Disulfide structure of the pheromone binding protein from the silkworm moth, *Bombyx mori*

Walter S. Leal*, Larisa Nikonova, Guihong Peng

Laboratory of Chemical Prospecting, National Institute of Sericultural and Entomological Science, 1-2 Ohwashi, Tsukuba 305-8634, Japan

Received 1 September 1999

Edited by Richard Cogdell

Abstract Disulfide bond formation is the only known post-translational modification of insect pheromone binding proteins (PBPs). In the PBPs from moths (Lepidoptera), six cysteine residues are highly conserved at positions 19, 50, 54, 97, 108 and 117, but to date nothing is known about their respective linkage or redox status. We used a multiple approach of enzymatic digestion, chemical cleavage, partial reduction with Tris-(2-carboxyethyl)phosphine, followed by digestion with endoproteinase Lys-C to determine the disulfide connectivity in the PBP from *Bombyx mori* (BmPBP). Identification of the reaction products by on-line liquid chromatography-electrospray ionization mass spectrometry (LC/ESI-MS) and protein sequencing supported the assignment of disulfide bridges at Cys-19-Cys-54, Cys-50-Cys-108 and Cys-97-Cys-117. The disulfide linkages were identical in the protein obtained by periplasmic expression in *Escherichia coli* and in the native BmPBP.

Key words: Enzymatic digestion; Disulfide linkage; ESI-MS; TCEP; Carboxyamidomethyl cysteine; PBP

1. Introduction

Pheromone binding proteins (PBPs) participate in the early olfactory processing of chemical signals in insect antennae by carrying the airborne hydrophobic pheromone molecules to the olfactory receptors, which are believed to be located in the dendritic surfaces of olfactory receptor neurons (ORNs). The transduction of the chemical signals (pheromones) into the brain language – electrical signals – is triggered by the binding of cognate ligands to olfactory receptors. PBPs are synthesized in auxiliary (trichogen and tormogen) cells [1,2] and secreted into the sensillar lymph surrounding the ORN dendrites. It has been suggested that PBPs act not merely as passive, non-selective carrier molecules, but that their role may involve serving as selective filters for molecules relevant to a given sensillum [3] and, consequently, contributing to the exquisite sensitivity and selectivity of insect pheromone detectors.

Since the discovery of a PBP in the giant silkworm moth, *Antheraea polyphemus* [4], homologous proteins have been identified in many species from several orders. A number of PBPs and odorant binding proteins (OBPs) have been cloned [5–16] and a few of them have been expressed in heterologous systems [17–22]. The whole family shows six highly conserved cysteine residues, and the formation of disulfide bonds is the only posttranslational modification of these proteins. PBP from the silkworm moth, *Bombyx mori* (hereafter BmPBP), as do other lepidopteran PBPs, contains six cysteine residues located at positions 19, 50, 54, 97, 108, and 117 (Fig. 1A). Previously, we have determined that all six cysteines form disulfide bridges in a recombinant BmPBP [22] so, strictly speaking, they are cystines (cysteines linked by disulfide bonds).

The formation of disulfide bonds, which is an important posttranslational modification of secretory proteins [23], plays a pivotal role in stabilizing and maintaining the three-dimensional structures of various biologically important molecules [24] including enzymes, protease inhibitors, plasma proteins, neurotoxins, and hormones. Determination of disulfide connectivity is an important step in their structural characterization. However, assignment of these disulfide bonds is increasingly difficult as the number of cystine residues increases. While two disulfide bonds give rise to three possible combinations, there are 15 and 105 possible isomers with three and four disulfide bonds, respectively. Moreover, disulfide bond exchange (scrambling) may take place during derivatization or cleavage of the protein. Earlier attempts to characterize disulfide connectivity in moth PBPs were unrewarding [25], probably because these proteins are not amenable to enzymatic degradation under non-reducing conditions. Using a combination of chemical and enzymatic cleavages, partial reduction and characterization of nascent peptides by electrospray ionization mass spectrometry and Edman degradation, we were able to determine for the first time the disulfide connectivity in an insect PBP. In addition, our findings confirm that native and recombinant BmPBPs are indistinguishable in their disulfide linkages.

2. Materials and methods

2.1. Materials

Endoproteinase Lys-C, chymotrypsin, proteinase K, elastase and pepsin were from Boehringer-Mannheim (Germany). Cyanogen bromide was purchased from Nakarai Tesque (Japan). Tris-(2-carboxyethyl)phosphine was from Fluka (Switzerland). Other chemicals were from Wako (Japan).

2.2. Protein sources

The recombinant protein was expressed in *Escherichia coli*, using the periplasmic expression system previously described [22]. Instead of osmotic shock, the protein was released from the harvested cells by freeze and thaw cycles [26]. The native protein was isolated from extracts of 100 male antennae, following our previous protocol [22]. In the last step, the isolated protein was further purified by HPLC.

2.3. Analytical procedures

Electrospray ionization mass spectra (ESI-MS) were obtained on a Hewlett-Packard on-line liquid chromatograph-mass spectrometer

*Corresponding author. Fax: (81)-298-38 6028.
E-mail: leal@nises.affrc.go.jp

PII: S0014-5793(99)01683-X

(LC/ESI-MS) system. HPLC separations were achieved on a Zorbax 300CB-C8 column (2.1×150 mm; 5 μm) using water and acetonitrile with 2% acetic acid as modifier in both solvents and a gradient from 0 to 80% acetonitrile in 15 min at a flow rate of 0.2 ml/min. Column effluent was monitored by UV at 214, 230 and 280 nm and MS on an HP 1100 series LC/MSD. The MSD parameters were: collision-induced dissociation voltage (fragmentor), 120 V; drying gas flow, 6 l/min; nebulizer pressure, 45 psig; nebulizer temperature, 300°C; capillary voltage, 3500 V. Throughout the paper, the molecular masses are given without experimental error because it was smaller than ±1 Da for the molecular species described in the text. Preparative HPLC was carried out on an HP series 1100 HPLC system with a Zorbax 300CB-C18 column (2.1×150 mm; 5 μm). Mobile phases were: A, 0.1% TFA in water and B, 0.1% TFA in acetonitrile; a linear gradient of 0–80% in 45 min was used at a flow rate of 0.2 ml/min. The column effluent was monitored at 214, 230 and 280 nm. Peak fractions were collected manually in ice-cold Eppendorf tubes. N-terminal amino acid sequences were obtained on a Hewlett-Packard Protein Sequencer Model 241 with PTH derivatives separated on an HP series 1100 HPLC system. Fractions collected from HPLC were directly loaded onto the sequencer column without any treatment. The standard retention time for the carboxyamidomethyl cysteine (Cam-Cys) was determined following the HP protocol [27].

2.4. Reactions

Digestion by endoproteinase Lys-C was carried out at 37°C in Tris-HCl 25 mM, pH 6.8, 1 mM EDTA, 5% acetonitrile and with a protein to enzyme ratio of 10:1. Chymotrypsin digestions of HPLC collected peaks were carried out in Tris-HCl, 100–250 mM, pH 6.8, $CaCl_2$ 10 mM using 0.05 to 0.25 μg of enzyme. Fractions were dried in a Speed Vac, reconstituted in buffer and incubated at 25°C. Cyanogen bromide reactions were carried out with 1–3 mg of CNBr per 30 μg of protein either in 70% formic acid or 70% TFA. The solution was flushed with argon and incubated in the dark for ca. 24 h. Reduction of BmPBP with TCEP followed by alkylation (iodoacetamide) was done according to a reported method [28]. Typically, 6.3 nmol of BmPBP was reacted with 10 mM TCEP in 0.2 M sodium acetate, pH 3, at 25°C for 30 min. The crude reaction mixture was subjected to preparative HPLC, and the collected fractions were added with rapid mixing to 2.2 M iodoacetamide [28].

3. Results and discussion

Our preliminary attempts to digest the unreduced recombinant BmPBP with chymotrypsin, elastase and proteinase K under standard conditions were unrewarding, probably because of the tight structure of the protein [25] with the disulfide bridges intact. At most, we were able to obtain two fragments with large molecular masses (>14 kDa), which were of little value for the assignment of disulfide bridges. The yield was even lower when the digestion was carried out at pH < 7, which is necessary to avoid scrambling the disulfide bridges. Under alkaline conditions, thiolate anion (RS^-) attacks the disulfide bridges and destroys the original bridges [29].

Prolonged incubation time (7–10 days) of BmPBP with endoproteinase Lys-C (pH 8) with a daily addition of fresh enzyme gave encouraging results because we detected small fragments by LC/ESI-MS. We decided to optimize the digestion under more forced conditions and at the desired pH (6.8). Trials to improve the digestion with denaturing agents showed that acetonitrile (5%) gave better results than SDS, urea, and guanidine hydrochloride. Although the amount of enzyme recommended by the supplier is 1/100 to 1/20 of the protein by weight, we achieved satisfactory digestion of BmPBP using a larger ratio of enzyme to protein (1/10) (Fig. 2).

Analysis of the reaction mixture on an HPLC coupled to an electrospray ionization mass spectrometer (ESI-MS) showed, in addition to other expected fragments, two peaks (5 and 6 at 19. 6 and 20.1 min, respectively) (Fig. 2), which account for

A

```
  1   *   10    *   20    *   30    *   40
  1 SQEVMKNLSLNFGKALDECKKEMTLTDAINEDFYNFWKEG
      *   50    *   60    *   70    *   80
 41 YEIKNRETGCAIMCLSTKLNMLDPEGNLHHGNAMEFAKKH
      *   90    *  100    *  110    *  120
 81 GADETMAQQLIDIVHGCEKSTPANDDKCIWTLGVATCFKA
      *  130    *  140
121 EIHKLNWAPSMDVAVGEILAEV
```

B — Endoproteinase Lys-C

Entry	MM(Da)	Location	Chain Data
#1	720.3	1–6	SQEVMK
#2	891.5	7–14	NLSLNFGK
#3	677.3	15–20	ALDECK
#4	146.1	21	K
#5	2137.0	22–38	EMTLTDAINEDFYNFWK
#6	737.4	39–44	EGYEIK
#7	1525.7	45–58	NRETGCAIMCLSTK
#8	2238.0	59–78	LNMLDPEGNLHHGNAMEFAK
#9	146.1	79	K
#10	2195.0	80–99	HGADETMAQQLIDIVHGCEK
#11	846.4	100–107	STPANDDK
#12	1340.7	108–119	CIWTLGVATCFK
#13	596.3	120–124	AEIHK
#14	1914.0	125–142	LNWAPSMDVAVGEILAEV

C — Cyanogen bromide (homoserine lactone)

Entry	MM(Da)	Location	Chain Data
#15	544.2	1–5	SQEVM
#16	2020.0	6–23	KNLSLNFGKALDECKKEM
#17	3524.6	24–53	TLTDAINEDFYNFWKEGYEIKNRE TGCAIM
#18	860.4	54–61	CLSTKLNM
#19	1355.6	62–74	LDPEGNLHHGNAM
#20	1314.6	75–86	EFAKKHGADETM
#21	4936.4	87–131	AQQLIDIVHGCEKSTPANDDKCIWTLG VATCFKAEIHKLNWAPSM
#22	1113.6	132–142	DVAVGEILAEV

Fig. 1. Primary structure of the pheromone binding protein from *B. mori* (A) and the predicted fragments by the digestion with endoproteinase Lys-C (B) and chemical cleavage with cyanogen bromide (C). Cysteine residues at positions 19, 50, 54, 97, 108 and 117 are underlined and in boldface (A, B, C), whereas the sites for cleavage with endoproteinase Lys-C are underlined; Met and Trp are highlighted in italics (A).

the peptides containing all three disulfide bridges. Peak 6 gave the molecular mass of 5733 Da, which was suggested to be to the fragments #3, #7, #10 and #12 (Fig. 1) linked by disulfide bridges (calculated, 5732.7 Da). We isolated that peak by preparative HPLC and applied the material to a protein sequencer. The amino acid residues observed in the first three cycles (cycle 1: A, N, H; 2: L, R, G, I; 3: D, E, A, W) supported the assignment. The molecular mass (5863 Da) for peak 5 suggested that it had a composition similar to that of peak 6, but with an extra Lys (calculated, 5861 Da). The relative ratio of the two peaks changed with the time of incubation; the longer the incubation time the larger was the relative amount of the later eluting peak. Moreover, amino acid sequence data showed that the material isolated from peak 6 had an extra Lys at the seventh position. These findings indicate that cleavage at Lys-21 is slower than at Lys-22 so that the peak 5 is similar to peak 6, but with an extra Lys in one of the peptides (ALDECKK), which was linked to the other peptides (#7, #10 and #12) by disulfide bridges. A similar pattern of complex bridged fragments was observed in one of the products of digestion with trypsin at pH 6.8. However, the enzyme did not cleave at Lys-99 and Lys-107 (Fig. 1B) and generated one of the peptides (residues 80–119) with three cystine residues, Cys-97, Cys-108, and Cys-117, which was bridged to fragments #3 and #7. In other words, trypsin

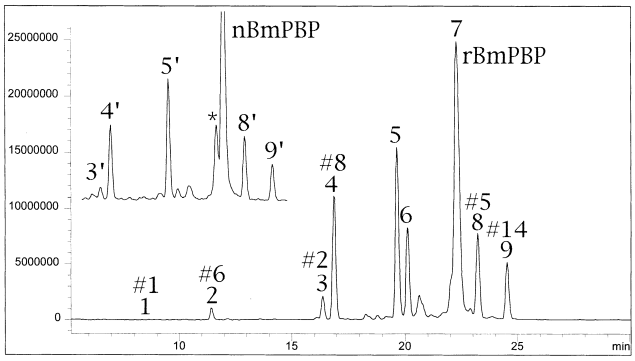

Fig. 2. Total ion current (TIC) profile of endoproteinase Lys-C digest of the recombinant (24 h) and native (insert, 13 h) BmPBP separated on a Zorbax C8 column. Fragment numbers are the same as in Fig. 1B. A minor impurity from the native protein (*) remained undigested.

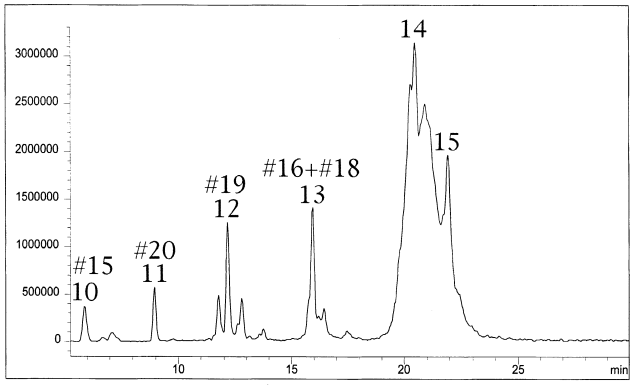

Fig. 3. TIC profile of the reaction of rBmPBP with cyanogen bromide. Fragment numbers are the same as in Fig. 1C.

gave a more complex pattern of bridged peptides. In addition, the reaction was too slow (even with a protein to enzyme ratio of 10:1); it took 10 days to reach the same level of digestion obtained in 24 h with endoproteinase Lys-C (Fig. 2).

The digestion of BmPBP with endoproteinase Lys-C gave us some insight into the location of the disulfide bridges. Given the fact that no dipeptide was observed and that the bridged fragments have all the three disulfide linkages, it became clear that the peptides containing only one cystine residue (fragments #3 and #10) were not linked to each other but rather to one of the two peptides with two cysteine residues (#7 and #12). Also, the peptides with two cystine residues were linked to each other by only one disulfide bridge. Considering that breaking the protein (or the nascent peptides) in between cystine residues would generate smaller bridged fragments, we tried to cleave BmPBP at Met and/or Trp residues. These residues are located between Cys-50 and Cys-54 (Met-53) and Cys-108 and Cys-117 (Trp-110) (Fig. 1B).

Attempts to achieve tryptophanyl cleavage of BmPBP with *o*-iodosobenzoic acid, according to a reported method [30], were unsuccessful. The crude reaction mixture was not clean enough to be analyzed directly by LC/ESI-MS, the yield was very low and oxidation of other residues generated a complicated mass spectrum profile. Analysis of the first eluting peptide (preparative HPLC, 23.3 min) emphasizes this view. MS data (observed, 1516 Da) and amino acid sequence (cycle 1: A; 2: P; 3: S) suggest that this was a fragment generated by cleavage at Trp-127 (Fig. 1A), i.e., APSMDVAVGEILAEV (residues 128–142). The calculated molecular mass for the intact peptide is 1499.8 Da, the difference in mass (16.2 Da) indicates further oxidation in the peptide. More importantly, larger fragments gave clusters of peaks differing in their molecular masses by 16. Because these complications could be misleading when analyzing bridged fragments (and the oxidation could destroy the original linkage), we did not proceed with this chemical cleavage.

Cleavage of Met with cyanogen bromide in formic acid gave a low yield because the Met-Thr peptide bond undergoes an N- to O-acyl shift [31–33]; the yield was higher in 70% TFA [31,32]. Clearly, the CNBr-derived peptides (Fig. 3) were homoserine lactones. Peak 10 (5.6 min) was assigned to fragment #15 (Fig. 1C, residues 1–5) (observed molecular mass, 544 Da; calculated 544.2 Da). Peak 11 (8.9 min) gave a molecular mass of 1315 Da, which was assigned to fragment #20 (residues 75–86) (Fig. 1C) (calculated for homoserine lactone, 1314.6 Da; homoserine, 1332.6 Da). Likewise, peak 12 (12.2 min) had a molecular mass of 1356 Da and was assigned to the homoserine lactone fragment #19 (residues 62–74; calculated for homoserine lactone, 1355.6 Da; homoserine, 1373.6 Da). Peak 13 (15.9 min) demonstrated the occurrence of a disulfide bridge between Cys-19 and Cys-54. The observed molecular mass of 2878 Da suggested that this was due to fragment #16 bridged to fragment #18 (calculated, 2878.4 Da). The amino acid sequence of the fragment isolated by HPLC (cycle 1: K; 2: N, L; 3: L, S; 4: S, T; 5: L, K) confirmed the assignment. These results indicate that Cys-19 and Cys-54 are covalently linked. The cluster of peaks at 19–23 min was separated by preparative HPLC under more aqueous conditions, and each fraction was monitored by LC/ESI-MS. The longer retention time peak (peak 15; 21.9 min) was due to fragments #17 and #21 (calculated 8459 Da; observed, 8457 Da), but this does not allow us to assign the linkage of Cys-50 either to Cys-108 or Cys-117. The Cys-50-Cys-97 linkage is ruled out on the basis of the lack of a tripeptide (fragments #3, #7 and #10) after digestion with endoproteinase Lys-C. The major fragment in the original sample (peak 14, Fig. 3) was due to a larger fragment (#16 and #17) bridged to fragments #18 and #21. Fragments #16 and #17 are linked by an ester bond (as opposed to the original peptide bond)

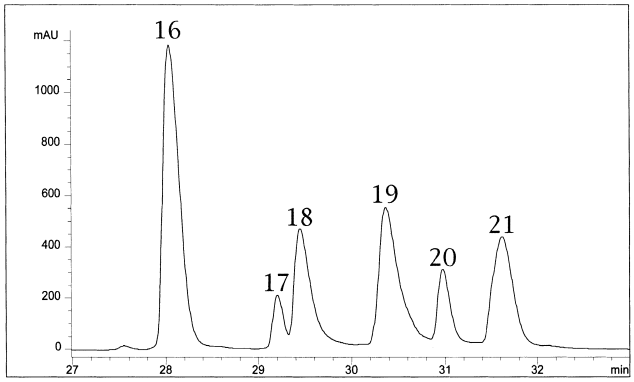

Fig. 4. Partial reduction of BmPBP with TCEP. HPLC trace (UV, 280 nm) of the effluent from a Zorbax C18 column. The partially reduced protein peaks appear in between the intact protein (peak 16) and the completely reduced form (peak 21).

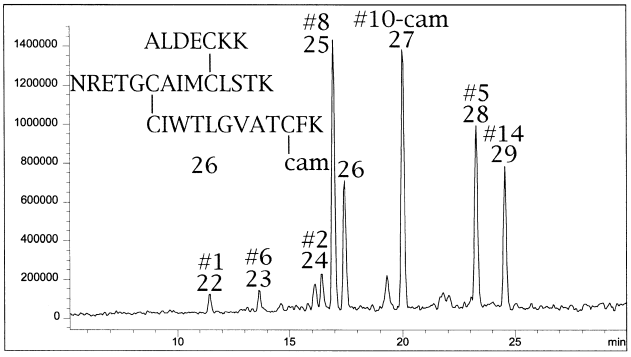

Fig. 5. TIC of endoproteinase Lys-C digest of reduced and alkylated BmPBP. Peak 26 was demonstrated to be a tripeptide linked by two disulfide bonds (insert).

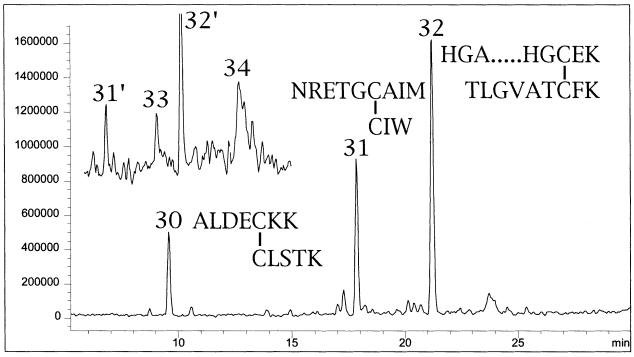

Fig. 7. TIC of the chymotrypsin digest of a fragment generated by previous digestion of BmPBP with endoproteinase Lys-C (peak 5, Fig. 2). The digest of the peptide originated from the native protein (peak 5′, Fig. 2) is also shown (insert). The reaction times were 2 h and 70 min, respectively. Peak 33 from the native protein was also observed in the beginning of the same reaction with the recombinant protein, but after 2 h it became a minor peak.

due to the N- to O-acyl shift (the Thr-24 OH replaces its NH in the original peptide bond) [33]. Attempts to cleave BmPBP at the tryptophanyl residues by treatment with cyanogen bromide in formic and heptafluorobutyric acids [34,35] did not generate any cleavage of tryptophan but gave a profile almost identical to that obtained by cleavage with CNBr in 70% TFA (Fig. 3).

Next, we explored the possibility of obtaining selective cleavage of the four bridged peptides (peak 6, Fig. 2) with Tris-(2-carboxyethyl)phosphine (TCEP) [36]. The reaction was too fast and complete reduction was observed in 12 min at 25°C. Within 2 min reaction time, one bridged fragment was observed, but this served only to corroborate that Cys-19 and Cys-54 are bridged. A peak at 15.2 min gave a molecular mass of 2201 Da, which was assigned to fragments #3 and #7 (Fig. 1A) linked by a disulfide bond (calculated 2201 Da). This dipeptide underwent some fragmentation as demonstrated by the detection of fragment #7 at 16.3 min (fragment #3 probably eluted with the solvent front). In addition, fragments #10 and #12 coeluted at 21.1 min. The occurrence of two peptides within the same peak was clearly demonstrated by deconvolution of the two clusters of peaks in the mass spectrum (fragment #10: observed 2195 Da; calculated, 2195 Da; fragment #12: observed, 1341 Da; calculated 1340.7 Da). We hypothesized that due to the tight structure of the intact protein, reduction would be slower with BmPBP than with peptides generated by enzymatic digestion. In fact, reduction at 25°C was slow enough; less than half of the starting material was partially reduced in 30 min. A prolonged reaction time (>2 h), however, generated the fully reduced protein. We isolated the reaction products of a larger sample by preparative HPLC (Fig. 4) and alkylated the free thiols with iodoacetamide under kinetically forced conditions [28]. The carboxyamidomethylated proteins derived from each nascent reduced protein (Fig. 4) were analyzed by LC/ESI-MS. Two peaks (ratio ca. 4:1) were obtained after alkylation of the single HPLC peak 18 (29.4 min; Fig. 4). Both alkylated products showed the incorporation of two carboxyamidomethyl groups. They showed the same molecular mass (15 993 Da; calculated, 15 993 Da), but had different retention times and mass profiles indicating that they were di-alkylated isomers. These data suggest that the two disulfide bonds remaining in the two proteins were different. On the other hand, the molecular mass of peak 19 (30.3 min) after alkylation

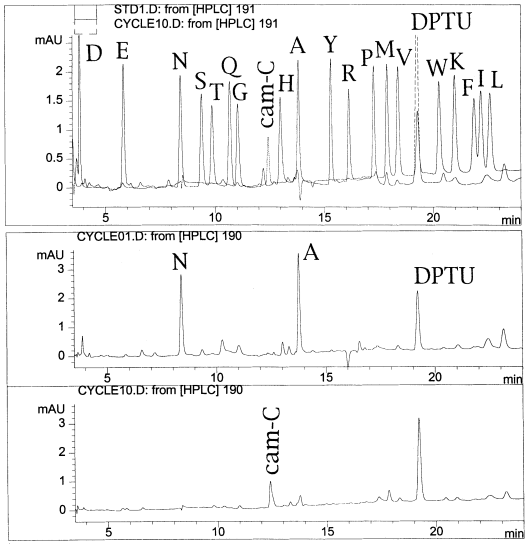

Fig. 6. Protein sequencing data for peak 26 (Fig. 5). Elution of PTH amino acids overlapped by the carboxyamidomethyl cysteine standard (Cam-C) (top). Results of the cycles 1 (middle) and 10 (bottom) demonstrate the occurrence of Asn and Ala and carboxyamidomethyl cysteine (Cam-C), respectively.

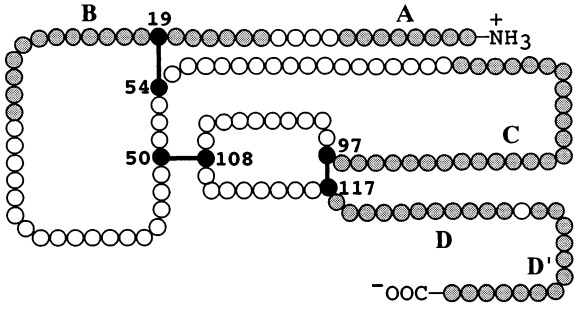

Fig. 8. Schematic diagram of the three disulfide linkages. Predicted helical regions (A, B, C and D) by the Chou-Fasman methods (Mac Vector software) are highlighted in gray, whereas the six cystines are shown in black. The prediction of the internal α-helical regions overlaps closely those by the Robison-Garnier method.

showed the incorporation of four carboxyamidomethyl groups (16 110 Da; calculated, 16 109 Da) so that the protein had only one disulfide bridge left. The same was observed with peak 20 (30.9 min; Fig. 4), but in this case two isomers with single disulfide links were observed but could not be completely separated by HPLC. The alkylation product of peak 21 (31.6 min) was mainly due to the complete reduction of BmPBP (observed, 16 225 Da; calculated, 16 225 Da).

Although reduction and alkylation per se did not solve the puzzle, enzymatic digestion of partially reduced proteins did. Digestion of the carboxyamidomethylated proteins with endoproteinase Lys-C was completed in ca. 3 h under the same conditions used for the native protein. The tetra-carboxyamidomethylated protein (only one disulfide bridge) gave a peak at 13.7 min with a molecular mass of 2388 Da. This was assigned to fragments #3 with Lys-21 remaining (ALDECKK) and #7 (Fig. 1B) linked by a disulfide bridge (calculated, 2387.1 Da). On the other hand, digestion of the major di-alkylated protein (originated from peak 18, Fig. 4) with endoproteinase Lys-C gave one fragment (peak 26, 17.4 min; Fig. 5) with a molecular mass of 3727 Da along with peak 27 (19.9 min) with molecular mass of 2252 Da. The MS data suggest that the latter is due to the carboxyamidomethylated fragment #10 (Fig. 1B) (calculated, 2253 Da). The former could be due to the fragments #3 (plus Lys-21), #7, and #12 linked by two disulfide bonds and having an S-carboxyamidomethyl cysteine (Fig. 5, insert) (calculated, 3727.8 Da). Given the fact that one way to detect cysteines in protein sequencing is to stabilize them for Edman degradation and detect their S-alkyl derivatives [27], we hypothesized that sequencing this three-peptide-containing fragment (Fig. 5, peak 26 and insert) would give us information regarding the existing cysteine bridges. Cys-19 is linked to Cys-54; therefore, the detection of an S-carboxyamidomethyl in the 10th cycle would indicate that Cys-117 is alkylated (Fig. 5, insert) and, consequently, that Cys-50 would be linked to Cys-108. Because Cys-54 is a cystine and consequently not detectable without derivatization, it would not be a problem even if it is at the same location (equivalent to cycle 10). On the other hand, if the S-carboxyamidomethyl would appear in the first cycle, this would support that Cys-50 is linked to Cys-117. The PTH derivative of the S-carboxyamidomethyl cysteine (Cam-Cys) standard (12.4 min) was well resolved from the other PTH standards (Fig. 6, top). The first cycle of the sequencing for the isolated peak 26 (Fig. 5) gave only two amino acids (Ala and Asn), whereas Cam-Cys alone appeared in cycle 10 (Fig. 6). These results indicate that Cys-50 was bound to Cys-108 and, consequently, Cys-97 was bridged to Cys-117.

Our next goal was to obtain direct evidence for the Cys-97-Cys-117 linkage. Digestion of the tetrapeptide containing all disulfide linkages in the original protein (Fig. 2, peak 5) with chymotrypsin generated three peaks (Fig. 7). Fortunately, each of the three fragments had one of the original disulfide bonds. The molecular mass (1354 Da) of peak 30 (9.6 min) suggests that it was derived from the cleavage at Met-54 generating the fragment ALDECKK and CLSTK (calculated, 1353.7 Da). Further cleavage at Trp-110 generated its counterpart, peak 31 (17.8 min) with a molecular mass of 1412 Da (calculated, 1411.6 Da). The identity of the fragment (NRETGCAIM, CIW), supported also by the amino acid sequence of the isolated peak (cycle 1: N; 2: R; 3: E; 4: T; 5: G; 6: X; 7: A, I), further confirmed that Cys-50 is linked to Cys-108. On the other hand, peak 32 (21.2 min) gave a molecular mass of 3004 Da suggesting that it was the counterpart of cleavage at Met-110 with another cleavage at Phe-118, i.e. TLGVATCF linked to HGADETMAQQLIDIVHGCEK (calculated, 3003.4 Da). Identity of the dipeptide, confirmed by the amino acid sequence of the isolated peak (cycle 1: T, H; 2: L, G; 3: G, A; 4: V, D; 5: A, E; 6: T; 7: M; 8: F, A; 9: Q; 10: Q; 11: L; 12: I; 13: D), showed that Cys-97 is bridged to Cys-117. All of these findings together demonstrate that, in the recombinant BmPBP, the cystines are linked as depicted in Fig. 8.

Next, we compared the disulfide connectivity of the recombinant to the native BmPBP. Digestion of purified nBmPBP with Lys-C gave the same profile as that for the recombinant protein (Fig. 2, insert) with the identity of the peaks confirmed by LC/ESI-MS. Chymotrypsin digestion of the isolated peak 5′ gave a profile similar to that of the recombinant counterpart (Fig. 7, insert). An extra peak at 20.1 min (peak 33) with a molecular mass of 3132 Da is a precursor of peak 32′ (preceding the cleavage at Phe-118). The same peak was also observed during similar incubation of the fragment derived from the recombinant protein. However, it decreased with time and became a minor peak after 2 h incubation (Fig. 7). With the scanty amount of native protein, we did not follow the digestion with time but analyzed the whole reaction mixture after 70 min. The relative amount of enzyme was very high so that the peak of chymotrypsin could be detected (peak 34, 23.7 min). In conclusion, the native protein is indistinguishable from rBmPBP in disulfide connectivity (Fig. 8).

As opposed to the β-sequence rich odorant binding proteins of vertebrates [37], insect PBPs are helical rich proteins. In BmPBP, this was demonstrated both by the circular dichroism spectrum [22] and predictions (53% α-helical) based on curve fitting of the CD data (Miyazawa and Leal, unpublished data). On the other hand, prediction of the secondary structure based on the amino acid sequence of BmPBP suggests that the protein has a 55% helical component divided into five regions, namely, a short one (Fig. 8, A; eight residues) at the N-terminal, two in the internal part of the protein (B, 19 residues; C, 26 residues) and another one at the C-terminal. This one is interrupted at the middle by Pro-129 (D, 12; D′, 13), resembling what has been observed in the crystal structure of interleukin-2 [38]. This prediction suggests that disulfide bonds play a pivotal role in the rigidity of the three-dimensional structure of the protein. The flexible segment between the two major internal helical regions (B and C, Fig. 8) is stabilized by two disulfide linkages (Cys-19-Cys-54 and Cys-50-Cys-108). Moreover, the disulfide bridge Cys-97-Cys-117 seems to determine the geometry of C and D helical regions by anchoring the carboxy-terminal of the former to the N-terminal of the latter; a similar case has been experimentally observed in the benign mouse prion [39].

Efforts now in progress to determine the three-dimensional crystal structure of BmPBP by X-ray crystallography and the solution structure by NMR spectroscopy may be facilitated by the disulfide bond assignment elucidated in this study. In addition, these structural studies will not only test the hypothesis put forward in this paper but also unveil the details of pheromone binding, transport and release of ligands to the receptors and other details of the events preceding the cascade of signal transduction.

Acknowledgements: This research was supported by the Program for Promotion of Basic Research Activities for Innovative Biosiences (BRAIN). We thank the members of our laboratory for helpful discussions and Dr. Shigeru Matsunaga (Yokogawa Analytical Systems) for his critique of an earlier draft version of the manuscript.

References

[1] Vogt, R.G., Köhne, A.C., Dubnau, J.T. and Prestwich, G.D. (1989) J. Neurosci. 9, 3332–3346.
[2] Steinbrecht, R.A., Ozaki, M. and Zeigelberger, G. (1992) Cell Tissue Res. 270, 287–302.
[3] Du, G. and Prestwich, G.D. (1995) Biochemistry 34, 8726–8732.
[4] Vogt, R.G. and Riddiford, L.M. (1981) Nature 293, 161–163.
[5] Györgyi, T.K., Roby-Shemkovitz, A.J. and Lerner, M.R. (1988) Proc. Natl. Acad. Sci. USA 85, 9851–9855.
[6] Raming, K., Krieger, J. and Breer, H. (1989) FEBS Lett. 256, 215–218.
[7] Raming, K., Krieger, J. and Breer, H. (1990) J. Comp. Physiol. B 160, 503–509.
[8] Breer, H., Krieger, J. and Raming, K. (1990) Insect Biochem. 20, 735–740.
[9] Krieger, J., Raming, K. and Breer, H. (1991) Biochim. Biophys. Acta 1088, 277–284.
[10] Vogt, R.G., Rybczynski, R. and Lerner, M.R. (1991) J. Neurosci. 11, 2972–2984.
[11] Krieger, J., Gänssle, H., Raming, K. and Breer, H. (1993) Insect Biochem. Mol. Biol. 23, 449–456.
[12] Krieger, J., von Nickisch-Rosenegk, E., Mameli, M., Pelosi, P. and Breer, H. (1996) Insect Biochem. Mol. Biol. 26, 297–307.
[13] Wojtasek, H., Hansson, B.S. and Leal, W.S. (1998) Biochem. Biophys. Res. Commun. 250, 217–222.
[14] Maibèche-Coisne, M., Jacquin-Joly, E., François, M.-C. and Nagnan-le Meillour, P. (1998) Insect Biochem. Mol. Biol. 28, 815–818.
[15] Vogt, R.G., Callahan, F.E., Rogers, M.E. and Dickens, J.C. Chem. Senses, in press.
[16] Wojtasek, H., Picimbon, J.-F. and Leal, W.S. Biochem. Biophys. Res. Commun., in press.
[17] Krieger, J., Raming, K., Prestwich, G.D., Frith, D., Stabel, S. and Breer, H. (1992) Eur. J. Biochem. 203, 161–166.
[18] Prestwich, G.D. (1993) Protein Sci. 2, 420–428.
[19] Feng, L. and Prestwich, G.D. (1997) Insect Biochem. Mol. Biol. 27, 405–412.
[20] Maibèche-Coisne, M., Longhi, S., Jacquin-Joly, E., Brunel, C., Egloff, M.-P., Gastinel, L., Cambillau, C., Tegoni, M. and Nagnan-le Meillour, P. (1998) Eur. J. Biochem. 258, 768–774.
[21] Merritt, T.J.S., LaForest, S., Prestwich, G.D., Quattro, J.M. and Vogt, R.G. (1998) J. Mol. Evol. 46, 272–276.
[22] Wojtasek, H. and Leal, W.S. J. Biol. Chem., in press.
[23] Freedman, R.B. (1984) Trends Biochem. Sci. 9, 438–441.
[24] Creighton, T.E. (1988) BioEssays 8, 57–63.
[25] Ziegelberger, G. (1995) Eur. J. Biochem. 232, 706–711.
[26] Johnson, B.H. and Hecht, M.H. (1994) Bio/Technology 12, 1357–1360.
[27] Anonymous (1995) Hewlett Packard Application Note 95-2, pp. 1–2.
[28] Gray, W.E. (1993) Protein Sci. 2, 1732–1748.
[29] Ryle, A.P. and Sanger, F. (1955) Biochem. J. 60, 535–540.
[30] Fontana, A., Dalzoppo, D., Grandi, C. and Zambonin, M. (1981) Biochemistry 20, 6997–7004.
[31] Schroeder, W.A., Shelton, J.B. and Shelton, J.R. (1969) Arch. Biochem. Biophys. 130, 551–556.
[32] Pappin, D.J.C. and Findlay, J.B.C. (1984) Biochem. J. 217, 605–613.
[33] Aitken, A., Geisow, M.J., Findlay, J.B.C., Holmes, C. and Yarwood, A. (1989) in: Protein Sequence, a Practical Approach (Findlay, J.B.C. and Geisow, M.J., Eds.), pp. 43–68, IRL Press, Tokyo.
[34] Ozols, J. and Gerard, C. (1977) J. Biol. Chem. 252, 5986–5989.
[35] Ozols, J. and Gerard, C. (1977) J. Biol. Chem. 252, 8549–8553.
[36] Burns, J.A., Butler, J.C., Moran, J. and Whitesides, G.M. (1991) J. Org. Chem. 56, 2648–2650.
[37] Spinelli, S., Ramoni, R., Grolli, S., Bonicel, J., Cambillau, C. and Tegoni, M. (1998) Biochemistry 37, 7913–7918.
[38] Brandhuber, B.J., Boone, T., Kenney, W.C. and McKay, D.B. (1987) Science 238, 1707–1709.
[39] Riek, R., Hornemann, S., Wider, G., Billeter, M., Glockshuber, R. and Wuthrich, K. (1996) Nature 382, 180–182.