred conformation, resulting in the S-S bond torsion angle to be within one of the measured conformational ranges. Torsion angles outside of the measured ranges leads to displacement of helices 3 and/or 6. This can be observed with ApolPBP where its unique torsion angle ranges results in helix 6 being oriented further from helix 3 as compared to the other PBPs. Such a relative positioning may influence pheromone binding strength or specificity.

The reactivity of the disulfides towards nucleophilic attack is controlled by two factors: 1) sterics (the more accessible the thiol the faster the reaction), 2) electronics (a better leaving group results in a faster reaction). The second disulfide bond is the most easily reducible of the three for both these reasons. First, this disulfide bridge is the most easily accessible of the three, located on the surface of the protein with the cysteine side chains being solvent exposed. In addition, the C2 of all insect PBPs, except for AmelASP1 and LmaPBP, is next to a glycine residue which may provide room and flexibility for the nucleophilic attack to occur (Figure 4.4 A). The side chains of the residues adjacent to C2 of AmelASP1 and LmaPBP also provide access for a nucleophilic attack because they are oriented away from the thiol group. Second, the conformation of the disulfide bond itself may be strongly influenced by electronic effects, such as a nearby acid stabilizing the leaving thiol (Figure 4.4 B, C). For example, measurements between C5 and a highly conserved aspartic acid two residues upstream reveal that the aspartic acid is within hydrogen bonding distance of the distal sulfur (C5) of the disulfide. When a reducing agent attacks C2, the negative charge which is directed

towards the sulfur atom of C5 can be stabilized by the nearby acid, resulting in proton transfer from aspartic acid to the thiol and formation of a good leaving group. Aspartic acid residues in all insect PBPs studied could form weak hydrogen bonds with the sulfur atom of C5, except for that of ApolPBP. C5 is the only cysteine residue which has a highly conserved aspartic acid residue that is consistently located two amino acids upstream from C5 (Figure 4.4 A, p. 116).

We hypothesize that the disulfide guides the protein structure into a range of preferred conformers by its unique electronic properties. The orientation of the lone pairs of this second disulfide bond in insect PBPs provides an explanation for the conformational ranges observed (Figure 4.12). Orbital interactions have been studied with small synthetic disulfides. These studies reveal that the torsion angles (or dihedral angles) of R-S-S-R are about 100° (Pauling 1949, Oae 1991), a similar value to the most frequent conformational range of the C2-C5 disulfide bond in insect PBPs (95° to 112°) and a common torsion angle range in other monomeric protein structures (Petersen *et al.* 1999). This configuration is stabilized by orbital interactions which occur between the nonbonded *p* electrons and the σ antibonding S-C orbital which it eclipses (Figure 4.12) (Oae 1991). As the torsion angle approaches 0°, there is repulsion between the lone electron pairs on the two sulfur atoms, which results in S-S bond lengthening and subsequent weakening (Oae 1991).

The solved PBP NMR structures reveal another two conformational ranges from -84° to -38° and 45° to 56°. These less frequent ranges lead to non-eclipsed *p* orbitals which lack the extra stabilizing influence of the σ antibonding S-C orbital overlap (Figure 4.12). The most infrequent range is that of the second disulfide bond from

ApolPBP, with torsion angle values from -151° to -145° and 130° to 163° (Figure 4.2 and Figure 4.3 B). These ranges should give partial eclipsing of lone pairs (Figure 4.12). This unfavorable interaction may be cancelled by opposing stabilizing forces, such as helix-helix interactions.

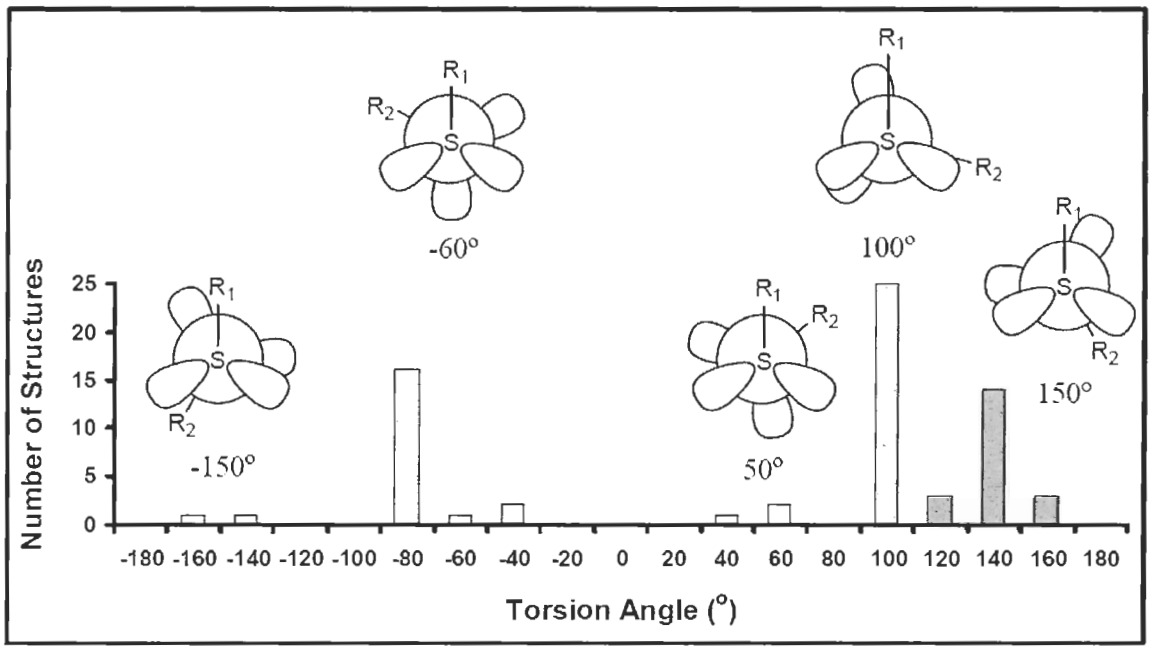



**Figure 4.12 Orientation of the C2-C5 disulfide bond in insect PBPs. The histogram illustrates the number of structures which possess the indicated disulfide bond torsion angles. Also shown are Newman projections along the S-S bond of the five torsion angle ranges. The Newman projection of the middle value of each range was chosen to represent each conformational range. R1 and R2 correspond to helix 3 and helix 6, respectively. The most frequent torsion angle of 100° is stabilized by orbital interactions which occur between the nonbonded *p* electrons and the σ antibonding S-C orbital which it eclipses.**

### 4.5.3 Biological Function of the C2-C5 Disulfide Bond

The C2-C5 disulfide bond in *L. dispar* PBP1 and PBP2 has unique steric and electronic properties. Comparisons with other PBPs have revealed interesting correlations between the torsion angles of this disulfide bond and the orientation of helix 3 and helix 6, which it interconnects. The C2-C5 disulfide linkage is the most easily reduced of the three, and this same linkage is missing among four cysteine-containing

insect OBPs. This suggests that this linkage may be important for PBP ligand selectivity by guiding helix 3 and helix 6 into specific orientations, thereby exposing important residues involved in ligand binding.

The secondary structural elements of PBP1 which lacks the C2-C5 disulfide bond were examined using CD. Both PBP1 with and without this disulfide bond had a high α-helical content of about 55.7%. However, PBP1 lacking the C2-C5 disulfide bond was more sensitive to thermal denaturation and showed slightly less α-helical content than unmodified PBP1. The GC binding assay revealed that the binding dissociation constant of PBP1 lacking the C2-C5 disulfide bond with **(-)-1** (the preferred enantiomer) did not differ significantly from that of the unmodified protein. Previous binding experiments with singly reduced PBP1 and PBP2 also revealed no change in ligand binding with their preferred enantiomer of **1** (Kowcun *et al.* 2001). However, cyanylated PBP1 resulted in very weak binding with **(+)-1** (the non-preferred enantiomer) when compared with unmodified PBP1. This suggests that the C2-C5 disulfide bond plays a role in ligand discrimination.

The biological significance of my findings is that under reducing conditions, the C2-C5 disulfide bond of PBPs can be reduced to produce PBPs with an altered conformation (helix 3 and helix 6 would be shifted) and slightly less α-helical content. PBPs have been found to exist in both oxidized and reduced forms *in vivo* (Ziegelberger 1995). The extracellular environment of the PBP dendritic nerve cell is oxidizing. However, a PBP could interact with any membrane bound protein embedded within the dendritic membrane to produce the reduced form. It may even be possible for the PBP to enter the reducing environment of the dendritic membrane. This reduced PBP could then

interact with the receptor, possibly removing odorants already bound to receptors and preventing prolonged hyperpolarization of the nerve cell (as suggested in 3.7.4).

### 4.5.4 Future Work

Cyanylation of PBP2 resulted in two disulfide bridges being reduced and cyanylated (see section 4.3.1). This was because of faster reduction of the C2-C5 disulfide bond of PBP2, when compared to reduction of PBP1 at neutral pH. In order to obtain only doubly cyanylated PBP1 and PBP2, site-directed mutagenesis can be employed to create mutants lacking this C2-C5 disulfide bond, where the cysteine residues are replaced by alanines.

The role of the conserved aspartic acid residue can also be tested by using site-directed mutagenesis. This aspartic acid residue can be mutated to an alanine to test if this residue assists reduction of the C2-C5 disulfide linkage.

# APPENDICES

## Appendix 1 Determination of Sensillar Lymph Volume

The average length of all types of sensilla is about 150 μm, with an inner diameter of 3 μm (Scheffler 1975). By estimating the shape of sensilla as cylinders, we can determine the approximate volume of lymph per sensillar hair:

$$V = L\pi r^2 \qquad (12)$$

$V = (150\ \mu m)\ \pi\ (1.5\ \mu m)^2 = 1060\ \mu m^3 = 1060 \times 10^{-12}\ cm^3$

$= 1.06 \times 10^{-12}\ L = 1\ pL/hair$

The following is an example calculation of the concentration of palmitic acid in the lymph from the extract of the aqueous whole antennal homogenate (Table 3.1):

1.46 μg/antenna ÷ (256.4 g/mol × 1 pL/hair × 30,000 hairs/antenna)

= 0.190 M

# Appendix 2 LC Traces of PBP1 and PBP2

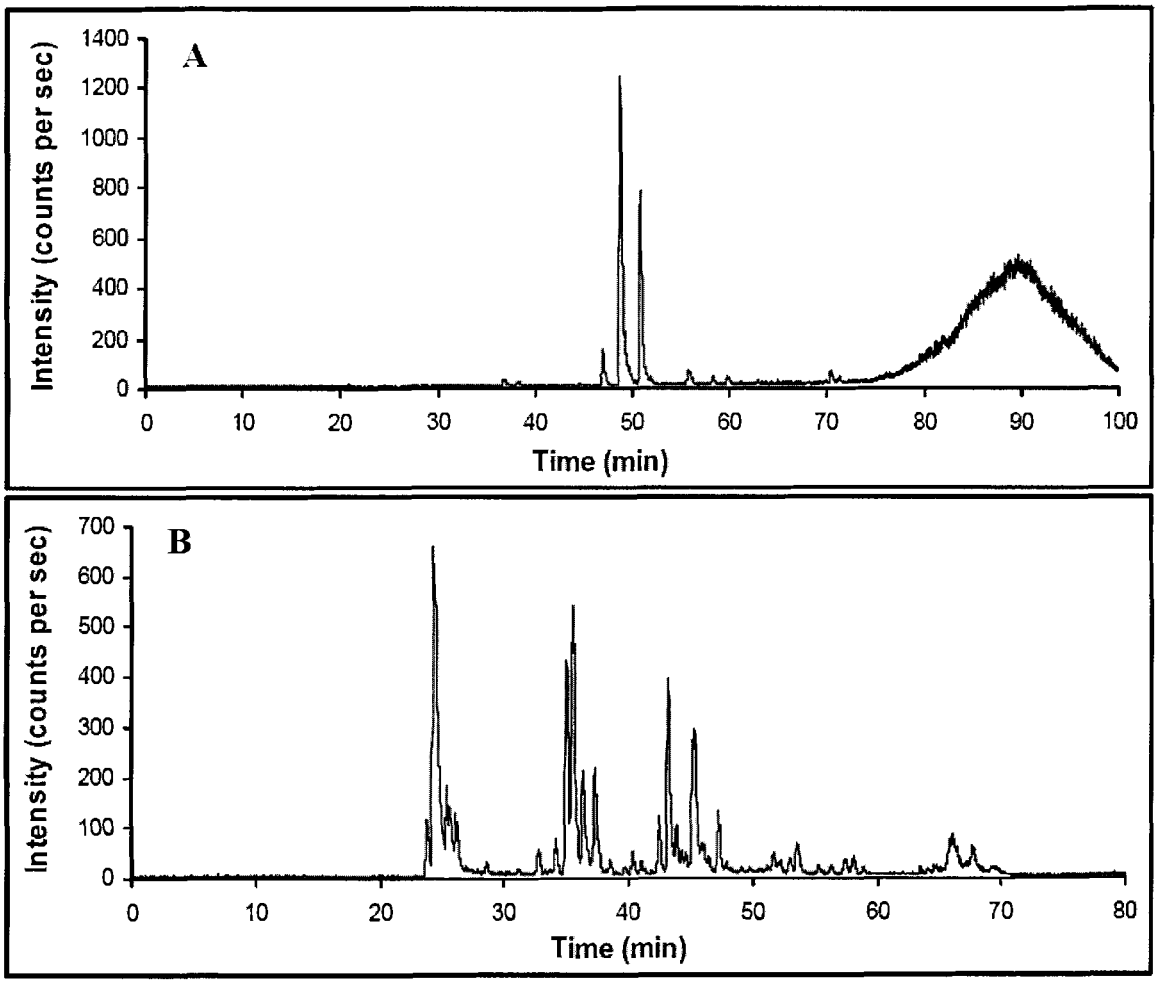



**Figure A.2.1 Liquid chromatograph traces of *L. dispar* PBPs. Elution was performed following a linear gradient at a flow rate of 200 nL/min (solution A is 98% H2O, 2% ACN, 0.1% formic acid; solution B is 15% H2O, 85% ACN, 0.1% formic acid): 5% solution B to 40% solution B in 60 minutes. (A) Base peak corrected chromatogram of PBP1. (B) Base peak chromatogram of PBP2. This figure is from reference (Honson and Plettner 2006), by permission of the journal ©Naturwissenschaften.**

# Appendix 3 Detailed CNBr Digest Assignments of all PBP1 and PBP2 Fragments (Excluding Fragments Containing a Disulfide Bond)

Assignments of all PBP1 fragments excluding fragments containing a disulfide bond are as follows. The N-terminal fragment residues 1-5 was detected as a very small peak at 546.50 Da (0.5% of the base peak, calculated mass of 545.30 Da). The fragment containing residues 6-8 (358.21 Da) was too small to be detected. Residues 9-16 were observed at 872.94 Da (0.5% of the base peak, calculated as 873.48 Da). Residues 34-66 were seen at 3860.19 Da (calculated as 3858.96 Da) and were a small peak (2.0% of the base peak). Loss of the amino acids FNFWKEG from the N-terminus of the 34-66 fragment resulted in a peptide fragment containing residues 41-66 observed at 2932.29 Da (13.95% of the base peak, calculated as 2932.53 Da). The same fragment appeared again with an additional 27.94 Da indicating formyl esterification of either a serine or threonine residue (Tarr and Crabb 1983, Kaiser and Metzka 1999). Amino acid residue Thr41 is the only residue in that fragment that could undergo the reaction with formic acid, resulting in a peak at 2960.23 Da (8.4% of the base peak, calculated as 2960.52 Da). Residues 53-74 resulted from cleavage at Met30 and Met74 and subsequent fragmentation at the N-terminus of the peptide, resulting in the loss of amino acids QDFFNFWKEGYQITNREAGCVI. The remaining residues (53-74) were detected as a small peak at 2475.19 Da (0.9% of the base peak, calculated mass of 2473.31 Da). Another peak was observed at 2493.79 Da (4.0% of the base peak, calculated as 2491.32 Da). The difference between these two peaks is 18.6 Da, suggesting that the latter peak was the homoseryl derivative (Smith 2002) of the peak appearing at 2475.19 Da. Yet another modification of this same (2475.19 Da) peak appeared at 2521.75 Da and is

thought to be the cysteic acid derivative (2.3% of the base peak, calculated mass of 2521.30 Da). The fragment containing residues 67-117 was observed as the $[M+2H]^{2+}$ ion at 2801.84 Da (1.1% of the base peak, calculated weight of 2800.38 Da), and it contains four missed cleavages at methionines 74, 78, 86, and 112. The identity of the three largest peaks was determined to be the fragment containing residues 75-117. The base peak appeared as the $[M+3H]^{3+}$ ion at 1574.58 Da (calculated as 1571.11 Da) and contains two threonines and one serine residue, with three missed cleavages at methionines 78, 86, and 112. The peak at 1592.66 Da is the same peptide fragment which has undergone formyl esterification on two of these residues (14.8% of the base peak, calculated mass of 1589.77 Da). The second largest peak appears at 1602.62 Da and is the same fragment as the previous peak but includes formyl esterification of all three residues (66.8% of the base peak, calculated mass of 1599.10 Da). The fragment containing residues 113-143 was observed as the $[M+2H]^{2+}$ ion as a very small peak at 1699.52 Da (0.5% of the base peak, calculated as 1699.38 Da) with one missed cleavage. However, loss of N-terminal amino acids LNLA resulted in three small peaks containing residues 117-143. The resulting N-terminal methionine is oxidized to yield methionine sulfoxide (Milzani *et al.* 2000, Grunert *et al.* 2003) seen as the $[M+H]^{+}$ peak at 3004.83 Da (9.7% of the base peak, calculated as 3002.50 Da). The same fragment, which contains two threonine residues, appears to undergo formyl esterification on one of these residues with the peak appearing at 3032.31 Da (7.5% of the base peak, calculated 3030.49 Da). Formyl esterification of both residues is seen at 3059.54 Da and 3060.69 Da (2.5% and 3.4% of the base peak, respectively, calculated as 3058.49 Da).

Assignments of all PBP2 fragments excluding fragments containing a disulfide bond are as follows. The N-terminal fragment containing residues 1-43 was observed as the $[M+2H]^{2+}$ ion at 2466.00 Da (8.4% of the base peak, calculated as 2464.24 Da), and contained two missed cleavages at methionines 5 and 8. Peptide fragment containing amino acids 44-51 (expected at 890.41 Da) was not detected. The fragment containing residues 52-118 was detected as three large $[M+2H]^{2+}$ peaks containing four missed cleavages at methionines 55, 57, 66, and 86. The peak observed at 3592.44 Da could arise from the formyl esterification of one of the two serine residues (serines 64 and 72), or one of the five threonine residues (threonines 73, 98, 99, 100, and 106) (23.9% of the base peak, calculated as 3591.70 Da). The largest of the three peaks at 3564.37 Da corresponds to the previously formylated peak after the loss of three water molecules, which could arise from aspartic acid isomerization in the presence of acid to yield a cyclic imide (Biemann and Scoble 1987, Ling *et al.* 1996) (66.4% of the base peak, expected mass of 3564.18 Da). The third peak at 3535.80 Da is due to N-terminal cleavage of the amide bond resulting in loss of leucine (47.5% of the base peak, expected mass of 3534.65 Da). The peak of residues 56-66 and the homoseryl derivative was observed at 1171.46 Da (2.5% of the base peak, expected mass of 1173.62 Da,) and 1189.48 Da (36.0% of the base peak, expected mass of 1191.63 Da,), respectively. Peptide fragment 58-66 and the formyl esterification of the fragment is observed at 971.11 Da (26.8% of the base peak, expected mass of 971.54 Da) and 999.26 Da (14.2% of the base peak, expected mass of 999.54 Da), respectively. Fragment 58-86 and the homoseryl derivative of this peptide was seen at 3078.85 Da (5.9% of base peak, calculated as 3079.49 Da) and 3096.93 Da (59.1% of the base peak, calculated as

3097.50 Da), respectively, and contains one missed cleavage at Met66. Peptide residues 67-86 were observed as four large peaks. The base peak appeared at 2079.48 Da (calculated 2078.97 Da). The $[M+2H]^{2+}$ ion was also present at 1039.13 Da (23.3% of the base peak, expected mass of 1040.49 Da). The formylated (both serine (Ser72) and threonine (Thr73) are present in the peptide) and homoseryl derivative of the base peak was also evident at 2107.52 Da (13.0% of the base peak, expected mass of 2106.96 Da) and 2125.52 Da (20.6% of the base peak, calculated 2124.97 Da), respectively. Finally, peptide fragment 87-118 was observed as a small peak at 3381.84 Da (2.3% of the base peak, expected mass of 3382.62 Da).

# Appendix 4 Confirmation of Disulfide Containing Fragments by Identification of Intramolecular Fragmentation Peptides for *L. dispar* PBPs

| PBP† | Parent peak | Fragment. product | $M_{monoiso}$ (Da) expected | observed (by ESI-MS) | Error (%) | Parent peptide location | Fragmentation peptide sequence cleaved from parent peptide |
|---|---|---|---|---|---|---|---|
| PBP1 | h | $[M+H]^{2+}$ | 2397.68 | 2397.02 | 0.03 | (87-112) + (118-143) | N-term - AKQLLDIK |
| | | $[M+H]^{2+}$ | 2696.39 | 2697.61 | 0.05 | | C-term - MaTQ |
| | | $[M+H]^{2+}$ | 2538.84 | 2539.06 | 0.01 | | C-term - MaTQCPD |
| | | $[M+H]^{2+}$ | 2445.81 | 2446.55 | 0.03 | | C-term - MaTQCPDDA |
| | | | | 2447.28 | 0.06 | | |
| | | | | 2447.40 | 0.07 | | |
| | | $[M+H]^{2+}$ | 2396.28 | 2397.02 | 0.03 | | C-term - MaTQCPDDAV |
| | | $[M+H]^{2+}$ | 2054.57 | 2053.89 | 0.03 | | C-term - MaTQCPDDAVITIVKE |
| | | $[M+H]^{2+}$ | 2801.94 | 2801.84 | 0.00 | | C-term - T |
| | | $[M+H]^{2+}$ | 2502.79 | 2503.34 | 0.02 | | C-term - TDALLEG |
| | | $[M+H]^{2+}$ | 2368.20 | 2367.28 | 0.04 | | C-term - TDALLEGVAV |
| | | | | 2367.44 | 0.03 | | |
| | | $[M+H]^{2+}$ | 2254.14 | 2254.31 | 0.00 | | C-term - TDALLEGVAVDL |
| | | $[M+H]^{2+}$ | 2203.62 | 2202.87 | 0.03 | | C-term - TDALLEGVAVDLT |
| | | $[M+H]^{2+}$ | 1658.84 | 1657.92 | 0.06 | | C-term –TDALLEGVAVD |
| | | | | 1658.00 | 0.05 | | LTPAWDLKHIE |
| | | | | 1658.30 | 0.03 | | |
| | | | | 1660.17 | 0.08 | | |
| | | $[M+H]^{2+}$ | 1623.32 | 1623.41 | 0.01 | | C-term - TDALLEGVA |
| | | | | 1624.36 | 0.06 | | VDLTPAWDLKHIEA |
| | | $[M+H]^{+}$ | 2970.47 | 2971.70 | 0.04 | | C-term – TDALLEGVA VDLTPAWDLKHIEAKF |
| | i | $[M+H]^{2+}$ | 2814.38 | 2812.64 | 0.06 | 17-66 | N-term - E |
| | | $[M+H]^{2+}$ | 2778.86 | 2777.98 | 0.03 | | N-term - EA |
| | | $[M+H]^{2+}$ | 2727.35 | 2727.40 | 0.00 | | N-term - EAC |
| | | | | 2727.53 | 0.01 | | |
| | | $[M+H]^{2+}$ | 2478.21 | 2479.36 | 0.05 | | N-term - EACKQEL |
| | | $[M+H]^{2+}$ | 2115.05 | 2114.42 | 0.03 | | N-term - EACKQELNVPDAVM |
| PBP2 | f | $[M+H]^{+}$ | 4086.04 | 4087.18 | 0.03 | (9-43) + (52-55) | N-term - ALK |
| | | | | 4087.70 | 0.04 | | |
| | | $[M+H]^{+}$ | 3566.86 | 3566.23 | 0.02 | | C-term - MaVYGDKW |
| | | $[M+H]^{+}$ | 3191.68 | 3192.58 | 0.03 | | C-term - MaVYGDKWFDL |
| | i | $[M+H]^{3+}$ | 3243.90 | 3242.60 | 0.04 | 56-145 | N-term - A |
| | | $[M+H]^{3+}$ | 3076.80 | 3078.85 | 0.07 | | N-term - AMKLE |
| | | $[M+H]^{2+}$ | 4400.60 | 4400.68 | 0.00 | | N-term – AMKLELLDS |
| | | $[M+H]^{3+}$ | 2866.68 | 2867.03 | 0.01 | | N-term - AMKLELLDSAM |
| | | $[M+H]^{2+}$ | 3969.41 | 3969.72 | 0.01 | | N-term - AMKLELLDSAMEIHHGS |
| | | $[M+H]^{2+}$ | 3636.73 | 3636.89 | 0.00 | | N-term – AMKLELLDSA MEIHHGSTFAFAK |
| | | $[M+H]^{2+}$ | 3601.21 | 3601.98 | 0.02 | | N-term – AMKLELLDSA |
| | | | | 3602.56 | 0.04 | | MEIHHGSTFAFAKA |

| PBP[†] | Parent peak | Fragment. product | $M_{monoiso}$ (Da) expected | observed (by ESI-MS) | Error (%) | Parent peptide location | Fragmentation peptide sequence cleaved from parent peptide |
|---|---|---|---|---|---|---|---|
| | | $[M+H]^{2+}$ | 3210.06 | 3211.98 | 0.06 | | N-term – AMKLELLDSA MEIHHGSTFAFAKAHGA DEAMA |
| | | $[M+H]^{2+}$ | 3146.03 | 3147.82 | 0.06 | | N-term – AMKLELLDSA |
| | | | | 3147.95 | 0.06 | | MEIHHGSTFAFAKAHGA |
| | | $[M+H]^{3+}$ | 2097.69 | 2097.47 | 0.01 | | DEAMAQ |
| | | $[M+H]^{3+}$ | 3091.49 | 3091.41 | 0.00 | | C-term - QSEAL |
| | | | | 3091.47 | 0.00 | | |
| | | $[M+H]^{2+}$ | 4087.45 | 4087.70 | 0.01 | | C-term - |
| | | | | 4088.18 | 0.02 | | QSEALFDAVLLEVDP |
| | | $[M+H]^{2+}$ | 3594.69 | 3594.59 | 0.00 | | C-term – QSEALFDAVLL EVDPAWNLKHVH |
| | | $[M+H]^{2+}$ | 3559.17 | 3559.41 | 0.01 | | C-term – QSEALFDAVLL EVDPAWNLKHVHA |

[†] Table is adapted from reference (Honson and Plettner 2006), by permission of the journal ©Naturwissenschaften.

[a] Methionine is in its homoserine lactone form.

# Appendix 5 Disulfide Bond Torsion Angles for C1-C3 and C4-C6

| Protein[a] | Ligand | PDB ID | Structure no. | C1-C3 τ (°) | C2-C5 τ (°) | C4-C6 τ (°) |
|---|---|---|---|---|---|---|
| BmorPBP | bombykol | IDQE | A | -86.53 | 108.75 | -78.00 |
| | | | B | -84.63 | 110.41 | -73.81 |
| BmorPBP | n/a | 1GM0 | 1 | -98.95 | 105.32 | -73.10 |
| | | | 2 | -82.35 | 98.26 | -79.40 |
| | | | 3 | -86.55 | 101.70 | -73.40 |
| | | | 4 | -89.14 | 104.68 | -76.82 |
| | | | 5 | -92.08 | 44.76 | -74.22 |
| | | | 6 | 114.26 | 102.03 | -70.49 |
| | | | 7 | -97.68 | -47.05 | -72.72 |
| | | | 8 | -108.77 | -82.45 | -74.89 |
| | | | 9 | -93.10 | 109.32 | -74.33 |
| | | | 10 | -98.11 | 102.56 | -70.98 |
| | | | 11 | -105.86 | -78.84 | -71.97 |
| | | | 12 | -90.00 | -70.83 | -75.41 |
| | | | 13 | -91.24 | -79.57 | -77.81 |
| | | | 14 | -92.17 | -80.51 | -68.94 |
| | | | 15 | -93.38 | 51.65 | -76.69 |
| | | | 16 | 137.18 | -38.27 | -68.89 |
| | | | 17 | -94.79 | 105.62 | -72.70 |
| | | | 18 | 152.84 | -61.98 | -72.93 |
| | | | 19 | -90.58 | -82.06 | -74.95 |
| | | | 20 | -97.91 | 101.84 | -73.16 |
| BmorPBP | n/a | 1LS8 | 1 | -89.30 | 105.93 | -84.67 |
| | | | 2 | -142.55 | 96.79 | -81.67 |
| | | | 3 | 109.77 | -78.85 | -82.95 |
| | | | 4 | 92.92 | 94.50 | -86.13 |
| | | | 5 | -87.01 | -76.77 | -79.72 |
| | | | 6 | -95.35 | -78.74 | -78.27 |
| | | | 7 | 114.34 | -76.47 | -78.29 |
| | | | 8 | -95.74 | 102.74 | -78.19 |
| | | | 9 | 94.09 | 102.33 | -81.74 |
| | | | 10 | 137.58 | 102.42 | -77.61 |
| | | | 11 | 120.87 | 107.05 | 69.04 |
| | | | 12 | 113.80 | -83.07 | -79.92 |
| | | | 13 | 93.40 | -78.23 | -81.45 |
| | | | 14 | 122.40 | -81.88 | -79.00 |
| | | | 15 | 117.72 | 56.03 | -75.49 |
| | | | 16 | 108.42 | 101.72 | -82.58 |
| | | | 17 | -87.40 | -75.77 | -84.41 |
| | | | 18 | -95.83 | -76.66 | -83.12 |
| | | | 19 | -92.47 | -84.39 | -77.94 |
| | | | 20 | 118.03 | 100.57 | -84.52 |

| Protein[a] | Ligand | PDB ID | Structure no. | C1-C3 τ (°) | C2-C5 τ (°) | C4-C6 τ (°) |
|---|---|---|---|---|---|---|
| LmaPBP | glycerol | 1ORG | A | -81.54 | 109.79 | -78.68 |
| | | | B | -81.72 | 101.14 | -77.84 |
| LmaPBP | ANS | 1OW4 | A | -81.25 | 109.13 | -81.62 |
| | | | B | -85.11 | 109.80 | -76.63 |
| LmaPBP | H3B2 | 1P28 | A | -79.52 | 109.26 | -78.51 |
| | | | B | -81.96 | 111.85 | -77.36 |
| ApolPBP | n/a | 1QWV | 1 | 146.29 | 140.21 | -63.44 |
| | | | 2 | -121.26 | 139.87 | -63.52 |
| | | | 3 | -60.30 | 152.06 | -63.68 |
| | | | 4 | -109.05 | 163.13 | 62.32 |
| | | | 5 | 86.82 | 141.43 | 65.01 |
| | | | 6 | 152.51 | 137.39 | -64.38 |
| | | | 7 | 110.76 | 142.67 | 112.96 |
| | | | 8 | 110.12 | 153.40 | -62.72 |
| | | | 9 | -35.27 | 144.52 | -68.16 |
| | | | 10 | 150.87 | 134.02 | -66.37 |
| | | | 11 | -81.80 | 143.03 | -64.36 |
| | | | 12 | -34.27 | 140.32 | -64.74 |
| | | | 13 | -118.13 | 149.81 | 62.83 |
| | | | 14 | 73.62 | 129.78 | 62.54 |
| | | | 15 | 152.59 | 147.80 | -65.91 |
| | | | 16 | -109.55 | -144.80 | 61.41 |
| | | | 17 | 146.89 | -151.10 | 60.30 |
| | | | 18 | -131.81 | 132.64 | -63.88 |
| | | | 19 | -86.79 | 133.40 | 60.01 |
| | | | 20 | -104.61 | 132.76 | -65.60 |
| AmelASP1 | NBBS | 1R5R | A | -79.64 | 101.91 | -78.78 |
| TmolThp12 | n/a | 1C3Y | 1 | 72.24 | n/a | -41.73 |
| | | | 2 | -78.20 | n/a | -50.81 |
| | | | 3 | -87.92 | n/a | -52.53 |
| | | | 4 | -86.66 | n/a | -52.72 |
| | | | 5 | -91.33 | n/a | -52.34 |
| | | | 6 | -100.42 | n/a | -53.79 |
| | | | 7 | 78.40 | n/a | -47.46 |
| | | | 8 | -91.19 | n/a | -49.42 |
| | | | 9 | -83.09 | n/a | -55.34 |
| | | | 10 | 87.59 | n/a | -54.89 |
| | | | 11 | 94.31 | n/a | -54.22 |
| | | | 12 | 87.23 | n/a | -59.53 |
| | | | 13 | -49.06 | n/a | -53.59 |
| | | | 14 | 94.38 | n/a | -53.07 |
| | | | 15 | 84.85 | n/a | -63.40 |
| | | | 16 | -98.37 | n/a | -54.88 |
| | | | 17 | -75.20 | n/a | -51.40 |

| Protein[a] | Ligand | PDB ID | Structure no. | C1-C3 τ (°) | C2-C5 τ (°) | C4-C6 τ (°) |
|---|---|---|---|---|---|---|
| | | | 18 | 70.69 | n/a | -50.20 |
| | | | 19 | 101.19 | n/a | -42.12 |
| | | | 20 | 78.49 | n/a | -51.96 |
| | | | 21 | 81.43 | n/a | -48.19 |
| | | | 22 | -69.34 | n/a | -52.26 |
| | | | 23 | -58.20 | n/a | -57.37 |

[a] Column headings are as follows: C1-C3 τ is the angle between the $C_\beta$ groups of the first and third cysteine residues (except for TmolTHP12, where the torsion angle is of the first and second cysteine residues) down the S-S bond measured with the lower numbered cysteine residue at the front upstream (1DQE (Sandler *et al.* 2000) 1GM0 (Horst *et al.* 2001), 1LS8 (Lee *et al.* 2002): C19-C54; 1ORG, 1OW4, 1P28 (Lartigue *et al.* 2003): C16-C47; 1QWV (Mohanty *et al.* 2004): C19-C54; 1R5R (Lartigue *et al.* 2004): C20-C51; 1C3Y (Rothemund *et al.* 1999): C14-C45), C2-C5 τ is the angle between the $C_\beta$ groups of the second and fifth cysteine residues down the S-S bond measured with the lower numbered cysteine residue at the front upstream (taken from Table 4.3 for comparison) (1DQE, 1GM0, 1LS8: C50-C108; 1ORG, 1OW4, 1P28: C43-C104; 1QWV: C50-C108; 1R5R: C47-C98); C4-C6 τ is the angle between the $C_\beta$ groups of the fourth and sixth cysteine residues (except for TmolTHP12, where the torsion angle is of the third and fourth cysteine residues) down the S-S bond measured with the lower numbered cysteine residue at the front upstream (1DQE, 1GM0, 1LS8: C97-C117; 1ORG, 1OW4, 1P28: C92-C113; 1QWV: C97-C117; 1R5R: C89-C107; 1C3Y: C85-C102).

# LITERATURE CITED


Ban L., Scaloni A., Brandazza A., Angeli S., Zhang L., Yan Y., Pelosi P. 2003. Chemosensory proteins of *Locusta migratoria*. *Insect Mol. Biol.* **12**: 125-134.

Ban L., Scaloni A., D'Ambrosio C., Zhang L., Yan Y., Pelosi P. 2003. Biochemical characterization and bacterial expression of an odorant-binding protein from *Locusta migratoria*. *Cell. Mol. Life Sci.* **60**: 390-400.

Baumann A., Frings S., Godde M., Seifert R., Kaupp U. B. 1994. Primary Structure and Functional Expression of a *Drosophila* Cyclic Nucleotide-Gated Channel Present in Eyes and Antennae. *Embo J.* **13**: 5040-5050.

Belles X., Martin D., Piulachs M. D. 2005. The mevalonate pathway and the synthesis of juvenile hormone in insects. *Annual Review of Entomology* **50**: 181-199.

Bette S., Breer H., Krieger J. 2002. Probing a pheromone binding protein of the silkmoth *Antheraea polyphemus* by endogenous tryptophan fluorescence. *Insect Biochem. Mol. Biol.* **32**: 241-246.

Biemann K., Scoble H. A. 1987. Characterization by Tandem Mass Spectrometry of Structural Modifications in Proteins. *Science* **237**: 992-998.

Bierl B. A., Beroza M., Collier C. W. 1970. Potent Sex Attractant of Gypsy Moth - Its Isolation, Identification, and Synthesis. *Science* **170**: 87-&.

Bohbot J., Sobrio F., Lucas P., Nagnan-Le Meillour P. 1998. Functional characterization of a new class of odorant-binding proteins in the moth *Mamestra brassicae*. *Biochem. Biophys. Res. Commun.* **253**: 489-494.

Borden J. H. 1993. Strategies and tactics for the use of semiochemicals against forest insect pests in North America. In: Lumsden R. D., Vaughn J. L., Lumsden R. D., Vaughn J. L.s. Pest Management: Biologically Based Technologies. ACS Conf. Proc. Washington, D.C.: Amer. Chem. Soc. p 265-279.

Breer H., Boekhoff I., Tareilus E. 1990. Rapid Kinetics of 2nd Messenger Formation in Olfactory Transduction. *Nature* **345**: 65-68.

Briand L., Nespoulous C., Huet J. C., Pernollet J. C. 2001. Disulfide pairing and secondary structure of ASP1, an olfactory-binding protein from honeybee (*Apis mellifera* L). *J. Pept. Res.* **58**: 540-545.

Butenandt A., Beckmann R., Stamm D., Hecker E. 1959. Uber Den Sexual-Lockstoff Des Seidenspinners *Bombyx Mori* - Reindarstellung Und Konstitution. *Zeitschrift Fur Naturforschung Part B-Chemie Biochemie Biophysik Biologie Und Verwandten Gebiete* **14**: 283-284.

Calvello M., Guerra N., Brandazza A., D'Ambrosio C., Scaloni A., Dani F. R., Turillazzi S., Pelosi P. 2003. Soluble proteins of chemical communication in the social wasp *Polistes dominulus*. *Cell. Mol. Life Sci.* **60**: 1933-1943.

Campanacci V., Krieger J., Bette S., Sturgis J. N., Lartigue A., Cambillau C., Breer H., Tegoni M. 2001. Revisiting the specificity of *Mamestra brassicae* and *Antheraea polyphemus* pheromone-binding proteins with a fluorescence binding assay. *J. Biol. Chem.* **276**: 20078-20084.

Chen I. A., Szostak J. W. 2004. A kinetic study of the growth of fatty acid vesicles. *Biophys. J.* **87**: 988-998.

Chen I. A., Szostak J. W. 2004. Membrane growth can generate a transmembrane pH gradient in fatty acid vesicles. *Proc. Natl. Acad. Sci. U. S. A.* **101**: 7965-7970.

Chirita C. N., Necula M., Kuret J. 2003. Anionic micelles and vesicles induce tau fibrillization *in vitro*. *J. Biol. Chem.* **278**: 25644-25650.

Cistola D. P., Hamilton J. A., Jackson D., Small D. M. 1988. Ionization and Phase-Behavior of Fatty-Acids in Water - Application of the Gibbs Phase Rule. *Biochemistry* **27**: 1881-1888.

Clyne P. J., Warr C. G., Freeman M. R., Lessing D., Kim J. H., Carlson J. R. 1999. A novel family of divergent seven-transmembrane proteins: Candidate odorant receptors in *Drosophila*. *Neuron* **22**: 327-338.

Clyne P. J., Warr C. G., Carlson J. R. 2000. Candidate taste receptors in *Drosophila*. *Science* **287**: 1830-1834.

Dahanukar A., Hallem E. A., Carlson J. R. 2005. Insect chemoreception. *Curr. Opin. Neurobiol.* **15**: 423-430.

Danty E., Briand L., Michard-Vanhee C., Perez V., Arnold G., Gaudemer O., Huet D., Huet J. C., Ouali C., Masson C., Pernollet J. C. 1999. Cloning and expression of a queen pheromone-binding protein in the honeybee: an olfactory-specific, developmentally regulated protein. *J. Neurosci.* **19**: 7468-7475.

Dauber-Osguthorpe P., Roberts V. A., Osguthorpe D. J., Wolff J., Genest M., Hagler A. T. 1988. Structure and Energetics of Ligand-Binding to Proteins - *Escherichia Coli* Dihydrofolate Reductase Trimethoprim, a Drug-Receptor System. *Proteins* **4**: 31-47.

Du G. H., Prestwich G. D. 1995. Protein-Structure Encodes the Ligand-Binding Specificity in Pheromone Binding Proteins. *Biochemistry* **34**: 8726-8732.

Fox A. N., Pitts R. J., Robertson H. M., Carlson J. R., Zwiebel L. J. 2001. Candidate odorant receptors from the malaria vector mosquito *Anopheles gambiae* and evidence of down-regulation in response to blood feeding. *Proc. Natl. Acad. Sci. U. S. A.* **98**: 14693-14697.

Fox A. N., Pitts R. J., Zwiebel L. J. 2002. A cluster of candidate odorant receptors from the malaria vector mosquito, *Anopheles gambiae*. *Chem. Senses* **27**: 453-459.

Gebicki J. M., Hicks M. 1973. Ufasomes Are Stable Particles Surrounded by Unsaturated Fatty-Acid Membranes. *Nature* **243**: 232-234.

Gietzen K., Xu Y. H., Galla H. J., Bader H. 1982. Multimers of Anionic Amphiphiles Mimic Calmodulin Stimulation of Cyclic-Nucleotide Phosphodiesterase. *Biochem. J.* **207**: 637-640.

Gohl T., Krieger J. 2006. Immunolocalization of a candidate pheromone receptor in the antenna of the male moth, *Heliothis virescens*. *Invertebr. Neurosci.* **6**: 13-21.

Graham L. A., Tang W., Baust J. G., Liou Y. C., Reid T. S., Davies P. L. 2001. Characterization and cloning of a *Tenebrio molitor* hemolymph protein with sequence similarity to insect odorant-binding proteins. *Insect Biochem. Mol. Biol.* **31**: 691-702.

Grant G. G., Langevin D., Liska J., Kapitola P., Chong J. M. 1996. Olefin inhibitor of gypsy moth, *Lymantria dispar*, is a synergistic pheromone component of nun moth, *L. monacha*. *Naturwissenschaften* **83**: 328-330.

Gries G., Gries R., Khaskin G., Slessor K. N., Grant G. G., Liska J., Kapitola P. 1996. Specificity of nun and gypsy moth sexual communication through multiple-component pheromone blends. *Naturwissenschaften* **83**: 382-385.

Grunert T., Pock K., Buchacher A., Allmaier G. 2003. Selective solid-phase isolation of methionine-containing peptides and subsequent matrix-assisted laser desorption/ionisation mass spectrometric detection of methionine- and of methionine-sulfoxide-containing peptides. *Rapid Commun. Mass Spectrom.* **17**: 1815-1824.

Guex N., Peitsch M. C. 1997. SWISS-MODEL and the Swiss-PDB Viewer: An environment for comparative protein modeling. *Electrophoresis* **18**: 2714-2723.

Hallem E. A., Ho M. G., Carlson J. R. 2004. The molecular basis of odor coding in the *Drosophila* antenna. *Cell* **117**: 965-979.

Hansen K. 1984. Discrimination and Production of Disparlure Enantiomers by the Gypsy Moth and the Nun Moth. *Physiol. Entomol.* **9**: 9-18.

Hargreaves W. R., Deamer D. W. 1978. Liposomes from Ionic, Single-Chain Amphiphiles. *Biochemistry* **17**: 3759-3768.

Hildebrand J. G. 1996. Olfactory control of behavior in moths: Central processing of odor information and the functional significance of olfactory glomeruli. *J. Comp. Physiol. A-Sens. Neural Behav. Physiol.* **178**: 5-19.

Hill C. A., Fox A. N., Pitts R. J., Kent L. B., Tan P. L., Chrystal M. A., Cravchik A., Collins F. H., Robertson H. M., Zwiebel L. J. 2002. G protein coupled receptors in *Anopheles gambiae*. *Science* **298**: 176-178.

Hoffmann E., Stroobant V. 2001. Mass spectrometry: principles and applications. Chichester: John Wiley & Sons.

Honson N., Johnson M. A., Oliver J. E., Prestwich G. D., Plettner E. 2003. Structure-activity studies with pheromone-binding proteins of the gypsy moth, *Lymantria dispar*. *Chem. Senses* **28**: 479-489.

Honson N. S., Plettner E. 2006. Disulfide connectivity and reduction in pheromone-binding proteins of the gypsy moth, *Lymantria dispar*. *Naturwissenschaften* **93**: 267-277.

Horst R., Damberger F., Luginbuhl P., Guntert P., Peng G., Nikonova L., Leal W. S., Wuthrich K. 2001. NMR structure reveals intramolecular regulation mechanism for pheromone binding and release. *Proc. Natl. Acad. Sci. U. S. A.* **98**: 14374-14379.

Hoyrup P., Davidsen J., Jorgensen K. 2001. Lipid membrane partitioning of lysolipids and fatty acids: Effects of membrane phase structure and detergent chain length. *J. Phys. Chem. B* **105**: 2649-2657.

Inkster J. A. H., Ling I., Honson N. S., Jacquet L., Gries R., Plettner E. 2005. Synthesis of disparlure analogues, using resolution on microcrystalline cellulose triacetate-I. *Tetrahedron: Asymmetry* **16**: 3773-3784.

Johnson B. H., Hecht M. H. 1994. Recombinant Proteins Can Be Isolated from *Escherichia Coli* Cells by Repeated Cycles of Freezing and Thawing. *Bio-Technology* **12**: 1357-1360.

Kaiser R., Metzka L. 1999. Enhancement of cyanogen bromide cleavage yields for methionyl-serine and methionyl-threonine peptide bonds. *Anal. Biochem.* **266**: 1-8.

Kaissling K. E., Priesner E. 1970. Olfactory Threshold of Silk Moths. *Naturwissenschaften* **57**: 23-28.

Kaissling K. E. 1977. Control of Insect Behavior *via* Chemoreceptor Organs. In: Chemical Control of Insect Behavior Theory and Application (Ed's H. H. Shorey, J. J. McKelvey Jr.). New York: John Wiley. p 45-65.

Kaissling K. E. 1987. R. H. Wright Lectures on Insect Olfaction. Burnaby, B.C., Canada: Simon Fraser University.

Kaissling K. E. 2004. Physiology of Pheromone Reception in Insects (an Example of Moths). *Anir* **6**: 73-91.

Kanicky J. R., Poniatowski A. F., Mehta N. R., Shah D. O. 2000. Cooperativity among molecules at interfaces in relation to various technological processes: Effect of chain length on the $pK_a$ of fatty acid salt solutions. *Langmuir* **16**: 172-177.

Kanicky J. R., Shah D. O. 2002. Effect of degree, type, and position of unsaturation on the $pK_a$ of long-chain fatty acids. *J. Colloid Interface Sci.* **256**: 201-207.

Keil T. A. 1999. Morphology and development of the peripheral olfactory organs. In: Insect Olfaction (Ed B. S. Hansson). Berlin: Springer. p 5-47.

Kim M. S., Repp A., Smith D. P. 1998. LUSH odorant-binding protein mediates chemosensory responses to alcohols in *Drosophila melanogaster*. *Genetics* **150**: 711-721.

Klein U. 1987. Sensillum-Lymph Proteins from Antennal Olfactory Hairs of the Moth *Antheraea Polyphemus* (Saturniidae). *Insect Biochemistry* **17**: 1193-1204.

Klusak V., Havias Z., Rulisek L., Vondrasek J., Svatos A. 2003. Sexual Attraction in the Silkworm Moth: Nature of Binding of bombykol in Pheromone Binding Protein - An *Ab Initio* Study. *Chem. Biol.* **10**: 331-340.

Koch A. L. 1996. What size should a bacterium be? A question of scale. *Annu. Rev. Microbiol.* **50**: 317-348.

Kowcun A., Honson N., Plettner E. 2001. Olfaction in the gypsy moth, *Lymantria dispar* - Effect of pH, ionic strength, and reductants on pheromone transport by pheromone-binding proteins. *J. Biol. Chem.* **276**: 44770-44776.

Krieger J., von Nickisch-Rosenegk E., Mameli M., Pelosi P., Breer H. 1996. Binding proteins from the antennae of *Bombyx mori*. *Insect Biochem. Mol. Biol.* **26**: 297-307.

Krieger J., Breer H. 1999. Olfactory reception in invertebrates. *Science* **286**: 720-723.

Krieger J., Strobel J., Vogl A., Hanke W., Breer H. 1999. Identification of a cyclic nucleotide- and voltage-activated ion channel from insect antennae. *Insect Biochem. Mol. Biol.* **29**: 255-267.

Krieger J., Raming K., Dewer Y. M. E., Bette S., Conzelmann S., Breer H. 2002. A divergent gene family encoding candidate olfactory receptors of the moth *Heliothis virescens*. *Eur. J. Neurosci.* **16**: 619-628.

Krieger M. J. B., Ross K. G. 2002. Identification of a major gene regulating complex social behavior. *Science* **295**: 328-332.

Kruse S. W., Zhao R., Smith D. P., Jones D. N. M. 2003. Structure of a specific alcohol-binding site defined by the odorant binding protein LUSH from *Drosophila melanogaster*. *Nat. Struct. Biol.* **10**: 694-700.

Lartigue A., Gruez A., Briand L., Pernollet J. C., Spinelli S., Tegoni M., Cambillau C. 2003. Optimization of crystals from nanodrops: crystallization and preliminary crystallographic study of a pheromone-binding protein from the honeybee *Apis mellifera* L. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **59**: 919-921.

Lartigue A., Gruez A., Spinelli S., Riviere S., Brossut R., Tegoni M., Cambillau C. 2003. The crystal structure of a cockroach pheromone-binding protein suggests a new ligand binding and release mechanism. *J. Biol. Chem.* **278**: 30213-30218.

Lartigue A., Riviere S., Brossut R., Tegoni M., Cambillau C. 2003. Crystallization and preliminary crystallographic study of a pheromone-binding protein from the cockroach *Leucophaea maderae*. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **59**: 916-918.

Lartigue A., Gruez A., Briand L., Blon F., Bezirard V., Walsh M., Pernollet J. C., Tegoni M., Cambillau C. 2004. Sulfur single-wavelength anomalous diffraction crystal structure of a pheromone-binding protein from the honeybee *Apis mellifera* L. *J. Biol. Chem.* **279**: 4459-4464.

Laue M., Steinbrecht R. A. 1997. Topochemistry of moth olfactory sensilla. *Int. J. Insect Morphol. Embryol.* **26**: 217-228.

Leal W. S., Nikonova L., Peng G. H. 1999. Disulfide structure of the pheromone binding protein from the silkworm moth, *Bombyx mori*. *FEBS Lett.* **464**: 85-90.

Leal W. S. 2000. Duality monomer-dimer of the pheromone-binding protein from *Bombyx mori*. *Biochem. Biophys. Res. Commun.* **268**: 521-529.

Lee D., Damberger F. F., Peng G. H., Horst R., Guntert P., Nikonova L., Leal W. S., Wuthrich K. 2002. NMR structure of the unliganded *Bombyx mori* pheromone-binding protein at physiological pH. *FEBS Lett.* **531**: 314-318.

Lee J. K., Strausfeld N. J. 1990. Structure, Distribution and Number of Surface Sensilla and Their Receptor-Cells on the Olfactory Appendage of the Male Moth *Manduca Sexta*. *J. Neurocytol.* **19**: 519-538.

Leite J. F., Hajivandi M. R., Diller T., Pope R. M. 2004. Removal of sodium and potassium adducts using a matrix additive during matrix-associated laser desorption/ionization time-of-flight mass spectrometric analysis of peptides. *Rapid Commun. Mass Spectrom.* **18**: 2953-2959.

Lima T. M., Cury-Boaventura M. F., Giannocco G., Nunes M. T., Curi1 R. 2006. Comparative toxicity of fatty acids on a macrophage cell line (J774). *Clin Sci (Lond)* Published June 1, **Online**.

Ling V. T., Eng M. L., Lee P.-J., Keck R. G., Keyt B. A., Canova-Davis E. 1996. Subtleties of peptide mapping in the analysis of protein pharmaceuticals. In: New Methods in Peptide Mapping for the Characterization of Proteins (Ed W. S. Hancock). Boca Raton: CRC Press. p 1-30.

Loboda A. V., Chernushevich I. V. 2005. Investigation of the mechanism of matrix adduct formation in MALDI at elevated pressure. *Int. J. Mass Spectrom.* **240**: 101-105.

Lopez-Nicolas J. M., Bru R., Sanchez-Ferrer A., Garcia-Carmona F. 1995. Use of 'Soluble Lipids' for Biochemical Processes: Linoleic Acid-Cyclodextrin Inclusion Complexes in Aqueous Solution. *Biochem. J.* **308**: 151-154.

McKenna M. P., Hekmatscafe D. S., Gaines P., Carlson J. R. 1994. Putative *Drosophila* Pheromone-Binding Proteins Expressed in a Subregion of the Olfactory System. *J. Biol. Chem.* **269**: 16340-16347.

Miller J. R., Mori K., Roelofs W. L. 1977. Gypsy Moth (Lepidoptera-Lymantriidae) Field Trapping and Electroantennogram Studies with Pheromone Enantiomers. *J. Insect Physiol.* **23**: 1447-1453.

Miller K. E., Bruno T. J. 2003. Isothermal Kovats retention indices of sulfur compounds on a poly(5% diphenyl-95% dimethylsiloxane) stationary phase. *J. Chromatogr. A* **1007**: 117-125.

Milzani A., Rossi R., di Simplicio P., Giustarini D., Colombo R., Dalledonne I. 2000. The oxidation produced by hydrogen peroxide on Ca-ATP-G-actin. *Protein Science* **9**: 1774-1782.

Mohanty S., Zubkov S., Campos-Olivas R. 2003. Letter to the Editor: $^{1}H$, $^{13}C$ and $^{15}N$ backbone assignments of the pheromone binding protein from the silk moth *Antheraea polyphemus* (ApolPBP). *J. Biomol. NMR* **27**: 393-394.

Mohanty S., Zubkov S., Gronenborn A. M. 2004. The solution NMR structure of *Antheraea polyphemus* PBP provides new insight into pheromone recognition by pheromone-binding proteins. *J. Mol. Biol.* **337**: 443-451.

Mohl C., Breer H., Krieger J. 2002. Species-specific pheromonal compounds induce distinct conformational changes of pheromone binding protein subtypes from *Antheraea polyphemus*. *Invertebr. Neurosci.* **4**: 165-174.

Mukhopadhyay S., Nayak P. K., Udgaonkar J. B., Krishnamoorthy G. 2006. Characterization of the formation of amyloid protofibrils from barstar by mapping residue-specific fluorescence dynamics. *J. Mol. Biol.* **358**: 935-942.

Nardi J. B., Miller L. A., Walden K. K. O., Rovelstad S., Wang L., Frye J. C., Ramsdell K., Deem L. S., Robertson H. M. 2003. Expression patterns of odorant-binding proteins in antennae of the moth *Manduca sexta*. *Cell Tissue Res.* **313**: 321-333.

Nation J. L. 2002. Insect Physiology and Biochemistry. Boca Raton: CRC Press LLC. p. 389-424.

Newcomb R. D., Sirey T. M., Rassam M., Greenwood D. R. 2002. Pheromone binding proteins of *Epiphyas postvittana* (Lepidoptera : Tortricidae) are encoded at a single locus. *Insect Biochem. Mol. Biol.* **32**: 1543-1554.

Oae S. 1991. Organic Sulfur Chemistry: Structure and Mechanism. Boca Raton: CRC Press, Inc. 1-30.

Oguri E., Steele J. E. 2003. A novel function of cockroach (*Periplaneta americana*) hypertrehalosemic hormone: translocation of lipid from hemolymph to fat body. *Gen. Comp. Endocrinol.* **132**: 46-54.

Pauling L. 1949. On the Stability of the S8 Molecule and the Structure of Fibrous Sulfur. *Proc. Natl. Acad. Sci. U. S. A.* **35**: 495-499.

Petersen M. T. N., Jonson P. H., Petersen S. B. 1999. Amino acid neighbours and detailed conformational analysis of cysteines in proteins. *Protein Engineering* **12**: 535-548.

Pikielny C. W., Hasan G., Rouyer F., Rosbash M. 1994. Members of a Family of *Drosophila* Putative Odorant-Binding Proteins Are Expressed in Different Subsets of Olfactory Hairs. *Neuron* **12**: 35-49.

Plettner E., Lazar J., Prestwich E. G., Prestwich G. D. 2000. Discrimination of pheromone enantiomers by two pheromone binding proteins from the gypsy moth *Lymantria dispar*. *Biochemistry* **39**: 8953-8962.

Prestwich G. D. 1993. Bacterial Expression and Photoaffinity Labeling of a Pheromone Binding Protein. *Protein Science* **2**: 420-428.

Riviere S., Lartigue A., Quennedey B., Campanacci V., Farine J. P., Tegoni M., Cambillau C., Brossut R. 2003. A pheromone-binding protein from the cockroach *Leucophaea maderae*: cloning, expression and pheromone binding. *Biochem. J.* **371**: 573-579.

Robertson H. M., Warr C. G., Carlson J. R. 2003. Molecular evolution of the insect chemoreceptor gene superfamily in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U. S. A.* **100**: 14537-14542.

Rogers M. E., Sun M., Lerner M. R., Vogt R. G. 1997. SNMP-1, a novel membrane protein of olfactory neurons of the silk moth *Antheraea polyphemus* with homology to the CD36 family of membrane proteins. *J. Biol. Chem.* **272**: 14792-14799.

Rogers M. E., Krieger J., Vogt R. G. 2001. Antennal SNMPs (sensor neuron membrane proteins) of lepidoptera define a unique family of invertebrate CD36-like proteins. *J. Neurobiol.* **49**: 47-61.

Rogers M. E., Steinbrecht R. A., Vogt R. G. 2001. Expression of SNMP-1 in olfactory neurons and sensilla of male and female antennae of the silkmoth *Antheraea polyphemus*. *Cell Tissue Res.* **303**: 433-446.

Rothemund S., Liou Y. C., Davies P. L., Sonnichsen F. D. 1997. Backbone structure and dynamics of a hemolymph protein from the mealworm beetle *Tenebrio molitor*. *Biochemistry* **36**: 13791-13801.

Rothemund S., Liou Y. C., Davies P. L., Krause E., Sonnichsen F. D. 1999. A new class of hexahelical insect proteins revealed as putative carriers of small hydrophobic ligands. *Struct. Fold. Des.* **7**: 1325-1332.

Sakurai T., Nakagawa T., Mitsuno H., Mori H., Endo Y., Tanoue S., Yasukochi Y., Touhara K., Nishioka T. 2004. Identification and functional characterization of a sex pheromone receptor in the silkmoth *Bombyx mori*. *Proc. Natl. Acad. Sci. U. S. A.* **101**: 16653-16658.

Sandler B. H., Nikonova L., Leal W. S., Clardy J. 2000. Sexual attraction in the silkworm moth: structure of the pheromone-binding-protein-bombykol complex. *Chem. Biol.* **7**: 143-151.

Scaloni A., Monti M., Angeli S., Pelosi P. 1999. Structural analysis and disulfide-bridge pairing of two odorant-binding proteins from *Bombyx mori*. *Biochem. Biophys. Res. Commun.* **266**: 386-391.

Scheffler H. J. 1975. Structure of Antennae in Wz and Zz Intersexes of Gypsy Moth *Lymantria Dispar* L (Insecta, Lepidoptera). *Zeitschrift Fur Morphologie Der Tiere* **80**: 203-227.

Schneider D., Kasang G., Kaissling K. E. 1968. Determination of Scent Threshold of *Bombyx Mori* with Tritium-Labeled Bombycol. *Naturwissenschaften* **55**: 395-&.

Shaked Z., Szajewski R. P., Whitesides G. M. 1980. Rates of Thiol-Disulfide Interchange Reactions Involving Proteins and Kinetic Measurements of Thiol $pK_a$ Values. *Biochemistry* **19**: 4156-4166.

Sharrow S. D., Novotny M. V., Stone M. J. 2003. Thermodynamic analysis of binding between mouse major urinary protein-I and the pheromone 2-sec-butyl-4,5-dihydrothiazole. *Biochemistry* **42**: 6302-6309.

Smith B. J. 2002. Chemical Cleavage of Proteins at Methionyl-X Peptide Bonds. In: The Protein Protocols Handbook, 2nd Edition (Ed J. M. Walker). Totowa, New Jersey: Humana Press. p 485-491.

Sreerama N., Venyaminov S. Y., Woody R. W. 2000. Estimation of protein secondary structure from circular dichroism spectra: Inclusion of denatured proteins with native proteins in the analysis. *Anal. Biochem.* **287**: 243-251.

Srinivasan N., Sowdhamini R., Ramakrishnan C., Balaram P. 1990. Conformations of disulfide bridges in proteins. *International Journal of Peptide Protein Research* **36**: 147-155.

Steinbrecht R. A., Ozaki M., Ziegelberger G. 1992. Immunocytochemical Localization of Pheromone-Binding Protein in Moth Antennae. *Cell Tissue Res.* **270**: 287-302.

Steiner T. 2002. The Hydrogen Bond in the Solid State. *Angewandte Chemie International Edition* **41**: 48-76.

Stengl M., Hatt H., Breer H. 1992. Peripheral Processes in Insect Olfaction. *Annu. Rev. Physiol.* **54**: 665-681.

Stengl M. 1994. Inositol-Trisphosphate-Dependent Calcium Currents Precede Cation Currents in Insect Olfactory Receptor Neurons *in-Vitro*. *J. Comp. Physiol. A-Sens. Neural Behav. Physiol.* **174**: 187-194.

Stengl M., Ziegelberger G., Boekhoff I., Krieger J. 1999. Perireceptor Events and Transduction Mechanisms in Insect Olfaction. In: Insect Olfaction (Ed B. S. Hansson). Berlin: Springer. p 50-66.

Stortkuhl K. F., Kettler R. 2001. Functional analysis of an olfactory receptor in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U. S. A.* **98**: 9381-9385.

Tarr G. E., Crabb J. W. 1983. Reverse-Phase High-Performance Liquid Chromatography of Hydrophobic Proteins and Fragments Thereof. *Anal. Biochem.* **131**: 99-107.

Tegoni M., Campanacci V., Cambillau C. 2004. Structural aspects of sexual attraction and chemical communication in insects. *Trends Biochem.Sci.* **29**: 257-264.

Vogt R. G., Riddiford L. M. 1981. Pheromone Binding and Inactivation by Moth Antennae. *Nature* **293**: 161-163.

Vogt R. G., Riddiford L. M., Prestwich G. D. 1985. Kinetic-Properties of a Sex Pheromone-Degrading Enzyme - the Sensillar Esterase of *Antheraea Polyphemus*. *Proc. Natl. Acad. Sci. U. S. A.* **82**: 8827-8831.

Vogt R. G., Kohne A. C., Dubnau J. T., Prestwich G. D. 1989. Expression of Pheromone Binding-Proteins During Antennal Development in the Gypsy Moth *Lymantria Dispar*. *J. Neurosci.* **9**: 3332-3346.

Vogt R. G., Prestwich G. D., Lerner M. R. 1991. Odorant-Binding-Protein Subfamilies Associate with Distinct Classes of Olfactory Receptor Neurons in Insects. *J. Neurobiol.* **22**: 74-84.

Vogt R. G., Rybczynski R., Lerner M. R. 1991. Molecular Cloning and Sequencing of General Odorant Binding Proteins GOBP1 and GOBP2 from the Tobacco Hawk Moth *Manduca Sexta* - Comparisons with Other Insect OBPs and Their Signal Peptides. *J. Neurosci.* **11**: 2972-2984.

Vogt R. G., Callahan F. E., Rogers M. E., Dickens J. C. 1999. Odorant binding protein diversity and distribution among the insect orders, as indicated by LAP, an OBP-related protein of the true bug *Lygus lineolaris* (Hemiptera, Heteroptera). *Chem. Senses* **24**: 481-495.

Vosshall L. B., Amrein H., Morozov P. S., Rzhetsky A., Axel R. 1999. A spatial map of olfactory receptor expression in the *Drosophila* antenna. *Cell* **96**: 725-736.

Vosshall L. B., Wong A. M., Axel R. 2000. An olfactory sensory map in the fly brain. *Cell* **102**: 147-159.

Wang J. W., Wong A. M., Flores J., Vosshall L. B., Axel R. 2003. Two-photon calcium imaging reveals an odor-evoked map of activity in the fly brain. *Cell* **112**: 271-282.

Wang Y. M., Lin D. S., Bolewicz L., Connor W. E. 2006. The predominance of polyunsaturated fatty acids in the butterfly *Morpho peleides* before and after metamorphosis. *J. Lipid Res.* **47**: 530-536.

Wetzel C. H., Behrendt H. J., Gisselmann G., Stortkuhl K. F., Hovemann B., Hatt H. 2001. Functional expression and characterization of a *Drosophila* odorant receptor in a heterologous cell system. *Proc. Natl. Acad. Sci. U. S. A.* **98**: 9377-9380.

Wiseman T., Williston S., Brandts J. F., Lin L. N. 1989. Rapid Measurement of Binding Constants and Heats of Binding Using a New Titration Calorimeter. *Anal. Biochem.* **179**: 131-137.

Wojtasek H., Leal W. S. 1999. Conformational Change in the Pheromone-binding Protein from *Bombyx mori* Induced by pH and by Interaction with Membranes. *J. Biol. Chem.* **274**: 30950 - 30955.

Wu J., Watson J. T. 1997. A novel methodology for assignment of disulfide bond pairings in proteins. *Protein Science* **6**: 391-398.

Xu P. X., Atkinson R., Jones D. N. M., Smith D. P. 2005. *Drosophila* OBP LUSH is required for activity of pheromone-sensitive neurons. *Neuron* **45**: 193-200.

Zhang S. G., Maida R., Steinbrecht R. A. 2001. Immunolocalization of odorant-binding proteins in noctuid moths (Insecta, Lepidoptera). *Chem. Senses* **26**: 885-896.

Zheng C. J., Yoo J. S., Lee T. G., Cho H. Y., Kim Y. H., Kim W. G. 2005. Fatty acid synthesis is a target for antibacterial activity of unsaturated fatty acids. *FEBS Lett.* **579**: 5157-5162.

Zhou J. J., Zhang G. A., Huang W. S., Birkett M. A., Field L. M., Pickett J. A., Pelosi P. 2004. Revisiting the odorant-binding protein LUSH of *Drosophila melanogaster:* evidence for odour recognition and discrimination. *FEBS Lett.* **558**: 23-26.

Ziegelberger G., Vandenberg M. J., Kaissling K. E., Klumpp S., Schultz J. E. 1990. Cyclic-GMP Levels and Guanylate Cyclase Activity in Pheromone-Sensitive Antennae of the Silkmoths *Antheraea Polyphemus* and *Bombyx Mori. J. Neurosci.* **10**: 1217-1225.

Ziegelberger G. 1995. Redox-Shift of the Pheromone-Binding Protein in the Silkmoth *Antheraea Polyphemus. Eur. J. Biochem.* **232**: 706-711.

Zubkov S., Gronenborn A. M., Byeon I. J. L., Mohanty S. 2005. Structural consequences of the pH-induced conformational switch in *A. polyphemus* pheromone-binding protein: Mechanisms of ligand release. *J. Mol. Biol.* **354**: 1081-1090.